CODEN: JASMAN

The Journal of the Acoustical Society of America

ISSN: 0001-4966

Vol. 127, No. 2 February 2010

JASA EXPRESS LETTERS

LETTERS TO THE EDITOR

(Continued)

CONTENTS—*Continued from preceding page*

(Continued)

CONTENTS—*Continued from preceding page*

(Continued)

CONTENTS—*Continued from preceding page*

SPEECH PRODUCTION [70]

SPEECH PERCEPTION [71]

SPEECH PROCESSING AND COMMUNICATION SYSTEMS [72]

MUSIC AND MUSICAL INSTRUMENTS [75]

BIOACOUSTICS [80]

(Continued)

CONTENTS—*Continued from preceding page*

Crosstalk mitigation using adaptive time reversal

H. C. Song,[a] **J. S. Kim, and W. S. Hodgkiss**
Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0238
hcsong@mpl.ucsd.edu, jeasoo@mpl.ucsd.edu, wsh@mpl.ucsd.edu

J. H. Joo
Maritime R&D Laboratory, LIG Nex1 Co., Ltd., Gyeonggi-do 446-912, Korea
joojaehoon@lignex1.com

Abstract: Although the spatial focusing property of the conventional time reversal approach facilitates multiuser communications, there always is residual crosstalk between users. A recent paper [Kim and Shin, J. Acoust. Soc. Am. **115**, 600–606 (2004)] proposed an adaptive active time reversal approach for simultaneous multiple focusing with minimal interference. This letter applies the adaptive approach to passive time reversal, multiuser communications for additional suppression of crosstalk among users. Experimental data at 3.5 kHz with a 1-kHz bandwidth demonstrate as much as 6.5-dB improvement per user in terms of output signal-to-noise ratio for three-user communications over a 20-km range in 120-m deep shallow water.

PACS numbers: 43.60.Dh, 43.60.Gk, 43.60.Fg [JC]
Date Received: September 27, 2009 **Date Accepted:** November 23, 2009

1. Introduction

Over the past several years, the time reversal approach has attracted attention in underwater acoustic communications due to its robustness and lower computational complexity as an alternative to multichannel equalizers.[1–3] Time reversal, either active or passive, exploits spatial diversity to mitigate the intersymbol interference (ISI) in channels with significant delay spread and can provide nearly optimal performance when combined with channel equalization to remove the residual ISI (see Fig. 1). Passive time reversal essentially is equivalent to active time reversal with the communications link being in the opposite direction.[4]

The spatial focusing property of time reversal has been applied to multiuser communications[4] where multiple users transmit information simultaneously to the passive time reversal array (similar to a base station in wireless communications). Crosstalk between users (e.g., Fig. 8 of Ref. 4), however, cannot be completely eliminated using conventional time reversal processing.

Recently Kim and Shin[5] proposed an adaptive active time reversal approach for simultaneous multiple focusing with minimal interference between the received transmissions. The objective of this paper is to apply the adaptive approach to passive time reversal, multiuser communications and investigate its performance improvement over the conventional approach using at-sea experimental data. In Sec. 2 we derive the adaptive time reversal filter in the context of the passive time reversal approach, followed by data analysis.

2. Adaptive time reversal filter

As depicted in Fig. 2, denote $h_i^j(t)$, $i=1,\ldots,M$, as known channel responses from a user j to a vertical receive array (VRA) consisting of M elements. Note that the channel response includes

[a]Author to whom correspondence should be addressed.

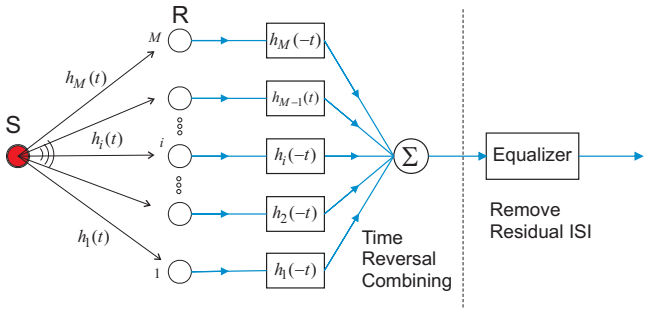


Fig. 1. (Color online) Passive time reversal receiver for the single user case. Time reversal combining is followed by a single channel equalizer to remove residual ISI.

the shaping pulse, transmit filter, channel impulse response, and receive filter. The conventional time reversal approach then extracts signals from each user j simply by matched-filtering the received signals $r_i(t)$ with each set of the channel responses $h_i^j(-t)$ followed by combining. The adaptive filter will be derived in the frequency domain and then converted into a time-domain filter using an inverse fast Fourier transform (IFFT). For simplicity, we consider a two-user case which easily can be generalized to more users.

Let us define a column vector $\mathbf{d}_j$ as the collective channel responses at frequency f from each user j such that

$$\mathbf{d}_1 = \begin{bmatrix} H_1^1(f) \\ \vdots \\ H_M^1(f) \end{bmatrix} \quad \text{and} \quad \mathbf{d}_2 = \begin{bmatrix} H_1^2(f) \\ \vdots \\ H_M^2(f) \end{bmatrix}. \tag{1}$$

The received signal vector $\mathbf{r}$ then is simply the addition of these two vectors excluding the noise component

$$\mathbf{r} = \mathbf{d}_1 + \mathbf{d}_2. \tag{2}$$

Denoting a column vector $\mathbf{w}$ as a weight vector (spatial filter) for combining the array element data, the output for User 1, u_1, becomes

$$u_1 = \mathbf{w}_1^\dagger \mathbf{r} = \mathbf{w}_1^\dagger \mathbf{d}_1 + \mathbf{w}_1^\dagger \mathbf{d}_2, \tag{3}$$

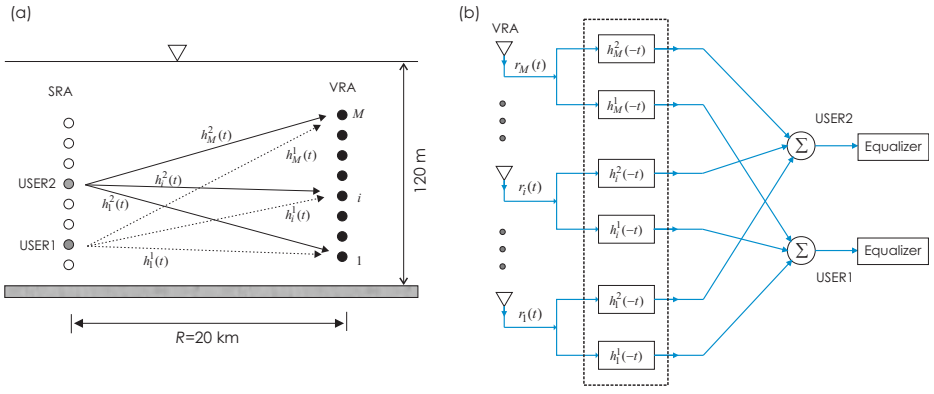


Fig. 2. (Color online) (a) Schematic diagram of multiuser communications using passive time reversal. A subset of elements of the SRA is selected as multiple users separated in depth which transmit information to the VRA. Time reversal approach requires knowledge of the channel responses $h_i^j(t)$ on the VRA from each user j. (b) Block diagram separating signals from different users in conventional passive time reversal. For an adaptive time reversal receiver, the space/time matched filters $h_i^j(-t)$ will be replaced by $w_i^j(-t)$ derived in Eq. (4).

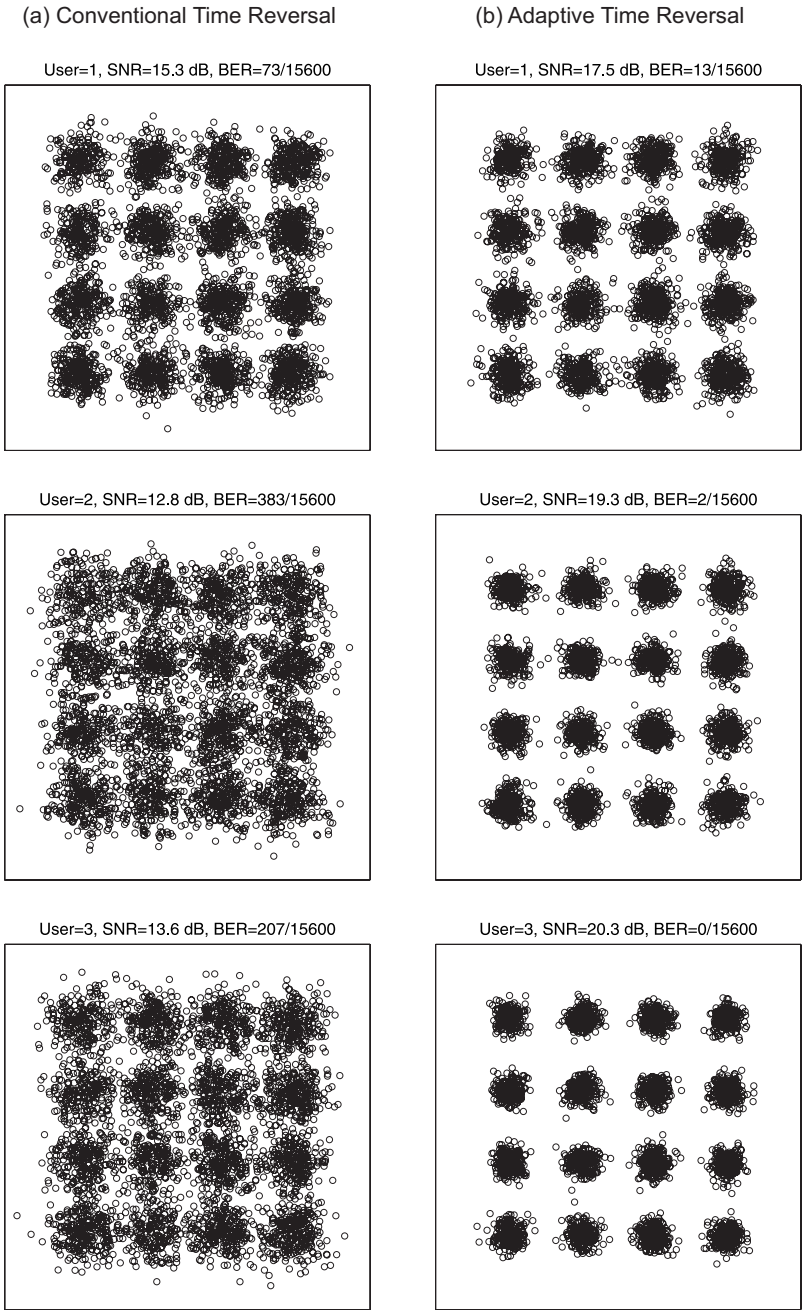


Fig. 3. Experimental results at 3.5 kHz with a 1-kHz bandwidth during FAF-05 experiment: (a) conventional time reversal and (b) adaptive time reversal. Scatter plots are shown for three-user communications using 16-QAM over a 20-km range in 120-m deep shallow water. The performance improvement by adaptive time reversal (bottom) is noticeable in terms of output SNR, i.e., 2, 6.5, and 6.7 dB, respectively. The three users are positioned at depths of 88, 101, and 113 m, respectively (about 12-m separation) and the bottom 20 elements ($M=20$, 40-m array aperture) of the VRA are used for spatial diversity. The symbol rate is 500 symbols/s and the aggregate data rate is 6 kbits/s.

where † denotes the complex conjugate operation.

For conventional time reversal processing, $\mathbf{w}_1=\mathbf{d}_1$ and the second term of Eq. (3), $\mathbf{d}_1^{\dagger}\mathbf{d}_2$, remains as crosstalk between the two users. To further suppress the crosstalk, an adaptive

time reversal filter[5] for User 1, $\mathbf{w}_1$, can be derived subject to the constraint that $\mathbf{w}_1^\dagger \mathbf{d}_2 = 0$, resulting in

$$\mathbf{w}_1 = \frac{\mathbf{R}^{-1}\mathbf{d}_1}{\mathbf{d}_1^\dagger \mathbf{R}^{-1}\mathbf{d}_1} \quad \text{where } \mathbf{R} = \mathbf{d}_1\mathbf{d}_1^\dagger + \mathbf{d}_2\mathbf{d}_2^\dagger + \sigma^2\mathbf{I}. \tag{4}$$

Note that $\mathbf{R}$ is a cross spectral density matrix exploiting the knowledge of channel responses at the receiving array and the denominator is a normalization constant. $\sigma^2\mathbf{I}$ is a small diagonal loading for a matrix inversion with $\mathbf{I}$ being an identity matrix. Similarly, an adaptive time reversal filter for User 2, $\mathbf{w}_2$, can be derived from Eq. (4) by substituting $\mathbf{d}_1$ with $\mathbf{d}_2$ while $\mathbf{R}$ remains the same.

Now we can construct the adaptive time-domain filter $w_i^j(-t)$ which replaces the conventional filter $h_i^j(-t)$ in Fig. 2 to eliminate the crosstalk between the users.

3. Multiuser communications with minimal crosstalk

In order to demonstrate the proposed adaptive time reversal approach, we revisit the time reversal experiment (FAF-05) conducted jointly with NATO Undersea Research Centre in July 2005, north of Elba Island off the west coast of Italy. A detailed description of the experiment can be found in Ref. 4, and here we review the main parameters relevant for our discussion.

The FAF-05 experiment was carried out in a flat region of 120-m deep water, and a 32-element VRA was deployed at two different ranges (4 and 20 km), spanning the water column from 48 to 110 m with 2-m spacing (Fig. 2). The multiple users (transmitters) were selected from the 29-element source/receive array (SRA). In this paper, we present results for three-user communications at 20-km range. Since most of the energy at 20-km range is captured by the receiver elements at deeper depths (see Fig. 10 of Ref. 4), here we use only the bottom 20 elements (i.e., $M = 20$) for time reversal processing. The probe signal was a 300-ms, 2.5–4.5-kHz linear frequency modulation (LFM) chirp with a Hanning window, resulting in an effective 200-ms, 3–4-kHz bandwidth chirp. The symbol rate was 500 symbols/s with an excess bandwidth of 100%, and the communication sequence was 9-s long. The aggregate data rate obtained with three users is then 6 kbits/s using 16 quadrature amplitude modulation (QAM). The transmitters (users) are positioned at depths of 88, 101, and 113 m, respectively (about 12-m separation).

The performance of three-user communications is shown in Fig. 3 as scatter plots. The left three panels (a) are a reproduction of the conventional time reversal approach presented in Ref. 4 with an overall bit error rate of 1.4% (refer to Fig. 13). Results of the proposed adaptive time reversal approach are shown in the right three panels (b), exhibiting almost error-free performance. Note that the output signal-to-noise ratios (SNRs) (reciprocal of mean-square-error) have increased to 17.5, 19.3, and 20.3 dB from 15.3, 12.8, and 13.6 dB, respectively, such that the improvement for Users 2 and 3 is up to 6.5 dB. This example clearly demonstrates that the adaptive time reversal approach can suppress the crosstalk among users, thus resulting in a significant performance enhancement in multiuser communications.

References and links

[1]H. C. Song, W. A. Kuperman, and W. S. Hodgkiss, "Basin-scale time reversal communications," J. Acoust. Soc. Am. **125**, 212–217 (2009).
[2]J. Gomes, A. Silva, and S. Jesus, "Adaptive spatial combining for passive time-reversed communications," J. Acoust. Soc. Am. **124**, 1038–1053 (2008).
[3]T. C. Yang, "Correlation-based decision-feedback equalizer for underwater acoustic communications," IEEE J. Ocean. Eng. **30**, 865–880 (2005).
[4]H. C. Song, W. S. Hodgkiss, W. A. Kuperman, T. Akal, and M. Stevenson, "Multiuser communications using passive time reversal," IEEE J. Ocean. Eng. **32**, 915–926 (2007).
[5]J. S. Kim and K. C. Shin, "Multiple focusing with adaptive time-reversal mirror," J. Acoust. Soc. Am. **115**, 600–606 (2004).

Effects of stimulus duration on amplitude modulation processing with cochlear implants

Xin Luo[a)]
Department of Speech, Language, and Hearing Sciences, Purdue University, 500 Oval Drive, West Lafayette, Indiana 47907
luo5@purdue.edu

John J. Galvin III and Qian-Jie Fu
Division of Communication and Auditory Neuroscience, House Ear Institute, 2100 West Third Street, Los Angeles, California 90057
jgalvin@hei.org, qfu@hei.org

Abstract: This study investigated temporal integration processes underlying cochlear implant (CI) users' amplitude modulation processing. Thresholds for modulation detection (AMDTs) and modulation frequency discrimination (AMFDTs) were measured for 50-, 100-, and 200-Hz modulation frequencies with stimulus durations from 50 to 400 ms in eight adult CI users. The results showed significant interactions between modulation frequency and stimulus duration for AMDTs and AMFDTs. The data suggest that temporal integration limits CI users' sensitivity to low temporal pitch over short durations, and that temporal integration over longer durations may not enhance CI users' sensitivity to high temporal pitch.

PACS numbers: 43.66.Ts, 43.66.Fe, 43.66.Hg [BLM]
Date Received: October 30, 2009 **Date Accepted:** December 2, 2009

1. Introduction

Due to the limited spectral cues available with cochlear implants (CIs), pitch information is primarily encoded by temporal amplitude modulation (AM) of electric pulse trains. CI users can typically perceive temporal envelope pitch for modulation frequencies up to ~300 Hz (Zeng, 2002). CI speech performance has been significantly correlated with thresholds for AM detection and AM frequency discrimination, suggesting the importance of temporal AM processing for CI users (Fu, 2002; Chatterjee and Peng, 2008; Luo *et al.*, 2008). Various stimulation parameters have been shown to affect CI users' AM detection thresholds (AMDTs), such as AM frequency (Shannon, 1992), stimulation level (Fu, 2002), electrode location (Pfingst *et al.*, 2008), and stimulation rate (Galvin and Fu, 2005, 2009; Pfingst *et al.*, 2007). However, AM frequency discrimination thresholds (AMFDTs) may be a more relevant measure of CI users' temporal pitch sensitivity, as it may better reflect sensitivity to pitch changes in dynamic stimuli such as speech (e.g., intonation, lexical tones, etc.). CI users' AMFDTs may also be affected by stimulation parameters, such as AM frequency, electrode location (Chatterjee and Peng, 2008), and stimulation level (Luo *et al.*, 2008).

For AM detection or AM frequency discrimination, listeners must optimally integrate information (stored in short-term memory) from multiple "looks" at the input modulations. A multiple-look model (Viemeister and Wakefield, 1991) has been used to explain increasing sensitivity to AM with increasing stimulus duration in normal-hearing (NH) listeners. As the stimulus duration is increased, more modulation periods are available, allowing NH listeners to better perceive AM. Viemeister (1979) measured NH listeners' AMDTs for AM frequencies ranging from 2 to 4000 Hz, using stimulus durations of 250, 500, and 1500 ms. Results showed that AM detection improved with stimulus duration, and that the improvement was larger for

[a)]Author to whom correspondence should be addressed.

Table 1. CI subject demographics.

Subject	Age	Gender	Etiology	Device	Strategy	Years with prosthesis
S1	50	M	Trauma	Nucleus-22	SPEAK	15
S2	64	F	Genetic	Nucleus-24	ACE	4
S3	76	M	Noise induced	Nucleus-22	SPEAK	10
S4	61	F	Congenital	Nucleus-24	SPEAK	9
S5	66	M	Trauma	Nucleus-22	SPEAK	17
S6	57	M	Genetic	Freedom	ACE	2
S7	74	F	Unknown	Nucleus-24	ACE	7
S8	44	M	Congenital	Nucleus-24	ACE	6

lower AM frequencies (which contained fewer than 16 modulation periods with short durations). The effect of stimulus duration on NH listeners' AMFDTs was investigated by Lee (1994) using sinusoidal carriers. In one condition, the carrier frequency was randomized across stimuli so that spectral sideband cues were unavailable, forcing listeners to attend to temporal cues. Results showed that AMFDTs decreased for durations up to about five modulation periods, but were largely unaffected by stimulus durations beyond this "critical" duration. Thus, there seems to be an integration period for NH listeners' temporal modulation processing.

In this study, we measured AMDTs and AMFDTs in CI users for a range of AM frequencies (50–200 Hz) and stimulus durations (50–400 ms). Given these experimental parameters, the number of modulation periods within the standard stimuli ranged from 2.5 to 80. We hypothesized that there is a critical temporal integration window that underlies CI users' AM processing, and that CI users' AM sensitivity will improve with increasing stimulus duration.

2. Methods

2.1 Subjects

Table 1 shows demographic details for the eight post-lingually deafened adult CI subjects. All subjects had extensive experience in psychophysical tests from previous studies. Informed consent was obtained from all subjects.

2.2 Stimuli

All stimuli were bi-phasic pulse trains delivered via custom research interface to electrode 10. The stimulation rate was 2000 pps (pps denotes pulses per second); this relatively high stimulation rate was selected to avoid aliasing effects with amplitude modulation. The AM frequency was 50, 100, or 200 Hz. The stimulus duration was 50, 100, 200, or 400 ms. Table 2 shows other

Table 2. Experimental stimulation parameters.

Subject	Stimulation mode	Phase duration (μs)	Inter-phase gap (μs)	Loudness-balanced levels at 50% of dynamic range (μA)			
				50 ms	100 ms	200 ms	400 ms
S1	BP+1	100	20	718	703	638	615
S2	MP1+2	50	8	213	215	182	187
S3	BP+1	100	20	600	579	463	448
S4	MP1+2	50	8	676	613	518	500
S5	BP+1	100	20	600	579	463	448
S6	BP+1	100	20	657	396	289	274
S7	BP+1	100	20	137	134	129	137
S8	BP+1	100	20	208	169	152	150

relevant stimulation parameters; note that different parameters were used for different subjects to achieve sufficient loudness for all experimental conditions.

Prior to the AMDT and AMFDT experiments, all stimuli were loudness-balanced across duration conditions. First, the electrode dynamic range (DR) was estimated for the 200-ms stimuli (no modulation), using methods similar to clinical fitting (i.e., "counting" thresholds, then slowly increasing the amplitude until achieving comfortable loudness). The 50% DR level (in linear microamperes) for the 200-ms stimuli was used as the reference for loudness-balancing. The 50-, 100-, and 400-ms stimuli (no modulation) were loudness-balanced to this reference using a two-interval, forced-choice adaptive double-staircase procedure (Jesteadt, 1980). The target stimulation levels (in linear microamperes) were averaged across two sequences and were used as the stimulation levels for the different duration conditions (see Table 2).

2.3 Procedures

All subjects (except S1) participated in the AMDT experiment. A 3–interval, forced–choice (3IFC), three-down/one-up adaptive procedure was used to track the modulation depth corresponding to 79.4% correct AM detection (Levitt, 1971). In each trial, two intervals contained un-modulated stimuli, and the other interval (randomly selected) contained sinusoidal (current) AM. The starting phase of the sinusoidal AM was randomly selected between 0 and 2π. Subjects were asked to identify which interval sounded different. The AM depth was adjusted according to subject response. The adaptive procedure terminated after 12 reversals or after 60 trials, and the AM depths (in percent) for the final 8 reversals were averaged to obtain the AMDT for each run; two to three runs were conducted for each condition.

AMFDTs were measured for a fixed 30% AM depth. This modulation depth was greater than most of the measured AMDTs. A 3IFC, three-down/one-up procedure was used to track the AM frequency corresponding to 79.4% correct AM frequency discrimination (Levitt, 1971). In each trial, two intervals contained the standard AM frequency, and the other interval (randomly selected) contained the target AM frequency that was always higher than the standard. Subjects were asked to identify which interval was different. The target AM frequency was adjusted according to subject response. The adaptive procedure terminated after 12 reversals or 60 trials, and the target AM frequencies for the final 8 reversals were averaged to obtain the AMFDT for each run; two to three runs were conducted for each condition. AMFDTs were expressed as Weber fractions ($\Delta f_{\mathrm{mod}}/f_{\mathrm{mod}}$). The test order of conditions was randomized within and across subjects. No preview, training, or feedback was provided during testing.

3. Results

Figure 1 shows AMDTs for individual subjects. There was a large inter-subject variability in AMDTs. A two-way repeated measures analysis of variance (RM ANOVA) showed that AMDTs were significantly affected by AM frequency [$F(2,12)=4.16$, $p=0.04$], but not by stimulus duration [$F(3,18)=1.14$, $p=0.36$]; there was a significant interaction between AM frequency and stimulus duration [$F(6,36)=4.35$, $p=0.002$]. According to *post-hoc* Bonferroni *t*-tests, 50- and 100-Hz AMDTs were not significantly affected by stimulus duration, but 200-Hz AMDTs were significantly higher for the 200- and 400-ms durations than for the 100-ms duration ($p<0.04$), driven mostly by the performance of subjects S2, S3, and S5. Performance for these same three subjects likely caused 200-Hz AMDTs to be significantly higher than 50- and 100-Hz AMDTs with the 200- and 400-ms durations ($p<0.05$).

Figure 2 shows AMFDTs for individual subjects. Again, there was large inter-subject variability in AMFDTs. For most subjects, 50-Hz AMFDTs decreased by half or more as the stimulus duration was increased from 50 to 100 ms, but remained constant thereafter, in agreement with the putative critical duration of five modulation periods. For subjects S3, S4, and S7, 50-Hz AMFDTs continued to decrease with durations up to 400 ms, and for subject S5, 50-Hz AMFDTs were similar across all durations. 100-Hz AMFDTs were lower than 50-Hz AMFDTs and were largely unchanged across all durations (except for subjects S1, S4, and S7). 200-Hz

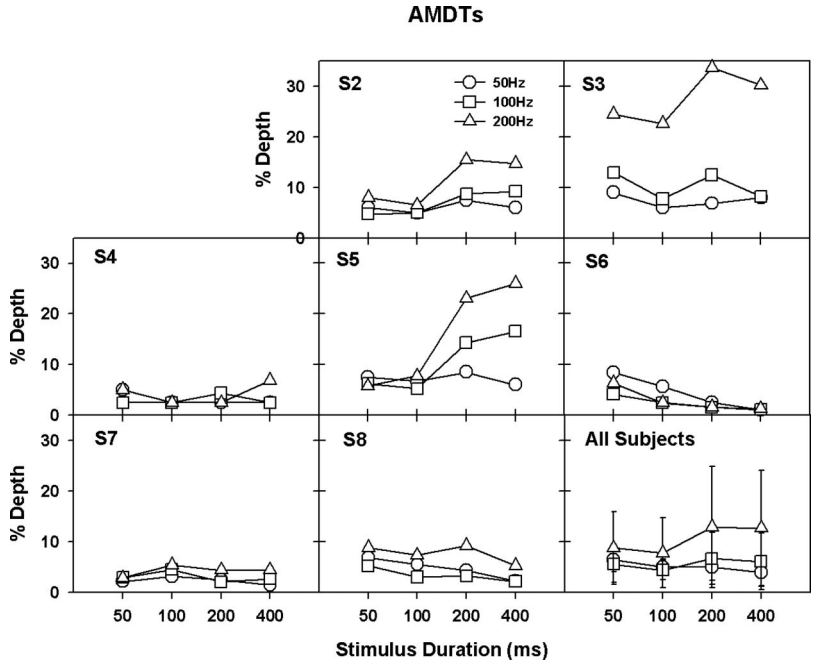


Fig. 1. AMDTs (in percent) as a function of stimulus duration. The results from individual subjects are shown in different panels. The right-bottom panel shows group mean data. The open circles, open squares, and open triangles represent the data for 50-, 100-, and 200-Hz AM frequencies, respectively.

AMFDTs were mostly similar to 100-Hz AMFDTs and were constant across all durations. Note that subjects S3 and S5 were unable to discriminate changes in AM frequency above 200 Hz, most likely due to the high AMDTs (see Fig. 1); due to the missing data, subjects S3 and S5 were excluded from the statistical analyses.

A two-way RM ANOVA showed that AMFDTs were significantly affected by stimulus duration [$F(3,15)=16.10$, $p<0.001$], but not by AM frequency [$F(2,10)=2.83$, $p=0.11$]; there was a significant interaction between stimulus duration and AM frequency [$F(6,30)=9.24$, $p<0.001$]. *Post-hoc* Bonferroni *t*-tests showed that 50-Hz AMFDTs were significantly higher for the 50-ms duration than for longer durations, and for the 100-ms duration than for the 400-ms duration ($p<0.03$). 100-Hz AMFDTs were significantly higher for the 50-ms duration than for the 200- and 400-ms durations ($p<0.05$). 200-Hz AMFDTs were not significantly different across all durations. *Post-hoc* analyses also showed that AMFDTs were significantly higher for 50 Hz than for 100 and 200 Hz with the 50-ms duration ($p<0.005$).

4. Discussion

The present AMDTs for the 200- and 400-ms durations are comparable to those from previous studies with similar stimulus durations (e.g., Shannon, 1992; Fu, 2002; Chatterjee and Peng, 2008; Luo *et al.*, 2008). Subjects S2, S3, and S5's AMDTs were higher for 200-Hz AM than for 50- and 100-Hz AMs with the 200- and 400-ms durations, reflecting the low-pass filter characteristics of CI users' temporal modulation transfer function (TMTF) (Shannon, 1992). However, AMDTs were largely unaffected by AM frequency for the remaining subjects with the 200- and 400-ms durations. AMDTs for all subjects (except for S3) were unaffected by AM frequency with the 50- and 100-ms durations. The results suggest that the cut-off frequency of the TMTF may be greater than 200 Hz for some CI users, and that the TMTF may be influenced by stimulus duration.

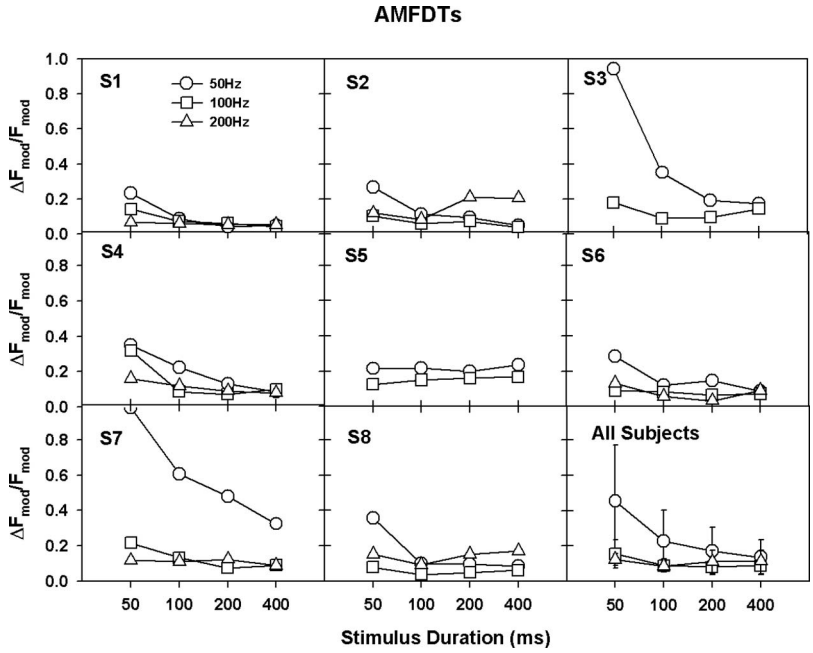


Fig. 2. AMFDTs (in $\Delta f_{mod}/f_{mod}$) as a function of stimulus duration. The results from individual subjects are shown in different panels. The right-bottom panel shows group mean data (including those from subjects S3 and S5). The open circles, open squares, and open triangles represent the data for 50-, 100-, and 200-Hz AM frequencies, respectively.

While NH listeners' AMDTs improve with increasing duration (Viemeister, 1979), the present CI subjects' 50- and 100-Hz AMDTs were largely unchanged across duration conditions. In contrast to NH results, CI users' mean 200-Hz AMDTs worsened with increasing duration (largely driven by the poorer performance of subjects S2, S3, and S5). According to the multiple-look model (Viemeister and Wakefield, 1991), if the information from each modulation period is mutually independent and optimally combined, the accumulative perceptual sensitivity should have improved with the greater number of modulation periods contained in the longer duration. The present pattern of results may be attributed to potential loudness cues used for AM detection. McKay and Henshall (2009) showed that AM stimuli were perceived louder than un-modulated stimuli with the same average current level. According to their AM loudness model, such loudness cues for AM detection would be similar across the present duration conditions because relative loudness is determined by a short temporal integration window (~7 ms). In the present study, loudness-balancing with un-modulated stimuli resulted in lower current levels for longer durations (see Table 2). However, according to McKay and Henshall's (2009) AM loudness model and their experimental data, the lower "absolute" current levels (while providing equal loudness for longer durations) may have reduced loudness difference between modulated and un-modulated stimuli, resulting in higher AMDTs. The reduced loudness cues for AM detection may have offset the benefits of longer duration (i.e., more looks at modulation), especially for subjects S2, S3, and S5 with 200-Hz AM detection.

The present AMFDTs for the 200-ms duration were similar to those measured by Chatterjee and Peng (2008). Performance patterns across different AM frequencies were also similar in both studies. With the 200-ms duration, 200-Hz AMFDTs were similar to 100-Hz AMFDTs; 50-Hz AMFDTs were slightly (but not significantly) higher than 100- and 200-Hz AMFDTs, possibly because the number of modulation periods was smaller for the 50-Hz AM. Also, the difference between 50-, 100-, and 200-Hz AMFDTs increased with shorter durations (i.e., 50 ms).

In line with the multiple-look model and with previous NH data (Lee, 1994), CI subjects' 50- and 100-Hz AMFDTs decreased with longer durations, showing the benefit of temporal integration when more modulation periods were available. These duration effects are consistent with pulse train rate discrimination performance in CI users (Tong *et al.*, 1982) as well as in NH subjects (Plack and Carlyon, 1995). Similar to NH data, the critical duration for 50-Hz AMFDTs was 100 ms (or five modulation periods) for subjects S1, S6, and S8; 50-Hz AMFDTs continued to decrease with longer durations for the remaining subjects. For 100- and 200-Hz AMFDTs, 50-ms duration (or five to ten modulation periods) was adequate for most CI subjects. The similar critical durations for AMFDTs between CI users and NH listeners suggest that similar temporal integration mechanisms and windows may be used for their AM processing, despite differences in auditory neuron inputs. Unlike AMDTs, AMFDTs are less likely to be affected by loudness cues, as McKay and Henshall's (2009) AM loudness model predicts little effects of AM frequency on overall loudness of AM stimuli. Chatterjee and Peng (2008) also found that AMFDTs were not significantly affected by 1-dB amplitude roving.

5. Summary

CI users' AM detection and AM frequency discrimination were measured as a function of stimulus duration. Significant interactions were found between AM frequency and stimulus duration. Lower-frequency AMFDTs (50 and 100 Hz) improved with stimulus durations up to 100 ms, suggesting that a critical duration (five to ten modulation periods) may be needed by CI users to process changes in AM frequency. Higher-frequency AMFDTs (200 Hz) were fairly constant across duration conditions (50–400 ms). Lower-frequency AMDTs (50 and 100 Hz) were not significantly affected by stimulus duration, while higher-frequency AMDTs (200 Hz) worsened with longer durations (200–400 ms), possibly due to reduced loudness cues used for AM detection. The results suggest that temporal integration limits CI users' sensitivity to low temporal pitch (50 Hz) over short durations (50–100 ms). In optimal listening conditions (e.g., clear speech with slow speaking rate), temporal integration over longer durations (400 ms) may not enhance CI users' sensitivity to high temporal pitch (200 Hz).

Acknowledgments

We are grateful to all subjects for their participation in the experiments. Research was supported in part by NIH (Grant Nos. R03-DC-008192 and R01-DC-004993).

References and links

Chatterjee, M., and Peng, S.-C. (**2008**). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition," Hear. Res. **235**, 143–156.
Fu, Q.-J. (**2002**). "Temporal processing and speech recognition in cochlear implant users," NeuroReport **13**, 1635–1639.
Galvin, J. J., and Fu, Q.-J. (**2005**). "Effects of stimulation rate, mode and level on modulation detection by cochlear implant users," J. Assoc. Res. Otolaryngol. **6**, 269–279.
Galvin, J. J., and Fu, Q.-J. (**2009**). "Influence of stimulation rate and loudness growth on modulation detection and intensity discrimination in cochlear implant users," Hear. Res. **250**, 46–54.
Jesteadt, W. (**1980**). "An adaptive procedure for subjective judgements," Percept. Psychophys. **28**, 85–88.
Lee, J. (**1994**). "Amplitude modulation rate discrimination with sinusoidal carriers," J. Acoust. Soc. Am. **96**, 2140–2147.
Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.
Luo, X., Fu, Q.-J., Wei, C.-G., and Cao, K.-L. (**2008**). "Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users," Ear Hear. **29**, 957–970.
McKay, C. M., and Henshall, K. R. (**2009**). "Amplitude modulation and loudness in cochlear implants," J. Assoc. Res. Otolaryngol. In press.
Pfingst, B. E., Burkholder-Juhasz, R. A., Xu, L., and Thompson, C. S. (**2008**). "Across-site patterns of modulation detection in listeners with cochlear implants," J. Acoust. Soc. Am. **123**, 1054–1062.
Pfingst, B. E., Xu, L., and Thompson, C. S. (**2007**). "Effects of carrier pulse rate and stimulation site on modulation detection by subjects with cochlear implants," J. Acoust. Soc. Am. **121**, 2236–2246.
Plack, C. J., and Carlyon, R. P. (**1995**). "Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics," J. Acoust. Soc. Am. **98**, 1355–1364.
Shannon, R. V. (**1992**). "Temporal modulation transfer functions in patients with cochlear implants," J. Acoust. Soc. Am. **91**, 2156–2164.

Tong, Y. C., Clark, G. M., Blamey, P. J., Busby, P. A., and Dowell, R. C. (**1982**). "Psychophysical studies for two multiple-channel cochlear implant patients," J. Acoust. Soc. Am. **71**, 153–160.
Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.
Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am. **90**, 858–865.
Zeng, F.-G. (**2002**). "Temporal pitch in electric hearing," Hear. Res. **174**, 101–106.

A simple method to account for size effects in the transfer matrix method

Dilal Rhazi[a] **and Noureddine Atalla**
Department of Mechanical Engineering, GAUS, Université de Sherbrooke, Sherbrooke, Quebec J1K 2R1, Canada
dilal.rhazi@usherbrooke.ca, noureddine.atalla@usherbrooke.ca

Abstract: The transfer matrix method based is extensively used and well-validated for predicting the transmission loss of multilayer structures. However, this method leads to poor results at low frequencies due to the infinite extent assumption it is based on. This paper presents an efficient implementation of a Rayleigh-integral based method to account for the finite size effects. The accuracy of the method is illustrated by various examples.

PACS numbers: 43.40.Rj, 43.25.Qp, 43.50.Cb [JM]
Date Received: April 8, 2009 **Date Accepted:** December 2, 2009

1. Introduction

The transfer matrix method (TMM) is widely used to predict the airborne transmission loss (TL) of flat infinite structures with added sound packages (Allard, 1993). To account for finite size effects and thus increase the accuracy of the method at low frequencies, Villot *et al.* (2001) proposed a wave approach based on spatial windowing to solve the response of multilayer structures with both airborne and a mechanical excitations. Ghinet and Atalla (2001) proposed a variant based on the Rayleigh integral to account for the size effect. The rationale behind both approaches is to replace the "infinite size" radiation efficiency in the receiving medium, by the radiation efficiency of an equivalent baffled window. The Rayleigh-integral based method, referred to by the finite transfer matrix method (FTMM), has been used by Atalla *et al.* (2006) to solve both the diffuse field transmission loss and absorption of flat multilayer structures. Various experimental examples were presented by the previous authors to corroborate the validity of the proposed methods. However, the main disadvantage of these methods is their high computational cost, especially for a structure-borne excitation. The aim of this note is to present an analytical simplification of the FTMM that allows for a fast and accurate computation of the radiated and transmitted power of multilayered panels.

2. Theory

For the sake of conciseness, the transmission problem is used to illustrate the method. The diffuse field transmission coefficient of a multilayer panel is given by

$$\tau = \frac{\int_0^{\theta_{\lim}} \int_0^{2\pi} \tau_\infty(\theta,\varphi) \sin\theta \cos\theta d\varphi d\theta}{\int_0^{\theta_{\lim}} \int_0^{2\pi} \sin\theta \cos\theta d\varphi d\theta}, \quad (1)$$

where $\tau_\infty(\theta,\varphi)$ is the transmission coefficient calculated using the TMM in the present context. To account for size effects, the FTMM proposes to replace the infinite size transmission coefficient by a finite size coefficient given by $\tau_f = \tau_\infty(\sigma_R \cos\theta)$, where σ_R is homogeneous to radiation efficiency and is given by (Atalla *et al.*, 2006)

[a]Author to whom correspondence should be addressed.

$$\sigma_R = \frac{\Re\mathfrak{e}(Z)}{\rho_0 c_0}, \tag{2}$$

with

$$Z = \frac{j\rho_0\omega}{S}\int_S\int_S \exp[-jk_t(x\cos\varphi + y\sin\varphi)]G(M,M_0) \times \exp[jk_t(x'\cos\varphi + y'\sin\varphi)]dxdydx'dy'. \tag{3}$$

Here S is the surface of the panel, $k_t = k_0 \sin\theta$ with k_0 the acoustic wavenumber, and ρ_0 and c_0 are the density and speed of sound in the receiver medium, respectively. G denotes the half space (baffled) Green's function:

$$G(x,y;x',y') = \frac{\exp[jkR]}{2\pi R}, \quad R = \sqrt{(x-y)^2 + (x'-y')^2}. \tag{4}$$

Note that the impedance Z as given by Eq. (3) only depends on the geometry of the panel, the angle of incidence θ, and heading angle φ. In the following, an analytical reduction in the order of integration in Eq. (3) from 4 to 1 is proposed. Let L_x and L_y denote the length and the width of the structure, respectively. Using the change in variables $\alpha = 2x/L_x$, $\beta = 2y/L_y$ and defining the plate's aspect ratio by $r = L_x/L_y$, Eq. (3) is first written in the form

$$Z = j\rho_0\omega\frac{L_y}{16\pi}\int_0^2\int_0^2\int_0^2\int_0^2 F_n(\alpha,\beta,\alpha',\beta')K(\alpha,\beta,\alpha',\beta')d\alpha d\beta d\alpha' d\beta' \tag{5}$$

with

$$K(\alpha,\beta,\alpha',\beta') = \frac{\exp[-jk_0R]}{\left[(\alpha-\alpha')^2 + \dfrac{(\beta-\beta')^2}{r^2}\right]^{1/2}}, \tag{6}$$

$$F_n(\alpha,\beta) = \exp\left[-j\frac{k_tL_x}{2}\left[(\alpha-\alpha')\cos\varphi + \frac{1}{r}(\beta-\beta')\sin\varphi\right]\right], \tag{7}$$

and

$$R = \frac{L_x}{2}\left[(\alpha-\alpha')^2 + \frac{1}{r^2}(\beta-\beta')^2\right]^{1/2}. \tag{8}$$

To reduce the quadruple integrals in Eq. (5) from 4 to 2, another change in variables is used. Let

$$u = \alpha - \alpha', \quad u' = \beta - \beta', \quad v = \alpha', \quad v' = \beta'. \tag{9}$$

Considering the variable α (the same formula can be written for β), a symbolic form of this change in variable reads

$$\int_0^2\int_0^2 d\alpha d\alpha' \rightarrow \int_0^2 du\int_0^{2-u} dv + \int_{-2}^0 du\int_{-u}^2 dv. \tag{10}$$

If K and F_n are rewritten in terms of u and u',

$$K(u,u') = \frac{\exp\left[-j\frac{k_0 L_x}{2}\left[u^2+\frac{u'^2}{r^2}\right]^{1/2}\right]}{\left[u^2+\frac{u'^2}{r^2}\right]^{1/2}}; \tag{11}$$

$$F_n(u,u') = \exp\left[-j\frac{k_t L_x}{2}u\cos\varphi\right]\exp\left[-j\frac{k_t L_x}{2}\frac{u'}{r}\sin\varphi\right].$$

Equation (10) leads to

$$\begin{aligned}\int_0^2\left(\int_0^{2-u}dv\right)K(u,u')\exp\left[-j\frac{k_t L_x}{2}u\cos\varphi\right]du \\ +\int_{-2}^0\left(\int_{-u}^2 dv\right)\mathrm{K}(u,u')\exp\left[-j\frac{k_t L_x}{2}u\cos\varphi\right]du \\ =2\int_0^2(2-u)K(u,u')\cos\left[\frac{k_t L_x}{2}u\cos\varphi\right]du.\end{aligned} \tag{12}$$

In consequence, the radiation impedance simplifies into

$$Z = j\rho_0\omega\frac{L_y}{4\pi}\int_0^2\int_0^2(2-u)(2-u')K(u,u')\cos\left[\frac{k_t L_x}{2}u\cos\varphi\right]\cos\left[\frac{k_t L_x}{2}\frac{u'}{r}\sin\varphi\right]dudu'. \tag{13}$$

Explicitly, Eq. (13) reads

$$\begin{aligned}Z = j\rho_0\omega\frac{L_y}{4\pi}\int_0^2\int_0^2\frac{(2-u)(2-u')}{r_0}\left[\cos\left(\frac{k_0 L_x r_0}{2}\right)\cos\left(\frac{k_t L_x}{2}u\cos\varphi\right)\cos\left(\frac{k_t L_x}{2}\frac{u'}{r}\sin\varphi\right)\right. \\ \left.-j\sin\left(\frac{k_0 L_x r_0}{2}\right)\cos\left(\frac{k_t L_x}{2}u\cos\varphi\right)\cos\left(\frac{k_t L_x}{2}\frac{u'}{r}\sin\varphi\right)\right]dudu',\end{aligned} \tag{14}$$

with $r_0 = \sqrt{u^2+(u'^2/r)}$.

To evaluate the radiation efficiency equation (2), only the real part of Z is needed. Making the change in variables, $X=u$ and $Y=u'/r$, the real part of radiation impedance equation (14) becomes

$$\begin{aligned}\Re\mathfrak{e}[Z] = \rho_0\omega\frac{L_y}{4\pi}\int_0^2\int_0^{2/r} r \\ \times\frac{(2-X)(2-rY)\sin\left(\frac{k_0 L_x r_0}{2}\right)\cos\left(\frac{k_t L_x}{2}X\cos\varphi\right)\cos\left(\frac{k_t L_x}{2}Y\sin\varphi\right)}{r_0}dXdY\end{aligned} \tag{15}$$

A second change of variables is used to work in polar coordinates:

$$\begin{cases}X = R\cos\psi \\ X = R\sin\psi\end{cases} \rightarrow dXdY = RdRd\psi \quad \text{and} \quad R=\sqrt{X^2+Y^2}=r_0. \tag{16}$$

Using the relation $\cos(\psi-\varphi)=\cos\psi\cos\varphi+\sin\psi\sin\varphi$, Eq. (15) becomes

$$\Re\mathfrak{e}[Z] = \rho_0 \omega \frac{L_y}{4\pi} \int_0^{\pi/2} r F_R(\psi) d\psi, \tag{17}$$

with

$$F_R(\psi) = \tfrac{1}{4}(F_R(\psi,\alpha_1) + F_R(\psi,\alpha_2) + F_R(\psi,\alpha_3) + F_R(\psi,\alpha_4)) \tag{18}$$

and

$$F_R(\psi,\alpha_i) = \int_0^{R_\theta} (2 - R\cos\psi)(2 - rR\sin\psi)\sin(R\alpha_i) dR,$$

$$\alpha_{1,2} = \frac{L_x}{2}[k_0 \pm k_t \cos(\psi - \varphi)], \quad \alpha_{3,4} = \frac{L_x}{2}[k_0 \pm k_t \cos(\psi + \varphi)]. \tag{19}$$

Integral $F_R(\psi,\alpha_i)$ can be performed analytically:

$$F_R(\psi,\alpha_i) = \frac{4}{\alpha_i} + \cos(\alpha_i R_\theta)\left[-\frac{4}{\alpha_i} + 2\frac{R_\theta}{\alpha_i}[r\sin\psi + \cos\psi] + r\cos\psi\sin\psi\left(\frac{2}{\alpha_i^3} - \frac{R_\theta^2}{\alpha_i}\right)\right] + \sin(\alpha_i R_\theta)\left[-\frac{2}{\alpha_i^2}(r\sin\psi + \cos\psi) + \frac{2r}{\alpha_i^2} R_\theta \cos\psi\sin\psi\right] - \frac{2}{\alpha_i^3} r\cos\psi\sin\psi, \tag{20}$$

where

$$\left.\begin{aligned} R_\theta &= \frac{2}{\cos\theta} \quad \text{if} \quad \theta \le \theta_l \\ R_\theta &= \frac{2}{r\sin\theta} \quad \text{if} \quad \theta > \theta_l \end{aligned}\right\} \quad \theta_l = \arctan\left(\frac{1}{r}\right). \tag{21}$$

Finally, the integration limits in Eq. (17) are changed using $\mu = (4\psi/\pi) - 1$ so that a Gauss numerical integration scheme will be used to evaluate the radiation efficiency equation (2):

$$\sigma_R(k_t,\varphi) = \frac{k_0 L_y}{16} \int_{-1}^{1} r F(\mu) d\mu. \tag{22}$$

For completeness, note that the imaginary part of radiation impedance can be evaluated using the same approach. It is given by

$$\Im\mathfrak{m}[Z] = \rho_0 \omega \frac{L_y}{4\pi} \int_0^{\pi/2} r F_I(\psi), \tag{23}$$

where

$$F_I(\psi) = \tfrac{1}{4}(F_I(\psi,\alpha_1) + F_I(\psi,\alpha_2) + F_I(\psi,\alpha_3) + F_I(\psi,\alpha_4)), \tag{24}$$

$$F_I(\psi,\alpha_i) = \int_0^{R_\theta} (2 - R\cos\psi)(2 - rR\sin\psi)\cos(R\alpha_i) dR. \tag{25}$$

3. Examples

To illustrate the application of the method, the estimation of the radiation efficiency of a thin plate is first considered. Note that by replacing in Eq. (22) the trace wavenumber k_t by the

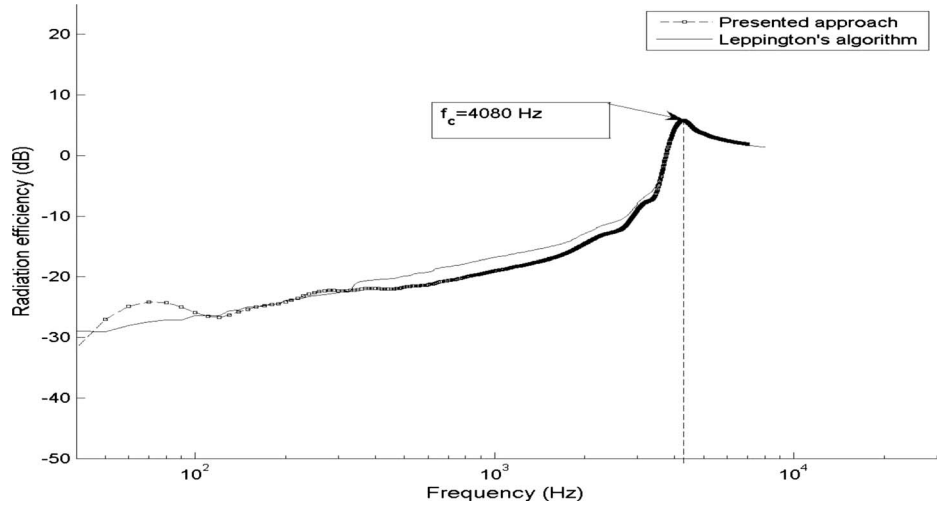


Fig. 1. Radiation efficiency of bare aluminum panel.

bending wavenumber of a structure ($\sqrt{\omega}\sqrt[4]{m/D}$ in the case of a thin plate where m is the mass per unit area and D is the bending stiffness) and averaging over heading angles, the radiation efficiency of the structure can be estimated:

$$\bar{\sigma}_R(k_b,\varphi)=\frac{1}{2\pi}\int_0^{2\pi}\sigma_R(k_b,\varphi)d\varphi. \tag{26}$$

Estimation of the radiation efficiency, calculated using Eq. (22), is compared in Fig. 1 to an estimation using Leppington's asymptotic formula (Leppington *et al.*, 1982) for a 3 mm thick 1×1 m^2 aluminum plate. Good agreement is observed apart from a slight underestimation at mid-frequencies; this is due to the differences theories used for both methods.

Next, to illustrate application of the method to a structure-borne problem, the estimation of the radiated power in the receiver side of a plate-fiber-plate system is considered. The radiated power of an infinite panel can be written in wavenumber space (Cremer *et al.*, 1988):

$$\Pi_{\text{rad}}=\frac{\langle F^2\rangle}{8\pi^2}\int_0^{2\pi}\int_0^{k_0}|v(\xi,\psi)|^2\mathfrak{Re}\ Z_\infty(\xi,\psi)\xi d\xi d\psi. \tag{27}$$

Here $\langle F^2\rangle$ denotes the excitation power spectrum and $v(\xi,\psi)$ denotes the plane wave normal velocity of the radiating panel (calculated using the TMM), and

$$Z_\infty(\xi,\psi)=\frac{k_0Z_0}{\sqrt{k_0^2-(\xi_1^2+\xi_2^2)}} \tag{28}$$

is the radiation impedance of the panel in air (the source domain), $Z_0=\rho_0c_0$ being the specific impedance in air. To account for the size effect, it is proposed here to use the same approach as for the transmission loss problem and replace the infinite panel radiation efficiency Z_∞ by Eq. (2). This transforms Eq. (27) into

$$\Pi_{\text{rad}}=\frac{\langle F^2\rangle}{8\pi^2}\int_0^{2\pi}\int_0^{\infty}|v(\xi,\psi)|^2Z_0\sigma_R(\xi,\psi)\xi d\xi d\psi. \tag{29}$$

Comparing Eq. (27) and Eq. (29), it is important to note that the infinite limit of integration for the wavenumber is important to correctly capture the behavior below the critical frequency of

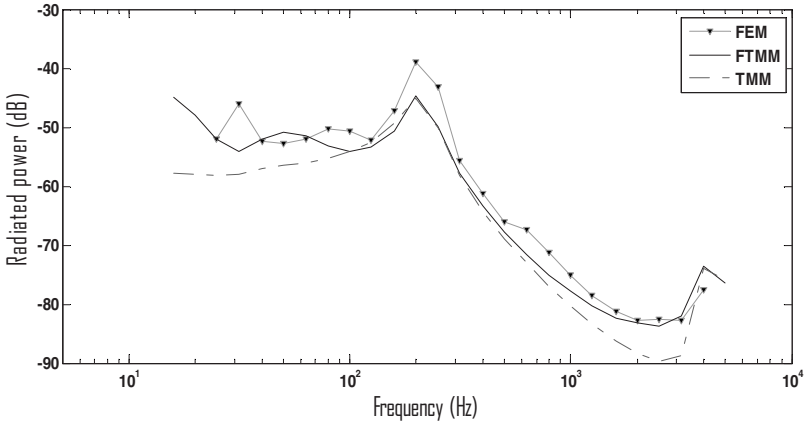


Fig. 2. Radiated power in the receiver side of a double wall structure.

the panel. Because of this double integration, the semi-analytical expression presented in Eq. (17) is important compared to a direct evaluation of Eq. (15) as done previously (Atalla *et al.*, 2006). Figure 2 shows a comparison between the radiated power calculated using Eq. (29) and finite element simulations for a 1×1 m^2 double wall structure made up a 3 cm thick fiber sandwiched between a 3 mm aluminum plate and 1 mm thick mass layer. The system is excited by a random-position point force (rain on the roof). In both the FTMM [given by Eq. (29)] and the Finite Element Method (FEM), the fiber is modeled as a limp porous layer. To approximate the rain on the roof excitation in the FEM, the results are one-third octave averaged over five randomly chosen locations. Good agreement is observed between the FEM and FTMM. The results obtained using the classical TMM [given by Eq. (27)] are also given to illustrate the importance of size correction for the radiation problem and accuracy of the presented method.

Finally, to illustrate the performance of the method for an airborne excitation, the diffuse field TL of a panel with attached foam is considered. It consists of a flat 1.64 m $\times$ 1.19 m $\times$ 1.016 mm aluminum panel with 7.62 cm thick attached foam. The description of the test and the finite element calculation are given in Atalla *et al.* (2006). The measured transmission loss and predictions with the FTMM are given in Fig. 3. Again, good agreement is observed over the whole frequency range.

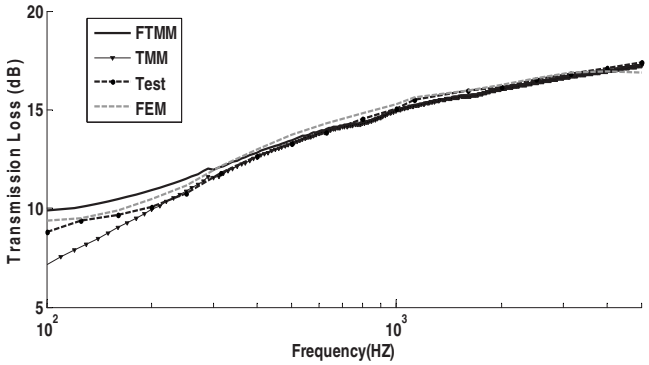


Fig. 3. Transmission loss of a plate-foam system.

4. Conclusion

A simple formulation to evaluate the Rayleigh-integral based size correction for the transfer matrix method is proposed. This simplification allows for an efficient and accurate estimation of the radiated power, radiation efficiency, and transmission loss using the classical TMM. The presented examples illustrate the relevance and accuracy of the presented method in relation to more exact and costly methods, such as the finite element method.

References and links

Allard, J. F. (**1993**). *Propagation of Sound in Porous Media: Modeling Sound Absorbing Materials* (Elsevier, New York).

Atalla, N., Sgard, F., and Amedin, C. K. (**2006**). "On the modeling of sound radiation from poroelastic materials," J. Acoust. Soc. Am. **120**, 1990–1995.

Cremer, L., Heckel, M., and Ungar, E. E. (**1988**). *Structure-Borne Sound*, 2nd ed. (Springer-Verlag, Berlin, Germany).

Ghinet, S., and Atalla, N. (**2001**). "Sound transmission loss of insulating complex structures," Can. Acoust. **29**, 26–27.

Leppington, F. G., Broadbent, E. G., and Heron, K. H. (**1982**). "The acoustic radiation efficiency from rectangular panels," Proc. R. Soc. London, Ser. A **382**, 245–271.

Villot, M., Guigou, C., and Gagliardini, L. (**2001**). "Predicting the acoustical radiation of finite size multi-layered structures by applying spatial windowing on infinite structures," J. Sound Vib. **245**, 433–455.

Recognition of interrupted sentences under conditions of spectral degradation

Monita Chatterjee[a] **and Fabiola Peredo**
Cochlear Implants and Psychophysics Laboratory, Department of Hearing and Speech Sciences, The University of Maryland, College Park, Maryland 20742
mchatterjee@hesp.umd.edu, fperedo@hesp.umd.edu

Desirae Nelson
Molecular Biology, Biochemistry, and Bioinformatics Program, Towson University, Towson, Maryland 21252
dnelso2@students.towson.edu

Deniz Başkent
Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen, Groningen, The Netherlands and School of Behavioral and Cognitive Neuroscience, University of Groningen, Groningen, The Netherlands
d.baskent@med.umcg.nl

Abstract: Cochlear implant (CI) and normally hearing (NH) listeners' recognition of periodically interrupted sentences was investigated. CI listeners' scores declined drastically when the sentences were interrupted. The NH listeners showed a significant decline in performance with increasing spectral degradation using CI-simulated, noise-band-vocoded speech. It is inferred that the success of top-down processes necessary for the perceptual reconstruction of interrupted speech is limited by even mild degradations of the bottom-up information stream (16 and 24 band processing). A hypothesis that the natural voice-pitch variations in speech would help in the perceptual reconstruction of the sentences was not supported by experimental results.

PACS numbers: 43.66.Ts, 43.71.Es, 43.71.Ky [JH]
Date Received: October 6, 2009 **Date Accepted:** December 4, 2009

1. Introduction

When normally hearing (NH) individuals listen to speech that has been periodically interrupted by silence, their performance declines relative to that with intact speech; however, speech recognition can sometimes still be remarkably good under these conditions (Miller and Licklider, 1950). If noise, or some other *plausible* masker, is now placed in the silent gaps, an illusory continuity in the target speech is perceived, and performance may actually improve (Miller and Licklider, 1950; Warren and Obusek, 1971; Powers and Wilcox, 1977; Bashford and Warren, 1987). This phenomenon is termed "phonemic restoration." The ability of the auditory system to recognize interrupted speech, the illusion of continuity in the presence of noise, and the ability to correctly reconstruct the interrupted speech clearly involve top-down processing, where the brain uses *a priori* knowledge to "fill in" the missing pieces.

It appears that this process of top-down restoration is more difficult when the target speech has been spectrally degraded or when the listener is hearing-impaired or uses a cochlear implant (CI). Thus, Nelson and Jin (2004) reported that CI listeners had great difficulty recognizing periodically interrupted sentences with gating frequencies from 1 to 32 Hz. The NH listeners attending to noise-band-vocoded speech also had difficulty with the task, although increasing the spectral resolution from 4 to 12 bands improved performance considerably. Nelson and Jin

[a]Author to whom correspondence should be addressed.

speculated that the lack of F0 resolution in spectrally degraded speech might contribute to the difficulty. Başkent (2007) also observed that CI listeners, in particular, had great difficulty in recognizing interrupted sentences, even when the interruptions were very brief. Additionally, Başkent *et al.* (in press) reported that hearing-impaired listeners had greater difficulty than NH listeners listening to interrupted sentences. Even with moderate levels of impairment, listeners were no longer able to benefit from phonemic restoration. These results suggest that the process of "filling in" is more successful when the bottom-up information is spectrally intact. In the present study, a particular focus of interest was in NH listeners' performance under conditions of *mild* spectral degradation [16 and 24 channels of noise-band-vocoded (NBV) speech, which normally allow near-perfect performance in sentence recognition tasks in quiet]. A group of CI listeners were also tested to confirm previous findings: only high-performing CI users were recruited for participation. As previous studies (e.g., Friesen *et al.*, 2001) have shown that the best CI listeners do not receive more than eight channels of spectral information, the NH listeners were also tested with eight-channel NBV speech to allow for a best-case comparison. In addition, the role of the fundamental frequency (F0) contour in listeners' ability to reconstruct the missing pieces of speech was investigated. Specifically, we hypothesized that speech intonation cues might contribute to auditory processes with which segments of interrupted speech are "strung together." Flattening the pitch contour has been shown to reduce accuracy in sentence recognition (Laures and Weismer, 1999). Here, it was hypothesized that the pitch contour helps NH listeners to "connect the dots," as it were, at either end of the intervening silence, and thus contribute to a sense of continuity in the sentence.

2. Methods

2.1 Participants

Two groups of listeners participated in this study. The first group consisted of 12 NH individuals with pure-tone thresholds less than or equal to 20 dB HL (hearing level) at audiometric frequencies of 250–8000 Hz for both ears. Their ages ranged from 20 to 30 years (mean=22). The second group consisted of six post-lingually deafened CI listeners (all users of the Cochlear Freedom or N24 devices). These participants were between ages 21–71 years (mean=60). All participants were monolingual native speakers of English.

2.2 Stimuli

Speech stimuli consisted of Hearing in Noise Test (HINT) sentence materials spoken by a single male talker (Nilsson *et al.*, 1994). PRAAT (Boersma, 2001) was used to create flat (100 Hz) F0 contours. The original HINT sentences and flattened F0-contour sentences were then noise-band-vocoded (NBV) offline into 8-, 16-, and 24-channel noise-bands using the TIGERCIS program developed by Qian-Jie Fu, Tigerspeech Technology, House Ear Institute (Los Angeles, CA; Fu, 2006). Noise-band vocoding methods were similar to those used by Shannon *et al.* (1995). The lowpass filter cutoff for temporal envelope extraction was 400 Hz (24 dB/oct) to preserve usable F0-based intonation cues from the male talker (Chatterjee and Peng, 2008; Peng *et al.*, 2009). Signal gating was applied (using MATLAB) to the NBV sentences (50% duty cycle, 100 ms on, 100 ms off, or 5 Hz gating, and 10 ms cosine-tapered rise/fall times). Preliminary work showed that the 5 Hz gating provides reasonable baseline performance with full-spectrum speech while avoiding floor and ceiling effects. Stimuli were routed from the computer sound card through a mixer and an amplifier to a single loudspeaker within the sound-attenuating booth.

2.3 Procedure

The 12 NH listeners were randomly assigned to two equal-sized groups. One group heard the original HINT sentences, while the others heard the HINT sentences with the flattened F0 contour. The task consisted of keyword-in-sentence recognition and was controlled using the I-STAR program developed by Qian-Jie Fu (Tigerspeech Technology, House Ear Institute). Sentences were presented at 65 dBA via a single loudspeaker located directly in front of the listener at a distance of 1 m. The participants were instructed to repeat back what they heard after hearing each

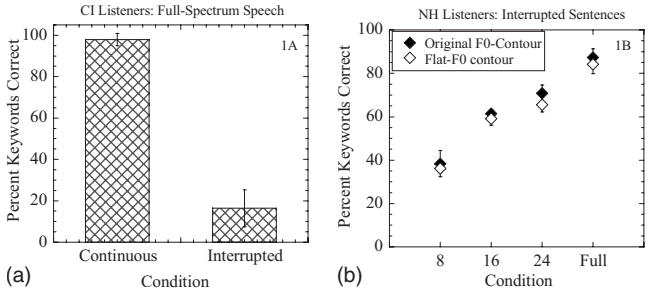


Fig. 1. (a) Sentence recognition scores (percent keyword correct) obtained by the CI listeners attending to continuous (left hand bar) and interrupted (right hand bar) sentences. Error bars show ±1 s.d. from the mean. (b) Recognition of interrupted sentences by NH listeners as a function of the degree of spectral resolution. Open and filled symbols represent scores obtained with the original and flattened F0 contours, respectively. Error bars show ±1 s.d. from the mean.

sentence. The experimenter sat outside the soundproof booth and manually marked the correct words on the computer screen. The percentage of key words correctly understood by the listener was automatically calculated by the computer program and written into a log file to be analyzed offline.

Normal-hearing listeners completed practice runs of two lists of interrupted sentences (ten sentences in each list) in each of the following four conditions: full-spectrum and 8, 16, and 24 channels of NBV speech. After the practice trial was completed, participants listened to one of the two sets of interrupted sentences (flattened F0-contour HINT or original HINT), again processed to have four levels of spectral degradation (8, 16, and 24 channels and full-spectrum). In each spectral processing condition, four lists were presented to the listener. Thus, NH participants listened to a total of 24 lists of sentences (including the 8 lists for practice trials). The order of presentation of stimulus type and condition was randomized for each participant. Care was taken to ensure that no listener heard the same sentence twice.

The CI users were tested with the original (full-spectrum) HINT sentences, either in the continuous or interrupted form, while wearing their everyday, clinically assigned speech processors. These listeners had the same practice as the NH participants and they also completed four trials of each stimulus type.

3. Results

Figure 1(a) shows the mean performance of six CI listeners attending to continuous and interrupted sentences with no F0 manipulation. It is apparent that (i) these listeners are excellent performers in open set sentence recognition and (ii) interrupting the sentences has a drastic effect on their performance. No notable differences in performance were observed between the youngest CI listeners (21 years old) and the others.

The filled symbols in Fig. 1(b) show mean results obtained with the six NH listeners who listened to the original-F0 HINT sentences, plotted as a function of the degree of spectral degradation. Note that even with eight channels of spectral information, the NH listeners' performance was better than that of the CI listeners. As the amount of spectral information increased, performance improved to about 87% correct with full-spectrum speech. However, even with 24 and 16 channels of spectral information, the NH listeners were only able to obtain on average 71% and 61% correct, respectively. With 16-channel uninterrupted speech, NH listeners are able to achieve near-perfect performance with ease (e.g., Friesen *et al.*, 2001). At best, if performance continued to improve at the same rate with increasing numbers of channels (i.e., 10% improvement per 8 channels or about 1.25% improvement per added channel), listeners would require about 37 channels to reach the same level of performance as with full-spectrum speech.

The open symbols in Fig. 1(b) show the results obtained with the listeners who attended to the flat-F0 HINT sentences. It is apparent that flattening the pitch contour had no appreciable impact on performance. A 4×2 repeated-measures analysis of variance was con-

ducted on the results obtained with the NH listeners, with the level of spectral information as a within-subjects factor and F0 contour as a between subjects factor. Not surprisingly, while the effect of spectral condition was highly significant [$F(3,30)=167.78$, $p<0.001$], no effect of F0 contour was observed, and no interaction between the two factors was found.

4. Discussion

The results presented here suggest that the top-down process that is presumably involved in filling in the missing portions of speech in interrupted-sentence recognition is significantly limited even with mild levels of spectral degradation. Thus, some minimal amount of bottom-up information is necessary for the process to succeed. It is somewhat surprising that even with 24 channels of spectral information, which would normally result in near-perfect speech recognition by NH listeners (e.g., Friesen *et al.*, 2001), the ability to restore the missing speech is impaired to the extent observed in these experiments. Even the relatively high degree of contextual cues available in HINT sentences did not appear to offset the negative effects of spectral degradation.

A second finding was that the intonation contour did not contribute to restoration of the interrupted speech under any level of spectral resolution. With 16–24 spectral channels and a 400 Hz envelope filter cutoff, sufficient F0 information is present in the speech signal to allow for good performance in tasks such as gender recognition and F0-based intonation recognition (e.g., Fu *et al.*, 2004; Gonzalez and Oliver, 2005; Chatterjee and Peng, 2008; Peng *et al.*, 2009). Thus, if the F0 contour is important in the process of reconstructing interrupted speech, we would minimally expect it to help in the perception of 16-channel, 24-channel, and full-spectrum speech and to some extent with 8-channel speech. The fact that flattening the contour did not influence performance in the task suggests that the continuity of the intonation pattern is not a contributing factor in the internal reconstruction of the interrupted speech. We note here that in unpublished preliminary results, we observed that introducing linearly increasing and decreasing F0 contours (one octave rise or fall from beginning to end) also did not impact performance in the task.

The results obtained with the CI listeners confirm the findings of Nelson and Jin (2004). The fact that even the high-performing CI patients recruited for the study found the task so difficult and underscores the difficulty of hearing with the prosthesis in everyday life. It is reasonable to speculate that part of the reason that CIs are successful is that top-down processes play an important role in reconstructing what is lost to poor spectral resolution and other potential degradations that occur in the sound transmission of the device. Further demands upon such processes, such as those that might be placed when attending to interrupted speech or speech in noise, may tax available cognitive resources excessively in most listeners.

The present results have important implications for HI listeners as well. Given the impact of even mild spectral degradation on performance with the task, it is to be expected that even listeners with mild or moderate hearing impairment would have difficulty with the perceptual restoration of interrupted sentences. Indeed, results obtained by Baskent *et al.* (in press) indicate that mildly and moderately HI listeners have increased difficulty in such tasks relative to NH individuals. Taken together, these results demonstrate clear limitations of the top-down process under conditions of mild spectral degradation and underscore the need for improved bottom-up information in auditory prostheses.

Acknowledgments

This work was supported by NIDCD Grant No. R01 DC004786 to M.C. We thank Qian-Jie Fu for the use of the software in the experiments. We are grateful to the participants for their listening time. The comments of the Associate Editor, two anonymous reviewers, and Kara C. Schvartz were very helpful in editing the manuscript.

References and links

Bashford, J. A., Jr., and Warren, R. M. (**1987**). "Multiple phonemic restorations follow the rules for auditory induction," Perception and Psychophyics **42**, 114–121.

Başkent, D. (**2007**). "Effects of amplitude ramps on phonemic restoration," in Proceedings of the 2007 Conference on Implantable Auditory Prostheses, July 16–20, Lake Tahoe, CA.

Başkent, D., Eiler, C. L., and Edwards, B. (in press, **2009**). "Perceptual restoration of speech by hearing-impaired listeners with mild to moderate sensorineural hearing loss," Hear. Res.
Boersma, P. (**2001**). "Praat, a system for doing phonetics by computer," Glot International **5**, 341–345.
Chatterjee, M., and Peng, S. C. (**2008**). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition," Hear. Res. **235**, 143–156.
Friesen, L. M., Shannon, R. V., Başkent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Am. **110**, 1150–1163.
Fu, Q.-J. (**2006**). TigerSpeech technology: Innovative speech software, version 1.05.02, available from http://www.tigerspeech.com/tst_tigercis.html/ (Last viewed Sept. 15, 2009).
Fu, Q.-J., Chinchilla, S., and Galvin, J. J. (**2004**). "The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users," J. Assoc. Res. Otolaryngol. **5**, 253–260.
Gonzalez, J., and Oliver, J. C. (**2005**). "Gender and speaker identification as a function of the number of channels in spectrally reduced speech," J. Acoust. Soc. Am. **118**, 461–470.
Laures, J. S., and Weismer, G. (**1999**). "The effects of a flattened fundamental frequency on intelligibility at the sentence level," J. Speech Lang. Hear. Res. **42**, 1148–1156.
Nelson, P. B., and Jin, S.-H. (**2004**). "Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **115**, 2286–2294.
Nilsson, M., Soli, S. I., and Sullivan, J. A. (**1994**). "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am. **95**, 1985–1099.
Miller, G. A., and Licklider, J. C. (**1950**). "The intelligibility of interrupted speech," J. Acoust. Soc. Am. **22**, 167–173.
Peng, S. C., Lu, N., and Chatterjee, M. (**2009**). "Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners," Audiol. Neuro-Otol. **14**, 327–337.
Powers, G. L., and Wilcox, J. C. (**1977**). "Intelligibility of temporally interrupted speech with and without intervening noise," J. Acoust. Soc. Am. **61**, 195–199.
Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.
Warren, R. M., and Obusek, C. J. (**1971**). "Speech perception and phonemic restorations," Percept. Psychophys. **9**, 358–362.

Ray-based blind deconvolution in ocean sound channels

Karim G. Sabra[a)]
School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405
karim.sabra@me.gatech.edu

Hee-Chun Song
Marine Physical Laboratory, Scripps Institution of Oceanography, La Jolla, California 92037
hcsong@ucsd.edu

David R. Dowling
Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109-2133
drd@engin.umich.edu

Abstract: This letter describes a ray-based blind deconvolution technique for ocean sound channels that produces broadband estimates of the source-to-array impulse response and the original source waveform from array-measured signals corrupted by (unknown) multipath propagation. The technique merely requires elementary knowledge of array geometry and sound speed at the array location. It is based on identifying a ray arrival direction to separate source waveform and acoustic-propagation phase contributions to the received signals. This technique successfully decoded underwater telecommunication sequences in the bandwidth 3–4 kHz that were broadcast 4 km in a 120-m-deep ocean sound channel without a-priori knowledge of sound channel characteristics.

PACS numbers: 43.30.Wi, 43.60.Pt, 43.30.Cq [GD]
Date Received: September 20, 2009 **Date Accepted:** December 10, 2009

1. Introduction

A signal that travels through an ocean sound channel is typically distorted when recorded by a remote hydrophone array because of interference arising from multiple propagation paths. Estimating both the source signal and the propagation characteristics (i.e., the Green's function or impulse response of the source-array geometry for a given ocean environment) from the array recordings alone—a process commonly referred to as blind deconvolution—is generally an ill-posed problem without any additional a-priori information or assumptions. This paper presents a simple and robust means to achieve blind deconvolution by extracting an acoustic ray arrival from the array measurements, using only a-priori knowledge of the sound speed profile at the array location. This ray-based blind deconvolution technique is an extension of the mode-based formulation (see Sabra and Dowling 2004), but here is applied to decoding of high-frequency (3–4 kHz) underwater telecommunication sequences recorded in shallow water on a vertical line array (VLA) (see Fig. 1).

2. Theoretical formulation of ray-based blind deconvolution

The mathematical formulation of the proposed ray-based blind deconvolution technique follows the prior mode-based formulation in Sabra and Dowling (2004) for the case of a point

[a)]Author to whom correspondence should be addressed.

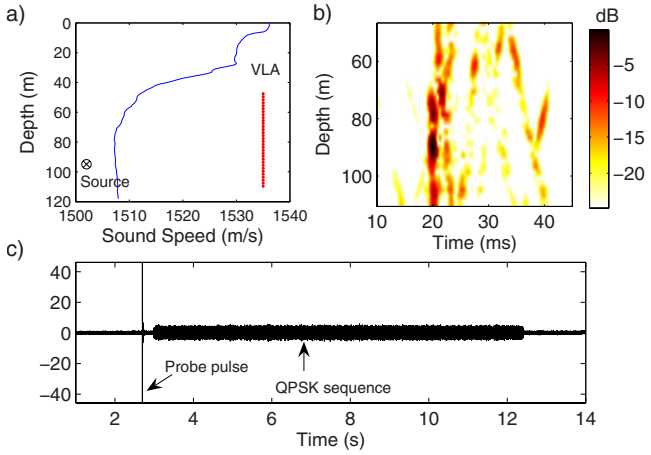


Fig. 1. (Color online) (a) Estimated depth-dependent sound-speed profile from conductivity-temperature-depth measurements between the source and VLA locations. The depths of the probe source (cross sign) and VLA receivers (filled circles, 2 m spacing) are also indicated. (b) Logarithm of the intensity of the channel impulse response received on the VLA plotted as a color image for depth and arrival time. (c) Typical QPSK communication sequence transmitted by the probe source and recorded here on one VLA receiver located at a depth of 66 m. The received sequence has been matched filtered.

source located at $\vec{r}_s$ which emits a signal $s(t)$ having a Fourier transform $\tilde{S}(\omega)=|\tilde{S}(\omega)|e^{i\Phi_s(\omega)}$, where $\Phi_s(\omega)$ is the source-signal's phase. The Fourier transform of the signal $p_j(t)$ recorded by the jth element located at $\vec{r}_j$ is given by

$$\tilde{P}_j(\omega)=\tilde{G}(\vec{r}_j,\vec{r}_s,\omega)\tilde{S}(\omega), \tag{1}$$

where $\tilde{G}(\vec{r}_j,\vec{r}_s,\omega)$ is the Fourier transform of the sound channel's impulse response (time-domain Green's function) between the source and jth element of the receiver array ($1\leq j\leq N$). The goal of the ray-based deconvolution technique is to compute estimates, $\tilde{G}_e(\vec{r}_j,\vec{r}_s,\omega)$ and $\tilde{S}_e(\omega)$, of the actual Fourier transform of the Green's function $\tilde{G}(\vec{r}_j,\vec{r}_s,\omega)$ and the actual signal $\tilde{S}(\omega)$, respectively, from $\tilde{P}_j(\omega)$ only.

For high-frequency recordings where the ray theory is applicable, incoming wavefronts associated with a specific arrival angle θ can be isolated at the output of a time-delay-beamformer $B(\omega;\theta;N)$ using N elements of the VLA:

$$B(\omega;\theta;N)=\sum_{j=1}^{N}\exp\{-i\omega\tau(\theta,\vec{r}_j)\}\tilde{P}_j(\omega)\approx|B(\omega;\theta;N)|\exp\{-i\omega T(\theta;N)+i\Phi_s(\omega)\}, \tag{2}$$

where the time-delay $\tau(\theta,\vec{r}_j)$ at the jth array element can be computed from plane-wave or more sophisticated beamforming, and $T(\theta)$ is approximately constant with respect to frequency and depends on the ray-travel time between the source and the receive array (Jensen *et al.*, 2000). For the case of a vertical line array in shallow water (see Fig. 1) with elements located at a depth z_j and range R from the source ($\vec{r}_j=[R,z_j]$), the time delays $\tau(\theta,\vec{r}_j)$ can be selected in accordance with ray group velocities to account for the curvature of the incoming wavefronts such that (Dzieciuch *et al.*, 2001; Roux *et al.*, 2008)

$$\tau(\theta,\vec{r}_j) = \int_{z_0}^{z_i} \sqrt{\frac{1}{c^2(z)} - \frac{\cos^2(\theta)}{c_0^2}}\, dz, \tag{3}$$

where c_0 is defined as the minimum of the depth-dependent sound speed profile $c(z)$ occurring at a depth z_0. Note that $\tau(\theta,\vec{r}_j)$ calculated from Eq. (3) does not require knowledge of the range R.

This ray-based deconvolution technique uses the phase $\phi(\omega;\theta;N)=\arg(B(\omega;\theta;N))$ $=\omega T(\theta;N)+\Phi_s(\omega)$ of $B(\omega;\theta;N)$ from Eq. (2) to phase-rotate the recorded data vector $\tilde{P}_j(\omega)$ $1 \le j \le N$ to remove its dependency on the unknown phase of the broadcast signal $\Phi_s(\omega)$ such that

$$\tilde{G}_e(\vec{r}_j,\vec{r}_s,\omega;\theta;N) = \tilde{P}_j(\omega)\exp\{-i\phi(\omega;\theta;N)\} = |\tilde{S}(\omega)|\tilde{G}(\vec{r}_j,\vec{r}_s,\omega)\exp\{i\omega T(\theta;N)\}. \tag{4}$$

The resulting rotated array measurements, denoted $\tilde{G}_e(\vec{r}_j,\vec{r}_s,\omega;\theta;N)$, provide an estimate of the Green's function, up to a linear frequency shift $\omega T(\theta;N)$, which is equivalent to a time-shift $T(\theta;N)$ in the time-domain, and an amplitude shading term $|\tilde{S}(\omega)|$ which may have minimal influence in the case of a (spectrally flat) broadband signal. Indeed for most underwater applications relying on the spatial coherence of the received signals across the receive array (e.g., telecommunications, source localization, and detection), obtaining an accurate estimate of the phase of $\tilde{G}(\vec{r}_j,\vec{r}_s,\omega)$ is of primary interest; amplitude estimation is typically not as crucial.

3. Application to underwater communication

An underwater telecommunication experiment (FAF-05) was conducted in July 2005 north of Elba Island, Italy with the same hardware and geometry as described in previous studies (Song *et al.*, 2007; Roux *et al.*, 2008). As shown in Fig. 1(a), the source was located at a depth of 95 m, 4.071 km away from a VLA of 32 elements (2 m spacing) with an aperture of 62 m (from 48 to 110 m depth). The communication sequence consisted of successively [see Fig. 1(c)] (1) a probe pulse [here a 150 ms linear frequency modulated (LFM) waveform from 2.5 to 4.5 kHz with a Hanning window], (2) a guard time of 150 ms (allowing the multipath to clear), and (3) a 9.4-s-long quadrature phase shift keying (QPSK) sequence of 4700 symbols (i.e., a rate of 500 symbols/s) where the basis symbol was a replica of the LFM probe. The recorded sequence on the VLA was matched filtered with the original LFM signal which yielded a gain of over 40 dB [see Fig. 1(c)]. Note that the autocorrelation of the original LFM pulse is embedded into the measured time-domain Green's functions $G(\vec{r}_j,\vec{r}_s,t)$ between the source and the VLA elements [see Fig. 1(b)] (Song *et al.*, 2007).

The measured Green's functions $G(\vec{r}_j,\vec{r}_s,t)$ reveal a complex wavefront-arrival structure with mixed reflected and refracted rays and significant multipath (over 30 ms) due to the specific nature of the sound speed profile at this site (see Roux *et al.*, 2008 for further physical description). Figure 2 (dotted line) displays the averaged energy output $\int_{\omega_1}^{\omega_2}|B(\omega;\theta;N)|^2 d\omega$ ($\omega_1/2\pi=2.5$ kHz and $\omega_2/2\pi=4.5$ kHz) of the time-delay beamformer $B(\omega;\theta;N)$ in Eq. (2) applied to the whole recorded QPSK sequence when using various sub-apertures of the VLA from $N=32$ to $N=8$ elements. The time delays $\tau(\theta,\vec{r}_j)$ were computed from Eq. (3) using the measured depth-dependent sound speed profile $c(z)$ shown in Fig. 1(a). Clear wavefronts can be identified at incoming angles around $\theta=-0.9°$, 1.8°, and 4.1° (see dashed lines in Fig. 2 for each sub-aperture of the VLA), and good agreement is obtained with the reference time-delay beamformer output (plain line) applied to the Green's function measured directly from the probe pulse [see Fig. 1(b)].

Figure 3 shows the estimated time-domain Green's function $G_e(\vec{r}_j,\vec{r}_s,t;\theta)$, computed by inverse Fourier transform of Eq. (4), using only the 9.4-s-long QPSK communication sequence and *not* the initial probe pulse. The phase of the time-delay beamformer $\phi(\omega;\theta,N)$ from Eq. (2) was used as a phase correction factor across frequencies [see Eq. (4)] for two different incoming angles identified from the beamformer output (see Fig. 2): stable low-grazing angles $\theta=-0.9°$ [Fig. 3(a)] and $\theta=-0.75°$ [Fig. 3(c)] and a less stable high-grazing angle $\theta=+4.1°$ [Figs.

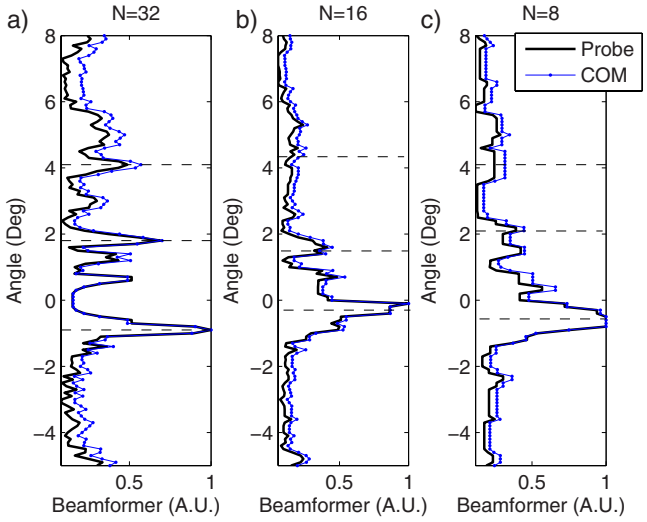


Fig. 2. (Color online) (a) Angular distribution of the recorded energy (normalized to peak value) on the VLA obtained from time-delay beamforming [see Eqs. (2) and (3)] by processing either the initial probe sequence (plain line) or the whole QPSK communication sequence (dotted line) [see Figs. 1(b) and 1(c)] using various sub-apertures of the VLA corresponding to (a) all 32 receivers, i.e., from 48 to 110 m, (b) the 16 deepest receivers, i.e., from 80 to 110 m, or (c) the 8 deepest receivers, i.e., from 100 to 110 m.

3(b) and 3(d)] that were influenced by the presence of internal waves (Roux *et al.*, 2008). Furthermore, in order to test the robustness of the ray-based blind deconvolution technique, both full [$N=32$, Figs. 3(a) and 3(b)] and partial [$N=8$, Figs. 3(c) and 3(d)] VLA apertures were used to compute the phase correction.

Overall, the estimated time-domain Green's function $G_e(\vec{r}_j,\vec{r}_s,t;\theta,N)$ (see Fig. 3) provides a $T(\theta)$-time shifted estimate of the original time-domain Green's function $G(\vec{r}_j,\vec{r}_s,t)$ that includes *all* incoming wavefronts on the receiver array [see Figs. 1(b) and 3], not just the spe-

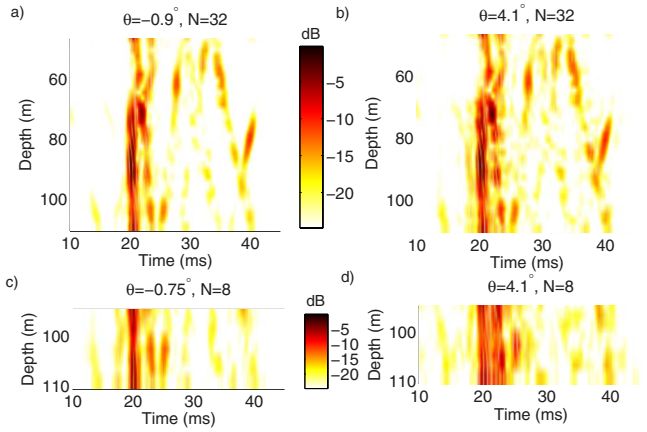


Fig. 3. (Color online) Logarithm of the intensity of the estimated channel impulse responses obtained by ray-based deconvolution of the transmitted QPSK communication sequence using Eq. (4). The intensity is plotted as a color image for depth and arrival time [see Fig. 1(b) for comparison with the measured probe source signal]. The ray-based deconvolution was performed using either the whole VLA aperture ($N=32$) for arrival angles (a) $\theta=-0.9°$ and (b) $\theta=4.1°$ or a shorter sub-aperture composed of the eight deepest VLA receivers ($N=8$) for the arrival angles (c) $\theta=-0.75°$ and (d) $\theta=4.1°$ (see Fig. 2).

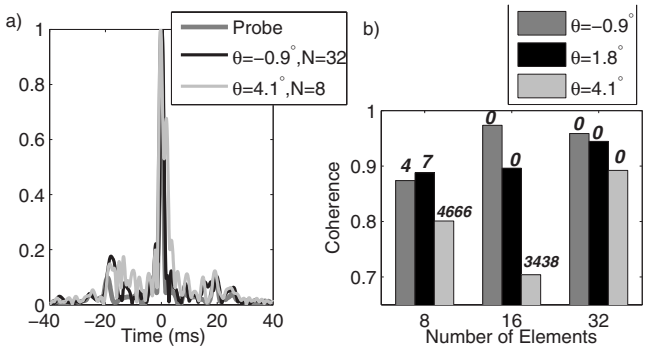


Fig. 4. (a) Normalized envelope of the summation across all $M=8$ deepest receivers located between 100 and 110 m [see Eq. (5)] of the autocorrelation time-functions of the measured channel impulse response from only the probe source signal, denoted $Q_P(t)$ (dotted line), and the cross-correlation time-functions between the channel impulse responses measured by the probe source and estimated from ray-based deconvolution ($Q_E(t;\theta;N)$ [see Eq. (5)] plain line and dashed line) using the respective two set of parameters [$N=32$, $\theta=-0.9°$, as in Fig. 3(a)] or [$N=8$, $\theta=4.1°$, as in Fig. 3(d)]. (b) Cross-correlation coefficient between the measured $Q_P(t)$ function and estimated $Q_E(t;\theta;N)$ for various combinations of the number N of elements composing the VLA sub-aperture and the arrival angle θ used to estimate the channel impulse responses from ray-based deconvolution. The number of bit errors (out of 4700 transmitted symbols) obtained when decoding the QPSK communication sequence is also indicated in italic for each configuration.

cific wavefront associated with the arrival angle θ in all cases investigated. Furthermore, as expected, better matches typically occurred when a stable ray arrival was used to compute the phase correction $\phi(\omega;\theta;N)$, especially when only a short VLA sub-aperture was used for beamforming (e.g., $N=8$), which is seen by comparing Fig. 3(c) for $\theta=-0.75°$ with Fig. 3(d) for $\theta=+4.1°$.

Finally, the quality of the Green's function estimates $G_e(\vec{r}_j,\vec{r}_s,t;\theta,N)$ was further quantified by using them for passive time-reversal (or passive phase conjugation, Dowling, 1994) of the received signal $P_j(t)$ in order to construct an estimate, $S_E(t;\theta)$, of the broadcast QPSK sequence $S(t)$:

$$S_E(t;\theta;N)=\sum_{j=1}^{M} P_j(t)*G_e(\vec{r}_j,\vec{r}_s,-t;\theta;N)=S(t)*\left[\sum_{j=1}^{M} G(\vec{r}_j,\vec{r}_s,t)*G_e(\vec{r}_j,\vec{r}_s,-t;\theta;N)\right]$$
$$=S(t)*Q_E(t;\theta,N), \quad (5)$$

where $M=8$ is the number of lowest VLA elements actually used for passive time-reversal, regardless of the number of VLA elements N (e.g., 32 or 8) used to compute the Green's function estimates $G_e(\vec{r}_j,\vec{r}_s,t;\theta;N)$. The term in the right bracket, denoted $Q_E(t;\theta;N)$, corresponds to the q function representing the summation of the cross-correlations between the measured and estimated Green's function $G_e(\vec{r}_j,\vec{r}_s,t;\theta;N)$ (Yang, 2003; Song *et al.*, 2007) [see Fig. 4(a)]. The classical q function $Q_p(t)$ is actually obtained when using the measured Green's function [see Fig. 1(b)] from the initial probe signal such as $Q_p(t)=[\Sigma_{j=1}^{M} G(\vec{r}_j,\vec{r}_s,t)*G(\vec{r}_j,\vec{r}_s,-t)]$. Figure 4(a) illustrates the impulsive nature $Q_p(t)$ which ensures the proper suppression of inter-symbol-interference due to multipath and yields ideal decoding of the broadcast QPSK sequence here (0 errors). The impulsive nature of $Q_E(t;\theta;N)$ for varying parameters N and θ is also illustrated in Fig. 4(a) and was further quantified by computing its cross-correlation coefficient with the q function $Q_p(t)$, see Fig. 4(b). High cross-correlation coefficient (>0.85) was achieved when using either a stable low-grazing angle ($\theta=-0.9°$ or $\theta=+1.8°$) or a sufficiently large aperture (e.g., $N=32$) for the time-delay beamforming so that higher-grazing angle wavefronts

($\theta = \pm 4.1°$) could be unambiguously extracted. Similar decoding performance was obtained using only the lowest four VLA elements ($M = 4$) for passive time-reversal (not shown here).

4. Summary and conclusions

The proposed ray-based blind deconvolution technique provides a means to estimate both source waveform and source-to-array impulse responses (or Green's functions) and merely requires elementary knowledge of array geometry and sound speed at the array location. Indeed only a short vertical aperture may be required if only high-frequency data are of interest. Furthermore, when successful, this technique eliminates the need for an initial probe signal and/or training sequence (used for channel estimation) as well as a guard time (used for multipath clearing) which typically precedes telecommunication sequences (Song *et al.*, 2007).

Acknowledgments

Part of this work was supported by the Office of Naval Research, Code 321. This work was performed in collaborative experiments with the NATO Underwater Research Center, La Spezia, Italy, with Mark Stevenson as Chief Scientist.

References and links

Dowling, D. R. (**1994**). "Acoustic pulse compression using passive phase conjugate processing," J. Acoust. Soc. Am. **95**, 1450–1458.

Dzieciuch, M. A., Worcester, P. F. W., and Munk, W. H. (**2001**). "Turning point filters: Analysis of sound propagation on a gyre-scale," J. Acoust. Soc. Am. **110**, 135–149.

Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H. (**2000**). *Computational Ocean Acoustics* (AIP, New York).

Roux, P., Cornuelle, B. D., Kuperman, W. A., and Hodgkiss, W. S. (**2008**). "The structure of raylike arrivals in a shallow-water waveguide," J. Acoust. Soc. Am. **124**, 3430–3439.

Sabra, K. G., and Dowling, D. R. (**2004**). "Blind deconvolution in oceanic waveguides using artificial time reversal," J. Acoust. Soc. Am. **116**, 262–271.

Song, H. C., Hodgkiss, W. S., Kuperman, W. A., Akal, T., and Stevenson, M. (**2007**). "Multiuser communications using passive time reversal," IEEE J. Ocean. Eng. **32**, 915–926.

Yang, T. (**2003**). "Temporal resolutions of time-reversed and passive-phase conjugation for underwater acoustic communications," IEEE J. Ocean. Eng. **28**, 229–245.

Structural damage identification in plates via nonlinear structural intensity maps

Fabio Semperlotti[a] **and Stephen C. Conlon**
Department of Aerospace Engineering, The Pennsylvania State University, State College, Pennsylvania 16802
fus119@psu.edu, scc135@only.arl.psu.edu

Abstract: A nonlinear structural intensity concept is presented as an approach for the identification of defects displaying nonlinear vibration behavior. The nonlinear structural dynamic response exhibited by a riveted joint with loosened fasteners connecting a stiffener with a flat panel is investigated. The excitation, generating elastic waves with dominant bending components, triggers the nonlinear contact between the plate and the stiffener inducing a dynamic response rich with nonlinear harmonics. Experimental structural intensity maps are evaluated at the super-harmonic frequencies. This technique provides an experimental approach for the characterization and two dimensional visualization of nonlinear types of defects.

PACS numbers: 43.40.Le, 43.40.Ga, 43.40.At [JM]
Date Received: August 7, 2009 **Date Accepted:** December 15, 2009

1. Introduction

Over the past 20 years, structural intensity (SI) has primarily been used as an effective tool for the development of vibration and noise control systems. The identification of preferential structural paths for optimal vibration energy transfer was regarded as powerful information to optimize the location of vibration and noise control devices. Since its original formulation[1,2] several numerical and experimental techniques were presented to evaluate the SI field in plates[3,4] and solids[5] in presence of deterministic or random excitations.[6] Only in recent years, SI was investigated as a possible tool for damage detection.[7–9] These preliminary studies investigated the relations between damage characteristics and the changes produced in the SI field. Previously, the SI damage detection approach has been implemented relying on the linear domain and on the use of a baseline signature of the healthy structure. In recent years, several studies conducted in the field of contact acoustic nonlinearity[10] investigated the potential of using the characteristic response of nonlinear behaving defects to increase the sensitivity to damage.[11,12] The nonlinear dynamic response of cracked structures under an external periodic excitation generates sub- or super-harmonics components which can be a more reliable indicator of the damage. By exploiting the nonlinear response of a cracked structure, Semperlotti *et al.*[13,14] presented a baseline free technique able to localize fatigue cracks in beams and plates. This letter is intended to present the application of the SI concept to the identification of damage exhibiting a nonlinear dynamic behavior. The evaluation of the SI at nonlinear harmonic frequencies is presented as an effective approach for the two dimensional (2D) mapping of nonlinear defects. This quantity will be referred to as nonlinear structural intensity (NSI).

2. Linear structural intensity: Theoretical background

SI is a vector field, described by magnitude and phase, indicating the path and the amplitude of the mechanical energy flowing through a vibrating structural component. The location where the mechanical energy enters the system is referred to as the *energy source* and where extracted the *energy sink*. From a general point of view, the structural power is defined as the active part associated with the vibration energy. In analytical terms, the power is represented by the time averaged product of a force with the in-phase component of the velocity in the direction of the

[a]Author to whom correspondence should be addressed.

force. The structural intensity at a prescribed point is defined as the power flow per unit area A or, equivalently, the stress σ multiplied by the complex conjugate of the velocity v^* and is measured in $\mathrm{W/m^2}$. In analytical terms, considering the harmonic response solution for a structure excited by an external periodic force:

$$I = \frac{P}{A} = \frac{1/2\ \mathrm{Re}[Fv^*]}{A} = \frac{1}{2}\ \mathrm{Re}\left[\frac{F}{A}v^*\right] = \frac{1}{2}\ \mathrm{Re}[\sigma v^*]. \tag{1}$$

The total intensity in a thin vibrating plate is due to the combined action of shear (Q_x and Q_y), bending (M_x and M_y), and twisting waves (M_{xy}), and the x component can be expressed as

$$I_x = \frac{\langle Q_x \dot{w}\rangle_t + \langle M_x \dot{\theta}_y\rangle_t - \langle M_{xy}\dot{\theta}_x\rangle_t}{h} = I_x^s + I_x^b + I_x^t, \tag{2}$$

where the superscripts s, b, and t represent the shear, bending, and twisting components, the subscript t indicates a time averaged quantity, h is the thickness of the plate, and $\dot{w}$ and $\dot{\theta}$ are the transverse and angular velocities (the orthogonal I_y component can be similarly described). Once the SI components are calculated in the two coordinate directions, the total structural intensity is given by their vector sum.

An additional quantity which, combined with the SI maps, provides further insight in the structural dynamic behavior as well as the location of the energy sources and sinks is the divergence of the SI. The divergence ∇ of the SI is a scalar quantity defined as

$$\mathrm{div(SI)} = \nabla(\mathrm{SI}) = \frac{\partial I_x}{\partial x} + \frac{\partial I_y}{\partial y}. \tag{3}$$

A positive value in the divergence field indicates an energy source while a negative value indicates an energy sink. The combined use of energy maps and divergence plots gives a clear assessment of the energy path, sources, and sinks in the whole structure.

2.1 SI experimental procedure

The experimental technique used to evaluate the SI maps is based on the methodology outlined by Arruda and Mas[4] and Daley and Hambric.[6] From a general point of view, the experimental approach is based on a combination of multi-input/multi-output (MIMO) system modeling techniques and a finite differencing scheme for the evaluation of the spatial derivatives in the SI formulation [moments and shears in Eq. (2)]. In order to apply this approach, the plate test structure is first discretized into a finite number of measurement grid points. The autospectrum of the driving force $G_{ff}(\omega)$ and the transfer functions $H_{fq}(\omega)$ and $H_{fr}(\omega)$ between the driving point f and each measurement point q, r are experimentally measured. Once these two quantities are acquired, the response cross-spectra G_{qr} between each pair of structural points can be obtained through a matrix multiplication. According to the MIMO system theory[15] the cross-spectra are given by

$$G_{qr}(\omega) = H_{fq}^{*T}(\omega) H_{fr}(\omega) G_{ff}(\omega), \tag{4}$$

where the superscript ^{*T} indicates the complex conjugate transpose. Once the experimental cross-spectra are evaluated, the data are processed using a finite differencing scheme used to calculate the spatial derivatives in Eq. (2) and, ultimately, the SI components. Several finite differencing schemes are available in the literature. In this research, a combination of a 13-point central and forward/backward finite differencing schemes was used.

3. Linear structural intensity: Stiffened panel results

The SI experimental measurements were conducted on an aluminum plate structure ($0.889 \times 0.584 \times 0.002\ 54$ m). The plate was reinforced with an I-beam stiffener located in the middle of the plate and oriented along the y axis. The plate was supported by a massive steel frame (Fig. 1). The

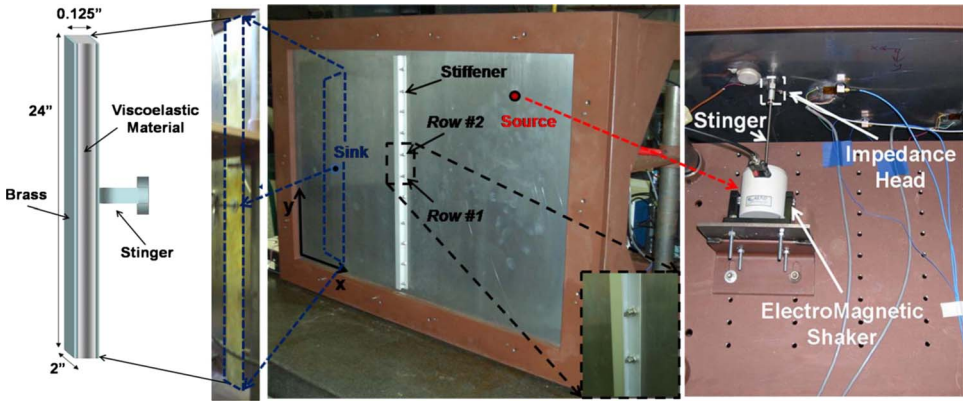


Fig. 1. (Color online) Experimental setup for the acquisition of nonlinear structural intensity maps. (Left) The energy sink was realized through a constrained layer damper. (Right) The driving condition was provided through an electromagnetic shaker acting in the out-of-plane direction. (Center) The damage was introduced by incrementally removing one and two rows of bolts.

excitation was provided in the out-of-plane direction (z axis) through an electromagnetic shaker (labeled as source in Fig. 1) while the energy sink was realized through a constrained layer damper (CLD) integrating a viscoelastic material.

The interrogation signal was a single tone periodic excitation centered at a specific frequency. In Fig. 2, examples of SI maps and the corresponding divergence fields at two different frequencies (f_1=166 Hz and f_2=343 Hz) are shown. Depending on the excitation frequency, different wave component combinations (bending, shear, and twist) are excited. The percentage of the energy transferred through the stiffener depends on both frequency and wave types. The energy flow is confined in the right hand side when the plate is driven at f_1. The off-resonance frequency f_1 excites a local bending response [Figs. 2(a) and 2(c)] which cannot transfer energy efficiently through the stiffener. At the off-resonance frequency f_2, a bending response affecting the whole plate is excited [Figs. 2(b) and 2(d)]. In this last case, vibrational energy is transferred across

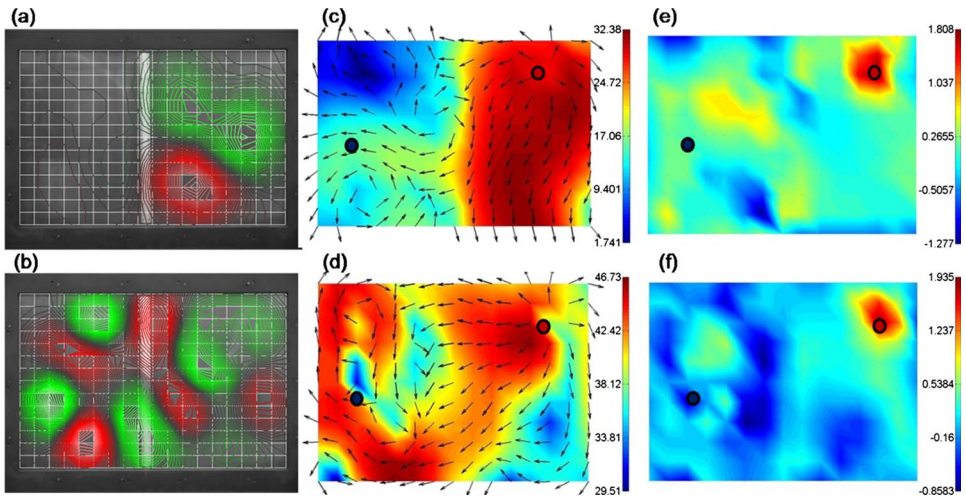


Fig. 2. (Color online) (Left) Operating deflection shapes. (Center) SI maps and (right) divergence plots corresponding to the frequencies (top) f=166 Hz and (bottom) f=343 Hz for the healthy panel (no fasteners removed). The dots indicate the location of the energy source (upper right) and sink (lower left), respectively.

the stiffener mainly by shear and bending wave types [Fig. 2(d)]. The divergence plots [Figs. 2(e) and 2(f)] clearly show the location of the energy source, but the energy sink is not clearly indicated as only a small percentage of the input vibrational energy reaches the left side for dissipation by the CLD.

4. Nonlinear structural intensity: Physical principle

In its linear formulation the SI is evaluated at the frequency of the excitation by extracting the active part of the vibrational energy at the steady state. The presence of certain types of structural damage can induce a redistribution of the vibrational energy over a broad bandwidth including super-harmonic and/or sub-harmonic frequencies. Damage such as breathing cracks, closing delaminations in composite materials, and damaged (riveted) joints are examples of nonlinear structural damage. In this study, a loosened bolted joint in a stiffened plate is used as a target fault. The damage is introduced in the form of loosened rows of bolts, as shown in Fig. 1.

Under these conditions the contact between the plate and the stiffener, in the damaged area, depends on the instantaneous relative displacement produced by the dynamic response of the two components. Under external excitation, the periodic opening and closing of the joint will generate a series of impacts between the plate and the stiffener which will result in a nonlinear structural dynamic response. The frequency content of the dynamic response is rich in nonlinear harmonics which are integer multiple (super-harmonics) or fractional multiple (sub-harmonics) of the driving frequency. The specific frequency content strictly depends on the characteristics of the excitation signal as well as the specific type of damage. In this scenario, driving the damaged plate with an out-of-plane sinusoidal excitation will generate an energy flow at multiple nonlinear frequencies. The evaluation of the structural intensity associated with the nonlinear harmonics is denominated nonlinear structural intensity (NSI). The elastic collisions periodically occurring between the stiffener and the plate at the damage location will act as an external source of excitation, injecting energy in the plate. This behavior suggests that the damage location should appear in the divergence plot as energy source (local maximum) acting at nonlinear frequencies. By analyzing the SI plots and the divergence maps, the number and the location of the energy sources could be accurately monitored to yield a 2D map of the nonlinear damage.

4.1 NSI experimental technique

The NSI concept was experimentally demonstrated by acquiring SI maps on the stiffened panel with increasing damage size. The plate was driven in the out-of-plane direction through the electromagnetic shaker, generating a sinusoidal force at a specific amplitude and frequency. The optimum amplitude and frequency able to trigger the nonlinear response were experimentally identified by sweeping through different excitation conditions. According to the experimental technique described in Sec. 2.1, the structural response was acquired in terms of accelerance frequency response functions. The experimental data were then processed through the finite difference based technique to evaluate the SI over the retained frequency bandwidth. If the system was perfectly linear, the result of a sinusoidal drive would be an energy flow only at the frequency of excitation. Due to the nonlinear damage mechanism, the vibration energy is also transferred to nonlinear frequencies, which should show evidence of energy flow. According to this mechanism, data were acquired over a wide bandwidth in order to include both the drive frequency and multiple nonlinear harmonics of the drive. Finally, data were processed at nonlinear super-harmonic frequencies following the techniques of Sec. 2.1.

5. Nonlinear structural intensity: Experimental results

Experimental results were acquired for different excitation frequencies and damage configurations. Three different conditions were considered: healthy, one row of bolts removed, and two rows of bolts removed. In this letter, the results for a $2f_e$ super-harmonic response to a driving signal at f_e=706 Hz are shown. In particular, the nonlinear SI maps and divergence plots for the different damage configurations are shown in Fig. 3. The SI results at the driving frequency are

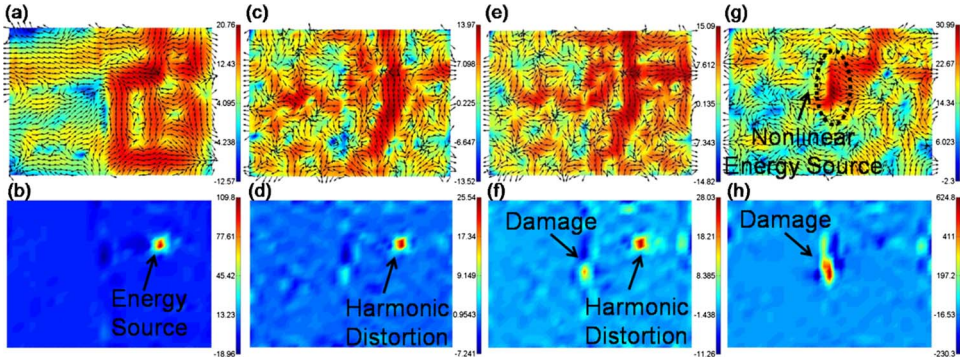


Fig. 3. (Color online) (a) SI map and (b) divergence plot at the driving frequency f_e=706 Hz for the healthy plate (no loosened bolts). SI map and divergence plot at $2f_e$=1412 Hz for the [(c) and (d)] healthy, [(e) and (f)] one-row, and [(g) and (h)] two-row damage conditions. The divergence plots show that energy is injected into the plate at the damage location due to the nonlinear contact with the stiffener. The HD effects are also clearly identified at the driving location [(e) and (f)].

shown for reference in Figs. 3(a) and 3(b). In the healthy configuration, the nonlinear contact mechanism cannot take place, and the resulting maps at super-harmonic frequencies are expected to follow the typical trend of the linear SI. This behavior is confirmed by the results in Fig. 3(c) which show only a reverberant flow. The divergence plot [Fig. 3(d)], however, shows an energy source located at the driving point. This means that energy is still injected into the system. This result can be explained by recalling the well known harmonic distortion (HD) phenomenon affecting any electronic equipment in the experimental chain. The HD alters the sinusoidal driving signal introducing super-harmonic frequencies which are not associated with the structural nonlinearities. The main result of HD is the injection of energy at frequencies which are super-harmonic components of the driving frequency. The energy associated with the HD is, however, many dB lower (40–90 dB) than the energy corresponding to the driving frequency [compare Figs. 3(b) and 3(d)]. By introducing damage in the form of one row of loosened bolts (labeled as row 1 in Fig. 1), the interrogation signal triggers the nonlinear contact response at the damage location. The periodic elastic collision occurring at the damage interface at the super-harmonic frequency results in an energy source which clearly appears in the divergence plot [Fig. 3(f)]. As for the healthy state, the harmonic distortion produces a local maximum at the driving point. The NSI maps provide the advantage of clearly differentiating the contribution of damage and HD to the overall nonlinear response. At the first damage increment, the SI map [Fig. 3(e)] shows a moderate increase in the SI at the damage interface even if the overall flow pattern is similar to the healthy state. Increasing the damage severity to two rows of loosened bolts (Fig. 1) produces a consistent increase in the amplitude of the nonlinear response. The damage is associated with a localized maximum in the SI field [Fig. 3(g)] with an intensity which is about 15 dB higher than the previous damage increment. In this configuration, energy is clearly injected at the damage interface. The divergence plot [Fig. 3(h)] clearly identifies the damage location as an energy source at the super-harmonic frequency $2f_e$. The damage is now associated with a larger footprint and a considerable increase in the divergence amplitude, highlighting the increase in the damage severity.

6. Summary and conclusions

This letter introduced the concept of NSI as a technique to produce a 2D mapping of nonlinear defects in vibrating plates. NSI maps were generated using experimentally acquired frequency response functions. The presence of a loosened structural joint makes the structural response nonlinear in nature due to the contact originating at the joint interface. The use of single tone periodic interrogation signals allows triggering the nonlinear response at the damage interface. If the SI map is acquired and plotted at super-harmonic frequencies, the resulting nonlinear SI map provides a useful tool to visualize the pattern of multiple nonlinear defects. The damage is

clearly visualized by a local maximum in the divergence plot and by an energy flow, in the SI map, originated at the damage location. The physical mechanism behind nonlinear damage results in a "virtual" energy source acting at nonlinear frequencies and originated by the elastic collisions taking place at the damage interface. The experimental results presented in this letter also confirmed the potential of using the nonlinear response characteristics of loosened structural joints as a possible indicator of structural damage. The divergence plot also provides a metric to estimate the damage and the HD contribution to the overall nonlinear response. By providing experimental evidence of the existence of nonlinear structural intensity, this study represented the first step in the development of a nonlinear structural intensity based damage detection system.

Acknowledgment

This material is based on work supported by the U.S. Army Aviation and Missile Research, Development and Engineering Center, Aviation Applied Technology Directorate under Contract No. W911W6-08-0036.

References and links

[1]D. U. Noiseux, "Measurements of power flow in uniform beams and plates," J. Acoust. Soc. Am. **47**, 238–247 (1970).
[2]G. Pavic, "Measurement of structure-borne wave intensity. I. Formulation of methods," J. Sound Vib. **49**, 221–230 (1976).
[3]J.-C. Pascal, T. Loyau, and X. Carniel, "Complete determination of structural intensity in plates using laser vibrometers," J. Sound Vib. **161**, 527–531 (1993).
[4]J. R. F. Arruda and P. Mas, "Predicting and measuring flexural power flow in plates," Proc. SPIE **2868**, 149–163 (1996).
[5]S. A. Hambric and R. P. Szwerc, "Predictions of structural intensity fields using solid finite elements," Noise Control Eng. J. **47**, 209–217 (1999).
[6]M. J. Daley and S. A. Hambric, "Simulating and measuring structural intensity fields in plates induced by spatially and temporally random excitation," J. Vibr. Acoust. **127**, 451–457 (2005).
[7]J. M. C. Dos Santos, J. R. F. Arruda, and R. M. da Cruz, "Localizing structural damage using structural intensity divergence plots," in Proceedings of the IMAC (1999), Vol. **1**, pp. 664–669.
[8]F. Semperlotti, W. T. Schmidt, E. C. Smith, S. C. Conlon, and J. C. Banks, "A parametric study for a structural intensity based damage detection technique," in AHS CBM Specialist Meeting, Huntsville, AL (2008).
[9]E. R. O. Santos, V. S. Pereira, J. R. F. Arruda, and J. M. C. Dos Santos, "Structural damage detection using energy flow models," Shock Vib. **15**, 217–230 (2008).
[10]I. Y. Solodov, "Ultrasonics of non-linear contacts: Propagation, reflection and NDE-applications," Ultrasonics **36**, 383–390 (1998).
[11]I. Solodov, J. Wackerl, K. Pfleiderer, and G. Busse, "Nonlinear self-modulation and subharmonic acoustic spectroscopy for damage detection and location," Appl. Phys. Lett. **84**, 5386–5388 (2004).
[12]V. V. Kazakov, A. Sutin, and P. A. Johnson, "Sensitive imaging of an elastic nonlinear wave-scattering source in solid," Appl. Phys. Lett. **81**, 646–648 (2002).
[13]F. Semperlotti, K. W. Wang, and E. C. Smith, "Localization of a breathing crack using super-harmonic signals due to system nonlinearity," AIAA J. **47**, 2076–2086 (2009).
[14]F. Semperlotti, "Structural damage detection via nonlinear system identification and structural intensity methods," Ph.D. thesis, The Pennsylvania State University, State College, PA (2009).
[15]J. S. Bendat and A. G. Piersol, *Random Data Analysis and Measurement Procedures* (Wiley, New York, 1986).

Orienting attention during phonetic training facilitates learning

Eric Pederson[a] **and Susan Guion-Anderson**
Department of Linguistics, University of Oregon, Eugene, Oregon 97403-1290
epederso@cas.uoregon.edu, guion@uoregon.edu

Abstract: The role of consciously directed attention toward speech input in learning has not yet been determined. Previous phonetic learning studies have manipulated acoustic signals and response feedback, but not conscious control over attentional orienting. This study tests whether directed attention facilitates learning of phonetic information. Two monolingual English-speaking groups were trained with feedback on the same auditory stimuli: Hindi words. One group was instructed to attend to the consonants and the other to the vowels. The consonant-oriented group, but not the vowel-oriented group, demonstrated post-training improvement in consonant perception, confirming a role for consciously directed attentional mechanisms during phonetic learning.

PACS numbers: 43.71.Hw, 43.71.Ft, 43.71.Es [JH]
Date Received: October 9, 2009 **Date Accepted:** December 17, 2009

1. Introduction

The purpose of the current study is to explore the role of attention in the learning of novel phonetic categories by adults. It is well known that listeners have difficulties with some non-native phonetic contrasts due to prior native-language experience. Presumably listeners have learned to allocate their attentional resources to best fit the particulars of their native language (Strange and Shafer, 2008). This acquired processing can lead to interference in learning novel phonetic contrasts. For example, Iverson and Kuhl (1995) proposed that linguistic experience warps perceptual space in that, with acoustically equidistant tokens, those categorized as good exemplars of a phonetic category cluster closely together in perceptual space, while poorer exemplars are more distant perceptually. Such warping of the perceptual space is thought to affect perceptual processing such that the relative weighting of specific acoustic cues is affected by linguistic experience. That is, perceptual sensitivity along an acoustic dimension is affected by the characteristics of one's native language (Iverson *et al.*, 2003).

It is widely believed that attention plays an important role in phonetic learning. Accurate perception of at least some non-native phonemes requires reorienting of attention. For example, Pisoni and colleagues demonstrated that laboratory training can effectively modify the ability to discriminate non-native phonetic categories and argued that sampling over the greater stimulus variability of multiple talkers helps focus the learner's attention on those acoustic cues most relevant to the novel categories (Bradlow *et al.*, 1997; Logan *et al.*, 1991). Other studies have manipulated characteristics of the stimuli on the assumption that such manipulation aids the listener in directing attention to some acoustic cues in preference to others (Jamieson and Morosan, 1986; McCandliss *et al.*, 2002). These authors have found that enhancement of phonetic features improves learning and hypothesize that this enhancement helps participants orient their attention to these features, thus improving learning. Similarly, some studies have manipulated training conditions to allow comparison of specific acoustic cues, with the assumption that listeners would attend to these cues. Christensen and Humes (1997) concluded that listeners shifted their attention according to the speech-like phonetic cues that were empha-

[a]Author to whom correspondence should be addressed.

sized by different training conditions. Francis and colleagues provided listeners with feedback about the category membership of the stimuli, and results demonstrated that trained acoustic cues were, in fact, weighted more heavily in post-training tests (Francis *et al.*, 2000; Francis and Nusbaum, 2002).

Note, however, that while each of the studies above posits an attentional component to learning and uses stimuli that promote attention to certain cues, none directly manipulate listeners' attention during training. Very few studies of phonetic learning have attempted to directly manipulate attention. Polka (1992) and Werker and Tees (1984) manipulated attention during *testing* and found no effect. In a first attempt, Guion and Pederson (2007) manipulated attention during *training* to either phonetic form or meaning and found that attention enhanced learning but they did not explore the effects of differential attention within the phonetic domain.

The current study manipulates endogenous orienting of attention through varying instructions that encourage participants to orient more strongly to one or another phonetic segment in identical stimuli. The control of orienting presumably increases (1) the likelihood of detection of the relevant phonetic information, (2) the processing in short term memory, and (3) the transfer to long-term memory (Posner, 1980; Posner and Peterson, 1990; Logan *et al.*, 1996).

Here, English monolinguals were randomly assigned to either of two groups: *consonant-attending* or *vowel-attending*. Phonetic categories were selected from the phonemic inventory of Hindi. It was hypothesized that greater attention to targeted segments of the stimuli during training would improve learning for those segments and this would be reflected in greater discrimination ability in post-training tests.

Importantly, both groups were exposed to identical auditory stimuli during training and testing. This ensures that any differences revealed by post-training testing are the results of differential attention during training rather than differential exposure during training. In order for the participants to maintain orienting either to consonants or to vowels during training, an identification task with feedback was used in which participants had to identify either the consonant or the vowel (depending on group assignment) for each presentation of a Hindi word.

2. Method

2.1 Participants

Forty-two (35 female) monolingual English speakers (mean age 22.4) participated for course credit. None had lived in a non-English-speaking region for more than 6 months, nor formally studied another language for more than 3 years, nor had any appreciable exposure to any South Asian language. All reported normal hearing. Participants were randomly assigned to one of two groups: consonant-attending ($n=21$) and vowel-attending ($n=21$).

2.2 Materials

2.2.1 Material for discrimination tests

Hindi has a number of phonetic contrasts not found in English among both vowels and consonants. Vowel discrimination was expected to be fairly accurate, as the contrasts employ duration and quality differences similar to those employed in English. On the other hand, the consonant discrimination was expected to be less accurate, as some of the contrasts employ novel phonetic distinctions known to be difficult for English speakers (Polka, 1991).

For the discrimination test, a female Hindi speaker produced minimal pair stimuli of the following Hindi initial stop contrasts: /b/-/b^{h}/, /d̪/-/ʈ/, /k/-/g/, /t̪/-/t̪^{h}/, and /t̪/-/ʈ/ and the following medial vowel contrasts: /a/-/aː/, /e/-/o/, /i/-/iː/, /o/-/a/, and /u/-/uː/. For example, /buːr/ "sawdust" vs /b^{h}uːr/ "fountain" was one of the minimal pairs recorded for the /b/-/b^{h}/ contrast and /ʈap/ "hood" vs /ʈaːp/ "hoof" was one of the minimal pairs recorded for the /a/-/aː/ contrast. Each word was said three times, and each production was edited into a separate file.

2.2.2 Materials for identification training

The training paradigm used four other native Hindi speakers (2 men, 2 women) reading 27 monosyllabic words beginning with one of eight consonants /k g ʈ ʈ^{h} ɖ b b^{h}/ and with one of eight medial vowels /a aː i iː u uː e o/. Thus, the same set of words was used for both consonant and vowel training. For example, for the word /b^{h}aːr/ "burden," the consonant-attending group was to select the symbol corresponding to /b^{h}/, while the vowel-attending group was to select the symbol corresponding to /aː/. None of the training words were the same as those used for the discrimination test. Each word was repeated three times by each speaker, digitally recorded, and edited into separate files.

2.3 Procedure

A discrimination pretest-posttest with intervening training design was used. Both groups of participants took the same pretest-posttest, which consisted of both vowel and consonant discrimination trials. Participants came to three sessions over the course of 2 weeks. In the first session, the pretest was administered, followed by identification training with feedback. The second session continued the training and in the final session, there was a final training, a short break, and then the posttest was administered. Identification training with discrimination testing was used for two reasons. (1) Identification testing was not possible because participants could not be expected to label unknown categories in the pretest nor in the posttest for the segment type (consonant or vowel) they were not trained on. (2) Identification training has been shown to improve phonetic discrimination (Wayland and Li, 2008).

2.3.1 Training

Both groups of participants were trained in an identification task using the *same* Hindi stimuli presented individually over headphones in a forced-choice identification task with feedback. The consonant-attending group was given explicit instructions to attend to word-initial consonants and the vowel-attending group to attend to word-medial vowels. The choices consisted of eight buttons displaying arbitrarily assigned non-orthographic characters. When the correct button was chosen, positive auditory feedback was given. When the incorrect button was chosen, the correct button flashed silently. There were three training sessions, each consisting of four blocks of 81 trials (the three repetitions of 27 words). Blocks alternated between the male and female speakers.

2.3.2 Discrimination test

An identical discrimination test was used for both the pre- and posttest with counterbalanced blocks. An AXB categorical discrimination procedure was used to test the five Hindi consonants and five Hindi vowel contrasts. Participants had to respond correctly to four practice trials, representing contrasts that were not subsequently tested, before taking the test.

The stimuli were presented in triads over headphones. Each triad consisted of three words, the first and third were two contrastive words constituting a minimal pair. The middle word was a different production by the same speaker of either the first or the third word. Half of these trials used minimal pairs differing by the initial consonant (e.g., buːr$_3$ buːr$_1$ b^{h}uːr$_3$) and the other half used minimal pairs differing by the medial vowel (e.g., ʈap$_2$ʈaːp$_2$ʈaːp$_1$). There was no signaling of the type of contrast tested for each trial. No feedback was given.

The test consisted of 160 trials (10 contrasts × 4 minimal pairs × 4 trials) in two pseudo-randomized, counterbalanced blocks. Each of the four possible orders (AAB, ABB, BBA, and BAA) was presented the same number of times for each contrast type. The inter-stimulus interval was 2 s and the inter-trial interval was 2 s after response. The participants were to decide whether the middle word was more like the first or the last word presented and to click with a computer mouse either of two buttons on the computer monitor labeled "first" and "last."

3. Results

3.1 Identification training

Both groups demonstrated improvement in the identification of Hindi segments across the three training sessions and, for both groups, some segment types were more difficult to identify than others. Analyses of variance (ANOVAs) ($\alpha = 0.05$) returned significant main effects of session (3) and segment type (8) for both groups. Neither group displayed an interaction between session and segment type [consonant-attending: session $F(2,480) = 81.33$, $p < 0.001$; segment type $F(7,480) = 58.58$, $p < 0.001$, session*segment type $F(14,480) = 0.59$; vowel-attending: session $F(2,480) = 80.55$, $p < 0.001$; segment type $F(7,480) = 23.88$, $p < 0.001$, session*segment type $F(14,480) = 0.23$], indicating that identification improved equally across all segments for both groups.

For the consonant-attending group, the mean correct response rate for session 1 was 57.3%, session 2 74.1%, and session 3 80.9%, with the segments ranking in increasing difficulty: /g/ /bʰ/ /b/ /d̪/ /t̪ʰ/ /k/ /ʈ/ /t̪/. For the vowel-attending group, the mean correct response rate for session 1 was 51.3%, session 2 69.5%, and session 3 75%, with the segments ranking in increasing difficulty: /iː/ /aː/ /i/ /o/ /e/ /a/ /uː/ /u/.

3.2 Pre-training discrimination test

Examining the pretest, the two groups showed no difference in their discrimination scores prior to training. Separate ANOVAs for vowel and consonant contrasts with the factors of group (consonant-attending vs vowel-attending) and contrast (five vowel and five consonant contrast types) showed no effect of group [vowels $F1(1,40) = 1.31$, $F2(1,150) = 0.71$; consonants $F1(1,40) = 2.22$, $F2(1,150) = 0.98$] nor was there an interaction between group and contrast [vowels $F1(4,160) = 1.68$, $F2(4,150) = 0.57$; consonants $F1(4,160) = 0.25$, $F2(4,150) = 0.12$]. However, the main effect of contrast was significant [vowels $F1(4,160) = 4.56$, $p = 0.002$ $F2(4,150) = 3.11$, $p = 0.017$; consonants $F1(4,160) = 49.78$, $p < 0.001$, $F2(4,150) = 18.03$, $p < 0.001$]. This indicates that some contrast types were more accurately discriminated than others for both vowel and consonant contrast types. As predicted, the combined groups responded to the vowel contrasts more accurately (97% correct) than the consonant contrasts (73% correct).[1]

3.3 Post-training evaluation

The effect of the two orienting conditions during training on discrimination of both vowels and consonants was investigated. The difference between the post- and pretest scores was submitted to a group (consonant-attending vs vowel-attending) by trial type (consonant vs vowel discrimination) ANOVA. A significant interaction between the factors was found [$F1(1,40) = 5.31$, $p = 0.026$, $F2(1,16) = 4.42$, $p = 0.052$], indicating that training conditions affected the discrimination of vowels and consonants differentially.

This interaction was investigated with separate ANOVAs for vowel and consonant discrimination contrasts. In these ANOVAS, contrast type was included as a factor to explore whether the different accuracy levels for the different contrasts observed in the pretest interacted with training condition.

For vowels, the mean posttest and pretest scores indicated no effect of training for either group (see Fig. 1). The difference in vowel scores between the posttest and the pretest was submitted to an ANOVA with factors of group and contrast ([a]-[aː], [e]-[o], [i]-[iː], [o]-[a], [u]-[uː]). There was no effect of group [$F1(1,40) = 0.16$, $F2(1,150) = 3.09$] nor contrast [$F1(4,160) = 1.38$, $F2(4,150) = 1.34$] and group and contrast did not interact [$F1(4,160) = 1.01$, $F2(4,150) = 0.59$]. This lack of a group effect for the vowels was expected given their already high pretest scores (97% correct).

However, for consonants, the difference scores did differ across the groups. The consonant-attending group demonstrated an improvement after training, whereas the vowel-attending group did not (see Fig. 1). The difference in consonant scores was submitted to an ANOVA with the factors of group and contrast ([b]-[bʰ], [d̪]-[ʈ], [k]-[g], [t̪]-[t̪ʰ], [t̪]-[ʈ]).

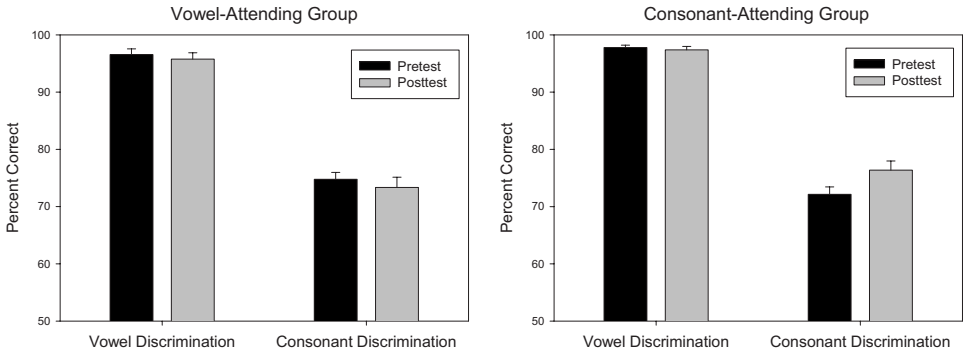


Fig. 1. Percentage correct on pretest and posttest for the consonant- and vowel-attending groups.

There was a significant effect of group [$F1(1,40)=5.03$, $p=0.03$, $F2(1,150)=4.15$, $p=0.04$] but no effect of contrast [$F1(4,160)=0.71$, $F2(4,150)=0.78$], nor was there an interaction between group and contrast [$F1(4,160)=0.47$, $F2(4,150)=1.02$]. These results indicate that orienting attention toward consonants, but not vowels, during training improved the subsequent consonant discrimination. Furthermore, the effect was not limited to individual contrast types, but was found across all contrasts.

To further test the effect of training condition on learning, each group was examined to determine whether there was a significant difference between pre- and posttest scores for consonants. The pretest and posttest consonant discrimination scores were submitted to paired, one-tailed *t*-tests. For the vowel-attending group, no significant difference between the post- and pretest scores was found [$t1(20)=-0.76$, $t2(79)=-0.26$]. For the consonant-attending group, a significant difference between the post- and pretest consonant discrimination scores was found in the predicted (positive) direction [$t1(20)=2.49$, $p=0.01$, $t2(79)=2.54$, $p=0.007$]. These results indicate that orienting attention toward consonants during training facilitated learning.

4. Discussion

Orienting attention is well known to facilitate grammatical aspects of language learning (see Schmidt, 2001 for a review). Attention has been hypothesized to facilitate phonetic learning as well (Christensen and Humes, 1997; Francis and Nusbaum, 2002), although this hypothesis has never been tested explicitly. In this study, the group instructed to attend to consonants during training demonstrated learning of consonant contrasts, whereas the group instructed to attend to vowels instead did not demonstrate learning of consonant contrasts. (Of course, since vowel discrimination was high at pretest, no learning for either group could be expected for vowels.) These results confirm that orienting attention during phonetic training facilitates learning of the specific class of stimuli to which the participants are instructed to attend.

The mechanisms of phonetic learning merit further exploration. Previous training studies have manipulated the types of stimuli (e.g., McCandliss *et al.*, 2002), but not directly manipulated the participants' attention to those stimuli. This study represents an early attempt to manipulate conscious attention during the training of phonetic categories by varying instructions to attend to certain class-specific stimuli. These instructions varied the endogenously controlled orienting of attention across the two groups. Orienting is known to generally have differential effects on the detection and manipulation of stimuli in working memory (Posner and Peterson, 1990) and/or the likelihood of transfer to long-term memory (Logan *et al.*, 1996). Thus, presumably, the attentional orienting manipulated in this study modulated the selective uptake and storage of phonetic information. Many processes are undoubtedly involved from auditory detection to final encoding and we do not yet know which specific processes are the more dependent on attentional orienting. Further development of this line of research could include the use of different phonetic contrasts, more extensive training periods, and different

methods of manipulating attention. Additionally, attention could be directed toward specific acoustic features, rather than whole segments. For now, however, it is noteworthy that the simple method of participant-directed orienting of attention can have a measurable effect on phonetic learning.

Acknowledgments

We would like to thank the five anonymous Hindi speakers for recording our speech materials, Megha Sundara for helpful comments, and Brad Biehl and Michael Peterson for their invaluable assistance running the experiment. Funding for this work was kindly provided by the University of Oregon.

References and links

[1]The higher accuracy in the vowel discrimination test than in the vowel identification training indicates that juxtaposition of different vowel tokens facilitates discrimination, whereas individually presented vowel tokens are relatively difficult to categorize.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (**1997**). "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production," J. Acoust. Soc. Am. **101**, 2299–2310.
Christensen, L. A., and Humes, L. E. (**1997**). "Identification of multidimensional stimuli containing speech cues and the effects of training," J. Acoust. Soc. Am. **102**, 2297–2310.
Francis, A. L., Baldwin, K., and Nusbaum, H. C. (**2000**). "Effects of training on attention to acoustic cues," Percept. Psychophys. **62**, 1668–1680.
Francis, A. L., and Nusbaum, H. C. (**2002**). "Selective attention and the acquisition of new phonetic categories," J. Exp. Psychol. Hum. Percept. Perform. **28**, 349–366.
Guion, S. G., and Pederson, E. (**2007**). "Investigating the role of attention in phonetic learning,"*Language Experience in Second Language Speech Learning*, edited by O.-S. Bohn and M. Munro (John Benjamins, Amsterdam), pp. 57–77.
Iverson, P., and Kuhl, P. K. (**1995**). "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling," J. Acoust. Soc. Am. **97**, 553–562.
Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., and Siebert, C. (**2003**). "A perceptual interference account of acquisition difficulties for non-native phonemes," Cognition **87**, B47–B57.
Jamieson, D. G., and Morosan, D. E. (**1986**). "Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/; contrast by francophones," Percept. Psychophys. **40**, 205–215.
Logan, G. D., Taylor, S. E., and Etherton, J. L. (**1996**). "Attention in the acquisition and expression of automaticity," J. Exp. Psychol. Learn. Mem. Cogn. **22**, 620–638.
Logan, J. S., Lively, S. E., and Pisoni, D. B. (**1991**). "Training Japanese listeners to identify English /r/ and /l/: A first report," J. Acoust. Soc. Am. **89**, 874–886.
McCandliss, B. D., Fiez, J. A., Proropapas, A., Conway, M., and McClelland, J. L. (**2002**). "Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception," Cogn. Affect. Behav. Neurosci. **2**, 89–108.
Polka, L. (**1991**). "Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions," J. Acoust. Soc. Am. **89**, 2961–2977.
Polka, L. (**1992**). "Characterizing the influence of native language experience on adult speech perception," Percept. Psychophys. **52**, 37–52.
Posner, M. I. (**1980**). "Orienting of attention," Q. J. Exp. Psychol. **32**, 2–25.
Posner, M. I., and Peterson, S. E. (**1990**). "The attention system of the human brain," Annu. Rev. Neurosci. **13**, 25–42.
Schmidt, R. (**2001**). "Attention," in *Cognition and Second Language Instruction*, edited by P. Robinson (Cambridge University Press, Cambridge), pp. 3–32.
Strange, W., and Shafer, V. L. (**2008**). "Speech perception in second language learners: The re-education of selective perception," in *Phonology and Second Language Acquisition*, edited by J. G. Hansen Edwards and M. L. Zampini (John Benjamins, Amsterdam), pp. 153–191.
Wayland, R. P., and Li, B. (**2008**). "Effects of two training procedures in cross-language perception of tones," J. Phonetics **36**, 250–267.
Werker, J. F., and Tees, R. C. (**1984**). "Phonemic and phonetic factors in adult cross-language speech perception," J. Acoust. Soc. Am. **75**, 1866–1878.

The neural code for interaural time difference in human auditory cortex

Nelli H. Salminen and Hannu Tiitinen
Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, P.O. Box 2200, FI-02015 TKK, Finland and BioMag Laboratory, Hospital District of Helsinki and Uusimaa HUSLAB, Helsinki University Central Hospital, P.O. Box 340, FI-00029 HUS, Finland
nelli.salminen@tkk.fi, hannu.tiitinen@tkk.fi

Santeri Yrttiaho
Department of Signal Processing and Acoustics, Helsinki University of Technology, P.O. Box 3000, FI-02015 TKK, Finland; Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, P.O. Box 2200, FI-02015 TKK, Finland; and BioMag Laboratory, Hospital District of Helsinki and Uusimaa HUSLAB, Helsinki University Central Hospital, P.O. Box 340, FI-00029 HUS, Finland
santeri.yrttiaho@tkk.fi

Patrick J. C. May
Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, P.O. Box 2200, FI-02015 TKK, Finland and BioMag Laboratory, Hospital District of Helsinki and Uusimaa HUSLAB, Helsinki University Central Hospital, P.O. Box 340, FI-00029 HUS, Finland
patrick.may@tkk.fi

Abstract: A magnetoencephalography study was conducted to reveal the neural code of interaural time difference (ITD) in the human cortex. Widely used crosscorrelator models predict that the code consists of narrow receptive fields distributed to all ITDs. The present findings are, however, more in line with a neural code formed by two opponent neural populations: one tuned to the left and the other to the right hemifield. The results are consistent with models of ITD extraction in the auditory brainstem of small mammals and, therefore, suggest that similar computational principles underlie human sound source localization.

PACS numbers: 43.64.Ri, 43.66.Pn, 43.64.Bt [BLM]
Date Received: December 4, 2009 **Date Accepted:** December 15, 2009

1. Introduction

Our ability to localize low-frequency sounds, for instance, speech, relies on the submillisecond interaural time difference (ITD) that results from the differential time delay for the sound to reach the two ears.[1,2] The current view on the neural extraction of ITD is based on a theoretical framework originally proposed by Jeffress.[3] According to this theory, the ITD is extracted by a system of neuronal delay lines and coincidence detectors. This results in an orderly place code representation of auditory space where each neuron is *narrowly tuned* to a preferred value of ITD. Neurophysiological studies have placed the coincidence detector system in the mammalian medial superior olive[4] (MSO).

Recently, the Jeffress model has been challenged by a hemifield model of ITD detection.[5,6] According to this view, ITD is extracted based on the relative timing of excitatory and inhibitory inputs arriving to the MSO.[5] This neuronal computation results in a *hemifield code* of auditory space where the location of a sound source is represented in the firing rates of two opponent populations: one tuned to the left and the other to the right hemifield. Neurophysiological evidence for this code has been found in small mammals[5,6] but it is unclear to what extent the results can be generalized to humans. Studies of spatial coding in the monkey[7] and human[8] auditory cortices have found receptive fields consistent with the hemifield code by utilizing free-field sound stimulation. It is, however, not clear what the contribution of ITD is in

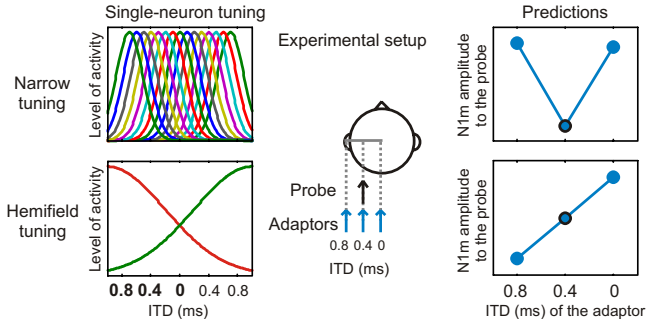


Fig. 1. (Color online) Models of ITD processing and their experimental predictions. Two alternative neuronal representations of ITD have been suggested: narrowly tuned receptive fields and hemifield tuning (left). In the present study, a probe sound was presented with ITD=0.4 ms, left-ear leading, and the amplitude of the N1m response was measured (middle). The probe presentation alternated with an adaptor sound presented with ITDs of 0, 0.4, and 0.8 ms. The alternative models of ITD coding predict distinct patterns of adaptation of the N1m response amplitude (right). The narrow tuning curves would lead to strongest attenuation when the probe and the adaptor have the same ITD (0.4 ms). The hemifield tuning, in contrast, predicts that adaptation is strongest when the adaptor is furthest to the left (0.8 ms).

these results. The free-field stimuli include a prominent interaural level difference (ILD), which is represented as a hemifield code already in the lateral superior olive (LSO) of the brainstem.[9]

The aim of the present magnetoencephalography (MEG) investigation was to find out whether the ITD is represented in the human auditory cortex by neurons narrowly tuned to ITD consistent with the output of the Jeffress model or by hemifield-tuned neurons as suggested by recent work on small mammals. For probing the ITD tuning of single neurons in the human brain, we used a stimulus-specific adaptation paradigm that capitalizes on the attenuation of activity of individual neurons after the presentation of a stimulus to which the neurons are sensitive.[10] We measured the amplitude of the N1m response that occurs at around 100 ms after stimulus onset in the event-related field and reflects the compound activity of auditory cortical areas. A probe sound with an ITD corresponding to a location in the left hemifield was presented in the context of adaptors with varying ITDs (Fig. 1). The narrow ITD tuning predicts that the strongest attenuation follows from the adaptor having the same ITD as the probe because they activate the same population of ITD-selective neurons. When the ITD of the adaptor is either shorter or longer than that of the probe, the two sounds activate distinct ITD-selective neurons and, thus, less attenuation is expected. The hemifield tuning, in contrast, predicts that the response to the probe sound arises mainly from a population tuned to the left hemifield and, therefore, the strength of the attenuation depends on the ability of the adaptor to activate this population. Consequently, the attenuation is the stronger the further to the left the adaptor is.

An additional goal of the experiment was to test the limits of ITD representation in the human auditory cortex. The ITD has a natural upper limit determined by head size. To test whether the neuronal representation reflects the limits of naturally occurring interaural delays, we measured ITD-selectivity both within and outside the upper bound of physiologically possible values.

2. Methods

Fourteen subjects participated in the study with written informed consent and the approval of the Ethical Committee of Helsinki University Central Hospital. The data of three subjects were discarded due to a poor signal-to-noise ratio. The remaining 11 subjects had a mean age of 27 (standard deviation 6, four females), reported no history of neurological or hearing disorders, and were all right-handed. During the measurements, the subjects sat in a reclining chair and were under the instruction to ignore the sound stimulation and to focus on watching a self-selected silent film.

The stimulus was a 200-ms low-pass filtered segment of white noise with the cut-off frequency at 3 kHz presented at 75 dB sound pressure level (SPL) (A). Stimuli with ITDs ranging from 0 to 2.4 ms were generated at intervals of 0.4 ms by delaying the right-ear with respect to the left-ear signal. Thus, both an onset and an ongoing ITD were imposed. The shortest two of the ITD values, 0 and 0.4 ms, are well within the physiologically determined range of delays (determined by the size of the human head). The third-smallest value, 0.8 ms, is at the physiological extreme or slightly longer. The ITD values larger than 0.8 ms are well outside the physiological range and, thus, represent ITDs that do not occur naturally. The subjects perceived all stimuli with an ITD >0 ms as lateralized the left. Sounds were presented through insert-tip earphones.

The stimulus presentation alternated between the probe and the adaptor with an onset-to-onset interstimulus interval of 1 s. In each stimulus block, a probe sound with ITD $=0.4$ ms and an adaptor with ITD $=0$, 0.4, or 0.8 ms was used (Fig. 1). Additionally, the setup was repeated for ITDs outside the physiological range, with the ITD of the probe being 0.8, 1.2, 1.6, or 2.0 ms. Again, for each probe, we used three adaptors that differed in their ITD values from that of the probe by −0.4, 0, or 0.4 ms (Fig. 3). Each probe-adaptor pair was presented within its own stimulation block with a duration of approximately 6 min. The presentation order of the blocks was counterbalanced across subjects.

Brain responses were acquired with a 306-channel whole-head MEG device (Vectorview 4-D, Neuromag Oy, Finland) with a passband from 0.03 to 200 Hz and a sampling rate of 600 Hz. A minimum of 150 responses to each sound were averaged online over a window from 100 before to 350 ms after stimulus onset. Eye-movements were monitored with electrodes, and absolute deviations larger than 150 μV or 3000 fT/cm led to automatic discarding of the epoch. The averaged brain responses were filtered at 2–30 Hz and baseline-corrected with respect to a 100-ms pre-stimulus period. To quantify the amplitude of the N1m response, the magnitudes of the vector sums from three gradiometer pairs showing the largest response amplitudes were averaged separately for the left and right hemispheres. The N1m response was identified as the waveform peak in the 80–140 ms latency region.

The peak amplitudes of the N1m responses to the probe sounds were submitted to a repeated-measures analysis of variance with cortical hemisphere, probe ITD, and adaptor ITD as the repeating factors. For further comparisons, Newman–Keuls *post-hoc* analyses and a set of planned comparisons were used. The planned comparisons were designed to evaluate the significance of the match between the experimental data and each of the model predictions and targeted responses measured to the probe with ITD $=0.4$ ms (Fig. 1). First, a contrast measuring the match to the prediction arising from narrow tuning gave weights of 1, −2, and 1 to adaptor values of 0, 0.4, and 0.8 ms, respectively. This results in large contrast values if the smallest responses are measured when the adaptor ITD is 0.4 ms and larger responses arise from other conditions. Second, a contrast measuring the match to the hemifield-tuning prediction gave weights of 1, 0, and −1 to adaptor conditions of 0, 0.4, and 0.8 ms, respectively. This contrast results in large values when the N1m response amplitude declines with increasing adaptor ITD.

3. Results

The adaptor ITD had a significant influence on the peak amplitude of the N1m response elicited by the probe (interaction between probe and adaptor ITD: $F[8,80]=2.7, p<0.02$). *Post-hoc* analyses revealed that the effect of the adaptor ITD was limited to the responses measured for the probe with an ITD of 0.4 ms, that is, to the stimulation conditions with ITDs within the physiologically plausible range (Figs. 2 and 3). No variation in the amplitude of the N1m response was found for probes with longer ITDs. This suggests that neuronal sensitivity in the cortex is limited to naturally occurring interaural delays.

For the probe sound presented within the physiological range (0.4 ms), the largest responses (i.e., the weakest attenuating effects of the adaptor) were measured with the adaptor ITD being 0 ms (Fig. 2). Response amplitudes were intermediate for the adaptor with an ITD identical to that of the probe. Finally, the smallest responses (i.e., the strongest attenuating effects) were measured when the adaptor ITD was 0.8 ms, that is, longer than the ITD of the probe

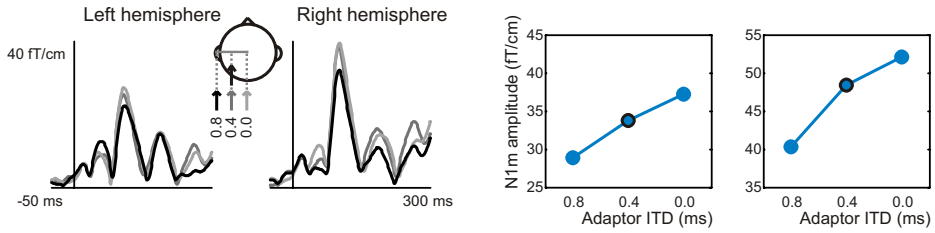


Fig. 2. (Color online) MEG results supporting hemifield tuning. MEG waveforms (left) and N1m response amplitudes (right) averaged over 11 subjects are presented for the left and right hemispheres. For the probe with an ITD of 0.4 ms, the largest N1m amplitudes were measured with the adaptor ITD of 0 ms (i.e., sound source in front) and the smallest with the adaptor ITD of 0.8 ms (i.e., sound source furthest to the left). This corresponds to the predictions based on hemifield tuning to ITD (see Fig. 1).

($p < 0.02$, compared to adaptor ITD being 0 ms). Thus, the adaptation of the response to the probe sound in the left hemifield was the stronger the further the adaptor was situated to the left. This result is consistent with the hemifield code of ITD (Fig. 1).

To further analyze whether ITD is represented in cortex through narrow or hemifield tuning, we conducted planned comparisons. These contrasts were designed to measure the extent to which the variation of the peak amplitude of the N1m response matches the hypothesized results arising from the two alternative models, with a significant contrast indicating a good match between the experimental result and the prediction. For the right-hemispheric N1m response, the contrast corresponding to hemifield tuning yielded significant results ($p < 0.02$) while that corresponding to narrow tuning was non-significant ($p > 0.4$). In the left hemisphere, the contrast corresponding to hemifield tuning approached significance ($p = 0.05$) and that corresponding to narrow tuning was again non-significant ($p > 0.8$). This provides further evidence that the results are in favor of hemifield tuning to ITD.

The ITD-specific adaptation of the N1m response in the left cortical hemisphere was similar to that in right hemisphere (Fig. 2). The right-hemispheric N1m responses were larger in amplitude than the left-hemispheric ones ($F[1,10] = 10.5, p < 0.01$) as is often found for left-lateralized sounds.[10] No interaction was found between the hemisphere and the stimulus conditions ($F[8,80] = 0.3$, $p =$ n.s.) suggesting that the two cortical hemispheres have similar neuronal selectivity to ITD.

4. Discussion

We conducted MEG measurements to reveal how the ITD cue of sound source location is represented by neurons in the human auditory cortex. The present findings are inconsistent with the narrow spatial receptive fields. If the human cortex implemented a narrow tuning to ITD, the strongest attenuation of brain responses should occur when the ITDs of the probe and the adap-

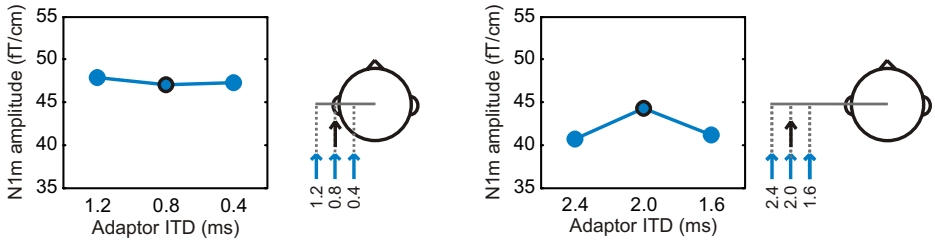


Fig. 3. (Color online) Neuronal tuning in the right hemisphere to ITDs outside the physiological range. The experimental setup was repeated for values outside the naturally occurring range of ITDs. The amplitude of the N1m response to the probes with large ITDs did not depend on the ITD of the adaptor. Note that the illustrations on the right do not depict the perceived locations of the stimuli but, rather, demonstrate the magnitudes of the delays.

tor are the same and the adaptation should be weaker when the ITDs of the two stimuli differ. Rather than observing this response pattern, we found that the attenuation was the stronger the more lateral the adaptor was. This finding is in line with an alternative, hemifield-tuning view of ITD representation. In hemifield tuning, neurons exhibit wide tuning curves with a flat maximum corresponding to sound sources either to the left or to the right. This type of tuning is consistent with the neural codes of auditory space found previously in single-unit studies of primate cortex[7] and MEG studies of the human brain[8] utilizing free-field stimuli. Therefore, it seems that in the human cortex the hemifield coding strategy applies to all horizontal sound source localization cues.

In the hemifield code, the ability of the neurons to discriminate between spatial locations is best when the sound source location corresponds to the steepest slopes of the tuning curves, that is, when the sound source is in front of the subject.[6,11] Due to the steepness of the tuning curve, a small separation between sound source locations in front produces a large change in the level of activity of the neurons. At lateral locations, however, sound sources with the same separation lead to a much smaller change in activity as the sources are mapped to the flat top portion of the tuning curve. The ITD-specific adaptation found in the present study is explained by the stimuli falling onto the slope of the tuning curve: The ITD values 0, 0.4, and 0.8 ms all corresponded to distinct heights of the ITD tuning curves and, consequently, were differently effective as adaptors. The larger ITD values extending outside the physiologically possible range, in contrast, were all similarly effective as adaptors. This suggests that the ITD tuning curves were at the same height for all of the adaptors with the unnaturally long ITDs. Thus, the ITD tuning curves seem to have their steepest slopes and therefore exhibit maximal discriminability for the range of delays that are relevant for sound source localization. This finding coincides with behavioral measures of ITD discrimination, which show a fast decline in performance when going beyond the limits of naturally occurring ITDs.[12]

The present findings have interesting parallels in recent psychophysical studies of adaptation in the human spatial hearing system. Exposure to an adaptor sound produces an aftereffect that alters the perceived source location of subsequently presented test sounds.[13,14] These effects are determined by the spatial location of the adaptor sound so that perception is always biased toward the hemifield opposite to the location of the adaptor. The effect can be induced by both ITD and ILD cues. For instance, when an adaptor is presented with an ITD corresponding to a location in the left hemisphere, the perceived locations of subsequent test sound sources are shifted toward the right from their perceived locations in conditions when no adaptors are presented. Importantly, depending on the location of the adaptor and the test sound sources (i.e., depending on whether the test sound is at a more lateral location than the adaptor), the aftereffect can lead to biases either toward or away from the adaptor location. Thus, it cannot be described in terms of the adaptor location repelling the perceived location of subsequent sounds. This can be explained in the hemifield-code framework by the laterally presented adaptor attenuating selectively one of the hemifield-tuned neuronal populations while the other one remains unaffected. Thus, the biased perception of the test sound results from the stronger neuronal activity arising from the population tuned to the hemifield opposite to the adaptor location. This finding would be very hard to account for with a model based on narrow ITD tuning but is well in line with hemifield tuning.

Here, we studied the neuronal receptive fields to ITD in the human auditory cortex. The extraction of ITD, however, occurs already in the auditory nuclei of the brainstem and, therefore, the cortical measures reflect the output of neuronal computations taking place at earlier stages of the auditory pathway. Recently, research on small mammals has led to detailed neural models of ITD extraction[5,6] whose output is consistent with the present findings on the receptive fields in human cortex. Nevertheless, the possibility remains that the hemifield tuning to ITD in the human cortex arises from some other, still unknown computational strategy realized in the auditory brainstem. To establish whether the models constructed based on small mammal data apply to human sound source localization, more detailed predictions will need to be formulated and tested in human brain imaging and behavioral studies.

Acknowledgment

This study was supported by the Academy of Finland (Project Nos. 111848, 217082, and 217113).

References and links

[1]A. W. Mills, "Lateralization of high-frequency tones," J. Acoust. Soc. Am. **32**, 132–134 (1960).
[2]F. L. Wightman and D. J. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," J. Acoust. Soc. Am. **91**, 1648–1661 (1992).
[3]L. A. Jeffress, "A place theory of sound localization," J. Comp. Physiol. Psychol. **41**, 35–39 (1948).
[4]P. X. Joris, P. H. Smith, and T. C. T. Yin, "Coincidence detection in the auditory system: 50 years after Jeffress," Neuron **21**, 1235–1238 (1998).
[5]B. Grothe, "New roles for synaptic inhibition in sound localization," Nat. Rev. Neurosci. **4**, 540–550 (2003).
[6]D. McAlpine, D. Jiang, and A. R. Palmer, "A neural code for low-frequency sound localization in mammals," Nat. Neurosci. **4**, 396–401 (2001).
[7]U. Werner-Reiss and J. M. Groh, "A rate code for sound azimuth in monkey auditory cortex: Implications for human neuroimaging studies," J. Neurosci. **28**, 3747–3758 (2008).
[8]N. H. Salminen, P. J. C. May, P. Alku, and H. Tiitinen, "A population rate code of auditory space in the human cortex," PLoS ONE **4**, e7600 (2009).
[9]J. C. Boudreau and C. Tsuchitani, "Binaural interaction in the cat superior olive S segment," J. Neurophysiol. **31**, 442–454 (1968).
[10]R. A. Butler, "The influence of spatial separation of sound sources on the auditory evoked response," Neuropsychologia **10**, 219–225 (1972).
[11]G. C. Stecker, I. A. Harrington, and J. C. Middlebrooks, "Location coding by opponent neural populations in the auditory cortex," PLoS Biol. **3**, e78 (2005).
[12]J. E. Mossop and J. F. Culling, "Lateralization of large interaural delays," J. Acoust. Soc. Am. **104**, 1574–1579 (1998).
[13]D. P. Phillips, M. E. Carmichael, and S. E. Hall, "Interactions in the perceptual processing of interaural time and level differences," Hear. Res. **211**, 96–102 (2006).
[14]B. K. Vigneault-MacLean, S. E. Hall, and D. P. Phillips, "The effects of lateralized adaptors on lateral position judgments of tones within and across frequency channels," Hear. Res. **224**, 93–100 (2007).

Verification of a near-field error sensor placement method in active control of compact noise sources

Benjamin M. Shafer,[a] Kent L. Gee, and Scott D. Sommerfeldt
Department of Physics and Astronomy, N283 ESC, Brigham Young University, Provo, Utah 84602
bshafer@seriousmaterials.com, kentgee@byu.edu, scott_sommerfeldt@byu.edu

Abstract: Recent experiments in active noise control (ANC) have used near-field error sensors whose locations are determined according to the minimization of sound power. Sensors should be placed in regions where the sound pressure reductions are the greatest during sound power minimization of the ANC system. Near-field pressure measurements of noise sources with and without ANC were made. With the error sensors in theoretically ideal locations, the measured near-field pressure map approximates the theoretical map created under the condition of minimized radiated power. Moving the error sensors to theoretically nonideal locations greatly reduces the attenuation of radiated sound power.

PACS numbers: 43.28.Ra, 43.50.Ki [MS]
Date Received: September 28, 2009 **Date Accepted:** October 29, 2009

1. Introduction

Active noise control (ANC) strategies for small axial cooling fans have demonstrated that significant global sound power reduction in blade passage tones can be achieved. Early work in this area by Quinlan[1] and Lauchle *et al.*[2] has been followed by several recent studies.[3–8] A fundamental requirement for fan ANC to result in reduced sound power radiated by the fan is that the control sources be located close to the primary source relative to a wavelength. Near collocation of the sources yields source coupling and an alteration of the power radiated by the system of sources. Nelson *et al.*[9] described a process by which the radiated power by a collection of primary and secondary point sources may be minimized by optimizing the control complex source strengths. This solution becomes a theoretical limit to the control achievable in practice with real sources modeled appropriately as collections of monopoles.

Another important consideration in achieving global control is the locations of error sensors. An adaptive algorithm, such as a multichannel filtered-x least-mean-squared algorithm (e.g., see Ref. 10), is used to minimize the squared-pressure at the error sensors. The algorithm does not ensure that the global radiation is minimized, but rather searches for the best solution that minimizes the error signals. It stands to reason that the same minimum error sensor noise level achieved for different sensor configurations can result in changed reductions in radiated power. This is true because the control signal(s) generated to minimize the noise at the error sensors will vary according to error sensor placement, resulting in different source coupling.

Where then, should error sensors be placed to minimize the noise radiated, not just at the sensor locations, but everywhere? Some authors[11,12] place the error sensors in the acoustic far field of the primary source. Hansen and Snyder[13] advocated far-field error sensing techniques by arguing that far-field noise reductions are often associated with increases in near-field levels. However, others have investigated near-field sensing techniques[3,4,8,14,15] in some cases because a compact ANC system is deemed more practical.

---

[a]Author to whom correspondence should be addressed. Also at Serious Materials, 1250 Elko Dr., Sunnyvale, CA 94089.

This paper details the results of an experimental investigation that stems directly from the previous work of Gee and Sommerfeldt.[3,4] In their work, they devised a means by which they could place error sensors in the extreme near field of the primary and secondary sources. They modeled each source as an ideal monopole and employed the method of Nelson *et al.*[9] to obtain the secondary source strengths that resulted in minimum total radiated power. They reasoned that the ideal locations for error sensors would be the locations in which the sound pressure levels experienced the greatest reductions[13] for the condition of minimized radiated power. Plots of the radiated pressure fields from the primary and optimized secondary sources revealed near-field minima in the plane of the sources. They hypothesized that placement of the error sensors in these locations would cause the squared-pressure to be minimized there, thus causing the sources to be driven in such a way so as to result in significant sound power reductions. Although Gee and Sommerfeldt demonstrated that their error sensor placement strategy resulted in greater average far-field reductions with greater consistency (see Table 1 in Ref. 4), they did not directly establish the relationship between actual near-field pressure variations and far-field reductions achieved. This letter is an experimental validation of that error sensor placement method.

2. Experimental method

To guide error sensor placement, a tool was developed that allows the user to select configurations of primary and secondary point sources, find the secondary source strengths that minimize radiated power, and create spatial pressure maps to locate potential error sensor locations.[16] This tool has been used to efficiently investigate the error sensor placement and global reduction potential for many different types of source configurations.

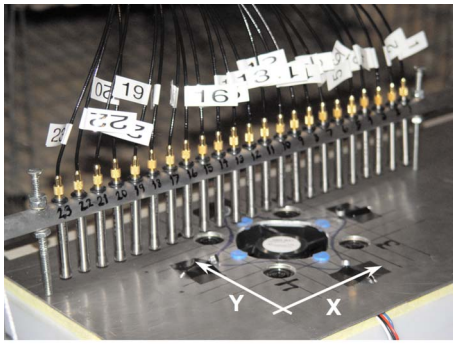
The ANC apparatus used was similar to those employed by Gee and Sommerfeldt[3,4] and Monson *et al.*[8] To explore the effect of error sensor placement on global sound power attenuation, both near-field and far-field acoustic measurements were made in a fully anechoic chamber with working dimensions of $8.71 \times 5.66 \times 5.74$ m^3. The primary sources used were a 50-mm-diameter loudspeaker and a five-bladed 60-mm axial cooling fan. Each primary source was installed in the geographic center of near- and far-field measurement grids. The fundamental radiation frequency of both sources was 622 Hz, yielding dimensionless source sizes of $ka = 0.28$ for the loudspeaker and $ka = 0.34$ for the fan. The control sources used were four 25-mm loudspeakers symmetrically located as in Fig. 1(a).

To explore the effect of error sensor placement on global sound power attenuation, both near-field and far-field acoustic measurements were made. A linear near-field measurement array consisting of 23 type-I 6.35-mm prepolarized microphones, with a 12.7-mm intermicrophone spacing, mounted in a bracket along the measurement x-axis. The array traversed along the measurement y-axis in 6.35-mm increments, resulting in a total of 1058 measurement locations over a 20.3×28.6-cm^2 aperture. The linear near-field array is shown in Fig. 1(a). The far-field array measurement was a stepper-controlled measurement boom with 13 type-I 12.7-mm prepolarized microphones in a semicircular configuration and positioned in equal area segments for calculating sound power, as shown in Fig. 1(b). The boom array was initially positioned along the x-axis (refer to Fig. 1) and was rotated in 10° increments for a total of 234 measurements.

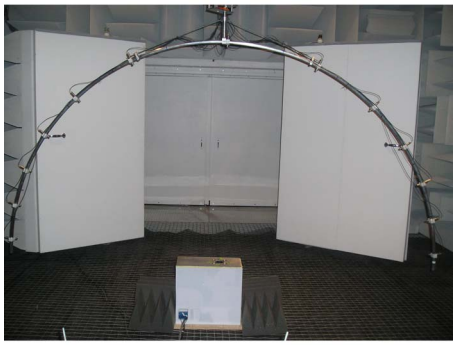
3. Measurement results

3.1 Loudspeaker near-field comparisons

For $ka = 0.28$, the loudspeaker primary source appeared to radiate essentially as a monopole source. When both the primary and control sources are treated as ideal monopoles and the geometrical configuration of the setup is input into the error sensor placement tool, the sound pressure level of the controlled field relative to the primary pressure field is shown in Fig. 2(a). What differentiates Fig. 2(a) from plots published in Refs. 4 and 8 is that the pressure fields in Fig. 2(a) are spatially averaged to approximate a 6.35-mm microphone measurement. This reduces the depth of the null that appears in between the sources (cf. Fig. 3 in Ref. 8) and shows that there are other locations in the near field outside the sources that can also lead to significant reductions, although they are not truly "ideal."



(a)



(b)

Fig. 1. (Color online) A photograph of (a) the linear near-field microphone array and aluminum plate used to install the primary and secondary sources, with white arrows and labels marking the directions of the *x*- and *y*-axes and (b) the angular far-field microphone array and noise source (bottom center).

Figures 2(b)–2(d) plot the near-field measurement of the loudspeaker and control actuators during ANC for three error sensor arrangements. Figure 2(b) shows the near-field sound pressure level relative to the primary field with ANC on and the four error sensors placed along the near-field null displayed in Fig. 2(a). The placement of the error sensors in these locations does, in fact, result in the desired coupling between the sources as evidenced by the similarity between Figs. 2(a) and 2(b). The region of closest resemblance is in and around the center of the primary source out to the locations of the control actuators. The variation between sound level reduction in Figs. 2(a) and 2(b) is greater outside of the primary source and control actuators. It is noted that the reduction near the ends of the measurement aperture was consistently less than the region in the vicinity of the sources for all tests and could be due to scattering and diffraction effects at the edges of the mock chassis.

In the test resulting in Fig. 2(c), one of the error sensors was removed from the theoretically ideal position and placed outside the sources where there should still be significant reductions. By comparing Figs. 2(b) and 2(c), this improved the control in the region to the right of the sources, while leaving the control achieved in the rest of the measured near-field region mostly the same. The variation in the sound level attenuation of Fig. 2(c) from the middle region of the theoretical map in Fig. 2(a) is greatest at the location of the theoretical pressure null to the right of the primary source.

In a more drastic test to examine the effects of error sensor placement on both near-field and global reductions, all four error sensors were placed in locations on the plate that were completely outside of any theoretically predicted pressure null. The near-field measurement for this error sensor placement, shown in Fig. 2(d), resulted in drastic changes in the near-field pressure when compared to Fig. 2(a). The common color scale reveals that although the sound

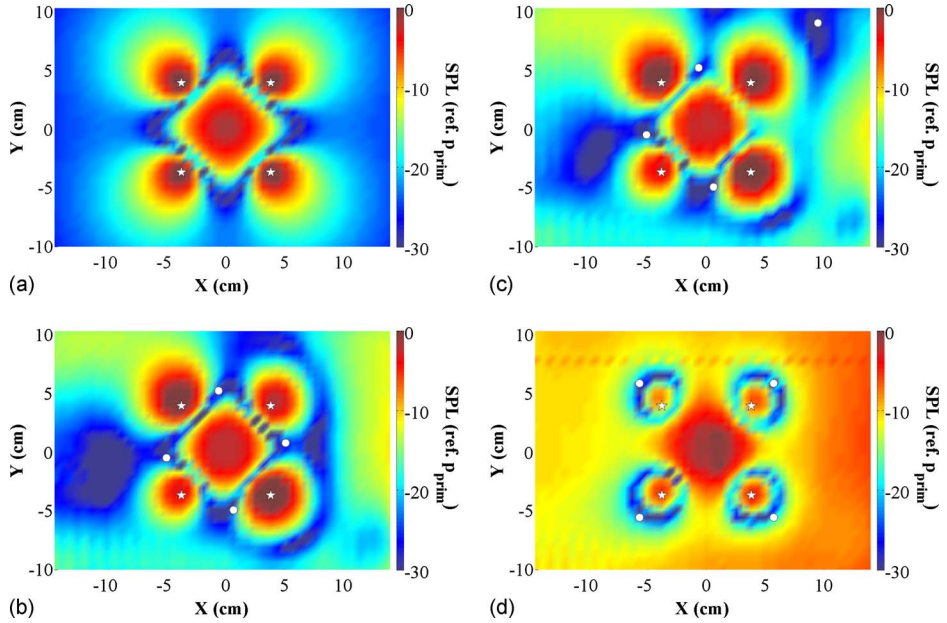


Fig. 2. (Color online) Near-field pressure maps (dB ref. 20 μPa) at 622 Hz—in a plane 0.6 cm above the source—of four-by-four ANC for (a) a theoretical point monopole and for the measurement of a 50-mm loudspeaker with (b) all error sensors placed in ideal locations, (c) three error sensors in ideal locations and one randomly placed, and (d) all error sensors in theoretical non-ideal near-field locations. Axis units are in centimeters. Circles and stars represent the locations of the error sensors and control actuators.

pressure level reduction is similar in the vicinity of the error sensors for both configurations, the average reduction in the near field is significantly greater with the microphones in theoretically ideal locations.

3.2 Far-field loudspeaker comparisons

The uncontrolled (wireframe) and controlled (surface) far-field sound pressure levels are shown in Fig. 3(a) for the loudspeaker primary source. The configuration corresponds to the near-field results shown in Fig. 2(b). In comparison, Fig. 3(b) shows the uncontrolled and controlled levels corresponding to the near-field results shown in Fig. 2(d). By moving the microphones to theoretically nonideal locations, not only is the near field significantly altered but there is a large difference in the sound power reductions achieved with the two configurations. With the error microphones in ideal locations [Fig. 3(a)], a sound power reduction of 17.1 dB was achieved.

With the error microphones in nonideal locations, the sound power was only reduced by 9.0 dB. These two results were obtained by averaging 12 trials, resulting in a standard deviation of 0.1 dB for both configurations. Not shown graphically are the angular sound pressure reductions for the configuration in Fig. 2(c), where one microphone was moved from the theoretically ideal location. Although there were some changes in the near-field and far-field sound pressure level patterns, the sound power reduction at 622 Hz was identical, 17.1 dB.

3.3 Near-field mapping and error sensor placement for an axial fan

A previous paper by Gee and Sommerfeldt[4] showed that greater global far-field reductions were achieved on average with the error microphones placed according to the theoretically ideal results of the near-field null mapping. To show that a near field similar to the theoretical field in Fig. 2(a) exists, the error sensor configuration in Fig. 2(b) was replicated for the 60-mm axial cooling fan used in Ref. 8. The near field of the axial fan without ANC was measured using the linear array and is shown in Fig. 4(a). This pressure map for 622 Hz indicates that the chassis-installed fan radiates similar to a monopole, but it is skewed in one direction.

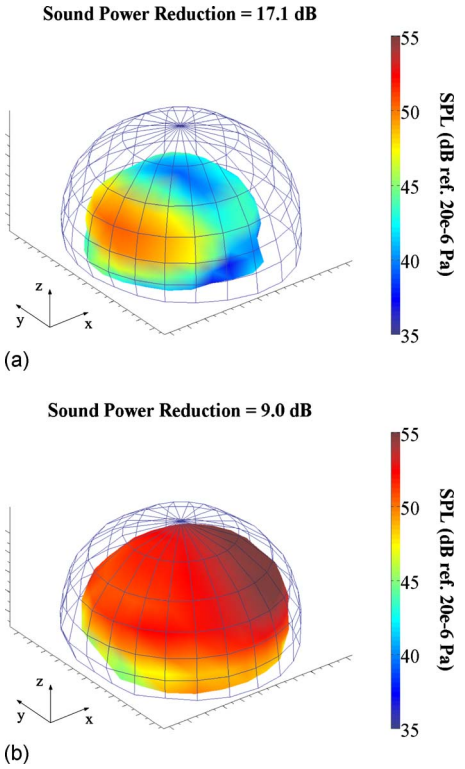


Fig. 3. (Color online) A far-field plot of the four-by-four ANC system when error sensors are located (a) according to theoretically predicted pressure nulls and (b) in theoretically non-ideal locations, with (color) and without (wire mesh) control for a 50-mm loudspeaker at 622 Hz. Axis units in dB ref. 20 μPa. The tick marks on the *z-axis* span from 0 to 55 dB (ref. 20×10^{-6} Pa). The center point on the *x*- and *y-axes* is 0 dB (ref. 20×10^{-6} Pa). The tick marks on each axis represent 5 dB (ref. 20×10^{-6} Pa).

Figure 4(b) is a measurement of the near field created by the fan and four 25-mm control loudspeakers during ANC and with each of the four error sensors placed in the same theoretically ideal locations as in Fig. 2(b). Although the overall shape of the near field resembles the theoretical prediction shown in Fig. 2(a), the nulls are not nearly as clearly accentuated as those measured for the loudspeaker in Fig. 2(b). This appears to be caused by the fact that the control at the blade passage frequency (BPF) is limited by the broadband noise floor of the fan. Nevertheless, the near field created is similar to that of the loudspeaker with the microphones in the same positions. For the sake of completeness, the sound power reduction for the fan, averaged over 12 trials, was 12.2 dB with a standard deviation of 0.3 dB.

4. Concluding discussion

The near-field measurements obtained during global ANC of a compact source have verified the error sensor positioning approach of Gee and Sommerfeldt.[4] In other words, placement of the error sensors in theoretically ideal near-field locations does, in fact, cause the adaptive control system to drive the control loudspeakers in such a way as to create the acoustic coupling that results in significant sound power reduction. Movement of the error sensors into nonideal locations can cause significant changes in the source coupling, and consequently, less attenuation in the near and far fields.

This is not to say that appreciable far-field attenuation cannot be achieved with the error sensors in theoretically nonideal locations. The modeling is based on mathematical monopoles; deviation of the sources from this behavior will cause some alteration of the near-field

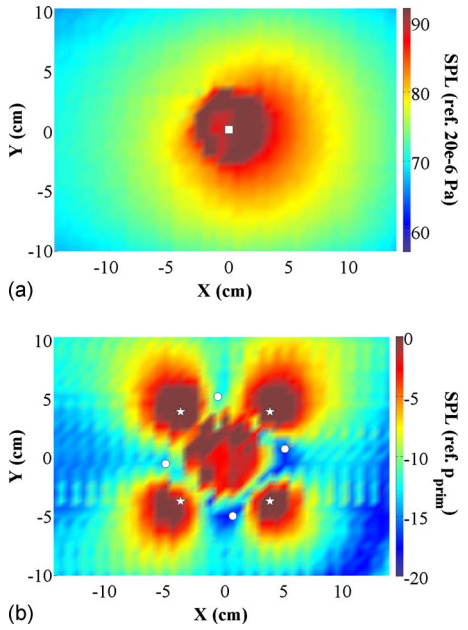


Fig. 4. (Color online) Measured near-field pressure maps at 622 Hz—in a plane 0.6 cm above the source—for (a) the 60-mm axial cooling fan without ANC (in dB ref. 20 μPa) and (b) the same fan and control actuators during four-by-four ANC (in dB ref. primary pressure field). The square marks the location of the source at the origin. Circles and stars represent the error sensors and control actuators. Note that the levels for the 2-cm radius circular region above the fan are contaminated by flow noise.

patterns. In addition, the finite widths of the error sensors will cause the region in which pressure is minimized to be somewhat different from the very thin nulls[4,8] predicted by theory. Thus, the method proposed by Gee and Sommerfeldt is useful as a systematic placement guide, but other error sensor locations [like the configuration in Fig. 2(c)] can also result in significant global reductions.

The approach of combining near-field and far-field measurements could be used as a diagnostic tool for a variety of ANC problems. In the case of the compact noise source, error sensor locations could be successively altered to demarcate regions in which global sound power reductions fall within a standard deviation of the ideal case. The near-field mapping could be used to understand in better detail the physics of the source coupling. Another possibility for this diagnostic technique would be to study the global control of noncompact noise sources. For example, although the axial fan appears to behave nominally as a monopole at the BPF, its radiation pattern will become more complex at higher harmonics. The noise source could be modeled with its multipolar characteristics and alternative ideal near-field locations sought analytically or numerically and verified experimentally.

References and links

[1]D. A. Quinlan, "Application of active control to axial flow fans," Noise Control Eng. J. **39**, 95–101 (1992).
[2]G. C. Lauchle, J. R. MacGillivray, and D. C. Swanson, "Active control of axial flow fan noise," J. Acoust. Soc. Am. **101**, 341–349 (1997).
[3]K. L. Gee and S. D. Sommerfeldt, "A compact active control implementation for axial cooling fan noise," Noise Control Eng. J. **51**, 325–334 (2003).
[4]K. L. Gee and S. D. Sommerfeldt, "Application of theoretical modeling to multichannel active control of cooling fan noise," J. Acoust. Soc. Am. **115**, 228–236 (2004).
[5]K. Homma, C. Fuller, and K. X. Man, "Broadband active-passive control of small axial fan noise emission," in Proceedings of the Noise-Con 2003 (2003), Paper No. nc03_206.
[6]Y.-J. Wong, R. Paurobally, and J. Pan, "Hybrid active and passive control of fan noise," Appl. Acoust. **64**, 885–901 (2003).
[7]J. Wang and L. Huang, "Active control of drag noise from a small axial flow fan," J. Acoust. Soc. Am. **120**,

192–203 (2006).
[8]B. B. Monson, S. D. Sommerfeldt, and K. L. Gee, "Improving compactness for active noise control of a small axial cooling fan," Noise Control Eng. J. **55**, 397–407 (2007).
[9]P. A. Nelson, A. R. D. Curtis, S. J. Elliot, and A. J. Bullmore, "The minimum power output of free field point sources and the active control of sound," J. Sound Vib. **116**, 397–414 (1987).
[10]S. D. Sommerfeldt, "Multi-channel adaptive control of structural vibration," Noise Control Eng. J. **37**, 77–89 (1991).
[11]A. J. Kempton, "The ambiguity of acoustic sources—A possibility for active control?," J. Sound Vib. **48**, 475–483 (1976).
[12]T. Martin and A. Roure, "Optimization of an active noise control system using spherical harmonics expansion of the primary field," J. Sound Vib. **201**, 577–593 (1997).
[13]C. H. Hansen and S. D. Snyder, *Active Control of Noise and Vibration* (E&FN Spon, London, 1997).
[14]A. Berry, X. Qiu, and C. H. Hansen, "Near-field sensing strategies for the active control of the sound radiated from a plate," J. Acoust. Soc. Am. **106**, 3394–3406 (1999).
[15]X. Qiu, C. H. Hansen, and X. Li, "A comparison of near-field acoustic error sensing strategies for the active control of harmonic free field sound radiation," J. Sound Vib. **215**, 81–103 (1998).
[16]S. D. Sommerfeldt and K. L. Gee, "Multi-channel active control system and method for the reduction of tonal noise from an axial fan," U.S. Patent No. 7,272,234 (September 18, 2007).

Auditory-model based robust feature selection for speech recognition

Christos Koniaris, Marcin Kuropatwinski, and W. Bastiaan Kleijn[a)]
Sound and Image Processing Laboratory, School of Electrical Engineering, KTH-Royal Institute of Technology, Osquldas väg 10, SE-100 44 Stockholm, Sweden
chris.koniaris@ee.kth.se, markurop@pg.gda.pl, bastiaan.kleijn@ee.kth.se

Abstract: It is shown that robust dimension-reduction of a feature set for speech recognition can be based on a model of the human auditory system. Whereas conventional methods optimize classification performance, the proposed method exploits knowledge implicit in the auditory periphery, inheriting its robustness. Features are selected to maximize the similarity of the Euclidean geometry of the feature domain and the perceptual domain. Recognition experiments using mel-frequency cepstral coefficients (MFCCs) confirm the effectiveness of the approach, which does not require labeled training data. For noisy data the method outperforms commonly used discriminant-analysis based dimension-reduction methods that rely on labeling. The results indicate that selecting MFCCs in their natural order results in subsets with good performance.

PACS numbers: 43.71.−k, 43.72.Qr, 43.72.Ne [DO]
Date Received: November 1, 2009 **Date Accepted:** December 7, 2009

1. Introduction

The extraction of acoustic features is an essential component of automatic speech recognition (ASR). It enables the classification of speech signals at a reasonable computational complexity based on training with speech databases of a practical size. However, the data processing inequality implies that the extraction of acoustic features from a speech signal at best preserves information relevant for phone discrimination. Thus, careful selection of the acoustic features is essential.

The existing approach for selecting features from a larger set of candidate features is based on direct optimization of classification performance, using labeled training databases. Many algorithms have been developed to select features for classification.[1–4] In ASR it is common to use *dimension-reduction* procedures,[4–6] a more general paradigm where input features are combined into a new set of lower cardinality. In general, existing feature-selection and dimension-reduction methods require classified training data. For ASR this means that dimension-reduction methods are sensitive to differences in training and testing conditions.

In this paper, a fundamentally different principle is proposed for feature *selection* for ASR: to exploit the knowledge implicit in the human auditory system. Importantly, this means that the new method does not require labeled training data. Humans perform better at speech recognition than machines, particularly for noisy environments. Recently, accurate models of the periphery have become available.[7,8] This motivates the selection of a subset of acoustic features from a larger set by maximizing the similarity of the Euclidean geometry of the selected feature set and the human auditory representation of the signal.

The implementation of our approach relies on perturbation theory. For two feature sets to perform similarly in classification, "small" Euclidean distances must be similar in the two domains (except for a scaling). The similarity of "large" distances is immaterial for the classification. The implementation is based on the so-called *sensitivity matrix*, which was first devel-

[a)]Author to whom correspondence should be addressed.

oped in the context of rate-distortion theory[9–11] and has been used for audio coding.[12]

The present work is related to the many studies on the usage of auditory models as a front-end for ASR, e.g., Refs. 13–16. The performance for such front-ends is generally robust to variations in environmental conditions. Importantly, the new approach removes the computational complexity associated with pre-processing the signal with an auditory model. It also avoids the difficulty of formatting the auditory-model output for classification.

A side outcome of our work is that it provides a measure of relative importance of a set of features. In this first study, the most commonly used set of static features, the mel-frequency cepstral coefficients (MFCCs),[17] is used. The results confirm that the human auditory model is a good guide for the selection of robust acoustic features. They also show that the initial set of MFCCs corresponds to perceptually important information.

This paper is organized as follows. Section 2 discusses a similarity measure for the perceptual and feature domains. Section 3 applies the method to ASR. Section 4 confirms with experiments that the selected features are effective and robust and Sec. 5 provides conclusions.

2. Maximizing similarity of feature and perceptual domain

Our objective is to select, from a larger set of features, a subset of features that provides a separation of sound classes that is close to that obtained by state-of-the-art auditory models. Ideally, this implies an isometry between the perceptual domain and the selected acoustic feature domain. The mapping from the perceptual domain to the acoustic feature domain would then be *distance preserving*. To obtain the best approximation to this ideal scenario, we define a new, objective criterion in this section. Thus, we avoid the *ad-hoc* nature of many auditory-system inspired features.

The motivation for the objective is that human recognition performance indicates that the human auditory periphery provides a relatively good separation of sound classes. We postulate that little information relevant for sound classification is lost in the mapping from the acoustic domain to the *perceptual domain*. It is, however, not clear if the representation is redundant.

2.1 A distance-preservation measure

It is not possible to design acoustic features that are a distance-preserving mapping from the perceptual domain. For accurate classification, the preservation of the data geometry near the class boundaries is most critical. The preservation of distances that are short relative to classification boundary curvature is important, whereas the preservation of "long" distances is not important.

Distance measures must be defined in both the perceptual and feature domains. Let $\mathbf{x}_j \in \mathbb{R}^N$ denote the N-dimensional speech signal vector characterizing a segment with time index $j \in \mathbb{Z}$ and let $\hat{\mathbf{x}}_{j,m}$ be a perturbation of $\mathbf{x}_j$ with perturbation index m. A perceptual-domain distortion is defined as a surjective mapping of two signals: $\Upsilon : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^+$, where $\mathbb{R}^+$ are the non-negative reals. Perceptual distortion measures are commonly based on the L^2 norm of the difference between the perceptual-domain signals $\mathbf{y}(\mathbf{x}_j)$ and $\mathbf{y}(\hat{\mathbf{x}}_{j,m})$, where $\mathbf{y} : \mathbb{R}^N \to \mathbb{R}^K$ is a mapping to the (K-dimensional) perceptual domain, $\Upsilon(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m}) = \|\mathbf{y}(\mathbf{x}_j) - \mathbf{y}(\hat{\mathbf{x}}_{j,m})\|^2$. This measure is the desired distance measure in the perceptual domain.

A similar distortion measure $\Gamma_i : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^+$ can be defined for the feature domain of feature set i. Let $\mathbf{c}_i : \mathbb{R}^N \to \mathbb{R}^L$ be the mapping from a signal segment $\mathbf{x}_j$ to a set of L features $\mathbf{c}_i(\mathbf{x}_j)$ with set index i. An L^2 norm based measure is then $\Gamma_i(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m}) = \|\mathbf{c}_i(\mathbf{x}_j) - \mathbf{c}_i(\hat{\mathbf{x}}_{j,m})\|^2$.

Given a finite sequence of frames $j \in \mathcal{J}$ and a finite set of acoustic perturbations $m \in \mathcal{M}_j$, the distance-preservation objective leads to the objective to find the particular set of features i that minimizes a measure of dissimilarity in the perceptual-domain distortion and the feature-domain distortion,

$$G(i) = \sum_{j \in \mathcal{J}, m \in \mathcal{M}_j} [\Upsilon(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m}) - \lambda \Gamma_i(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m})]^2, \quad (1)$$

where $\lambda = \Sigma_{j \in \mathcal{J}, m \in \mathcal{M}_j} \Upsilon(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m}) \Gamma_i(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m}) / \Sigma_{j \in \mathcal{J}, m \in \mathcal{M}_j} \Gamma_i(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m})^2$ is an optimal scaling of the acoustic feature criterion. Equation (1) can be interpreted as a measure of proximity to isometry.

2.2 Perturbation analysis

While it is possible to evaluate Eq. (1) directly even for complex distortion measures, this can be computationally expensive. For short distances, the perceptual distortion measure $\Upsilon(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m})$ and the feature-based distortion measure $\Gamma(\mathbf{x}_j, \hat{\mathbf{x}}_{j,m})$ can be approximated with simpler quadratic measures, reducing the computational complexity. The approach is based on the sensitivity matrix framework.[9–12]

The perturbation analysis for the perceptual domain and the feature domain is identical, and we only describe the first case. For notational brevity we omit the subscripts indicating frame number and perturbation where no ambiguity exists. First, let us consider $\Upsilon(\mathbf{x}, \hat{\mathbf{x}})$ to be known. We assume that $\Upsilon(\mathbf{x}, \mathbf{x}) = 0$ and that this forms a minimum. We furthermore assume that $\Upsilon(\mathbf{x}, \hat{\mathbf{x}})$ is analytic in $\hat{\mathbf{x}}$. Then, for sufficiently small perturbations $\hat{\mathbf{x}} - \mathbf{x}$, we can make the approximation

$$\Upsilon(\mathbf{x}, \hat{\mathbf{x}}) \approx [\hat{\mathbf{x}} - \mathbf{x}]^T \mathbf{D}_\Upsilon(\mathbf{x}) [\hat{\mathbf{x}} - \mathbf{x}], \quad (2)$$

where $\mathbf{D}_{\Upsilon,ij}(\mathbf{x}) = \partial^2 \Upsilon(\mathbf{x}, \hat{\mathbf{x}}) / \partial \hat{x}_i \partial \hat{x}_j |_{\hat{\mathbf{x}}=\mathbf{x}}$ is the *sensitivity matrix*.

It is common that the mapping from $\mathbf{x}$ to the perceptual or feature domain is given, rather than the distortion criterion. Consider the mapping $\mathbf{c}$ to the feature domain. If the mapping $\mathbf{c}$ is analytic, the Taylor series can be used to make a local approximation around $\mathbf{x}$:

$$\mathbf{c}(\hat{\mathbf{x}}) \approx \mathbf{c}(\mathbf{x}) + \mathbf{A}[\hat{\mathbf{x}} - \mathbf{x}], \quad (3)$$

where $\mathbf{A} = \partial \mathbf{c}(\mathbf{x}) / \partial \hat{\mathbf{x}} |_{\hat{\mathbf{x}}=\mathbf{x}}$. An L^2 distance measure in the feature domain then leads to a signal domain sensitivity matrix

$$\mathbf{D}_\Gamma(\mathbf{x}) = \mathbf{A}^T \mathbf{A}. \quad (4)$$

Thus, we can write the distortion $\Gamma(\mathbf{x}, \hat{\mathbf{x}})$ in the form of Eq. (2). The sensitivity matrix based expressions facilitate a fast evaluation of Eq. (1).

3. Application to speech recognition

In our algorithm the perceptual domain is the domain of the output vectors of the auditory model used. This section illustrates the application of the method to a specific auditory model and specific type of acoustic features.

3.1 van de Par auditory model

The van de Par[8] auditory model is a static psycho-acoustic masking model. As it uses the magnitude spectrum as input, the vector $\mathbf{x}_j$ characterizing speech segment j is now the (square-root) periodogram. The model consists of channels f, in each of which the ratio of the distortion $\hat{\mathbf{x}} - \mathbf{x}$ to masker $\mathbf{x}$ is estimated, where $\mathbf{x}$ denotes the magnitude spectrum of speech. In the end, all ratios are combined together to account for the spectral integration property of the human auditory system. The complete model is

$$\Upsilon(\mathbf{x}, \hat{\mathbf{x}}) = C_s L_e \sum_{g \in \mathcal{G}} \frac{1/N \Sigma_{f=0,\ldots,N-1} |h_{\mathrm{om}}(f)|^2 |\gamma_i(f)|^2 |x(f) - \hat{x}(f)|^2}{1/N \Sigma_{f=0,\ldots,N-1} |h_{\mathrm{om}}(f)|^2 |\gamma_i(f)|^2 |x(f)|^2 + C_a}, \quad (5)$$

where C_s and C_a are constants calibrated using measurement data, L_e is the effective duration of the segment according to the temporal integration time of the human auditory system, the inte-

ger g labels the gamma-tone filter and $\mathcal{G}$ the set of gammatone filters considered, h_{om} is the outer and middle ear transfer function which is the inverse of the threshold in quiet, and γ_i is the ith gammatone filter.

Combining Eqs. (2) and (5), the sensitivity matrix $\mathbf{D}_{\Upsilon}(\mathbf{x})$ can be obtained. It is a diagonal matrix with the diagonal element for row and column f given by

$$\mathbf{D}_{\Upsilon,ff}(\mathbf{x}) \approx 2C_s L_e \sum_i \frac{1/N\Sigma_f |h_{\mathrm{om}}(f)|^2 |\gamma_i(f)|^2}{1/N\Sigma_f |h_{\mathrm{om}}(f)|^2 |\gamma_i(f)|^2 |x(f)|^2 + C_a}. \tag{6}$$

3.2 Local linearization of the MFCCs

In our experiments, the MFCCs (Ref. 17) were used since they are the most commonly used acoustic features. The MFCCs are defined as

$$\mathbf{c}(q) = \sum_{m=0}^{M-1} \ln\left\{\sum_{n=0}^{N-1} \mathbf{x}(n)\mathbf{H}_m(n)\right\} \cos\left\{q\left[m-\frac{1}{2}\right]\frac{\pi}{M}\right\}, \quad q = 1, \ldots, Q, \tag{7}$$

where $\mathbf{x}(n)$ is the periodogram, $\mathbf{H}_m(n)$ is the mth triangular mel-filter, m is the filter index, M is the number of triangular bandpass filters used, and Q is the number of cepstrum coefficients.

Section 2.2 introduced the matrix $\mathbf{A}$ that characterizes the local relation between the features and the signal $\mathbf{x}$. For the MFCCs, the matrix $\mathbf{A}$ is

$$\mathbf{A}_{qn} = \sum_{m=0}^{M-1} \cos\left\{q\left[m-\frac{1}{2}\right]\frac{\pi}{M}\right\} \frac{\mathbf{H}_m(n)}{\Sigma_{n=0}^{N-1}\mathbf{x}(n)\mathbf{H}_m(n)}. \tag{8}$$

3.3 Overview of the algorithm

We now outline the computation of the measure of proximity to isometry. Given an unlabeled database of speech, we compute for each of a large set of frames the periodogram $\mathbf{x}$. We compute the sensitivity matrix $\mathbf{D}_{\Upsilon}(\mathbf{x})$ using Eq. (6). We also compute the sensitivity $\mathbf{D}_{\Gamma}(\mathbf{x})$ matrix equation (4) using the Jacobian $\mathbf{A}$ of Eq. (8). For each frame we then create a large number of small random perturbations $\hat{\mathbf{x}}$ of $\mathbf{x}$ and evaluate $\Upsilon(\mathbf{x},\hat{\mathbf{x}})$ using Eq. (2) and $\Gamma(\mathbf{x},\hat{\mathbf{x}})$ using a corresponding equation. Finally we evaluate Eq. (1) for all perturbations and all frames simultaneously.

4. Experimental results

This section examines the plausability of the linearity assumption used in the perturbation method and verifies the robust performance of the selected feature sets. All experiments were performed on MFCCs.

The MFCCs were extracted from the AURORA2 (Ref. 18) database, sampled at 8 kHz, using a Hamming window of 25 ms with an overlap of 12.5 ms. The discrete Fourier transform (DFT) dimensionality was 256 and the number of filters used was 23. A set of 12 conventional MFCCs was extracted.

4.1 Range of linearity

The range of validity for the linearization assumption between the cepstrum and the speech was examined first. The speech was distorted with independently identically distributed (i.i.d.) Gaussian noise at different signal-to-noise ratios (SNRs) ranging from 30 to 90 dB with a step of 1. We compared the change in the features computed from the linearized relation [Eq. (3)] with the true difference between the cepstra of the original and distorted signals for all the MFCCs. We found that the linearity assumption is accurate at a scale that is meaningful for sound discrimination. Outliers occur but they result from time segments where the power of the signal is low, as can be seen from Eq. (8).

Table 1. AURORA2 clean training results. The amfs selected coefficients with indices 1, 2, 3, 4, 9, 10, 11, and 12 for $n=8$ and 1, 2, 3, and 12 for $n=4$, respectively.

Feature set		Data test set A				
		Clean 1	Clean 2	Clean 3	Clean 4	Avg.
12+E	Full, static	98.1	97.6	97.7	98.1	97.9
12+E+V+A	Full, dynamic	98.9	99.1	98.8	99.2	99.0
8+E	amfs	97.8	97.6	97.1	97.8	97.6
	lda	98.3	97.8	98.4	98.3	98.2
	hlda	98.7	98.1	98.5	99.0	98.6
	5-rsfs	95.8	95.7	95.5	96.0	95.8
	initial	97.5	97.3	97.6	97.9	97.6
8+E+V+A	amfs	99.1	98.9	98.8	99.2	99.0
	lda	98.4	98.1	98.2	98.6	98.3
	hlda	98.6	98.5	98.5	99.1	98.7
	initial	98.9	98.9	98.8	99.1	98.9
4+E	amfs	96.9	96.5	96.6	97.3	96.8
	lda	96.3	95.5	95.7	96.7	96.1
	hlda	95.9	95.0	96.0	96.5	95.9
	5-rsfs	85.5	85.3	85.6	85.6	85.5
	initial	97.0	96.5	96.7	97.3	96.9
4+E+V+A	amfs	98.6	98.9	98.7	99.0	98.8
	lda	98.3	97.9	98.2	98.2	98.2
	hlda	98.7	98.3	98.3	98.9	98.6
	initial	98.6	98.8	98.6	99.2	98.8

4.2 Speech recognition experiments

We performed recognition experiments on features derived from the standard set of 12 MFCCs. We compared five types of feature sets for identical dimensionalities $n<12$. The first set of features results from the auditory-model based feature selection (amfs) method introduced in this paper (we use Gaussian perturbations with $|\mathcal{M}_j|=100$ and a SNR of 100 dB). The second set of features was obtained using standard (homoscedastic) linear discriminant analysis (lda).[4] The third set of features was obtained using standard heteroscedastic linear discriminant analysis (hlda).[5] The average performance of five randomly selected MFCC feature subsets is displayed as 5-rsfs. The fifth and final set is simply the set of the first n MFCCs, denoted as *initial*.

Note that lda and hlda have two advantages over amfs: (*i*) they are *dimension-reduction* methods rather than *subset-selection* methods and (*ii*) they require classified data as training input. The amfs method has as advantage that it can rely on knowledge inherent in the auditory periphery.

To build the recognizer we used the hidden Markov model toolkit (HTK) (Ref. 19). The digits were modeled as whole word hidden Markov models with 16 states (HTK notation is 18 states including the beginning and end states) and three Gaussian mixture components per state. To minimize modeling artifacts, the results are for full covariance matrices, but the use of diagonal covariance matrices gives essentially the same results. An initial model with global data means and covariances, identical for each digit, was used and 16 iterations were used to build the final model.

Table 1 shows the recognition accuracy for training and testing on clean data for dimensionality $n=8$ and for $n=4$. The caption of the table provides the MFCC subsets selected by amfs. For the experiments we added the energy feature "+E." We also performed the experi-

Table 2. AURORA2 multi-conditioning noisy training results. The amfs selected coefficients with indices 1, 2, 3, 4, 5, 6, 11, and 12 for $n=8$ and 1, 2, 3, and 4 for $n=4$, respectively.

Feature set		Data Test Set A				
		20 dB	15 dB	10 dB	5 dB	0 dB
12+E	Full, static	94.8	93.1	88.4	75.0	47.4
12+E+V+A	Full, dynamic	97.4	96.5	94.0	86.0	59.6
8+E	amfs	93.6	91.4	85.7	65.1	34.9
	lda	84.0	78.1	63.2	40.8	17.8
	hlda	85.0	80.9	71.7	55.7	33.2
	5-rsfs	88.0	84.6	75.6	53.0	24.6
	initial	93.1	91.0	86.0	69.2	38.1
8+E+V+A	amfs	97.1	96.0	93.4	84.1	56.8
	lda	94.4	92.8	87.8	76.5	52.0
	hlda	94.8	93.4	89.4	79.3	58.4
	initial	97.5	96.1	92.9	83.1	53.2
4+E	amfs	91.1	87.4	76.9	50.9	20.9
	lda	40.5	32.5	19.6	9.9	7.7
	hlda	44.0	38.8	28.6	16.6	8.4
	5-rsfs	66.3	59.3	43.6	24.0	11.9
	initial	91.1	87.4	76.9	50.9	20.9
4+E+V+A	amfs	96.0	94.6	90.9	79.4	50.1
	lda	93.6	92.1	86.3	73.9	48.9
	hlda	94.2	92.8	88.9	78.8	58.5
	initial	96.0	94.6	90.9	79.4	50.1

ments with feature sets that were augmented with their velocity ("+V") and acceleration ("+A"). For clarity we note that the subset-selection and dimension-reduction operations were always performed on the static features. Training was performed on the clean training set of 8440 sentences and the testing on the 4004 clean data of test set A.

For clean data the amfs selected feature set performs similar to the lda and hlda feature sets and to the initial set of MFCCS. All these feature sets perform significantly better than the average of randomly selected feature sets 5-rsfs. It is interesting to note that multiple distinct MFCC subsets perform well. Consistent with the recognition results shown in Table 1, the score of the measure of proximity to isometry, given by Eq. (1), is similar for the set initial and for the amfs selected features.

Table 2 shows the recognition accuracy for noisy data. Again the table caption provides the MFCC subsets selected by amfs. The training was performed on the multi-conditioned noisy training set consisting of 6752 files and the testing on the 24 024 noisy data of test set A. The results shown in Table 2 are averaged over subway, babble, car, and exhibition additive noise for several SNR values. The 5-rsfs configuration is the same as for the clean case (the same MFCC subsets were considered).

For the noisy data, the performance of the amfs subsets is in all cases better than lda, hlda, and 5-rsfs. The performance of initial is similar to that of amfs, although it does not consistently use the same subset. The results indicate that the new amfs method is more robust to environmental noise than lda and hlda. This increased robustness was confirmed in other experiments where we trained and tested on different environmental conditions. This result is not unexpected as amfs is based on an auditory periphery that is robust over a large range of environmental conditions. In contrast, lda and hlda must rely on the training data only.

Our results indicate that the natural ordering of the MFCCs is perceptually highly relevant. In both Tables 1 and 2 the initial set of MFCCs, "initial," performs as well as the amfs selected set and significantly better than a typical randomly selected set. While amfs does, in general, not simply select the initial set of MFCCs, it always includes the low frequency MFCCs, indicating that they represent an important component of the perceived information.

5. Conclusions

We conclude that the selection of speech features based on human perception results in robust features that perform well for speech recognition over a range of environmental conditions. Our results suggest that the method results in features that are more robust to noise than either homoscedastic or heteroscedastic discriminant analysis. This implies that effective dimension-reduction of feature sets for speech recognition is possible without knowledge of the meaning of the signal (without the availability of classified data). Our results indicate that the human auditory periphery has a parsimonious output representation, as significant redundancy would have made the measure of proximity to isometry, Eq. (1), ineffective for classification.

References and links

[1]H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," IEEE Trans. Knowl. Data Eng. **17**, 491502 (2005).

[2]H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy," IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1226–1238 (2005).

[3]J. H. Plasberg and W. B. Kleijn, "Feature selection under a complexity constraint," IEEE Trans. Multimedia **11**, 566–571 (2009).

[4]R. Haeb-Umbach and H. Ney, "Linear discriminant analysis for improved large vocabulary continuous speech recognition," in IEEE International Conference on Acoustics, Speech and Signal Processing (1992) Vol. **1**, pp. 13–16.

[5]N. Kumar and A. Andreou, "Heteroscedastic discriminant analysis and reduced rank hmms for improved speech recognition," Speech Commun. **267**, 283–297 (1998).

[6]F. Valente and C. Wellekens, "Maximum entropy discrimination (MED) feature subset selection for speech recognition," in IEEE Workshop on ASRU (2003) pp. 327–332.

[7]T. Dau, D. Püschel, and A. Kohlrausch, "A quantitative model of the "effective" signal processing in the auditory system. I. Model structure," J. Acoust. Soc. Am. **99**, 3615–3622 (1996).

[8]S. van de Par, G. Charestan, and R. Heusdens, "A gammatone-based psychoacoustical modeling approach for speech and audio coding," in Proceedings of the ProRISC (2001) Veldhoven, The Netherlands, pp. 321–326.

[9]W. R. Gardner and B. D. Rao, "Theoretical analysis of the high-rate vector quantization of LPC parameters," IEEE Trans. Speech Audio Process. **3**, 367–381 (1995).

[10]T. Linder, R. Zamir, and K. Zeger, "High-resolution source coding for non-difference distortion measures: Multidimensional companding," IEEE Trans. Inf. Theory **45**, 548–561 (1999).

[11]J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with a perceptual distortion measure," IEEE Trans. Inf. Theory **45**, 1082–1091 (1999).

[12]J. H. Plasberg and W. B. Kleijn, "The sensitivity matrix: Using advanced auditory models in speech and audio processing," IEEE Trans. Audio, Speech, Lang. Process. **15**, 310–319 (2007).

[13]S. Seneff, "A joint synchrony/mean rate model of auditory speech processing," J. Phonetics **16**, 55–76 (1988).

[14]O. Ghitza, "Auditory nerve representation as a basis for speech processing," *Advances in Speech Signal Processing* (Marcel Dekker, New York, 1991) pp. 453–485.

[15]W. Jeon and B.-H. Juang, "A study of auditory modeling and processing for speech signals," in IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (2005), Vol. **1**, pp. 929–932.

[16]S. Haque, R. Togneri, and A. Zaknich, "A temporal auditory model with adaptation for automatic speech recognition," in IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (2007), Vol. **4**, pp. 1141–1144.

[17]S. B. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," IEEE Trans. Acoust., Speech, Signal Process. **28**, 357–366 (1980).

[18]H. G. Hirsch and D. Pearce, "The AURORA experimental framework for the performance evaluation of speech recognition systems under noise conditions," in ISCA ITRW ASR2000 Automatic Speech Recognition: Challenges for the Next Millenium, Paris (2000).

[19]S. Young,D. Kershaw, J. Odell, D. Ollason, V. Valtchev, and P. Woodland, *The HTK Book (for HTK Version 3.2)* (Cambridge University, Cambridge, 2002).

A non-linear three-dimensional model for quantifying microbubble dynamics

Abhay V. Patil
Department of Biomedical Engineering, University of Virginia, 415 Lane Road, Room 2127, Charlottesville, Virginia 22908
avp2b@virginia.edu

Paul Reynolds
Weidlinger Associates, 399 West El Camino Real, Mountain View, California 94040–2607
reynolds@wai.com

John A. Hossack[a)]
Department of Biomedical Engineering, University of Virginia, 415 Lane Road, Room 2127, Charlottesville, Virginia 22908
jh7fj@virginia.edu

Abstract: A three-dimensional non-linear model for simulating microbubble response to acoustic insonation is presented. A 1 μm radius microbubble stimulated using positive and inverted 2.4 MHz pulses produced radius-time curves that matched (error $<10\%$) with the experimental observation. A bound 2.3 μm radius microbubble insonated using 2.25 MHz 6 cycle pulse was observed to oscillate with max/min oscillations 45% lower than that of the free microbubble, this correlated ($\sim 10\%$ error) with the observations of Garbin *et al.* [Appl. Phys. Lett. **90**, 114103 (2007)]. The adherent microbubble oscillated asymmetrically in the plan view and symmetrically in the elevation view, consistent with the previous experimental results.

PACS numbers: 43.80.Qf [CC]
Date Received: August 28, 2009 **Date Accepted:** December 10, 2009

1. Introduction

Ultrasound contrast agents are gas filled microspheres encapsulated by lipid, protein, or polymer shells. The linear resonance frequencies of the microbubbles are within the diagnostic ultrasound limit, thus making them valuable in medical imaging. In the past decade, the therapeutic applications of microbubbles have emerged, primarily motivated by the need of a suitable drug delivery/gene delivery agent.[1]

One-dimensional (1D) radially symmetric models have been widely reported in the literature for quantifying the microbubble's behavior in response to acoustic insonation. However, most 1D (Ref. 2) models are incapable of predicting phenomena such as higher order nonspherical microbubble oscillations and a consequent shell rupture. Also, phenomena such as the microbubble-microbubble interaction and the microbubble-cell interactions are poorly understood. A thorough understanding of these phenomena can aid in selectively guiding, imaging, and destroying adherent microbubbles, thus enhancing local drug delivery. Criteria such as the Blake threshold,[3] the peak negative velocity,[4] and the surface kinetic energy of a collapsing microbubble and its cutoff shell stress/strain[5] have been suggested in the literature as thresholds for predicting the onset of an inertial microbubble cavitation and a successive fragmentation. Although some of the above formulations have been validated using high speed camera experiments, *ad hoc* modifications in these 1D models have been suggested to fully or partially explain them. Here, we propose a full non-linear three-dimensional (3D) finite element analysis

[a)]Author to whom correspondence should be addressed.

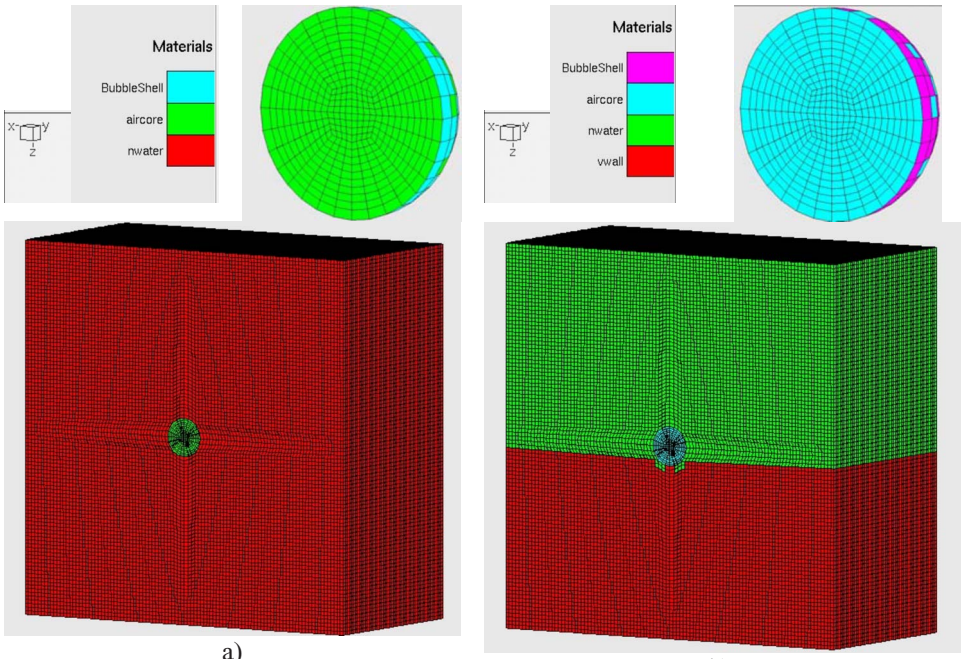


Fig. 1. (Color online) (a) Cross-section view of a free microbubble enclosed by a meshed cube representing saline. The microbubble shell is 1 nm thick. The microbubble shell encloses perfluorobutane gas. (b) Cross-section view of an adherent microbubble attached to a lipid wall. The external excitation is applied to *XY* plane near the origin. The acoustic wave propagates in positive *Z* direction. The radii of the free and the adherent microbubbles are 1 and 2.3 μm, respectively.

(FEA) model for quantifying the microbubble dynamics. This model estimates the coupled 3D oscillatory-translational motion, the shell stress/strain, and the backscattered acoustic pressure.

2. Methods

PZFlex (Weidlinger Associates Inc., Mountain View, CA), a 3D FEA software was used in this work.[6] In PZFlex, the coupled acoustomechanical problems can be simulated by precisely accounting for the non-linear propagation of the acoustic waves. Free (1 and 2.3 μm radii) and adherent (2.3 μm radius) microbubbles were simulated in PZFlex. A free microbubble [Fig. 1(a)] is a microbubble that is surrounded by saline or blood and is not in contact with any other entity (vessel wall or other microbubble). An adherent microbubble is a microbubble that is attached to an arterial vessel wall [Fig. 1(b)]. Each microbubble was simulated in PZFlex by meshing a sphere. The microbubble was encompassed by a cube that simulated the surrounding media (e.g., saline or blood). The meshed cube was ten times the diameter of the meshed microbubble. Eight-node skewed hexahedral elements[7] were used to mesh the model. A target element size of approximately 1/100 of the length of the cube was chosen for meshing the model. The mesh was refined and the model was executed until the output was mesh density invariant and convergent. A final node density of approximately 2×10^6 nodes per model was used in the simulations reported in this study. In the case of an adherent microbubble, the lipid[8] wall properties were assigned to a column of elements to simulate a lumen-vessel wall interface [Fig. 1(b)]. The adhered microbubble was simulated to be in contact with the arterial wall such that a zero relative motion condition was imposed on the contact shell elements (2×2 or 4, element size $\sim140\times140\times1$ nm^3) between the wall and the microbubble shell [Fig. 1(b)].

Table 1 lists the values corresponding to various parameters used in the simulations. The local material properties of the media and the microbubble gas (continuum elements) were defined using the material's bulk modulus, shear modulus, Poisson's ratio, and density. External

Table 1. List of all parameters used in the 3D simulations.

	Saline	C_4F_{10}	Lipid wall	Bubble shell
Density ρ	1000 kg/m^3	11.2 kg/m^3	1000 kg/m^3	1000 kg/m^3
Poisson's ratio υ	0.499	⋯	0.499	0.499
Shear modulus G	0	0	$E/(2(1+\upsilon))$	$E/(2(1+\upsilon))$
Elastic modulus E	⋯	⋯	1 MPa	χ/ε
Sound speed c	1480 m/s	108 m/s	⋯	⋯
Dilation modulus χ	⋯	⋯	⋯	0 N/m, 0.3 N/m
Bulk modulus K	ρc^2	ρc^2	$E/(3(1-2\upsilon))$	$E/(3(1-2\upsilon))$
B/A	5	0.006	⋯	⋯

acoustic stimulation was applied to a side of the model (along XY plane at $Z=0$, Fig. 1). Boundary conditions on all the other sides were defined such that they absorbed the acoustic energy to prevent reverberations and undesired reflection. The coupled differential equations that govern the transient behavior of the model are as follows:

$$M_{\mu\mu}\frac{d^2\mu}{dt^2}+C_{\mu\mu}\frac{d\mu}{dt}+C_{\mu\psi}\frac{d\psi}{dt}+K_{\mu\mu}\mu=F, \tag{1}$$

$$M_{\psi\psi}\frac{d^2\psi}{dt^2}-C^T_{\mu\psi}\frac{d\mu}{dt}+C_{\psi\psi}\frac{d\psi}{dt}+K_{\psi\psi}\psi=0. \tag{2}$$

Equation (1) defines the mechanical behavior of the system and Eq. (2) defines the acoustic behavior of the system. The global unknowns, μ and Ψ, are the elastic displacement vector and the velocity potential vector, respectively. F is the external excitation function applied to a side of the model. M, C, and K are mass, damping, and stiffness matrices, respectively. The above equations are applicable to all the material phases' continuum and shells of the FEA model. The equations are integrated step by step using an explicit time-domain approach.[6]

In non-linear wave propagation, the wave speed of the propagating wave is inherently pressure dependent by the virtue of the pressure and density relationship,

$$c=\sqrt{\frac{dp}{d\rho}},\quad p=p_o+A\left[\theta+\frac{B}{A}\frac{\theta^2}{2}\right],\quad \theta=\frac{\rho}{\rho_o}-1, \tag{3a}$$

where θ is dilatation or expansion and B/A (Ref. 9) is the parameter of nonlinearity in the truncated Taylor series, c is the velocity of sound through a medium, ρ is the material density, and p is the wave pressure. The incrementally linear, second order accurate,[10] explicit algorithm used in PZFlex for simulating the non-linear propagation is based on the velocity and force update equations for a computational element and is stated as

$$v^n=v^{n-1}+\frac{\Delta t F^{n-1/2}}{m}\Rightarrow \mu^{n+1/2}+v^n\Delta t\Rightarrow \theta^{n+1/2}=div\,.\,u^{n+1/2}, \tag{3b}$$

$$p^{n+1/2}=p_o+A\left[\theta^{n+1/2}+\frac{B}{A}\frac{(\theta^{n+1/2})^2}{2}\right]\Rightarrow F^{n+1/2}, \tag{3c}$$

where v^n, $\mu^{n-1/2}$, and $F^{n-1/2}$ are the nodal velocity, displacement, and force, respectively, m is the lumped mass, Δt is the time step, and superscripts refer to the time level. This non-linear propagation model is referred to as "tisu" model in PZFlex. The non-linear pressure-volume relationship in the gas domain of a microbubble can be modeled in PZFlex using the tisu material model and the B/A parameter. The B/A values for a given gas can be calculated as follows,

$$\frac{B}{A} = \frac{\gamma - 1}{\rho}, \tag{4}$$

where γ is the polytropic expansion constant of the gas and ρ is the density. The B/A values of air, perflurobutane (C_4F_{10}), and water are 0.4, 0.006, and 5.0, respectively. The gas core of the microbubble was modeled as perflurobutane gas (C_4F_{10}) with density[11] 11.2 kg/m^3 and dilatational wave speed of approximately 108 m/s. The shell parameters (shell elements) such as shell elastic (E), bulk (K), and shear moduli (G) can be estimated from the area dilatational modulus or shell stiffness modulus. The area dilatational modulus is defined as, $\chi = E\varepsilon$, where E and ε are Young's modulus and the thickness of the microbubble shell. For all simulations reported in this work, the microbubble's shell was assumed to be of 1 nm thickness.[12] The Poisson's ratio for biomaterials such as lipid and protein is in the range[5] $0.49 < \upsilon < 0.5$. For all simulations reported in this work, the Poisson's ratio was assumed to be 0.499. The surrounding water media was modeled as a viscous Newtonian fluid. The density and the viscosity of water were assumed to be 1000 kg/m^3 and 0.001 Pa s, respectively.

3. Results

Free and adherent 2.3 μm radius microbubbles were simulated in the 3D FEA software. The microbubbles were stimulated using a windowed 6 cycle 2.25 MHz sine pulse with a peak negative pressure of 240 kPa [Fig. 2(d)]. Similar experimental parameters were used in works published by Garbin *et al.*,[13] Zhao *et al.*,[14] and Vos *et al.*[15] In the adherent microbubble model, the vessel wall was simulated using the mechanical properties of a stiff arterial wall[8] with the stiffness of ($E = 1$ MPa) and the density of 1000 kg/m^3. The arterial wall was assumed to be incompressible. Figure 2(a) illustrates the superimposed radial oscillations of the free and adherent microbubbles. The adherent microbubble exhibits reduced radial excursion when compared with the free microbubble. The maximum positive and negative excursions of the adherent microbubble are approximately 0.8 and 0.6 μm, respectively [Fig. 2(a)], whereas the maximum excursions of the free microbubble are 1.5 and 1.3 μm, respectively. The radial excursions of the adherent microbubble are approximately 45% smaller than that of the free microbubble. This observation correlates with the results published by Garbin *et al.*,[13] who reported a 50% reduction. As illustrated in Fig. 2(b), the center of the adherent microbubble follows its radial oscillation, but with reduced amplitude, implying that the microbubble is continuously attached to the simulated wall. The center of the free microbubble follows a uniform oscillatory-translational motion in the direction away from the excitation source [Fig. 2(b)]. Similar experimental results have been demonstrated by Vos *et al.*[15] The spectrum of the echo received from the adherent microbubble exhibits reduced power in the higher order harmonics and an increased ratio in the power of the fundamental to harmonics (20% fundamental second and 50% fundamental third) when compared with the spectrum of the free microbubble [Fig. 2(c)]. Experimental results reporting the same were first published by Zhao *et al.*[16] (30% fundamental second and 70% fundamental third at 2 MHz/210 kPa). It should be noted that the experimental results were reported for a cloud of microbubble rather than a single microbubble.

Figure 3 illustrates the 3D oscillatory motion of an adherent microbubble in two different orthogonal planes, azimuth-range (XZ plane) and azimuth-elevation (XY plane). Here after, these planes are referred to as plan view and elevation view, respectively. The adherent microbubble oscillates asymmetrically (second order mode) in the plan view and symmetrically in the elevation view. Thus, our 3D simulations validate the experimental results independently reported by Vos *et al.*[15] and Zhao *et al.*[14] 3D FEA simulations may provide insight in designing pulse codes capable of uniquely enhancing signals from adherent microbubbles and suppressing signals from free microbubbles, 1D models are incapable of providing such valuable information.

Figure 4(a) illustrates the R-T (radius-time) curves of a free 1 μm radius microbubble stimulated using a 2.4 MHz Gaussian modulated sine wave with 50% fractional bandwidth (BW) and a peak negative pressure of 360 kPa. The 3D FEA R-T curves obtained using the above parameters and the viscoelastic model published by Church *et al.*[17] are superimposed on

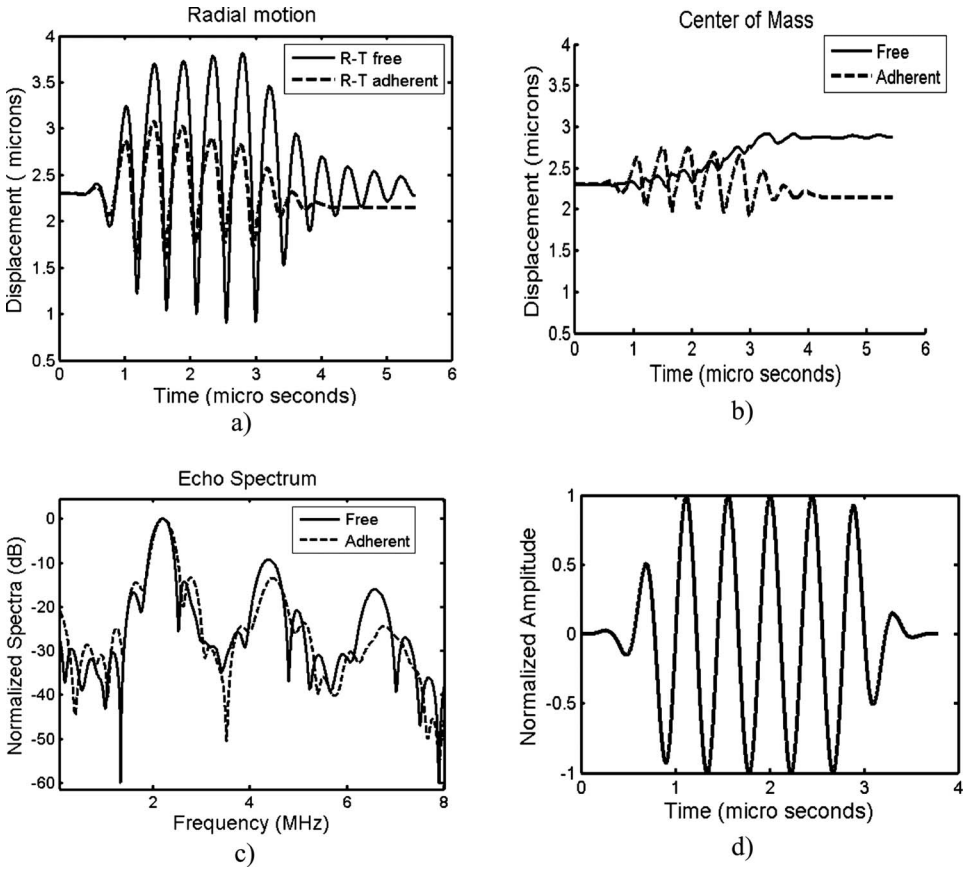


Fig. 2. A 2.3 μm radius microbubble (0.3 N/m area dilation modulus) stimulated using a 2.25 MHz, 6 cycle sine pulse with a peak negative pressure of 240 kPa. (a) Superimposed radius-time curves of free and adherent microbubble, (b) superimposed motion of the center of mass (translational motion) of free and adherent microbubble, (c) superimposed spectra of the received echoes from free and adherent microbubble, and (d) excitation pulse.

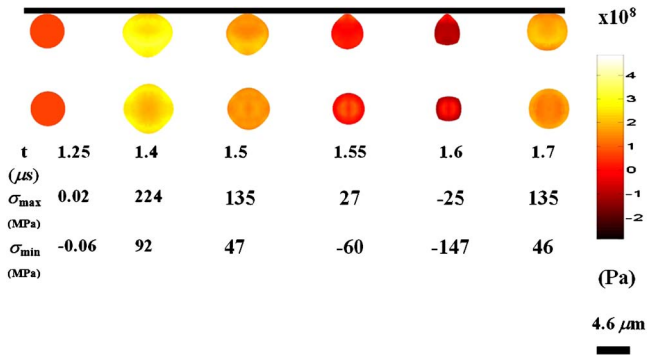


Fig. 3. (Color online) Time snapshots of the 2.3 μm adherent microbubble. The microbubble was stimulated using a 2.25 MHz 6 cycle sine pulse (240 kPa) along the plane parallel to azimuth-elevation plane. Plan view (azimuth-range plane) of the adherent microbubble (first row). Elevation view (azimuth elevation) of the adherent microbubble (second row). The last two rows list the corresponding maximum and minimum shell stresses at various time instants in the excitation cycle. The elastic modulus (E) of the microbubble shell was simulated to be 300 MPa. The peak negative normal shell stress or compressive shell stress (at t=1.6 μs) was approximately 147 MPa ($<$0.5 E).

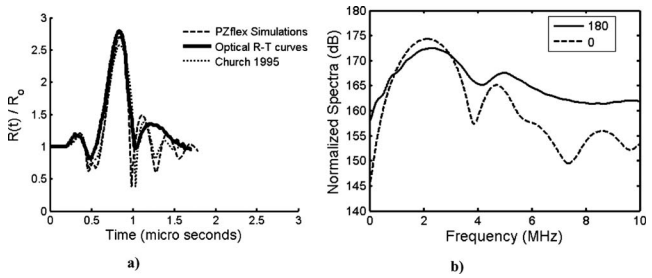


Fig. 4. A 1 μm radius microbubble was stimulated using a 2.4 MHz 50% fractional BW Gaussian modulated sine wave with 0° phase pulse at 360 kPa. (a) Superimposed PZFlex simulations, optical *R-T* curves (Ref. 12), and 1D simulations obtained using viscoelastic model published in Ref. 17. (b) Spectrum of the echoes received from 0° and 180° phase pulses. $R(t)$ is the instantaneous microbubble radius, whereas R_o is the resting microbubble radius.

the optical *R-T* curves obtained from Morgan *et al.*[12] [Fig. 4c in Ref. 12]. The rms error between the 3D FEA simulations and the optical *R-T* curves is 9%. Figure 4(b) illustrates the superimposed spectra of the echo response obtained by insonating a 1 μm microbubble with 0° and 180° phase pulses. A 0° phase pulse is a Gaussian modulated sine wave and a 180° pulse is an inverted Gaussian modulated sine wave. The shift between the centroids of the spectra is approximately 0.5 MHz and is within the range of the frequency shifts reported by Morgan *et al.*[12]

4. Conclusion

In this letter, we introduced and validated a 3D non-linear microbubble simulation framework. This model has great potential in addressing complex problems such as analyzing 3D asymmetric mode of vibrations and providing complementary information in the form of shell stress and strain that is not available from the conventional 1D models. The complementary information provided by this model may further aid in deciphering the microbubble-cell interaction, formulating microbubble cavitation limits, and analyzing multiple microbubble-microbubble interactions that may aid in selectively imaging microbubble and guiding local drug delivery.

Acknowledgments

The authors are grateful to Dr. John Mold, Jr. and Mr. David Milner from Weidlinger Associates, Mountain View, CA. A.V.P. and J.A.H. acknowledge support from the NIH Grant Nos. NIBIB EB002185 and NHBLI HL 090700.

References and links

[1]R. Unger, T. P. McCreery, R. Sweitzer, V. Caldwell, and Y. Wu, "Acoustically active lipospheres containing paclitaxel: A new therapeutic ultrasound contrast agent," Invest. Radiol. **33**, 886–892 (1998).
[2]J. S. Allen, J. M. Donovan, and K. W. Ferrara, "Dynamics of therapeutic ultrasound contrast agents," Ultrasound Med. Biol. **28**, 805–816 (2002).
[3]C. E. Brennen, *Cavitation and Bubble dynamics* (Oxford University Press, New York, 1995).
[4]J. E. Chomas, P. A. D. Dayton, and K. W. Ferrara, "Threshold of fragmentation for ultrasonic contrast agents," J. Biomed. Opt. **6**, 141–150 (2001).
[5]M. Postema and G. Schmitz, "Ultrasonic bubbles in medicine: Influence of the shell," Ultrason. Sonochem. **14**, 438–444 (2007).
[6]G. L. Wojcik, D. K. Vaughan, N. Abboud, and J. Mould, "Electromechanical modeling using explicit time-domain finite elements," Proc.-IEEE Ultrason. Symp. **2**, 1107–1112 (1993).
[7]PZFlex Manual, Weidlinger Associates, Mountain View, CA (2007).
[8]D. H. Bergel, "The dynamic elastic properties of the arterial wall," J. Physiol. **156**, 458–469 (1961).
[9]T. L. Szabo, *Diagnostic ultrasound imaging: Inside out* (Elsevier, New York/Academic, New York, 2004), Chap. 12, p. 382.
[10]G. L. Wojcik, J. Mould, N. Lizzi, M. Abboud, and V. D. Ostromogilsky, "Non-linear modeling of therapeutic ultrasound," Proc.-IEEE Ultrason. Symp. **2**, 1617–1622 (1995).
[11]J. E. Chomas, P. A. Dayton, J. S. Allen, K. E. Morgan, and K. W. Ferrara, "Mechanism of contrast agent destruction," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **48**, 232–248 (2001).
[12]K. E. Morgan, J. S. Allen, P. A. Dayton, J. E. Chomas, A. L. Klibanov, and K. W. Ferrara, "Experimental and theoretical evaluation of microbubble behavior: Effect of transmitted phase and bubble size," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 1494–1509 (2000).

[13]V. Garbin, D. Cojoc, E. Ferrari, E. D. Fabrizio, M. L. J. Overvelde, M. S. Van der Meer, N. de Jong, D. Lohse, and M. Versluis, "Changes in microbubble dynamics near a boundary revealed by combined optical manipulation and high-speed camera," Appl. Phys. Lett. **90**, 114103 (2007).
[14]S. Zhao, K. W. Ferrara, and P. A. Dayton, "Asymmetric oscillations of adherent ultrasound contrast agents," Appl. Phys. Lett. **87**, 13410 (2005).
[15]H. J. Vos, M. Versluis, and N. de Jong, "Orthogonal observations of vibrating microbubbles," Proc.-IEEE Ultrason. Symp. **1**, 765–768 (2007).
[16]S. Zhao, D. E. Kruse, K. W. Ferrara, and P. A. Dayton, "Acoustic response from adherent contrast agents," J. Acoust. Soc. Am. **120**, EL63–EL69 (2006).
[17]C. Church, "The effects of an elastic solid surface layer on the radial pulsation of gas bubble," J. Acoust. Soc. Am. **97**, 1510–1521 (1995).

LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

Scattering by an arrangement of eccentric cylinders embedded on a coated cylinder with applications to tomographic density imaging (L)

Roberto J. Lavarello[a] and Michael L. Oelze

Department of Electrical and Computer Engineering, Bioacoustics Research Laboratory, University of Illinois at Urbana-Champaign, 405 North Matthews, Urbana, Illinois 61801

(Received 10 August 2009; revised 22 November 2009; accepted 23 November 2009)

The solution to the scattering of an incident pressure wave by an arrangement of eccentric cylinders embedded inside a pair of concentric cylinders is derived here using a combination of *T*-matrix and mode-matching approaches. This method allows the generation of synthetic data from relatively complex structures to be used for the validation of acoustic tomography methods. An application of the solution derived here is illustrated by reconstructing sound speed and density profiles from a complex phantom using inverse scattering. © *2010 Acoustical Society of America.*
[DOI: 10.1121/1.3277154]

PACS number(s): 43.20.Fn, 43.35.Wa [TDM] Pages: 645–648

I. INTRODUCTION

The scattering of a cylindrical wave by an arrangement of eccentric cylinders is derived here in order to validate inverse scattering routines with scatterers more complex than a single homogeneous cylinder. Works dealing with the scattering by two concentric fluid cylinders,[1] two[2] and multiple[3] rigid parallel cylinders, two[4] and multiple[5] fluid parallel cylinders, two eccentric fluid cylinders,[6,7] and multiple eccentric cylinders embedded in a circular cylinder[8] can be found in the literature. The work presented here extends the scope of the work in Ref. 8 by studying the scattering by *N* circular cylinders embedded inside a coated cylinder, as shown in Fig. 1.

The approach presented here is a combination of the *T*-matrix formulation and mode-matching techniques, and considers changes in compressibility, density, and attenuation (unlike the work in Ref. 8 that only considers changes in refractive index).

II. CALCULATION OF THE SCATTERING COEFFICIENTS

In the following derivation, the background has a wave number k_0 and acoustic impedance Z_0. For the other cylinders, the complex wave numbers k_n and impedances Z_n are defined as $k_n=(\omega/c_n)+i\alpha_n$ and $Z_n=\rho_n c_n/(1+i\alpha_n c_n/\omega)$, where c_n, ρ_n, and α_n are the speed of sound, density, and attenuation coefficient in the *n*-th subregion. The radius of the *n*-th cylinder is denoted by a_n. The wave number, impedance, and outer radius of the coating are denoted by k_c, Z_c, and a_c, respectively. The acoustic field inside cylinder 1 can be written as

$$p_1(\vec{r}_o)=\sum_{m=-\infty}^{\infty}A_m J_m(k_1 r_o)e^{im\theta_o}+\sum_{n=2}^{N+1}\sum_{p=-\infty}^{\infty}B_{p,n}H_p(k_1 r_{on})e^{ip\theta_{on}}, \quad (1)$$

where $J_m(\cdot)$ is the *m*-th order Bessel function, $H_m(\cdot)$ is the *m*-th order Hankel function of the first kind, $\vec{r}_o=(r_o,\theta_o)$ are the polar coordinates of the observation point, $\vec{r}_n=(r_n,\theta_n)$ are the polar coordinates of the center of the *n*-th cylinder, and $\vec{r}_{on}=\vec{r}_o-\vec{r}_n$. The second sum in Eq. (1) represents the fields produced by the embedded cylinders. The *T*-matrix approach[9] can be used to relate the amplitudes of the A_m and $B_{p,n}$ terms. First, a matrix $\bar{T}_s$ is used to express the Bessel waves relative to the center of each one of the *N* embedded cylinders. The matrix $\bar{T}_s$ is composed of *M* blocks $\bar{T}_s^h$ of size $N\times M_p$, where *M* and M_p are the number of terms used to expand the first and second infinite sums in Eq. (1). This can be expressed in matrix form as

$$\bar{e}_t=\bar{T}_s\cdot\bar{a}, \quad (2)$$

$$\bar{T}_s=[\cdots\ (\bar{T}_s^{m=-1})^T\ (\bar{T}_s^{m=0})^T\ (\bar{T}_s^{m=1})^T\ \cdots]^T, \quad (3)$$

[a]Author to whom correspondence should be addressed. Electronic mail: lavarell@illinois.edu

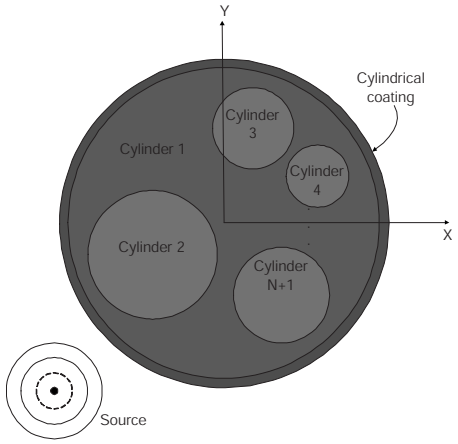


FIG. 1. Scatterer formed by an arrangement of circular cylinders embedded inside a coated circular cylinder.

$$[\bar{T}_s^m]_{np} = J_{p-m}(k_1 r_n) e^{-i(p-m)(\theta_n+\pi)}, \tag{4}$$

where $\bar{a}$ is an M_p length vector with the amplitudes of the Bessel wave harmonics and $\bar{e}_t$ is an $(M \times N)$ vector with the translated amplitudes of the Bessel waves. In the notation above, $\bar{X}^T$ represents the transpose of a matrix $\bar{X}$. The harmonics $\bar{e}_t$ can be related to the Hankel harmonics in Eq. (1) as

$$\bar{b}_t = [\bar{I} - \mathcal{D}(\bar{R}) \cdot \bar{A}]^{-1} \cdot \mathcal{D}(\bar{R}) \cdot \bar{e}_t. \tag{5}$$

where $\bar{b}_t$ is an $(M \times N)$ vector with the $B_{p,n}$ coefficients and $\mathcal{D}(\cdot)$ is an operator that transforms a vector into a diagonal matrix. The $(M \times N)$ vector $\bar{R}$ contains the single-cylinder scattering coefficients given by

$$\bar{R} = [(\bar{R}^{n=2})^T\ (\bar{R}^{n=3})^T\ \cdots\ (\bar{R}^{n=N+1})^T]^T, \tag{6}$$

$$[\bar{R}^n]_p = \frac{\frac{1}{Z_{rn}} J_p(k_1 a_n) J_p'(k_n a_n) - J_p(k_n a_n) J_p'(k_1 a_n)}{J_p(k_n a_n) H_p'(k_1 a_n) - \frac{1}{Z_{rn}} J_p'(k_n a_n) H_p(k_1 a_n)}, \tag{7}$$

where the " $'$ " symbol denotes derivative with respect to the total argument and $Z_{rn} = Z_n / Z_1$. The elements of the $(M \times N) \times (M \times N)$ T-matrix $\bar{A}$ are given by

$$\bar{A} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} & \cdots & \bar{A}_{1M} \\ \bar{A}_{21} & \bar{A}_{22} & \cdots & \bar{A}_{2M} \\ & \vdots & & \\ \bar{A}_{M1} & \bar{A}_{M2} & \cdots & \bar{A}_{MM} \end{bmatrix}, \tag{8}$$

$$[\bar{A}_{mn}]_{pq} = \begin{cases} 0 & \text{if } p = q \\ H_{m-n}(k_0 r_{pq}) e^{-i(m-n)\theta_{pq}}, & \text{else,} \end{cases} \tag{9}$$

where (r_{pq}, θ_{pq}) is the polar representation of the vector $(\vec{r}_p - \vec{r}_q)$ containing the location of the center of cylinder p relative to the center of cylinder q, with $p, q \in [2, N+1]$. Enforcing mode-matching at the cylindrical coating becomes simpler if Eq. (1) is expressed in terms of Bessel and Hankel fields centered at the origin. In particular, if the observation point $\vec{r}_o$ satisfies $|\vec{r}_o| > |\vec{r}_n|$, $\forall n \in [2, N+1]$, the addition theorem of Hankel functions[10] allows the pressure field to be expressed as

$$p_1(\vec{r}_o) = \sum_{m=-\infty}^{\infty} (A_m J_m(k_1 r_o) + B_m H_m^{(1)}(k_1 r_o)) e^{im\theta_o}, \tag{10}$$

where the relationship between the B_m and $B_{p,n}$ coefficients is given by

$$\bar{b} = \bar{T} \cdot \bar{b}_t, \tag{11}$$

$$\bar{T} = [\cdots\ \bar{T}^{h=-1}\ \bar{T}^{h=0}\ \bar{T}^{h=1}\ \cdots], \tag{12}$$

$$[\bar{T}^{h=m}]_{pn} = J_{p-m}(k_1 r_n) e^{-i(p-m)\theta_n}, \tag{13}$$

where $\bar{T}$ is an $(M \times P) \times (M \times M)$ transition matrix, and $\bar{b}$ is an M_p length vector containing the coefficients B_m. Overall, the relationship between the A_m and B_m coefficients can be expressed as

$$\bar{b} = \bar{P} \cdot \bar{a},$$

$$\bar{P} = \bar{T} \cdot [\bar{I} - \mathcal{D}(\bar{R}) \cdot \bar{A}]^{-1} \cdot \mathcal{D}(\bar{R}) \cdot \bar{T}_s. \tag{14}$$

The expression in Eq. (14) is valid, in particular, for points close to the edge of cylinder 1. In the cylindrical coating, the field can be expressed as

$$p(\vec{r}_o) = \sum_{m=-\infty}^{\infty} (C_m J_m(k_c r_o) + D_m H_m(k_c r_o)) e^{im\theta_o}, \tag{15}$$

where the first sum is the incident field and the second one is the scattered field. The magnitudes of the A_m, B_m, C_m, and D_m coefficients can be related using the continuity of both pressure and normal particle velocity at $r_o = a_1$. Therefore, one can express

$$\mathcal{D}(\bar{R}_{jj}) \cdot \bar{a} + \mathcal{D}(\bar{R}_{hj}) \cdot \bar{b} = \bar{c}, \tag{16}$$

$$\mathcal{D}(\bar{R}_{jh}) \cdot \bar{a} + \mathcal{D}(\bar{R}_{hh}) \cdot \bar{b} = \bar{d}, \tag{17}$$

where $\bar{c}$ and $\bar{d}$ are M_p length vectors with the coefficients C_m and D_m, respectively, and $\bar{R}_{jj}$, $\bar{R}_{hj}$, $\bar{R}_{jh}$, and $\bar{R}_{hh}$ are M_p length vectors containing the transmission coefficients given by

$$[\bar{R}_{jj}]_m = \frac{J_m(k_1 a_1) H_m'(k_c a_1) - \frac{1}{Z_{rc}} J_m'(k_1 a_1) H_m(k_c a_1)}{J_m(k_c a_1) H_m'(k_c a_1) - J_m'(k_c a_1) H_m(k_c a_1)}, \tag{18}$$

$$[\bar{R}_{hj}]_m = \frac{H_m(k_1 a_1) H_m'(k_c a_1) - \frac{1}{Z_{rc}} H_m'(k_1 a_1) H_m(k_c a_1)}{J_m(k_c a_1) H_m'(k_c a_1) - J_m'(k_c a_1) H_m(k_c a_1)}, \tag{19}$$

$$[\bar{R}_{hj}]_m = -\frac{J_m(k_1 a_1) J_m'(k_c a_1) - \frac{1}{Z_{rc}} J_m'(k_1 a_1) J_m(k_c a_1)}{J_m(k_c a_1) H_m'(k_c a_1) - J_m'(k_c a_1) H_m(k_c a_1)}, \tag{20}$$

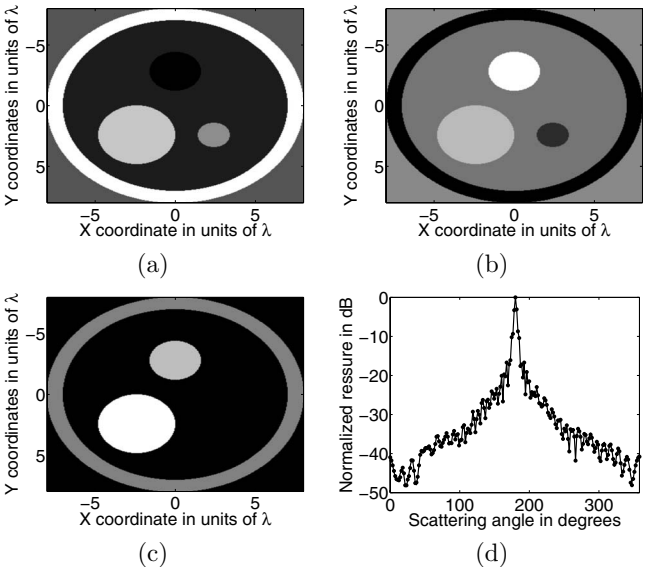


FIG. 2. Numerical validation of the scattering solution. (a) Δc of the phantom. (b) $\Delta\rho$ of the phantom. (c) Ratio α/k_0 of the phantom. (c) Scattered field produced by the phantom calculated using the solution presented in Sec. II (solid) and the numerical solver from Ref. 11 (dots).

$$[\bar{R}_{hh}]_m = \frac{\frac{1}{Z_{rc}}H'_m(k_1a_1)J_m(k_ca_1) - H_m(k_1a_1)J'_m(k_ca_1)}{J_m(k_ca_1)H'_m(k_ca_1) - J'_m(k_ca_1)H_m(k_ca_1)}, \quad (21)$$

where $Z_{rc}=Z_1/Z_c$. Similarly, the field in the background can be written as

$$p(\vec{r}_o) = \sum_{m=-\infty}^{\infty} (E_mJ_m(k_0r_o) + F_mH_m(k_0r_o))e^{im\theta_o}, \quad (22)$$

where the first term represents the incident field and the second term represents the scattered field. The coefficients E_m are known and depend on the type of illumination used. For example, when using a line source located at $\vec{R}_s(R_s,\theta_s)$ the coefficients E_m are given by $E_m=H_m^{(1)}(k_0R_s)e^{-im\theta_s}$. The coefficients F_m are the unknowns that need to be solved in order to calculate the scattered field. The relationships between the coefficients C_m, D_m, E_m, and F_m are given by

$$\mathcal{D}(\bar{R}_{jj2})\cdot\bar{c} + \mathcal{D}(\bar{R}_{hj2})\cdot\bar{d} = \bar{e}, \quad (23)$$

$$\mathcal{D}(\bar{R}_{jh2})\cdot\bar{c} + \mathcal{D}(\bar{R}_{hh2})\cdot\bar{d} = \bar{f}, \quad (24)$$

where $\bar{e}$ and $\bar{f}$ are M_p length vectors containing the E_m and F_m coefficients, respectively, and the transmission coefficients $\bar{R}_{jj2}$, $\bar{R}_{hj2}$, $\bar{R}_{jh2}$, and $\bar{R}_{hh2}$ can be found using Eqs. (18)–(21) replacing a_1 by a_c, k_c by k_0, k_1 by k_c, and Z_{rc} by Z_c/Z_0. Therefore, the A_m coefficients can be found by using

$$\bar{a} = [\mathcal{D}(\bar{R}_{jj2})\cdot\bar{M}_1 + \mathcal{D}(\bar{R}_{hj2})\cdot\bar{M}_2]^{-1}\cdot\bar{e}, \quad (25)$$

$$\bar{M}_1 = \mathcal{D}(\bar{R}_{jj}) + \mathcal{D}(\bar{R}_{hj})\cdot\bar{P}, \quad (26)$$

$$\bar{M}_2 = \mathcal{D}(\bar{R}_{jh}) + \mathcal{D}(\bar{R}_{hh})\cdot\bar{P}. \quad (27)$$

Finally, the scattering coefficients F_m can be obtained by using

$$\bar{f} = [\mathcal{D}(\bar{R}_{jh2})\cdot\bar{M}_1 + \mathcal{D}(\bar{R}_{hh2})\cdot\bar{M}_2]\cdot\bar{a}. \quad (28)$$

These coefficients F_m are the ones needed to calculate the scattered field using the second term in Eq. (22).

TABLE I. Properties of the scatterer used for numerical validation of the scattering solution from Sec. II.

Cylinder	Center	Radius	Δc (%)	$\Delta\rho$ (%)	α/k
1	[0,0]	$8\lambda_0$ (outer) $7\lambda_0$ (inner)	6	−6	0.1
2	[0,0]	$7\lambda_0$	−2	−1	0
3	$[0,2.8\lambda_0]$	$1.6\lambda_0$	−3	5	0.15
4	$[-2.4\lambda_0,-2.4\lambda_0]$	$2.4\lambda_0$	4	2	0.2
5	$[2.4\lambda_0,-2.4\lambda_0]$	λ_0	2	−4	0

III. NUMERICAL VALIDATION

An example of the scattering solution derived in this work is shown in Fig. 2. The dimensions, speed of sound, and density contrasts Δc and $\Delta\rho$, and α/k ratios of all cylinders are given in Table I. The incident field was produced by a line source located at $x=300\lambda$. The values of M and M_p for the scattering solution were set to 51 and 121, respectively. The scattered field was also calculated using the numerical solver presented in Ref. 11 with a grid size of $\lambda/20$. The root mean square error between both calculated scattered fields was only 0.4%, which suggests a proper convergence of the solution derived in this manuscript even in the presence of attenuation. The results are shown in Fig. 2.

Although in principle there should be no restrictions on the sizes and acoustic properties of the embedded cylinders as long as they remain nonoverlapping, special care must be taken when reconstructing scatterers consisting of regions

TABLE II. Mean speed of sound and density values corresponding to the reconstruction of a complex scatterer using the multiple frequency T-matrix approach with different $f_{\min}$ values.

Cylinder No.	Center position	Radius	$(\Delta c,\Delta\rho)$ ideal	$(\Delta c,\Delta\rho)$ $f_{\min}=f_0$	$(\Delta c,\Delta\rho)$ $f_{\min}=f_0/16$	$(\Delta c,\Delta\rho)$ $f_{\min}=f_0/64$
1	[0,0]	$8\lambda_0$ (outer) $7\lambda_0$ (inner)	(1.8%, −1.5%)	(1.68%, −2.5%)	(1.6%, −1.31%)	(1.6%, −1.51%)
2	[0,0]	$7\lambda_0$	(−1.8%, 1.5%)	(−1.78%, 2.88%)	(−1.76%, 2.09%)	(−1.76%, 1.67%)
3	$[1.8\lambda_0,1.8\lambda_0]$	$2\lambda_0$	(2.5%, −2%)	(2.36%, −3.56%)	(2.32%, −1.61%)	(2.32%, −2.13%)
4	$[-2.4\lambda_0,-2.4\lambda_0]$	$1.6\lambda_0$	(−2.5%, 2%)	(−2.49%, 4.11%)	(−2.48%, 2.99%)	(−2.49%, 2.34%)

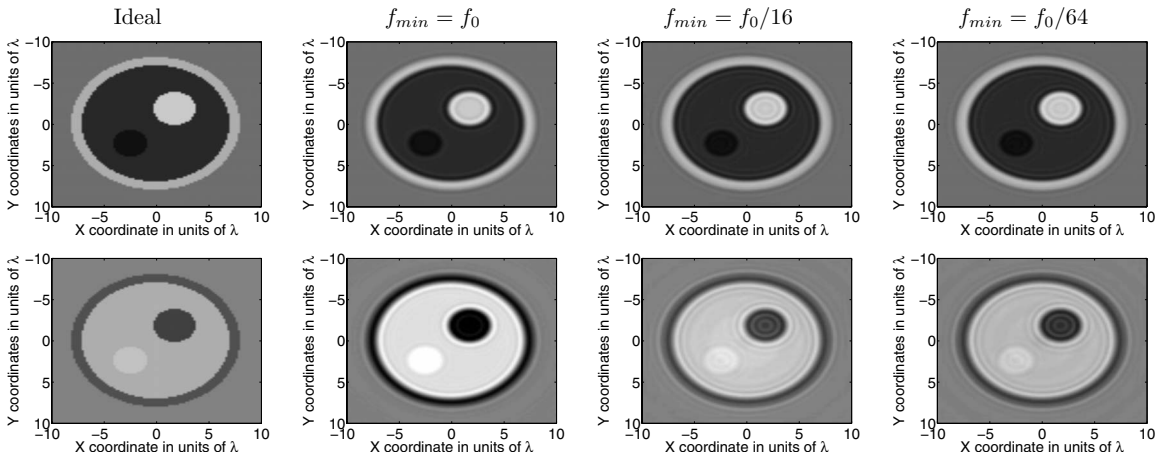


FIG. 3. Speed of sound (top row) and density (bottom row) images obtained using the multiple frequency T-matrix approach. First column: ideal profiles. Second columns: reconstructions using $f_{\min}$ values of f_0, $f_0/16$, and $f_0/64$, respectively.

that exhibit markedly different sizes or acoustic contrasts. In these cases, the finite precision of currently available floating-point arithmetic systems may prevent the proper computation and inversion of all the matrices involved in the solution presented in Sec. II.

IV. APPLICATION TO INVERSE SCATTERING PROBLEMS

The scattering solution presented in Sec. II was used to generate synthetic scattered data from a complex object in order to perform density imaging using the multiple frequency T-matrix approach.[11] The performance of this method when imaging homogeneous cylindrical objects has been reported previously.[12] Data were generated at frequencies $f_0, f_0/2, f_0/4, \ldots, f_{\min}$ and processed sequentially starting from the minimum frequency $f_{\min}$. Details of the implementation of the algorithm are given in Ref. 12. The properties and mean reconstructed values for all cylinders using $f_{\min}$ values of f_0, $f_0/16$, and $f_0/64$ are given in Table II. The reconstructions are shown in Fig. 3.

The speed of sound reconstructions exhibited high numerical accuracy independently of the value of $f_{\min}$. As for the density reconstructions, the maximum reduction in the bias between the ideal and reconstructed density values for a homogeneous circular cylinder of radius a should occur when $k_{\min}a \approx 1$ according to the results in Ref. 12. Therefore, the cylindrical inclusions of radius $1.6\lambda_0$ and $2\lambda_0$ should already exhibit the minimum achievable bias when using $f_{\min} = f_0/16$. However, the bias was reduced even further when using $f_{\min} = f_0/64$. Therefore, these results suggest that the absolute density values of a complex imaging target may not be obtained unless convergence is guaranteed for the overall structure when using the T-matrix approach.

V. CONCLUSIONS

The solution for the scattering produced by multiple parallel cylinders embedded inside a coated cylinder taking into account changes in compressibility, density, and acoustic attenuation has been presented. The applicability of the scattering solution presented here was demonstrated by analyzing the convergence of a method for variable density inverse scattering when imaging objects with multiple levels of spatial variations.

[1] J. Sinai and R. C. Waag, "Ultrasonic scattering by two concentric cylinders," J. Acoust. Soc. Am. **83**, 1728–1735 (1988).
[2] J. W. Young and J. C. Bertrand, "Multiple scattering by two cylinders," J. Acoust. Soc. Am. **58**, 1190–1195 (1975).
[3] S. E. Sherer, "Scattering of sound from axisymetric sources by multiple circular cylinders," J. Acoust. Soc. Am. **115**, 488–496 (2004).
[4] J. A. Roumeliotis, A.-G. P. Ziotopoulos, and G. C. Kokkorakis, "Acoustic scattering by a circular cylinder parallel with another of small radius," J. Acoust. Soc. Am. **109**, 870–877 (2001).
[5] V. Twersky, "Multiple scattering of radiation by an arbitrary configuration of parallel cylinders," J. Acoust. Soc. Am. **24**, 42–46 (1952).
[6] N. K. Uzunoglu and J. G. Fikioris, "Scattering from an infinite dielectric cylinder embedded into another," J. Phys. A **12**, 825–834 (1979).
[7] R. P. Parrikar, A. A. Kishk, and A. Z. Elsherbeni, "Scattering from an impedance cylinder embedded in a nonconcentric dielectric cylinder," IEE Proc. Microwaves, Antennas Propag. **138**, 169–175 (1991).
[8] L. G. Stratigaki, M. P. Ioannidou, and D. P. Chrissoulidis, "Scattering from a dielectric cylinder with multiple eccentric cylindrical dielectric inclusions," IEE Proc. Microwaves, Antennas Propag. **143**, 505–511 (1996).
[9] P. C. Waterman, "New formulation of acoustic scattering," J. Acoust. Soc. Am. **45**, 1417–1429 (1969).
[10] W. C. Chew, *Waves and Fields in Inhomogeneous Media* (IEEE, Piscataway, NJ, 1995).
[11] J. Lin and W. Chew, "Ultrasonic imaging by local shape function method with CGFFT," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 956–969 (1996).
[12] R. Lavarello and M. Oelze, "Density imaging using inverse scattering," J. Acoust. Soc. Am. **125**, 793–802 (2009).

Theoretical model for the threshold onset of contrast microbubble oscillations (L)

Alexander A. Doinikov[a] and Ayache Bouakaz
INSERM U930 CNRS ERL 3106, Université François Rabelais, CHU Bretonneau, 2 Boulevard Tonnellé, 37044 Tours Cedex 9, France

(Received 15 June 2009; revised 6 October 2009; accepted 6 December 2009)

It has been reported recently through high speed optical observations of phospholipid-coated contrast microbubbles that there is a threshold value for the acoustic pressure amplitude below which the radial oscillation of the microbubbles does not occur. In this Letter, it is suggested that this threshold behavior results from the fact that a phospholipid layer, as a physical material, has a certain value of the limiting shear stress so that its deformation does not start until this limiting value is exceeded. A theoretical model is proposed for the description of this phenomenon. The model explains the experimentally observed dependence of the threshold onset of microbubble oscillation on the initial bubble radius. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3278607]

PACS number(s): 43.25.Yw, 43.35.Ei, 43.80.Qf [AJS] Pages: 649–651

I. INTRODUCTION

Emmer *et al.* (2007a) investigated experimentally the onset of radial oscillation of phospholipid-coated contrast microbubbles (BR14) with radii ranging from 1 to 5.5 μm. The microbubbles were insonified at a driving frequency of 1.7 MHz and acoustic pressure amplitudes ranging from 20 to 250 kPa. It has been found that for smaller microbubbles, with resting radii $R_0 < 2.5$ μm, there is a threshold value for the acoustic pressure amplitude below which little or no oscillation is observed. The microbubble oscillation was not detected until the acoustic driving pressure reached the threshold value, and as the resting radius of the microbubbles decreased, the threshold pressure increased. In this study, it is hypothesized that the observed threshold behavior results from the fact that an encapsulating phospholipid layer, as a physical material, has a certain value of the limiting shear stress so that its deformation does not start until this limiting value is exceeded. A theoretical model is proposed for the description of this behavior. The model explains the experimentally observed dependence of the threshold onset of microbubble oscillation on the initial bubble radius.

II. THEORETICAL MODEL

It is well known that many elastic solids and viscous fluids start moving (deforming or flowing, respectively) only if the applied stress exceeds a critical value, which is called the limiting shear stress (Reiner, 1958). Let us denote this threshold by τ_{el} for an elastic solid and by τ_{vis} for a viscous fluid. If a medium possesses both elasticity and viscosity, it will begin to move when the applied stress exceeds the greater of the thresholds τ_{el} and τ_{vis}. Let us denote this greater threshold by τ_0.

The commonly accepted opinion is that phospholipid possesses both elasticity and viscosity. Therefore one can assume that phospholipid, as a physical material, has a certain value of the threshold shear stress τ_0. We suggest that it is just this general property inherent in most physical materials that is responsible for the threshold onset of radial oscillation of a phospholipid-coated microbubble. It should be emphasized, however, that τ_0 is treated here as the threshold shear stress for phospholipid as a material, not as the threshold value for a phospholipid spherical shell. The importance of this distinction will become clear below.

The fact that the phospholipid shell of a contrast microbubble is spherical and ultrathin poses the following question: How to relate the bulk physical properties of phospholipid to the behavior of a phospholipid shell, which is usually described by zero-thickness encapsulation models such as modified Rayleigh–Plesset equations (de Jong *et al.*, 1994; Morgan *et al.*, 2000; Sarkar *et al.*, 2005; Marmottant *et al.*, 2005)? This problem can be solved using the results of Church (1995) and Doinikov *et al.* (2009).

Generalizing the Rayleigh–Plesset equation for a free bubble (Plesset, 1949) to the case of a bubble enclosed in a finite-thickness shell, Church (1995) derived the following equation:

$$R_1\ddot{R}_1\left[1+\left(\frac{\rho_L}{\rho_S}-1\right)\frac{R_1}{R_2}\right]+\frac{3}{2}\dot{R}_1^2\left[1+\left(\frac{\rho_L}{\rho_S}-1\right)\times\left(\frac{4R_2^3-R_1^3}{3R_3^3}\right)\frac{R_1}{R_2}\right]=\frac{1}{\rho_S}\left[P_{g0}\left(\frac{R_{10}}{R_1}\right)^{3\gamma}-\frac{2\sigma_1}{R_1}-\frac{2\sigma_2}{R_2}-4\eta_L\frac{R_1^2\dot{R}_1}{R_2^3}-P_0-P_{\mathrm{ac}}(t)-S\right], \quad (1)$$

where $R_1(t)$ and $R_2(t)$ are the inner and the outer radii of the encapsulating shell, respectively, the overdot denotes the time derivative, ρ_L and ρ_S are the equilibrium densities of the surrounding liquid and the shell, respectively, P_{g0} is the equilibrium gas pressure within the bubble, γ is the ratio of specific heats of the gas, R_{10} and R_{20} are the inner and the outer radii of the shell at rest, σ_1 and σ_2 are the surface tension

[a]Author to whom correspondence should be addressed. Electronic mail: doinikov@bsu.by

coefficients for the gas-shell and the shell-liquid interfaces, respectively, η_L is the shear viscosity of the liquid, P_0 is the hydrostatic pressure in the liquid, and $P_{ac}(t)$ is the driving acoustic pressure at the location of the bubble. The effect of encapsulation is described by the term S, which is given by

$$S = -3\int_{R_1}^{R_2} \frac{\tau_{rr}(r,t)}{r} dr, \tag{2}$$

where r is the radial coordinate of a spherical coordinate system with the origin at the center of the bubble, and $\tau_{rr}(r,t)$ is the radial component of the stress deviator of the shell. For thin-shelled bubbles, such as phospholipid-coated microbubbles considered here, Eq. (1) reduces to

$$R\ddot{R} + \frac{3}{2}\dot{R}^2 = \frac{1}{\rho_L}\left[P_{g0}\left(\frac{R_0}{R}\right)^{3\gamma} - \frac{2\sigma}{R} - 4\eta_L\frac{\dot{R}}{R} - P_0 - P_{ac}(t) - S \right], \tag{3}$$

where $R(t)$ is the radius of the microbubble, R_0 is the resting value of $R(t)$, and σ is the surface tension at the gas-liquid interface. Doinikov *et al.* (2009) showed that in this limit the term S becomes

$$S = -(3\varepsilon/R)\tau_{rr}(r,t)|_{r=R}, \tag{4}$$

where ε denotes the shell thickness. This equation allows one to transform existing constitutive equations for the stress tensor τ_{ij}, which are normally specified in the bulk form, into a surface form that is required in Eq. (3). Equation (4) can be used to establish a criterion for the onset of oscillation of phospholipid-coated microbubbles.

Normally, it is assumed that S includes two parts, S_{el} and S_{vis}, which describe the shell elasticity and the shell viscosity. From Eq. (4) it follows that if a phospholipid layer, as a physical material, has the threshold shear stress equal to τ_0, then S should include a third part, S_0, which can be written as

$$S_0 = 3\varepsilon\tau_0/R_0. \tag{5}$$

The proposed criterion is that the microbubble oscillation starts when the acoustic pressure amplitude P_a exceeds the magnitude equal to S_0; i.e., the critical value of the acoustic pressure amplitude, P_a^{cr}, is equal to S_0. Note that, as follows from Eq. (5), the smaller the bubble radius, the greater the value of S_0, which is in agreement with Emmer *et al.*'s (2007a) experimental observations.

III. RESULTS AND DISCUSSION

Using the measurements of Emmer *et al.* (2007a), one can estimate the magnitude of τ_0 and thereby check the proposed theoretical model by comparing τ_0 with the value of the bulk elastic modulus reported in literature for lipid. Figure 1 displays the acoustic pressure threshold as a function of microbubble initial radius. The circles in Fig. 1 correspond to experimental data presented in Fig. 4 in Emmer *et al.* (2007a). The solid line shows a fit to the experimental points by Eq. (5). The fitting was made using the program package MATHEMATICA (Wolfram Research, Inc., Champaign, IL). The equation of the fitting curve is given by

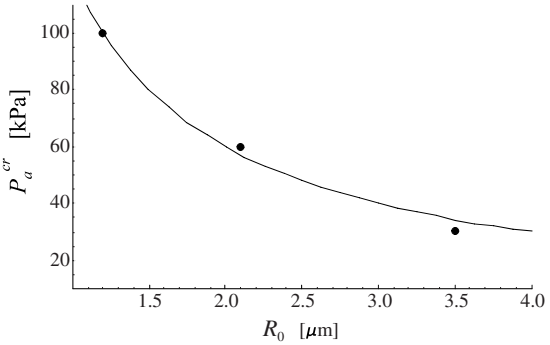


FIG. 1. Threshold amplitude of the driving pressure as a function of the resting bubble radius. Circles show experimental data from Emmer *et al.* (2007a).

$$P_a^{cr} = 0.12/R_0. \tag{6}$$

It follows that $\varepsilon\tau_0 \approx 0.04$ N/m. Assuming that the thickness of the phospholipid encapsulating layer is on the order of 1 nm (Morgan *et al.*, 2000), the magnitude of the critical shear stress for the phospholipid layer as a material should be of the order of 40 MPa. According to the experimental data of Marmottant *et al.* (2005), the bulk elastic modulus of a 1 nm thick phospholipid layer is estimated to be equal to 333.3 MPa; see also Doinikov and Dayton (2006). Thus the value of τ_0 is found to be smaller by almost one order of magnitude than the bulk elastic modulus. This is a reasonable difference for such material parameters, and hence this result confirms the plausibility of the proposed theory.

To sum up, the assumption that the threshold onset of radial oscillation of phospholipid-coated microbubbles is a consequence of the rheological properties of a phospholipid layer allows one to suggest a simple physical explanation for this experimental observation. This explanation is that phospholipid, as a physical material, has the limiting shear stress, which in turn results in the threshold onset of microbubble oscillation. The derived expression for the critical acoustic pressure, Eq. (5), reveals that the observed dependence of the threshold pressure on the resting bubble radius is a geometrical factor that results from the spherical shape of the phospholipid coating.

In conclusion, we should mention that experiments reveal that the threshold behavior of phospholipid-coated microbubbles is dependent on the ultrasound parameters of the excitation pulse, such as the driving frequency f, the acoustic pressure amplitude P_a, the number of pulses in the imposed ultrasound signal N, etc. (Emmer *et al.*, 2007b, 2007c). In this connection it should be noted that Eq. (5) alone cannot explain such effects since they result from the rheological properties of a phospholipid layer as a material. That is, they are accounted for in the quantity τ_0, which accordingly should be considered as a function of the parameters of excitation, $\tau_0 = \tau_0(f, P_a, N, \ldots)$. Further experiments are thus required in order to understand the specific rheological laws that can be applied to describe the dependence of τ_0 on the excitation parameters.

ACKNOWLEDGMENT

A.A.D. gratefully acknowledges the financial support from the le STUDIUM®, Institute for Advanced Studies (Orleans, France).

Church, C. C. (**1995**). "The effect of an elastic solid surface layer on the radial pulsations of gas bubbles," J. Acoust. Soc. Am. **97**, 1510–1521.

de Jong, N., Cornet, R., and Lancée, C. T. (**1994**). "Higher harmonics of vibrating gas-filled microspheres. Part one: Simulations," Ultrasonics **32**, 447–453.

Doinikov, A. A., and Dayton, P. A. (**2006**). "Spatio-temporal dynamics of an encapsulated gas bubble in an ultrasound field," J. Acoust. Soc. Am. **120**, 661–669.

Doinikov, A. A., Haac, J. F., and Dayton, P. A. (**2009**). "Modeling of nonlinear viscous stress in encapsulating shells of lipid-coated contrast agent microbubbles," Ultrasonics **49**, 269–275.

Emmer, M., Matte, G., van Neer, P., van Wamel, A., and de Jong, N. (**2007c**). "Improved ultrasound contrast agent detection in a clinical setting," in Proceedings of the IEEE Ultrasonics Symposium, New York, pp. 2235–2238.

Emmer, M., van Wamel, A., Goertz, D. E., and de Jong, N. (**2007a**). "The onset of microbubble vibration," Ultrasound Med. Biol. **33**, 941–949.

Emmer, M., Vos, H. J., van Wamel, A., Goertz, D. E., Versluis, M., and de Jong, N. (**2007b**). "Vibrating microbubbles at low acoustic pressures," in The Abstract Book of the 12th European Symposium on Ultrasound Contrast Imaging, Rotterdam, The Netherlands, p. 84.

Marmottant, P., van der Meer, S., Emmer, M., Versluis, M., de Jong, N., Hilgenfeldt, S., and Lohse, D. (**2005**). "A model for large amplitude oscillations of coated bubbles accounting for buckling and rupture," J. Acoust. Soc. Am. **118**, 3499–3505.

Morgan, K. E., Allen, J. S., Dayton, P. A., Chomas, J. E., Klibanov, A. L., and Ferrara, K. W. (**2000**). "Experimental and theoretical evaluation of microbubble behavior: Effect of transmitted phase and bubble size," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 1494–1509.

Plesset, M. S. (**1949**). "The dynamics of cavitation bubbles," ASME J. Appl. Mech. **16**, 277–282.

Reiner, M. (**1958**). *Rheology* (Springer-Verlag, Berlin).

Sarkar, K., Shi, W. T., Chatterjee, D., and Forsberg, F. (**2005**). "Characterization of ultrasound contrast microbubbles using in vitro experiments and viscous and viscoelastic interface models for encapsulation," J. Acoust. Soc. Am. **118**, 539–550.

The influence on predicted harmonic and distortion product generation of the position of the nonlinearity within cochlear micromechanical models (L)

Jacqueline A. How,[a] Stephen J. Elliott, and Ben Lineton
Institute of Sound and Vibration Research, University of Southampton, Southampton, Hampshire SO17 1BJ, United Kingdom

(Received 8 October 2009; revised 27 November 2009; accepted 7 December 2009)

Numerical techniques are used to explore the influence on the predicted basilar membrane (BM) response of the position of the nonlinearity within the micromechanical feedback loop of an active nonlinear cochlear model. This position is found to influence both the harmonic and distortion product spectra of the predicted BM response. The BM motion at the fundamental or primary frequencies is not significantly altered by the position of the nonlinearity, however, provided that the gain is appropriately adjusted. The observed effects are explained in terms of the frequency responses of the elements within the micromechanical feedback loop.
 [DOI: 10.1121/1.3279812]

PACS number(s): 43.64.Kc, 43.64.Bt, 43.64.Jb [BLM] Pages: 652–655

I. INTRODUCTION

Discrete models of the cochlea have been used extensively to describe and investigate the mechanics of the cochlea (Neely and Kim, 1986; Viergever and de Boer, 1987; Kanis and de Boer, 1993; Ku, 2008). In a linear active variant of this model, each micromechanical element can be described by a feedback diagram, as shown in Fig. 1(a) (Neely, 1985), where $p_{\mathrm{OHC}}^{\mathrm{lin}}$ is the pressure arising from the active outer hair cells (OHCs) and is assumed to act in the same way on the basilar membrane (BM) as the transmembrane pressure difference p.

Filter 1 corresponds to the passive mechanical admittance of the BM. Filter 2 is a simplified representation of the impedance of the active OHC complex, incorporating the many factors that contribute to the action of the OHCs, including linear components of the mechanoelectrical and electromechanical transduction processes associated with the cells. For example, filter 2 may encompass contributions from hair bundle motility (Martin and Hudspeth, 1999), processes within the OHC (Santos-Sacchi, 1989), extra-cellular processes, such as resonance of the tectorial membrane (TM) and stereocilia (Kanis and de Boer, 1993), and frequency-dependent phase shifts between the OHC pressure output and the radial displacement of the BM and TM (Neely and Kim, 1986). A nonlinearity, attributed to the OHC transduction process, can be positioned either *before* (Cooper, 1998) or *after* (Kanis and de Boer 1993) filter 2, as shown in Figs. 1(b) and 1(c), respectively. In other physical systems, such as chemical and thermodynamic processes, the position of the nonlinearity within a model is known to influence the predicted results (Aguirre *et al.*, 2005). As this effect has yet to be examined for cochlear models, we use numerical modeling to investigate the effect on the predicted BM response of changing the position of the nonlinearity within the micromechanical feedback loop. This is initially achieved by predicting the fundamental and low order harmonic components of the BM displacement evoked by single tone stimulation of the model. In addition, the influence on distortion products (DPs) arising from two tone excitation of the alternative micromechanical arrangements is also quantified.

II. MODEL AND METHOD

A discrete long-wave cochlear model is used, where the longitudinal coupling is mediated by the cochlear fluid and 1000 micromechanical elements are used to represent the 35 mm length of the human cochlea. The assumed passive admittance of the BM (filter 1) and the dynamics of the OHC complex (filter 2) correspond to those suggested by Kanis and de Boer (1993) for a human cochlear model. The frequency response functions for the filters, observed at the 4 kHz characteristic place in the cochlear model, are shown in Fig. 2. Details of the boundary conditions and the two port model used to represent the ear canal and middle ear are described by Elliott *et al.* (2007) and Ku (2008).

The nonlinearity within each micromechanical element is assumed to be a first order Boltzmann function (Cooper, 1998), defined as follows:

$$f(u) = \frac{\alpha}{\eta}\left(\frac{1}{1+\beta e^{-\eta u/\gamma}} - \frac{1}{1+\beta}\right). \quad (1)$$

The value of α, which determines the maximum output of the nonlinearity and has the same dimensions as the input u, increases from base to apex so that the coupled model exhibits a compression knee-point at 30 dB sound pressure level (SPL) for frequencies between 0.5 and 6 kHz. The spatial distribution of α differs between the two alternative micromechanical arrangements to ensure that this assumed condition is satisfied in both cases. The dimensionless constant η is assigned a value of either 2.5×10^{10} or 1, when the nonlinearity is positioned either before or after filter 2, respec-

[a]Author to whom correspondence should be addressed. Electronic mail: jh@isvr.soton.ac.uk

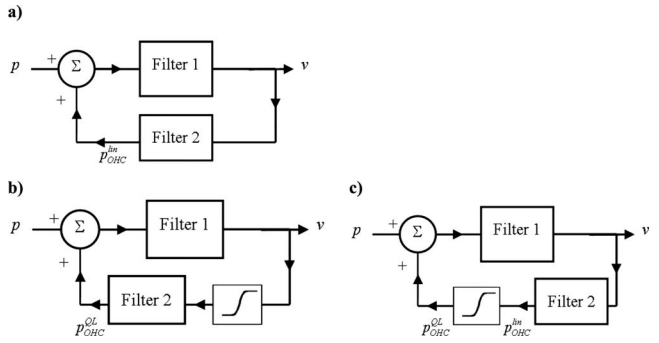


FIG. 1. Block diagrams representing the micromechanical feedback loop within each element of the discrete cochlear model. Examples of (a) a linear active model and [(b) and (c)] two alternative active nonlinear models are shown. The pressure difference across the BM and the BM velocity are denoted by p and v in each diagram. The linear pressure output and the quasilinear pressure output of the OHCs are represented by $p_{\mathrm{OHC}}^{\mathrm{lin}}$ and $p_{\mathrm{OHC}}^{\mathrm{QL}}$, respectively. Filters 1 and 2 represent the passive admittance of the basilar membrane and the impedance of the OHC complex, respectively.

tively, to compensate for the change in the units of the input to the nonlinearity. The dimensionless constant β controls the asymmetry of the function and is set equal to 1.2. The value of γ, which has the same dimensions as the input, is selected so that the nonlinearity provides 0 dB gain to the fundamental component at low stimulus levels. The parameters β and γ influence the amplitude of even- and odd-order

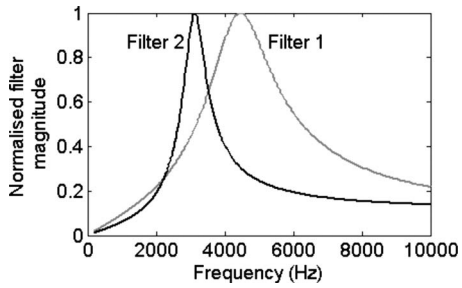


FIG. 2. Assumed frequency response functions for filters 1 and 2, associated with the passive BM admittance and the impedance of OHC complex, respectively, observed at the 4 kHz characteristic place in the cochlear model.

harmonic components of the nonlinear output, respectively (Cooper, 1998).

The fundamental response of the model is estimated using an iterative, quasilinear method devised by Kanis and de Boer (1993). This approach, with modifications described in the online appendix to improve the convergence, was also applied to estimate harmonic components and DPs.

III. RESULTS

A. Single tone excitation

For single tone stimulus levels of 60 and 80 dB SPL between 1 and 7 kHz, the fundamental component of the

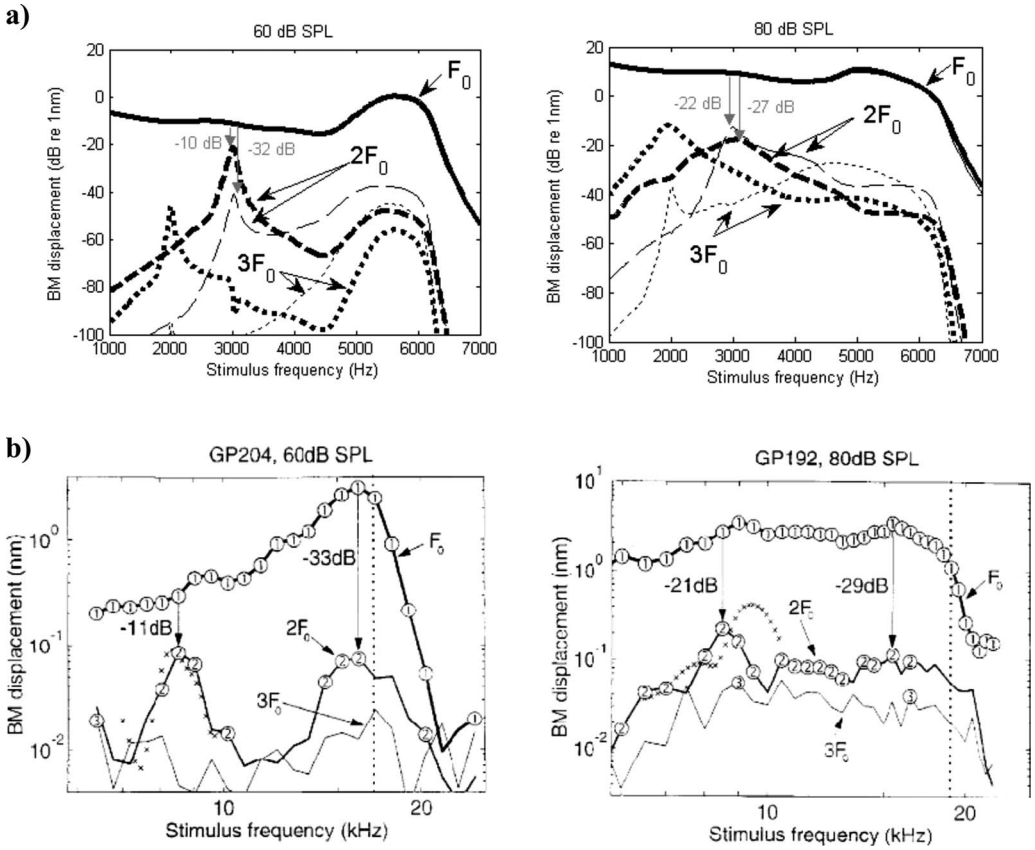


FIG. 3. (a) The predicted spectra of BM displacement evoked by single stimulus tones at the fundamental (solid line), second harmonic (dashed line), and third harmonic (dotted line) frequencies, observed at a distance of 81 mm from the stapes, which corresponds to the 6 kHz characteristic place. The results are shown for two cochlear models: one where the nonlinear is positioned before filter 2 (thick lines), and one where the nonlinearity is positioned after filter 2 (thin lines). Stimulus levels of 60 and 80 dB SPL have been used. (b) The observations of Cooper (1998) of the BM displacement response to single tone stimuli in the guinea pig cochlea at the 17–19 kHz characteristic place, used with permission. The crosses indicate the fundamental responses to low level stimuli and can be neglected when comparing (a) and (b).

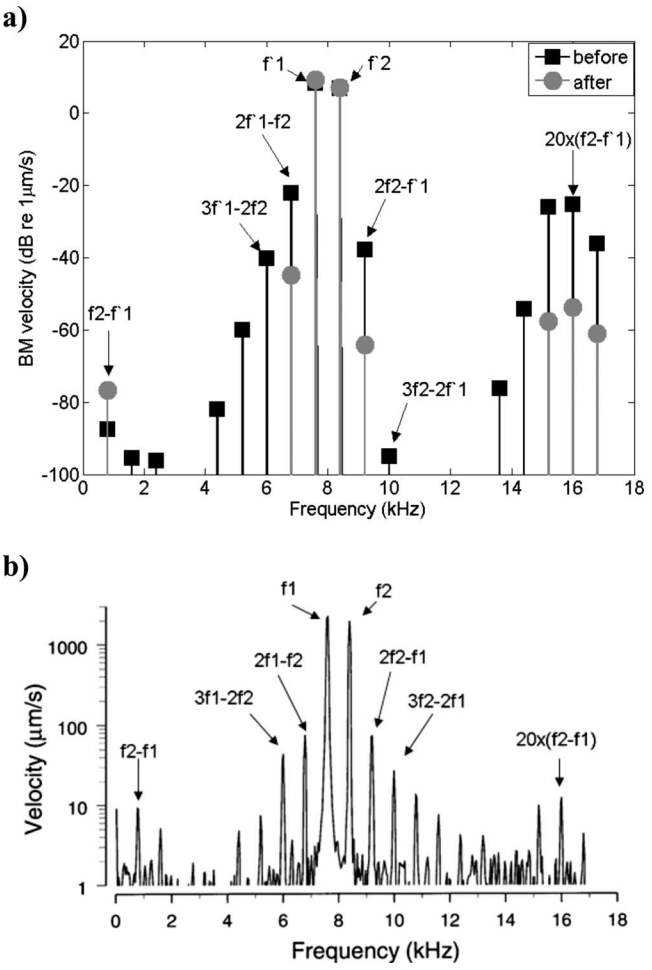


FIG. 4. (a) The predicted spectrum of the BM velocity at the f_2 characteristic place in the cochlear model ($L_1=L_2=70$ dB SPL) with the nonlinearity positioned before (squares) or after (circles) filter 2, and (b) the physiological measurement of Robles *et al.* (1997) at the 8 kHz characteristic place in the chinchilla cochlea ($L_1=L_2=80$ dB SPL), used with permission. In both cases, stimuli were presented at $f_1=7.6$ kHz and $f_2=8.4$ kHz.

predicted BM displacement differs by less than 5 dB between the two micromechanical arrangements [Fig. 3(a)]. However, the amplitude of the estimated harmonic components is significantly influenced by the position of the nonlinearity within the micromechanical feedback loop. For example, the relative amplitude of the predicted fundamental and second harmonic responses to a 3 kHz tone at 60 dB SPL decreases by 22 dB, from −10 to −32 dB relative to the fundamental response, when the nonlinearity is moved from before to after filter 2.

B. Two tone stimulation

When two tones with frequencies f_1 and f_2 are presented simultaneously to the model, DPs are predicted in the BM response, as shown in Fig. 4(a). The amplitude of the predicted $2f_1-f_2$ and $2f_2-f_1$ DPs is decreased by 23 and 26 dB, respectively, when the nonlinearity is moved from before to after filter 2. In contrast, the predicted BM velocity at the primary frequencies does not change by more than 2 dB for the same change in position.

IV. DISCUSSION AND CONCLUSIONS

The simulations of pure tone responses are compared to the experimental results obtained by Cooper (1998) from a guinea pig cochlea shown in Fig. 3(b). There are amplitude differences between the simulations and the physiological observations. These differences may arise because the assumed middle ear response of the model differs from that of the guinea pig, resulting in different effective stimulus levels, or because the parameters of the model are primarily adjusted to replicate the response of the human cochlea. However, it is interesting to note that the relative difference between the amplitude of the fundamental and second harmonic responses to a 60 dB stimulus observed by Cooper (1998) is 11 dB when the second harmonic frequency is approximately equal to the best frequency of the observation location, and that this is best replicated by the model when the nonlinearity is positioned before, rather than after, filter 2. The position of the nonlinearity within the micromechanical feedback loop does not have a substantial effect on the fundamental response. The micromechanical arrangement affects the harmonic response because the impedance of the OHC complex, represented by filter 2 in Fig. 1, is frequency dependent (Fig. 2). For example, when considering the effect of the feedback loop on the second harmonic response, the signal of most interest is at the fundamental frequency f_0 before the nonlinearity, but changes to $2f_0$ in the nonlinearity's output. Therefore if the nonlinearity is positioned before filter 2, filter 2 operates on a signal at frequency $2f_0$, but if the nonlinearity is positioned after filter 2, then filter 2 acts on a signal at frequency f_0. The extent to which the position of the nonlinearity influences the harmonic responses is thus expected to depend on the nature of the impedance used in the model to represent the OHC complex.

The frequency dependence of filter 2 is also responsible for the influence of the position of the nonlinearity on the predicted DP amplitudes since it acts on either the DP frequency, or the primary frequencies individually, when the nonlinearity is placed either before or after filter 2, respectively. For example, Robles *et al.* (1997) found that the amplitude of the $2f_1-f_2$ DP is approximately 28 dB below the amplitude of the primary frequencies in their observation of the chinchilla cochlea shown in Fig. 4(b). The simulation uses a lower stimulus level than the physiological measurement in order to achieve a similar degree of saturation, as a consequence of the middle ear differences. The relative amplitude of the predicted $2f_1-f_2$ DP to the primaries decreases from −31 to −54 dB when the nonlinearity is moved from before to after filter 2, and so the former model again appears to be most consistent with the measured data. Despite the considerable effect on the DPs, the amplitudes of the primary responses are not altered significantly by the placement of the nonlinearity.

In summary, the position of the nonlinearity within the micromechanical feedback loop can substantially affect simulations of harmonic responses and distortion products, without significantly influencing the fundamental response if the gain is adjusted appropriately. The extent of this effect will depend on the form of the cochlear model used in the

simulations, as it is attributed to the assumed frequency dependence of the impedance attributed to the OHC complex.

Aguirre, L. A., Coelho, M. C. S., and Correa, M. V. (**2005**). "On the interpretation and practice of dynamical differences between Hammerstein and Wiener models," IEE Proc.: Control Theory Appl. **152**, 349–356.

Cooper, N. P. (**1998**). "Harmonic distortion on the basilar membrane in the basal turn of the guinea-pig cochlea," J. Physiol. (London) **509**, 277–288.

Elliott, S. J., Ku, E. M., and Lineton, B. (**2007**). "A state space model for cochlear mechanics," J. Acoust. Soc. Am. **122**, 2759–2771.

Kanis, L., and de Boer, E. (**1993**). "Self-suppression in a locally active nonlinear model of the cochlea: A quasilinear approach," J. Acoust. Soc. Am. **94**, 3199–3206.

Ku, E. M. (**2008**). "Modelling the human cochlea," Ph.D. thesis, University of Southampton, United Kingdom.

Martin, P., and Hudspeth, A. J. (**1999**). "Active hair-bundle movements can amplify a hair cell's response to oscillatory mechanical stimuli," Proc. Natl. Acad. Sci. U.S.A. **96**, 14306–14311.

Neely, S. T. (**1985**). "Mathematical modelling of cochlear mechanics," J. Acoust. Soc. Am. **78**, 345–352.

Neely, S. T., and Kim, D. O. (**1986**). "A model for active elements in cochlear biomechanics," J. Acoust. Soc. Am. **79**, 1472–1480.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1997**). "Two-tone distortion on the basilar membrane of the chinchilla cochlea," J. Neurophysiol. **77**, 2385–2399.

Santos-Sacchi, J. (**1989**). "Asymmetry in voltage-dependent movements of isolated outer hair cells from the organ of Corti," J. Neurosci. **9**, 2954–2962.

Viergever, M. A., and de Boer, E. (**1987**). "Matching impedance of a non-uniform transmission line: Application to cochlear modelling," J. Acoust. Soc. Am. **81**, 184–186.

Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users (L)

Heather A. Kreft,[a] Andrew J. Oxenham,[b] and David A. Nelson
Department of Otolaryngology, Clinical Psychoacoustics Laboratory, University of Minnesota, Rm 8-323 Phillips-Wangensteen Building, 420 Delaware Street Southeast, Minneapolis, Minnesota 55455

(Received 1 September 2009; revised 3 December 2009; accepted 9 December 2009)

Detection and modulation rate discrimination were measured in cochlear-implant users for pulse-trains that were either sinusoidally amplitude modulated or were modulated with half-wave rectified sinusoids, which in acoustic hearing have been used to simulate the response to low-frequency temporal fine structure. In contrast to comparable results from acoustic hearing, modulation rate discrimination was not statistically different for the two stimulus types. The results suggest that, in contrast to binaural perception, pitch perception in cochlear-implant users does not benefit from using stimuli designed to more closely simulate the cochlear response to low-frequency pure tones. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3282947]

PACS number(s): 43.66.Ts, 43.66.Fe, 43.66.Hg [JCM] Pages: 656–659

I. INTRODUCTION

Cochlear-implant (CI) users often have difficulty with speech recognition in noisy situations (e.g., Friesen *et al.*, 2001; Nelson *et al.*, 2003) as well as with pitch and music perception (e.g., McDermott, 2004; Gfeller *et al.*, 2007; Galvin *et al.*, 2007). A number of studies have suggested that pitch coding deficits might underlie some of the speech-perception difficulties faced by CI users in noisy situations by impairing CI users' ability to make use of F0-based segregation cues (Stickney *et al.*, 2004; Turner *et al.*, 2004; Qin and Oxenham, 2005). Some efforts have been made to improve F0 representation in CIs, so far with relatively little success (Geurts and Wouters, 2001; Green *et al.*, 2005; Vandali *et al.*, 2005; Laneau *et al.*, 2006). Most recently, Landsberger (2008) compared pitch discrimination using a number of different waveforms, including sine, sawtooth, sharpened sawtooth, and square waves. The results suggested no great benefit of any of the waveforms, leading Landsberger to conclude that the waveform shapes may be used interchangeably and their use in future speech processing strategies may ultimately depend on other engineering issues.

It remains possible that a more "optimal" waveform, or class of waveforms, exists. For instance, it may be that CI algorithms should strive to approximate as closely as possible the temporal structure of neural responses that are produced by the normal ear in response to temporal fine structure. Conveying temporal fine-structure information via an envelope of a high-frequency carrier has been attempted in normal-hearing listeners using so-called "transposed stimuli" (TS). van der Par and Kohlrausch (1997) proposed modulating a high-frequency sinusoidal carrier with a half-wave rectified and lowpass-filtered low-frequency stimulus to produce a temporal envelope that is intended to produce a response in the auditory nerve that resembles the response to the low-frequency stimulus itself. In this way the neural response to the low-frequency stimulus is "transposed" to a higher-frequency cochlear location. This approach has been used to show that some aspects of binaural hearing, such as binaural masking level differences, and interaural time resolution and lateralization for frequencies below 150 Hz, can be represented as well in high-frequency (4000 Hz and above) as in low-frequency cochlear regions (e.g., Henning, 1974; van der Par and Kohlrausch, 1997; Bernstein and Trahiotis, 2002). Most encouragingly, Long *et al.* (2006) used half-wave rectified (HWR) stimuli, mimicking the TS used in normal-hearing listeners, to show that CI users can benefit from information normally carried by the temporal fine structure to obtain a binaural masking release, thereby demonstrating the potential advantages of bilateral stimulation for spatial hearing and speech masking release.

It remains to be seen whether HWR stimuli can improve pitch perception for CI users. Results from acoustic studies in normal-hearing listeners showed that TS produced somewhat better pitch discrimination than sinusoidally amplitude modulated (SAM) stimuli (Oxenham *et al.*, 2004), although performance did not match that found using low-frequency pure tones. The aim of this study was to test pitch perception, quantified as modulation rate discrimination of high-rate pulse trains, by comparing the more traditional SAM with HWR stimuli. The hypothesis was that HWR stimuli may provide more accurate pitch perception by simulating more accurately the neural temporal patterns normally evoked by low-frequency pure tones.

II. METHODS

A. Subjects and implants

The subjects were seven postlingually deafened adults (two males, five females) with either a Clarion C-II or HiRes90K cochlear implant. Two subjects were bilaterally im-

[a]Author to whom correspondence should be addressed. Electronic mail: plumx@umn.edu
[b]Also affiliated with the Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455.

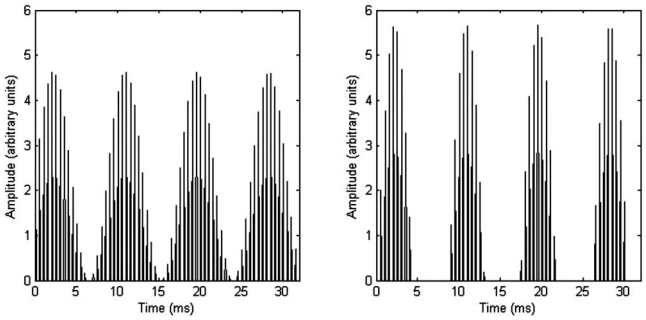


FIG. 1. Schematic diagram of the SAM stimulus (left panel) and HWR sinusoidal stimulus (right panel) with the same rms level. In this example, the pulse rate is 2000 Hz and the modulation rate is 115 Hz.

planted, and each ear was tested separately, resulting in a total of nine test ears. Full insertion of the electrode array (25 mm) was achieved in all cases. The age of the subjects ranged from 49 to 76 (mean age: 58 yr); the duration of CI use prior to the experiment ranged from 4 months to 5 years (mean duration: 2.8 yr); and the duration of deafness prior to implantation ranged from 1 to 28 years (mean duration: 14 yr). For the present study, stimulation was monopolar, with the active intracochlear electrode referenced to an electrode in the case of the internal receiver-stimulator. Electrodes were selected at three points across the array (apical, middle, and basal). In all cases, except for one subject's apical position (electrode 1), the test electrodes were 2, 8, and 14.

B. Stimuli and procedures

Experiments were controlled by a personal computer running custom programs written for the Bionic Ear Data Collection System (BEDCS; Advanced Bionics, 2003) and MATLAB (The Mathworks, Natick, MA). Stimuli were 300-ms trains of 32 μs/phase, cathodic-first biphasic pulses, presented at a rate of 2000 pulses per second. The amplitude of the pulses was modulated either by sinusoidal amplitude modulation (SAM condition) or by HWR sinusoidal modulation (HWR condition). Both the SAM and HWR conditions utilized 100% modulation depth with zero starting phase on each trial. Base modulation rates of 80, 115, 160, 230, and 320 Hz were tested. A schematic diagram of the two stimulus types is provided in Fig. 1. The levels of the stimuli were adjusted for each subject individually to correspond to either the 40% or 80% point of the subject's dynamic range (DR). The DR was defined as the range of currents (in μA) between absolute threshold (THS) and maximum acceptable loudness (MAL).

The THS and MAL were measured individually for each stimulus type at each modulation rate. The THS was measured using a two-interval, two-alternative forced-choice (2IFC/2AFC) procedure with a three-down, one-up adaptive tracking rule (Levitt, 1971). Correct-answer feedback was provided after each trial. The THS estimates from three to five tracks were averaged to obtain a final THS value for each subject. The MAL was measured using a one-up one-down adaptive tracking procedure in which the sound was presented, followed by the question "Was it too loud?." A subject's "no" and "yes" choices led to increases and decreases in signal level, respectively. The track terminated when the subject had responded that the intensity was too loud six times, and the average current level at the last six reversals was calculated. The MAL estimates from three tracks were averaged to obtain a final value of MAL for each subject.

Once the THS and MAL had been established for each subject, modulation rate discrimination thresholds were measured at each of the base modulation rates (80, 115, 160, 230, and 320 Hz) for both SAM and HWR modulations, using levels corresponding to 40% and 80% of the subject's DR. Discrimination thresholds were obtained using a 2IFC/2AFC procedure. No training was provided; however, 4 of the 7 subjects had extensive previous psychophysical testing experience. In order to reduce any potential loudness cues, the current level was roved (jittered) across intervals by $\pm 10\%$ DR for all subjects except one. For the remaining subject (also the subject for whom electrode 1 was tested), the jitter was set to $\pm 3\%$ DR, as this subject reported discomfort at jitter amounts greater than this. The two 300-ms stimulus intervals in each trial were separated by a silent interval of 500 ms. Within each of the two stimulus intervals, the pulse trains were modulated at rates of f and $f+\Delta f$, with the two modulation rates geometrically centered around the base modulation rate. The order of presentation of the two rates was selected randomly on each trial, and the subjects were instructed to select the interval containing the stimulus with the higher pitch. Correct-answer feedback was provided after each trial.

The value of Δf, expressed as a percentage of f, was varied adaptively using a three-down one-up procedure. Threshold was defined as the geometric mean value of Δf at the last six reversal points. Three such threshold estimates were averaged to obtain a single modulation rate discrimination threshold at each level for each electrode site. Conditions were tested in a blocked randomized order, with all conditions tested once before any were repeated. The order of presentation was selected randomly between subjects and between repetitions.

III. RESULTS AND DISCUSSION

A. THS and MAL

The THS and MAL measures for the apical, middle, and basal electrode locations, averaged across subjects, are plotted in terms of rms current as a function of modulation base rate in the left, middle, and right panels of Fig. 2, respectively. Generally, the THS values were similar for both HWR and SAM conditions, with average values (across modulation rate, electrode location, and subject) of 89.4 and 95.8 μA, respectively. The MAL values, in terms of rms current level, were higher for the SAM than for the HWR conditions. There was also a tendency for MAL values to increase from the apical to the basal electrode location, but the effect of modulation rate was less clear.

The fact that SAM and HWR produced similar THS values but different MAL values is in line with the results of Zhang and Zeng (1997). They used analog stimuli and found that detection thresholds were determined primarily by the

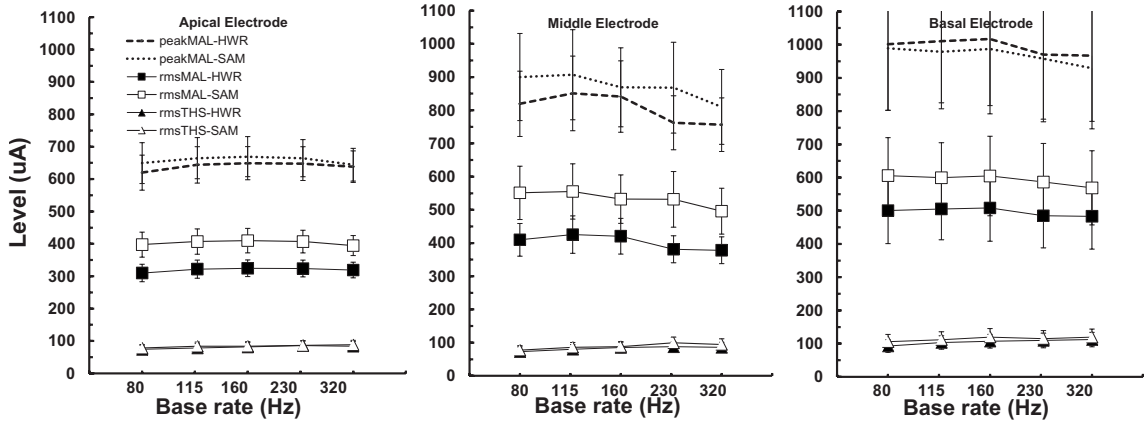


FIG. 2. THS and MAL as a function of base modulation rate averaged across subjects. Each panel represents the data for a different electrode location across the array. Open and filled symbols represent SAM and HWR stimuli, respectively. Triangles show THS values, squares show MAL values, and circles replot MAL values in terms of peak (rather than rms) amplitude. Error bars represent 1 s.e. of the mean.

rms amplitude of dynamic waveforms, whereas MAL was determined primarily by the peak amplitude. Using pulsatile stimuli, McKay and Henshall (2009) concluded that the relationship between modulated and unmodulated stimuli depends not only on level but also on pulse rate, in line with the model predictions of McKay *et al.* (2003). For a given rms amplitude, the peak amplitude of the HWR waveform is about 1.76 dB higher than that of the SAM waveform. This relationship is illustrated in Fig. 2 by the filled and open circles, which represent peak amplitudes at MAL for the HWR and SAM waveforms, respectively. Plotting MAL in terms of peak amplitude yields more similar results; in fact, the difference between HWR and SAM MAL values is not statistically significant when measured in terms of peak amplitude $[F(1,8)=1.912,\ p=0.204]$.

B. Modulation rate discrimination

Figure 3 shows the just-noticeable change in modulation rate (Δf), expressed as a percentage of the lower rate (f), as a function of base modulation rate at current levels corresponding to 40% (squares) and 80% (triangles) of the subjects' dynamic range. Open and filled symbols represent thresholds using SAM and HWR waveforms, respectively. As in Fig. 2, the left, middle, and right panels show data from the apical, middle, and basal electrodes, respectively. There was no obvious advantage for the HWR stimuli compared to the SAM stimuli. The size of the difference limens (DLs) for both HWR and SAM conditions was 6%–10% at the lowest base rate. This finding is consistent with other studies of modulation rate discrimination in cochlear-implant users (Geurts and Wouter, 2001; Baumann and Nobbe, 2004; Chatterjee and Peng, 2008; Landsberger, 2008).

In general, modulation rate DLs were lower (better) for the higher level stimulation (80% DR) than for the lower level (40% DR). There was also a clear trend for thresholds to increase (degrade) with increasing base modulation rate. In fact, thresholds were often so high for the highest two base rates (>30% of the base rate) that it seems unlikely that the stimuli evoked a reliable difference in pitch height. For instance, a 30% DL at a base rate of 230 Hz implies modulation rates for f and $(f+\Delta f)$ of 198 and 267 Hz, respectively. It may be that such large differences were based on selecting the stimulus that produced the weaker, rather than higher, pitch percept. Modulation rate DLs were unaffected by electrode location. A repeated-measures analysis of variance formally tested all of the parameters associated with the

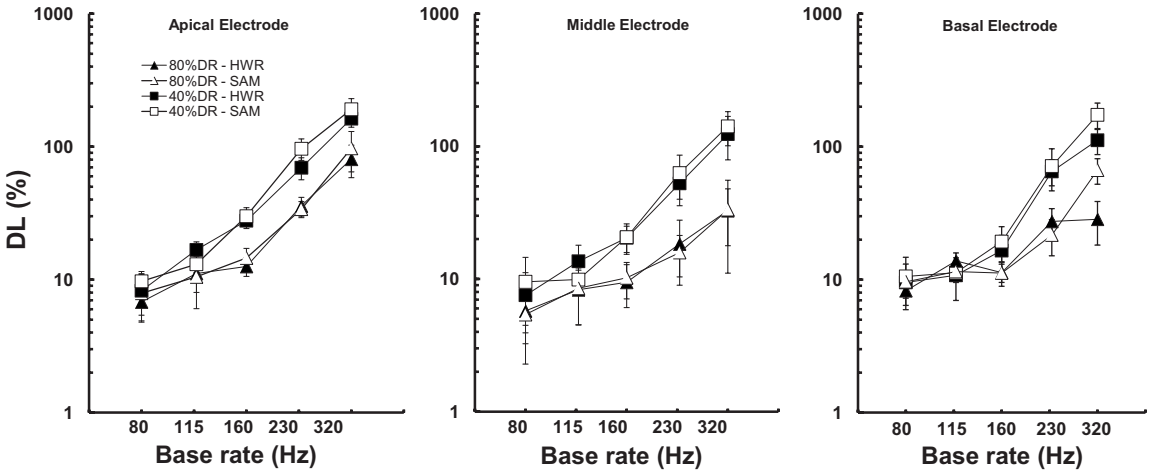


FIG. 3. Modulation rate DLs expressed as a percentage of the base rate averaged across subjects. Each panel represents the data for a different electrode location across the array. Error bars represent 1 s.e. of the mean.

modulation rate DLs. The effect of modulation rate was significant [$F(4,32)=40.864$, $p<0.001$], as was the effect of level [$F(1,8)=25.839$, $p=0.001$], whereas the effect of electrode location [$F(2,16)=2.732$, $p=0.127$] and waveform type [$F(1,8)=2.052$, $p=0.190$] was not. The only interaction term that was statistically significant was level $\times$ modulation rate [$F(4,32)=8.767$, $p<0.005$]. This pattern of results is in line with earlier studies of pulse-rate or modulation rate discrimination (for a review, see Moore and Carlyon, 2005).

The results do not support our initial hypothesis that HWR sinusoidal waveforms might provide more accurate modulation rate information than SAM waveforms. Our conclusion is consistent with that of Landsberger (2008), who tested a number of different waveform shapes and found that there were no significant differences in terms of rate discrimination abilities. The relatively poor performance in rate discrimination with temporal cues is in contrast to the relatively good performance in binaural masking level differences found by Long *et al.* (2006) also using the principle of TS. The fact that a benefit was found for spatial hearing but not for pitch is consistent with the idea that the temporal coding required for accurate pitch is different in nature from that required by the binaural system (e.g., Oxenham *et al.*, 2004). However, although the pattern of results is in the same direction as that found for normal-hearing listeners with acoustic stimulation, the details are somewhat different: normal-hearing listeners showed some benefit of TS over SAM in rate discrimination (Oxenham *et al.*, 2004), whereas CI users appear to show none. Taken together with previous results, the results suggest that accurate pitch discrimination in normal-hearing listeners requires place and/or place-time information (e.g., Loeb *et al.*, 1983; Shamma and Klein, 2000), and that accurate pitch coding in CI users may not be achieved by temporal patterns alone, but may require accurate spatial and temporal representation of sound along the tonotopic array.

ACKNOWLEDGMENTS

This research was supported by NIDCD Grant No. R01 DC 006699 and the Lions 5M International Hearing Foundation. The authors thank Advanced Bionics Corporation for providing the BEDCS research interface and Leo Litvak for providing advice and assistance in its implementation. The authors thank the Associate Editor, John Middlebrooks, and three anonymous reviewers for their helpful comments on an earlier version of this manuscript. The authors wish to extend special thanks to the subjects who participated in this study.

Advanced Bionics (**2003**). Bionic Ear Data Collection System, Version 1.16 User Manual.

Baumann, U., and Nobbe, A. (**2004**). "Pulse rate discrimination with deeply inserted electrode arrays," Hear. Res. **196**, 49–57.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," J. Acoust. Soc. Am. **112**, 1026–1036.

Chatterjee, M., and Peng, S.-C. (**2008**). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition," Hear. Res. **235**, 143–156.

Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Am. **110**, 1150–1163.

Galvin, J. J., III, Fu, Q. J., and Nogaki, G. (**2007**). "Melodic contour identification by cochlear implant listeners," Ear Hear. **28**, 302–319.

Geurts, L., and Wouters, J. (**2001**). "Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants," J. Acoust. Soc. Am. **109**, 713–726.

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., and Olszewski, C. (**2007**). "Accuracy of cochlear implant recipients on pitch perception, melody recognition and speech reception in noise," Ear Hear. **28**, 412–423.

Green, T., Faulkner, A., Rosen, S., and Macherey, O. (**2005**). "Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification," J. Acoust. Soc. Am. **118**, 375–385.

Henning, G. B. (**1974**). "Detectability of interaural delay in high-frequency complex waveforms," J. Acoust. Soc. Am. **55**, 84–90.

Landsberger, D. M. (**2008**). "Effects of modulation wave shape on modulation frequency discrimination with electric hearing," J. Acoust. Soc. Am. **124**, EL21–EL27.

Laneau, J., Wouters, J., and Moonen, M. (**2006**). "Improved music perception with explicit pitch coding in cochlear implants," Audiol. Neuro-Otol. **11**, 38–52.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Loeb, G. E., White, M. W., and Merzenich, M. M. (**1983**). "Spatial cross correlation: A proposed mechanism for acoustic pitch perception," Biol. Cybern. **47**, 149–163.

Long, C. J., Carlyon, R. P., Litovsky, R. Y., and Downs, D. H. (**2006**). "Binaural unmasking with bilateral cochlear implants," J. Assoc. Res. Otolaryngol. **7**, 352–360.

McDermott, H. J. (**2004**). "Music perception with cochlear implants: A review," Trends Amplif. **8**, 49–82.

McKay, C. M., and Henshall, K. R. (**2010**). "Amplitude modulation and loudness in cochlear implants," J. Assoc. Res. Otolaryngol. (in press).

McKay, C. M., Henshall, K. R., Farrell, R. J., and McDermott, H. J. (**2003**). "A practical method of predicting the loudness of complex electrical stimuli," J. Acoust. Soc. Am. **113**, 2054–2063.

Moore, B. C. J., and Carlyon, R. P. (**2005**). "Perception of pitch by people with cochlear hearing loss and by cochlear implant users," in *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, R. Fay, and A. N. Popper (Springer, New York).

Nelson, P. B., Jin, S.-H., Carney, A. E., and Nelson, D. A. (**2003**). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 961–968.

Oxenham, A. J., Bernstein, J. G. W., and Penagos, H. (**2004**). "Correct tonotopic representation is necessary for complex pitch perception," Proc. Natl. Acad. Sci. U.S.A. **101**, 1421–1425.

Qin, M. K., and Oxenham, A. J. (**2005**). "Effects of envelope-vocoder processing on F0 discrimination and concurrent-vowel identification," Ear Hear. **26**, 451–460.

Shamma, S., and Klein, D. (**2000**). "The case of the missing pitch templates: How harmonic templates emerge in the early auditory system," J. Acoust. Soc. Am. **107**, 2631–2644.

Stickney, G. S., Zeng, F. G., Litovsky, R., and Assmann, P. (**2004**). "Cochlear implant speech recognition with speech maskers," J. Acoust. Soc. Am. **116**, 1081–1091.

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., and Henry, B. A. (**2004**). "Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing," J. Acoust. Soc. Am. **115**, 1729–1735.

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W., and McDermott, H. J. (**2005**). "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies," J. Acoust. Soc. Am. **117**, 3126–3138.

van der Par, Z., and Kohlraush, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," J. Acoust. Soc. Am. **101**, 1671–1680.

Zhang, C., and Zeng, F. G. (**1997**). "Loudness of dynamic stimuli in acoustic and electric hearing," J. Acoust. Soc. Am. **102**, 2925–2934.

Psychophysical tuning curves and recognition of highpass and lowpass filtered speech for a person with an inverted V-shaped audiogram (L)

Vinay[a)]
All India Institute of Speech and Hearing, Manasagangothri, Mysore 570006, India

Brian C. J. Moore[b)]
Department of Experimental Psychology, Cambridge University, Downing Street, Cambridge CB2 3EB, England

(Received 26 October 2009; revised 26 November 2009; accepted 3 December 2009)

A single subject whose audiogram resembled an inverted V shape (good hearing at 4000 Hz, and poorer hearing at other frequencies) was tested. Results of the threshold-equalizing noise test suggested that a dead region (DR) in the cochlea was present at all test frequencies from 500 to 3000 Hz, but no DR was present at 4000 Hz. Psychophysical tuning curves (PTCs) obtained using signal frequencies of 2000, 3000, 4000, and 6000 Hz showed upward shifted tips for the lowest two signal frequencies, and a downward shifted tip for the highest frequency. The results of the PTCs suggested a functioning region extending from 3900 to 5100 Hz, with DRs outside that range. The identification of nonsense syllables, amplified according to the "Cambridge formula," was measured as a function of lowpass or highpass filter cutoff frequency. The results suggested that useful speech information could only be extracted from a limited frequency range around 4000 Hz. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277218]

PACS number(s): 43.71.Ky, 43.66.Ts, 43.71.Gv [CJP] Pages: 660–663

I. INTRODUCTION

Sometimes, but rather rarely, a hearing-impaired individual presents with an audiogram which has a form resembling an inverted neural or psychophysical tuning curve (PTC); the absolute threshold is relatively low (good) at one audiometric frequency, or over a very small range of frequencies, and worsens for frequencies on either side; Moore and Alcántara (2001) and Halpin (2002) presented examples of such cases. Halpin (2002) called such audiograms "tuning curve audiograms," which we will abbreviate as TCA. He suggested that they might be associated with dead regions (DRs) in the cochlea. A DR is a place in the cochlea where there are no or very few functioning inner hair cells (IHCs) and/or neurons (Moore and Glasberg, 1997; Moore, 2001, 2004). Halpin (2002) proposed that a TCA might be a result of a restricted region where there are surviving IHCs and/or neurons, with extensive DRs on either side. Indeed, he argued that a TCA could result from a single surviving IHC. The audiogram would simply reflect the tuning of that IHC.

There are several ways of diagnosing DRs. A simple clinical test was described by Moore *et al.* (2000). The test is based on a masking noise, called "threshold-equalizing noise" (TEN), which is spectrally shaped so that, for normally hearing subjects, the masked threshold of a pure tone is almost independent of frequency. The test involves measuring the detection threshold for a sinusoidal signal in the TEN. When there is no DR at the test frequency, the detection threshold is usually within 10 dB of the noise level per ERB_N, where ERB_N refers to the equivalent rectangular bandwidth of the auditory filter, as determined using listeners with normal hearing at moderate sound levels (Glasberg and Moore, 1990). A threshold, which is at least 10 dB higher than "normal," and also at least 10 dB above the absolute threshold, is taken as indicating a DR at the test frequency. In the original version of the TEN test (Moore *et al.*, 2000), all levels were specified and calibrated in dB sound pressure level (SPL). In a more recent version of the test, levels are calibrated in dB hearing level (HL), and the test is referred to as the TEN(HL) test (Moore *et al.*, 2004). However, the same criteria apply for both tests.

An alternative, but more precise way of diagnosing DRs, applicable to co-operative listeners, is to measure PTCs. These are obtained using a fixed signal presented at a low sensation level (SL). The level of a narrowband masker required to mask the signal is determined as a function of the masker center frequency. When the tip of the PTC is shifted away from the signal frequency, this is taken as indicating a DR at the signal frequency. The frequency at the tip of the PTC is assumed to indicate the frequency corresponding to the boundary of the DR (Moore and Alcántara, 2001).

Halpin (2002) described a TCA for a 52-year old man who developed a sudden sensorineural hearing loss after a respiratory infection. His hearing was relatively good over the range 3000–4000 Hz, but worsened outside this range. Halpin (2002) did not directly confirm the presence of DRs for frequencies below 3000 Hz or above 4000 Hz. However, he did measure the intelligibility of monosyllables at two sound levels. For the lower sound level (63 dB HL), it was argued that only frequencies over a narrow range centered around 3000–4000 Hz would have been audible. For the

[a)]Present address: Acoustics Research Center, Department of Electronics and Telecommunications, Norwegian University of Science and Technology (NTNU), O. S. Bragstads plass 2B, NO-7491 Trondheim, Norway.
[b)]Author to whom correspondence should be addressed. Electronic mail: bcjm@cam.ac.uk

higher level (98 dB HL), a much wider frequency range should have been audible. Scores were low for both levels and did not differ significantly for the low level (20%) and the high level (24%). Halpin (2002) concluded that the results were "in good agreement with the hypothesis that only frequencies in the octave surrounding 3 kHz are stimulated." However, Halpin (2002) also described the case of a man showing a TCA who achieved high speech intelligibility (92%). Halpin (2002) argued that "it is not possible that the central region alone could be responsible for the observed score, and so more of the cochlea must also be contributing." This illustrates that it is not safe to base diagnosis of the underlying problem on the audiogram alone.

In this paper, we describe a case of an individual presenting with a TCA. The TEN(HL) test was used to provide an initial diagnosis of DRs. PTCs were used to confirm the presence of a localized surviving region in the cochlea, with DRs on either side. The intelligibility of nonsense syllables was also measured as a function of cutoff frequency for lowpass and highpass filtered speech. The results show that speech information could only be extracted from a limited frequency region, and that the presence of amplified low frequencies led to poor speech discrimination.

II. METHOD

A. Audiometric evaluation

The subject was a 54-year old male who attended an audiology clinic for the purpose of being fitted with a (replacement) hearing aid. The subject reported no history of tinnitus, ear discharge, or ear pain. The subject reported that hearing loss had first been experienced 15 years earlier. The cause of the hearing loss was unknown. Audiometric thresholds were measured using the modified Hughson–Westlake procedure described by Carhart and Jerger (1959) with a Madsen OB922 audiometer equipped with Telephonics TDH39 headphones and Radio ear B-71 bone vibrator. A diagnosis of sensorineural hearing loss was made, based on an air-bone gap in audiometric threshold less than or equal to 10 dB for frequencies from 250 to 4000 Hz, and normal results for immittance measurements (Grason–Stadler Tympstar). The audiometric thresholds for the left ear were 70, 70, 75, 70, 70, 65, 60, 20, 45, and 70 dB HL for frequencies of 250, 500, 750, 1000, 1500, 2000, 3000, 4000, 6000, and 8000 Hz, respectively. These values are plotted as the dashed lines in Fig. 1 (see below); note that in this figure, HL is plotted increasing upwards, in the opposite direction to a conventional audiogram. The right ear had a profound hearing loss.

B. Application of the TEN(HL) test

The TEN(HL)-test CD was replayed via a Philips 729K CD player connected to a Madsen OB922 audiometer equipped with TDH39 earphones. The level of the signal and the TEN(HL) were controlled using the attenuators in the audiometer. The TEN(HL) test includes test frequencies from 500 to 4000 Hz. The TEN(HL) level was 70 dB HL/ERB_N. The signal level was varied in 2-dB steps to determine the masked thresholds, as recommended by Moore *et al.* (2004).

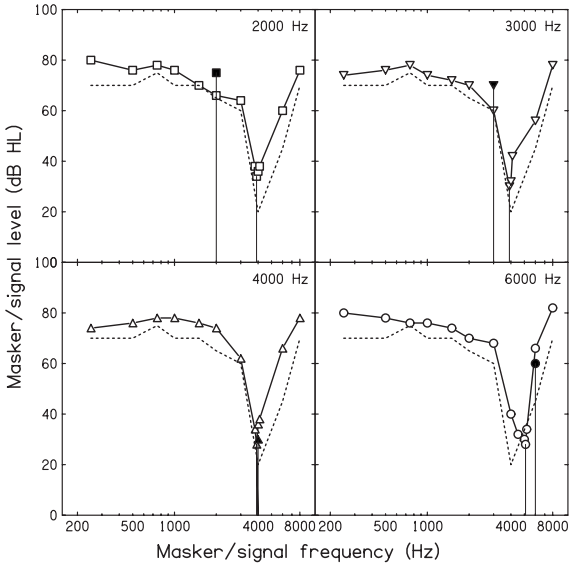


FIG. 1. PTCs for signal frequencies of 2000, 3000, 4000, and 6000 Hz, as indicated in the key for each panel. The signal and masker level are plotted in dB HL. Filled symbols show the signal level and frequency, and open symbols show the masker level required for threshold. The dashed line in each panel shows the audiometric threshold as a function of frequency. The thin vertical lines indicate the signal frequency and the tip frequency of the PTC.

A "no response (NR)" was recorded when the subject did not indicate hearing the signal at the maximum output level of the audiometer, which was 86 dB HL for the signals derived from the TEN(HL) CD. The presence or absence of a DR at a specific frequency was based on the criteria suggested by Moore *et al.* (2004) and described in the Introduction.

C. Measurement of PTCs

PTCs were measured using a Maico 53 dual-channel audiometer equipped with TDH39 headphones. The audiometer was set to dual-frequency mode. The signal tone was generated in one channel. It was presented at 10 dB SL, and was pulsed on and off in a regular sequence (0.25 s on and 0.25 s off). A narrowband noise masker was selected in the other channel. The noise conformed to the specifications given in ANSI-S3.6 (ANSI, 2004), and had a bandwidth between 1/3 and 1/2 octaves. This relatively large bandwidth helps to reduce the influence of beats on the PTCs (Kluk and Moore, 2004, 2005). The outputs of the two channels were mixed and presented to the left ear. The subject was asked to respond when he could hear the tone in the presence of the noise. The minimum noise level (in dB HL) required to mask the tone was determined by manual adjustment of the noise level. This was repeated for several masker frequencies placed below and above the signal frequency. Signal frequencies of 2000, 3000, 4000, and 6000 Hz were used. At least 12 masker frequencies were used for each signal frequency, and the masker frequencies were closely spaced around the tip of the PTC, in order to define the tip frequency precisely.

D. Measurement of the intelligibility of filtered speech

Speech intelligibility was measured using vowel-consonant-vowel (VCV) nonsense syllables, spoken by a

man. There were ten lists of 63 VCV syllables, each recorded with a different order. For full details, see Vinay and Moore (2007b). Speech stimuli with a nominal level of 65 dB SPL were amplified according to the Cambridge formula (Moore and Glasberg, 1998), which is intended to provide approximately equal loudness/ERB_N over the speech frequency range from 500 to 4000 Hz. Following amplification, stimuli were lowpass or highpass filtered, with cutoff frequencies of 610, 862, 1220, 1725, 2440, 3450, 4880, and 6900 Hz; these are spaced at 0.5-octave intervals.

A Tucker–Davis technologies (TDT) System 3 was used to digitize the stimuli and to implement the Cambridge formula filter (Camfilter) and highpass or lowpass filter. Stimuli were presented to one ear only using Sennheiser HD580 headphones. The Camfilter took into account the frequency response of the HD580 headphones, as measured on a KEMAR manikin (Burkhard and Sachs, 1975). For full details, see Vinay and Moore (2007b).

The subject was initially given examples of the stimuli, including stimuli with different lowpass and highpass cutoff frequencies. Following each presentation, the subject was instructed to write down the consonant heard and was asked to make a response, even if he thought he was guessing. Feedback as to the correct response was given. Identification of the vowels was not required. Once the subject was familiar with the stimuli and the task, testing proper commenced. For testing proper, no feedback was given, to minimize further learning effects. Testing was conducted in four sessions, each lasting 1 h. In a given session, each cutoff frequency was used once with a specific type of filtering (highpass or lowpass). The order of cutoff frequencies was randomized for the first session with a given type of filtering, and the order was reversed for the second session, in order to counterbalance effects of practice and fatigue. Scores presented are the means (and standard deviations) for each cutoff frequency across the two sessions for each type of filtering.

III. RESULTS

A. TEN(HL) test and PTCs

For the test frequencies of 500, 750, and 1000 Hz, the test tone was not detected at the maximum output level, which was 86 dB HL. Thus, the tone threshold was above 86 dB HL, and the criteria for diagnosing a DR were met. For the test frequencies of 1500, 2000, and 3000 Hz, the tone thresholds were 86, 84, and 82 dB HL, again meeting the criteria for diagnosing a DR. For the test frequency of 4000 Hz, the tone threshold was 72 dB HL, which was only 2 dB above the TEN level/ERB_N, and did not meet the criteria for diagnosing a DR. Thus, the results of the TEN(HL) test suggested the presence of a DR at all test frequencies, except 4000 Hz.

The PTCs are presented in Fig. 1. Each panel shows results for one signal frequency. The level and frequency of the signal are indicated in each panel by the filled symbol. The audiometric thresholds are shown by the dashed line in each panel. For the 4000-Hz signal (bottom left), the tip of the PTC fell very close to the signal frequency, consistent with a functioning region of the cochlea tuned to 4000 Hz.

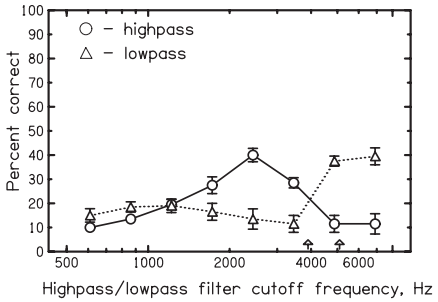


FIG. 2. Results of the VCV test. The percent correct score is plotted as a function of highpass (circles) or lowpass (triangles) filter cutoff frequency. Error bars indicate ±1 standard deviation across repeated runs. The up-pointing arrows indicate the estimated boundaries of the functioning region in the cochlea.

For the 2000- and 3000-Hz signals, the tips of the PTCs were shifted upwards to 3900 Hz, suggesting the presence of DRs at 2000 and 3000 Hz. For the 6000-Hz signal, the tip of the PTC was shifted downwards, to 5100 Hz, suggesting the presence of a DR at 6000 Hz. The PTCs were generally very similar in shape to the audiometric thresholds, consistent with the idea of a small functioning region centered around 4000–5000 Hz. The tip of the PTC for the 6000-Hz signal fell slightly below the dashed curve around 5000 Hz, but this probably happened because the audiometric threshold was not measured at 5000 Hz, so the dashed curve is plotted as a straight line between 4000 and 6000 Hz. The actual audiometric threshold was probably a little lower than indicated by the dashed curve at 5000 Hz. Overall, the results for the PTCs are consistent with those of the TEN(HL) test. The tip frequencies of the PTCs suggest that the surviving region extends from 3900 to 5100 Hz.

B. Identification of VCVs

The scores for identification of lowpass and highpass filtered VCVs are shown in Fig. 2. For the lowpass filtered speech, scores were very low for cutoff frequencies up to 3450 Hz, indicating that very little speech information could be extracted from frequency components falling in the DR extending from low frequencies up to about 3900 Hz. Performance improved when the lowpass cutoff frequency was increased to 4880 Hz, and then remained almost the same when the cutoff frequency was increased further to 6900 Hz. For the highpass filtered speech, scores were very low for the lowest cutoff frequency, but improved with increasing cutoff frequency up to 2440 Hz. Thus, removing low-frequency speech components led to better performance. As the highpass filter cutoff frequency was increased above 2440 Hz, performance worsened again, becoming very low for the cutoff frequencies of 4880 and 6900 Hz.

The results for the highpass filtered speech resemble those for subjects with DRs restricted to low frequencies (Vinay and Moore, 2007b; Vinay *et al.*, 2008). Vinay and Moore (2007b) found that performance improved with increasing cutoff frequency until the cutoff frequency was about 0.57fe, where fe denotes the edge frequency of the DR. For the subject tested here, if the value of fe is taken as corresponding to the lower boundary of the functioning region, i.e., 3900 Hz, then the optimum highpass cutoff fre-

quency predicted by this rule is about 2200 Hz, which is close to the obtained optimum value of 2440 Hz.

There is one apparent anomaly in the results. The data for highpass filtering suggest that useful speech information could be extracted for frequencies between 2440 and 3450 Hz. However, the data for lowpass filtering suggest that little useful information could be extracted for frequencies below 3450 Hz. A possible explanation is that, for the condition with the lowpass filter cutoff of 3450 Hz, amplified frequency components below 2440 Hz had a deleterious effect, preventing effective use of information between 2440 and 3450 Hz.

IV. DISCUSSION

The results presented here confirm the proposal of Halpin (2002) that a TCA can be associated with a localized surviving region in the cochlea, with DRs on either side. The subject tested here had an audiometric threshold of only 20 dB HL at 4000 Hz, with thresholds of 45 dB HL or higher for all other audiometric frequencies. The TEN(HL) test results suggested the presence of a DR for all test frequencies from 500 to 3000 Hz, but no DR at 4000 Hz. The PTC results suggested that the functioning region extended from 3900 to 5100 Hz. Note that the audiometric thresholds at low frequencies were only about 70 dB HL, despite there being a DR at these frequencies. The audiometric thresholds at 3000 and 6000 Hz were only 60 and 45 dB HL, respectively, even though the PTC results suggested the presence of DRs at these frequencies. These results confirm that a DR can be present even for audiometric frequencies where the hearing loss would normally be classified as mild or moderate. They also confirm that the presence or absence of a DR cannot be reliably determined from the audiometric threshold alone (Aazh and Moore, 2007; Vinay and Moore, 2007a).

As pointed out by Halpin (2002), a TCA may not always be associated with a localized functioning region in the cochlea, with DRs on either side. Thus, when a patient with a TCA appears in the clinic, it would seem prudent to apply a test for the presence of DRs, such as the TEN(HL) test or PTCs. In practice, the "fast" method of measuring PTCs with a sweeping masker may be useful for this purpose (Sęk *et al.*, 2005; Kluk and Moore, 2006). A person who does have a localized surviving region in the cochlea will probably have a poor ability to understand speech (Halpin, 2002; Huss and Moore, 2005). However, speech reception may be optimized by appropriate filtering. For example, if a DR exists for frequencies up to X Hz and frequencies above Y Hz, with a functioning region between X and Y Hz, our results, combined with earlier results on the effects of lowpass and highpass filtering of speech for subjects with DRs (Vickers *et al.*, 2001; Baer *et al.*, 2002; Vinay and Moore, 2007b; Vinay *et al.*, 2008), suggest that bandpass filtering between about $0.57X$ and $1.5Y$ may lead to optimum intelligibility. However, further work is needed to verify the applicability of this rule. It is possible that the deleterious effect of amplification of frequencies below $0.57X$ results from upward spread of masking or from the resulting abnormal spectral shape of the speech. If so, then it might be appropriate not to remove frequency components below $0.57X$ entirely, but rather to present them unamplified, or amplified to a lesser extent than prescribed by the Cambridge formula.

ACKNOWLEDGMENTS

The work of B.C.J.M. was supported by the Medical Research Council (UK). We thank Tom Baer for designing the digital filters used in this study and the Director, All India Institute of Speech and Hearing, for providing the facilities to carry out this work. We also thank Tom Baer and two anonymous reviewers for helpful comments on an earlier version of this paper.

Aazh, H., and Moore, B. C. J. (**2007**). "Dead regions in the cochlea at 4 kHz in elderly adults: Relation to absolute threshold, steepness of audiogram, and pure tone average," J. Am. Acad. Audiol. **18**, 96–107.

ANSI (**2004**). *ANSI S3.6–2004 Specification for Audiometers* (American National Standards Institute, New York).

Baer, T., Moore, B. C. J., and Kluk, K. (**2002**). "Effects of lowpass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies," J. Acoust. Soc. Am. **112**, 1133–1144.

Burkhard, M. D., and Sachs, R. M. (**1975**). "Anthropometric manikin for acoustic research," J. Acoust. Soc. Am. **58**, 214–222.

Carhart, R., and Jerger, J. F. (**1959**). "Preferred method for clinical determination of pure-tone thresholds," J. Speech Hear Disord. **24**, 330–345.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103–138.

Halpin, C. (**2002**). "The tuning curve in clinical audiology," J. Am. Audiol Soc. **11**, 56–64.

Huss, M., and Moore, B. C. J. (**2005**). "Dead regions and pitch perception," J. Acoust. Soc. Am. **117**, 3841–3852.

Kluk, K., and Moore, B. C. J. (**2004**). "Factors affecting psychophysical tuning curves for normally hearing subjects," Hear. Res. **194**, 118–134.

Kluk, K., and Moore, B. C. J. (**2005**). "Factors affecting psychophysical tuning curves for hearing-impaired subjects," Hear. Res. **200**, 115–131.

Kluk, K., and Moore, B. C. J. (**2006**). "Detecting dead regions using psychophysical tuning curves: A comparison of simultaneous and forward masking," Int. J. Audiol. **45**, 463–476.

Moore, B. C. J. (**2001**). "Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids," Trends Amplif. **5**, 1–34.

Moore, B. C. J. (**2004**). "Dead regions in the cochlea: Conceptual foundations, diagnosis and clinical applications," Ear Hear. **25**, 98–116.

Moore, B. C. J., and Alcántara, J. I. (**2001**). "The use of psychophysical tuning curves to explore dead regions in the cochlea," Ear Hear. **22**, 268–278.

Moore, B. C. J., and Glasberg, B. R. (**1997**). "A model of loudness perception applied to cochlear hearing loss," Aud. Neurosci. **3**, 289–311.

Moore, B. C. J., and Glasberg, B. R. (**1998**). "Use of a loudness model for hearing aid fitting. I. Linear hearing aids," Br. J. Audiol. **32**, 317–335.

Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (**2004**). "New version of the TEN test with calibrations in dB HL," Ear Hear. **25**, 478–487.

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcántara, J. I. (**2000**). "A test for the diagnosis of dead regions in the cochlea," Br. J. Audiol. **34**, 205–224.

Sęk, A., Alcántara, J. I., Moore, B. C. J., Kluk, K., and Wicher, A. (**2005**). "Development of a fast method for determining psychophysical tuning curves," Int. J. Audiol. **44**, 408–420.

Vickers, D. A., Moore, B. C. J., and Baer, T. (**2001**). "Effects of lowpass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies," J. Acoust. Soc. Am. **110**, 1164–1175.

Vinay, and Moore, B. C. J. (**2007a**). "Prevalence of dead regions in subjects with sensorineural hearing loss," Ear Hear. **28**, 231–241.

Vinay, and Moore, B. C. J. (**2007b**). "Speech recognition as a function of highpass filter cutoff frequency for subjects with and without low-frequency cochlear dead regions," J. Acoust. Soc. Am. **122**, 542–553.

Vinay, Moore, B. C. J., and Baer, T. (**2008**). "Speech recognition in noise as a function of highpass-filter cutoff frequency for people with and without low-frequency cochlear dead regions," J. Acoust. Soc. Am. **123**, 606–609.

The scattering of sound by a long cylinder above an impedance boundary

Wai Keung Lui
Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Kai Ming Li[a)]
School of Mechanical Engineering, Ray W. Herrick Laboratories, Purdue University, 140 South Martin Jiskche Drive, West Lafayette, Indiana 47907-2011

(Received 26 August 2009; revised 18 November 2009; accepted 19 November 2009)

The classical problem of sound scattering by an acoustically hard cylinder due to a point monopole and a line airborne source is extended in the present study. The solution to the homogeneous Helmholtz equation is expressed in a cylindrical coordinate system and represented by an expansion of Fourier integrals. Incorporating the image source method and the Bessel function addition theorem, the analytical solution is derived for the prediction of multiple scattering of sound by a hard cylinder placed above a ground surface of finite impedance. The total sound field can be expressed as a sum of four components: the incident field, the reflected wave, and the scattered fields from the cylinder and its image. The total far-field scattered potential was evaluated asymptotically by the method of stationary phase. Experimental measurements by using a point source were conducted in an anechoic chamber to validate the theoretical formulations. The numerical predictions of using a point source model give good agreements with all the experimental data but there are obvious discrepancies in the spectral magnitudes between the calculation and experimental results when a line source model is used to simulate the scattering problem due to a point source excitation. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3273891]

PACS number(s): 43.20.Bi, 43.20.El, 43.20.Fn, 43.28.Js [RMW] Pages: 664–674

NOMENCLATURE

$k=\omega/c$ = Wave number in the space
f,ω = Frequency and angular frequency of the source
c = Sound speed in the space
a = Radius of the cylinder
d = Distance between the axes of the cylinder and its image
$\phi^d,\phi^s,\phi_r^d,\phi_r^s$ = Sound field contributed from direct waves, scattered waves, reflected waves emanating from the image source, and the image cylinder
ϕ^T = Total sound field at the field point
$R,\bar{R}$ = Distance from the field point to the source and image source
$r,\theta,z;\ r_o,\theta_o,z_o$ = Cylindrical coordinates of the field point and the source
$\bar{r},\bar{\theta},z$ = Field point in the imaginary coordinate system
$\bar{r}_o$ = The shortest distance measured from the centerline of the cylinder to the image source
θ_1 = Polar angle subtended from the image source to the x-axis
J_n = The nth order Bessel function of the first kind
$H_n\equiv H_n^{(1)}$ = The nth order Hankel function of the first kind
γ,α = Radial and axial wave numbers in free space
A_n,B_n = Spectral amplitudes of the scattered waves from the real and the image cylinder
ϕ_{Tf}^s = Total far-field scattered potential
$\hat{R},\tilde{R}$ = Separations from the field point to the centers of the long cylinder and its image
Q_1,Q_2 = Spherical wave reflection coefficients for the point source and scattering wave reflection on the impedance ground
V_1,V_2 = Plane wave reflection coefficients
α_1,α_2 = The angles of incidence of the reflected waves
w_1,w_2 = Numerical distances
β_g = Specific normalized admittance of the reflecting ground
σ_e = Effective flow resistivity of the impedance ground
α_e = Effective rate of change in porosity with depth of the impedance ground

I. INTRODUCTION

The problem of sound scattering by an obstacle has been a subject of investigations for the past few decades in various scientific and engineering research areas. In particular, the scattering of sound waves by an infinite cylinder has attracted considerable interests because of its broad applica-

a)Author to whom correspondence should be addressed. Electronic mail: mmkmli@purdue.edu

tions in underwater acoustics, geophysics, atmospheric acoustics, medical imaging, and nondestructive testing. The scattering of an incident plane wave by a long cylinder, which is often treated as a canonical solution, has been studied extensively both theoretically and experimentally.[1–4] Simple analytical and numerical solutions have been developed. These numerical solutions sometimes serve as benchmark results for other more sophisticated numerical schemes solving a more difficult problem with the canonical problem as a special case.[5,6] However, the assumption of an incident plane wave is not appropriate in many situations because the source is close to the obstacle. Hence, the assumption of plane wave propagation from the source to the scatterer cannot be established.

In addition to the study on incident plane waves, there were many other studies that focused on the scattering of a spherical sound wave by a long cylinder. Their solutions were expressed either in an approximate form which holds in a very restricted region or in an integral form where further numerical analyses were needed.[7–9] For instance, Skelton and James[10] represented the sound pressure by a Fourier integral. The far-field sound pressure is then approximated by evaluating the Fourier integral asymptotically.

For the problem of sound scattered by a long cylinder above a plane boundary, it is treated as the plane wave incident on a long cylinder near an elastic plane.[11,12] Ye and Chen[13] considered a line source placed in the vicinity of a long cylinder near a pressure release surface. Shenderov[14] solved the problem for a line source placed near an elastic cylinder and above an elastic halfspace.

One of the prime objectives of the current work is to develop a theoretical model to predict the scattering of sound by a long cylinder placed above an impedance ground. To the best of our knowledge, this problem has not been tackled in the past despite the fact that the consideration of a point source has a broader application. The exact analytical solution for this relatively simple case is needed because it will provide a benchmark solution for comparison with other more complicated and theoretically intractable numerical schemes. The analytical model may also be useful in other related fields such as acoustical oceanography and scattering theory of electromagnetic waves.

This paper is arranged as follows. The theoretical development for the scattered sound field for a hard cylinder irradiated by a point source is given in Sec. II. In particular, the application of Fourier integral and Fourier series for representing the variations in axial and radial pressures is discussed. The physical parameters used in the mathematical model are elucidated.

In Sec. III, the closed-form solutions for the sound pressure in the far-field above an impedance ground are presented. Section IV examines the analogous analytical solutions for a long hard cylinder irradiated by a line source placed in the vicinity of a hard ground or an impedance ground. Section V presents indoor experimental results where they will be compared with different numerical prediction schemes using either a point source or a line source. The conclusion is given in Sec. VI.

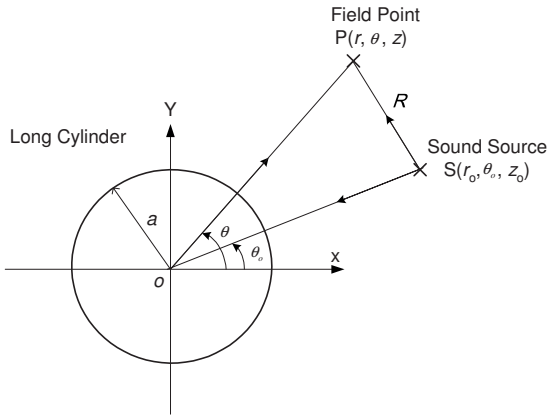


FIG. 1. Illustration of the geometrical configuration of the scattering problem by a long cylinder irradiated by a point source or line source.

II. SOUND SCATTERING BY A LONG CYLINDER IN FREE FIELD

Suppose a long cylinder of radius a is placed initially in a free space. A point monopole source is placed exterior to the cylinder at a point S with cylindrical coordinates of (r_o, θ_o, z_o) and the field point P is situated at (r, θ, z), as shown in Fig. 1. The total sound field is governed by the Helmholtz equation expressed in terms of a scalar velocity potential ϕ as

$$\nabla^2\phi + k^2\phi = -4\pi\delta(\mathbf{x} - \mathbf{x}_o), \tag{1}$$

where $\mathbf{x}$ is the field point, $\mathbf{x}_o$ is the source point, $k=\omega/c$ is the wave number, ω is the angular frequency of the source, and c is the sound speed in the space. In Eq. (1), the time dependent parameter $e^{-i\omega t}$ is understood and suppressed.

The total sound field due to the presence of a long cylinder in an otherwise free space can be expressed as

$$\phi = \phi^d + \phi^s, \tag{2}$$

where ϕ^d and ϕ^s are the free-field and the scattered field velocity potentials due to the source of strength 4π. Using the standard Fourier transform, the free-field velocity potential can be expanded in terms of its Fourier integral representation as follows:[10]

$$\phi^d = \frac{e^{ikR}}{R}$$

$$= \frac{i}{2}\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_o)}\int_{-\infty}^{+\infty} J_n(\gamma r)H_n(\gamma r_o)e^{i\alpha(z-z_o)}d\alpha \quad \text{for } r_o \geqslant r, \tag{3}$$

where R is the distance from the source to receiver, J_n is the nth order Bessel function of the first kind, H_n is the nth order Hankel function of the first kind, and $\gamma=+\sqrt{k^2-\alpha^2}$ and α are the respective radial and axial wave numbers in free space. We remark that Eq. (3) can be treated as a special case for the addition theorem of the Bessel function. The absolute signs imposed on the order n of the Bessel functions are found unnecessary [cf. Eq. (4.3.11) in Ref. 10] and removed in Eq. (3) for brevity.

The scattered velocity potential is given in a convenient form as[10]

$$\phi^s = \sum_{n=-\infty}^{\infty} e^{in\theta} \int_{-\infty}^{+\infty} A_n H_n(\gamma r) e^{i\alpha z} d\alpha, \tag{4}$$

where A_n are the unknown constant coefficients of the scattered wave which can be determined by imposing appropriate boundary conditions on the surface of the cylinder. The form of Eq. (4) is modified (the absolute sign imposed on the order n of the Hankel function) and the polar angle θ_o of the incident wave is removed [cf. Eq. (4.3.12) in Ref. 10] to facilitate the subsequent formulations of sound scattering above a ground surface. In Eq. (3), the roles of the source and receiver may be interchanged if $r \geqslant r_o$ because the condition of the reciprocity holds in this situation. For a hard cylinder, the total particle displacement normal to the surface vanishes, which leads to a solution for the scattering constant A_n in Eq. (4) as follows:

$$A_n = -\frac{i}{2} \frac{J_n'(\gamma a)}{H_n'(\gamma a)} \cdot H_n(\gamma r_o) e^{-i\alpha z_o} e^{-in\theta_o}, \tag{5}$$

where the primes represent the derivative of the respective Bessel functions with respect to their arguments.

III. SOUND SCATTERING ABOVE A GROUND SURFACE

Suppose the long cylinder, which has been aligned with its longitudinal axis placed parallel to the ground surface, is positioned with its center placed at a height of $d/2$ above the flat ground. Again, a cylindrical coordinate system is used in the analysis where the point source is positioned at a finite distance above the ground surface. The geometrical configuration of the problem is shown in Fig. 2. The total sound field above the ground is contributed to by the direct radiation from the source, the diffracted sound from the cylinder, and the reflected sound from the ground surface. There is an additional contribution due to the interaction between the cylinder and the ground surface which will be included in the subsequent analysis. We shall give the solution for the sound fields above a hard ground in Sec. III A. In Sec. III B, the problem will be generalized to the situation where the cylinder is placed above an impedance ground.

A. Scattering of sound by a cylinder above a hard ground

The sound field due to a long cylinder insonified by a point source in a free space can be computed by using Eq. (2) with the free-field potential ϕ^d computed by Eq. (3) and the scattered field potential ϕ^s computed by Eq. (4). The scattering constant A_n in Eq. (4) can be calculated by Eq. (5) for a cylinder that has a hard surface.

The sound field scattered by a cylinder above a hard ground can be found readily by applying the image source method in accordance with the theory of geometrical acoustics. It is straightforward to see that the total sound field can be expressed as a sum of four components: direct waves (ϕ^d), scattered waves (ϕ^s) from the cylinder, reflected waves

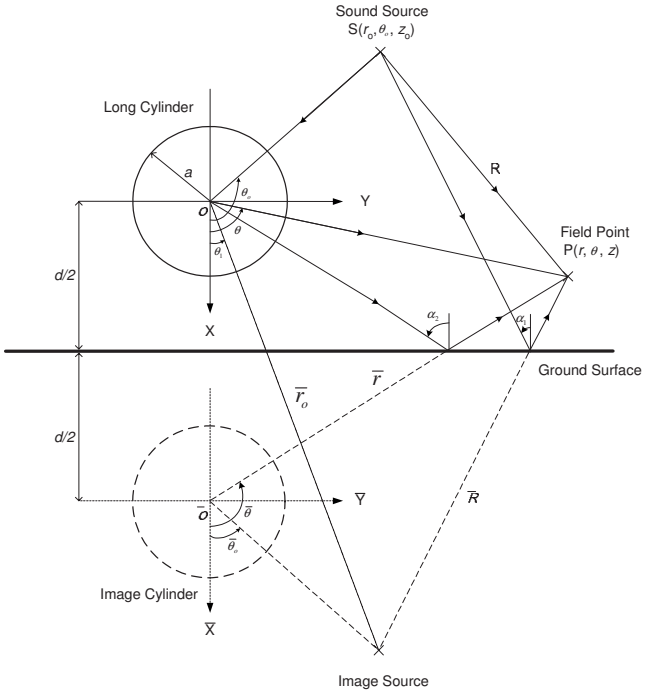


FIG. 2. Illustration of the geometrical configuration of the scattering problem by a long cylinder close to a flat ground excited by a point source or line source.

emanating from the image source (ϕ_r^d), and the image cylinder (ϕ_r^s). Figure 2 shows a schematic diagram of the contribution from these four components. In light of Eq. (3), the sound field contributed by an image source above a hard ground ϕ_r^d can be expressed in terms of a Fourier series as

$$\phi_r^d = \frac{e^{ik\bar{R}}}{\bar{R}} = \frac{i}{2} \sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_1)} \int_{-\infty}^{+\infty} J_n(\gamma r) H_n(\gamma \bar{r}_o) e^{i\alpha(z-z_o)} d\alpha \quad \text{for } \bar{r}_o \geqslant r, \tag{6}$$

where $\bar{R}$ is the distance from the image source to receiver and $\bar{r}_o$ is the shortest distance measured from the centerline of the long cylinder to the image source. Using the cylindrical coordinate system centered on the image cylinder, the scattered sound field contributed by the image cylinder, ϕ_r^s, can be written as

$$\phi_r^s = \sum_{n=-\infty}^{\infty} e^{in\bar{\theta}} \int_{-\infty}^{\infty} B_n H_n(\gamma \bar{r}) e^{i\alpha z} d\alpha, \tag{7}$$

where $\bar{r}$ is the shortest distance measured from the centerline of the image cylinder to the receiver and $\bar{\theta}$ is the corresponding polar angle subtended from the receiver to the imaginary x-axis ($\bar{x}$), as shown in Fig. 2. When we impose the boundary condition at the flat ground surface for the total scattered sound field, $\phi_T^s = \phi^s + \phi_r^s$, we have $\partial \phi_T^s / \partial x = 0$ at $x = -d/2$ for a hard ground. Since

$$\partial \phi_T^s / \partial x = (\partial \phi_T^s / \partial r)(\partial r / \partial x) + (\partial \phi_T^s / \partial \theta)(\partial \theta / \partial x), \tag{8}$$

we can show that A_n and B_n are related by

$$B_n = A_{-n}. \tag{9}$$

To obtain the above relationship between A_n and B_n, we have used the identities of $\bar{r}=r$, $\bar{\theta}=\pi-\theta$, $e^{in\pi}=(-1)^n$, and $H_{-n}(z)=(-1)^n H_n(z)$ at the ground surface where $x=-d/2$.

It is important to note that the scattered sound field due to the image cylinder is expressed in the transformed coordinate system centered on the image cylinder. However, it is more convenient to use the coordinate system centered on the long cylinder placed above the ground surface. The scattered sound field (ϕ_r^s) due to the image cylinder can be expressed in terms of the original coordinate system by employing the translational addition theorem for the cylindrical wave function.[15] For the case of a vertical translation, the Hankel function can be expressed as

$$H_n(\gamma\bar{r})e^{in\bar{\theta}} = \sum_{m=-\infty}^{\infty} H_{m-n}(\gamma d)J_m(\gamma r)e^{im\theta} \quad \text{for } r<d. \tag{10}$$

The total sound field of the acoustic excitation due to a point source on a long cylinder above a hard ground can now be expressed as

$$\phi^T(r,\theta,z) = (\phi^d + \phi_r^d) + (\phi^s + \phi_r^s). \tag{11a}$$

Substituting Eqs. (3), (6), (4), and (7) into Eq. (11a) for the respective terms, we obtain

$$\begin{aligned}\phi^T(r,\theta,z) = {} & \frac{i}{2}\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_o)}\int_{-\infty}^{+\infty} J_n(\gamma r)H_n(\gamma r_o)e^{i\alpha(z-z_o)}d\alpha \\ & + \frac{i}{2}\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_1)}\int_{-\infty}^{+\infty} J_n(\gamma r)H_n(\gamma \bar{r}_o)e^{i\alpha(z-z_o)}d\alpha \\ & + \sum_{n=-\infty}^{\infty} e^{in\theta}\int_{-\infty}^{+\infty} A_n H_n(\gamma r)e^{i\alpha z}d\alpha \\ & + \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} e^{in\theta}\int_{-\infty}^{+\infty} A_{-m}H_{n-m}(\gamma d)J_n(\gamma r)e^{i\alpha z}d\alpha.\end{aligned} \tag{11b}$$

The amplitudes of the constant coefficients of the scattered waves A_n can be determined by applying the Neumann boundary condition on the surface if the cylinder is hard. A set of infinite and coupled linear complex equations can be formed by solving the constant coefficients of the scattered waves. The set of complex equations can be written in a convenient form as follows:

$$\begin{aligned}A_n + \sum_{m=-\infty}^{\infty} A_{-m}\frac{J_n'(\gamma a)}{H_n'(\gamma a)}H_{n-m}(\gamma d) \\ = -\frac{i}{2}\frac{J_n'(\gamma a)}{H_n'(\gamma a)}e^{-i\alpha z_o}[H_n(\gamma r_o)e^{-in\theta_o} + e^{-in\theta_1}H_n(\gamma \bar{r}_o)].\end{aligned} \tag{12}$$

We can write the above equations in a matrix form as

$$\mathbf{MA} = \mathbf{D}, \tag{13}$$

where the elements of the square matrix $\mathbf{M}=\{M_{nm}\}$ are given as follows:

$$M_{nm} = \begin{cases} \dfrac{J_n'(\gamma a)}{H_n'(\gamma a)}H_{n+m}(\gamma d) & \text{if } n \neq m \\ 1 + \dfrac{J_n'(\gamma a)}{H_n'(\gamma a)}H_{2n}(\gamma d) & \text{if } n = m \end{cases}. \tag{14}$$

the elements for the column vector $\mathbf{A}$ are the respective scattering constants A_n, and the elements for the column vector $\mathbf{D}$ are determined by

$$D_n = -\frac{i}{2}\frac{J_n'(\gamma\alpha)}{H_n'(\gamma a)}e^{-i\alpha z_o}[H_n(\gamma r_o)e^{-in\theta_o} + e^{-in\theta_1}H_n(\gamma \bar{r}_o)]. \tag{15}$$

The required accuracy of the numerical solution for A_n can be achieved by truncating the system of complex equations in Eq. (11a) and (11b) to an order of N, where N is the maximum number of terms required in the summation series. The order of N can be determined by using a stepwise approach until the numerical results are converged. Hence, in Eq. (13), $\mathbf{M}$ is a $(2N+1)\times(2N+1)$ square matrix where its elements are complex variables. The complex column vectors, $\mathbf{A}$ and $\mathbf{D}$, have dimensions of $(2N+1)$ for each value of n varying from $-N$ to N.

We are particularly interested in obtaining the far-field sound pressure in the present study. For a point source excitation, the sound field due to a point source or its image source in the far-field is simply computed by using the Green function such as $\phi=e^{ik|\mathbf{x}-\mathbf{x}_o|}/|\mathbf{x}-\mathbf{x}_o|$, where $|\mathbf{x}-\mathbf{x}_o|$ is the distance between the point source or its image source to the field point. The integrals of the scattered potential in Eq. (11a) and (11b) can be evaluated analytically by the method of stationary phase if the receiver is located at far-field. The details of the asymptotic method can be found in many textbooks (see, for example, Ref. 10) and they will be omitted here. The total far-field scattered potential can be approximated by

$$\phi_{Tf}^s = -2i\sum_{n=-\infty}^{\infty}(-i)^n\left[A_n e^{in\theta}\frac{e^{ik\hat{R}}}{\hat{R}} + B_n e^{in\bar{\theta}}\frac{e^{ik\tilde{R}}}{\tilde{R}}\right], \tag{16}$$

where $\hat{R}$ and $\tilde{R}$ denote the separations from the field point to the centers (origins) of the long cylinder and its image, respectively. There two centers are marked as O and $\bar{\text{O}}$ in Fig. 2.

For an absorbing cylinder, the sound field can also be computed by imposing appropriate boundary conditions on the cylindrical surface which is either a locally reacting surface or an extended reaction one. In an earlier publication,[16] we have developed an analytical method to solve a problem involving an absorbing sphere. The same method can be applied in the present study for the absorbing cylinder. However, the details of the analysis are not provided here for brevity.

In Eqs. (13)–(15), the magnitudes of the Hankel functions decrease with increasing γd and $\gamma \bar{r}_o$ as evidenced by the following asymptotic approximation;[17]

$$H_\nu(z) = \sqrt{2/(\pi z)}\,\exp[i(z - \nu\pi/2 - \pi/4)]. \tag{17}$$

As a result, it is reasonable to assume that all orders of the Hankel functions, $H_n(\gamma \bar{r}_o)$ and $H_{n+m}(\gamma d)$, become negligibly small when the cylinder is placed far from the ground. In this case, the magnitudes for γd and $\gamma \bar{r}_o$ are large where the asymptotic approximation for the Hankel function [cf. Eq. (17)] can be used. The constant scattering coefficients A_n obtained from Eq. (13) can then be reduced to those given in Eq. (5) which are the corresponding coefficients for the sound scattered by a hard cylinder in free space.

B. Scattering of sound by a cylinder above an impedance ground

The outdoor sound field is inevitably affected by the presence of an absorbing ground.[18] The acoustical characteristic of the ground surface affects the phase and amplitude of the reflected wave, which interferes with the direct wave to produce an intricate directivity pattern of the total sound field. An analytical model for determining the acoustical properties of ground is therefore crucial in addition to a scattering model for predicting the sound field due to the presence of a long cylinder above outdoor ground surfaces. The effect of ground impedance can be precisely modeled by an approximate analytical solution, which is known as the Weyl–van der Pol formula.[19,20] Without the scattering sound field due to the long cylinder, the total sound field in the presence of an absorbing ground is simply given by

$$\phi_T^d = \phi^d + Q_1\phi_r^d = \frac{e^{ikR}}{4\pi R} + Q_1\frac{e^{ik\bar{R}}}{4\pi \bar{R}}, \tag{18}$$

where Q_1 is the spherical wave reflection coefficient for the point source reflection on the impedance ground and R is the distance between the source and the field point. Knowing the distance $\bar{R}$ between the image source and the field point, the angle of incidence of the reflected wave α_1 (see Fig. 2), and the specific normalized admittance of the reflecting ground β_g, the spherical wave reflection coefficient can be determined by

$$Q_1 = V_1 + (1 - V_1)F(w_1), \tag{19}$$

where the plane wave reflection coefficient is given by

$$V_1 = \frac{\cos\alpha_1 - \beta_g}{\cos\alpha_1 + \beta_g} \tag{20}$$

and the boundary loss factor $F(w_1)$ is given by

$$F(w_1) = 1 + i\sqrt{\pi}w_1 e^{-w_1^2}\,\mathrm{erfc}(-iw_1), \tag{21}$$

w_1 is the numerical distance approximated by

$$w_1 = +\sqrt{\tfrac{1}{2}ik\bar{R}}(\cos\alpha_1 + \beta_g), \tag{22}$$

and erfc is the complementary error function.

Using an analogous approach, a heuristic formulation for the multi-scattering sound field due to the presence of a long cylinder placed above an impedance ground ϕ_T^s can be expressed as[21]

$$\phi_T^s = \phi^s + Q_2\phi_r^s, \tag{23}$$

where ϕ^s and ϕ_r^s are given by Eqs. (4) and (7), respectively. The spherical wave reflection coefficient Q_2 for a scattering wave reflection on an impedance ground can be derived by the standard Weyl–van der Pol formula, cf. Eq. (19). However, the angle of incidence α_2 and the distance between the centerline of the image cylinder to the receiver $\bar{r}$ are necessarily different from those used in Eq. (19). Hence, the plane wave reflection coefficient V_2 is modified to

$$V_2 = \frac{\cos\alpha_2 - \beta_g}{\cos\alpha_2 + \beta_g}, \tag{24}$$

and the boundary loss factor is calculated according to Eq. (21) where the numerical distance w_2 becomes

$$w_2 = +\sqrt{\tfrac{1}{2}ik\bar{r}}(\cos\alpha_2 + \beta_g). \tag{25}$$

The total sound field near an impedance ground can be expressed exactly in analogous form as Eq. (11a) but with the addition of Q_1 and Q_2 as multipliers for the second and the fourth terms, respectively. It is given by

$$\phi^T(r,\theta,z) = \phi_T^d + \phi_T^s = (\phi^d + Q_1\phi_r^d) + (\phi^s + Q_2\phi_r^s). \tag{26a}$$

We can substitute the appropriate terms for ϕ^d, ϕ_r^d, ϕ^s, and ϕ_r^s in Eq. (26a) to yield

$$\begin{aligned}\phi^T(r,\theta,z) = {} & \frac{i}{2}\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_o)}\int_{-\infty}^{+\infty} J_n(\gamma r)H_n(\gamma r_o)e^{i\alpha(z-z_o)}d\alpha \\ & + \frac{i}{2}Q_1\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_1)}\int_{-\infty}^{+\infty} J_n(\gamma r)H_n(\gamma \bar{r}_o)e^{i\alpha(z-z_o)}d\alpha \\ & + \sum_{n=-\infty}^{\infty} e^{in\theta}\int_{-\infty}^{+\infty} A_n H_n(\gamma r)e^{i\alpha z}d\alpha \\ & + Q_2\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} e^{in\theta}\int_{-\infty}^{+\infty} A_{-m}H_{n-m}(\gamma d)J_n(\gamma r)e^{i\alpha z}d\alpha.\end{aligned} \tag{26b}$$

A new set of coupled linear equations for solving the scattering coefficients A_n can be established. They can be expressed in a compact form similar to Eq. (13). The elements for the column vector $\mathbf{D}$ are calculated by

$$D_n = -\frac{i}{2}\frac{J_n'(\gamma\alpha)}{H_n'(\gamma a)}e^{-i\alpha z_o}[H_n(\gamma r_o)e^{-in\theta_o} + Q_1 e^{-in\theta_1}H_n(\gamma \bar{r}_o)], \tag{27}$$

where the spherical wave reflection coefficient Q_1 appears in the second term of the square bracket. The elements for the square matrix $\mathbf{M}$ are determined by

$$M_{nm} = \begin{cases} Q_2 \dfrac{J_n'(\gamma a)}{H_n'(\gamma a)} H_{n+m}(\gamma d) & \text{if } n \neq m \\ 1 + Q_2 \dfrac{J_n'(\gamma a)}{H_n'(\gamma a)} H_{2n}(\gamma d) & \text{if } n = m \end{cases}. \quad (28)$$

Comparing with Eq. (14), an extra factor Q_2 appears in the term involving the Hankel functions.

In the special case of a hard ground, the spherical wave reflection coefficients Q_1 and Q_2 are equal to 1. Hence, Eqs. (27) and (28) can be reduced to Eqs. (15) and (14), respectively. Although it is possible to consider the case of an absorbing cylinder, we only derive expressions for sound scattered by a hard cylinder because it is a more common situation in the context of the outdoor sound propagation. However, it is straightforward to extend the current model to include the scattering effect due to an absorbing cylinder.[16]

IV. SOUND SCATTERING BY A LONG CYLINDER DUE TO A LINE SOURCE NEAR AN ABSORBING GROUND

For the sake of completeness, we consider a two-dimensional scattering problem analogous to the three-dimensional problem presented in Secs. II and III. In this section, a similar analysis is presented for the sound scattered by a cylinder due to a line source above an absorbing ground. Mathematically speaking, the two-dimensional scattering problem is less involved as compared with the analytical model for scattering of spherical wave by a long cylinder because of a more complicated geometrical configuration.

The analytical solutions can be obtained readily by using the same approaches described earlier in Secs. II and III. For simplicity, the line source is placed parallel to the axial direction of the long cylinder. A two-dimensional coordinate system is used with the line source S and the field point P located at (r_o, θ_o) and (r, θ), respectively, see Figs. 1 and 2.

Using the Graf addition theorem (which is a special case of the Bessel function addition theorem),[17] the free-field sound pressure due to a line source can be expanded as

$$\phi^d = -\frac{i}{4} H_o(kR) = -\frac{i}{4} \sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_o)} J_n(kr) H_n(kr_o) \quad \text{for } r_o \geqslant r. \quad (29)$$

The scattered sound pressure can be written as

$$\phi^s = \sum_{n=-\infty}^{\infty} A_n H_n(kr) e^{in\theta}, \quad (30)$$

where A_n is the scattering coefficient which can be determined by applying the appropriate boundary condition on the surface of the long cylinder. For a hard cylinder, the expression for A_n takes the form

$$A_n = \frac{i}{4} \frac{J_n'(ka)}{H_n'(ka)} \cdot H_n(kr_o) e^{-in\theta_o}. \quad (31)$$

The free-field and scattered field potentials given in Eqs. (30) and (31) are in contrast with the spherical wave representation where the integral representation (integration with respect to the axial wave number) is not needed for the two-dimensional case.

Similarly, we can use the image source method to calculate the contributions due to the scattering of sound from the hard ground. The contributions of the reflected sound wave from the image source (ϕ_r^d) and image cylinder (ϕ_r^s) are expressed in terms of the Bessel functions as

$$\phi_r^d = -\frac{i}{4} H_o(k\bar{R}) = -\frac{i}{4} \sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_1)} J_n(kr) H_n(k\bar{r}_o) \quad \text{for } \bar{r}_o \geqslant r \quad (32)$$

and

$$\phi_r^s = \sum_{n=-\infty}^{\infty} B_n H_n(k\bar{r}) e^{in\bar{\theta}}, \quad (33)$$

where B_n is the scattering coefficients due to the reflection from the ground surface. By imposing the Neumann boundary condition [cf. Eq. (8)] at the ground surface for the total scattered sound field, we can get the same relation between A_n and B_n, as given in Eq. (9).

Suppose that the ground surface is an impedance plane with the specific normalized admittance of β_g. It has been shown that in Ref. 20 the "cylindrical" wave reflection coefficient has the same form as the spherical wave reflection coefficient. Hence the expressions, V_1 and V_2, given earlier can be used in the two-dimensional problem. As a result, the total sound field scattered by a long cylinder due to a line source excitation above an impedance ground can then be written as

$$\begin{aligned} \phi^T(r,\theta) &= \phi_T^d + \phi_T^s = (\phi^d + Q_1 \phi_r^d) + (\phi^s + Q_2 \phi_r^s) \\ &= -\frac{i}{4} \sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_o)} J_n(kr) H_n(kr_o) \\ &\quad - \frac{i}{4} Q_1 \sum_{n=-\infty}^{\infty} e^{in(\theta-\theta_1)} J_n(kr) H_n(k\bar{r}_o) \sum_{n=-\infty}^{\infty} A_n H_n(kr) e^{in\theta} \\ &\quad + Q_2 \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} A_{-m} H_{n-m}(kd) J_n(kr) e^{in\theta}. \end{aligned} \quad (34)$$

The constant coefficient of the scattered waves A_n can be determined by applying the Neumann boundary condition on the surface for an acoustically hard cylinder. Another set of coupled equations in a similar form as Eq. (13) can be constructed for the determination of A_n. The column vector $\mathbf{D}$ can be formed where the nth element is given by

$$D_n = \frac{i}{4} \frac{J_n'(ka)}{H_n'(ka)} [H_n(kr_o) e^{-in\theta_o} + Q_1 e^{-in\theta_1} H_n(k\bar{r}_o)]. \quad (35)$$

The square matrix $\boldsymbol{M}$ can be found by noting that

$$M_{nm} = \begin{cases} Q_2 \dfrac{J_n'(ka)}{H_n'(ka)} H_{n+m}(kd) & \text{if } n \neq m \\ 1 + Q_2 \dfrac{J_n'(ka)}{H_n'(ka)} H_{2n}(kd) & \text{if } n = m \end{cases}. \quad (36)$$

Equation (34) represents a general solution for an impedance ground. However, the solution can be reduced to the case for the sound field above a hard ground by setting $\beta_g = 0$. In Sec. V, we shall present experimental data to validate the analytical solution for the scattered sound field due to a point source. The experimental results will also compare with an approximation scheme for a two-dimensional problem using a line source.

V. EXPERIMENT

Indoor experiments were conducted in a $6 \times 6 \times 4$ m^3 (high) anechoic chamber for the validation of the theoretical models developed in Secs. II–IV. The anechoic chamber used in the experiments was designed with a cut-off frequency of 75 Hz and a Noise Criteria (NC) rating of NC-25. The point source was simulated by a Tannoy loudspeaker which was connected to a 1 m tube of 3 cm internal diameter. The loudspeaker was driven by a B&K 2713 power amplifier and was connected to a maximum length sequence system analyzer (MLSSA) with an MLS card installed in a desktop computer. The MLSSA system was used as a signal generator for the source and also functioned as a signal processing analyzer for all the experimental measurements. A BSWA TECH MK224 $\frac{1}{2}$ in. condenser microphone coupled with a BSWA TECH MA201 preamplifier was used as a signal receiver. The receiver system was connected to the MLS card through a BSWA TECH MC102 amplifier.

A. Scattering by a hard cylinder in free field

An initial set of measurements in a free space has been carried out which was used to validate the theoretical formulations in order to validate the theoretical model for the scattering of sound by a hard cylinder. A cylinder of radius 0.058 m and of length 2 m was made of cast iron which was painted to provide a smooth hard surface. The cylinder, which was supported at both ends by means of iron stands, was placed in the anechoic chamber. A rig was designed to measure the angular and radial positions of the source and receiver. The source and reception point were located near the mid-point along the longitudinal axis of the cylinder in order to reduce the scattering effects of the end caps of the cylinder at both sides.

The measured data for different geometrical configurations are plotted in Figs. 3 and 4. The theoretical predictions of using point source and line source for the same geometrical configurations are also shown in the figures for comparison. The relative sound pressure level (SPL), which is defined as the ratio of the total sound field relative to the reference free-field level measured at 1 m from the source, is used to present the numerical and experimental results. In the figures, the solid line, dotted line, and dashed line are used to represent the experimental results and theoretical predictions due to a point source and a line source, respectively, throughout the paper.

In Fig. 3, we show the spectra of relative SPLs with the source located at (0.35 m, 90°, 0 m) and receiver located at two locations of (0.2 m, 150°, 0 m) and (0.2 m, 120°, 0 m), respectively. The spectra in Fig. 4 were obtained with the

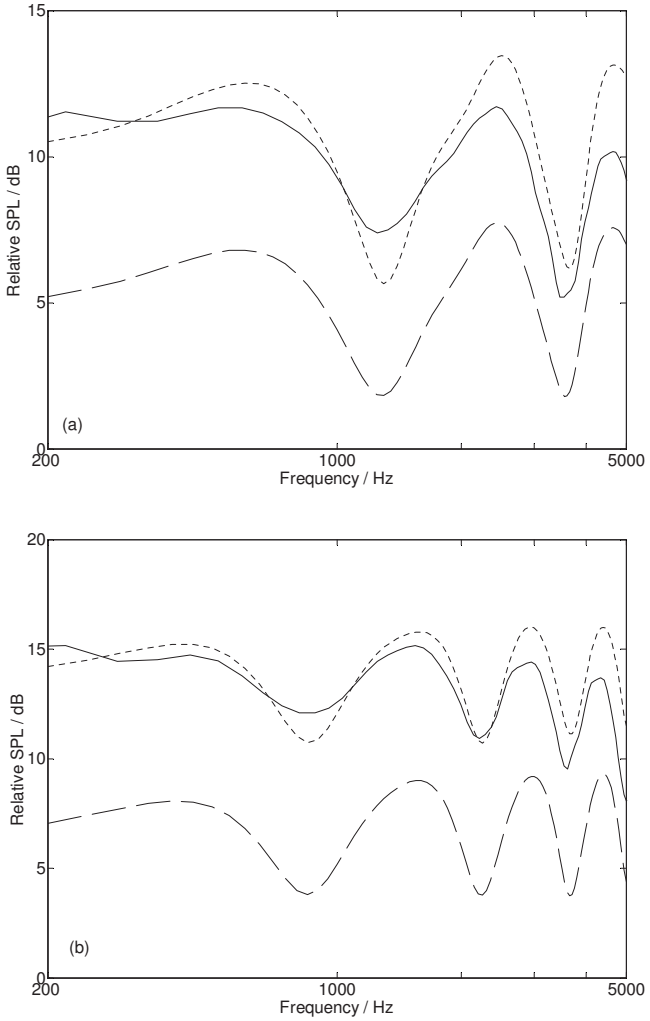


FIG. 3. Experimental results and theoretical predictions for acoustic scattering by a hard cylinder of radius 0.058 m. Source: $r_o = 0.35$ m, $\theta_o = 90°$, $z_o = 0$ m. Receiver: (a) $r = 0.2$ m, $\theta = 150°$, $z = 0$ m and (b) $r = 0.2$ m, $\theta = 120°$, $z = 0$ m. (Solid line: experimental results, dotted line: theoretical predictions due to a point source, and dashed line: theoretical predictions due to a line source.)

source located at (0.3 m, 0°, 0 m) and receiver located at (0.25 m, 180°, 0.1 m) and (0.25 m, 150°, 0.1 m), respectively. For the numerical computations of this section, the maximum number of terms N required in the summation series is set at 5. This ensures the convergence of the numerical results for calculating the scattered sound fields. In general, the theoretical predictions of using point source are in accord with the experimental measurements for different experimental configurations despite the fact that the experimental results show a general trend of noticeable differences in magnitude (the experimental data are always lower than the predicted results) at the high frequency regime above 2 kHz. These discrepancies are mainly due to the limitations of our experimental set-up because a long tube was used to represent a point source. This set-up was inadequate especially at high frequencies because the point source was neither omnidirectional nor adequate in producing sufficient signal levels for accurate indoor measurements.

As shown in Figs. 3 and 4, the two-dimensional model can give reasonably good prediction of the trend of the in-

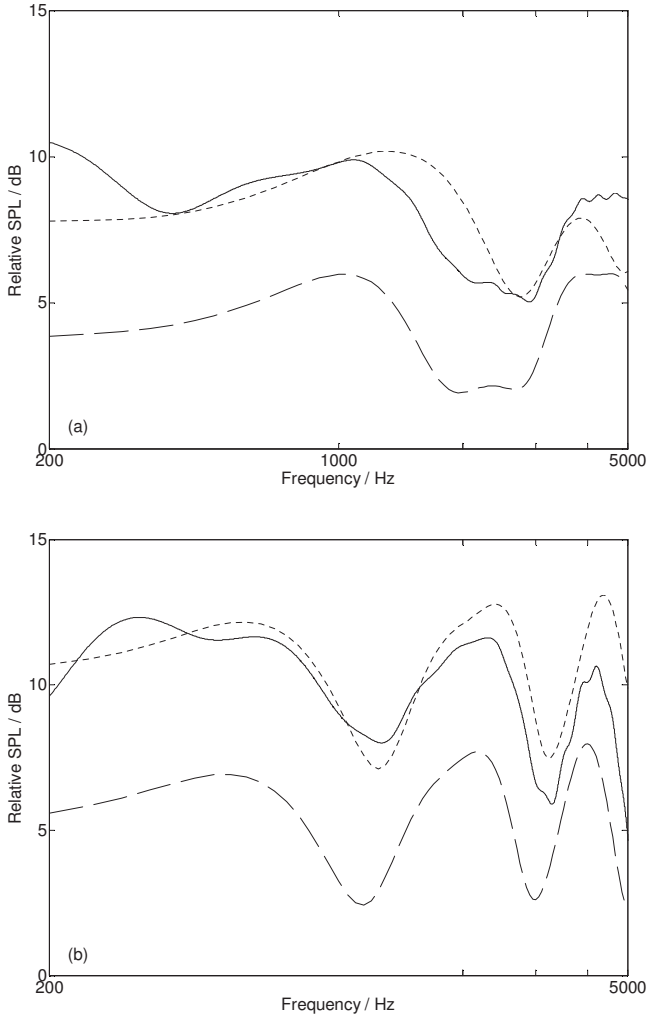


FIG. 4. Experimental results and theoretical predictions for acoustic scattering by a hard cylinder of radius 0.058 m. Source: r_o=0.3 m, θ_o=0°, z_o=0 m. Receiver: (a) r=0.25 m, θ=180°, z=0.1 m and (b) r=0.25 m, θ=150°, z=0.1 m. (Solid line: experimental results, dotted line: theoretical predictions due to a point source, and dashed line: theoretical predictions due to a line source.)

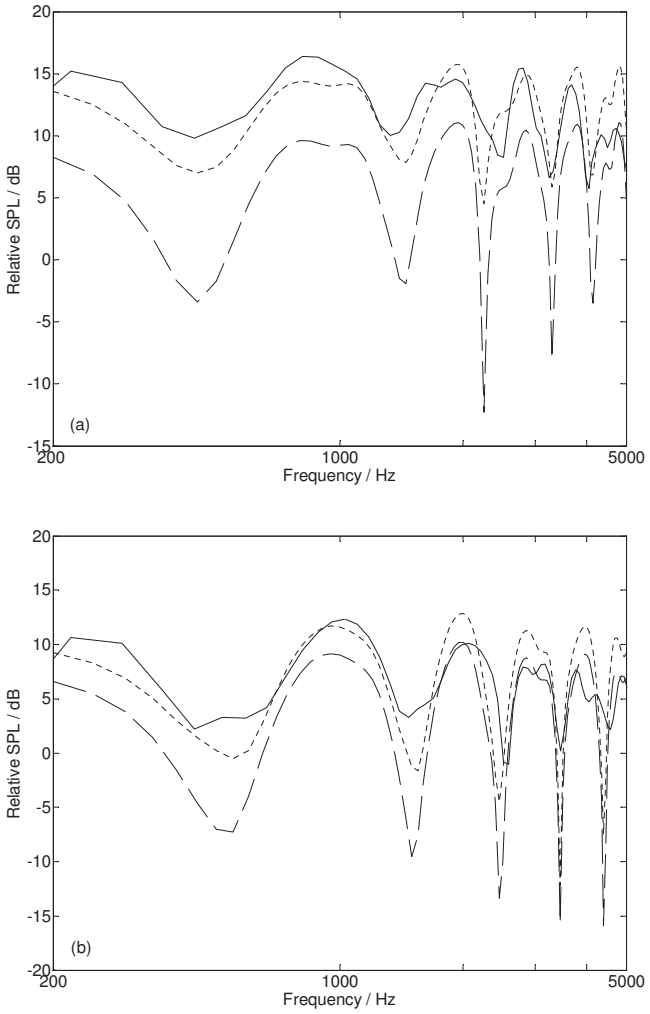


FIG. 5. Experimental results and theoretical predictions for acoustic scattering by a hard cylinder of radius 0.045 m above a hard ground. The cylinder center is 0.2 m above the hard ground. Source: r_o=0.5 m, θ_o=90°, z_o=0 m. Receiver: (a) r=0.4 m, θ=120°, z=0 m and (b) r=0.4 m, θ=150°, z=0 m. (Solid line: experimental results, dotted line: theoretical predictions due to a point source, and dashed line: theoretical predictions due to a line source.)

terference pattern. The locations of the frequency "peaks" and "dips" match quite well with the experimental results and the prediction according to the three-dimensional model. Nevertheless, there are significant discrepancies in the magnitude of the relative SPL.

B. Scattering by a hard cylinder above a hard ground

A varnished hardwood board of 0.02 m thick was used to model the hard ground. Two cast-iron cylinders of respective radii 0.045 and 0.058 m and of length 2 m were used. Both cylinders were painted to provide smooth surfaces. The cylinders were put above the hard ground in the same way as described in Sec. V A. The first set of experiments was done with the center of the cylinder at 0.2 m above the hard ground surface and the source was situated at (0.5 m, 90°, 0 m); the calculation and experimental results were plotted in Fig. 5.

The second set of experiments was conducted with the source offset by 0.16 m along the longitudinal axis of the cylinder and the receiver offset by 0.1 m. In these cases, the center of the cylinder was placed at 0.1 m above the hard ground and the source was situated at (0.55 m, 90°, 0 m). The theoretical predictions and experimental results were presented in Fig. 6. In the numerical analyses, N was set at 20 and the far-field approximations were used in the calculations. We can observe that the numerical results according to the point source model can generally give good agreements with the experimental data within the frequency range of interest. Again, the two-dimensional model can only give reasonably good agreements with the locations of dips and peaks of the spectra.

C. Scattering by a hard cylinder above an impedance ground

The ground with impedance was simulated by laying a thick carpet on the hardwood board. The carpet can be regarded as a locally reacting ground, and a two-parameter model[22] was used to predict its specific normalized admittance as follows:

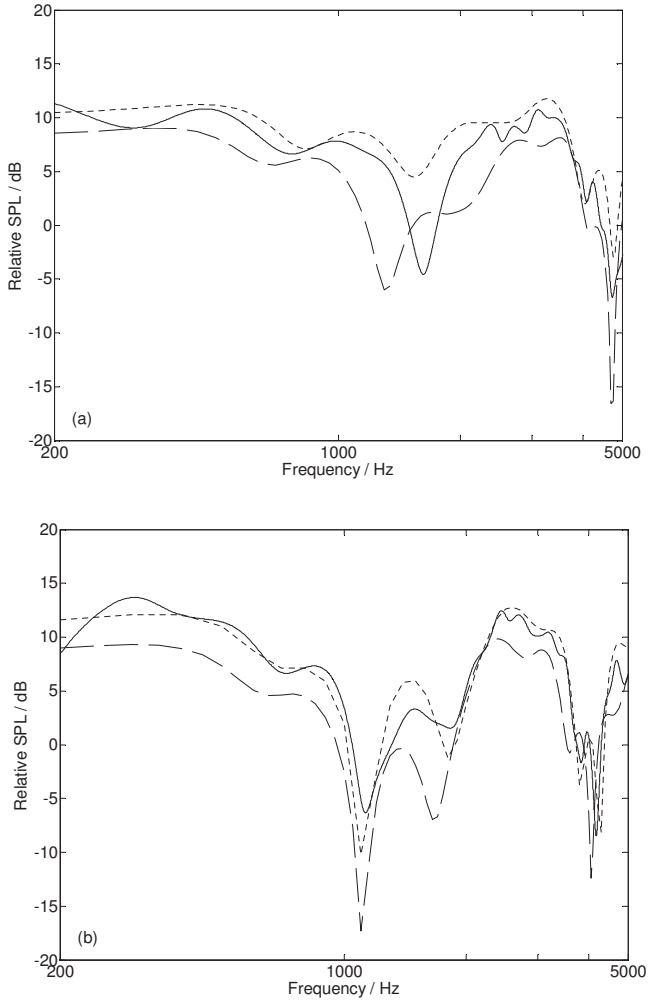


FIG. 6. Experimental results and theoretical predictions for acoustic scattering by a hard cylinder of radius 0.058 m above a hard ground. The cylinder center is 0.1 m above the hard ground. Source: $r_o=0.55$ m, $\theta_o=90°$, $z_o=0$ m. Receiver: (a) $r=0.24$ m, $\theta=160°$, $z=0.16$ m and (b) $r=0.28$ m, $\theta=150°$, $z=0.1$ m. (Solid line: experimental results, dotted line: theoretical predictions due to a point source, and dashed line: theoretical predictions due to a line source.)

$$\beta_g = \frac{1}{0.436(1+i)(\sigma_e/f)^{1/2} + 19.48 i \alpha_e/f}, \tag{37}$$

where f is the frequency, σ_e is the effective flow resistivity, and α_e is the effective rate of change in porosity with depth of the material. Prior measurements were conducted to characterize the acoustical properties of the carpet. Figure 7 shows a typical result of theoretical prediction by the Weyl–van der Pol formula and experimental measurements for characterizing the impedance (admittance) of the carpet. The measurement was carried out with a point source near an impedance surface and without the presence of a scattering cylinder. Both the source and receiver were located at 0.16 m above the ground and at a horizontal separation of 1 m. These measurement results were subsequently fitted with the two-parameter model, and the best-fit parameters were found to be $\sigma_e=10$ kPa s m^{-2} and $\alpha_e=80$ m^{-1}. These acoustical parameters were used in our subsequent predictions of acoustic scattering by a hard cylinder above an impedance ground.

Two sets of experiments with different sizes of cast-iron cylinders and source-receiver configurations were used in this case, as shown in Figs. 8 and 9. The first set of experiments was performed with the cylinder of 0.045 m in radius and its center was placed at 0.2 m above the impedance ground. The second set of experiments was conducted with the cylinder of 0.058 m in radius and its center was located at 0.085 m above the impedance ground. The receiver was offset by 0.1 m to the source along the longitudinal axis of the cylinder in this case.

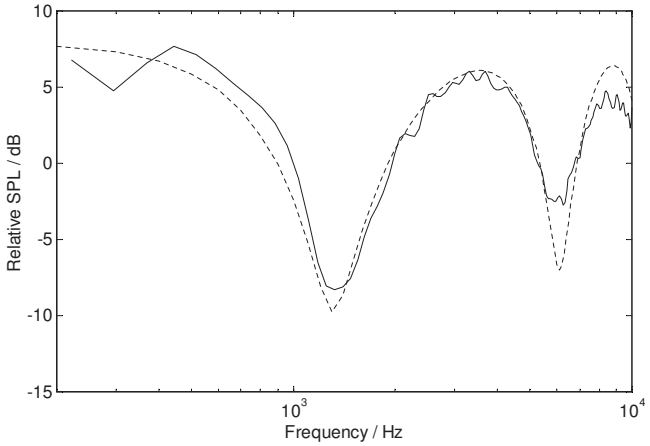


FIG. 7. Characterization of a thick carpet. Source height=0.16 m, receiver height=0.16 m, and range=1 m. The best-fit parameters are found to be $\sigma_e=10$ kPa s m^{-2} and $\alpha_e=80$ m^{-1}. (Dotted line: theoretical predictions by Weyl–van der Pol formula and two-parameter model and solid line: experimental results.)

In general, the predictions of using a point source agree satisfactorily with the experimental results for various source-receiver configurations irrespective of whether the receiver was in the illuminated zone or in the penumbra zone, as illustrated in Fig. 9. Once again, we can see similar discrepancies between the predictions of using different source incidences as concluded in Secs. V A and V B. However, it is noticeable that the differences in the shifts and magnitudes of the spectra diminish when the receiver migrates from the illuminated zone to the penumbra zone, see Fig. 9. It is rational to have this observation as the total sound pressure in the penumbra or shadow zone is mainly contributed by the interfering wavefronts spreading from the scatterer at the plane of receiver irrespective whether the source is finite or infinite. In this case, the wavefronts of the diffracted waves either due to a spherical wave incident or a line wave incident on the object do not differ too much as long as the positions of the monopole source and the receiver along the longitudinal axis of the long cylinder are close.

VI. CONCLUSIONS

The acoustic scattering by a long cylinder placed above a ground surface was studied theoretically and experimentally. The multi-scattering between the cylinder and the ground surface can be tackled tactfully by considering the reflected waves from the image cylinder. The theoretical formulations are represented in Fourier integrals, and the integrals of the scattered sound fields due to a point source excitation can be evaluated analytically by the method of stationary phase. The effect of an impedance ground in the

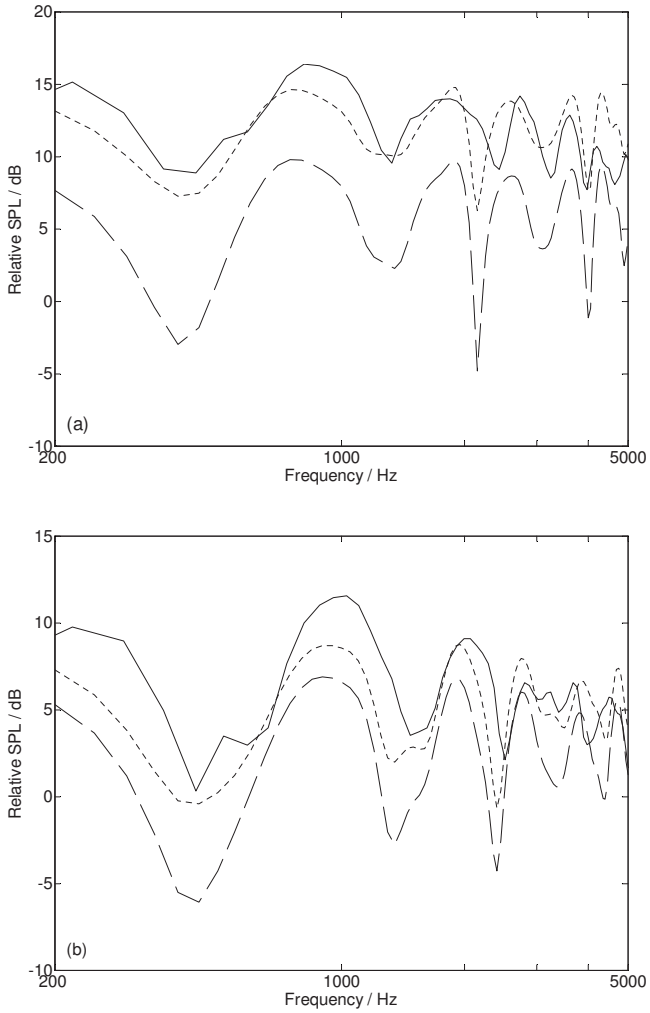


FIG. 8. Experimental results and theoretical predictions for acoustic scattering by a hard cylinder of radius 0.045 m above an impedance ground. The cylinder center is 0.2 m above the impedance ground. Source: r_o=0.5 m, θ_o=90°, z_o=0. Receiver: (a) r=0.4 m, θ=120°, z=0 m and (b) r=0.45 m, θ=160°, z=0 m. (Solid line: experimental results, dotted line: theoretical predictions due to a point source, and dashed line: theoretical predictions due to a line source.)

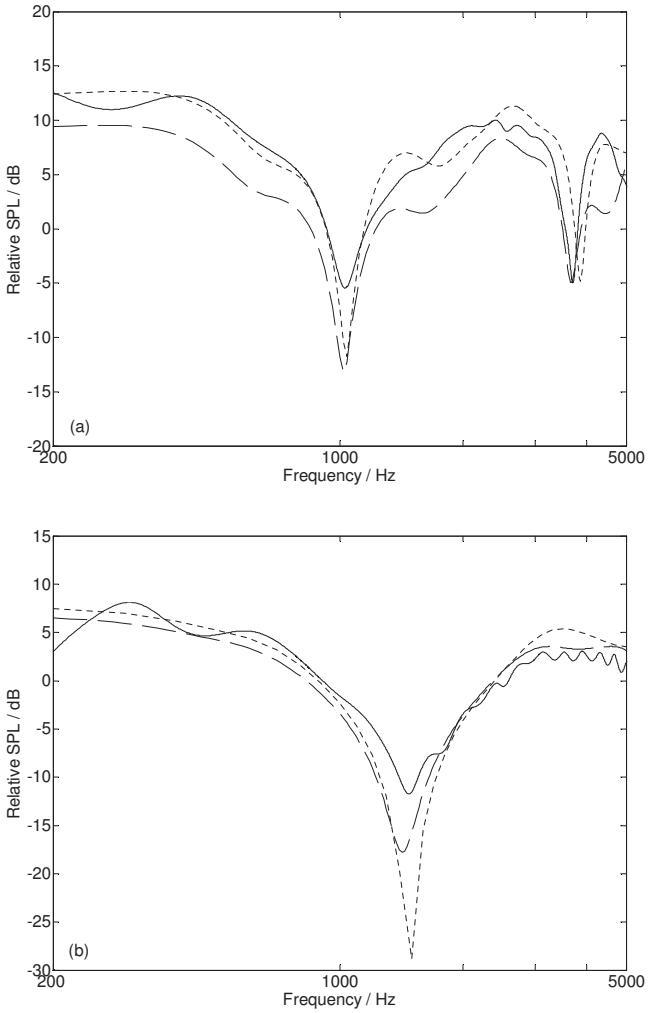


FIG. 9. Experimental results and theoretical predictions for acoustic scattering by a hard cylinder of radius 0.058 m above an impedance ground. The cylinder center is 0.08 m above the impedance ground. Source: r_o=0.5 m, θ_o=90°, z_o=0. Receiver: (a) r=0.29 m, θ=150°, z=0.1 m and (b) r=0.35 m, θ=240°, z=0.1 m. (Solid line: experimental results, dotted line: theoretical predictions due to a point source, and dashed line: theoretical predictions due to a line source.)

presence of a long cylinder can successfully be handled by using the image source method and the Weyl–van der Pol formula. The numerical results according to the point source model can generally give good agreements with the experimental data within the frequency range of interest. On the other hand, numerical solutions for the line source model only can give the overall interference pattern for the spectra of the relative sound pressure levels. However, this two-dimensional model generally underpredicts the measured sound pressure levels. But the discrepancies are less obvious when the receiver is moved from the illuminated zone to the penumbra zone because of the fact that the total sound field in the penumbra or shadow zone is primarily contributed by the intrusive waves scattering from the object at the plane of field point irrespective if the airborne source is a finite or infinite. Generally speaking, a two-dimensional sound scattering model cannot readily be used to simulate a three-dimensional scattering problem.

ACKNOWLEDGMENTS

The initial part of this research was undertaken while one of the authors (W.K.L.) was Research Associate at Ray W. Herrick Laboratories, School of Mechanical Engineering, Purdue University. The authors thank the financial, technical, and administrative support offered by the School of Mechanical Engineering, Purdue University. W.K.L. also thanks Hong Kong Polytechnic University for partial funding support (PolyU Research Grant No. G-U547) on this project.

[1]F. M. Wiener, "Sound diffraction by rigid spheres and circular cylinders," J. Acoust. Soc. Am. **19**, 444–451 (1947).
[2]J. J. Faran, "Sound scattering by solid cylinders and spheres," J. Acoust. Soc. Am. **23**, 405–418 (1951).
[3]L. Flax, V. K. Varadan, and V. V. Varadan, "Scattering of an obliquely incident acoustic wave by an infinite cylinder," J. Acoust. Soc. Am. **68**, 1832–1835 (1980).
[4]T. Li and M. Udea, "Sound scattering of a plane obliquely incident on a cylinder," J. Acoust. Soc. Am. **86**, 2363–2368 (1989).
[5]H. D. Maniar and J. N. Newman, "Wave diffraction by a long array of cylinders," J. Fluid Mech. **339**, 309–330 (1997).
[6]Y. Y. Chen and Z. Ye, "Acoustic attenuation by two-dimensional arrays of

rigid cylinders," Phys. Rev. Lett. **87**, 184301 (2001).
[7]J. C. Piquette, "Spherical wave scattering by an elastic solid cylinder of infinite length," J. Acoust. Soc. Am. **79**, 1248–1259 (1986).
[8]T. Li and M. Udea, "Sound scattering of a spherical wave incident on a cylinder," J. Acoust. Soc. Am. **87**, 1871–1879 (1990).
[9]C. R. Fuller, "Free-field correction factor for spherical acoustic waves impinging on cylinders," AIAA J. **27**, 1722–1726 (1989).
[10]E. A. Skelton and J. H. James, *Theoretical Acoustics of Underwater Structures* (Imperial College Press, London, 1997).
[11]J. C. Bertrand and J. W. Young, "Multiple scattering between a cylinder and a plane," J. Acoust. Soc. Am. **60**, 1265–1269 (1976).
[12]A. Coatanhay, "Scattering near a plane interface using a generalized method of images approach," J. Comput. Acoust. **12**, 233–256 (2004).
[13]Z. Ye and Y. Y. Chen, "Acoustic scattering by a cylinder near a pressure release surface," http://arxiv.org/abs/physics/0103070 (Last viewed July, 2009).
[14]E. L. Shenderov, "Diffraction of sound by an elastic cylinder near the surface of an elastic halfspace," Acoust. Phys. **48**, 225–234 (2002).
[15]W. C. Chew, *Waves and Fields in Inhomogeneous Media* (IEEE, New York, 1995), p. 591.
[16]K. M. Li, W. K. Lui, and G. H. Frommer, "The diffraction of sound by an impedance sphere in the vicinity of a ground surface," J. Acoust. Soc. Am. **115**, 42–56 (2004).
[17]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
[18]T. F. W. Embleton, "Tutorial on sound propagation outdoors," J. Acoust. Soc. Am. **100**, 31–48 (1996).
[19]L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1980).
[20]K. Attenborough, K. M. Li, and K. Horoshenkov, *Predicting Outdoor Sound* (Taylor & Francis, London, 2007).
[21]S. M. Hasheminejad and M. Azarpeyvand, "Modal vibrations of a cylindrical radiator over an impedance plane," J. Sound Vib. **278**, 461–477 (2004).
[22]K. Attenborough, "Ground parameter information for propagation modeling," J. Acoust. Soc. Am. **92**, 418–427 (1992).

A residual-potential boundary for time-dependent, infinite-domain problems in computational acoustics

Thomas L. Geers[a)]
Department of Mechanical Engineering, University of Colorado, Boulder, Colorado 80309-0427

Michael A. Sprague
School of Natural Sciences, University of California, Merced, California 95343

(Received 1 July 2009; revised 19 November 2009; accepted 23 November 2009; publisher error corrected 9 March 2010)

A theoretically exact computational boundary is introduced that is based on modal residual potentials for the spherical geometry. The boundary produces a set of first-order, uncoupled ordinary differential equations for nodal boundary responses, and a set of uncoupled time-stepping equations for modal boundary responses. The two sets are coupled through nodal-modal transformation based on the orthogonal surface functions for the spherical boundary. Numerical results generated with the boundary are presented for a step-wave-excited, elastic, spherical shell submerged in an infinite acoustic medium. Extension of the method to other separable geometries for partial differential equations defined in unbounded domains is mentioned.
 [DOI: 10.1121/1.3273900]

PACS number(s): 43.20.Px [LLT] Pages: 675–682

I. INTRODUCTION

Following the pioneering papers of Engquist and Majda (1977) and Bayliss and Turkel (1980), many researchers have documented the development and application of computational boundaries for infinite-domain problems (see, e.g., Geers, 1998; Givoli, 2004; Givoli and Harari, 1998; Tsynkov, 1998; Turkel, 1998; Wolf and Song, 1996). With the aid of Fig. 1, a three-dimensional computational (or artificial or non-reflecting) boundary might be defined as "a surface that replaces the infinite exterior volume surrounding an interior volume of interest." This definition excludes infinite elements (Bettess, 1992; Astley, 1996) and perfectly matched layers (Berenger, 1994; Qi and Geers, 1998), although both are widely used to accomplish the purpose of a computational boundary.

Typically, the infinite-domain problem involves a body (or bodies) of interest embedded in a uniform infinite medium (Fig. 1). Ideally, the computational boundary would reside on the body's surface, thereby eliminating the need for any medium to be included in the interior volume. In many cases, however, the complex geometry of the body precludes this option.

The essential property of a computational boundary is that it enables one to obtain satisfactorily accurate solutions in the interior volume; this implies, of course, that the boundary does not introduce numerical instabilities. Desirable properties of a computational boundary are as follows:

1. *Geometric locality.* The boundary hugs the body, which minimizes the size of the interior volume.
2. *Spatial locality.* Only adjacent surface elements interact at a given time, which produces banded coefficient matrices characteristic of finite-element and finite-difference models.
3. *Temporal locality.* Only recent data are required to advance the solution in time, which admits efficient computation.
4. *Ease of implementation.* Users can readily understand and implement the boundary, which facilitates the development of "ownware."

The acoustic fields in a radiation problem are the total fields in the medium; hence, a computational boundary for acoustic radiation is formulated in terms of total-field response variables. In contrast, the acoustic fields in a scattering problem are the sums of the respective incident and scattered fields; hence, a computational boundary for acoustic scattering is formulated in terms of scattered-field response variables.

It is convenient to sort computational boundaries into two categories: *absorbing boundaries*, which are designed to absorb waves propagating outward from the body, and *impedance boundaries*, which are designed to represent the exterior volume to the interior volume for both propagating- and evanescent-wave components. It has been found that an absorbing boundary can provide properties 2–4, but not property 1 because it absorbs evanescent waves as well as propagating waves. Impedance boundaries can provide properties 1, 3, and 4, but not property 2, because surface representations for evanescent waves are inherently non-local. Hagstrom *et al.* (2008) approached bridging the gap between absorbing and impedance boundaries by introducing evanescent modes into the formulation of high-order local absorbing boundaries.

This is a good place to mention that a perfectly matched layer corresponds to an absorbing boundary, as it incorporates high dissipation in otherwise compatible layers that en-

[a)]Author to whom correspondence should be addressed. Electronic mail: Thomas.Geers@Colorado.edu.

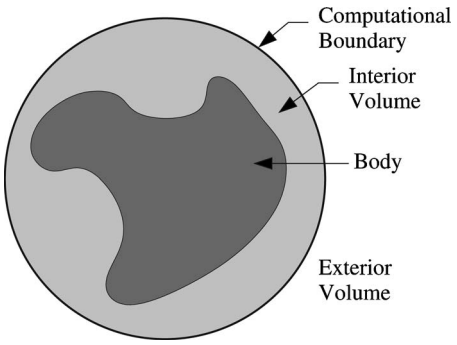


FIG. 1. Schematic of an infinite-domain problem.

close the interior volume. Hence, such a boundary is successfully used well away from the body, where the evanescent fields have decayed to negligibility. An array of infinite elements corresponds to an impedance boundary that can provide properties 1–3. It does this by utilizing nodes on "porcupine quills" that extend out into the infinite exterior volume. Although infinite elements are widely used, they have the following limitations: (1) the geometry of the infinite-element mesh must conform to a separable coordinate system for the wave equation and (2) improving solution accuracy requires increasing the order of the elements, which may lead to ill-conditioning problems (Astley, 2000).

A. Early- and late-time boundaries

Theoretically exact and approximate computational boundaries may be constructed through the use of Kirchhoff's integral equation (KIE), which may be written for pressure p at position $\mathbf{R}$ on the boundary and time t as (Baker and Copson, 1939)

$$2\pi p(\mathbf{R},t) = \int_S^{\mathbf{R}'\neq\mathbf{R}} \left\{ \rho \frac{\dot{v}(\mathbf{R}',t_r)}{|\mathbf{R}'-\mathbf{R}|} - \frac{\mathbf{n}'\cdot(\mathbf{R}'-\mathbf{R})}{|\mathbf{R}'-\mathbf{R}|^3}\left[p(\mathbf{R}',t_r) + \frac{|\mathbf{R}'-\mathbf{R}|}{c}\dot{p}(\mathbf{R}',t_r)\right]\right\} dS', \tag{1}$$

where $v(\mathbf{R}',t)$ is the normal fluid velocity at the boundary, the overdot denotes a time derivative, $t_r = t - |\mathbf{R}'-\mathbf{R}|/c$ is the retarded time, $\mathbf{n}'$ is the outward unit normal to the boundary at $\mathbf{R}'$, and c is the speed of sound. Laplace transformation of this equation yields

$$\int_S \frac{\mathbf{n}'\cdot(\mathbf{R}'-\mathbf{R})}{|\mathbf{R}'-\mathbf{R}|^3}\left(1+|\mathbf{R}'-\mathbf{R}|\frac{s}{c}\right)e^{-|\mathbf{R}'-\mathbf{R}|s/c}P(\mathbf{R}',s)dS' = \rho s\int_S \frac{e^{-|\mathbf{R}'-\mathbf{R}|s/c}}{|\mathbf{R}'-\mathbf{R}|}V(\mathbf{R}',s)dS', \tag{2}$$

in which the singularity that produces the term on the left side of Eq. (1) has been retained in the integral on the left side of Eq. (2). Although KIE possesses property 1, it does not possess properties 2–4. In addition, the time-retarded terms in Eq. (1) generate numerical-instability problems. Hence, KIE is not an attractive foundation for a computational boundary.

However, KIE *is* a useful vehicle for the construction of early-time-approximation (ETA) and late-time-approximation (LTA) boundaries. Felippa (1980) and Geers (1991) systematically constructed three ETA boundaries by (1) parametrizing the surface around a field point, (2) noting the region of influence at a prescribed early time, (3) expanding the surface fields in Fourier series, and (4) retaining terms of appropriate orders. The results were

$$\text{ETA1:}\quad p(\mathbf{R},t) = \rho c v(\mathbf{R},t), \tag{3a}$$

$$\text{ETA2:}\quad \dot{p}(\mathbf{R},t) + c\kappa(\mathbf{R})p(\mathbf{R},t) = \rho c\dot{v}(\mathbf{R},t), \tag{3b}$$

$$\text{ETA3:}\quad \ddot{p}(\mathbf{R},t) + c\kappa(\mathbf{R})\dot{p}(\mathbf{R},t) = \rho c\ddot{v}(\mathbf{R},t) + \tfrac{1}{2}c^2[\kappa^2(\mathbf{R}) - \tau(\mathbf{R}) + \nabla_S^2]v(\mathbf{R},t), \tag{3c}$$

where κ and τ are mean and total curvature (Struik, 1961). ETA1 is the well-known plane-wave approximation (PWA). It is a point relation, which is the quintessence of spatial locality. ETA2 is known as the curved-wave approximation (CWA), as it introduces the effects of local curvature into ETA1; it is also a point relation. ETA3 adds additional curvature terms, as well as the surface Laplacian, which destroys the point-relation attribute but maintains spatial locality. Equations (3a)–(3c) also proceed from ray theory (Marston, 1997), which pertains to high acoustic wavenumbers. Hence, they are also high-frequency approximations for continuous surface fields. Early-time boundaries are highly absorptive and spatially local.

Geers and Zhang (1994) and Geers and Toothaker (2000) systematically constructed three LTA boundaries from the Laplace-transformed KIE by (1) expanding the exponentials in Maclauren series, (2) retaining terms of appropriate powers of s, and (3) inverse-transforming the results. This produced

$$\text{LTA1:}\quad \gamma p(\mathbf{R},t) = \rho\beta\dot{v}(\mathbf{R},t), \tag{4a}$$

$$\text{LTA2:}\quad \gamma p(\mathbf{R},t) = \rho\left[\beta\dot{v}(\mathbf{R},t) - \frac{1}{c}\alpha\ddot{v}(\mathbf{R},t)\right], \tag{4b}$$

$$\text{LTA3:}\quad \gamma p(\mathbf{R},t) - \frac{1}{c^2}\eta\ddot{p}(\mathbf{R},t) = \rho\left[\beta\dot{v}(\mathbf{R},t) - \frac{1}{c}\alpha\ddot{v}(\mathbf{R},t) + \frac{1}{c^2}\mu\dddot{v}(\mathbf{R},t)\right], \tag{4c}$$

where the Greek symbols (with the exception of ρ) denote non-local, spatial integral operators.

LTA1 is the well-known added-mass approximation, which governs hydrodynamic flow. LTA2 introduces the first-order effects of compressibility, embodied in the second term on the right, and LTA3 includes both first- and second-order compressibility effects. LTAs are accurate when the initial conditions for the late-time fields are essentially hydrodynamic and the excitations during the late period produce only low-frequency response. This typically occurs

when the radiation or scattering problem is dominated by low-frequency (long-acoustic-wavelength) components from the start. Late-time boundaries are minimally absorptive and spatially non-local.

B. Impedance boundaries

Doubly asymptotic approximations (DAAs) produce approximate impedance boundaries for convex, piecewise-smooth surfaces (Geers, 1978; Nicolas-Vullierme, 1991; Geers and Zhang, 1994; Wolf and Song, 1996; Geers and Toothaker, 2000). They are highly accurate at both early and late times, and have been found to be surprisingly accurate at intermediate time. As developed with a matching technique that employs ETAs and LTAs, the first three members are

$$\text{DAA1:}\quad \dot{p}(\mathbf{R},t)+c\beta^{-1}\gamma p(\mathbf{R},t)=\rho c\dot{v}(\mathbf{R},t), \tag{5a}$$

$$\text{DAA2:}\quad \ddot{p}(\mathbf{R},t)+c(\chi+\kappa)\dot{p}(\mathbf{R},t)+c^2\chi\beta^{-1}\gamma p(\mathbf{R},t) = \rho c[\ddot{v}(\mathbf{R},t)+c\chi\dot{v}(\mathbf{R},t)], \tag{5b}$$

$$\text{DAA3:}\quad \dddot{p}(\mathbf{R},t)+c(\lambda\beta^{-1}\alpha+\zeta+\kappa)\ddot{p}(\mathbf{R},t) + c^2\zeta\beta^{-1}\gamma\dot{p}(\mathbf{R},t)+c^3\lambda\beta^{-1}\gamma p(\mathbf{R},t)=\rho c[\dddot{v}(\mathbf{R},t) + c(\lambda\beta^{-1}\alpha+\zeta)\ddot{v}(\mathbf{R},t)+c^2\lambda\dot{v}(\mathbf{R},t)], \tag{5c}$$

where, again, the Greek symbols (with the exceptions of ρ and κ) denote spatial integral operators. The principal advantage of DAA boundaries is their geometric versatility. Their principal limitation is that improving solution accuracy requires increasing DAA order, the difficulty of which increases dramatically after DAA2 (Geers and Toothaker, 2000). DAA boundaries are spatially non-local.

The Dirichlet-to-Neumann (DtN) boundary (Givoli, 1992; Grote and Keller, 1995a, 1995b, 1996) is theoretically exact in the context of generalized Fourier series. Although conceptually derivable for any geometry in which the wave equation is separable, only a circular boundary for two-dimensional computations and a spherical boundary for three-dimensional computations have been developed. Consider an imaginary spherical surface of radius R in an infinite acoustic domain. On and outside that surface, an outgoing acoustic wave may be described in spherical coordinates (r,θ,φ) by a *displacement potential* $\phi(r,\theta,\varphi,t)$, where the acoustic pressure and fluid displacement are obtained from ϕ as $p=\rho(\partial^2\phi/\partial t^2)$ and $\mathbf{u}=-\nabla\phi$, respectively. It is well established that the potential field on the surface may be expressed as (Morse and Ingard, 1968)

$$\phi(\theta,\varphi,t)=\sum_{n=0}^{\infty}\sum_{m=0}^{n}\phi_{nm}^{c}(t)P_n^m(\cos\theta)\cos(m\varphi) + \sum_{n=1}^{\infty}\sum_{m=1}^{n}\phi_{nm}^{s}(t)P_n^m(\cos\theta)\sin(m\varphi), \tag{6}$$

where P_n^m is the Legendre function of degree n and order m. The spherical DtN boundary for $n_{\max}=N$ then has the form (Grote and Keller, 1995a)

$$\dot{\phi}(\theta,\varphi,t)+\frac{c}{R}\phi(\theta,\varphi,t)=cu(\theta,\varphi,t)-\left(\frac{c}{R}\right)\left[\sum_{n=0}^{N}\sum_{m=0}^{n}\sum_{j=1}^{n}c_{nmj}^{c}\left(\frac{R}{c}\right)^{n-j}\frac{d^{n-j}}{dt^{n-j}}w_{nm}^{c}(t)P_n^m(\cos\theta)\cos(m\varphi) + \sum_{n=1}^{N}\sum_{m=1}^{n}\sum_{j=1}^{n}c_{nmj}^{s}\left(\frac{R}{c}\right)^{n-j}\frac{d^{n-j}}{dt^{n-j}}w_{nm}^{s}(t)P_n^m(\cos\theta)\sin(m\varphi)\right], \tag{7}$$

in which $u(\theta,\varphi,t)$ is the radial fluid displacement at the surface, c_{nmj}^{c} and c_{nmj}^{s} are constants, and $w_{nm}^{c}(t)$ and $w_{nm}^{s}(t)$ are solutions to the ordinary differential equation (ODE):

$$\sum_{n=1}^{n}\alpha_{nj}\left(\frac{R}{c}\right)^{n-j}\frac{d^{n-j}}{dt^{n-j}}w_{nm}(t)=(-1)^n\phi_{nm}(t), \tag{8}$$

where the α_{nj} are constants, and superscripts c and s have been suppressed. In Grote and Keller, 1996, the high-order ODE [Eq. (8)] is replaced by a set of first-order state equations, but the high-order nature of Eq. (8) remains. The problem has, however, been ameliorated by Huan and Thompson (2000) and Thompson and Huan (2000).

The principal advantage of the DtN boundary is that accuracy may be easily and systematically improved by merely increasing N. The principal disadvantages are as follows: (1) the high order of Eq. (8) for large n, which raises ill-conditioning questions for large N (as in the case of infinite elements), (2) the boundary can only hug bodies that are geometrically compact, and (3) the formulation is sufficiently complicated that extension to more versatile separable geometries is problematic.

II. RESIDUAL-POTENTIAL BOUNDARY

In this section, we formulate and evaluate a time-dependent impedance boundary that possesses the advantages of a DtN boundary but avoids some of its disadvantages. It is based on the concept of the residual-potential, which has been previously employed to solve a variety of canonical problems (Geers, 1969, 1971, 1972; Akkaş and Engin, 1980; Akkaş and Bauld, 1981; Geers and Yen, 1981, 1989; Akkaş, 1985). This residual-potential (RP) boundary is spherical and spatially non-local.

Again, we consider an imaginary spherical surface of radius R. It can be shown that each of the modal responses $\phi^c_{nm}(t)$ and $\phi^s_{nm}(t)$ in Eq. (6) is the solution to the equation (Zhang and Geers, 1993)

$$\dot{\phi}_{nm}(t) + \frac{c}{R}\phi_{nm}(t) = cu_{nm}(t) - \frac{c}{R}\psi_{nm}(t), \tag{9}$$

where

$$\psi_{nm}(t) = \frac{c}{R}\int_0^t r_n(t-t')\phi_{nm}(t')dt'. \tag{10}$$

In Eqs. (9) and (10) the superscripts c and s are suppressed. Each $u_{nm}(t)$ is to $u(\theta,\varphi,t)$ and each $\psi_{nm}(t)$ is to $\psi(\theta,\varphi,t)$, as each $\phi_{nm}(t)$ is to $\phi(\theta,\varphi,t)$ [see Eq. (6)]. Finally, $\psi(\theta,\varphi,t)$ is the surface *residual potential*, and each $r_n(t)$ is a nondimensional *residual function*. Apparently unaware of the work of Geers and his co-workers (1969, 1971, 1972, 1981, 1989, 1994) and Akkaş and his co-workers (1980, 1981, 1985), Alpert *et al.* (2000, 2002) rediscovered these functions, but named them "nonreflecting boundary kernels."

The Laplace transform of $r_n(t)$ is given by the following ratio of polynomials (Zhang and Geers, 1993):

$$R_n(s) = \frac{\sum_{\ell=1}^{n} \ell\Gamma_{n\ell}(c/R)^{\ell-1}s^{n-\ell}}{\sum_{\ell=0}^{n} \Gamma_{n\ell}(c/R)^{\ell}s^{n-\ell}}, \tag{11}$$

where $\Gamma_{n\ell} = (n+\ell)!/2^{\ell}\ell!(n-\ell)!$. For example, the first three residual functions in the s- and t-domains are

$$R_0(s) = 0, \quad r_0(t) = 0,$$

$$R_1(s) = (s + c/R)^{-1}, \quad r_1(t) = e^{-ct/R},$$

$$R_2(s) = \frac{3(s+2c/R)}{s^2 + 3(c/R)s + 3(c/R)^2},$$

$$r_2(t) = e^{-(3ct/2R)}\left[3\cos\left(\frac{1}{2}\sqrt{3}ct/R\right) + \sqrt{3}\sin\left(\frac{1}{2}\sqrt{3}ct/R\right)\right]. \tag{12}$$

In general, an even-indexed residual function consists of $n/2$ pairs of complex-conjugate exponentials, and an odd-indexed residual function consists of one real exponential and $(n-1)/2$ pairs of complex-conjugate exponentials. A single $r_n(t)$ serves to determine $(2n+1)$ $\psi_{nm}(t)$. Finally, $\phi_{nm}(0) = \dot{\phi}_{nm}(0) = 0$ and $\psi_{nm}(0) = \dot{\psi}_{nm}(0) = 0$. The residual functions for $n = 1-6$ are shown in Fig. 2.

Now Eq. (10) is non-local in time, so we must find an alternate expression for it. One possibility is to take the Laplace transform to get $\Psi_{nm}(s) = (c/R)R_n(s)\Phi_{nm}(s)$, introduce Eq. (11), multiply through by the denominator of Eq. (11), and inverse-transform to get an nth-order ODE for $\psi_{nm}(t)$ in terms of $\phi_{nm}(t)$ and its derivatives (Zhang and Geers, 1993). However, this parallels the DtN method, leading to the first disadvantage mentioned at the end of Sec. I. A second possibility, *the path taken here*, is to express $R_n(s)$ as a partial-fraction expansion and inverse-transform the result to obtain

$$r_n(t) = \begin{cases} 2\sum_{k=1}^{n/2} e^{-\beta_n^k ct/R}[\beta_n^k \cos(\omega_n^k ct/R) + \omega_n^k \sin(\omega_n^k ct/R)], & n \text{ even} \\ \beta_n^0 e^{-\beta_n^0 ct/R} + 2\sum_{k=1}^{(n-1)/2} e^{-\beta_n^k ct/R}[\beta_n^k \cos(\omega_n^k ct/R) + \omega_n^k \sin(\omega_n^k ct/R)], & n \text{ odd}, \end{cases} \tag{13}$$

where the constants β_n^k and ω_n^k emerge from the partial-fraction expansion. We calculated these constants with the residue function in MATLAB (Mathworks, Inc., Natick, MA); they are listed in Table I for $1 \le n \le 8$.

Introducing Eq. (13) into Eq. (10) and reversing the order of integration and summation, we obtain

$$\psi_{nm}(t) = \begin{cases} \sum_{k=1}^{n/2} \psi_{nm}^k(t), & n \text{ even} \\ \sum_{k=0}^{(n-1)/2} \psi_{nm}^k(t), & n \text{ odd}, \end{cases} \tag{14}$$

where

$$\psi_{nm}^0(t) = \beta_n^0 \frac{c}{R}\int_0^t e^{-\beta_n^0 c(t-t')/R}\phi_{nm}(t')dt', \tag{15a}$$

$$\psi_{nm}^k(t) = \frac{2c}{R}\int_0^t e^{-\beta_n^k c(t-t')/R}\{\beta_n^k \cos[\omega_n^k c(t-t')/R] + \omega_n^k \sin[\omega_n^k c(t-t')/R]\}\phi_{nm}(t')dt', \quad k \ge 1. \tag{15b}$$

We see that a modal RP response $\psi_{nm}(t)$ is, in general, a sum of partial responses $\psi_{nm}^k(t)$. For $n=0$, the modal RP response is zero for all time. For $n=1$ and $n=2$, there is only one partial response, which *is* the modal RP response. For $n=3$ and $n=4$, there are two partial responses for each modal RP

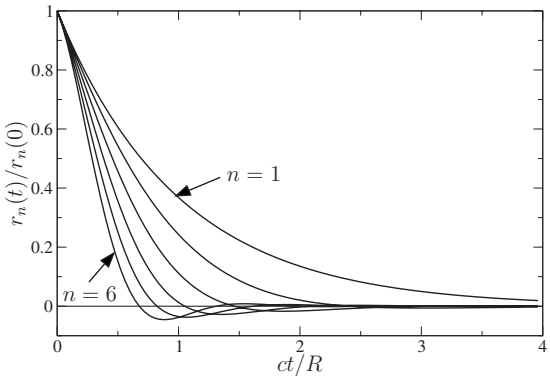


FIG. 2. Normalized residual functions for $n=1-6$ $[r_n(0)=n(n+1)/2]$.

response. In general, for $n=Q-1$ and $n=Q$ (Q even), there are $Q/2$ partial responses for each modal RP response.

From Eq. (15a), we write

$$\psi^0_{nm}(t+\Delta t) = e^{-\beta^0_n c\Delta t/R}\psi^0_{nm}(t) + \beta^0_n \frac{c}{R}\int_t^{t+\Delta t} e^{-\beta^0_n c(t+\Delta t-t')/R}\phi_{nm}(t')dt' \quad (16)$$

and then let $t'=t+\Delta t-\xi$ and $\Delta t=(R/c)\Delta\hat{t}$ to get the *exact stepping formula*

$$\psi^0_{nm}(t+\Delta t) = e^{-\beta^0_n\Delta\hat{t}}\psi^0_{nm}(t) + \beta^0_n\frac{c}{R}\int_0^{\Delta t} e^{-\beta^0_n c\xi/R} \times \phi_{nm}(t+\Delta t-\xi)d\xi, \quad (17)$$

which is local in time. For $k\geq 1$, a somewhat more complicated development (but only requiring exponential and trigonometric identities, plus Leibnitz' rule) produces from Eq. (15b) the exact stepping formula,

$$\psi^k_{nm}(t+\Delta t) = C^k_n\psi^k_{nm}(t) + S^k_n[\Delta t\dot{\psi}^k_{nm}(t) + \beta^k_n\Delta\hat{t}\psi^k_{nm}(t) - 2\beta^k_n\Delta\hat{t}\phi_{nm}(t)] + \frac{2c}{R}\int_0^{\Delta t} e^{-\beta^k_n c\xi/R}[\beta^k_n\cos(\omega^k_n c\xi/R) + \omega^k_n\sin(\omega^k_n c\xi/R)]\phi_{nm}(t+\Delta t-\xi)d\xi, \quad k\geq 1, \quad (18)$$

where $C^k_n=e^{-\beta^k_n\Delta\hat{t}}\cos(\omega^k_n\Delta\hat{t})$ and $S^k_n=e^{-\beta^k_n\Delta\hat{t}}\sin(\omega^k_n\Delta\hat{t})/(\omega^k_n\Delta\hat{t})$.

Although Eqs. (17) and (18) are exact, implementation is approximate, as only a finite number of modes can be considered, and numerical interpolation and differentiation are required. The last two are easily done by taking $\phi_{nm}(t+\Delta t-\xi)=\phi_{nm}(t+\Delta t)-[\phi_{nm}(t+\Delta t)-\phi_{nm}(t)]\xi/\Delta t+O(\Delta t^2)$ for $0\leq\xi\leq\Delta t$, and $\dot{\psi}^k_{nm}(t)=[\psi^k_{nm}(t+\Delta t)-\psi^k_{nm}(t-\Delta t)]/(2\Delta t)+O(\Delta t^2)$. Introducing these into Eqs. (17) and (18), we obtain the *numerical stepping formulas*,

$$\psi^0_{nm}(t+\Delta t) = e^{-\beta^0_n\Delta\hat{t}}\psi^0_{nm}(t) + \mathcal{I}^0_{Bn}\phi_{nm}(t) + (\mathcal{I}^0_{An}-\mathcal{I}^0_{Bn})\phi_{nm}(t+\Delta t) + O(\Delta t^3), \quad (19a)$$

TABLE I. Residual-function constants for $1\leq n\leq 8$.

n	k	β^k_n	ω^k_n
1	0	1.000 000 000 000 00	
2	1	1.500 000 000 000 00	0.866 025 403 784 439
3	0	2.322 185 354 626 09	
	1	1.838 907 322 686 96	1.754 380 959 783 72
4	1	2.896 210 602 820 36	0.867 234 128 934 47
	2	2.103 789 397 179 64	2.657 418 041 856 76
5	0	3.646 738 595 329 47	
	1	3.351 956 399 153 65	1.742 661 416 183 18
	2	2.324 674 303 181 63	3.571 022 920 337 96
6	1	4.248 359 395 863 47	0.867 509 673 231 54
	2	3.735 708 356 325 66	2.626 272 311 447 02
	3	2.515 932 247 810 88	4.492 672 953 653 97
7	0	4.971 786 858 532 45	
	1	4.758 290 528 151 33	1.739 286 061 128 90
	2	4.070 139 163 639 41	3.517 174 047 710 76
	3	2.685 676 878 943 10	5.420 694 130 716 49
8	1	5.587 886 043 263 42	0.867 614 445 359 12
	2	5.204 840 790 635 60	2.616 175 152 640 34
	3	4.368 289 217 203 13	4.414 442 500 471 21
	4	2.838 983 948 897 45	6.353 911 298 605 03

$$\psi^k_{nm}(t+\Delta t) = \left(1-\tfrac{1}{2}S^k_n\right)^{-1}\left[-\tfrac{1}{2}S^k_n\psi^k_{nm}(t-\Delta t) + (C^k_n + \beta^k_n\Delta\hat{t}S^k_n)\psi^k_{nm}(t) + 2(\mathcal{I}^k_{Bn}-\beta^k_n\Delta\hat{t}S^k_n)\phi_{nm}(t) + 2(\mathcal{I}^k_{An}-\mathcal{I}^k_{Bn})\phi_{nm}(t+\Delta t)\right] + O(\Delta t^3) \quad \text{for } k\geq 1, \quad (19b)$$

where

$$\mathcal{I}^0_{An} = \beta^0_n\int_0^{\Delta\hat{t}} e^{-\beta^0_n\xi}d\xi, \quad (20a)$$

$$\mathcal{I}^0_{Bn} = \frac{\beta^0_n}{\Delta\hat{t}}\int_0^{\Delta\hat{t}}\xi e^{-\beta^0_n\xi}d\xi, \quad (20b)$$

$$\mathcal{I}^k_{An} = \int_0^{\Delta\hat{t}} e^{-\beta^0_n\xi}[\beta^k_n\cos(\omega^k_n\xi) + \omega^k_n\sin(\omega^k_n\xi)]d\xi \quad \text{for } k\geq 1, \quad (20c)$$

$$\mathcal{I}^k_{Bn} = \frac{1}{\Delta\hat{t}}\int_0^{\Delta\hat{t}}\xi e^{-\beta^0_n\xi}[\beta^k_n\cos(\omega^k_n\xi) + \omega^k_n\sin(\omega^k_n\xi)]d\xi \quad \text{for } k\geq 1. \quad (20d)$$

These integrals are readily evaluated in closed form.

III. NUMERICAL IMPLEMENTATION

To examine the performance of the RP boundary, we introduced it into a spectral-finite-element code previously developed for analysis of the transient interaction between an elastic structure and a surrounding infinite acoustic fluid (Sprague and Geers, 2004, 2006). The code admits the possible occurrence of fluid cavitation, which is suppressed here. Our computational procedure exploits the fact that the modal superposition of Eq. (6) may be applied to both sides of Eq. (9) to obtain the point relation

$$u(\theta,\varphi,t)=\frac{1}{c}\dot{\phi}(\theta,\varphi,t)+\frac{1}{R}\phi(\theta,\varphi,t)+\frac{1}{R}\psi(\theta,\varphi,t). \quad (21)$$

With $u(\theta,\varphi,t)$, $\phi(\theta,\varphi,t)$, $\dot{\phi}(\theta,\varphi,t)$, $\ddot{\phi}(\theta,\varphi,t)$, and $\psi(\theta,\varphi,t)$ known at all surface nodes of the spectral-finite-element mesh, advancement from time t to time $t+\Delta t$ proceeds as follows.

(1) At all boundary nodes, calculate

$$\dot{\psi}^*(\theta,\varphi,t)=\frac{1}{\Delta t}[\psi(\theta,\varphi,t)-\psi(\theta,\varphi,t-\Delta t)], \quad (22)$$

$$\dot{u}^*(\theta,\varphi,t)=\frac{1}{c}\ddot{\phi}(\theta,\varphi,t)+\frac{1}{R}\dot{\phi}(\theta,\varphi,t)+\frac{1}{R}\dot{\psi}^*(\theta,\varphi,t). \quad (23)$$

(2) At all boundary nodes, compute predicted values of the radial fluid displacement with

$$u^*(\theta,\varphi,t+\Delta t)=u(\theta,\varphi,t)+\Delta t\,\dot{u}^*(\theta,\varphi,t). \quad (24)$$

(3) Input the results of Eq. (24) into the spectral-finite-element code to calculate all interior-domain nodal responses at time $t+\Delta t$, including nodal values of $\phi(\theta,\varphi,t+\Delta t)$, $\dot{\phi}(\theta,\varphi,t+\Delta t)$, and $\ddot{\phi}(\theta,\varphi,t+\Delta t)$ on the boundary.

(4) Compute the modal displacement potentials at time $t+\Delta t$ with

$$\phi_{nm}^{c/s}(t+\Delta t)=\frac{(2n+1)(n-m)!}{2(n+m)!}\int_0^{2\pi}d\varphi\int_0^{\pi}\phi(\theta,\varphi,t+\Delta t)\times P_n^m(\cos\theta)\begin{Bmatrix}\cos(m\varphi)\\ \sin(m\varphi)\end{Bmatrix}\sin\theta\,d\theta. \quad (25)$$

for $0\le m\le n$ and $0\le n\le N$; the integrals are evaluated with Gauss–Lobatto–Legendre quadrature consistent with that employed for the spectral finite elements.

(5) Calculate the partial RP responses at time $t+\Delta t$ with Eqs. (19).

(6) Compute the modal RP responses at time $t+\Delta t$ with Eq. (14).

(7) Calculate $\psi(\theta,\varphi,t+\Delta t)$ at all boundary nodes by modal superposition, as in Eq. (6).

(8) Determine corrected nodal values $u(\theta,\varphi,t+\Delta t)$ with Eq. (21).

IV. NUMERICAL RESULTS: STEP-WAVE-EXCITED SPHERICAL SHELL

We present here results for the axisymmetric ($m=0$) response of a submerged, empty spherical shell subjected to a plane step wave (Huang, 1969; Zhang and Geers, 1993; Huang and Mair, 1996). System properties correspond to a steel shell in water. The ratio of shell thickness to shell radius is $h/a=0.01$, the ratio of shell to fluid density is $\rho_s/\rho=7.505$, the ratio of plate velocity to acoustic fluid velocity is $c_p/c=3.517$, in which $c_p=[E/\rho_s(1-\nu^2)]^{1/2}$; E is Young's Modulus, and ν is Poisson's ratio. A quarter-symmetry discrete model was employed in which the structure was composed of 150 bilinear shell elements surrounded by a base mesh of hexagonal spectral-finite-elements. Figure 3 shows a representative mesh with the computational boundary located at $R=5a$. All fluid elements employed 12th-order basis functions (i.e., each element was associated with $13^3=2197$ nodes) and non-dimensional time increments of 2.25×10^{-3} for the structure and 2.50×10^{-4} for the fluid were used.

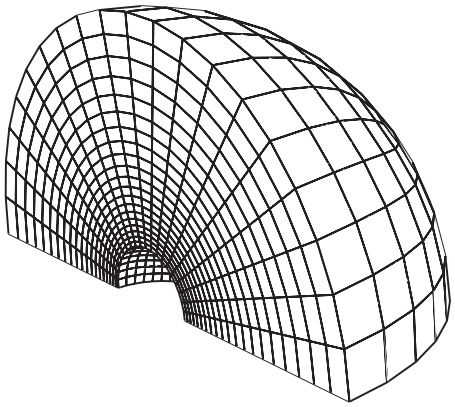


FIG. 3. Representative spectral-finite-element model used for the fluid surrounding the spherical shell; for the model shown, the computational boundary is located at $R=5a$.

Figure 4 shows radial-velocity histories for $0\le ct/a\le8$ at the front ($\theta=0°$) and back ($\theta=180°$) of the shell calculated with CWA (ETA2) boundaries at $R=1.08a$ (one layer of fluid elements), $R=3a$, and $R=5a$. For $R=5a$, reflections from the boundary are not felt by the shell over the duration $0\le ct/a\le8$, so the histories for $R=5a$ may be considered benchmarks. Shell response for $R=3a$ is affected by boundary reflections after $ct/a=4$, but the close agreement between the $R=3a$ histories and their benchmarks demonstrates that the CWA boundary performs well at $R=3a$. Unfortunately, the volume of fluid contained within the $R=3a$

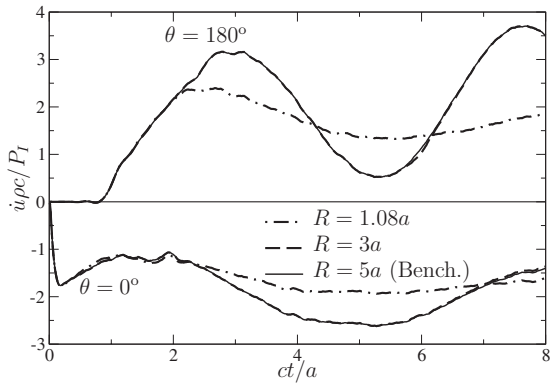


FIG. 4. Non-dimensional radial-velocity histories at the front and back of the shell calculated with a CWA boundary for three values of R.

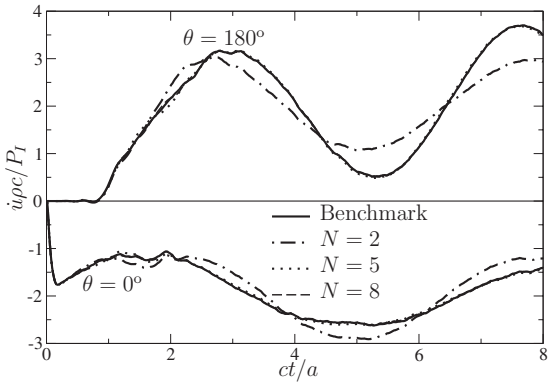


FIG. 5. Non-dimensional radial-velocity histories at the front and back of the shell calculated with the RP boundary located at $R=1.08a$ for three values of N, compared with their benchmark histories.

boundary is 26 times the volume contained within the shell. For $R=1.08a$, this ratio drops to 0.26, but CWA-boundary performance at this location is poor after $ct/a=2$.

Shown in Fig. 5 are the benchmark histories in Fig. 4 accompanied by velocity-response histories calculated with the RP boundary located at $R=1.08a$ for maximum n values $N=2$, 5, and 8. The $N=2$ histories are closer than the corresponding CWA histories in Fig. 4 to their benchmark histories, but are still not acceptable. However, the $N=5$ histories are nearly coincident with their benchmarks, and the $N=8$ histories are indistinguishable from their benchmarks.

Figure 6 shows the normalized root-mean-square (RMS) error,

$$\varepsilon_{\mathrm{RMS}} = \left\{ \frac{\int_0^{8a/c} [\dot{u}_N(\theta,t) - \dot{u}_B(\theta,t)]^2 dt}{\int_0^{8a/c} [\dot{u}_B(\theta,t)]^2 dt} \right\}^{1/2}, \tag{26}$$

where $\dot{u}_N(\theta,t)$ is a velocity history calculated with a RP boundary at $R=1.08a$ for $n_{\max}=N$ and $\dot{u}_B(\theta,t)$ is its benchmark history. Convergence is seen to be slow between $N=0$ and $N=3$, rapid between $N=3$ and $N=6$, and slow again after that. From Fig. 5, RMS errors below 0.01 indicate that RP histories are indistinguishable from their benchmark histories. This point is reached at $N=7$.

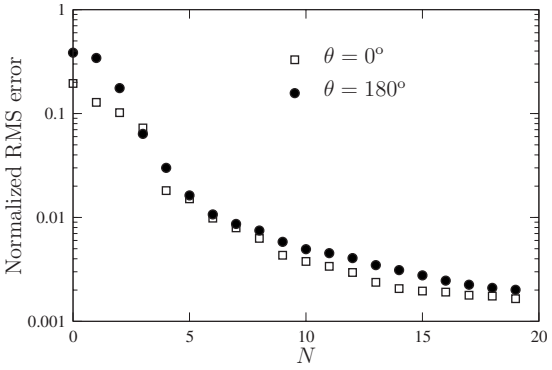


FIG. 6. RMS error of radial-velocity histories calculated for $0 \le ct/a \le 8$ normalized to the respective RMS values of their benchmark histories.

V. CONCLUSION

We have formulated and evaluated a spherical, theoretically exact, computational boundary for transient acoustics based on the residual-potential method, which was introduced by the first author some 40 years ago. The boundary is (1) geometrically local for a compact body, (2) spatially nonlocal but computationally efficient, (3) temporally local, and (4) easy to implement. The uncoupled equations (19) remain of low order for arbitrarily large values of the modal indices, thereby avoiding ill-conditioning issues. The principal limitation of the boundary is that it loses geometric locality for non-compact bodies. This may be alleviated by developing residual-potential boundaries for other separable coordinate systems, such as the spheroidal and ellipsoidal systems. The feasibility of such an endeavor derives from the benign characteristics of residual-potential functions.

Akkaş, N. (**1985**). "The residual variable method and its applications," Acta Mech. **55**, 203–217.

Akkaş, N., and Bauld, R. (**1981**). "Residual concentration method for the problem of diffusion on the surface of a spherical cavity," Int. J. Eng. Sci. **19**, 1341–1351.

Akkaş, N., and Engin, A. E. (**1980**). "Transient response of a spherical shell in an acoustic medium—Comparison of exact and approximate solutions," J. Sound Vib. **73**, 447–460.

Alpert, B., Greengard, L., and Hagstrom, T. (**2000**). "Rapid evaluation of nonreflecting boundary kernels for time-domain wave propagation," SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **37**, 1138–1164.

Alpert, B., Greengard, L., and Hagstrom, T. (**2002**). "Nonreflecting boundary conditions for the time-dependent wave equation," J. Comput. Phys. **180**, 270–296.

Astley, R. J. (**1996**). "Transient wave envelope elements for wave problems," J. Sound Vib. **192**, 245–261.

Astley, R. J. (**2000**). "Infinite elements for wave problems: A review of current formulations and an assessment of accuracy," Int. J. Numer. Methods Eng. **49**, 951–976.

Baker, B. B., and Copson, E. T. (**1939**). *The Mathematical Theory of Huygens Principle* (Oxford University Press, Oxford).

Bayliss, A., and Turkel, E. (**1980**). "Radiation boundary conditions for wave-like equations," Commun. Pure Appl. Math. **33**, 707–725.

Berenger, J. P. (**1994**). "A perfectly matched layer for the absorption of electromagnetic waves," J. Comput. Phys. **114**, 185–200.

Bettess, P. (**1992**). *Infinite Elements* (Penshaw, Sunderland).

Engquist B., and Majda, A. (**1977**). "Absorbing boundary conditions for the numerical simulation of waves," Math. Comput. **31**, 629–651.

Felippa, C. A. (**1980**). "A family of early-time approximations for fluid-structure interaction," J. Appl. Mech. **47**, 703–708.

Geers, T. L. (**1969**). "Excitation of an elastic cylindrical shell by a transient acoustic wave," J. Appl. Mech. **36**, 459–469.

Geers, T. L. (**1971**). "Residual potential and approximate methods for three-dimensional fluid-structure interaction," J. Acoust. Soc. Am. **49**, 1505–1510.

Geers, T. L. (**1972**). "Scattering of a transient acoustic wave by an elastic cylindrical shell," J. Acoust. Soc. Am. **37**, 1091–1106.

Geers, T. L. (**1978**). "Doubly asymptotic approximations for transient motions of submerged structures," J. Acoust. Soc. Am. **64**, 1500–1508.

Geers, T. L. (**1991**). "A fully consistent formulation of early-time approximations for acoustic media," in *The Finite Element Method in the 1990's*, edited by E. Oñate, J. Periaux, and A. Samuelsson (Springer-Verlag, Berlin), pp. 521–528.

Geers, T. L., ed. (**1998**). *IUTAM Symposium on Computational Methods for Unbounded Domains* (Kluwer, Dordrecht).

Geers, T. L., and Toothaker, B. J. (**2000**). "Third-order doubly asymptotic approximations for computational acoustics," J. Comput. Acoust. **8**, 101–120.

Geers, T. L., and Yen, C. L. (**1981**). "Nonlinear response of an elastic cylindrical shell to a transient acoustic wave," J. Appl. Mech. **48**, 15–24.

Geers, T. L., and Yen, C. L. (**1989**). "Inelastic response of an infinite cylindrical shell to transient acoustic waves," J. Appl. Mech. **56**, 900–909.

Geers, T. L., and Zhang, P. (**1994**). "Doubly asymptotic approximations for submerged structures with internal fluid volumes," J. Appl. Mech. **61**, 893–906.
Givoli, D. (**1992**). "A spatially exact non-reflecting boundary condition for time dependent problems," Comput. Methods Appl. Mech. Eng. **95**, 97–113.
Givoli, D. (**2004**). "High-order local non-reflecting boundary conditions: A review," Wave Motion **39**, 319–326.
Givoli, D., and Harari, I. (**1998**). "Exterior problems of wave propagation (special issue)," Comput. Methods Appl. Mech. Eng. **164**, 1–2.
Grote, M. J., and Keller, J. B. (**1995a**). "Exact nonreflecting boundary conditions for the time dependent wave equation," SIAM J. Appl. Math. **55**, 280–297.
Grote, M. J., and Keller, J. B. (**1995b**). "On nonreflecting boundary conditions," J. Comput. Phys. **122**, 231–243.
Grote, M. J., and Keller, J. B. (**1996**). "Boundary conditions for time-dependent scattering," J. Comput. Phys. **127**, 52–65.
Hagstrom, T., Mar-Or, A., and Givoli, D. (**2008**). "High-order local absorbing conditions for the wave equation: Extensions and improvements," J. Comput. Phys. **227**, 3322–3357.
Huan, R., and Thompson, L. L. (**2000**). "Accurate radiation boundary conditions for the time dependent wave equation on unbounded domains," Int. J. Numer. Methods Eng. **47**, 1569–1603.
Huang, H. (**1969**). "Transient interaction of plane acoustic waves with a spherical elastic shell," J. Acoust. Soc. Am. **45**, 661–670.
Huang, H., and Mair, H. U. (**1996**). "Neo-classical solution of transient interaction of plane acoustic waves with a spherical elastic shell," Shock Vib. **3**(2), 85–89.
Marston, P. L. (**1997**). "Quantitative ray methods for scattering," in *Encyclopedia of Acoustics*, edited by M. J. Crocker (Wiley, New York), Chap. 43, pp. 483–492.
Morse, P. M., and Ingard, K. U. (**1968**). *Theoretical Acoustics* (McGraw-Hill, Princeton, NJ).
Nicolas-Vullierme, B. (**1991**). "A contribution to doubly asymptotic approximations: An operator top-down derivation," J. Vib. Acoust. **113**, 409–415.
Qi, Q., and Geers, T. L. (**1998**). "Evaluation of the perfectly matched layer for computational acoustics," J. Comput. Phys. **139**, 166–183.
Sprague, M. A., and Geers, T. L. (**2004**). "A spectral-element method for modeling cavitation in transient fluid-structure interaction," Int. J. Numer. Methods Eng. **60**, 2467–2499.
Sprague, M. A., and Geers, T. L. (**2006**). "A spectral/finite-element analysis of a ship-like structure subjected to an underwater explosion," Comput. Methods Appl. Mech. Eng. **195**, 2149–2167.
Struik, J. D. (**1961**). *Lectures on Classical Differential Geometry*, 2nd ed. (Addison-Wesley, Reading, MA).
Thompson, L. L., and Huan, R. (**2000**). "Implementation of exact non-reflecting boundary conditions in the finite element method for the time-dependent wave equation," Comput. Methods Appl. Mech. Eng. **187**, 137–159.
Tsynkov, S. V. (**1998**). "Numerical solution of problems on unbounded domains. A review," Appl. Numer. Math. **27**, 465–532.
Turkel, E. (**1998**). "Introduction to the special issue on absorbing boundary conditions," Appl. Numer. Math. **27**, 327–329.
Wolf, J. P., and Song, C. (**1996**). *Finite-Element Modelling of Unbounded Media* (Wiley, West Sussex).
Zhang, P., and Geers, T. L. (**1993**). "Excitation of a fluid-filled, submerged spherical shell by a transient acoustic wave," J. Acoust. Soc. Am. **93**, 696–705.

Nonlinear reflection of shock shear waves in soft elastic media

Gianmarco Pinton
Laboratoire Ondes et Acoustique, Institut Langevin, ESPCI ParisTech, CNRS UMR 7587, INSERM ERL U979, 10 rue Vauquelin, 75231 Paris Cedex 05, France

François Coulouvrat[a)]
Institut Jean Le Rond d'Alembert, Université Pierre et Marie Curie-Paris 6, CNRS UMR 7190, 4 place Jussieu, 75252 Paris Cedex 05, France

Jean-Luc Gennisson and Mickaël Tanter
Laboratoire Ondes et Acoustique, Institut Langevin, ESPCI ParisTech, CNRS UMR 7587, INSERM ERL U979, 10 rue Vauquelin, 75231 Paris Cedex 05, France

(Received 13 July 2009; revised 14 November 2009; accepted 30 November 2009)

For fluids, the theoretical investigation of shock wave reflection has a good agreement with experiments when the incident shock Mach number is large. But when it is small, theory predicts that Mach reflections are physically unrealistic, which contradicts experimental evidence. This von Neumann paradox is investigated for shear shock waves in soft elastic solids with theory and simulations. The nonlinear elastic wave equation is approximated by a paraxial wave equation with a cubic nonlinear term. This equation is solved numerically with finite differences and the Godunov scheme. Three reflection regimes are observed. Theory is developed for shock propagation by applying the Rankine–Hugoniot relations and entropic constraints. A characteristic parameter relating diffraction and non-linearity is introduced and its theoretical values are shown to match numerical observations. The numerical solution is then applied to von Neumann reflection, where curved reflected and Mach shocks are observed. Finally, the case of weak von Neumann reflection, where there is no reflected shock, is examined. The smooth but non-monotonic transition between these three reflection regimes, from linear Snell–Descartes to perfect grazing case, provides a solution to the *acoustical* von Neumann paradox for the shear wave equation. This transition is similar to the quadratic non-linearity in fluids.
 [DOI: 10.1121/1.3277202]

PACS number(s): 43.25.Cb, 43.25.Dc, 43.25.Jh [OAS] Pages: 683–691

I. INTRODUCTION

Finite-amplitude waves propagate at a speed that depends on the instantaneous amplitude, as opposed to the constant speed of linear waves. For sound waves in fluids, the modification of the sound speed is proportional to the wave amplitude.[1] Since fluids have a small Mach number, the nonlinear effects are locally small but they can accumulate over long distances, which eventually leads to the formation of shock waves. Nevertheless, local nonlinear phenomena may occur for shock wave fronts propagating in almost parallel directions. This behavior can be observed in fold[2] or cusp caustics[3] or irregular reflection of grazing shocks.[4] In this last case, the nonlinear interaction of the incident and reflected shocks give rise to various regimes of reflection patterns, which can be separated into two types: regular and irregular reflections.[5] For regular reflection, at a sufficiently large grazing angle or for a sufficiently weak shock, the incident and reflected shocks intersect right on the rigid surface in a way similar to linear Snell–Descartes laws. However, the reflection angle may be smaller than the incident one, and the reflected amplitude larger than the incident one. As the grazing angle decreases, or the shock amplitude increases, the point of intersection of the incident and reflected shocks detaches from the surface and a new shock emerges to ensure contact with the surface. This three shock pattern, called a *Mach shock* or *Mach stem*, is known as *irregular reflection*. For a strong incident shock, this is known as *Mach reflection* and is named after its first experimental observation by Ernst Mach[6] (see also Ref. 7). Theoretical investigation of shock wave reflection was first carried out by John von Neumann[8] for inviscid perfect gases, with a transition criterion between the two types of reflection.[9] However, the three-shock theory has a good agreement with experiments only when the incident shock Mach number is greater than 1.47. For weaker shocks, the three-shock theory leads to the conclusion that the Mach reflection is physically unrealistic. This runs contrary to experimental evidence supported by numerical simulations based on the inviscid Euler equations, which shows that Mach reflection is indeed possible for such weak shocks. The discrepancy between three-shock theory and experimental studies has been referred to as the *von Neumann paradox* and the corresponding observed regime of reflection is nowadays referred to as von Neumann type. It is characterized by a continuous slope of the shock front along the incident and the Mach shock.[10] Various theories and observations have been proposed to elucidate, in particular, the local behavior

[a)]Author to whom correspondence should be addressed. Electronic mail: francois.coulouvrat@upmc.fr

at the triple point.[7,11–13] Extremely weak shock waves exist in the acoustical regime corresponding to Mach numbers not much larger than 1.001. Nonlinear regular and irregular reflection patterns have been observed experimentally[14] and investigated numerically[4] for various profiles of acoustical shock waves, such as ideal step shocks, N-waves, or periodic saw-tooth waves. In particular, the smooth but nonlinear and nonmonotonic transition between Snell–Descartes laws and perfectly grazing incidence has been described, thus elucidating the so-called *acoustical paradox* (another paradox also proposed by von Neumann).

Acoustical shock waves are limited to neither fluids nor compression waves. Transverse shear shock waves have been observed in soft elastic media,[15,16] quite some time after theoretical predictions.[17–19] Conventional weak shock theory can be extended to the resulting modified. Burgers equation, but contrary to fluid expansion shocks are permitted.[20] Indeed, nonlinear shear effects are one order of magnitude smaller than longitudinal ones in crystals or metals. However, the acoustical Mach number, and hence nonlinear effects, can be several orders of magnitude larger in soft solids such as biological media (muscle, fat, breast, etc.) or phantoms (agar or gelatin) because low frequency (on the order of 100 Hz) shear (SH) waves have a very low propagation speed (on the order of 1 m/s). In biological media viscosity plays an important role because as the wave propagates it preferentially attenuates high frequencies, thus affecting shock properties such as the rise time. However, if measurements are made sufficiently for a sufficiently large amplitude source, shock behavior can be observed at least up to the seventh harmonic.[15] Moreover, experiments in fluids in the megahertz regime[14] indicate that nonlinear reflection phenomena are insensitive to thermoviscous absorption. The symmetric properties of isotropic media lead to the cancellation of quadratic nonlinear effects, so that leading terms are cubic. Note that this cancellation is not generalizable to other polarizations, as only SH polarization does not lead to the interaction with compression waves at the interface.[21] Reformulated elastic theories[22,23] propose a new expression for the coefficient of the cubic nonlinearities in terms of three (one, respectively, at second, third, and fourth orders) elastic coefficients. Recent experiments[24,25] have measured these coefficients.

II. GENERALIZED KZ EQUATION FOR TWO-DIMENSIONAL NONLINEAR SHEAR WAVES

We consider a nonlinear SH wave with $u(x_1,x_2,T)$ the horizontal displacement in the direction Ox_3 propagating in the horizontal plane (Ox_1x_2) in a homogeneous and isotropic solid. In this case, a reflected wave will have the same polarization as the incident shear wave so that there is no coupling with compression waves and both incidence and reflection can be described by a scalar wave equation. The equation of motion in Lagrangian coordinates is

$$p_0\frac{\partial^2 u}{\partial T^2}=\frac{\partial \sigma_{3j}}{\partial x_j}, \tag{1}$$

where ρ_0 is the density of the undeformed material and σ_{ij} is the stress tensor. We use here the expression of the stress tensor obtained in Ref. 23 [Eq. (10)], and based on Landau and Lifshitz' choice of invariants for the strain energy density[21] that allows a separation between shear and compression. For nonlinear SH waves, this reduces $(j=1,2)$ to

$$\sigma_{3j}=\frac{\partial u}{\partial x_j}\left[\mu+\left(\mu+\frac{A}{2}+D\right)\left(\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right)\right]. \tag{2}$$

Here, μ is the shear modulus associated with the quadratic terms in the expansion of the strain energy (linear motion), while coefficients A and D are, respectively, associated with the cubic and fourth-order terms of the energy expansion. Because of the medium isotropy, quadratic nonlinear terms vanish by symmetry, and the leading nonlinear terms are cubic. Substituting Eq. (2) into the equation of motion, Eq. (1), yields the nonlinear wave equation for SH waves:

$$\begin{aligned}\rho_0\frac{\partial^2 u}{\partial T^2}-\mu\left(\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}\right)&=\left(\mu+\frac{1}{2}A+D\right)\left[\frac{\partial}{x_1}\left(\left[\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]\frac{\partial u}{\partial x_1}\right)\right.\\&\left.+\frac{\partial}{x_2}\left(\left[\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]\frac{\partial u}{\partial x_2}\right)\right].\end{aligned} \tag{3}$$

A more tractable equation can be obtained by making a paraxial approximation. We assume that the wave amplitude is small (as measured by the small parameter $\epsilon\ll 1$), and that it propagates primarily in the Ox_1 direction. By omitting the first order nonlinear terms [on the right-hand side of Eq. (4)] and diffraction terms (that depend on x_2), we obtain the usual one-dimensional (1D) wave equation, for which a traveling wave solution is $u=\epsilon U(T-x/c_T)$ and where $c_T=\sqrt{\mu/\rho_0}$ is the linear shear wave velocity. A paraxial approximation, valid for waves propagating in directions at a small angle from the Ox_1 axis, is obtained by searching for approximate solutions of the form

$$u(x_1,x_2,T)=\epsilon U(\tau=T-x_1/c_T,X=\epsilon^2 x_1,Y=\epsilon x_2). \tag{4}$$

By inserting the above form of the solution into the full nonlinear wave equation, Eq. (4), the zeroth order terms (ϵ^0) vanish, and by equating leading order terms (of order ϵ^2), one obtains[21]

$$\frac{\partial^2 U}{\partial X\partial\tau}=\frac{c_T}{2}\frac{\partial^2 U}{\partial Y^2}+\frac{1}{2c_T^3}\left(1+\frac{A/2+D}{\mu}\right)\frac{\partial}{\partial\tau}\left[\left(\frac{\partial U}{\partial\tau}\right)^3\right], \tag{5}$$

with the boundary condition

$$\sigma_{32}=0\Leftrightarrow\frac{\partial U}{\partial Y}=0. \tag{6}$$

Introducing the medium velocity $V=\partial U/\partial\tau$ instead of the displacement, the above equation can be equivalently set to

$$\frac{\partial^2 V}{\partial X \partial \tau} = \frac{c_T}{2}\frac{\partial^2 V}{\partial Y^2} + \frac{\beta}{c_T^3}\frac{\partial}{\partial \tau}\left[V^2 \frac{\partial V}{\partial \tau}\right], \tag{7}$$

where

$$\beta = \frac{3}{2}\left(1 + \frac{\frac{A}{2} + D}{\mu}\right) \tag{8}$$

is the coefficient of non-linearity, involving a combination of the three coefficients of the elastic energy expansion. In the one-dimensional case, the above equation reduces to the usual nonlinear convection equation with cubic nonlinearities (see introduction for a review of the topic). In the paraxial case, Eq. (7) is similar to the KZ equation for compressional waves,[19,26] but with quadratic nonlinearities replaced by cubic ones.

That equation is nondimensionalized by introducing the following scaling:

$$V = V_0 v, \quad t = \omega_0 \tau, \quad X = Lx = \frac{6}{\beta k_T M^2} x,$$

$$Y = Hy = \sqrt{\frac{3}{\beta k_T^2 M^2}}\, y, \tag{9}$$

where V_0 is the velocity wave amplitude, ω_0 is some reference angular frequency, $k_T = \omega_0 / c_T$ is the corresponding shear wave number, $M = V_0 / c_T$ is the shear Mach number, and L is the shock formation distance of a sine wave of amplitude V_0 and angular frequency ω_0. This distance is dependent on the square of the amplitude because nonlinearities here are cubic. In the absence of any physical transverse characteristic distance, H is chosen so as to equate all other coefficients in the dimensionless equation, so that

$$\frac{\partial^2 v}{\partial x \partial t} = \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v^3}{\partial t^2}. \tag{10}$$

Note that here the symbol t refers as the dimensionless retarded time, which is different from the physical time, T, used only in Eqs. (1) and (3).

III. SHOCK CONDITIONS

An equivalent conservative form of the cubic KZ equation, Eq. (10), can be written by introducing $q = \partial u / \partial y$, the linear and dimensionless part of the stress tensor σ_{32}. Then, one obtains an equivalent first-order system for the coupled variables (q, v) evolving in the plane (y, t) with variable x:

$$0 = \frac{\partial v}{\partial y} - \frac{\partial q}{\partial t},$$

$$\frac{\partial v}{\partial x} = \frac{\partial q}{\partial y} + \frac{\partial v^3}{\partial t}. \tag{11}$$

With this form, the associated Rankine–Hugoniot shock relations can be easily deduced, using the same methods as for the quadratic KZ equation.[27] The result is identical, except that the quadratic terms v^2 have to be replaced by cubic ones v^3:

$$0 = N_y[v] - N_t[q],$$

$$-W[v] = N_y[q] + N_t[v^3], \tag{12}$$

where $[q]$ and $[v]$ are, respectively, the jumps of quantities q and v through the shock, $\mathbf{N} = (N_y, N_t)$ is the unit vector normal to the shock front, and W is the front shock normal speed. Here, the shock front is a surface in the (y, t) plane moving with variable x which plays the role of time. Eliminating the jump $[q]$ between the two, another relation can be obtained for the velocity jump through the shock

$$-W[v]N_t = N_y^2[v] + N_t^2[v^3]. \tag{13}$$

Eliminating the continuous case $[v]=0$ by factoring Eq. (13), yields

$$-WN_t = N_y^2 + N_t^2 s(v), \tag{14}$$

where $s(v) = v_+^2 + v_+ v_- + v_-^2$ with v_- (respectively, v_+) is the velocity amplitude just before (respectively, after) the shock. In the one-dimensional case, with no dependence on y, the above equation reduces to $W = dt_s/dx = -s(v)$, relating the inverse of the shock velocity occurring at time t_s to the left- and right-hand side values of the shock. Note that the above relations are not sufficient enough to fully solve the problem, and that they must be complemented by an entropy condition through the shock. For quadratic nonlinearities, the entropy condition implies that the shock is compressive ($v_+ > v_-$). On the contrary, for cubic nonlinearities, the shock can be either in compression or expansion, and the entropy condition is $|v_+| > |v_-|$[6], admissible shocks having an absolute value after the shock that is larger than before the shock. A particular solution for the 1D case (for which $N_t = 1$, $N_y = 0$, $W = W_t = dt_s/dx$) is the step shock of amplitude 1 at time $t_s(x)$:

$$v(x,t) = \begin{cases} v_- = 0 & \text{if } t < t_s(x) \\ v_+ = \pm 1 & \text{if } t > t_s(x). \end{cases} \tag{15}$$

According to Eq. (14) and assuming that the shock occurs at time 0 for the initial position $x=0$, one necessarily has $t_s(x) = -x$. For the particular solution described by Eq. (15), the shock can be either in compression (+ sign) or expansion (− sign). The two solutions are equivalent due to the invariance of the nonlinear wave equation, Eq. (10), and jump relations Eq. (13) with the change of sign $(v, q) \to (-v, -q)$.

IV. REGULAR REFLECTION OF SHOCK SHEAR WAVES

A. The control parameter

We now consider a plane, perfect step-shock wave impinging on a semi-infinite free surface located at the position $x_2 = 0$ with a grazing angle θ (Fig. 1). We denote by x_1' the dimensional spatial coordinate in the propagation direction of this incident shock. One clearly has $x_1' = \cos\theta x_1 - \sin\theta x_2$, so that $x' = x_1'/L = \cos\theta x - \sin\theta (H/L) y$. We can now compute the retarded, dimensionless time associated with this plane shock:

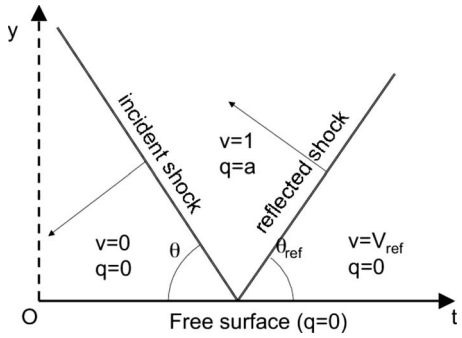


FIG. 1. Geometry of regular reflection of shear shock waves.

$$t' = \omega_0(t - x_1'/c_T)$$
$$= \omega_0 t - k_T L x + k_T H \sin\theta y + k_T L(1-\cos\theta)x. \quad (16)$$

This coordinate rotation by an angle θ introduces the control parameter:

$$a = k_T H \sin\theta = \sqrt{\frac{3}{\beta}}\frac{\sin\theta}{M}, \quad (17)$$

which is a measurement of the diffraction effects ($\sin\theta$) relative to the nonlinear ones (M). Note that for a given value of a the values of β and M must fall within a range that allows θ to respect the conditions set by the paraxial approximation. Typically, the paraxial approximation is valid for angles less than 15°, corresponding here to an acoustical Mach number M of about 0.26, which is large. When $a \gg 1$, diffraction is dominant and nonlinear effects are negligible. This happens far from the grazing case. We expect this limit case to satisfy the usual linear Snell–Descartes laws of reflection. Indeed, the directions of the incident and reflected shocks differ sufficiently so that no cumulative nonlinear interactions between the two occur. On the contrary, when $a = O(1)$, these directions are sufficiently close to one another so that nonlinear cumulative interactions between incident and reflected waves can occur. This is the regime we are currently studying. The limit case $a \to 0$ corresponds to the perfectly grazing case, for which the incident shock plane travels parallel to the interface so that it does not interact with it. By making this parameter evolve from 0 to large values, we can characterize the various regimes of reflection and quantify the continuous transition, thus elucidating the von Neumann *acoustical paradox* in the particular case of shear shock waves. The last term in Eq. (16) involves the parameter

$$k_T L(1-\cos\theta) = 0.5 k_T L \sin^2\theta + O(k_T L \sin^4\theta)$$
$$= a^2 + O(a^2 \sin\theta^2) = a^2 + O(a^2 M^2), \quad (18)$$

the last identity coming from the assumption $a = O(1)$. The last term can therefore be neglected because it involves higher-order nonlinear terms that are neglected. Also note that this is consistent with the paraxial approximation, which assumes that propagation is in some principal direction (here x) and for which $\sin\theta = O(M)$ has to be small. Hence one has the important formula:

$$t' = t + ay + a^2 x, \quad (19)$$

relating the dimensionless retarded time t' in the direction of the incident shock to the coordinate system (t,x,y) adapted to the interface geometry. Following this reasoning one can compute

$$x' = x - \sin\theta \frac{H}{L} y - (1-\cos\theta)x$$
$$= x - O(\sin\theta M)y - O(\sin^2\theta)x = x + O(M^2). \quad (20)$$

B. Self-similar solutions

The theoretical problem to be solved is therefore the following. The total velocity field v satisfies the cubic KZ equation, Eq. (7), in the domain $[x \geqslant 0;\ 0 \leqslant y < +\infty;\ -\infty < t < +\infty]$. On the free surface $y=0$, the normal stress component σ_{32}, Eq. (2), has to vanish, hence so does its time derivative also: $\partial v / \partial y(x, y=0, t) = 0$. Finally, far from the interface, the wave field can be separated into an incident shock v_{inc} and a reflected field v_{ref} which is unknown. The incident shock is given by Eq. (15) but is rewritten in terms of the oriented coordinates (t', x'). Returning to the original coordinates, the incident shock is thus given by

$$v_{\text{inc}}(x,y,t) = \begin{cases} 0 & \text{if } t + ay + a^2 x < -x \\ \pm 1 & \text{if } t + ay + a^2 x > -x. \end{cases} \quad (21)$$

Note that the case for an *expansion* shock ($-$ sign) can be deduced from the compression case ($+$ sign) by a $v \to -v$ symmetry, so that only the first will be considered. The shock takes place at time $t_s = -(1+a^2)x - ay$. The normal vector to the wave front in the (t,y) plane is therefore $\mathbf{N} = (N_y, N_t) = (a,1)/\sqrt{1+a^2}$. The normal stress q has to vanish before the shock ($t < t_s$) to satisfy the boundary condition. Using the shock relations, Eq. (13), one can deduce the value of the normal stress after the shock ($t > t_s$) $q = a$.

The parameters ω_0, L, and H used to make the problem dimensionless all involve the reference frequency ω_0, which measures the characteristic time variation of the wavefield. However, a step shock has no intrinsic characteristic time. Hence, the problem should not involve such a parameter. Therefore, a self-similar solution involving a combination of variables that suppress this dependence with ω_0 can be sought. A simple verification shows that the convenient choice is to seek the solution of the form

$$v(x,y,t) = v\left(\frac{t}{x}, \frac{y}{x}\right). \quad (22)$$

Note that the control parameter a as defined by Eq. (17) does not depend on ω_0 and hence is fully relevant, even in the case of a step shock. The self-similar solution in Eq. (22) shows that the combined nonlinear and diffraction effects will take place in a diffraction boundary layer along the interface. The thickness of this boundary layer increases linearly along the reflective plate with the variable x. Of course, the solution is singular right at the tip of the plate ($x \to 0$), because the paraxial approximation is not be valid locally. Note that self-similar solutions for step shocks also exist for various nonlinear diffraction problems in fluids.[4,27,2]

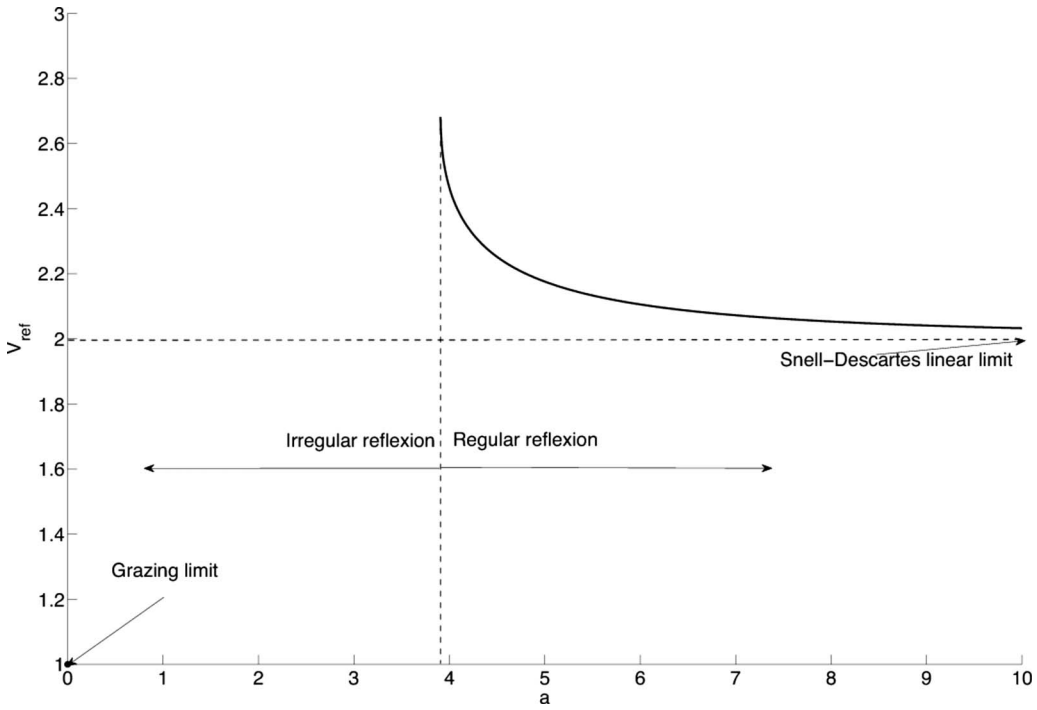


FIG. 2. Nonlinear regular reflection of shear shock waves.

C. Regular reflection and von Neumann paradox

An exact solution can be found for the case where the reflected field is also a straight step shock. This is the so-called *regular* regime or two-shock regime. We assume the angle of reflection (relative to the interface) is equal to $\theta_{\text{ref}} \in]0, \pi/2[$ (Fig. 1). By introducing the related parameter $b = \sqrt{3/\beta}\sin\theta_{\text{ref}}/M > 0$, which measures this angle in dimensionless variables in a way similar to the control parameter a, Eq. (17), the shock position necessarily has the form $t_s^{\text{ref}} = f(x) + by$. This is a form similar to the incident one, but with a change of sign in front of the y term as the wave front is now going *away* from the interface. The function $f(x)$ has to match the incident shock at the interface $y=0$ and hence has to be $f(x) = -(1+a^2)x$. The vector normal to the reflected wave front in the (t,y) plane is therefore $\mathbf{N}_{\text{ref}} = (N_y^{\text{ref}}, N_t^{\text{ref}}) = (-b, 1)/\sqrt{1+b^2}$. The velocity field behind the reflected shock is assumed to be of constant amplitude V_{ref} so that the reflected shock amplitude is $[v]_{\text{ref}} = V_{\text{ref}} - 1$. To satisfy the boundary condition, the quantity q has to be equal to 0 after the reflected shock so that $[q]_{\text{ref}} = -a$ (Fig. 1). The wave front velocity is uniform along the shock so that it is equal to the one at the free surface $\mathbf{W}_{\text{ref}} = (W_y^{\text{ref}}, W_t^{\text{ref}}) = (0, -(1+a^2))$. Now satisfying the Rankine–Hugoniot equations through the shock yields, for the Eq. (13) $V_{\text{ref}} = 1 + a/b$ and for Eq. (14) $V_{\text{ref}}^2 + V_{\text{ref}} + b^2 - a^2 = 0$. Note that the entropy condition will be satisfied for the reflected shock as $|V_{\text{ref}}| > 1$ because a and b are positive. Substituting the value $V_{\text{ref}} = 1 + a/b$, one gets the following equation relating the dimensionless reflected angle b to the incident one a:

$$(b+a)(b^3 - ab^2 + 2b + a) = 0. \tag{23}$$

The solution $b = -a < 0$ is impossible. Hence b has to be a real and positive root of the third degree polynomial $p(x) = x^3 - ax^2 + 2x + a$. An analysis of the roots of this polynomial (see Appendix) shows that it has two positive roots if $a > a_{\text{cr}} \simeq 3.908\,167\,555$, and no positive roots if $a < a_{\text{cr}}$. The regime $a > a_{\text{cr}}$ is called the *regular* regime. In this case, the two-shock problem indeed has a unique solution. Indeed, among the two roots that exist, one converges for large values of a toward $b = a - 3/a + o(1/a)$ and one toward $b = 1 + 3/2a + o(1/a)$. In the limit $a \to +\infty$, we expect to recover the linear behavior, for which the reflected angle b is equal to the incident one a according to Snell–Descartes law. For this limit behavior, we also recover the fact that $V_{\text{ref}} = 1 + b/a \to 2$ (in the linear case, one has a field doubling at the interface or a reflection coefficient equal to +1). Hence, the admissible solution is the *maximum* positive root of the polynomial, when it exists. This so-called regular regime is similar to the linear reflection (two straight shocks connected at the free surface) but is nevertheless nonlinear in the sense that (i) the reflection angle b is smaller than the incident one a and consequently (ii) the amplitude of the field behind the reflected field $V_{\text{ref}} = 1 + a/b$ is larger than 2. This behavior is illustrated in Fig. 2 where the total field behind the shock is plotted as a function of the dimensionless incident angle a. Departure from Snell–Descartes law ($V_{\text{ref}} = 2$) is obvious when a is not very large. Note also that this two-shock solution is self-similar.

For values of the angle below the critical value a_{cr}, the two-shock regime does not exist anymore: this is the *irregular reflection*. As expected from the von Neumann paradox and quadratic case, a three-shock analytical solution is unlikely. Therefore, this regime will now be explored numerically.

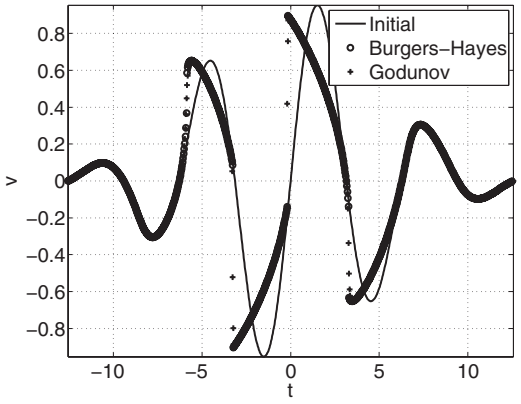


FIG. 3. Verification of Godunov's method for cubic non-linearity with a Burgers–Hayes' solution.

V. NUMERICAL OBSERVATION

The numerical solution of the parabolic wave equation with cubic nonlinearities, Eq. (10), was determined with operator splitting, which applies finite difference to each of the terms in the equation independently over a propagation step. It has been used successfully in the context of the nonlinear acoustical waves with quadratic nonlinearities.[28,29] Here it is applied to Eq. (10) and expressed in integral form

$$\frac{\partial v}{\partial x} = \mathcal{L}_D(v) + \mathcal{L}_N(v), \tag{24}$$

where

$$\mathcal{L}_D = \int_{-\infty}^{t} \frac{\partial^2 v}{\partial y^2} dt', \quad \mathcal{L}_N = \frac{\partial v^3}{\partial t}. \tag{25}$$

The operators $\mathcal{L}_D$ and $\mathcal{L}_N$ represent diffraction and nonlinearity, respectively. That equation is completed by boundary conditions applied over the numerical domain $[0 \leqslant y \leqslant y_{\max}] \times [t_{\min} \leqslant t \leqslant t_{\max}]$. On the free surface, one has $\partial v / \partial y(x, y=0, t)=0$. The size of the numerical domain is chosen so that the incident (respectively, the reflected) shocks do not intersect the upper boundary $y=y_{\max}$. As a consequence, one can impose $v(x, y, t=t_{\min})=0$, and $v(x, y_{\max}, t)=v_{\text{inc}}(x, y_{\max}, t)$. There is no boundary condition to write on $t=t_{\max}$. The initial solution at $x=0$ is the incident shock $v_{\text{inc}}(0, y, t)$, Eq. (21), before it interacts with the free surface.

To calculate the diffraction operator, a standard implicit second order finite difference scheme was applied to the partial derivative and the trapezoidal rule was used to calculate the integral. This results in an easily solvable tridiagonal system.[28] The equality composed of the left hand side of Eq. (24) together with the nonlinear operator $\mathcal{L}_N$ is a hyperbolic conservation law and it was solved with Godunov's scheme,[30] which generalizes upwind differencing for nonlinear problems:

$$v_i^{n+1} = v_i^n - \frac{\Delta x}{\Delta t}[f_{i+1/2}^{n+1/2} - f_{i-1/2}^{n+1/2}], \tag{26}$$

where $f_{i+1/2}^{n+1/2}$ is the discrete approximation of the flux, $f(v) = v^3$. The characteristics, as determined by the derivative of the flux $f'(v)=3v^2$, are all greater than or equal to zero so that the flux in Godunov's method can be written simply as

$$f_{i+1/2} = (v_i)^3. \tag{27}$$

This method has been amply characterized,[31] it is known to satisfy the Rankine–Hugoniot relation, Eq. (14), and it has the nice feature of converging to the correct solution whenever it converges.

Figure 3 compares the numerical solution of Godunov's method (shown with plus symbols) with an analytical solution based on a Burgers–Hayes method (shown with small circles). Note that here we are solving the nonlinear equation without diffraction. The Burgers–Hayes solution is based on a method originally proposed by Burgers. It combines Poisson's solution of the nonlinear equation written for displacement u rather than velocity v, with shock localization. It is evaluated in a single time step and the only approximation it makes is interpolation to regain the appropriate grid spacing.[20] The initial waveform (shown as a solid line) was propagated a distance of $x=2$ with a propagation step of $\Delta x=0.0036$ and a time step of $\Delta t=0.0500$. Both positive and negative shocks are correctly characterized by the numerical method. At this mesh size, numerical diffusion introduces approximately two points within the shock wavefront.

In general there are no known analytical solutions to Eq. (10), but the self-similarity property described in Eq. (22) can be used to corroborate the validity of the full numerical solution, as the self-similarity property comes out from a coupling between non-linearity and diffraction. Figure 4 shows a step wave incident on the free surface (the $\nu=0$ axis) and its associated reflection for three propagation distances of $x=1, 2, 3$. Here the rescaling is defined by the self-similar variables $\xi=t/x$ and $\eta=y/x$, and the grid parameters are the same as those used in the verification of the nonlinear operator (discussed previously). Visually the three plots are indis-

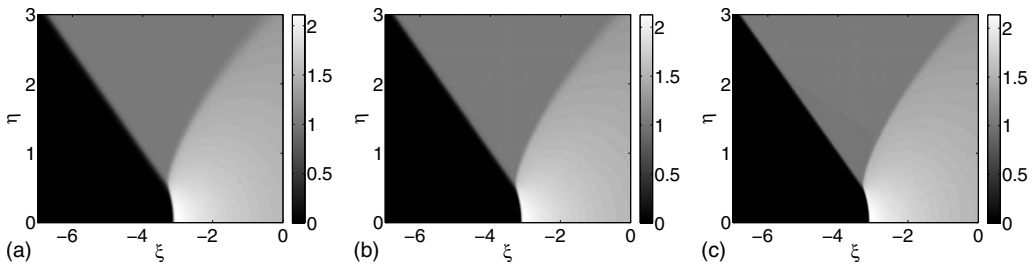


FIG. 4. Numerical demonstration of self similarity. The wave velocity (in gray levels) is shown for three propagation distances $x=1$, 2, and 3 (from left to right). The axes are in self-similar coordinates.

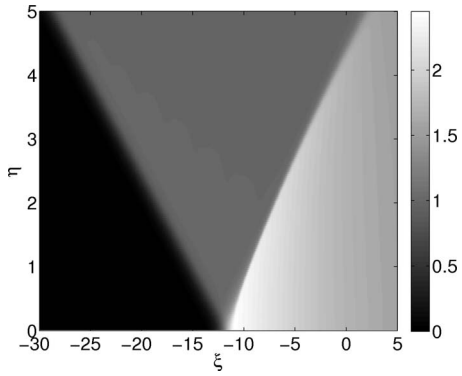


FIG. 5. Regular reflection: the wave velocity is shown at $x=1$ in gray levels for a shock impinging the free surface at a dimensionless angle $a=5.8$.

tinguishable from each other, confirming that the numerical solution is indeed self-similar, as the equations predict.

Numerical simulations were used to illustrate what occurs when the nondimensional angle a is less than the critical angle, 3.91. The numerical solution in Fig. 5 was calculated for an angle of $a=5.8$. As predicted by theory, it shows a two-shock solution with the reflection occurring directly at the interface. Because of non-linearity, however, the incident and reflection angles are slightly different, and the maximum velocity at the free surface is larger than 2.

As was described in the theory section, for values of a that are inferior to the critical angle, the two-shock regime (regular reflection) ceases to exist. In particular, there are two types of irregular reflection that emerge, depending on how much smaller a is compared to the critical angle.

Figure 6 illustrates the behavior for the first kind of irregular reflection. Here $a=3$ and in this regime three shocks that meet at a triple point can clearly be observed: the incident shock, the reflected shock, and the Mach shock connecting the previous two to the interface. The incident shock is straight, but contrary to classical Mach theory, the Mach shock and reflected shock are curved. Maximum amplitude is well above 2. This type of behavior is referred to as a von Neumann reflection, in a way similar to observations for quadratic nonlinear waves in fluids.[4,14]

As the value of a is decreased, the magnitude of the reflected shock also decreases until the three-shock solution is lost. Figure 7 illustrates this type of behavior for a nondimensional angle $a=1.5$. There is no reflected shock at all and

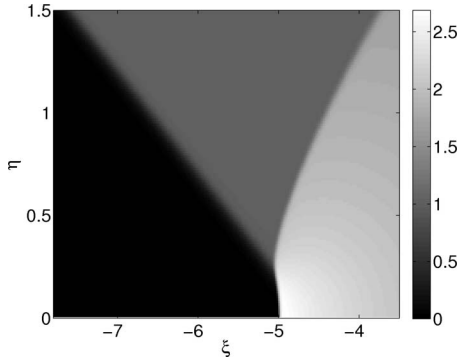


FIG. 6. Reflection in the von Neumann regime, with $a=3$ and a characteristic curved Mach shock and reflected shock.

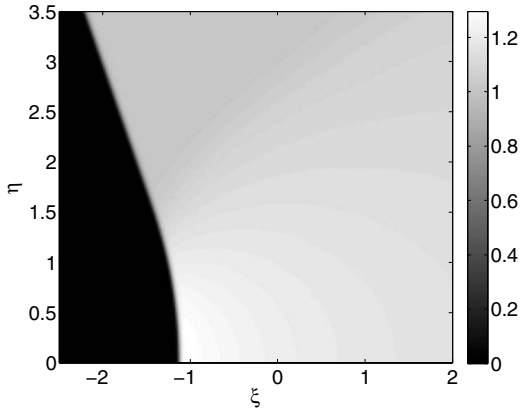


FIG. 7. Reflection in the weak von Neumann regime, with $a=1.5$. There is no reflected shock due to the almost grazing angle.

the reflected wave field is a smooth compression wave. As a consequence, maximum velocity is between 1 and 2. This is called a weak von Neumann regime in analogy again with the fluid case.

The maximum wave velocity as a function of the parameter a is explored in Fig. 8. For a equal to the critical angle, the theoretical maximum value of v is 3.68, which is close to the numerically predicted value of 3.77.

Figure 8 summarizes the various reflection regimes: for large values of parameter a, the shock reflection behavior is almost linear and linear doubling at the interface is recovered. When decreasing the angle down to the critical value, reflection remains regular but, because of non linear effects, the maximum velocity increases. Maximum deviation is observed around the critical value. Below a_{cr}, the von Neumann three-shock regime is observed, until the reflected shock fully disappears. This is the weak von Neumann regime which converges smoothly towards the perfect grazing case at $a=0$ for which the incident shock remains undisturbed. As for shock waves in fluids,[4] that transition solves the acoustical von Neumann paradox for shear shock waves in a smooth but nonmonotone way.

VI. CONCLUSION

The equation for a polarized nonlinear shear wave was simplified with the paraxial approximation, which is valid for

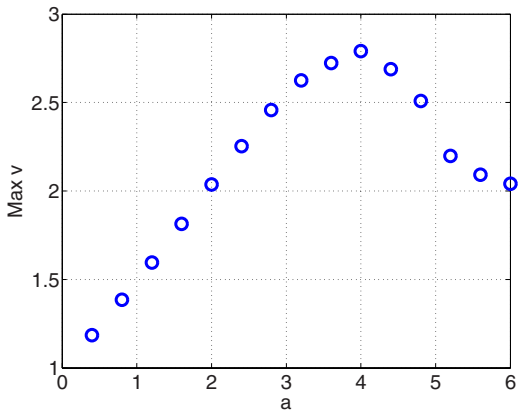


FIG. 8. (Color online) Maximum velocity as a function of the critical parameter a illustrating the acoustical von Neumann paradox for nonlinear shear shocks. The transition between irregular and regular reflection occurs at $a=3.91$.

small grazing angles, to obtain an equation that is similar to the KZ equation for compressional waves but where the nonlinear term is cubic rather than quadratic. Due to the invariance of the equation under a sign change in v, the cubic term has important implications for the types of shocks that are admissible, and contrary to acoustics, both compressive and expansive shocks may occur. Nevertheless the nonmonotonic transition between reflection regimes waves mimic the properties of reflected shock compressive waves.

To study the reflection at a free interface, a control parameter a was introduced that describes the relationship between the diffractive and nonlinear effects. This parameter is essential for the categorization of the reflection behaviors that emerged. The step-shock case was studied to facilitate the analysis of these different reflection regimes because of its self-similarity properties and the ease with which the Rankine–Hugoniot shock relations apply.

Three types of reflection were observed theoretically and numerically. For large values of a, an exact solution can be found where the incident and reflected shocks are straight (similar to classical Snell–Descartes reflection) but the angle of incidence is different from the angle of reflection and the reflected velocity magnitude is larger than what is predicted by classical theory. For a nondimensional angle that is less than the critical value, $a=3.91$, the theory breaks down and numerical simulations were used to observe the two remaining types of reflection. For angles that are not much less the critical value, a three shock solution emerged that differed from a classical Mach solution because the reflected shock and the Mach shock were curved. However, the incident shock remained straight and there was a triple point where the three shocks met. As the value of a decreased further, the reflected shock disappeared. These last two modes are known as von Neumann and weak von Neumann reflection.

By solving and characterizing the cubically nonlinear shear wave equation as it is reflected from a free surface and as it transitions between the different types of reflection regimes, we have described a solution to the acoustical von Neumann paradox for elastic materials. Further studies would investigate shock reflection for nonideal pulse waveforms, and experimental demonstration.

APPENDIX

The objective is to determine the number of real and positive roots of the polynomial $p(x)=x^3-ax^2+2x+a$. We first investigate the roots of its derivative $p'(x)=3x^2-2ax+2$, which vanish at $x_\pm=a\pm\sqrt{a^2-6}/3$. If $a<\sqrt{6}$, these roots are complex; hence $p'(x)$ does not vanish and has the same sign as $p'(0)=2>0$. As a consequence, $p(x)>p(0)=a>0$ remains positive and has no positive root. If $a>\sqrt{6}$, $p'(x)$ reaches its minimum value either at 0 or at x_+, and a local maximum at x_-. As $p(0)=a>0$, the polynomial will have real positive roots only if

$$g(a)=p(x_+)=\left(\frac{a+\sqrt{a^2-6}}{3}\right)-a\left(\frac{a+\sqrt{a^2-6}}{3}\right)^2+2\left(\frac{a+\sqrt{a^2-6}}{3}\right)+a<0. \tag{28}$$

It is easy to check that $g(a)$ is monotonously decreasing, with $g(\sqrt{6})>0$ and $g(+\infty)\to-\infty$. Hence the function $g(a)$ has a single root, which can be evaluated numerically $a_{\mathrm{cr}}=3.908\,167\,555\,1324\,29\simeq 3.91$.

[1] M. F. Hamilton and D. T. Blackstock, *Nonlinear Acoustics* (Academic, San Diego, 1998).
[2] R. Marchiano, F. Coulouvrat, and R. Grenon, "Numerical simulation of shock wave focusing at fold caustics, with application to sonic boom," J. Acoust. Soc. Am. **114**, 1758–1771 (2003).
[3] R. Marchiano, F. Coulouvrat, and J.-L. Thomas, "Nonlinear focusing of acoustic shock waves at a caustic cusp," J. Acoust. Soc. Am. **117**, 566–577 (2005).
[4] S. Baskar, F. Coulouvrat, and R. Marchiano, "Nonlinear reflection of grazing acoustical shock waves: Unsteady transition from von Neumann to Mach to Snell-Descartes reflections," J. Fluid Mech. **575**, 27–55 (2007).
[5] G. Ben-Dor, *Shock Wave Reflection Phenomena* (Springer, New York, 1992).
[6] E. Mach, "Über den Verlauf von Funkenwellen in der Ebene und im Raume (On the propagation of radio waves in a plane and in space)," Sitzungsbr. Akad. Wiss. Wien **78**, 819–838 (1878).
[7] G. Ben-Dor and K. Takayama, "The phenomena of shock wave reflection—A review of unsolved problems and future research needs," Shock Waves **2**, 211–223 (1992).
[8] J. von Neumann, "Oblique reflection of shocks," *John von Neumann Collected Work*, edited by A. H. Taub (Pergamon, New York, 1943), Vol. **6**, pp. 238–299.
[9] L. F. Henderson, "Region and boundaries for diffracting shock wave systems," Z. Angew. Math. Mech. **67**, 73–86 (1987).
[10] S. Catheline, J.-L. Gennisson, M. Tanter, and M. Fink, "Observation of shock transverse waves in elastic media," Phys. Rev. Lett. **91**, 164301 (2003).
[11] J. K. Hunter and M. Brio, "Weak shock reflection," J. Fluid Mech. **410**, 235–261 (2000).
[12] B. W. Skews and J. T. Ashworth, "The physical nature of weak shock wave reflection," J. Fluid Mech. **542**, 105–114 (2005).
[13] A. M. Tesdall and J. K. Hunter, "Self-similar solutions for weak shock reflection," SIAM J. Appl. Math. **63**, 42–61 (2002).
[14] R. Marchiano, S. Baskar, F. Coulouvrat, and J.-L. Thomas, "Experimental evidence of deviation from mirror reflection for acoustical shock waves," Phys. Rev. E **76**, 056602 (2007).
[15] P. Colella and L. F. Henderson, "The von Neumann paradox for the diffraction of a weak shock waves," J. Fluid Mech. **213**, 71–94 (1990).
[16] X. Jacob, S. Catheline, J. L. Gennisson, Ch. Barrière, D. Royer, and M. Fink, "Nonlinear shear wave interaction in soft solids," J. Acoust. Soc. Am. **122**, 1917–1926 (2007).
[17] R. W. Lardner, "Nonlinear effects on transverse shear waves in an elastic medium," J. Elast. **15**, 53–57 (1985).
[18] I. P. Lee-Bapty and D. G. Crighton, "Nonlinear wave motion governed by the modified Burgers equation," Philos. Trans. R. Soc. London, Ser. A **323**, 173–209 (1987).
[19] E. A. Zabolotskaya, "Sound beams in a nonlinear isotropic solid," Sov. Phys. Acoust. **32**, 296–299 (1986).
[20] F. Coulouvrat, "A quasi exact shock fitting algorithm for general nonlinear progressive waves," Wave Motion **49**, 97–107 (2009).
[21] M. Wochner, M. Hamilton, Y. Ilinksii, and E. Zabolotskaya, "Cubic nonlinearity in shear wave beams with different polarizations," J. Acoust. Soc. Am. **123**, 2488–2495 (2008).
[22] M. F. Hamilton, Y. A. Ilinskii, and E. A. Zabolotskaya, "Separation of compressibility and shear deformation in the elastic energy density," J. Acoust. Soc. Am. **116**, 41–44 (2004).
[23] E. A. Zabolotskaya, M. F. Hamilton, Y. A. Ilinskii, and G. Douglas Meegan, "Modeling of nonlinear shear waves in soft solids," J. Acoust. Soc. Am. **116**, 2807–2813 (2004).
[24] J. L. Gennisson, M. Rénier, S. Catheline, Ch. Barrière, J. Bercoff, M. Tanter, and M. Fink, "Acoustoelasticity in soft solids: Assessment of the

nonlinear shear modulus with the acoustic radiation force," J. Acoust. Soc. Am. **122**, 3211–3219 (2007).
[25]M. Rénier, J. L. Gennisson, Ch. Barrière, D. Royer, and M. Fink, "Fourth-order shear elastic constant assessment in quasi-incompressible soft solids," Appl. Phys. Lett. **93**, 101912 (2008).
[26]E. A. Zabolotskaya and R. V. Khokhlov, "Quasi-plane waves in the nonlinear acoustics of confined beams," Sov. Phys. Acoust. **15**, 35–40 (1969).
[27]F. Coulouvrat and R. Marchiano, "Nonlinear Fresnel diffraction of weak shock waves," J. Acoust. Soc. Am. **114**, 1749–1757 (2003).
[28]Y. S. Lee and M. F. Hamilton, "Time-domain modeling of pulsed finite-amplitude sound beams," J. Acoust. Soc. Am. **97**, 906–917 (1995).
[29]R. J. Zemp, J. Tavakkoli, and R. S. Cobbold, "Modeling of nonlinear ultrasound propagation in tissue from array transducers," J. Acoust. Soc. Am. **113**, 139–152 (2003).
[30]S. K. Godunov, "A finite difference method for the numerical computation and discontinuous solutions of the equations of fluid dynamics," Mat. Sb. **47**, 271–306 (1959).
[31]J. Trangenstein, *Numerical Solution of Hyperbolic Partial Differential Equations* (Cambridge University Press, Cambridge, 2009).

Nonlinear Biot waves in porous media with application to unconsolidated granular media

Olivier Dazel and Vincent Tournat[a)]
LAUM, CNRS, Université du Maine, Avenue Olivier Messiaen, 72085 Le Mans, France

(Received 3 November 2008; revised 15 October 2009; accepted 24 November 2009)

The nonlinear propagation through porous media is investigated in the framework of Biot theory. For illustration, and considering the current interest for the determination of the elastic properties of granular media, the case of nonlinear propagation in "model" granular media (disordered packings of noncohesive elastic beads of the same size embedded in a visco-thermal fluid) is considered. The solutions of linear Biot waves are first obtained, considering the appropriate geometrical and physical parameters of the medium. Then, making use of the method of successive approximations of nonlinear acoustics, the solutions for the second harmonic Biot waves are derived by considering a quadratic nonlinearity in the solid frame constitutive law (which takes its origin from the high nonlinearity of contacts between grains). The propagation in a semi-infinite medium with velocity dispersion, frequency dependent dissipation, and nonlinearity is first analyzed. The case of a granular medium slab with rigid boundaries, often considered in experiments, is then presented. Finally, the importance of mode coupling between solid and fluid waves is evaluated, depending on the actual fluid, the bead diameter, or the applied static stress on the beads. The application of these results to other media supporting Biot waves (porous ceramics, polymer foams, etc.) is straightforward. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277190]

PACS number(s): 43.25.Dc, 43.20.Jr, 43.20.Gp, 43.20.Bi [ROC] Pages: 692–702

I. INTRODUCTION

The acoustics of granular media has been widely investigated among different fields such as geophysics, underwater, or airborne acoustics (sound proofing, shock wave absorption, etc.). Recently in physics, granular materials have attracted a strong interest because they exhibit unusual behaviors.[1,2] These behaviors are sometimes comparable to glassy media, from the point of view of the aging process, for instance,[3] or regarding the prediction of the soft modes and their relation to the boson peak.[4,5] The transition from one state to another, and, in particular, the unjamming transition from a solid-like to a fluid-like phase, is intensively studied.[6] One key point for the related experimental and theoretical investigations is to obtain the elastic properties of the granular materials. Their geometrical properties being better characterized thanks to several existing methods such as x-ray tomography, γ-ray transmission, and iso-index optical imaging, for instance.[3] In this context, acoustic waves are the only one able to probe accurately the elastic properties in the bulk of three-dimensional media and should contribute to the experimental characterization of the aforementioned processes.

Frequently, due to the complexity in behaviors and structures of granular media (sand, sediments, etc.) and due to the subsequent lack of realistic modeling, it is of interest to study ''model'' granular media such as disordered packings of noncohesive spherical beads of the same size. In this type of model media, there exists a characteristic scale, the diameter d of the beads, and one can use a constitutive law at the level of the contacts such as the Hertz or the Hertz–Mindlin contact laws.[7,8] The intrinsic nonlinearities ensure a parameter of quadratic nonlinearity for these media 100–1000 times larger than in homogeneous fluids and solids.[9] As a consequence, various nonlinear acoustic effects have been observed for moderate excitation amplitudes.[10–16] They are able to give complementary information on the elastic properties of model granular media. In particular, nonlinear waves exhibit a preferential sensitivity to the most weakly stressed contacts of the granular packing[14,16] (a distribution of contact stresses exists due to the disordered structure of the granular packing[17]) compared to the linear waves.

The results on the elastic wave propagation in the solid frame of model granular media contain several features, such as a dependence of longitudinal and shear wave velocities on the applied static pressure,[18,19,14] a frequency dependent attenuation, a velocity dispersion, multiple scattering effects at wavelengths of the order of the bead diameter,[19] memory effects,[20] hysteresis,[10] response fluctuations,[4] dynamic dilatancy,[14] and nonlinearities.

Most of the studies on wave propagation through model granular media either consider the propagation through the solid network, neglecting the saturating fluid, or consider the propagation through the fluid phase saturating the rigid bead packing, which is assumed to be motionless.

For waves propagating in the fluid saturating the bead packing, the equivalent fluid model (i.e., the solid is motionless) has been successfully applied to the case of long-wavelength propagation in air saturating glass beads, for instance,[21–23] and provides the evaluation of useful parameters for the present work. Theoretical investigations on air-saturated granular media in the frame of the equivalent fluid model have also been performed.[24,22,25] Nonlinear effects

[a)]Author to whom correspondence should be addressed. Electronic mail: vincent.tournat@univ-lemans.fr

have been investigated in this equivalent fluid model approximation in the case of intense sound waves.[26] For shorter wavelengths, becoming of the order of the bead diameter, scattering occurs and its modeling for such close scatterers is a hard task even if the geometry is well-known.[27] In water, owing to the lower acoustic energy dissipation than in air, several pioneering experiments have been carried out on the acoustic energy diffusion process and showed that most of the energy transports through water and that the associated waves are multiply scattered by the glass beads.[28]

There are only few experiments where both waves in the fluid and in the bead structure have been shown to play a role and to couple. In Ref. 29, the nonlinear self-demodulation process has been experimentally observed with a transmission toward the air of the frequency component nonlinearly generated inside the solid frame.

The aim of this article is to use the Biot theory,[30,31] generally applied to linear acoustics, to model the nonlinear propagation in porous media, with an application to granular media.

Biot theory describes the deformation of a porous deformable elastic solid in which the nonsolid part is saturated by a compressible fluid. Even if Biot theory is not able to model the behavior of every type of porous materials (porous rocks, etc.), it has been successfully applied to a wide range of materials[32,33] using homogenization techniques[34,35] or volume averaging methods.[36,37] The application of Biot theory to sound absorbing materials takes its origin in the beginning of 1980s. For acoustical materials, the saturating fluid is air. The description of the viscous and thermal properties of air saturating a porous immobile solid has been a wide research topic and many models have been published.[38–41,36] These models consist in introducing a frequency dependent complex density (respectively, compressibility) to take into account viscous (respectively, thermal) effects. Extensive details on these aspects can be found in Refs. 38, 42, and 43.

In this article, an extension of the Biot theory is first presented to describe the nonlinear acoustic propagation in a nonlinear porous solid, where the motion of the solid and fluid phases with viscous and thermal effects is considered. Inertial and elastic couplings between both constituents are considered. This work is mainly based on two references in the literature[44,45] about nonlinear wave propagation in porous media. The originality of the present study is first to propose a new formulation to describe the nonlinear Biot wave propagation, which allows for significant simplifications of the existing expressions without any additional assumptions, second to take into account the viscous and thermal effects through Johnson–Allard's models, and finally, to apply the theory to unconsolidated granular media. In the case of interest, the parameters of the Biot theory corresponding to "model" granular media are used, but the theoretical results could be applied to other media supporting Biot waves, such as sediments, porous ceramics, polymer foams, and trabecular bone,[46] possibly with other saturating fluids. After the presentation of the theory, we analyze the effect of second harmonic generation in a one-dimensional geometry for a semi-infinite medium and for a slab configuration. We describe the influence of some parameters on the mode coupling and on the nonlinear effects (such as the bead diameter, for instance, or parameters of the solid frame and of the saturating fluid). This allows to evaluate the importance of the often neglected effects of coupling in the linear and nonlinear acoustic experiments on model granular media.

II. BIOT THEORY WITH QUADRATIC NONLINEARITY

A. Nonlinear Biot equations in $\{\mathbf{u}^s,\mathbf{u}^W\}$ formulation

Biot theory[31] is able to model the acoustic propagation through homogenized porous media using six fields, which can be, depending on the chosen formulation, the three displacements of each homogenized phase ($\mathbf{u}^s$ and $\mathbf{u}^f$, respectively, for the solid and fluid displacements) or any free linear combination[30,47] of these two fields. As the 1956 (Ref. 31) formulation was not valid for porous material with an inhomogeneous porosity ϕ, Biot[30] derived a second formulation in which the considered fields are the solid displacement $\mathbf{u}^s$ and the relative flow $\mathbf{w}=\phi(\mathbf{u}^f-\mathbf{u}^s)$. Recently, Dazel *et al.*[47] published a strain decoupled formulation (equivalent to 1962 formulation), which simplifies the Biot formalism without additional assumptions. One goal of this article is to show that similar simplifications can also be obtained for nonlinear porous materials. The fields involved in the present formulation are the solid displacement $\mathbf{u}^s$ and the strain decoupled displacement field $\mathbf{u}^W=\mathbf{u}^t+(\alpha-1)\mathbf{u}^s$, where $\alpha=1-K_s/K_r$ (K_r is the bulk modulus of the bead matter, glass for the example in Sec. II B, and K_s is the bulk modulus of the porous solid in vacuum) and where $\mathbf{u}^t=(1-\phi)\mathbf{u}^s+\phi\mathbf{u}^f$ is the total displacement of the porous medium.

In $\{\mathbf{u}^s,\mathbf{u}^W\}$ formulation, the equations of motion at cyclic frequency ω read as[47]

$$\sigma^s_{ij,j}(\mathbf{u}^s)=-\omega^2\tilde{\rho}_s\mathbf{u}^s-\omega^2\tilde{\rho}_{\mathrm{eq}}\tilde{\gamma}\mathbf{u}^W, \tag{1a}$$

$$-\boldsymbol{\nabla}p=-\omega^2\tilde{\rho}_{\mathrm{eq}}\tilde{\gamma}\mathbf{u}^s-\omega^2\tilde{\rho}_{\mathrm{eq}}\mathbf{u}^W, \tag{1b}$$

where $\tilde{\rho}_{\mathrm{eq}}$ and $\tilde{\rho}_s$ are equivalent densities and $\tilde{\gamma}$ is a coupling coefficient. The tilde symbol "$\tilde{}$" is associated with frequency dependent coefficients at frequency ω, $\boldsymbol{\nabla}$ denotes the partial derivative nabla operator, p corresponds to the interstitial pressure, and σ^s_{ij} is the *in-vacuo* stress tensor of the solid phase. For a linear constitutive law, it has been shown that σ^s_{ij} only depends on $\mathbf{u}^s$ and that p only depends on $\mathbf{u}^W$,[47]

$$\sigma^s_{ij}=2\mu\varepsilon_{ij}+\lambda\varepsilon\delta_{ij},\quad p=-M\xi, \tag{2}$$

where λ and μ are the Lamé coefficients of the solid, M is the equivalent compressibility of the equivalent fluid model, ε_{ij} is the strain tensor of the solid phase, δ_{ij} is the Kronecker tensor, $\varepsilon=\boldsymbol{\nabla}\cdot\mathbf{u}^s$, and $\xi=\boldsymbol{\nabla}\cdot\mathbf{u}^W$. In the two other formulations[30,31] the conjugated stresses depend on both displacement fields. Moreover, σ^s_{ij} is related to the total stress tensor σ^t_{ij} by the following relation:

$$\sigma^s_{ij}=\sigma^t_{ij}+\alpha p\delta_{ij}. \tag{3}$$

Stress-strain relations (2) are modified in the case of a nonlinear porous material. As mentioned in Sec. I, the nonlinearity of granular media mainly comes from the contacts between beads and is several orders of magnitude higher than

in homogeneous solids and fluids.[9] Even at moderate excitation amplitudes, nonlinear effects have been observed.[10–16] As a consequence, in the present case, we neglect Forchheimer's nonlinearity, which has been shown to play a role in some air-saturated porous media at high acoustic levels in the context of a rigid solid frame (equivalent fluid approximation).[26] A way to consider nonlinear terms in the constitutive laws has been described by Biot[44] and Donskoy *et al.*[45] and consists in introducing a nonlinear potential H defined by

$$H = D\left(\frac{\bar{I}_1^3}{3} - \bar{I}_1\bar{I}_2 + \bar{I}_3\right) + F(\bar{I}_1\bar{I}_2 - 3\bar{I}_3) + G\bar{I}_3 \tag{4}$$

where the $\bar{I}_i$ are the strain invariants associated with the modified strain[45] $\bar{\varepsilon}_{ij}$ defined by

$$\bar{\varepsilon}_{ij} = \varepsilon_{ij} + p\theta\delta_{ij}, \tag{5}$$

where $\theta = K_r/3$ is the fluid compliance coefficient. The strain invariants are

$$\bar{I}_1 = \mathrm{tr}(\bar{\varepsilon}_{ij}), \quad \bar{I}_2 = \bar{\varepsilon}_{11}\bar{\varepsilon}_{22} + \bar{\varepsilon}_{11}\bar{\varepsilon}_{33} + \bar{\varepsilon}_{22}\bar{\varepsilon}_{33} - \bar{\varepsilon}_{12}\bar{\varepsilon}_{21} - \bar{\varepsilon}_{13}\bar{\varepsilon}_{31} - \bar{\varepsilon}_{23}\bar{\varepsilon}_{32}, \quad \bar{I}_3 = \det(\bar{\varepsilon}_{ij}). \tag{6}$$

The nonlinear constitutive laws with the nonlinear potential H are then written [Eq. (2)] as

$$\sigma_{ij}^s = 2\mu\varepsilon_{ij} + \lambda\varepsilon\delta_{ij} + \partial_{\bar{\varepsilon}_{ij}}H, \tag{7a}$$

$$p = -M\xi + M\theta\delta_{ij}\partial_{\bar{\varepsilon}_{ij}}H. \tag{7b}$$

These two equations [Eqs. (7a) and (7b)] correspond to relations (12) in Ref. 45. Equation (7b) is equivalent to the second relation (12) of Ref. 45 but one can notice that relation (7b) has the advantage of involving in its linear part only the divergence of $\mathbf{u}^W$ instead of the mixture law $\boldsymbol{\nabla}(\mathbf{w}+\alpha\mathbf{u}^s)$. The use of the *in-vacuo* stress instead of the total stress tensor of the porous medium implies that the pressure p does not appear in Eq. (7a) as it is the case for the first relation (12) in Ref. 45 for the total stress tensor. The problem of elimination of the pressure from σ^t was mentioned in Ref. 45 and does not stand in our formulation of the poroelasticity equations. It is an advantage of the present formulation.

The nonlinear contributions of H to the *in-vacuo* stress and pressure can be expanded and read as

$$\partial_{\bar{\varepsilon}_{ij}}H = F\bar{I}_1^2\delta_{ij} + (F-D)(\delta_{ij}\bar{I}_2 - \bar{I}_1\bar{\varepsilon}_{ij}) + (G+D-3F)\mathrm{cof}(\bar{\varepsilon}_{ij}), \tag{8a}$$

$$\delta_{ij}\partial_{\bar{\varepsilon}_{ij}}H = 3F\bar{I}_1^2 + (F-D)(3\bar{I}_2 - \bar{I}_1^2) + (G+D-3F)\bar{I}_2 = P\bar{I}_1^2 + (G-2D)\bar{I}_2, \tag{8b}$$

with $P=D+2F$ and $\mathrm{cof}(\bar{\varepsilon}_{ij})$ corresponds to the matrix of cofactors.

B. One-dimensional problem

Only one-dimensional problems are considered in the following, which implies simplifications that are now detailed for the nonlinear terms. The displacement fields are along the x direction, and all the fields of the problem only depend on x. The space derivation with respect to x is denoted by subscript $_{,x}$ ($_{,xx}$ for the second derivative). Hence x corresponds to the direction 1 of Sec. II A and directions 2 and 3 are not considered. In this case, the strain invariants read as

$$\bar{I}_1 = u_{,x}^s + 3\theta p, \quad \bar{I}_2 = (2u_{,x}^s + 3\theta p)\theta p, \quad \bar{I}_3 = (u_{,x}^s + \theta p)\theta^2 p^2, \tag{9}$$

and the nonlinear contributions to the stress are

$$\partial_{\bar{\varepsilon}_{ij}}H = D(u_{,x}^s)^2 + 2P(\theta p u_{,x}^s) + K(\theta p)^2, \tag{10a}$$

$$\delta_{ij}\partial_{\bar{\varepsilon}_{ij}}H = P(u_x^s)^2 + 2K\theta p u_x^s + 3K(\theta p)^2, \tag{10b}$$

with $K=D+6F+G$. In the following, the method of successive approximation is used, and the forces corresponding to the nonlinear terms contain only the linear solutions previously computed (and denoted with an l index) u_l^s and u_l^W. Consequently, the pressure p_l in the nonlinear terms is defined by relation (2) and $p_l = -Mu_{l,x}^W$. The forces associated with the nonlinear potential are then written as

$$\partial_{\bar{\varepsilon}_{ij}}H_{l,x} = 2Du_{l,x}^s u_{l,xx}^s - 2MP\theta(u_{l,xx}^W u_{l,x}^s + u_{l,x}^W u_{l,xx}^s) + 2KM^2\theta^2 u_{l,x}^W u_{W,xx}^t, \tag{11a}$$

$$\delta_{ij}\partial_{\bar{\varepsilon}_{ij}}H_{l,x} = 2Pu_{l,x}^s u_{l,xx}^s - 2MK\theta(u_{l,xx}^W u_{l,x}^s + u_{l,x}^W u_{l,xx}^s) + 6KM^2\theta^2 u_{l,x}^W u_{l,xx}^W. \tag{11b}$$

These two last equations are compared to Eqs. (20) and (21) of Ref. 45 obtained with $\{\mathbf{u}^s, \mathbf{w}\}$ formulation. In comparison, it is important to note the drastic simplifications of the nonlinear forces in Eq. (11) obtained using our formulation in $\{\mathbf{u}^s, \mathbf{u}^W\}$ without loss of generality.

C. Simplifications associated with the stiffness ratio between the solid frame and the frame material

Additional simplifications can be made in the case of granular media taken here as an illustration, and for other porous media supporting Biot waves as long as the following conditions on compressibility are fulfilled. The bead material is much stiffer than the contacts between beads (and consequently the solid frame),[7] which leads to the following strong inequalities: $K_r \gg K_s$ and $K_r \gg K_f$. Consequently, $\alpha \approx 1$, which leads to $u^W = u^t$. Also $\theta M \approx K_f/3\phi K_r \ll 1$. Hence, the only remaining nonlinear term in Eq. (11a) is the one containing $u_{l,x}^s u_{l,xx}^s$ and the constitutive law [Eq. (7b)] can be approximated by its linear part.

For model granular media, the quadratic elastic nonlinearity is a reasonably good first approximation[11,14,15] as for other porous media, in order to explain the harmonic generation process, at least for some moderate range of excitation amplitude.[13,14] Consequently, we keep this type of nonlinearity as the starting point of the present study. The value of the parameter of quadratic nonlinearity has been measured to be several orders of magnitude higher than in homogeneous

fluid or solids,[9,12] but can vary over a wide range depending on the external conditions of static stress or on the particular bead arrangement, for instance.[48]

With the previous assumptions, the one-dimensional equations of motion for a model granular porous medium read as

$$\lambda_c u^s_{,xx} + 2Du^s_{,x}u^s_{,xx} = -\omega^2\tilde{\rho}_s u^s - \omega^2\tilde{\gamma}\tilde{\rho}_{\mathrm{eq}}u^t, \tag{12a}$$

$$Mu^t_{,xx} = -\omega^2\tilde{\gamma}\tilde{\rho}_{\mathrm{eq}}u^s - \omega^2\tilde{\rho}_{\mathrm{eq}}u^t, \tag{12b}$$

where $\lambda_c=\lambda+2\mu$. In the following, the value $2D=100\lambda_c$ is taken everywhere. Considering this value for the parameter of quadratic nonlinearity, the expansion of the stress-strain relationship up to the quadratic term is valid if $|u^s_{,x}| \ll 1/50$. In this manuscript, dispersive viscous and thermal effects of the fluid are taken into account based on Johnson *et al.*[40] and Champoux–Allard[39] models. It is not the case in Refs. 44 and 45. Details on these models can be found in Appendix B.

D. Case of the pure elastic solid

It is instructive to consider the pure elastic solid case, which corresponds to the propagation of a single mode in the solid frame. This allows to compare our results with other studies of acoustic propagation in granular media, which neglect the influence of the saturating fluid. Note that inertial coupling effects with the fluid are taken into account but not the thermal effects, unlike the so-called *in-vacuo* stress, which neglects completely the saturating fluid. This case is called "elastic" in the following (unlike poroelastic for the Biot modes), and the comparison of the elastic and poroelastic solutions should reveal the influence of the saturating fluid. The unknown elastic displacement field is denoted by u^e and the motion equation [Eq. (12)] becomes

$$\lambda_c u^s_{,xx} + 2Du^e_{,x}u^e_{,xx} = -\omega^2\tilde{\rho}u^e. \tag{13}$$

The method of successive approximations is also applied for this problem and is detailed in Appendix A 2.

III. RESOLUTION OF THE PROBLEM

The method of successive approximations is used to solve the problem of second harmonic generation at the leading order of nonlinearity.[49,50] It is equivalent to solving a so-called "weak nonlinear" problem, where the nonlinear acoustic effects are quadratic and the nonlinear interactions only take place once for the fundamental wave (or primary radiated wave). This gives the limit of applicability for this quadratic approximation: the primary wave amplitude should be weak enough so that the nonlinearly generated components (such as the second harmonic wave) exhibit negligible nonlinear effects themselves.

The first step of this method is to find the solution $\{u^s_l, u^t_l\}$ of the linear problem at frequency ω for which the nonlinearity is omitted. The second step is to write the problem [Eq. (12)] at frequency 2ω and to substitute the solution of the linear problem at frequency ω in the nonlinear term $2Du^s_{,x}u^s_{,xx}$ of Eq. (12). This term then acts as a source term for the second harmonic wave. A second linear problem, at frequency 2ω, is solved. Simple mathematics shows that the solution of this problem is the superposition of the particular and the general solutions. It is important to note that in this weakly nonlinear quadratic approximation, no nonlinear effect on wave speed (in average over a wave period) and attenuation occur.[49]

A. General form of the solutions

It is straightforward by the methods of characteristics[38,31] to show that the solution of the linear problem at frequency ω is a superposition of two compressional waves, where the solid and total displacements read as

$$u^s_l = (u_1 e^{-jk_1 x} + u_2 e^{-jk_2 x} + u'_1 e^{jk_1 x} + u'_2 e^{jk_2 x})e^{j\omega t}, \tag{14a}$$

$$u^t_l = (\mu_1 u_1 e^{-jk_1 x} + \mu_2 u_2 e^{-jk_2 x} + \mu_1 u'_1 e^{jk_1 x} + \mu_2 u'_2 e^{jk_2 x})e^{j\omega t}, \tag{14b}$$

where u_1, u_2, u'_1, and u'_2 are coefficients determined by the boundary conditions of the problem, and k_1 and k_2 are the wave numbers of the two compressional waves. For each one of these modes, μ_i is the ratio of the total displacement amplitude over the solid one. Their expressions are given at the end of Appendix B.

The particular solution of the problem at 2ω can be written in the following forms:

$$u^s_p = \alpha_0 + \left(\sum_{i=1}^{4} \alpha_i e^{-2jk_i x} + \alpha'_i e^{2jk_i x}\right)e^{2j\omega t}, \tag{15}$$

$$u^t_p = \beta_0 + \left(\sum_{i=1}^{4} \beta_i e^{-2jk_i x} + \beta'_i e^{2jk_i x}\right)e^{2j\omega t}, \tag{16}$$

where $2k_3=k_1+k_2$, $2k_4=k_1-k_2$, and α_i and β_i are obtained through an identification method presented in Appendix A. The form of the general solution of the problem at 2ω is quite similar to the linear solution

$$u^s_g = (\gamma_1 e^{-j\ddot{k}_1 x} + \gamma_2 e^{-j\ddot{k}_2 x} + \gamma'_1 e^{j\ddot{k}_1 x} + \gamma'_2 e^{j\ddot{k}_2 x})e^{2j\omega t}, \tag{17a}$$

$$u^t_g = (\ddot{\mu}_1\gamma_1 e^{-j\ddot{k}_1 x} + \ddot{\mu}_2\gamma_2 e^{-j\ddot{k}_2 x} + \ddot{\mu}_1\gamma'_1 e^{j\ddot{k}_1 x} + \ddot{\mu}_2\gamma'_2 e^{j\ddot{k}_2 x})e^{2j\omega t}. \tag{17b}$$

The double dot symbol "¨" denotes the 2ω values of the coefficients (for example, $\ddot{k}_1$ corresponds to the wave number of the first Biot wave at cyclic frequency 2ω). The γ_i are amplitudes determined by the boundary conditions for the nonlinear solution of the problem at 2ω.

B. Case of a semi-infinite porous medium

The case of a semi-infinite porous medium is first considered. Accordingly, there is no counter-propagating wave. This imposes that the amplitudes u'_1 and u'_2 are zero. A sinusoidal displacement with amplitude u_0 and frequency ω is imposed at $x=0$ on the solid and the fluid. The associated boundary condition is consequently $u^l_s(0)=u^l_t(0)=u_0$.[38,47] Then, u_1 and u_2 are solutions of the following system:

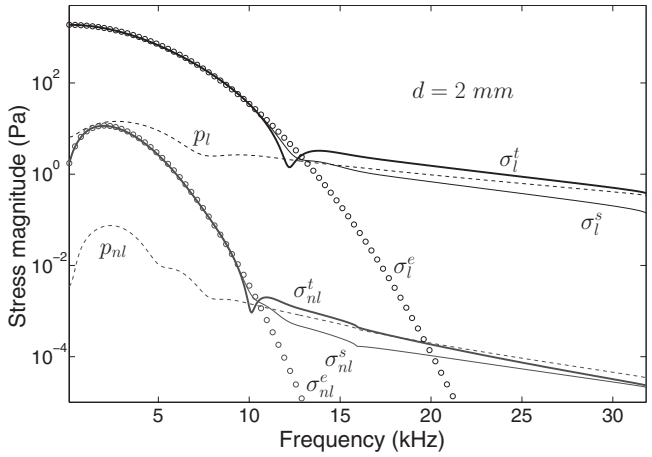


FIG. 1. Linear (at frequency ω) and nonlinear (at frequency 2ω) stress magnitudes at a distance $x=10$ cm from the excitation boundary in the semi-infinite configuration. σ^s, σ^f, σ^t, and σ^e are, respectively, the solid, fluid, total, and elastic stresses. Indices l and "nl" are, respectively, used for the linear and the nonlinear contributions. Note that the frequency axis indicates the fundamental frequency, i.e., the magnitude for the second harmonic at 2ω frequency should be read at ω. Case of air-saturated elastic beads of diameter 2 mm.

$$\begin{bmatrix} 1 & 1 \\ \mu_1 & \mu_2 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \end{Bmatrix} = \begin{Bmatrix} u_0 \\ u_0 \end{Bmatrix}, \tag{18}$$

which solutions read as

$$u_1 = \frac{\mu_2 - 1}{\mu_2 - \mu_1} u_0, \quad u_2 = \frac{\mu_1 - 1}{\mu_1 - \mu_2} u_0. \tag{19}$$

For the particular and general solutions of the problem at 2ω, the prime coefficients associated with the counter-propagating waves are also zero. This corresponds to consider only cumulative nonlinear effects. Local nonlinear effects, neglected in this theory, could in principle allow for the generation of counter-propagating second harmonic wave, but with a much lower efficiency. The particular solution is obtained using the method in Appendix A. The boundary conditions relative to the nonlinear displacements are the nullity of the solid and total displacements at $x=0$. Hence, γ_1 and γ_2 are solutions of

$$\begin{bmatrix} 1 & 1 \\ \ddot{\mu}_1 & \ddot{\mu}_2 \end{bmatrix} \begin{Bmatrix} \gamma_1 \\ \gamma_2 \end{Bmatrix} = - \begin{Bmatrix} \alpha_1 + \alpha_2 + \alpha_3 \\ \beta_1 + \beta_2 + \beta_3 \end{Bmatrix}, \tag{20}$$

and their expressions are

$$\gamma_1 = \frac{\ddot{\mu}_2(\alpha_1 + \alpha_2 + \alpha_3) - (\beta_1 + \beta_2 + \beta_3)}{\ddot{\mu}_1 - \ddot{\mu}_2}, \tag{21a}$$

$$\gamma_2 = \frac{-\ddot{\mu}_1(\alpha_1 + \alpha_2 + \alpha_3) + (\beta_1 + \beta_2 + \beta_3)}{\ddot{\mu}_1 - \ddot{\mu}_2}. \tag{21b}$$

In Fig. 1, the magnitude of the acoustic stress at a 10 cm distance from the source is plotted as a function of frequency. The excitation displacement at $x=0$ is imposed in the form $u=u_0 \sin(\omega t)$ with $u_0=0.001\times 2\pi/\omega$, which corresponds to an imposed strain amplitude independent of frequency and of realistic magnitude of 3×10^{-5} for a medium without dispersion. This is also close to the efficiency of wide-band ultrasonic transducers in experiments.[51] The stress in the solid frame σ^s (supported by the beads and their contacts) and the fluid pressure p both contribute to the total stress σ^t. The pure elastic stress is denoted by σ^e. Linear and nonlinear stresses, corresponding here to the contributions at frequencies ω and 2ω, respectively, are plotted separately for clarity because the linear stress is most of the time dominant over the nonlinear one for this process of second harmonic generation. The linear stresses are indexed with l while the nonlinear stresses are indexed with "nl."

For all the figures in this article, when it is not specified, the wave and medium parameters are those in Appendix B. These parameters are chosen in agreement with available experimental results[51,14,15] for weakly stressed granular packings of glass beads. The equivalent fluid parameters of the glass bead packing and their scaling laws with the bead radius are mostly taken from Ref. 21 (see Appendix B). The longitudinal wave velocity in the solid frame in the long wavelength limit is 200 m/s. It decreases with frequency as observed in sand and model granular media, and due to the absence of models of velocity dispersion in disordered three-dimensional bead packings, the dispersion law of a one-dimensional granular chain is used. The corresponding wavelength is 2 cm at 10 kHz (ten times larger than the bead diameter) and ~0.7 cm at 30 kHz. Due to the occurrence of scattering, the limit of applicability of the Biot theory is reached for the highest frequencies of the plots. However, this frequency limit of applicability can be increased if the coherent wave is considered by performing a spatial averaging of the acoustic field (over different configurations of the medium or by transducers with a sensitive surface larger than the wavelength, for instance). Scattering in this case would result in additional attenuation and velocity dispersion.

The linear elastic stress σ_l^e monotonously decreases in magnitude with increasing frequency because of the attenuation proportional to ω. Such as in a homogeneous solid or fluid with a quadratic nonlinearity and an increasing attenuation with frequency, the nonlinear stress σ_{nl}^e first increases with frequency, reaches a maximum, and then decreases. The linear solid stress σ_l^s has a different behavior than σ_l^e, showing that under the considered conditions, the coupling between the solid frame and the saturating fluid cannot be neglected here. It is possible to define a characteristic frequency f_c, from which σ_l^s deviates significantly from σ_l^e. This corresponds to a crossover between the two Biot waves. Below this frequency, the dynamics is mainly solid controlled, and above this frequency, it is fluid controlled. Around this frequency f_c, the solid and fluid waves are of comparable magnitude, and their interferences produce the notch at 13 kHz. It is important to mention that this crossover is not always characterized by a local minimum in σ_l^t, as shown in Fig. 1. The same qualitative behavior can be observed for the nonlinear stresses at frequency 2ω.

In Fig. 2, the same quantities as in Fig. 1 are plotted but as a function of distance from the excitation boundary at the fixed frequency of 10 kHz. At small distance from the source, the linear total stress σ_l^t is mainly supported by the solid and decreases exponentially with distance as does the pure elastic stress σ_l^e for the whole frequency range (a linear

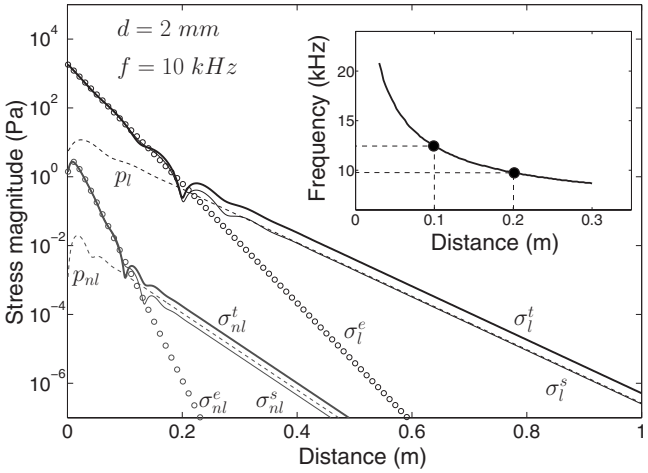


FIG. 2. Linear and nonlinear stress magnitudes as a function of distance from the source in the semi-infinite configuration. Same notations as in Fig. 1. The inset shows the relation between the excitation frequency and the propagation distance for the occurrence of the crossover between σ_l^s and p_l.

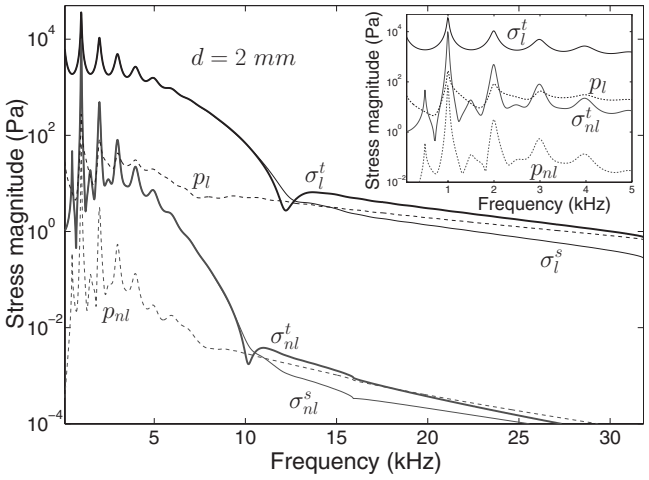


FIG. 3. Linear and nonlinear stress magnitudes at $x=0$ for a 10 cm slab configuration. Same notations and parameters as in Fig. 1. The inset shows a zoom of the low-frequency region.

decrease in the semilog plot). At a distance of approximately 20 cm from the source a crossover occurs: the fluid and solid stresses are of comparable magnitude. For longer distances, the behavior of the total stress is determined dominantly by the fluid properties as it clearly deviates from the behavior of the pure elastic stress σ_l^e. The reason of this crossover is that the solid Biot wave is more damped than the fluid wave. Consequently, for long enough distances, the acoustic energy initially injected in the solid is strongly damped, and the only remaining wave, able to propagate in the fluid, releases energy to the solid frame by coupling. The response of the system cannot be accurately predicted by a model, neglecting the solid-fluid interactions, which justifies the use of the Biot theory.

C. Harmonic generation in the slab configuration

The slab configuration consisting of an excitation boundary at $x=-d$ and a rigid boundary at $x=0$ is now investigated. The slab configuration qualitatively corresponds to the one used in some available experimental results.[48,51] The solid and total displacements are imposed and equal at $x=-d$ and vanish at the rigid boundary $x=0$. They can be written using trigonometric functions as

$$u^s = (A_1 \sin(k_1 x) + A_2 \sin(k_2 x) + B_1 \cos(k_1 x) + B_2 \cos(k_2 x))e^{j\omega t}, \tag{22}$$

$$u^t = (\mu_1 A_1 \sin(k_1 x) + \mu_2 A_2 \sin(k_2 x) + \mu_1 B_1 \cos(k_1 x) + \mu_1 B_2 \cos(k_2 x))e^{j\omega t}. \tag{23}$$

Both displacements are zero at $x=0$ and $B_1=B_2=0$. At $x=-d$,

$$u_0 = -A_1 \sin(k_1 d) - A_2 \sin(k_2 d), \tag{24}$$

$$u_0 = -\mu_1 A_1 \sin(k_1 d) - \mu_2 A_2 \sin(k_2 d), \tag{25}$$

which leads to

$$A_1 = \frac{(\mu_2 - 1)u_0}{(\mu_1 - \mu_2)\sin(k_1 d)}, \quad A_2 = \frac{(\mu_1 - 1)u_0}{(\mu_2 - \mu_1)\sin(k_2 d)}. \tag{26}$$

Note that the solid and total displacements can be written with Eq. (14) representation as

$$u_1' = \frac{A_1}{2i} = -u_1, \quad u_2' = \frac{A_2}{2i} = -u_2. \tag{27}$$

The particular solution can be obtained with the method in Appendix A. The general solution corresponds to a zero nonlinear displacement at $x=-d$ and $x=0$, and its amplitudes are solutions of the system

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ \ddot{\mu}_1 & \ddot{\mu}_2 & \ddot{\mu}_1 & \ddot{\mu}_2 \\ e_1 & e_2 & e_1' & e_2' \\ \ddot{\mu}_1 e_1 & \ddot{\mu}_2 e_2 & \ddot{\mu}_1 e_1' & \ddot{\mu}_2 e_2' \end{bmatrix} \begin{Bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_1' \\ \gamma_2' \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^{4} \alpha_i + \alpha_i' \\ \sum_{i=1}^{4} \beta_i + \beta_i' \\ \sum_{i=1}^{4} \alpha_i e^{2jk_i d} + \alpha_i' e^{-2jk_i d} \\ \sum_{i=1}^{4} \beta_i e^{2jk_i d} + \beta_i' e^{-2jk_i d} \end{Bmatrix}. \tag{28}$$

It is straightforward to find numerical values of the amplitudes γ_i using Cramer's method.

In Fig. 3, the stress magnitude at $x=0$ for a slab of 10 cm thickness is plotted as a function of frequency. The main difference compared to the semi-infinite case is the presence of resonances both for the linear fluid pressure p_l and the solid stress σ_l^s, originating from the limited size of the slab. The first resonance of the total stress σ_l^t occurs at frequency of 1 kHz, which corresponds to the longitudinal velocity in the solid frame over twice the slab thickness. The higher

resonance frequencies are almost multiple of the first resonance frequency but not exactly due to the velocity dispersion (see the zoom in the low-frequency region in the inset). Due to the increasing attenuation with frequency in the solid frame, these resonances have decreasing maximum values with increasing frequency. For the linear fluid pressure p_l, some resonances also occur, but due to the higher contribution of the solid stress compared to the fluid pressure at these low frequencies, there is a competition between the resonances associated with the equivalent fluid slab and the resonances corresponding to the solid frame slab. This results in asymmetric resonance curves, with lower quality factors than for the solid stress. The other features of these transmitted stresses through a slab of porous medium are comparable to the semi-infinite case in Fig. 1. Note that close to the resonance frequencies, for which the stress magnitudes are higher than in the semi-infinite case, the assumptions of the developed theory could be locally violated. One should notice that these results are in qualitative agreement with experimental results obtained in the same configuration in the context of acoustic probing of the granular compaction process.[48,51] First, the acoustic transfer function between two transducers placed on opposite sides of the granular container (the present granular slab) has the same shape as the total stress σ_l^t, including the local minimum observed at the crossover for the characteristic frequency f_c. The compaction process, which consists in applying vertical discrete vibrations to a cell containing a packing of beads in order to reduce its volume, modifies strongly the solid frame elastic properties but only slightly for the geometrical properties (few %). What is observed along a compaction process is a strong increase in the acoustic transfer function magnitude of the medium slab for the frequencies lower than f_c, corresponding to a region where the solid stress dominates over the fluid pressure, while the acoustic transfer function amplitude for frequencies higher than f_c slightly diminishes. In this higher-frequency region, the fluid pressure dominates over the solid stress, and the decrease in the acoustic transfer function magnitude is explained by the few % change in the geometrical properties of the packing.

Considering the nonlinear contributions, we remind that the frequency axis corresponds to the fundamental frequency at ω; therefore, the nonlinear magnitude plotted at frequency of 5 kHz, for instance, corresponds to the one of the second harmonic at 10 kHz. In this respect, the observed resonances for the nonlinear stresses are located roughly at each 500 Hz, at least for the first ones, with differences in their magnitudes (see the inset in Fig. 3). There is actually a difference when both the fundamental wave at frequency ω is at a resonance and the second harmonic wave at 2ω is also at a resonance compared to the case where only the wave at 2ω is at a resonance of the slab. When the fundamental wave is at a resonance of the slab, the nonlinear source term in Eq. (12) is also at a local maximum in frequency, and the second harmonic wave is efficiently generated. Moreover, due to the almost regularly spaced resonances, the second harmonic wave is itself at a resonance of the slab. This results in a strong resonance for the second harmonic wave, such as frequencies of 1, 2, and 3 kHz. When the second harmonic wave is at a resonance of the slab but not the fundamental wave, the nonlinear source term is not maximum and the second harmonic is less efficiently generated; being at a resonance of the slab, it still exhibits a resonance (fundamental frequencies of 500 Hz, 1.5 kHz, and 2.5 kHz, for instance) but smaller than in the previous case.

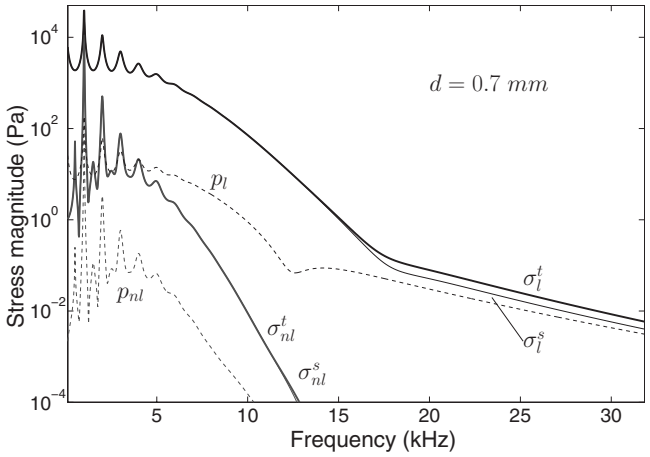


FIG. 4. Stress magnitude at $x=0$ for the same conditions as in Fig. 3, except that the bead diameter is decreased down to $d=0.7$ mm.

IV. INFLUENCE OF SOME MEDIUM PARAMETERS

In this section, the influence of three medium parameters on the previous results is analyzed: the bead diameter, the fluid properties, and the static stress applied on the granular solid frame, with the fluid static pressure being kept the same.

A. Influence of the bead diameter

The role of the bead diameter on the previous results is difficult to understand completely. To the authors' knowledge, the relation between the bead size and the acoustic energy dissipation in the solid frame is not yet understood. In addition, the relation between the bead size and the velocity dispersion in three-dimensional disordered granular media is not straightforward. However, the bead diameter role on the other parameters of the Biot model is quite well understood in the literature and follows the scaling laws given in Appendix B. Consequently, we choose to keep the same attenuation law in the solid frame and to scale the velocity dispersion law according to its chosen expression for this article, which depends on the bead diameter (see Appendix B). The nonlinear quadratic parameter D is considered independent of the bead diameter. Considering the current lack of understanding on this aspect, the possible dependence on the bead diameter is beyond the scope of the present article. In the assumptions of the developed model, a modification of D would have the only effect to shift up or down the nonlinear stress magnitudes proportional to D.

In Fig. 4, the stress magnitude at $x=0$ is plotted for the slab configuration as in Fig. 3 for a bead diameter $d=0.7$ mm. The main effect of decreasing the bead diameter is observed on the frequency region where the fluid pressure rules the total stress behavior, i.e., above the characteristic

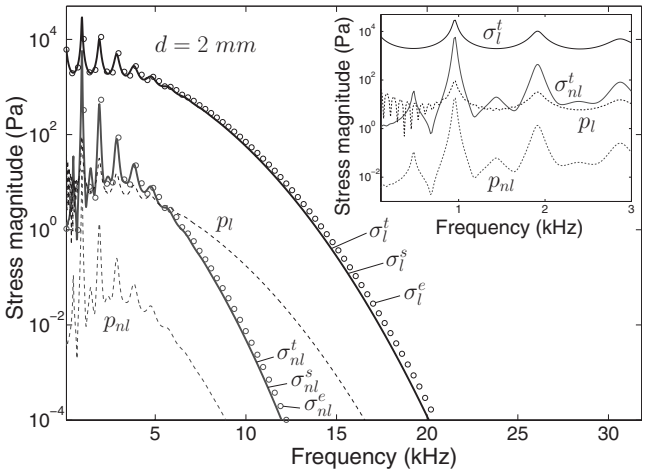


FIG. 5. Stress magnitudes in the same configuration as in Fig. 3 except that the saturating fluid is water. The inset is a zoom of the low-frequency region, showing oscillations of the fluid pressure magnitude.

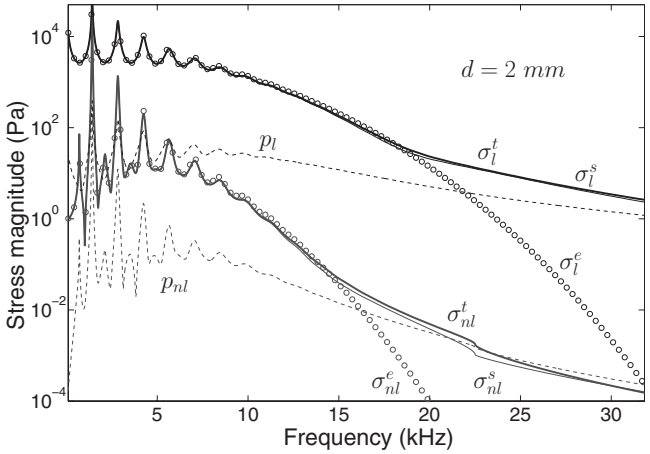


FIG. 6. Stress magnitudes in the same configuration as in Fig. 3 except that the elastic properties of the solid frame are modified due to the application of a stronger static solid stress (100 kPa).

frequency f_c. Note that in this region, in contrast to Fig. 3, the solid stress still dominates over the fluid pressure. The resistivity $\sigma \sim d^{-2}$ is increased by one order of magnitude when d is changed from 2 to 0.7 mm (see Appendix B) and the characteristic lengths decrease. This ensures a stronger attenuation of the acoustic energy in the fluid. This is clearly visible when comparing the solid and total stress magnitudes in Figs. 3 and 4 above the characteristic frequency. This behavior has also been confirmed in experiments.[48] For the second harmonic generation, the same remarks apply.

B. Influence of the saturating fluid properties

In order to modify the fluid properties and to stay in a realistic situation from the point of view of the possible experiments, we change the air considered previously for water. We neglect the possible effects of water on the solid frame properties, including the possible lubrication at the contacts, and the consequently modified attenuation law inside the solid frame.

In this case, the main difference is associated with the appearance of a strong asynchronism between the wave velocity in the equivalent fluid, close to that of water ($\sim$1500 m/s), and the much lower wave velocity in the solid frame ($\sim$200 m/s). This was not the case for air where both velocities were comparable.

The consequence is the occurrence of oscillations in the fluid pressure magnitude as a function of frequency in the low-frequency region of the presented results (see the inset in Fig. 5). These oscillations are visible here due to the better coupling between the solid frame and the fluid than in Fig. 3 for air, and the characteristic distance between two successive minima is well explained by the difference in the two wave numbers k_1 and k_2.

Another feature related to the saturating fluid properties is associated with the efficient coupling between water and the granular solid frame. In contrast to Fig. 3, there is no crossover between a region ruled by the solid stress behavior and a region ruled by the fluid pressure behavior as a function of frequency. Due to the stronger coupling, the attenuation in the solid frame (the one observed for the elastic stress σ_l^e) rules the attenuation of the fluid pressure too. Consequently, the solid stress dominates the fluid pressure in magnitude for the whole frequency range in Fig. 5 and is close to the elastic stress. This observation should be checked carefully in experiments in the future because it is in contradiction with the idea that for water-saturated granular media, most of the acoustic energy propagates through water.

C. Influence of the static solid stress

In several existing experimental results, a static stress of the order of 100 kPa is applied on the granular slab.[19,13,14] In this case, while the geometrical properties of the packing are almost not modified, the elastic properties of the solid frame are strongly different. The acoustic attenuation decreases both due to decreasing dissipation and decreasing scattering, and the wave velocity increases with some power law, mainly due to the nonlinearity of the contacts.[18,19,7,14] Consequently, in order to take into account an increase in the applied static stress on the solid frame properties, we keep constant the geometrical properties and modify only the real and imaginary parts of the elastic modulus λ_c for the solid frame. We choose $\Re e(\lambda_c)$ two times larger than previously mentioned (corresponding roughly to an increase of one order of magnitude of the applied static stress) and a two times smaller $\Im m(\lambda_c)$. Each parameter is identical to that used in Fig. 3. Results are given in Fig. 6. The main effect of increasing the static solid stress is to shift to higher frequencies the occurrence of the crossover between the frequency regions ruled by the solid stress and the fluid pressure behaviors. It means that the higher the static solid stress is, the wider the frequency region dominated by the solid stress is. This effect has been confirmed by recent experiments.[48,51]

V. CONCLUSIONS

An extension of the Biot theory for nonlinear propagation in unconsolidated model granular media or more generally in porous media supporting Biot waves is presented. The method of successive approximations is used to determine

the nonlinear solutions for the particular case of a quadratic nonlinearity of the solid. Both linear and second harmonic results can be used in future investigations on the acoustics of granular media. In particular, it has been shown that there exists a crossover frequency between a solid-controlled and a fluid-controlled behavior. This result can guide future experiments on the probing of the acoustical properties of granular packings, especially in the context where the linear or nonlinear elastic properties of the solid frame are studied. The influence of some medium parameters (bead diameter, saturating fluid, and application of a static solid stress) has also been studied.

Future works among which some are currently in progress concern the study of the nonlinear self-demodulation process or difference frequency generation, the self-action process, and the consideration of other types of nonlinearities in the solid or in the fluid. The application of the developed approach to other porous media including polymer foams, trabecular bone,[46] or other media supporting Biot waves is in principle possible.

ACKNOWLEDGMENTS

We would like to thank D. Lafarge, V. Gusev, and J.-F. Allard for stimulating discussions. This work has been supported by ANR project "grANuLar" under Grant No. NT05-3_41989.

APPENDIX A: PARTICULAR SOLUTION

This appendix presents a method to calculate the particular solution of Eq. (12) in the one-dimensional case.

1. Case of the porous media

The spatial derivatives of the solid displacement field [Eq. (14)] are

$$\frac{\partial u_l}{\partial x} = -jk_1u_1e^{-jk_1x} - jk_2u_2e^{-jk_2x} + jk_1u_1'e^{jk_1x} + jk_2u_2'e^{jk_2x}, \tag{A1a}$$

$$\frac{\partial^2 u_l}{\partial x^2} = -k_1^2u_1e^{-jk_1x} - k_2^2u_2e^{-jk_2x} - k_1^2u_1'e^{jk_1x} - k_2^2u_2'e^{jk_2x}. \tag{A1b}$$

Hence, with $2k_3=k_1+k_2$ and $2k_4=k_1-k_2$,

$$\begin{aligned}\frac{\partial u_l}{\partial x}\frac{\partial^2 u_l}{\partial x^2} = {} & jk_1^3u_1^2e^{-2jk_1x} - jk_1^3u_1'^2e^{2jk_1x} + jk_2^3u_2^2e^{-2jk_2x} \\ & - jk_2^3u_2'^2e^{2jk_2x} + 2jk_1k_2k_3u_1u_2e^{-2jk_3x} \\ & - 2jk_1k_2k_3u_1'u_2'e^{2jk_3x} - 2jk_1k_2k_4u_1'u_2e^{-2jk_4x} \\ & + 2jk_1k_2k_4u_2'u_1e^{2jk_4x}.\end{aligned} \tag{A2}$$

The particular solution reads as

$$u_p^s = \alpha_0 + \sum_{i=1}^{8} \alpha_i e^{-2jk_ix}, \tag{A3}$$

$$u_p^t = \beta_0 + \sum_{i=1}^{8} \beta_i e^{-2jk_ix}. \tag{A4}$$

For $i>4$, $k_i=-k_{i-4}$, $\alpha_i=\alpha_i'$, and $\beta_i=\beta_i'$. As the product [Eq. (A2)] does not involve constant terms, one has

$$\ddot{\rho}_s\alpha_0 + \tilde{\gamma}\ddot{\rho}_E\beta_0 = 0, \quad \ddot{\gamma}\tilde{\rho}_{\mathrm{eq}}\alpha_0 + \ddot{\rho}_{\mathrm{eq}}\beta_0 = 0, \tag{A5}$$

and thereby $\alpha_0=\beta_0=0$. The other amplitudes are the solutions of the following system:

$$\begin{bmatrix} 4\omega^2\ddot{\rho}_s[\mathbf{I}_8] - 4\lambda_c[\mathbf{k}]^2 & 4\omega^2\ddot{\gamma}\ddot{\rho}_{\mathrm{eq}}[\mathbf{I}_8] \\ 4\omega^2\ddot{\gamma}\ddot{\rho}_{\mathrm{eq}}[\mathbf{I}_8] & 4\omega^2\ddot{\rho}_{\mathrm{eq}}[\mathbf{I}_8] - 4\ddot{K}_{\mathrm{eq}}[\mathbf{k}]^2 \end{bmatrix} \begin{Bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}' \\ \boldsymbol{\beta} \\ \boldsymbol{\beta}' \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{\nu} \\ \mathbf{0}_8 \end{Bmatrix}, \tag{A6}$$

where $[\mathbf{k}]$ is the diagonal matrix of the k_i wave numbers and $\boldsymbol{\nu}$ is defined by

$$\begin{aligned}\boldsymbol{\nu} = j\{ & k_1^3u_1^2, k_2^3u_2^2, 2k_1k_2k_3u_1u_2, -2k_1k_2k_4u_1'u_2, -k_1^3u_1'^2, \\ & -k_2^3u_2'^2, -2k_1k_2k_3u_1'u_2', 2k_1k_2k_4u_2'u_1\}^t.\end{aligned} \tag{A7}$$

It is possible to find analytical expressions of the solutions

$$\alpha_i = \frac{\nu_i}{4\omega^2\ddot{\rho}_{\mathrm{eq}} - 4\lambda_c k_i^2 - \dfrac{4\omega^2\ddot{\gamma}\ddot{\rho}_{\mathrm{eq}}}{4\omega^2\ddot{\rho}_{\mathrm{eq}} - 4\ddot{K}_{\mathrm{eq}}k_i^2}}, \tag{A8}$$

$$\beta_i = -\frac{4\omega^2\ddot{\gamma}\ddot{\rho}_{\mathrm{eq}}}{4\omega^2\ddot{\rho}_{\mathrm{eq}} - 4\ddot{K}_{\mathrm{eq}}k_i^2}\alpha_i. \tag{A9}$$

2. Case of the pure elastic solid

In the case of a pure elastic solid, the linear elastic displacement is

$$u_l^e = u_fe^{-jk_0x} + u_f'e^{jk_0x}, \tag{A10}$$

with $k_0=\omega\sqrt{\tilde{\rho}/\lambda_c}$. The partial derivatives of the displacement [Eq. (A10)] are

$$\frac{\partial u_l^e}{\partial x} = -jk_0u_fe^{-jk_0x} + jk_0u_f'e^{jk_0x}, \tag{A11a}$$

$$\frac{\partial^2 u_l^e}{\partial x^2} = -k_0^2u_fe^{-jk_0x} - k_0^2u_f'e^{jk_0x}, \tag{A11b}$$

where k_0 corresponds to the wave number of the compressional waves and

$$\frac{\partial u_l}{\partial x}\frac{\partial^2 u_l}{\partial x^2} = jk_0^3u_f^2e^{-2jk_0x} - jk_0^3u_f'^2e^{2jk_0x}. \tag{A12}$$

The particular solution is

$$u_p^s = \alpha_0e^{-2jk_0x} + \alpha_0'e^{2jk_0x}, \tag{A13}$$

where α_0 and α_0' are solutions of

$$(4\omega^2\rho_0 - 4\lambda_c k_0^2)(\alpha_0 e^{-2jk_0x} + \alpha_0' e^{2jk_0x}) = \alpha j k_0^3 u_f^2 e^{-2jk_0x} - jk_0^3 u_f'^2 e^{2jk_0x}. \quad \text{(A14)}$$

Finally, we obtain

$$\alpha_0 = \frac{\alpha j k_0^3 u_f^2}{4\omega^2\rho_0 - 4\lambda_c k_0^2}, \quad \text{(A15)}$$

$$\alpha_0' = -\frac{\alpha j k_0^3 u_f'^2}{4\omega^2\rho_0 - 4\lambda_c k_0^2}. \quad \text{(A16)}$$

APPENDIX B: PARAMETERS OF THE MODEL

This appendix provides the expressions of the inertial and constitutive parameters of the model for a circular frequency ω. The geometrical and physical parameters of the equivalent fluid model have been derived previously in the literature for air saturating disordered packings of solid spheres.[21] In the general case they depend on the fluid properties and on the bead diameter.

The density terms are

$$\rho_1 = (1-\phi)\rho_s, \quad \rho_2 = \phi\rho_0, \quad \rho_{12} = -\phi\rho_0(\alpha_\infty - 1), \quad \text{(B1)}$$

where ϕ is the porosity, ρ_s is the frame material density, ρ_0 is the interstitial fluid density, and α_∞ is the geometric tortuosity. ρ_{12} accounts for the interaction between the inertia forces of the solid and fluid phases. The apparent coupling density can be introduced in the form

$$\tilde{\rho}_{12} = \rho_{12} - \frac{\tilde{b}}{j\omega}.$$

The viscous effects are modeled through the following coefficient:

$$\tilde{b} = j\omega\phi\rho_0(\tilde{\alpha} - \alpha_\infty), \quad \text{(B2)}$$

where $\tilde{\alpha}$ is the dynamic tortuosity defined by

$$\tilde{\alpha} = 1 - \frac{j\phi\sigma}{\alpha_\infty\rho_0\omega}\sqrt{1 - \frac{4j\alpha_\infty^2\eta_a\rho_0\omega}{(\sigma\Lambda\phi)^2}}. \quad \text{(B3)}$$

Here, σ is the flow resistivity, η_a is the dynamic viscosity of the saturating fluid, and Λ is the viscous characteristic length. The equivalent density $\tilde{\rho}_{\mathrm{eq}}$ and the coupling coefficient $\tilde{\gamma}$ are then given by

$$\tilde{\rho}_{\mathrm{eq}} = \frac{\rho_0\tilde{\alpha}}{\phi^2}, \quad \tilde{\gamma} = \phi\left(\frac{\tilde{\rho}_{12}}{\rho_0\tilde{\alpha}} - \frac{1-\phi}{\phi}\right). \quad \text{(B4)}$$

The thermal effects are introduced through the dynamic compressibility

$$K_f = \frac{\phi\gamma P_0}{\gamma - (\gamma-1)\left[1 + \frac{8\eta_a}{j\Lambda'\mathrm{Pr}\omega\rho_0}\sqrt{1 + \frac{j\rho_0\omega\mathrm{Pr}\Lambda'^2}{16\eta_a}}\right]^{-1}}, \quad \text{(B5)}$$

where Λ' is the thermal characteristic length, Pr is the Prandtl number, P_0 is the ambient pressure, and γ is the specific heat ratio of the fluid.

The compressibility of the fluid M is then obtained as

$$M = \frac{K_r^2}{K_r(1 + \phi K_r/(K_f - 1)) - K_s}. \quad \text{(B6)}$$

The porosity ϕ of the medium can be evaluated for random close packings of spheres[3] as $\phi \simeq 0.4 \pm 0.04$. The tortuosity for disordered packings of beads is evaluated as $\alpha_\infty = 1.36$ and is independent of the bead diameter. The thermal characteristic length reads as $\Lambda' = \phi d/3(1-\phi)$. The viscous characteristic length is estimated as $\Lambda \simeq \Lambda'/3$. The flow resistivity is $\sigma = 4\eta_a F/\Lambda^2$, where $F = 3.4$ is a formation factor independent of the bead diameter.

In summary, the bead size scaling laws for the equivalent fluid parameters are $\alpha_\infty \sim \phi \sim F \sim d^0$, $\Lambda \sim \Lambda' \sim d^1$, and $\sigma \sim d^{-2}$.

Concerning the parameters of the solid frame, they are much less characterized and can vary by several orders of magnitude depending on the applied external conditions and the state of the packing. They have been chosen here to correspond to the experimental conditions of some available experimental results,[48,51] i.e., a weak external static stress of ~ 3 kPa, giving a measured wave velocity of ~ 200 m/s at low frequencies. Due to the existing velocity dispersion in experiments, we considered an elastic modulus for the solid frame of the form $\lambda_c = \lambda_{c0}[(\omega/\omega_c)/\arcsin(\omega/\omega_c)]^2 + i\xi\lambda_{c0}\omega$. The real part corresponds to the dispersion relation of a one-dimensional granular chain[52] with $\omega_c = d/2\sqrt{\lambda_{c0}/\rho_s}$, $\lambda_{c0} = 60$ MPa, and $\rho_s = 1500$ kg/m^3. The imaginary part is chosen to be proportional to the frequency in order to increase with frequency. Another dependency can in principle be chosen. A value of $\xi = 4 \times 10^{-6}$ provides an attenuation close to the one observed in experiments.[48,51] Note that these dispersion and attenuation laws may not fulfill the Kramers–Kronig relations and this could provide causality problems if the results obtained in the frequency domain in this article are transformed into the time domain.

The two compressional waves of the porous medium are defined by their wave number k_i and the ratio μ_i of the total displacement over the solid one. They are defined by

$$k_i^2 = \frac{(\delta_{s2}^2 + \delta_{\mathrm{eq}}^2) \pm \sqrt{(\delta_{s2}^2 + \delta_{\mathrm{eq}}^2)^2 - 4\delta_{\mathrm{eq}}^2\delta_{s1}^2}}{2}, \quad \text{(B7)}$$

with

$$\delta_{\mathrm{eq}} = \omega\sqrt{\frac{\tilde{\rho}_{\mathrm{eq}}}{M}}, \quad \delta_{s1} = \omega\sqrt{\frac{\tilde{\rho}}{\lambda_c}}, \quad \delta_{s2} = \omega\sqrt{\frac{\tilde{\rho}_s}{\lambda_c}}, \quad \text{(B8)}$$

where

$$\tilde{\rho} = \rho_1 - \tilde{\rho}_{12} - \frac{\tilde{\rho}_{12}^2}{\rho_2 - \tilde{\rho}_{12}}, \quad \tilde{\rho}_s = \tilde{\rho} + \tilde{\gamma}^2\tilde{\rho}_{\mathrm{eq}}, \quad \text{(B9)}$$

$$\mu_i = \tilde{\gamma}\frac{(k_i^2 - \delta_{s2}^2)}{\delta_{s2}^2 - \delta_{s1}^2} = \tilde{\gamma}\frac{\delta_{\mathrm{eq}}^2}{k_i^2 - \delta_{\mathrm{eq}}^2}. \quad \text{(B10)}$$

[1] P. G. de Gennes, "Granular matter: A tentative view," Rev. Mod. Phys. **71**, S374–S382 (1999).
[2] H. M. Jaeger and S. R. Nagel, "Physics of the granular state," Science **255**, 1523–1531 (1992).

[3]P. Richard, M. Nicodemi, R. Delannay, P. Ribière, and D. Bideau, "Slow relaxation and compaction of granular systems," Nature Mater. **4**, 121–128 (2005).
[4]H. A. Makse, N. Gland, D. L. Johnson, and L. Schwartz, "Granular packings: Nonlinear elasticity, sound propagation, and collective relaxation dynamics," Phys. Rev. E **70**, 061302 (2004).
[5]N. Xu, N. Wyart, A. J. Liu, and S. R. Nagel, "Excess vibrational modes and the boson peak in model glasses," Phys. Rev. Lett. **98**, 175502 (2007).
[6]C. S. O'Hern, L. E. Silbert, A. J. Liu, and S. R. Nagel, "Jamming at zero temperature and zero applied stress: The epitome of disorder," Phys. Rev. E **68**, 011306 (2003).
[7]K. L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985).
[8]R. D. Mindlin and H. Deresiewicz, "Elastic spheres in contact under varying oblique forces," J. Appl. Mech. **21**, 327–344 (1953).
[9]I. Y. Belyaeva, V. Y. Zaitsev, and E. M. Timanin, "Experimental study of nonlinear elastic properties of granular media with nonideal packing," Sov. Phys. Acoust. **40**, 789–793 (1994).
[10]P. A. Johnson and X. Jia, "Nonlinear dynamics, granular media and dynamic earthquake triggering," Nature (London) **437**, 871–874 (2005).
[11]V. Tournat, B. Castagnède, V. E. Gusev, and Ph. Béquin, "Self-demodulation acoustic signatures for non-linear propagation in glass beads," C. R. Mec. **331**, 119–125 (2003).
[12]V. Tournat, V. E. Gusev, and B. Castagnède, "Subharmonics and noise excitation in transmission of acoustic wave through unconsolidated granular medium," Phys. Lett. A **326**, 340–348 (2004).
[13]V. Tournat, V. E. Gusev, V. Yu. Zaitsev, and B. Castagnède, "Acoustic second harmonic generation with shear to longitudinal mode conversion in granular media," Europhys. Lett. **66**, 798–804 (2004).
[14]V. Tournat, V. Yu. Zaitsev, V. E. Gusev, V. E. Nazarov, P. Béquin, and B. Castagnède, "Probing weak forces in granular media through nonlinear dynamic dilatancy: Clapping contacts and polarization anisotropy," Phys. Rev. Lett. **92**, 085502 (2004).
[15]V. Tournat, V. Yu. Zaitsev, V. E. Nazarov, V. E. Gusev, and B. Castagnède, "Experimental study of nonlinear acoustic effects in a granular medium," Acoust. Phys. **51**, 543–553 (2005).
[16]V. Yu. Zaitsev, V. E. Nazarov, V. Tournat, V. Gusev, and B. Castagnède, "Luxembourg-Gorky effect in a granular medium: Probing perturbations of the material state via cross-modulation of elastic waves," Europhys. Lett. **70**, 607–613 (2005).
[17]D. M. Mueth, H. M. Jaeger, and S. R. Nagel, "Force distribution in a granular medium," Phys. Rev. E **57**, 3164–3169 (1998).
[18]J. D. Goddard, "Nonlinear elasticity and pressure-dependent wave speeds in granular media," Proc. R. Soc. London, Ser. A **430**, 105–131 (1990).
[19]X. Jia, C. Caroli, and B. Velicky, "Ultrasound propagation in externally stressed granular media," Phys. Rev. Lett. **82**, 1863–1866 (1999).
[20]C. Josserand, A. V. Tkachenko, D. M. Mueth, and H. M. Jaeger, "Memory effects in granular materials," Phys. Rev. Lett. **85**, 3632–3635 (2000).
[21]J.-F. Allard, M. Henry, J. Tizianel, L. Kelders, and W. Lauriks, "Sound propagation in air-saturated random packings of beads," J. Acoust. Soc. Am. **104**, 2004–2007 (1998).
[22]K. V. Horoshenkov and J. Swift, "The acoustic properties of granular materials with pore size distribution close to log-normal," J. Acoust. Soc. Am. **110**, 2371–2378 (2001).
[23]G. Pispola, K. V. Horoshenkov, and F. Asdrubali, "Transmission loss measurement of consolidated granular media," J. Acoust. Soc. Am. **117**, 2716–2719 (2005).
[24]K. Attenborough, "Models for the acoustical characteristics of air filled granular materials," Acta Acust. **1**, 213–226 (1993).
[25]O. Umnova, K. Attenborough, and K. M. Li, "A cell model for the acoustical properties of packings of spheres," J. Acoust. Soc. Am. **87**, 226–235 (2001).
[26]O. Umnova, K. Attenborough, H.-C. Shin, and A. Cummings, "Response of multiple rigid porous layers to high levels of continuous acoustic excitation," J. Acoust. Soc. Am. **116**, 703–712 (2004).
[27]C. Ayrault and S. Griffiths, "Separation of viscothermal losses and scattering in ultrasonic characterization of porous media," Ultrasonics **45**, 40–49 (2006).
[28]H. P. Schriemer, M. L. Cowan, J. H. Page, P. Sheng, Z. Liu, and D. A. Weitz, "Energy velocity of diffusing waves in strongly scattering media," Phys. Rev. Lett. **79**, 3166–3169 (1997).
[29]A. Moussatov, B. Castagnède, and V. Gusev, "Observation of nonlinear interaction of acoustic waves in granular materials: Demodulation process," Phys. Lett. A **283**, 216–223 (2001).
[30]M. Biot, "Mechanics of deformation and acoustic propagation in porous media," J. Appl. Phys. **33**, 1482–1498 (1962).
[31]M. A. Biot, "Theory of propagation of elastic waves in a fluid-filled-saturated porous solid," J. Acoust. Soc. Am. **28**, 168–191 (1956).
[32]J. G. Berryman, "Confirmation of Biot's theory," Appl. Phys. Lett. **37**, 382–384 (1980).
[33]T. J. Plona, "Observation of a second bulk compressional wave in a porous medium at ultrasonic frequencies," Appl. Phys. Lett. **36**, 259–261 (1980).
[34]J. L. Auriault, L. Borne, and R. Chambon, "Dynamics of porous saturated media, checking of the generalized law of Darcy," J. Acoust. Soc. Am. **77**, 1641–1650 (1985).
[35]R. Burridge and J. B. Keller, "Poroelasticity equations derived from microstructure," J. Acoust. Soc. Am. **70**, 1140–1146 (1981).
[36]S. R. Pride and J. G. Berryman, "Connecting theory to experiments in poroelasticity," J. Mech. Phys. Solids **46**, 719–747 (1998).
[37]S. R. Pride, A. F. Gangi, and F. D. Morgan, "Deriving the equations of motion for porous isotropic media," J. Acoust. Soc. Am. **92**, 3278–3290 (1992).
[38]J. F. Allard, *Propagation of Sound in Porous Media, Modelling Sound Absorbing Materials* (Elsevier Applied Science, London and New York, 1993).
[39]Y. Champoux and J. F. Allard, "Dynamic tortuosity and bulk modulus in air-saturated porous media," J. Appl. Phys. **70**, 1975–1979 (1991).
[40]D. L. Johnson, J. Koplik, and R. Dashen, "Theory of dynamic permeability and tortuosity in fluid-saturated porous media," J. Fluid Mech. **176**, 379–402 (1987).
[41]D. Lafarge, P. Lemarinier, J. F. Allard, and V. Tarnow, "Dynamic compressibility of air in porous structures at audible frequencies," J. Acoust. Soc. Am. **102**, 1995–2006 (1997).
[42]J. M. Carcione, "Wave propagation in anisotropic, saturated porous media: Plane-wave theory and numerical simulation," J. Acoust. Soc. Am. **99**, 2655–2666 (1996).
[43]O. Coussy, *Mechanics of Porous Continua* (John Wiley, New York, 1995).
[44]M. A. Biot, "Nonlinear and semilinear rheology of porous solids," J. Geophys. Res. **78**, 4924–4937 (1973).
[45]D. M. Donskoy, K. Khashanah, and T. G. McKee, Jr., "Nonlinear acoustic waves in porous media in the context of Biot's theory," J. Acoust. Soc. Am. **102**, 2521–2528 (1997).
[46]G. Renaud, S. Calle, J.-P. Remenieras, and M. Defontaine, "Exploration of trabecular bone nonlinear elasticity using time-of-flight modulation," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **55**, 1497–1507 (2008).
[47]O. Dazel, B. Brouard, C. Depollier, and S. Griffiths, "An alternative Biot's displacement formulation for porous materials," J. Acoust. Soc. Am. **121**, 3509–3516 (2007).
[48]C. Inserra, "Characterization of granular compaction by linear and nonlinear acoustic methods," Ph.D. thesis, Université du Maine, Le Mans, France (2007).
[49]M. F. Hamilton and C. L. Morfey, "Model equations," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, San Diego, CA, 1998), pp. 52–56.
[50]B. K. Novikov, O. V. Rudenko, and V. I. Timoshenko, *Nonlinear Underwater Acoustics* (American Institute of Physics, New York, 1987).
[51]C. Inserra, V. Tournat, and V. E. Gusev, "Characterization of granular compaction by nonlinear acoustic resonance method," Appl. Phys. Lett. **92**, 191916 (2008).
[52]V. Tournat, V. E. Gusev, and B. Castagnède, "Self-demodulation of elastic waves in a 1D granular chain," Phys. Rev. E **70**, 056603 (2004).

Acoustic microstreaming around a gas bubble

Alexander A. Doinikov[a] and Ayache Bouakaz
INSERM U930 CNRS ERL 3106, Université François Rabelais, CHU Bretonneau, 2 Boulevard Tonnellé, 37044 Tours Cedex 9, France

(Received 30 September 2009; revised 7 December 2009; accepted 7 December 2009)

The problem of acoustic microstreaming that develops around a gas bubble in an ultrasound field is considered. It is shown that the solutions obtained previously by Wu and Du [(1997). J. Acoust. Soc. Am. **101**, 1899–1907], which are based on the assumption that viscous effects are essential only within a thin boundary layer while beyond the boundary layer the liquid can be considered to be inviscid, lead to a severe underestimation of the power of acoustic streaming. An improved theory is suggested that corrects the errors of the previous theory and extends its limits. The proposed theory treats the entire bulk of the liquid outside the bubble and the gas inside the bubble as viscous heat-conducting fluids. No restrictions are imposed on the size of the bubble relative to the viscous, thermal, and sound wavelengths in the ambient liquid and those in the internal gas medium. All modes of the bubble's motion (volume pulsation, translation, and shape oscillations) are taken into account. Expressions for the radial and tangential stresses produced by the acoustic streaming are also derived. Numerical examples for parameters of interest are presented.
 [DOI: 10.1121/1.3279793]

PACS number(s): 43.25.Nm, 43.25.Yw [RMW] Pages: 703–709

I. INTRODUCTION

Interest to acoustic streaming that develops around a gas bubble in an ultrasound field is supported by the variety of technological and biomedical applications of this phenomenon. Extensive theoretical work has been done to calculate the velocity field of this kind of acoustic streaming (Nyborg, 1958, 1965; Elder, 1959; Davidson and Riley, 1971; Lee and Wang, 1990; Wu and Du, 1997; Longuet-Higgins, 1998). Wu and Du (1997) developed an analytical approach to calculate the microstreaming velocity for an isolated free gas bubble, taking into account two modes of the bubble's motion: a monopole (volume pulsation) and a dipole (translational harmonic vibration). Their theory also assumes that (1) the wavelength of the incident sound field is much greater than the radius of the bubble, (2) the thickness of the viscous boundary layer surrounding the bubble is much smaller than the radius of the bubble, (3) the surface tension on the gas-liquid interface is neglected, and (4) the surrounding liquid is non-heat-conducting. Recently, the same approach has been applied to an encapsulated bubble (Liu and Wu, 2009).

The method of calculation of acoustic streaming is known to consist of two stages. The linear (first-order) scattered field is first calculated. This field is then used to solve the equations of acoustic streaming. These are time-averaged equations of fluid motion that are taken with accuracy up to second-order terms in the amplitude of the incident acoustic wave. Wu and Du (1997) assumed that viscous effects are essential only within a thin boundary layer while beyond the boundary layer the fluid can still be considered to be inviscid. Based on this approach, they obtained approximate solutions for the first-order velocity field, which are valid only inside the boundary layer, and then used these to find solutions to the equations of acoustic streaming.

In this paper, we show that the approach suggested by Wu and Du (1997) may be applied only to very low-viscosity fluids, whose viscosity is far less than that of water. Under normal conditions, their solutions lead to a severe underestimation of the power of acoustic streaming. Instead, we suggest a theory that treats the entire bulk of the liquid outside the bubble and the gas inside the bubble as viscous heat-conducting fluids. Moreover, no restrictions are imposed on the size of the bubble relative to the viscous, thermal, and sound wavelengths, and all modes of the bubble's motion, including the surface ones, are taken into account.

II. ANALYSIS OF WU AND DU'S SOLUTIONS

To understand why Wu and Du's (1997) solutions lead to underestimating the velocity field of acoustic streaming, let us consider the expression that was obtained in their work for the first-order tangential velocity outside the bubble. According to Eq. (10) of their paper, this velocity in the field of a plane traveling wave is given by

$$u_{1\theta o} = -3u_0 \sin\theta \exp(i\omega t)\left[\frac{\tilde{\rho}}{\rho} + \left(\frac{1-\tilde{\rho}/\rho}{1+\eta\tilde{\delta}_v/\tilde{\eta}\delta_v}\right) \times \exp[(i-1)(r-R_0)/\delta_v]\right], \quad (1)$$

where r and θ are the coordinates of the spherical coordinate system (r,θ,ε) whose origin is at the equilibrium center of the bubble and the z axis (θ=0) lies in the direction of the incident wave propagation, u_0 is the velocity amplitude of the incident wave, ω is the angular frequency, ρ is the equilibrium density of the ambient liquid, η is the shear viscosity

[a]Author to whom correspondence should be addressed. Electronic mail: doinikov@bsu.by

of the liquid, $\delta_v=\sqrt{2\nu/\omega}$ is the viscous penetration depth in the liquid (the thickness of the viscous boundary layer), $\nu=\eta/\rho$ is the kinematic viscosity of the liquid, the tilde denotes the same quantities for the gas inside the bubble, and R_0 is the equilibrium radius of the bubble.

The maximum amplitude of $u_{1\theta o}$ is reached on the bubble surface, at $r=R_0$,

$$u_{1\theta o}(r=R_0)=-3u_0\sin\theta\exp(i\omega t)\left(\frac{\tilde{\rho}}{\rho}+\frac{1-\tilde{\rho}/\rho}{1+\eta\tilde{\delta}_v/\tilde{\eta}\delta_v}\right). \tag{2}$$

Let us compare this expression to what is given by exact solutions for the first-order scattered field from a gas bubble in a viscous heat-conducting liquid. Such solutions can be found in Doinikov, 1997a, 1997c. We call them exact because they are valid within the entire bulk of the ambient liquid and the entire bulk of the gas inside the bubble, whereas the viscous terms of Wu and Du's (1997) solutions are valid only within the thin boundary layers outside and inside the bubble.

From the solutions obtained by Doinikov (1997a, 1997c), it follows that, in the limit used by Wu and Du (1997), i.e., $\delta_v/R_0,\tilde{\delta}_v/R_0\ll 1$ and $kR_0,\tilde{k}R_0\ll 1$, where $k=\omega/c$ and $\tilde{k}=\omega/\tilde{c}$ are, respectively, the wave numbers in the liquid and in the gas, the first-order tangential velocity in the field of a plane traveling wave, taken on the bubble surface, are written as

$$v_\theta^{(1)}(r=R_0)=-3u_0\sin\theta\exp(-i\omega t)\left[\frac{1}{3}-\frac{i\alpha_{11}}{(kR_0)^3}-\frac{i\alpha_{31}}{kR_0}\exp[-(1-i)R_0/\delta_v]\right], \tag{3}$$

where the coefficients α_{11} and α_{13} are given by

$$\alpha_{11}=-\frac{i}{3}(kR_0)^3\left(\frac{1-\tilde{\rho}/\rho}{1+2\tilde{\rho}/\rho}\right)\left[1-\frac{3(1+i)(1-\tilde{\rho}/\rho)\delta_v}{(1+2\tilde{\rho}/\rho)R_0}\times\left(1+\frac{\rho\delta_v/\tilde{\rho}\tilde{\delta}_v}{1+(1+i)\eta\tilde{\delta}_v/\tilde{\eta}R_0}\right)^{-1}\right], \tag{4}$$

$$\alpha_{31}=ikR_0\left(\frac{1-\tilde{\rho}/\rho}{1+2\tilde{\rho}/\rho}\right)\exp[(1-i)R_0/\delta_v]\times\left(1+\frac{\rho\delta_v/\tilde{\rho}\tilde{\delta}_v}{1+(1+i)\eta\tilde{\delta}_v/\tilde{\eta}R_0}\right)^{-1}. \tag{5}$$

Equation (3) is calculated from Eqs. (13), (33), (35), and (37) of Doinikov, 1997a, and Eqs. (4) and (5) are taken from Appendix C of Doinikov, 1997c. It should be noted that Eqs. (4) and (5) have been derived for a liquid drop suspended in another liquid, assuming both liquids to be viscous, heat-conducting, and *compressible*. Therefore, it is clear that the same expressions can be applied to a gas bubble as well, just substituting proper values for the internal density $\tilde{\rho}$ and the internal viscosity $\tilde{\eta}$.

Substituting Eqs. (4) and (5) into Eq. (3) and taking into account that the gas density and the gas viscosity are small compared to those of the liquid, one has, with accuracy up to $O(\delta_v/R_0)$,

$$v_\theta^{(1)}(r=R_0)=-3u_0\sin\theta\exp(-i\omega t)\left[\frac{\tilde{\rho}}{\rho}+\frac{1-\tilde{\rho}/\rho}{1+\eta\tilde{\delta}_v/\tilde{\eta}\delta_v}\left(1+(1+i)\frac{\delta_v}{R_0}\sqrt{\frac{\rho\eta}{\tilde{\rho}\tilde{\eta}}}\right)\right]. \tag{6}$$

Comparison of Eqs. (2) and (6) reveals that Wu and Du's (1997) theory ignores the term $\Lambda=(1+i)(\delta_v/R_0)\sqrt{\rho\eta/\tilde{\rho}\tilde{\eta}}$ with respect to unity. Let us estimate the magnitude of this term, assuming that the ambient liquid is water, the gas inside the bubble is air, $f=100$ kHz, and $R_0=33$ μm [these parameters are used in the paper by Wu and Du (1997)]. The result is $|\Lambda|\approx 16.5$. For biomedical applications, smaller bubbles and higher frequencies are of interest. If we set, for example, $f=1$ MHz, and $R_0=5$ μm, we obtain $|\Lambda|\approx 34.3$. Thus Eq. (2) loses the dominant viscous term. Note also that we obtain these values of Λ in the limit $\delta_v/R_0\ll 1$. For larger values of δ_v/R_0, the disagreement between Eq. (2) and the exact solution will be even greater. The same problem is with the radial velocity component obtained by Wu and Du (1997). These first-order solutions are then used to calculate the velocity field of acoustic streaming. As a result, the power of acoustic streaming given by Wu and Du's (1997) theory is found to be considerably understated.

III. IMPROVED THEORY

Doinikov (1997a, 1997b, 1997c) published a series of papers where the acoustic radiation force on a spherical particle of arbitrary internal structure, suspended in a viscous heat-conducting fluid, is calculated. In the process of this calculation, solutions to the equations of acoustic streaming have been derived. In distinction to the previous studies, the above solutions impose no restrictions on the size of the particle with respect to the sound, viscous, and thermal wavelengths in the ambient fluid and take into account all modes of the particle's motion: volume pulsation, translation, and shape modes. The velocity field of acoustic streaming is expressed in terms of linear scattering coefficients which are determined by the type of the particle's internal medium. In Doinikov, 1997c, the linear scattering coefficients are derived assuming the internal medium to be a compressible viscous heat-conducting fluid. These results can be immediately used to calculate acoustic streaming around a gas bubble. In Secs. III A and III B, we present all the equations necessary for this calculation.

A. The first-order liquid velocity

The first-order velocity in the ambient liquid is represented as

$$\boldsymbol{v}^{(1)}=\boldsymbol{v}_I^{(1)}+\boldsymbol{v}_S^{(1)}, \tag{7}$$

where $\boldsymbol{v}_I^{(1)}$ and $\boldsymbol{v}_S^{(1)}$ are the velocities of the incident and the scattered sound wave fields. The velocity of the incident field

is defined as $\boldsymbol{v}_I^{(1)}=\nabla\varphi_I^{(1)}$ with the velocity potential given by

$$\varphi_I^{(1)}=\exp(-i\omega t)\sum_{n=0}^{\infty}A_n j_n(k_1 r)P_n(\cos\theta), \tag{8}$$

where $j_n(x)$ is the spherical Bessel function of order n, P_n is the Legendre polynomial of degree n, k_1 is the sound wave number, and r and θ are the coordinates of the spherical coordinate system (r,θ,ε) whose origin is at the equilibrium center of the particle under consideration and the z axis ($\theta=0$) lies in the direction of wave propagation. It should be emphasized that Eq. (8) describes an axially symmetric wave and the particle is assumed to be located on the axis of the wave. The coefficients A_n specify the type of incident wave. For a plane traveling wave, A_n are given by

$$A_n=A(2n+1)i^n, \tag{9}$$

where A denotes the amplitude of the velocity potential.

The sound wave number k_1 is defined by

$$k_1=\left\{\frac{\beta_1+\beta_2}{2}\left[1-\left(1-\frac{4i\omega\beta_1}{\gamma\chi(\beta_1+\beta_2)^2}\right)^{1/2}\right]\right\}^{1/2}, \tag{10}$$

where

$$\beta_1=\gamma\omega^2/(c^2\beta_3),\quad \beta_2=i\omega(\gamma-1+\beta_3)/(\gamma\chi\beta_3),$$

$$\beta_3=1-i\omega\gamma(\xi+4\eta/3)/(\rho c^2), \tag{11}$$

γ is the ratio of specific heats of the ambient liquid, χ is the thermal diffusivity, c is the sound speed, η is the liquid shear viscosity, ξ is the liquid bulk viscosity, and ρ is the liquid density at rest.

The velocity of the scattered field is represented as

$$\boldsymbol{v}_S^{(1)}=\nabla\varphi_S^{(1)}+\nabla\times\boldsymbol{\psi}_S^{(1)}, \tag{12}$$

where the scalar and vorticity velocity potentials are given by

$$\varphi_S^{(1)}=\exp(-i\omega t)\sum_{n=0}^{\infty}A_n[\alpha_{1n}h_n^{(1)}(k_1 r)+\alpha_{2n}h_n^{(1)}(k_2 r)]P_n(\cos\theta), \tag{13}$$

$$\boldsymbol{\psi}_S^{(1)}=\exp(-i\omega t)\boldsymbol{e}_\varepsilon\sum_{n=1}^{\infty}A_n\alpha_{3n}h_n^{(1)}(k_3 r)P_n^1(\cos\theta). \tag{14}$$

Here, α_{1n}, α_{2n}, and α_{3n} are the linear scattering coefficients, which are determined by the type of the internal medium of the particle (fluid, solid, etc.), $h_n^{(1)}(x)$ is the spherical Hankel function of the first kind, k_2 is the thermal wave number, $\boldsymbol{e}_\varepsilon$ is the unit azimuth vector of the spherical coordinate system introduced above, k_3 is the viscous wave number, and P_n^1 is the associated Legendre polynomial of the first order and degree n. Note that P_n^1 is defined here as

$$P_n^1(\mu)=(1-\mu^2)^{1/2}dP_n(\mu)/d\mu. \tag{15}$$

The wave numbers k_1 and k_3 are given by

$$k_2=\left\{\frac{\beta_1+\beta_2}{2}\left[1+\left(1-\frac{4i\omega\beta_1}{\gamma\chi(\beta_1+\beta_2)^2}\right)^{1/2}\right]\right\}^{1/2}, \tag{16}$$

$$k_3=\frac{1+i}{\delta_v}, \tag{17}$$

where $\delta_v=\sqrt{2\nu/\omega}$ is the viscous penetration depth in the ambient liquid and $\nu=\eta/\rho$ is the liquid kinematic viscosity.

From Eqs. (8), (13), and (14), the r and θ components of the first-order liquid velocity are found to be

$$v_r^{(1)}=e^{-i\omega t}\sum_{n=0}^{\infty}A_n\left[k_1 j_n'(k_1 r)+\alpha_{1n}k_1 h_n^{(1)\prime}(k_1 r)+\alpha_{2n}k_2 h_n^{(1)\prime}(k_2 r)+\frac{n(n+1)}{r}\alpha_{3n}h_n^{(1)}(k_3 r)\right]P_n(\cos\theta), \tag{18}$$

$$v_\theta^{(1)}=-e^{-i\omega t}\frac{1}{r}\sum_{n=1}^{\infty}A_n[j_n(k_1 r)+\alpha_{1n}h_n^{(1)}(k_1 r)+\alpha_{2n}h_n^{(1)}(k_2 r)+\alpha_{3n}(h_n^{(1)}(k_3 r)+k_3 r h_n^{(1)\prime}(k_3 r))]P_n^1(\cos\theta), \tag{19}$$

where the prime denotes the differentiation with respect to the argument in parentheses.

The calculation of acoustic streaming also requires the expression for the first-order perturbation of the liquid density. When the particle is absent, the liquid density perturbation is given by

$$\rho_I^{(1)}=-\frac{i\rho}{\omega}\nabla\cdot\boldsymbol{v}_I^{(1)}=\frac{i\rho k_1^2}{\omega}\varphi_I^{(1)}. \tag{20}$$

When the particle is present, the liquid density perturbation is written as

$$\rho^{(1)}=-\frac{i\rho}{\omega}\nabla\cdot\boldsymbol{v}^{(1)}=e^{-i\omega t}\frac{i\rho}{\omega}\sum_{n=0}^{\infty}A_n[k_1^2 j_n(k_1 r)+\alpha_{1n}k_1^2 h_n^{(1)}(k_1 r)+\alpha_{2n}k_2^2 h_n^{(1)}(k_2 r)]P_n(\cos\theta). \tag{21}$$

In Doinikov, 1997c, the expressions for the linear scattering coefficients are presented, which have been calculated assuming that the internal medium of the particle is a compressible viscous heat-conducting fluid, i.e., the particle may be a liquid drop or a gas bubble depending on the values of the internal density and viscosity used. These expressions are as follows.

The linear scattering coefficients α_{mn} $(m=1,2,3)$ are calculated as

$$\alpha_{mn}=\det(a_{ij}^{(nm)})/\det(a_{ij}^{(n)}), \tag{22}$$

where $a_{ij}^{(n)}$ is a square matrix $(i,j=1,2,3)$ and $a_{ij}^{(nm)}$ is the same matrix in which the mth column is replaced with the elements $b_i^{(n)}$. The matrix elements $a_{ij}^{(n)}$ and $b_i^{(n)}$ are calculated by

$$a_{ij}^{(n)}=\mu_{ij}^{(n)}h_n^{(1)}(x_j)+\nu_{ij}^{(n)}x_j h_n^{(1)\prime}(x_j), \tag{23}$$

$$b_i^{(n)}=-\mu_{i1}^{(n)}j_n(x_1)-\nu_{i1}^{(n)}x_1 j_n'(x_1), \tag{24}$$

where $x_j=k_j R_0$, R_0 is the equilibrium radius of the particle, and the functions $\mu_{ij}^{(n)}$ and $\nu_{ij}^{(n)}$ are defined by

$$\mu_{11}^{(n)}=f_1^{(n)}(k_1),\quad \mu_{12}^{(n)}=f_1^{(n)}(k_2),\quad \mu_{13}^{(n)}=n(n+1)(g-g_{1n}),$$

$$\mu_{21}^{(n)} = f_2^{(n)}(k_1), \quad \mu_{22}^{(n)} = f_2^{(n)}(k_2), \quad \mu_{23}^{(n)} = n(n+1)[1-\rho/\tilde{\rho} - (n^2+n-1)g],$$

$$\mu_{31}^{(n)} = \mu_{32}^{(n)} = 1-\rho/\tilde{\rho} + g(g_{2n}+1-n-n^2), \quad \mu_{33}^{(n)} = 1+n(n+1)(g-g_{1n}) - (1+g_{2n})[\rho/\tilde{\rho} + (n^2+n-1)g], \tag{25}$$

$$\nu_{11}^{(n)} = f_3^{(n)}(k_1), \quad \nu_{12}^{(n)} = f_3^{(n)}(k_2), \quad \nu_{13}^{(n)} = -n(n+1)g,$$

$$\nu_{21}^{(n)} = f_4^{(n)}(k_1), \quad \nu_{22}^{(n)} = f_4^{(n)}(k_2), \quad \nu_{23}^{(n)} = n(n+1)g,$$

$$\nu_{31}^{(n)} = \nu_{32}^{(n)} = g(1-g_{2n}) - g_{1n}, \quad \nu_{33}^{(n)} = 1+g(g_{2n}+1-n-n^2), \tag{26}$$

$$f_1^{(n)}(k_l) = (\beta_1 - k_l^2)g_{3n}[(\tilde{\beta}_1 - \tilde{k}_1^2)\tilde{x}_1 j_n'(\tilde{x}_1) j_n(\tilde{x}_2) - (\tilde{\beta}_1 - \tilde{k}_2^2)\tilde{x}_2 j_n'(\tilde{x}_2) j_n(\tilde{x}_1)] - n(n+1)g - \rho/\tilde{\rho},$$

$$f_2^{(n)}(k_l) = n(n+1)g + (\beta_1 - k_l^2)g_{3n}(\tilde{k}_1^2 - \tilde{k}_2^2)\tilde{x}_1 j_n'(\tilde{x}_1)\tilde{x}_2 j_n'(\tilde{x}_2),$$

$$f_3^{(n)}(k_l) = (\beta_1 - k_l^2)g_{3n}(\kappa/\tilde{\kappa})(\tilde{k}_1^2 - \tilde{k}_2^2) j_n(\tilde{x}_1) j_n(\tilde{x}_2) + 2g - g_{1n},$$

$$f_4^{(n)}(k_l) = 1 - n(n+1)g - (\beta_1 - k_l^2)g_{3n}(\kappa/\tilde{\kappa})[(\tilde{\beta}_1 - \tilde{k}_2^2)\tilde{x}_1 j_n'(\tilde{x}_1) j_n(\tilde{x}_2) - (\tilde{\beta}_1 - \tilde{k}_1^2)\tilde{x}_2 j_n'(\tilde{x}_2) j_n(\tilde{x}_1)], \tag{27}$$

$$g = 2(\tilde{\eta} - \eta)/(\tilde{\eta}\tilde{x}_3^2), \quad g_{1n} = \sigma(n-1)(n+2)/(\tilde{\rho}\omega^2 R_0^3), \quad g_{2n} = \tilde{x}_3 j_n'(\tilde{x}_3)/j_n(\tilde{x}_3)$$

$$g_{3n} = \tilde{\alpha}\beta_3\{\alpha\tilde{\beta}_3(\tilde{\beta}_1 - \tilde{k}_1^2)(\tilde{\beta}_1 - \tilde{k}_2^2)[\tilde{x}_1 j_n'(\tilde{x}_1) j_n(\tilde{x}_2) - \tilde{x}_2 j_n'(\tilde{x}_2) j_n(\tilde{x}_1)]\}^{-1}. \tag{28}$$

Here, κ is the thermal conductivity of the ambient liquid, σ is the surface tension, α is the volume thermal expansion coefficient of the ambient liquid, and the tilde denotes quantities that concern the internal medium of the particle.

The equations for the first-order velocity field look cumbersome but their numerical calculation is straightforward and does not bring any problems. Note also that the solutions are expressed as series, which makes it possible to take account of any number of surface modes.

B. Acoustic streaming

As follows from Doinikov, 1997a, the velocity of acoustic streaming can be written as

$$\boldsymbol{V} = \boldsymbol{\nabla}\Phi + \boldsymbol{\nabla} \times \boldsymbol{\Psi}, \tag{29}$$

where the scalar and vorticity potentials are given by

$$\Phi = \sum_{n=0}^{\infty} \Phi_n(r/R_0) P_n(\cos\theta), \tag{30}$$

$$\boldsymbol{\Psi} = \boldsymbol{e}_\varepsilon \sum_{n=1}^{\infty} \Psi_n(r/R_0) P_n^1(\cos\theta). \tag{31}$$

Substituting Eqs. (30) and (31) into Eq. (29), one finds the r and θ components of the acoustic streaming velocity to be

$$V_r = \frac{1}{R_0}\sum_{n=0}^{\infty}\left[\Phi_n'(x) + \frac{n(n+1)}{x}\Psi_n(x)\right]P_n(\cos\theta), \tag{32}$$

$$V_\theta = -\frac{1}{R_0}\sum_{n=1}^{\infty}\left[\frac{1}{x}\Phi_n(x) + \frac{1}{x}\Psi_n(x) + \Psi_n'(x)\right]P_n^1(\cos\theta), \tag{33}$$

where $x = r/R_0$ and the prime denotes the differentiation with respect to the argument in parentheses.

The functions $\Phi_n(x)$ and $\Psi_n(x)$ are calculated by

$$\Phi_n(x) = \frac{1}{x^{n+1}}[K_n(x,n+2) - C_{\varphi n}] - x^n[K_n(x,1-n) - K_n(\infty,1-n)], \tag{34}$$

$$\Psi_n(x) = \frac{1}{2(2n+3)}\left\{\frac{1}{x^{n+1}}N_n(x,n+3) - x^{n+2}[N_n(x,-n) - N_n(\infty,-n)]\right\} + \frac{1}{2(2n-1)}\{x^n[N_n(x,2-n) - N_n(\infty,2-n)] - x^{1-n}[N_n(x,n+1) - C_{\psi m}]\}, \tag{35}$$

where

$$K_n(x,m) = \int_1^x y^m \mu_n(y)dy, \tag{36}$$

$$\mu_n(r/R_0) = \frac{R_0^2}{4\rho}\int_0^\pi \mathrm{Re}\{\boldsymbol{\nabla}\cdot(\rho^{(1)*}\boldsymbol{v}^{(1)} - \rho_I^{(1)*}\boldsymbol{v}_I^{(1)})\}P_n(\cos\theta)\sin\theta d\theta, \tag{37}$$

$$N_n(x,m) = \int_1^x y^m[\chi_{rn}(y) + m\chi_{\theta n}(y)]dy, \tag{38}$$

$$\chi_{rn}(r/R_0) = \int_0^\pi W_r(r/R_0,\theta)P_n(\cos\theta)\sin\theta d\theta, \tag{39}$$

$$\chi_{\theta n}(r/R_0) = -\frac{1}{n(n+1)}\int_0^\pi W_\theta(r/R_0,\theta)P_n^1(\cos\theta)\sin\theta d\theta, \tag{40}$$

$$\boldsymbol{W}(r/R_0,\theta) = \frac{R_0^3}{4\nu}\mathrm{Re}\{\boldsymbol{v}^{(1)*}(\boldsymbol{\nabla}\cdot\boldsymbol{v}^{(1)}) + (\boldsymbol{v}^{(1)*}\cdot\boldsymbol{\nabla})\boldsymbol{v}^{(1)} - \boldsymbol{v}_I^{(1)*}(\boldsymbol{\nabla}\cdot\boldsymbol{v}_I^{(1)}) - (\boldsymbol{v}_I^{(1)*}\cdot\boldsymbol{\nabla})\boldsymbol{v}_I^{(1)}\}, \tag{41}$$

$$C_{\varphi n} = \frac{n(2n-1)}{2(n+1)} K_n(\infty, 1-n) + \frac{n(2n+1)}{4(2n+3)} N_n(\infty, -n) - \frac{n}{4} N_n(\infty, 2-n) + \frac{(2n+1)(n-2)}{2(n+1)} a_n + \frac{n(2n+1)}{2} b_n, \tag{42}$$

$$C_{\psi m} = (4n^2-1)\left[\frac{N_n(\infty, 2-n)}{2(2n-1)} - \frac{N_n(\infty, -n)}{2(2n+1)} - \frac{K_n(\infty, 1-n) + a_n}{n+1} - b_n \right], \tag{43}$$

$$a_n = \int_0^\pi U_r(R_0, \theta) P_n(\cos\theta) \sin\theta d\theta, \tag{44}$$

$$b_n = -\frac{1}{n(n+1)} \int_0^\pi U_\theta(R_0, \theta) P_n^1(\cos\theta) \sin\theta d\theta, \tag{45}$$

$$\boldsymbol{U}(r, \theta) = \frac{R_0}{2} \boldsymbol{V}_I - \frac{R_0}{4\omega^2} \mathrm{Re}\left\{ \left(\frac{\partial v_r^{(1)}}{\partial t} \right)^* \frac{\partial \mathbf{v}^{(1)}}{\partial r} \right\}. \tag{46}$$

Here, the asterisk denotes the complex conjugate, Re means "the real part of," and $\boldsymbol{V}_I$ is the velocity produced by the incident sound field in the ambient liquid when the particle is absent. This velocity results from the space attenuation of the incident sound wave. For a plane traveling wave, $\boldsymbol{V}_I$ is given by

$$\boldsymbol{V}_I = -\frac{|k_1 A|^2}{2\omega} \mathrm{Re}\{\boldsymbol{k}_1\} \exp(i\boldsymbol{k}_1 \cdot \boldsymbol{r} - i\boldsymbol{k}_1^* \cdot \boldsymbol{r}). \tag{47}$$

According to Eq. (60) in Doinikov, 1997a, the radial stress produced by acoustic streaming can be calculated by

$$\sigma_r = -\frac{1}{2\rho}\left(\xi + \frac{4\eta}{3} \right) \mathrm{Re}\{\nabla \cdot (\rho^{(1)*} \boldsymbol{v}^{(1)} - \rho_I^{(1)*} \boldsymbol{v}_I^{(1)})\} - \frac{\eta}{R_0^2} \sum_{n=0}^{\infty} Q_{rn}(r/R_0) P_n(\cos\theta), \tag{48}$$

where

$$Q_{rn}(x) = \frac{n}{x^{n+1}} [N_n(x, n+1) - C_{\psi m}] + (n+1) x^n [N_n(x, -n) - N_n(\infty, -n)]. \tag{49}$$

Substituting Eqs. (32) and (33) into the formula

$$\sigma_\theta = \eta\left(\frac{1}{r} \frac{\partial V_r}{\partial \theta} + \frac{\partial V_\theta}{\partial r} - \frac{V_\theta}{r} \right), \tag{50}$$

one finds the tangential stress produced by acoustic streaming to be

$$\sigma_\theta = \frac{\eta}{R_0^2} \sum_{n=1}^{\infty} Q_{\theta n}(r/R_0) P_n^1(\cos\theta), \tag{51}$$

where

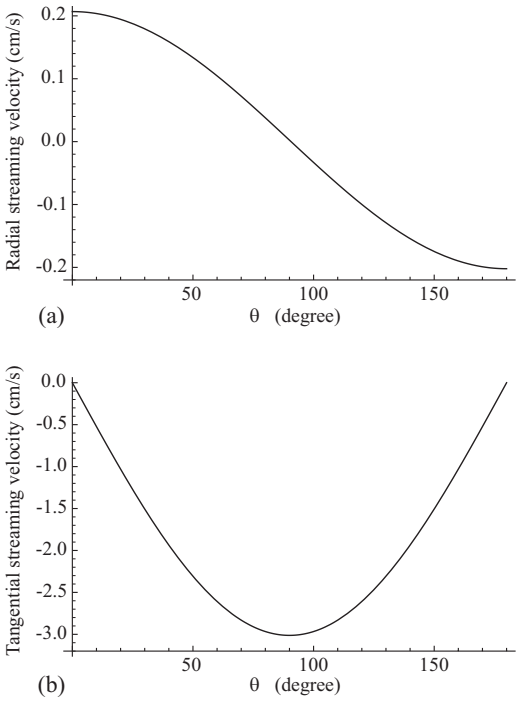


FIG. 1. Acoustic streaming velocity versus θ for an air bubble in water with $R_0 = 33$ μm. Excitation is a plane traveling wave with $f = 100$ kHz and the sound intensity 1000 W/m^2.

$$Q_{\theta n}(x) = \frac{2}{x^2}\Phi_n(x) - \frac{2}{x}\Phi_n'(x) + \frac{2 - 2n(n+1)}{x^2}\Psi_n(x) + \frac{2}{x}\Psi_n'(x) + x^n[N_n(x, -n) - N_n(\infty, -n)] - \frac{1}{x^{n+1}}[N_n(x, n+1) - C_{\psi m}]. \tag{52}$$

Like the first-order velocity equations, the equations for acoustic streaming seem to be cumbersome; however their numerical calculation is straightforward as well.

IV. NUMERICAL EXAMPLES

As a first example, let us consider an air bubble in water when $R_0 = 33$ μm and $f = 100$ kHz. The incident sound field is a plane traveling wave. The sound intensity is $I_s = 1000$ W/m^2. According to the known formula $I_s = P_a^2/2\rho c$, where P_a denotes the amplitude of the driving acoustic pressure, this sound intensity corresponds to the acoustic pressure amplitude 54.77 kPa. This is the case considered by Wu and Du (1997). The values of the other parameters used in this calculation are the following: $P_0 = 101.3$ kPa, $\rho = 1000$ kg/m^3, $\eta = 0.001$ Pa s, $\xi = 0.0026$ Pa s, $\sigma = 0.072$ N/m, $c = 1500$ m/s, $\kappa = 0.58$ J/s m K, $\chi = 1.4 \times 10^{-7}$ m^2/s, $\alpha = 2.1 \times 10^{-4}$ K^{-1}, $\gamma = 1.007$, $\tilde{\rho} = 1.2$ kg/m^3, $\tilde{\eta} = 1.8 \times 10^{-5}$ Pa s, $\tilde{\xi} = 1.6 \times 10^{-5}$ Pa s, $\tilde{c} = 330$ m/s, $\tilde{\kappa} = 0.034$ J/s m K, $\tilde{\chi} = 2.8 \times 10^{-5}$ m^2/s, $\tilde{\alpha} = 3.665 \times 10^{-3}$ K^{-1}, and $\tilde{\gamma} = 1.4$. Figure 1 shows the radial and the tangential streaming velocities versus θ, which are calculated at $r = R_0 + 2\delta_v$ as in Wu and Du's (1997) paper. The maximum magnitude of the radial velocity is 0.21 cm/s and that of the tangential velocity is 3.0 cm/s, whereas Wu and Du's (1997)

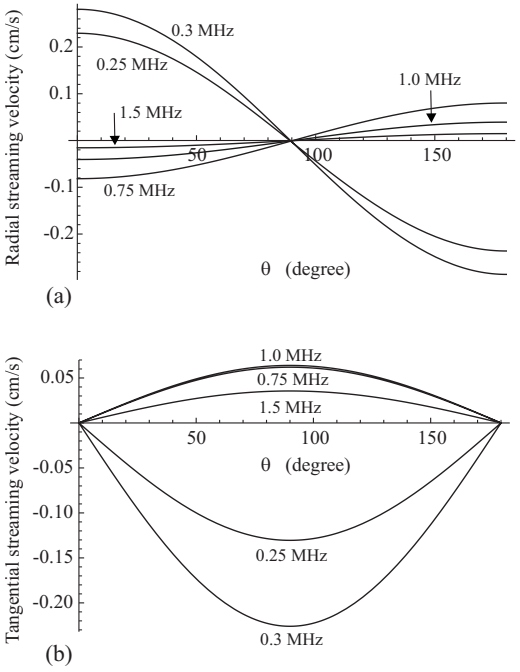


FIG. 2. Acoustic streaming velocity versus θ at different values of the driving frequency. The bubble radius is 5 μm and the acoustic pressure amplitude is 50 kPa.

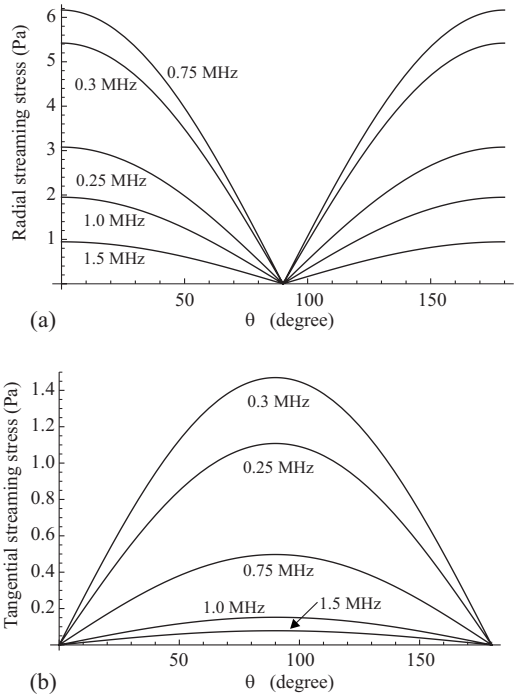


FIG. 3. Acoustic streaming stresses versus θ at different values of the driving frequency. The bubble radius is 5 μm and the acoustic pressure amplitude is 50 kPa.

calculation gives about 0.06 and 0.05 cm/s, respectively. This comparison confirms that Wu and Du's (1997) theory strongly underestimates the intensity of acoustic streaming.

Let us now consider examples of greater interest to biomedical applications. It is assumed that $R_0=5$ μm, f lies in the range 0.25–1.5 MHz, and the acoustic pressure amplitude is 50 kPa. The results of the calculation are presented in Figs. 2 and 3. Figure 2 shows the radial and tangential streaming velocities versus θ at different values of the driving frequency. Figure 3 demonstrates the behavior of the radial and tangential stresses produced by the acoustic streaming. The values of the velocities and the stresses are calculated at $r=R_0+2\delta_v$. One can see that the intensity of acoustic streaming increases as the driving frequency approaches the resonant frequency of the bubble, which is about 607 kHz in the case under consideration. It is interesting to note that, as follows from Fig. 2, the direction of acoustic streaming reverses when the driving frequency passes through the resonant frequency of the bubble. In other words, the direction of acoustic streaming around a bubble depends on whether the bubble is driven below or above resonance.

Our calculations show that the main contribution to acoustic streaming comes from the volume and translational modes of the bubble oscillation. The contribution of surface modes is negligible. However, in the case of a liquid drop, the effect of surface modes becomes decisive. As an example, let us consider a benzene drop suspended in water. The data for benzene are $\tilde{\rho}=880$ kg/m^3, $\tilde{\eta}=6.5\times10^{-4}$ Pa s, $\tilde{c}=1321$ m/s, $\tilde{\kappa}=0.14$ J/s m K, $\tilde{\chi}=1.18\times10^{-7}$ m^2/s, $\tilde{\alpha}=1.24\times10^{-3}$ K^{-1}, and $\tilde{\gamma}=1.1$. The equilibrium radius of the drop is set to be equal to 10 μm, the driving frequency is 90 kHz, and the acoustic pressure amplitude is 200 kPa. Note that the oscillations of liquid drops remain small up to quite high pressures, which allows one to apply the present theory to such pressures as 200 kPa and even higher. The velocity field of acoustic streaming that arises around the drop is shown in Fig. 4. The radial and the tangential streaming velocities are calculated at $r=R_0+2\delta_v$. The calculation reveals that the main contribution comes from the translational motion and the surface quadrupole

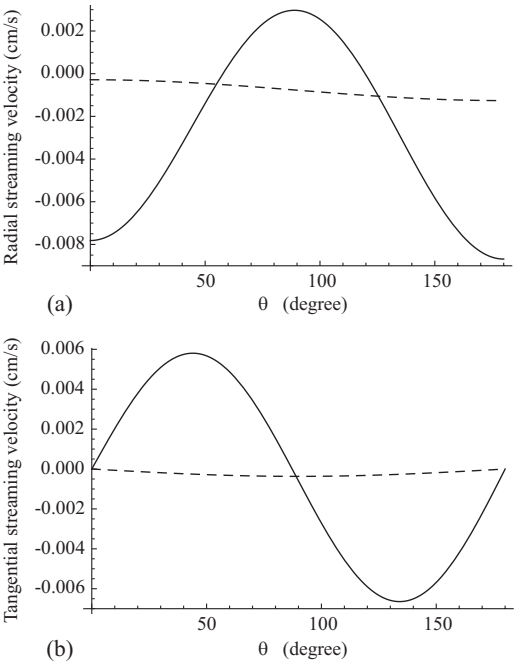


FIG. 4. Acoustic streaming velocity versus θ for a benzene drop in water. The equilibrium radius of the drop is 10 μm. Excitation is a plane traveling wave with the frequency of 90 kHz and the pressure amplitude of 200 kPa. The dashed lines show the result of the calculation when the surface quadrupole mode is ignored.

mode, whereas the contribution of the radial motion and surface modes of higher order is insignificant. This conclusion is illustrated by the dashed lines in Fig. 4 which show what happens if the quadrupole mode is ignored.

V. CONCLUSIONS

In this paper, the problem of acoustic microstreaming that develops around a gas bubble in an ultrasound field has been re-considered. The approach proposed previously by Wu and Du (1997), based on the assumption that viscous effects are essential only within a thin boundary layer while beyond the boundary layer the liquid can still be considered to be inviscid, has been analyzed. It has been shown that the solutions obtained by Wu and Du (1997) lead to a severe underestimation of the power of acoustic streaming because dominant viscous terms are determined incorrectly. In this paper, an improved theory has been suggested that corrects the errors of the previous theory and extends its limits. The proposed theory treats the entire bulk of the liquid outside the bubble and the gas inside the bubble as viscous heat-conducting fluids. No restrictions are imposed on the size of the bubble relative to the viscous, thermal, and sound wavelengths. All modes of the bubble's motion (volume pulsation, translation, and shape oscillations) are taken into account. Analytical expressions for the radial and tangential stresses produced by the acoustic streaming have been obtained as well. The proposed theory can be applied to both gas bubbles and liquid drops. By means of numerical examples, it has been confirmed that the main contribution to the acoustic streaming generated by a gas bubble comes from the volume and translational modes of the bubble oscillation. However, it has been shown that in the case of a liquid drop, due to low compressibility of the internal medium, the volume pulsation is insignificant, and therefore the contribution of surface modes becomes dominant.

ACKNOWLEDGMENT

A.A.D. gratefully acknowledges the financial support from the le STUDIUM®, Institute for Advanced Studies (Orléans, France).

Davidson, B. J., and Riley, N. (**1971**). "Cavitation microstreaming," J. Sound Vib. **15**, 217–233.

Doinikov, A. A. (**1997a**). "Acoustic radiation force on a spherical particle in a viscous heat-conducting fluid. I. General formula," J. Acoust. Soc. Am. **101**, 713–721.

Doinikov, A. A. (**1997b**). "Acoustic radiation force on a spherical particle in a viscous heat-conducting fluid. II. Force on a rigid sphere," J. Acoust. Soc. Am. **101**, 722–730.

Doinikov, A. A. (**1997c**). "Acoustic radiation force on a spherical particle in a viscous heat-conducting fluid. III. Force on a liquid drop," J. Acoust. Soc. Am. **101**, 731–740.

Elder, S. A. (**1959**). "Cavitation microstreaming," J. Acoust. Soc. Am. **31**, 54–64.

Lee, C. P., and Wang, T. G. (**1990**). "Outer acoustic streaming," J. Acoust. Soc. Am. **88**, 2367–2375.

Liu, X., and Wu, J. (**2009**). "Acoustic microstreaming around an isolated encapsulated microbubble," J. Acoust. Soc. Am. **125**, 1319–1330.

Longuet-Higgins, M. S. (**1998**). "Viscous streaming from an oscillating spherical bubble," Proc. R. Soc. London, Ser. A **454**, 725–742.

Nyborg, W. L. (**1958**). "Acoustic streaming near a boundary," J. Acoust. Soc. Am. **30**, 329–339.

Nyborg, W. L. (**1965**). "Acoustic streaming," in *Physical Acoustics*, edited by W. Mason (Academic, New York), Vol. **IIB**, pp. 266–331.

Wu, J., and Du, G. (**1997**). "Streaming generated by a bubble in an ultrasound field," J. Acoust. Soc. Am. **101**, 1899–1907.

A frequency domain linearized Navier–Stokes equations approach to acoustic propagation in flow ducts with sharp edges

Axel Kierkegaard[a] and Susann Boij
The Marcus Wallenberg Laboratory of Sound and Vibration Research (MWL), KTH, SE-100 44 Stockholm, Sweden

Gunilla Efraimsson
Department of Aeronautical and Vehicle Engineering, Linné FLOW Centre, KTH, SE-100 44 Stockholm, Sweden

(Received 22 June 2009; revised 19 November 2009; accepted 23 November 2009)

Acoustic wave propagation in flow ducts is commonly modeled with time-domain non-linear Navier–Stokes equation methodologies. To reduce computational effort, investigations of a linearized approach in frequency domain are carried out. Calculations of sound wave propagation in a straight duct are presented with an orifice plate and a mean flow present. Results of transmission and reflections at the orifice are presented on a two-port scattering matrix form and are compared to measurements with good agreement. The wave propagation is modeled with a frequency domain linearized Navier–Stokes equation methodology. This methodology is found to be efficient for cases where the acoustic field does not alter the mean flow field, i.e., when whistling does not occur.
 [DOI: 10.1121/1.3273899]

PACS number(s): 43.28.Py, 43.20.Fn, 43.20.Mv, 43.28.Ra [AH] Pages: 710–719

I. INTRODUCTION

Traffic is a major source of environmental noise in modern day society. Subsequently, development of new vehicles is subject to heavy governmental legislations. The major noise sources on common road vehicles are engine noise, transmission noise, tire noise, and, at high speeds, wind noise. At low speeds (<30–50 km/h), intake and exhaust noise are particularly important during acceleration. One way to reduce intake and exhaust noise is to attach mufflers to the exhaust pipes. However, to develop prototypes of mufflers for evaluation is a costly and time-consuming process. As a consequence, in recent years, so-called virtual prototyping has emerged. However, more efficient methods are needed in order to make this a useful alternative.

In predictions of the acoustic performance of duct systems, there is a strong need of knowledge on how incoming acoustic waves are transmitted and reflected in the system. This is sometimes referred to as the acoustic characterization of the system.[1] Often, linear conditions can be assumed, such that a frequency domain approach is adopted.

Different levels of approximations in the underlying equations can be applied in order to characterize an acoustical system by numerical simulations from the most simplistic wave equation with[2] or without[3,4] convective effects considered to the complexity of Euler or Navier–Stokes equations,[5,6] linearized about an arbitrary mean flow. The latter is needed to model a significant coupling between acoustic waves and vortical motion, such as when acoustic energy is transformed into vorticity around sharp edges.[7–9] For simulations of sound generation, the full non-linear equations are needed.[10–12]

The overall objective of this paper is to present an efficient methodology for the acoustic characterization of duct systems in order to quickly predict the effect of different geometry alterations. One of the advantages with simulations, as compared to experiments, is their ability to perform parameter studies with full control over the alterations made. In addition, the acoustic and hydrodynamic fields, respectively, are obtained in the whole computational domain. A possible drawback is that the computational time can be very long for very detailed simulations. Hence, for a method to be useful in practical applications, there is a balance between universality and computational cost. Previous numerical studies on wave propagation in-duct systems, in which the acoustic field and the flow field interact, include studies on flows through orifices or perforated plates. In general, advanced simulation techniques such as direct numerical simulations (DNS) or large eddy simulations (LES) are employed, as in the work by Tam and co-workers,[13,14] Leung *et al.*,[15] Mendez and Eldredge,[16] and Föller *et al.*[10] In these simulations, the non-linear time-dependent compressible Navier–Stokes equations are solved, with no or a small fraction of the turbulent eddies modeled, which in turn causes a high computational cost. The advantage is a full coupling between the acoustic field and the flow field, including a possible feedback on the flow field by the acoustic field. However, these simulations are beyond reach for practical industrial-like applications. The transmission and reflection of plane waves in a duct with an orifice were studied by Efraimsson *et al.*[17] via Reynolds averaged Navier-Stokes (RANS) simulations, which are less computationally expensive compared to

[a]Author to whom correspondence should be addressed. Electronic mail: axelk@kth.se

LES and DNS. It was concluded that an excessive damping of the acoustic waves occurs in the jet region. The damping was found to be due to an over-prediction of the turbulent viscosity by the turbulence model used. This implied that the characterization of the orifice duct gave results that were not in agreement with experimental results, and the conclusion that RANS simulations of plane wave propagation in separated flows need to be performed with great care. Too large dissipation was also found in simulations of scattering at an area expansion by the use of LES.[10]

We present an efficient methodology for the acoustic characterization of ducts with sudden area variations. The ambition is to include relevant physical effects but still have a computationally cheap method. In order to yield reliable results, the method needs to capture the coupling between the acoustic field and the hydrodynamic field that can occur at an edge.[18,19] In the methodology, the acoustic waves are modeled as solutions to the linearized Navier–Stokes equations. This is believed to be sufficient for flows with a fixed point of separation, since theoretical work with linear models predicts the coupling between the acoustic and hydrodynamic fields, respectively, very well, see, e.g., the work by Boij and Nilsson[18,19] and references therein. In addition, a methodology based on the linearized equations is computationally cheaper, as compared to using the full non-linear equations since it enables the usage of a coarser computational grid, which is designed for the length scales of the acoustic field.

An isentropic relation between the acoustic pressure and density fluctuations is assumed according to general linear acoustics,[20] such that the energy equation is redundant and can be omitted from the system of equations. Also, we make the ansatz of harmonic time dependence yielding a frequency domain formulation. This considerably reduces the computational cost.

In the methodology, the Navier–Stokes equations are linearized about an arbitrary stationary mean flow. Thus, the methodology is general and any realistic time-independent mean flow field can, in principle, be handled.

In this paper, we present the methodology and evaluate its ability to acoustically characterize a duct element with an orifice. Comparisons are made with experimental results. This paper is organized as follows. In Sec. II, the linearized Navier–Stokes equations in frequency domain are derived. In Sec. III, the test case of a duct with an orifice is described. Thereafter, in Sec. IV, a frequency scaling in order to compare the experimental results in three space dimensions to the numerical results in two space dimensions are discussed. The computational methodology is presented in Sec. V. In Sec. VI, results for plane waves in a quiescent medium are presented, followed by a presentation of the mean flow and acoustic field, respectively, in Secs. VII. The plane wave decomposition and a presentation of the acoustic two-ports are shown in Sec. VIII. Finally, the results and conclusions of the paper are presented in Secs. IX and X, respectively.

II. THE LINEARIZED NAVIER–STOKES EQUATIONS

The basic idea for the proposed methodology is to efficiently model acoustic wave propagation with inclusion of the effects of the flow field. In this section, an appropriate form of the governing equations is derived. The scope is acoustic wave propagation in flow duct geometries with sharp edges, where flow separation, production of vorticity, and viscous effects can be assumed to play a non-negligible role. We aim for a formulation where the viscous effects at an edge are inherent in the equations. It is then appropriate to apply the Navier–Stokes equations. Hence, no explicit edge condition needs to be employed.

For the derivation, the starting point is the full compressible Navier–Stokes equations. These are in dimensional form in a Cartesian coordinate system:[21]

$$\text{Continuity:}\quad \frac{D\rho}{Dt}+\rho\frac{\partial u_k}{\partial x_k}=0,$$

$$\text{Momentum:}\quad \rho\frac{Du_i}{Dt}=-\frac{\partial p}{\partial x_i}+\frac{\partial \tau_{ij}}{\partial x_j}+\rho F_i,$$

$$\text{Energy:}\quad \rho\frac{De}{Dt}=-p\frac{\partial u_k}{\partial x_k}+\Phi+\frac{\partial}{\partial x_k}\left(\kappa\frac{\partial T}{\partial x_k}\right), \tag{1}$$

with

$$\Phi=\tau_{ij}\frac{\partial u_i}{\partial x_j},\quad \tau_{ij}=\mu\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}-\frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij}\right), \tag{2}$$

$e=e(p,T)$, and $p=\rho RT$, where ρ is the density, p is the pressure, R is the universal gas constant, T is the absolute temperature, u_i is the velocity component in the ith direction (where $i=1$ is the x-direction and $i=2$ is the y-direction), τ_{ij} is the viscous stress tensor, F_i is a volume force in the ith direction, e is the internal energy, κ is the thermal conductivity, μ is the dynamic viscosity, Φ is the dissipation function, D/Dt is the convective derivative, δ_{ij} is the Kronecker delta function, and the Einstein summation convention is used.

We first assume that the solution of Eq. (1) can be divided into an *a priori* given time-independent mean flow and unknown time-dependent perturbations. For a two-dimensional (2D) case, this yields

$$\rho(\boldsymbol{x},t)=\rho_0(\boldsymbol{x})+\rho'(\boldsymbol{x},t),$$

$$u(\boldsymbol{x},t)=u_0(\boldsymbol{x})+u'(\boldsymbol{x},t),$$

$$v(\boldsymbol{x},t)=v_0(\boldsymbol{x})+v'(\boldsymbol{x},t),$$

$$p(\boldsymbol{x},t)=p_0(\boldsymbol{x})+p'(\boldsymbol{x},t), \tag{3}$$

where $u=u_1$ and $v=u_2$. The quantities ρ_0, u_0, v_0, and p_0 represent the mean flow components, and the prime denotes the time-varying perturbations.

Thereafter, we assume that linear acoustics applies. The assumption of linear conditions is based on measurements presented in the work by Ingård and Ising,[22] where linear acoustic relations for an orifice were found up to sound pressure levels of 125 dB in up to 0.2 Mach number mean flows.

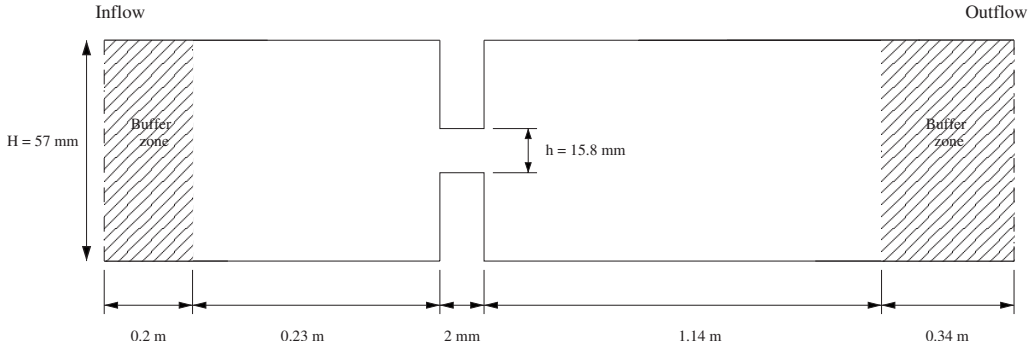


FIG. 1. An overview of the geometry of the duct in the calculation. Note that the sketch is not to scale.

The shape and dimensions of the orifice in the study[22] did, however, differ from those of the orifice treated in this paper.

Furthermore, we assume that the relation between pressure and density can be regarded as isentropic. This is not applicable for all flows, e.g., combustion, but early work indicates that this assumption is appropriate for the type of problem treated here.[18] The assumption decreases the implementation and computational effort. In this case, the pressure and density perturbations are related as[20]

$$\frac{\partial p'}{\partial x_i} = c^2 \frac{\partial \rho'}{\partial x_i}, \tag{4}$$

where $c^2(x) = \gamma p_0 / \rho_0$ is the local adiabatic speed of sound, and γ is the ratio of specific heats. With this relation, one dependent variable is removed from the system, and the continuity and momentum equations are decoupled from the energy equation, which in turn can be omitted from the system of Eq. (1). In this way, the size of the computational problem is considerably reduced.

We assume harmonic time dependence for the perturbation quantities considered, and a frequency domain approach is hence adopted. Any perturbation quantity q' can then be represented as

$$q'(\boldsymbol{x},t) = \mathrm{Re}\{\hat{q}(\boldsymbol{x}) e^{-i\omega t}\}, \tag{5}$$

where $\hat{q}$ is a complex quantity and ω is the angular frequency.

Substituting Eqs. (3)–(5) into Eq. (1) and linearizing, the linearized Navier–Stokes equations can be written in the following forms:

$$\hat{\rho}: \quad (u_0 v_0)\, \nabla \hat{\rho} + \left(\frac{\partial u_0}{\partial x} + \frac{\partial v_0}{\partial y} - i\omega \right) \hat{\rho} = - \left(\frac{\partial \rho_0 \hat{u}}{\partial x} + \frac{\partial \rho_0 \hat{v}}{\partial y} \right), \tag{6}$$

$$\hat{u}: \quad \nabla^T \left[- \begin{pmatrix} \frac{4}{3}\mu & 0 \\ 0 & \mu \end{pmatrix} \nabla \hat{u} \right] + \rho_0 (u_0 \quad v_0)\, \nabla \hat{u} + \rho_0 \left(\frac{\partial u_0}{\partial x} - i\omega \right) \hat{u} = \rho_0 \hat{F}_x - \left(u_0 \frac{\partial u_0}{\partial x} + v_0 \frac{\partial u_0}{\partial y} \right) \hat{\rho} - c^2 \frac{\partial \hat{\rho}}{\partial x} + \frac{1}{3}\mu \frac{\partial^2 \hat{v}}{\partial x\, \partial y} - \rho_0 \frac{\partial u_0}{\partial y} \hat{v}, \tag{7}$$

$$\hat{v}: \quad \nabla^T \left[- \begin{pmatrix} \mu & 0 \\ 0 & \frac{4}{3}\mu \end{pmatrix} \nabla \hat{v} \right] + \rho_0 (u_0 \quad v_0)\, \nabla \hat{v} + \hat{\rho}_0 \left(\frac{\partial v_0}{\partial y} - i\omega \right) \hat{v} = - c^2 \frac{\partial \hat{\rho}}{\partial y} - \left(u_0 \frac{\partial v_0}{\partial x} + v_0 \frac{\partial v_0}{\partial y} \right) \hat{\rho} - \rho_0 \frac{\partial v_0}{\partial x} \hat{u} + \frac{1}{3}\mu \frac{\partial^2 \hat{u}}{\partial x\, \partial y}, \tag{8}$$

where $\nabla = (\partial/\partial x \quad \partial/\partial y)^T$. This is the form used for the simulations throughout this work.

III. TEST CASE

In order to evaluate the proposed methodology, both from a modeling point of view and also from computational aspects, the propagation of acoustic plane waves in a straight two-dimensional duct with an orifice plate mounted inside is studied. The results are then compared to the experimental data presented in the work by Åbom *et al.*[23] The parameter values in the test case correspond to those of the flow case and the geometry in the experiment.

A schematic overview of the geometry in the calculation is shown in Fig. 1. The duct length of the full experimental rig is around 7 m, but here only a small section of less than 2 m is simulated to reduce computational costs. The height of the duct is $H = 57$ mm, the same as the diameter of the duct in the experiment. The diameter of the orifice in the experiment is $d = 30$ mm. This yields an area aspect ratio $\eta = 30^2/57^2 \approx 0.277$. As the area ratio is of fundamental importance for the acoustic propagation properties, this ratio needs to be maintained, and thus in the 2D rectangular case, an orifice height of $h = 0.277 \times 57$ mm$^2 = 15.8$ mm is chosen. The thickness of the orifice plate is 2 mm in both cases, and the orifice edges are sharp with an angle of 90°. A mean flow of $M = 0.068$, averaged over the duct cross-section, is present.

In order to efficiently solve Eqs. (6)–(8) for the test case, three different boundary conditions are introduced for the surfaces of the orifice plate, the duct walls, and the in- and outflow boundaries, respectively.

On the surface of the orifice plate, rigid wall, no-slip boundary conditions,

$$\hat{\boldsymbol{u}} = \mathbf{0} \quad \text{and} \quad \hat{\boldsymbol{n}} \cdot \nabla \hat{\rho} = 0, \tag{9}$$

are imposed, where $\hat{\boldsymbol{n}}$ is a unit vector normal to the walls. It should be noted that as the velocity components are restricted to zero on all sharp edges, no explicit Kutta edge condition is needed.

At the duct walls in between the orifice plate and the in- and outflow boundaries, respectively, rigid wall, slip boundary conditions,

$$\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}} = 0 \quad \text{and} \quad \hat{\boldsymbol{n}} \cdot \nabla \hat{\rho} = 0, \tag{10}$$

are used. The benefit of this choice is that the acoustic boundary layers along the duct walls are not resolved by the computational mesh, which in turn considerably decreases the computational effort. On the other hand, this implies that the viscous dissipation associated with the damping effect at these walls is not included in the model governing equations. For the evaluation of scattering at the orifice, this damping, however, has no influence.

At the in- and outflow boundaries, only plane waves are assumed to be present, and thus non-reflecting boundary conditions

$$\hat{\boldsymbol{n}} \cdot \nabla \hat{\rho} = ik\hat{\rho}, \quad \hat{\boldsymbol{n}} \cdot \nabla \hat{u} = ik\hat{u}, \quad \hat{v} = 0 \tag{11}$$

for a vertical boundary are employed. Here, k is the wave number, and $k = \omega/(c - u_0)$ at the upstream boundary and $k = \omega/(c + u_0)$ at the downstream boundary. These non-reflecting boundary conditions break down in the presence of vorticity, where the Helmholtz equation is not satisfied. Therefore, additional variable viscosity is applied in the buffer zones of Fig. 1. This can be written as

$$\mu = \mu_{\text{physical}} + \mu_{\text{artificial}}, \tag{12}$$

where the artificial dynamic viscosity $\mu_{\text{artificial}} = 0$ inside the computational domain is then ramped up as a cubic polynomial through the buffer zone to a specific value of, in this case, $\mu_{\text{artificial}} = 10$ at the boundaries.

As an acoustic source, a time-harmonic body force function $\hat{F}(\boldsymbol{x})$ is applied over a region in the domain. The advantage with employing forcing functions inside a domain is that non-reflective boundary conditions can then be readily applied at the in- and outflow boundaries. The body force functions are implemented as bell-shaped piece-wise cubic interpolation polynomials with peak amplitudes 10^2. The source regions are located at $-0.26 \leq x \leq -0.22$ and $1.03 \leq x \leq 1.14$ for the two cases with sources up- or downstream of the orifice, respectively. This methodology is analogous to measurement techniques in which loudspeakers, instead of volume forces, are used as sound sources.

IV. FREQUENCY SCALING

The simulation model is for a two-dimensional duct, whereas the experimental data to be compared are for a three-dimensional cylindrical geometry. In order to compare the simulation results to the measurement results, a relevant frequency scaling has to be adopted; here, the scaling proposed by Boij and Nilsson[18] is chosen.

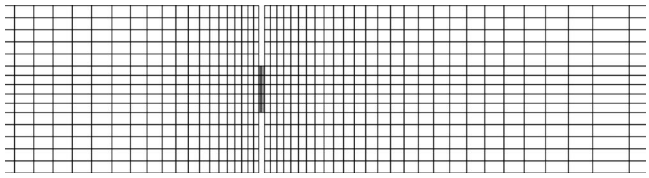


FIG. 2. Proportional overview of the mesh around the orifice.

To relate the results of two-dimensional model to those of the three-dimensional problem, the so-called normalized Helmholtz number He^* of the two cases should be equal. The normalized Helmholtz numbers of a system is the Helmholtz number divided by the Helmholtz number of the frequency of onset of the first higher order mode of the system, the so-called cut-on frequency, such that $\text{He}^* = 1$ at cut-on. This is defined as

$$\text{He}^* = \frac{(kH)_{2\text{D}}}{\pi} = \frac{(ka)_{\text{cyl}}}{\kappa_0}, \tag{13}$$

where π and $\kappa_0 \approx 3.832$ are the cut-on Helmholtz numbers for a 2D duct and a cylindrical duct, respectively, and a is the duct radius in the cylindrical case. This determines the relation between the frequencies of the two-dimensional calculations and those of the measurements of a cylindrical duct.

V. COMPUTATIONAL METHODOLOGY AND MESH

The linearized Navier–Stokes equations (6)–(8) were implemented in the commercial finite element method code, Comsol MULTIPHYSICS v3.5, and solved with the UMFPACK direct solver. The typical time consumption for one frequency was around 20 s on a 3 GHz, 2 Gbyte RAM workstation. A mesh of 7120 quadrilateral elements of third order Lagrangian interpolation polynomials were used, yielding a problem of 195 069 degrees of freedom. Grid stretching with a 3% growth ratio was applied in both directions from the orifice, such that the grid was finer where small scale vorticity was present and coarser in regions of only plane wave propagation. The largest elements, located at the outlet, were 1.8 cm, and the wavelength of the highest frequency ($\approx$2600 Hz) was $\lambda_{\text{ac}} = c/f = 343/2600 = 13.2$ cm, which yields an element per wavelength ratio of $13.2/1.8 = 7.3$ for the worst case. This is not high, but acceptable when using higher order interpolation polynomials.

The smallest elements are located at the orifice and are 0.2 mm in length. The approximate length scale of the vorticity waves are $\lambda_{\text{vort}} = \lambda_{\text{ac}} M = 8.8$ mm for the highest frequency. The length scale of a typical element in the vorticity region is about 1.2 mm, which gives $8.8/1.2 = 7.3$ elements per vortical wavelength for the highest frequency. For the lowest frequency, a similar approximation yields more than 380 elements per vortical wavelength. The acoustic boundary layer thickness is given by $\delta_{\text{ac}} = \sqrt{2\nu/\omega}$,[24] and is not accurately resolved by the mesh used. To resolve the acoustic boundary layer would, however, require an extremely fine mesh. It is seen in Sec. IX that this is a valid assumption. An overview of the mesh in the vicinity of the orifice is shown in Fig. 2.

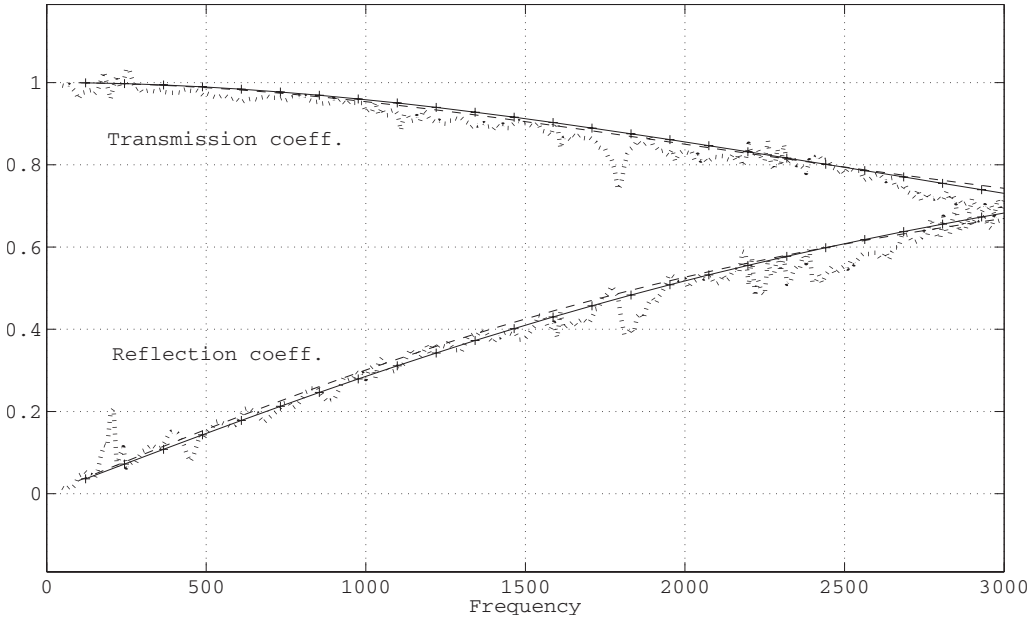


FIG. 3. Plane wave reflection and transmission coefficients for an orifice in a duct with a quiescent medium, comparison between measurements and theory. Experimental data: (Ref. 23) ···; simulations (present model): −+−; analytical model: (Ref. 25) −−.

VI. SIMULATION IN A QUIESCENT MEDIUM

A first test to evaluate the methodology is to determine the acoustic scattering data for the orifice in a duct without mean flow, for which both corresponding experimental data[23] and an analytical solution[25] exist. Thus, an indication of the accuracy of both the simulations and the measurements can be obtained. Simulations were carried out for frequencies in a sweep from 50 to 3050 Hz with a 100 Hz frequency step to correspond to the frequencies measured in the experiments.

Plane wave reflection and transmission coefficients for the orifice are shown in Fig. 3. It is clear that the result of all three methods is in excellent agreement.

VII. SIMULATION WITH MEAN FLOW

In the derivations of the linearized Navier–Stokes equations, Eqs. (6)–(8), the main assumption on the mean flow field is that it is time-independent. The mean flow field for the test case was calculated by averaging the time exact solutions from the DNS calculation described by Kierkegaard and Efraimsson.[26] The mean Mach number, averaged over the duct cross-section of the mean flow, is $M=0.068$, and the Reynolds number based on the duct height is $\mathrm{Re}_H=10^4$.

In Fig. 4, the mean flow is shown in a region -0.14 m $\leq x \leq 0.3$ m $(-2.5 \leq x/H \leq 5)$. The shear layer downstream of the orifice is easily identified in the u_0 component. Also,

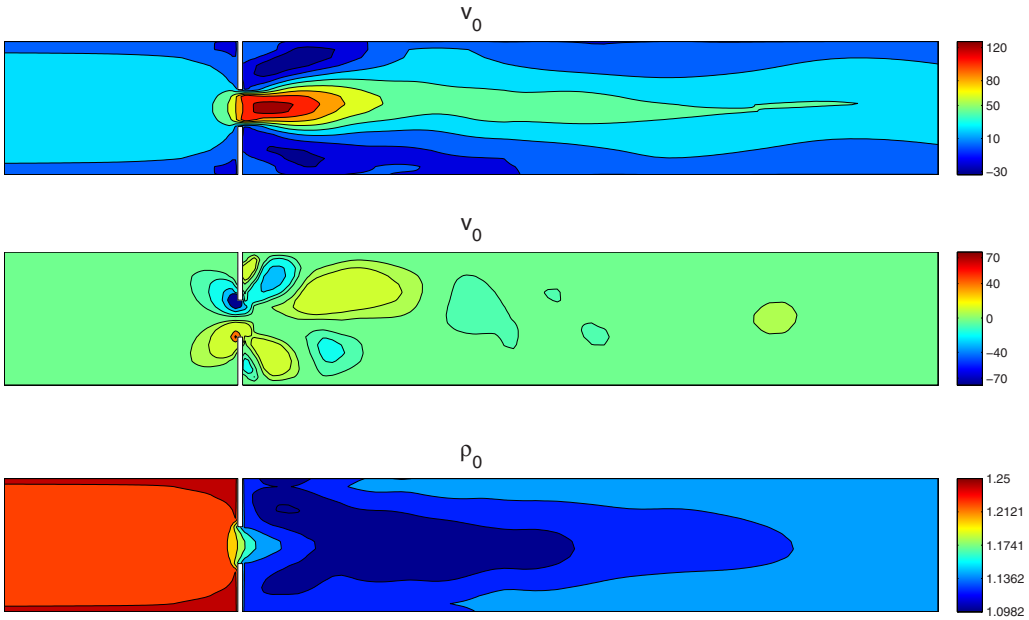


FIG. 4. (Color online) The mean flow components axial velocity u_0, transverse velocity v_0, and density ρ_0 used for linearization.

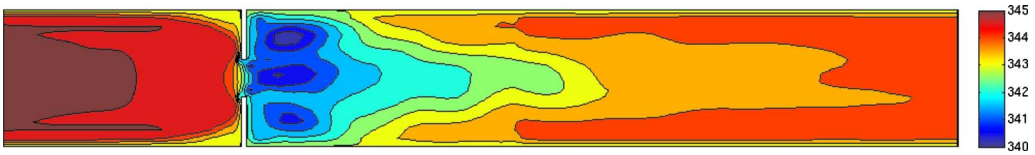


FIG. 5. (Color online) Spatial variation in the speed of sound c in the duct (m/s).

the recirculation zones are present in the u_0 and v_0 components. The slight asymmetry of the jet, as seen in the figure for u_0, could be due to the fact that the time span of the DNS solution is not long enough for the mean quantities to stabilize or could be due to a Coanda effect, similar to the mean flow reported by Gloerfelt and Lafon.[12]

The speed of sound is shown in Fig. 5. The local speed of sound c in Eqs. (7) and (8) is calculated by $c=\sqrt{\gamma p_0/\rho_0}$, where p_0 and ρ_0 are given from the DNS. It can be noted that that the sound speed is rapidly varying in the vicinity of the orifice.

The calculations of the acoustic field in the duct with the orifice were carried out in a frequency range $50 \leq f \leq 3150$ Hz with a frequency step of $\Delta f=100$ Hz. All frequencies are below the cut-on frequency of the first higher order mode, which in a circular duct with radius $r=2.85$ mm and rigid walls is[27]

$$f^c = \frac{c_0 k_{1,0}}{2\pi}\sqrt{1-M^2} \approx 3.5 \text{ kHz}, \tag{14}$$

where $k_{1,0} \approx 1.84/r$. Due to computational limitations, a higher dynamic viscosity than that of air was chosen, such that a Reynolds number of $\mathrm{Re}=cH/\nu=10^4$ was obtained. This reduces spurious waves arising in the simulation and allows for coarser meshes throughout the domain without noticeable effects on the acoustic solution. The density field at $f=1550$ Hz is shown in Fig. 6. In the plot, the plane wave character of the acoustic field is clearly visible. A close-up of all the field quantities in the vicinity of the orifice is shown in Figs. 7 and 8 for $f=1000$ Hz and $f=2500$ Hz, respectively, for comparison of the perturbation fields of a low and a high frequency case. Since the same forcing function is used in both cases, the power input to the acoustic waves differs between the frequencies.

In Figs. 7 and 8, the solution for the perturbations at 1000 and 2500 Hz are displayed, with the source located far upstream of the orifice. Upstream of the orifice, a strong standing wave pattern, due to the reflections at the orifice, is observed. Downstream, transmitted plane waves are propagating away from the orifice. Vortical structures in the solutions are concentrated to the regions downstream of the orifice, as can be seen in Fig. 9; however, at around three to five duct diameters downstream of the orifice, the flow field is dominated by acoustic plane wave motion.

VIII. PLANE WAVE DECOMPOSITION AND ACOUSTIC TWO-PORTS

In order to obtain the acoustic scattering properties for the orifice and to compare the solution with experimental results, we first need to decompose the wave solution into upstream and downstream propagating acoustic waves. Then, the acoustic properties are formulated as a scattering matrix.

To obtain the up- and downstream propagating acoustic waves from the simulated solution, we apply a plane wave decomposition. Hence, we assume that the acoustic field quantities can be written as a sum of its up- and downstream propagating components as

$$\hat{\rho} = \hat{\rho}_+ + \hat{\rho}_-, \quad \hat{u} = \hat{u}_+ + \hat{u}_-. \tag{15}$$

For a plane wave, the relation $\hat{\rho}=\pm\rho_0/c_0\hat{u}$ is valid, and thus, the wave decomposition can be written as

$$\hat{\rho}_+(x) = \frac{1}{2}\left(\hat{\rho}_{\text{mean}} + \frac{\rho_0}{c_0}\hat{u}_{\text{mean}}\right),$$

$$\hat{\rho}_-(x) = \frac{1}{2}\left(\hat{\rho}_{\text{mean}} - \frac{\rho_0}{c_0}\hat{u}_{\text{mean}}\right), \tag{16}$$

where *mean* is an averaging of the quantities over the duct cross section, i.e.,

$$\hat{\rho}_{\text{mean}}(x) = \frac{1}{H}\int_0^H \hat{\rho}(x,y)dy, \tag{17}$$

and correspondingly for the velocity perturbations. In Fig. 10, the magnitudes of the up- and downstream propagating waves and the magnitude of the total density perturbation field are shown as a function of x. The results shown in the figure is for the case with $f=1550$ Hz and with the forcing function applied upstream of the orifice plate. The orifice plate is located at $x=0$. It can be seen that the decomposition method is not valid inside the region $-0.02<x<0.1$, where vorticity and near field effects are significant. The oscilla-

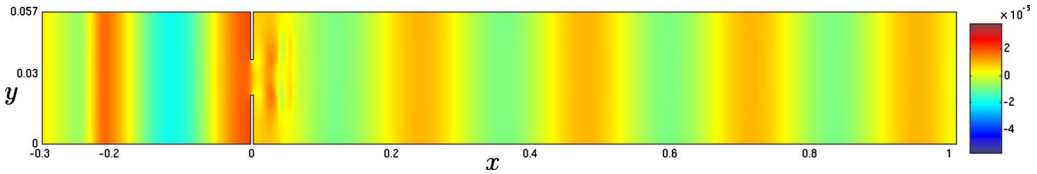


FIG. 6. (Color online) Overview of the real part of the density field at $f=1550$ Hz (corresponding to a duct Helmholtz number of $\mathrm{He}=1.6$) with the acoustic source located upstream of the orifice. Note that the figure is not to scale.

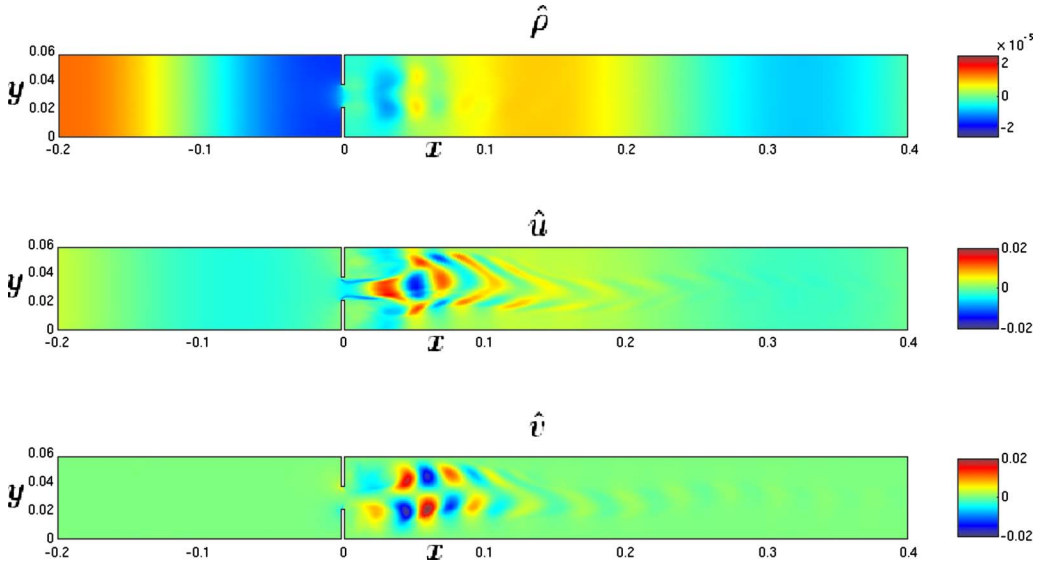


FIG. 7. (Color online) The real part of the perturbed quantities at $f=1000$ Hz (corresponding to a duct Helmholtz number of He=1.0). The acoustic source is located upstream of the orifice: $\hat{\rho}$ (upper), $\hat{u}$ (middle), and $\hat{v}$ (lower).

tions of $|\hat{\rho}|$ at $x<0$ is due to a standing wave phenomena by the interaction between the incident wave and the reflected wave. As a side effect, the decomposition enables the estimate of the numerical reflections from the boundary conditions. By dividing the magnitude of the waves incident on in- or outflow boundaries with those propagating from the corresponding boundary, the reflections from the non-reflecting boundaries can be estimated as $\hat{\rho}_-/\hat{\rho}_+ \approx 0.03$, $x \gtrsim 0.1$. It is thus observed that the reflections from the non-reflecting boundary conditions are about a third of a percent.

For acoustic characterization of in-duct components, it is often convenient to describe the properties of the acoustic element with scattering matrix formulation.[1] If no sources of sound are present in the duct element, the scattering matrix **S** can be written as

$$\begin{pmatrix} \hat{\rho}_{a-} \\ \hat{\rho}_{b-} \end{pmatrix} = \mathbf{S} \begin{pmatrix} \hat{\rho}_{a+} \\ \hat{\rho}_{b+} \end{pmatrix}. \tag{18}$$

Here, $\hat{\rho}_{a\pm}$ and $\hat{\rho}_{b\pm}$ are the density perturbation amplitudes of the waves, as depicted in Fig. 11.

To determine the four unknown quantities in **S**, two independent acoustic fields, denoted by *I* and *II*, are needed.[1] This is realized with a version of the two source location method,[28] where in this case a time-harmonic body force is applied first at the inflow and then at the outflow to obtain two independent acoustic cases. The forcing functions are described in Sec. III. The scattering matrix can then be calculated from

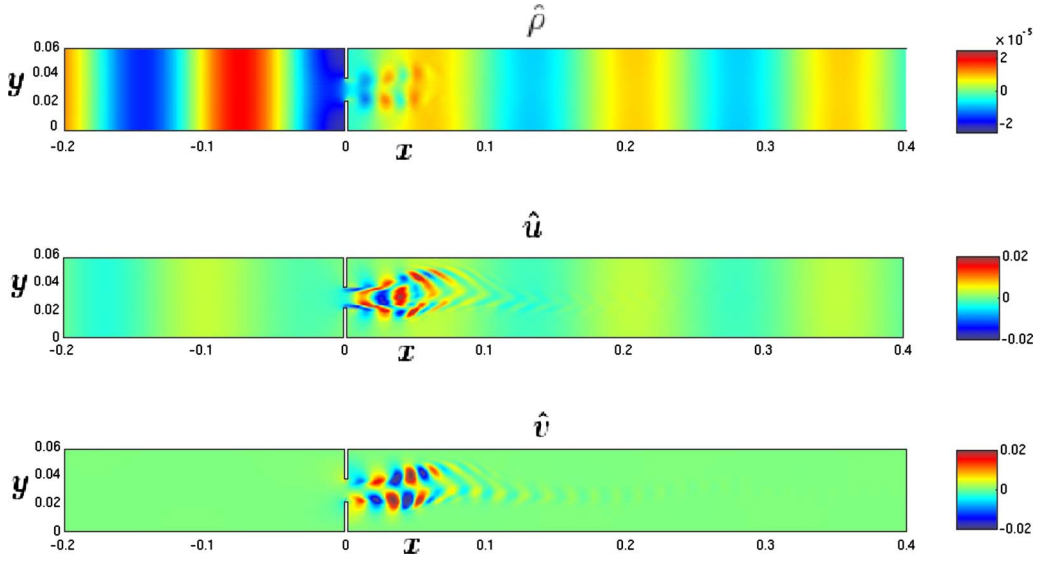


FIG. 8. (Color online) The real part of the perturbed quantities at $f=2500$ Hz (corresponding to a duct Helmholtz number of He=2.6). The acoustic source is located upstream of the orifice: $\hat{\rho}$ (upper), $\hat{u}$ (middle), and $\hat{v}$ (lower).

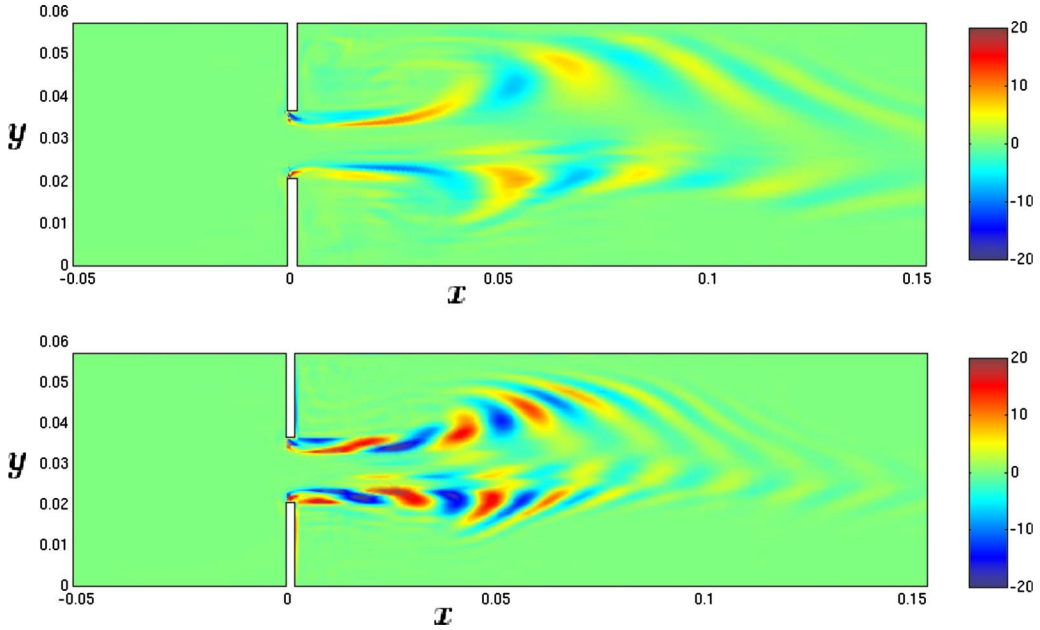


FIG. 9. (Color online) The real parts of the acoustic-generated vorticity fields at $f=1000$ Hz (above) and $f=2500$ Hz (below), corresponding to the perturbation field of Figs. 7 and 8, respectively.

$$\mathbf{S} = \begin{bmatrix} R^+ & T^- \\ T^+ & R^- \end{bmatrix} = \begin{bmatrix} \hat{\rho}^I_{a-} & \hat{\rho}^{II}_{a-} \\ \hat{\rho}^I_{b-} & \hat{\rho}^{II}_{b-} \end{bmatrix} \begin{bmatrix} \hat{\rho}^I_{a+} & \hat{\rho}^{II}_{a+} \\ \hat{\rho}^I_{b+} & \hat{\rho}^{II}_{b+} \end{bmatrix}^{-1}, \tag{19}$$

where R^+ and R^- represent the upstream side and downstream side plane wave reflection coefficients, respectively, and T^+ and T^- represent the upstream-to-downstream side and downstream-to-upstream side plane wave transmission coefficients, respectively.

An incompressible quasi-steady theory for the acoustic response of an orifice in a flow duct is presented by Åbom *et al.*[23] and Hofmans,[29] and will be used for comparison. The scattering matrix, as defined in Eq. (19), is obtained from

$$\begin{pmatrix} \hat{\rho}_{1-} \\ \hat{\rho}_{2-} \end{pmatrix} = \frac{1}{2+MC_L} \begin{bmatrix} MC_L & 2 \\ 2 & MC_L \end{bmatrix} \begin{pmatrix} \hat{\rho}_{1+} \\ \hat{\rho}_{2+} \end{pmatrix}, \tag{20}$$

where M is the Mach number of the mean flow, $C_L = (1/(\eta\Gamma_0)-1)^2$ is the pressure loss coefficient, and Γ_0 is the vena contracta contraction ratio.

Here, a low-frequency, low Mach number approximation

$$\Gamma_0 = \frac{1}{1+\sqrt{0.5(1-\eta)}} \tag{21}$$

for the vena contracta ratio is used, see, for example, Ref. 30. The reflection and transmission coefficients are then given by

$$R = \frac{1}{1+2/MC_L} \approx 0.44 \quad \text{and} \quad T = \frac{1}{1+MC_L/2} \approx 0.56. \tag{22}$$

In this model, the coefficients are neither dependent on frequency nor direction of propagation relative to the mean flow direction.

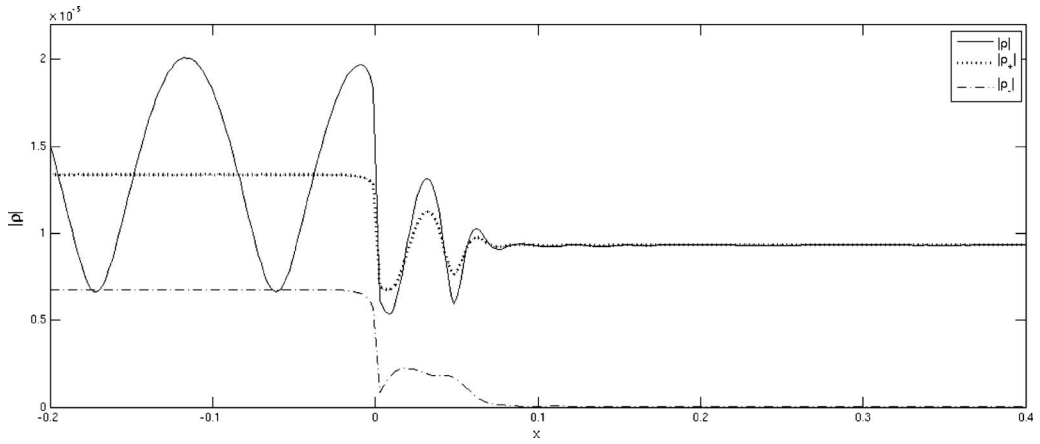


FIG. 10. Plane wave decomposition of simulated data at $f=1550$ Hz. Magnitude of the total density field $|\hat{\rho}|$: –; magnitude of the density of the downstream propagating wave $|\hat{\rho}^+|$: – – –; magnitude of the density of the upstream propagating wave $|\hat{\rho}^-|$: – · – · –.

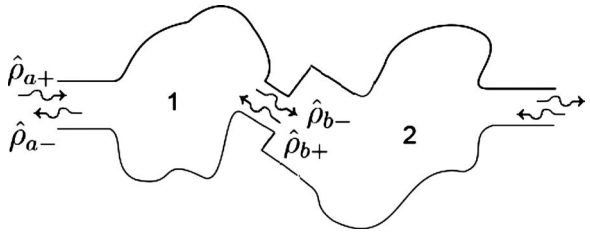


FIG. 11. An acoustic two-port and the definition of the wave propagation directions.

IX. COMPARISON WITH EXPERIMENTAL DATA

Now, simulated results obtained from the methodology proposed in this paper are compared to experimental data and results obtained with the quasi-stationary model described above. In Fig. 12, the simulated results for the scattering matrix elements are compared to experimental data from measurement by Åbom *et al.*[23] Results for the limit of zero frequency from the quasi-stationary model[23,29] are also displayed. The results for simulations and experiments are in excellent agreement for most frequencies. The matrix element $\mathbf{S}_{11}$, i.e., the upstream side reflection coefficient R^+, shows a discrepancy of about 4%–10% at frequencies above 1.5 kHz. This might be due to inaccuracies in the simulation of the mean flow. The discrepancies in $\mathbf{S}_{11}$ and $\mathbf{S}_{21}$ at very low frequencies could be due to the short section of the duct, which is simulated upstream of the orifice. For the matrix element $\mathbf{S}_{22}$, i.e., the downstream side reflection coefficient R^-, variation in both experimental data and the simulation is observed. It should be noted that parts of the discrepancies are probably due to measurement errors.

Calculations on finer meshes and with dynamic viscosity closer to that of air were carried out in a limited frequency range, but neither refinements showed significant improvements of the results compared to those presented in Fig. 12.

As can be seen, the elements of the scattering matrix have a rather weak dependence on the frequency for this flow case. In addition, the accuracy does not seem to be reduced at high frequencies even though a coarse mesh is used. The assumption of isentropic relations between pressure and density seems to be justified, considering these results, but a further investigation on this issue is needed.

X. CONCLUSIONS

A new methodology for efficient simulations of acoustic wave propagation in ducts with sharp edges have been proposed and evaluated. Acoustic wave propagation calculations have been performed with a frequency domain linearized Navier–Stokes equation solver. The methodology takes into account the main effects of the mean flow field on the acoustic waves. The mean flow is assumed steady, but otherwise no further assumptions are made on the mean flow field. As a test case, the acoustic response of an in-duct orifice plate with mean flow is simulated. The simulated results are compared to measurements and theory with excellent agreement. The results justify the assumptions of isentropic relations. In addition, it demonstrates the possibility to perform calculations for rectangular two-dimensional geometry in order to study the acoustic properties of a cylindrical duct. We can conclude that a frequency domain formulation of the linearized Navier–Stokes equations is a promising methodology for simulation of acoustic propagation properties in-duct systems and that this is a way to efficiently determine acoustic two-ports from numerical simulations.

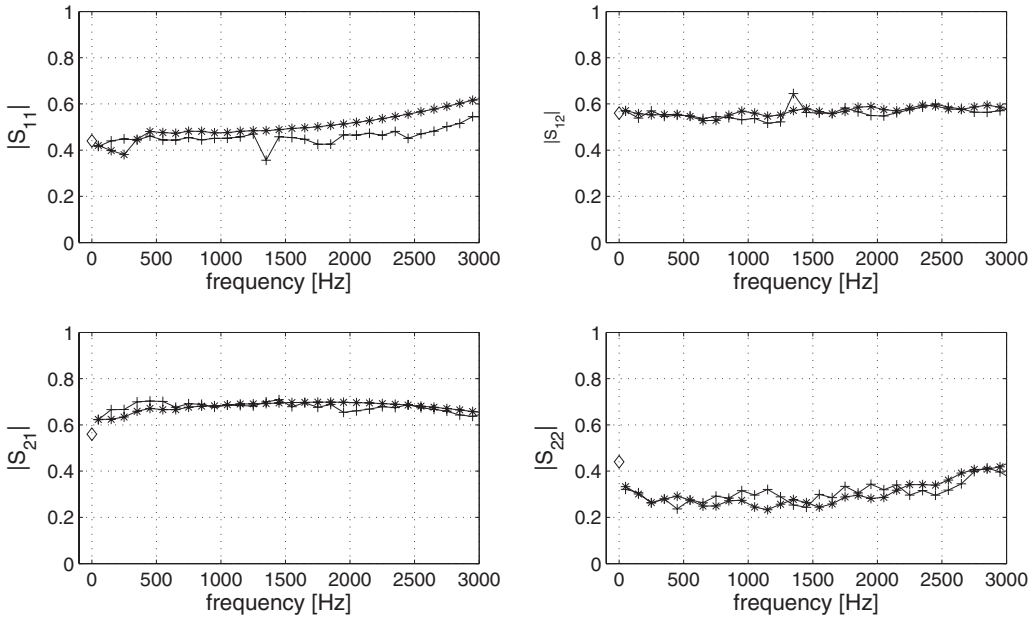


FIG. 12. Calculated scattering matrix for the orifice. Comparison to measurements and theory with $M=0.068$. The maximum frequency is $f=3000$ Hz, which corresponds to a duct Helmholtz number of He=3.1. Measurements: (Ref. 23) −+−; simulations: −∗−; low-frequency incompressible model: (Refs. 23, 29, and 30) ◇.

ACKNOWLEDGMENTS

The authors would like to thank Professor M. Åbom for valuable discussions on the fundamental aspects of acoustics and the contents of this paper. This work has been performed within the research project "Reduction of external noise for diesel propulsed passenger cars," sponsored by the EMFO (Grant No. Dnr AL 90B2004: 156660) and the Vinnova/Green Car (Grant No. Dnr 2005-00059) in which the participating partners are Saab Automobile AB, KTH, and Chalmers. The funding is gratefully acknowledged.

[1]M. Åbom, "Measurement of the scattering-matrix of acoustical two-ports," Mech. Syst. Signal Process. **5**, 89–104 (1991).
[2]K. S. Peat, "Evaluation of four-pole parameters for ducts with flow by the finite element method," J. Sound Vib. **84**, 389–395 (1982).
[3]R. Glav, "The transfer matrix for a dissipative silencer of arbitrary cross-section," J. Sound Vib. **236**, 575–594 (2000).
[4]C. J. Young and M. Crocker, "Prediction of transmission loss in mufflers by the finite-element method," J. Acoust. Soc. Am. **57**, 144–148 (1975).
[5]Y. Özyörük, E. Dizemeny, S. Kayay, and A. Aktürkz, "A frequency domain linearized Euler solver for turbomachinery noise propagation and radiation," Proceedings of the 13th AIAA/CEAS Aeroacoustics Conference (2007).
[6]P. P. Rao and P. J. Morris, "Use of finite element methods in frequency domain aeroacoustics," AIAA J. **44**, 1643–1652 (2006).
[7]D. W. Bechert, "Sound absorption caused by vorticity shedding, demonstrated with a jet flow," J. Sound Vib. **70**, 389–405 (1980).
[8]Y. Fukumoto and M. Takay, "Vorticity production at the edge of a slit by sound waves in the presence of a low-Mach-number bias flow," Phys. Fluids A **3**, 3080–3082 (1991).
[9]J. C. Wendoloski, "Sound absorption by an orifice plate in a flow duct," J. Acoust. Soc. Am. **104**, 122–132 (1998).
[10]S. Föller, R. Kaess, and W. Polifke, "Reconstruction of acoustic transfer matrices from large-eddy-simulations of compressible turbulent flows," Proceedings of the 14th AIAA/CEAS Aeroacoustics Conference (2008), AIAA Paper No. 2008-3046.
[11]R. Arina, R. Malvano, A. Piccato, and P. G. Spazzini, "Numerical and experimental analysis of sound generated by an orifice," Proceedings of the 13th AIAA/CEAS Aeroacoustics Conference (2007).
[12]X. Gloerfelt and P. Lafon, "Direct computation of the noise induced by a turbulent flow through a diaphragm in a duct at low Mach number," Comput. Fluids **37**, 388–401 (2008).
[13]C. K. W. Tam, K. A. Kurbatskii, K. K. Ahuja, and R. J. Gaeta, "A numerical and experimental investigation of the dissipation mechanisms of resonant acoustic liners," J. Sound Vib. **245**, 545–557 (2001).
[14]C. K. W. Tam, H. Ju, and B. E. Walker, "Numerical simulation of a slit resonator in a grazing flow under acoustic excitation," J. Sound Vib. **313**, 449–471 (2008).
[15]R. C. K. Leung, R. M. C. So, M. H. Wang, and X. M. Li, "In-duct orifice and its effect on sound absorption," J. Sound Vib. **299**, 990–1004 (2007).
[16]S. Mendez and J. D. Eldredge, "Acoustic modeling of perforated plates with bias flow for large-eddy simulations," J. Comput. Phys. **228**, 4757–4772 (2009).
[17]G. Efraimsson, T. Pieper, M. Åbom, and S. Boij, "Simulation of wave scattering at an orifice by using a Navier–Stokes solver," Proceedings of the 13th AIAA/CEAS Conference (2007), AIAA Paper No. 2007-3402.
[18]S. Boij and B. Nilsson, "Reflection of sound at area expansions in a flow duct," J. Sound Vib. **260**, 477–498 (2003).
[19]S. Boij and B. Nilsson, "Scattering and absorption of sound at flow duct expansions," J. Sound Vib. **289**, 577–594 (2006).
[20]S. Temkin, *Elements of Acoustics* (Acoustical Society of America, Melville, New York, 2001).
[21]D. S. Henningson and M. Berggren, *Fluid Dynamics: Theory and Computation* (KTH, Stockholm, Sweden, 2005).
[22]U. Ingård and H. Ising, "Acoustic nonlinearity of an orifice," J. Acoust. Soc. Am. **42**, 6–17 (1967).
[23]M. Åbom, S. Allam, and S. Boij, "Aero-acoustics of flow duct singularities at low mach numbers," Proceedings of the 12th AIAA/CEAS Aeroacoustics Conference (2006).
[24]A. D. Pierces, *Acoustics: An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981).
[25]F. C. Karal, "The analogous acoustical impedance for discontinuities and constrictions of circular cross section," J. Acoust. Soc. Am. **25**, 327–334 (1953).
[26]A. Kierkegaard and G. Efraimsson, "Generation and propagation of sound waves in low mach number flows," Proceedings of the 13th AIAA/CEAS Aeroacoustics Conference (2007).
[27]M. Åbom, *An Introduction to Flow Acoustics* (KTH, Stockholm, Sweden, 2006).
[28]M. L. Munjal and A. G. Doice, "Theory of a two source-location method for direct experimental evaluation of the four-pole parameters of an aeroacoustic element," J. Sound Vib. **141**, 323–333 (1990).
[29]G. C. J. Hofmans, "Vortex sound in confined flows," Ph.D. thesis, Eindhoven University of Technology, Eindhoven, The Netherlands (1998).
[30]P. Durrieu, G. Hofmans, G. Ajello, R. Boot, Y. Aurégan, A. Hirschberg, and M. C. A. M. Peters, "Quasisteady aero-acoustic response of orifices," J. Acoust. Soc. Am. **110**, 1859–1872 (2001).

Inverse problem in anisotropic poroelasticity: Drained constants from undrained ultrasound measurements

James G. Berryman and Seiji Nakagawa
Earth Sciences Division, Lawrence Berkeley National Laboratory, One Cyclotron Road MS90-R1116, Berkeley, California 94720

(Received 21 July 2009; revised 20 November 2009; accepted 23 November 2009)

Poroelastic analysis has traditionally focused on the relationship between dry and drained constants, which are assumed known, and the saturated or undrained constants, which are assumed unknown. However, there are many applications in this field of study for which the main measurements can only be made on the saturated/undrained system, and then it is uncertain what the effects of the fluids were on the system, since the drained constants remain a mystery. The work presented here shows how to deduce drained constants from undrained constants for anisotropic systems having symmetries ranging from isotropic to orthotropic. Laboratory ultrasound data are then inverted for the drained constants in three granular packings: one of glass beads, and two others for distinct types of more or less angular sand grain packings. Experiments were performed under uniaxial stress, which resulted in hexagonal (transversely isotropic) symmetry of the poroelastic response. One important conclusion from the general analysis is that the drained constants are uniquely related to the undrained constants, assuming that porosity, grain bulk modulus, and pore fluid bulk modulus are already known. Since the resulting system of equations for all the drained constants is linear, measurement error in undrained constants also propagates linearly into the computed drained constants. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277162]

PACS number(s): 43.30.Ma, 43.30.Pc, 43.20.Jr [AIT] Pages: 720–729

I. INTRODUCTION

Analysis of poroelastic systems typically focuses on how poroelastic moduli from measurements of undrained samples are related to the pore fluid and drained system elastic properties.[1] These technical terms in poroelasticity sometimes have slightly different names and meanings throughout literature, but the approximate correspondence follows this recipe: undrained ≃ jacketed ≃ unrelaxed, while drained ≃ relaxed and unjacketed ≃ solid ≃ mineral. There are many useful and pertinent references.[2–13] But the inverse problem of deducing drained properties from undrained properties has seldom been addressed by analytical means.[14,15] Nevertheless, this important issue arises many times, especially in oil and gas reservoir field measurements, because it may not be practical to determine the drained moduli separately, as it is typically impossible to drain the system—at least over the timescales when this information would be most useful. Another example is that of ocean sediment measurements made in place,[5,12,16] and similarly in soil fluidization during earthquakes.[17,18] Further examples include medical applications, particularly studies of bone and osteoporosis.[19,20]

One purpose of the present work is therefore to show how to deduce these hard-to-measure drained quantities from the easier-to-measure undrained moduli, without requiring the fairly common (but always tedious and uncertain) iterative and/or fitting techniques that are typically used for such purposes at the present time.

We begin by discussing isotropic systems and then progress to orthotropic systems. We ultimately concentrate on transversely isotropic ones as these are the ones for which we currently have laboratory ultrasound data to invert.

II. POROUS MATERIALS CONTAINING FLUIDS

We first consider isotropic and microhomogeneous elastic materials containing voids or pores, where the pores are connected (having finite permeability) and fluid-filled. This situation is the one originally studied by Gassmann.[21] One simplification that arises immediately is due to the fact that the presence of pore fluids—in microhomogeneous and overall isotropic porous media[11,21]—has no mechanical effect on shear moduli that are not coupled to the principal stresses, so $\mu^u = \mu^d$, which means that the undrained (u superscript) shear modulus is the same as the drained (d superscript) shear modulus, for the overall porous system. There can be (and often may be in practice) other effects on the shear moduli due to the presence of pore fluids, such as softening of cementing materials or expansion of interstitial clays, which we shall term "chemical" effects to distinguish them from the purely mechanical effects to be considered here. We neglect all such chemical effects in the following analysis. This means that mechanical and acoustical analyses for the effective shear moduli are typically unaltered by the presence of fluids.[22] Thus, we may simplify our system of equations in order to focus only on those significant parts of the analysis, i.e., the ones that change due to the mechanical properties of the pore fluids.

A. Isotropy

We start the analysis by showing Gassmann's equation,[21] which for isotropic systems is sometimes written in the following form:

$$K^u = K^d + \alpha^2/[(\alpha - \phi)/K_s + \phi/K_f], \tag{1}$$

where $\alpha \equiv 1 - K^d/K_s$ is the effective stress or Biot–Willis coefficient,[4] K_s is the solid modulus of the grains (assumed to be a homogeneous collection composed of grains of the same mineral), K_f is the fluid modulus, and ϕ is the porosity. Now, by introducing a modulus for a fluid suspension having the same solid and fluid components, but the corresponding drained modulus is $K^d \equiv 0$, we find that the effective modulus is given by

$$K_{\text{susp}} = \left[\frac{1-\phi}{K_s} + \frac{\phi}{K_f} \right]^{-1}. \tag{2}$$

In fact, this result follows directly from Gassmann's formula by setting $K^d = 0$ everywhere in Eq. (1). But of course, it is also a well-known exact result[23] in mechanics and acoustics for these types of fluid-solid suspensions.

Rewriting Gassmann's formula in these terms, we find first that

$$K^u = K^d + \frac{(1 - K^d/K_s)^2}{1/K_{\text{susp}} - K^d/K_s^2}. \tag{3}$$

Now, if we simply multiply through by the denominator on the right hand side, we find

$$K^u \left(\frac{1}{K_{\text{susp}}} - \frac{K^d}{K_s^2} \right) = 1 - 2\frac{K^d}{K_s} + \frac{K^d}{K_{\text{susp}}}. \tag{4}$$

Note that two terms of the form $(K^d/K_s)^2$ have been canceled from this expression. Once these convenient cancellations have occurred, K^d now appears only linearly in this expression. The equation can therefore be solved immediately for K^d in terms of the undrained modulus K^u, and the other factors that are also assumed to be known (and, in fact, these other factors are typically easier to measure than either K^u or K^d). Finally, we find

$$K^d = \left(\frac{K^u}{K_{\text{susp}}} - 1 \right) [1/K_{\text{susp}} - 2/K_s + K^u/K_s^2]^{-1}. \tag{5}$$

This form is very useful in our present applications, but so far it applies only to the fully isotropic case. We show next that a very similar (but nevertheless distinct) formula applies to the anisotropic cases under consideration here. This version of the formula (for the isotropic case only) has been derived previously by Zhu and McMechan.[14,15]

B. Orthotropy

If the overall porous medium is anisotropic, either due to some preferential alignment of the constituent particles or due to externally imposed stress (such as a gravity field and weight of overburden, for example), we consider the orthorhombic anisotropic version of the poroelastic equations

$$\begin{pmatrix} e_{11} \\ e_{22} \\ e_{33} \\ -\zeta \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} & -\beta_1 \\ s_{12} & s_{22} & s_{23} & -\beta_2 \\ s_{13} & s_{23} & s_{33} & -\beta_3 \\ -\beta_1 & -\beta_2 & -\beta_3 & \gamma \end{pmatrix} \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ -p_f \end{pmatrix}. \tag{6}$$

The e_{ii} are strains in the $i=1,2,3$ directions. The σ_{ii} are the corresponding stresses. The fluid pressure is p_f. The increment of fluid content is ζ. The drained compliances are $s_{ij} = s_{ij}^d$. Undrained compliances (not yet shown) are symbolized by s_{ij}^u. Coefficients $\beta_i = s_{i1} + s_{i2} + s_{i3} - 1/3K_s$, where K_s is again the solid modulus of the grains (assumed uniform for simplicity here). The Reuss average bulk modulus[24,25] is defined by

$$\frac{1}{K_R^d} = \sum_{ij=1,2,3} s_{ij}^d. \tag{7}$$

A similar definition of K_R^u, with undrained compliances replacing drained compliances, will also be needed later in our discussion. The alternative Voigt average[25,26] of the stiffnesses ($K_V^d = \Sigma_{ij=1,2,3} c_{ij}^d/9$) will play no significant role in our present discussion, although we will show some of these values later with the examples. And, finally, $\gamma = \Sigma_{i=1-3} \beta_i / B K_R^d$, where B is Skempton's coefficient,[27] which will be defined carefully later in our discussion.

The shear terms due to twisting motions (i.e., strains e_{23}, e_{31}, and e_{12}, and stresses σ_{23}, σ_{31}, and σ_{12}) are excluded from the present discussion since they typically do not couple to the modes of interest for anisotropic systems having orthotropic symmetry, or more symmetric systems such as transversely isotropic or isotropic. We have also assumed that we know the true axes of symmetry, and have made use of them in our formulation of the problem; this helps to eliminate coupling between the modes shown here and those shear modes not displayed (which can happen for systems less symmetric than orthorhombic). Note that s_{ij} represent the elements of the compliance matrix $\mathbf{S}$ and are all independent of the fluid, and therefore would be the same if the medium were treated as elastic (i.e., by ignoring the fluid pressure, or assuming that the fluid saturant is air). We typically term these compliances as the drained compliances, and the corresponding matrix as the drained compliance matrix $\mathbf{S}^d$, since the fluids do not contribute to the mechanical energy if they are free to drain into a surrounding reservoir containing the same type of fluid. In contrast, the undrained compliance matrix $\mathbf{S}^u$ presupposes that the fluid is trapped (unable to drain from the system into an adjacent reservoir), and therefore contributes in a significant and measurable way to the compliance and stiffness ($\mathbf{C}^u = [\mathbf{S}^u]^{-1}$) of the undrained system.

The poroelasticity matrix in Eq. (6) is in compliance form and has extremely simple poroelastic behavior in the sense that *all the fluid mechanical effects* appear only in the single coefficient γ. We can simplify the notation a little more by lumping some coefficients together, combining the 3×3 submatrix in the upper left corner of the matrix in Eq. (6) into $\mathbf{S}$, and defining the column vector $\mathbf{b}$ by

$$\mathbf{b}^T \equiv (\beta_1, \beta_2, \beta_3). \tag{8}$$

The resulting 4×4 matrix and its inverse are now related by

$$\begin{pmatrix} \mathbf{S} & -\mathbf{b} \\ -\mathbf{b}^T & \gamma \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{q} \\ \mathbf{q}^T & z \end{pmatrix}^{-1}, \tag{9}$$

where the elements of the inverse matrix can be shown to be written in terms of drained stiffness matrix $\mathbf{C}^d = \mathbf{C} = \mathbf{S}^{-1}$ by introducing three components: (a) the scalar $z = [\gamma - \mathbf{b}^T \mathbf{C} \mathbf{b}]^{-1}$, (b) the column vector $\mathbf{q} = z\mathbf{C}\mathbf{b}$, and (c) the un-

drained 3×3 stiffness matrix (i.e., the pertinent one connecting the principal strains to principal stresses) is given by $\mathbf{A}=\mathbf{C}+z\mathbf{C}\mathbf{b}\mathbf{b}^T\mathbf{C}=\mathbf{C}^d+z^{-1}\mathbf{q}\mathbf{q}^T\equiv\mathbf{C}^u$, since $\mathbf{C}^d$ is the drained stiffness and $\mathbf{A}=\mathbf{C}^u$ is clearly the undrained stiffness by construction. This fact is most easily recognized by noting the inverse relationship, showing that

$$\begin{pmatrix}\mathbf{A} & \mathbf{q}\\ \mathbf{q}^T & z\end{pmatrix}\begin{pmatrix}\mathbf{E}\\ -\zeta\end{pmatrix}=\begin{pmatrix}\mathbf{\Sigma}\\ -p_f\end{pmatrix}, \tag{10}$$

where

$$\mathbf{E}=(e_{11},e_{22},e_{33})^T \quad \text{and} \quad \mathbf{\Sigma}=(\sigma_{11},\sigma_{22},\sigma_{33})^T. \tag{11}$$

When $\zeta=0$ on the left hand side, the only possible interpretation of the equation is that the nonzero contributions are for the undrained case, and so $\mathbf{A}=\mathbf{C}^u$ is necessarily the undrained stiffness by inspection.

The results contained in Eqs. (9) and (10) are also equivalent to those of Gassmann[21] for both isotropic and anisotropic porous media, as we shall show later in this paper, although Gassmann's results were presented in a form somewhat harder to scan. Also, note the important fact that the observed decoupling of the fluid effects occurs only in the compliance form (6) of the equations, and never in the stiffness (inverse) form for the poroelasticity equations, since all terms in $\mathbf{A}$, $\mathbf{q}$, and z contain fluid-dependent components.

From these results, it is not hard to show that

$$\mathbf{S}^d=\mathbf{S}^u+\gamma^{-1}\mathbf{b}\mathbf{b}^T. \tag{12}$$

This result shows the remarkably simple result that the drained compliance matrix can be found directly from knowledge of the inverse of undrained compliance, and the still unknown, but sometimes relatively easy to estimate, values of γ and the three distinct orthotropic β_i coefficients, for $i=1,2,3$.

Now we make further progress by considering again the Reuss averages of both of the drained and undrained orthotropic compliances

$$\frac{1}{K_R^d}\equiv s_{11}^d+s_{22}^d+s_{33}^d+2(s_{12}^d+s_{13}^d+s_{23}^d) \tag{13}$$

and

$$\frac{1}{K_R^u}\equiv s_{11}^u+s_{22}^u+s_{33}^u+2(s_{12}^u+s_{13}^u+s_{23}^u). \tag{14}$$

These effective moduli are the Reuss averages of the nine compliances in the upper left 3×3 of the full (including the uncoupled shear components) 6×6 compliance matrix for the two cases, respectively, when the pore fluid is allowed to drain from the porous system, and when the pore fluid is trapped by a jacketing material and therefore undrained.

Although the significance of the formula is somewhat different now, we find again that

$$\beta_1+\beta_2+\beta_3=\frac{1}{K_R^d}-\frac{1}{K_s}=\frac{\alpha_R}{K_R^d} \tag{15}$$

if we also choose to define a Reuss average effective stress coefficient (or Biot–Willis[4] coefficient) by

$$\alpha_R\equiv 1-K_R^d/K_s, \tag{16}$$

in analogy to the isotropic case. Furthermore, we have

$$\gamma=\frac{\beta_1+\beta_2+\beta_3}{B}=\frac{\alpha_R}{K_R^d}+\phi\left(\frac{1}{K_f}-\frac{1}{K_s}\right), \tag{17}$$

since we have the rigorous result in this notation (see Ref. 10) that Skempton's B coefficient is given by

$$B\equiv\frac{1-K_R^d/K_R^u}{1-K_R^d/K_s}=\frac{\alpha_R/K_R^d}{\alpha_R/K_R^d+\phi(1/K_f-1/K_s)}. \tag{18}$$

We emphasize that all these formulas are rigorous statements based on the preceding anisotropic analysis. Note that the appearance of the Reuss averages K_R^d and α_R is not an approximation, but merely a useful choice of notation made here because it will enable us to see clearly the similarity between the rigorous anisotropic formulas and the corresponding isotropic ones.

C. Reuss average moduli

We are now in position to develop the analogy between the isotropic and anisotropic Gassmann equations. In particular, the equation for the suspension modulus in Eq. (2) does not change at all. The equation of the effective undrained bulk modulus K^u, as shown in both Eqs. (1) and (3), changes only in that the relationship is now between the Reuss averages K_R^u and K_R^d of these quantities. The result is completely analogous to Eq. (3), and is given for pertinent Reuss average quantities by

$$K_R^u=K_R^d+\frac{(1-K_R^d/K_s)^2}{1/K_{\text{susp}}-K_R^d/K_s^2}. \tag{19}$$

The remainder of the argument is virtually identical to the isotropic case, and the final result for the drained modulus is

$$K_R^d=\left(\frac{K_R^u}{K_{\text{susp}}}-1\right)[1/K_{\text{susp}}-2/K_s+K_R^u/K_s^2]^{-1}. \tag{20}$$

This formula shows how to invert for drained Reuss bulk modulus K_R^d from knowledge of K_R^u, ϕ, K_f, and K_s in an anisotropic (up to orthotropic) poroelastic system.

Clearly, this formula does not yet give us the individual matrix elements s_{ij}^d directly. Nevertheless, Eq. (20) was the hardest step in the procedure that follows, and the rest of the steps are easily taken, once we have this general result.

First note that, from Eqs. (15) and (17), it also follows that $\gamma^{-1}=BK_R^d/\alpha_R$. So we can now rearrange Eq. (12) to give

$$s_{ij}^d=s_{ij}^u+\frac{BK_R^d}{\alpha_R}\beta_i\beta_j \quad \text{for} \quad i,j=1,2,3. \tag{21}$$

At this point in the analysis, we know everything needed for the inversion except for the three coefficients β_i, for $i=1,2,3$. But, by taking appropriate sums of Eq. (21) and recalling Eq. (15), we find that

$$\beta_i = s_{i1}^d + s_{i2}^d + s_{i3}^d - \frac{1}{3K_s} = s_{i1}^u + s_{i2}^u + s_{i3}^u - \frac{1}{3K_s} + B\beta_i. \tag{22}$$

We used Eq. (15) again to show that the product term resulting from Eq. (21) is $(K_R^d/\alpha_R)(\beta_1+\beta_2+\beta_3)=1$. Rearranging Eq. (22), we find that

$$\beta_i = \left[s_{i1}^u + s_{i2}^u + s_{i3}^u - \frac{1}{3K_s} \right] / (1-B). \tag{23}$$

Formula (18) for Skempton's coefficient[27] determines B, exactly in terms of previously known quantities. So, all three β_i (which are themselves drained constants) and also γ [using Eq. (17)] are now precisely determined. All the remaining constants s_{ij}^d can then be found directly from Eq. (21). Note that all the steps in this inversion procedure are linear; there is no need to solve any quadratic (or higher order) equation in this formulation of the undrained-to-drained inversion problem. As we shall demonstrate in the examples, it is sufficient to know the c_{ij}^u, porosity ϕ, and the two bulk moduli K_s and K_f, in order to determine first the β_i from Eq. (23), and then the s_{ij}^d from Eq. (21). Once all the s_{ij}^d for all $i,j=1,2,3$ are known, then the corresponding stiffnesses are found by inverting the pertinent 3×3 matrix of compliances.

III. EXAMPLES

Recently, one of us [together with other colleagues at LBNL (Ref. 28)] developed a new technology for measuring the seismic/ultrasonic anisotropy of soft sediments during their compaction. This method is based on a specially designed compaction cell, equipped with three pairs of ultrasonic transducers, including one *P*-wave and two *S*-wave transducer pairs. The setup also includes a phased *P*-wave piezoelectric source array with 32 elements, and a sensor to generate steerable ultrasonic plane waves. The central (or resonance) frequency of the transducers is in the range 500 kHz–1 MHz. Compared to previously existing methods, this method has the capability to (a) determine all five of the independent elastic constants in a transversely isotropic (TI) medium, (b) to conduct measurements continuously while the sedimentary material is compacted *in-situ* (i.e., in the uniaxially pressurized chamber), and (c) does not require single or multiple directionally cored samples (such coring, when needed, actually releases stress of the compacted sediments and therefore does not give accurate measurements of *in-situ* conditions). Uniaxial loading was taken to a maximum of 5 MPa.

The spectral content of the measured ultrasonic waves changes significantly during compaction of fluid-saturated sediments. These changes occur primarily because of the scattering of high-frequency acoustic energy, which can be large—particularly for low-velocity *S*-waves. To reduce the phase-velocity measurement errors related to the changes in the central frequency of the waves, and to reduce the impact of the high-frequency wave scattering in the examples presented in this paper, amplitude spectrum of the measured wave was filtered to have the shape of a Ricker wavelet with a central frequency of 500 kHz for *P*-waves and 50 kHz for *S*-waves. From each measurement, the propagation time was determined from the arrival time of a dominant phase (e.g., first peak or trough), compared to the reference measurements (face-to-face measurements of the transducer couples). Four out of the five elastic constants of a TI medium can be determined directly from the wave velocities along the axes of the sample (one *P*-wave and two *S*-wave measurements along the horizontal plane, and one *P*-wave measurement in the vertical direction). The last constant is determined through a numerical inversion method[29] using phase slownesses of the *P*-waves steered in a range of directions.

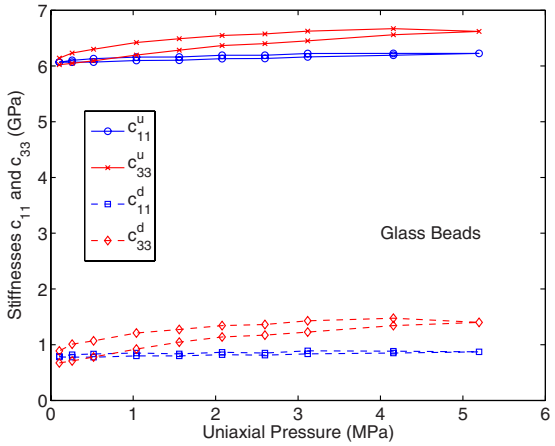


FIG. 1. (Color online) Plots of measured undrained (u) and inverted drained (d) constants for the stiffnesses c_{11} and c_{33} in glass bead samples under uniaxial pressure. The abscissa values are those of the uniaxial stress applied to the anisotropic system.

Further details of these experimental methods will be presented later in a companion paper. The present work will concentrate on showing examples of the inversion results for drained constants from the measured values of undrained constants.

A. Examples for transversely isotropic glass bead samples

Experiments were performed on glass bead samples (potter industry, soda lime glass, size 212–250 μm). Figures 1 and 2 display plots of measured undrained (u) and inverted drained (d) constants for stiffnesses obtained by performing ultrasound experiments and measuring the wave speeds in different directions. For example, stiffness $c_{33}^u=\rho v_p^2$, where ρ is the fluid-saturated system density and compressional wave speed v_p is measured along the symmetry axis in our glass bead samples under uniaxial pressure. Figure 1 shows both drained and undrained constants c_{11} and c_{33}, while Fig. 2 shows the same types of results for c_{12} and c_{13}. The abscissa values are the applied uniaxial stress. These glass beads had $K_s=40.7$ GPa, $\mu_s=29.7$ GPa, and density $\rho_s=2.46\times10^3$ kg/m^3 (for comparison, also see Ref. 30). Bulk modulus $K_f=2.2$ GPa for the water saturant. Porosity varied with stress level and was monitored, with values in the range $0.370\leq\phi\leq0.375$. Two (or more for the sands) distinct curves are visible for each stiffness as the stress was first cycled up and then down again (repeatedly for the sands), while measurements were made during both (multiple)

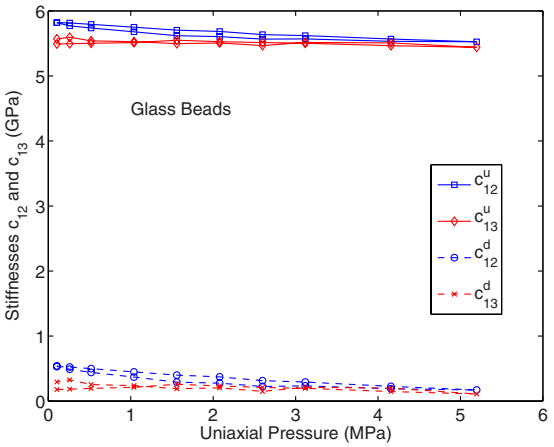


FIG. 2. (Color online) Plots of measured undrained (u) and inverted drained (d) constants for the stiffnesses c_{12} and c_{13} in glass bead samples under uniaxial pressure. The abscissa values are the uniaxial stresses applied to the anisotropic system.

phases of the test. Figure 3 shows the results for the computed Voigt and Reuss bulk moduli for both measured undrained (K_V^u and K_R^u) and the computed drained cases (K_V^d and K_R^d). Figure 4 shows the measured values of porosity for the cases in Figs. 1–3, and also the uniaxial effective stress coefficient X_3 (see Appendix B).

B. Examples for sand samples: K and OK

The next two sets of figures follow the same sequencing as that used for presenting the glass bead data. The bulk modulus for the sands was taken as that for quartz, so K_s =38.0 GPa, and G_s=44.0 GPa in both cases. The density of the sand grains was $\rho=2.65\times10^3$ kg/m^3. Visual inspection shows that the K sand grains are more angular, and the OK

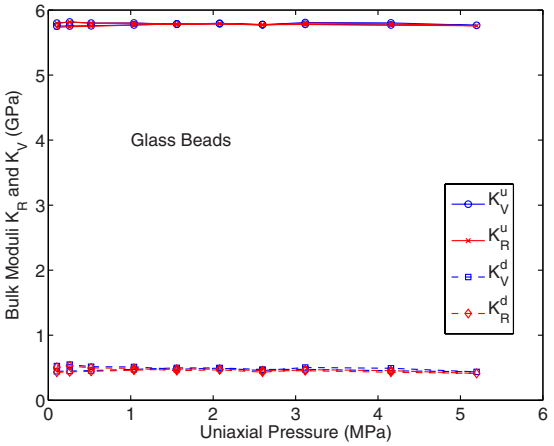


FIG. 3. (Color online) For glass bead samples, examples of the undrained and drained bulk modulus estimators K_R and K_V, being the appropriate Reuss and Voigt averages of the compliance and stiffness matrices, respectively. Undrained values were measured, and drained values were computed using the formulas in the text. The drained Voigt averages are found by first computing the drained compliances, then inverting for the drained stiffnesses, and finally averaging. The abscissa values are the uniaxial stresses applied to the anisotropic system. These glass beads had K_s=40.7 GPa, μ_s=29.7 GPa, and density $\rho_s=2.46\times10^3$ kg/m^3. Bulk modulus K_f =2.2 GPa for water. Porosity varied with stress level and was monitored, with values in the range $0.370\leq\phi\leq0.375$.

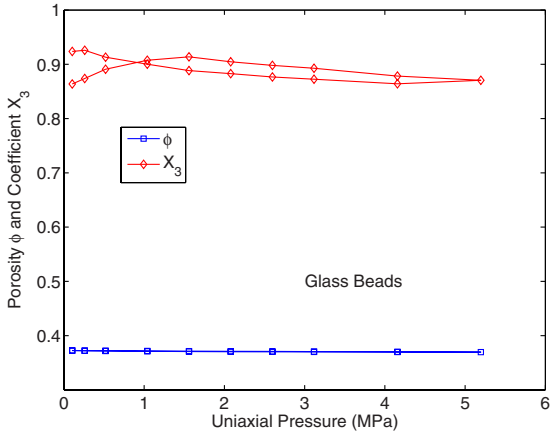


FIG. 4. (Color online) Plots of measured porosity ϕ and computed effective uniaxial stress coefficient X_3 in glass bead samples.

No. 1 sand grains are significantly less angular. K sand results are presented in Figs. 5–8, while those for OK sand are seen in Figs. 9–12.

C. Summary of the inversion procedure

The first step of the inversion procedure is to compute the undrained Reuss bulk modulus K_R^u using Eq. (14) from measured undrained compliances (or stiffnesess—when these are the measured data as in ultrasound data—by first converting them to compliances). The suspension bulk modulus K_{susp} is also computed from Eq. (2) using known values of porosity ϕ, solid grain bulk modulus K_s, and pore fluid modulus K_f. These values and K_s again are the only ones needed in formula (20). Once K_R^d is known, then α_R $=1-K_R^d/K_s$ is known, and Skempton's coefficient can be computed from Eq. (18). Once B is known, then β_1, β_2, and β_3 can all be obtained from Eqs. (21) and (22). Once all these constants are known, Eq. (12) is used to compute the individual drained compliance matrix elements. Then, if so desired, the drained stiffness matrix elements may also be computed by inverting the drained compliance matrix. The remaining compliance and stiffness coefficients (i.e., s_{44}, s_{55},

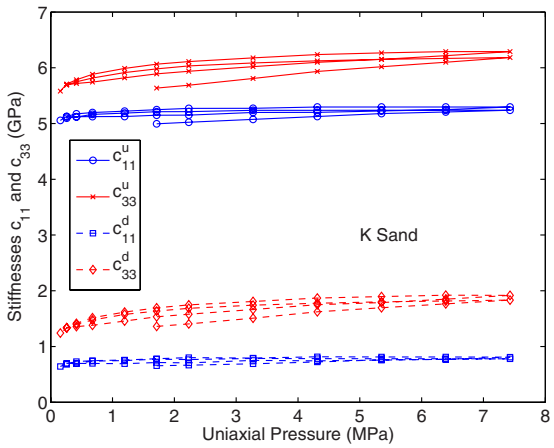


FIG. 5. (Color online) Plots of measured undrained (u) and inverted drained (d) constants for the stiffnesses c_{11} and c_{33} in K sand samples under uniaxial pressure. The abscissa values are the uniaxial stresses applied to the anisotropic system.

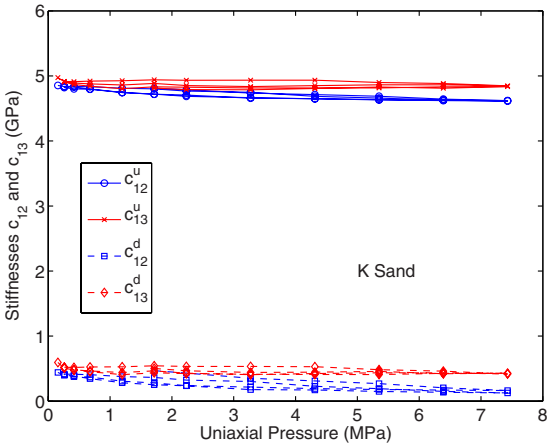


FIG. 6. (Color online) Plots of measured undrained (u) and inverted drained (d) constants for the stiffnesses c_{12} and c_{13} in K sand samples under uniaxial pressure. The abscissa values are the uniaxial stresses applied to the anisotropic system.

and s_{66}—and, therefore, c_{44}, c_{55}, and c_{66}) are not affected by the presence of the fluid, in transversely isotropic (or more generally orthotropic) materials, so their values are the same for both drained and undrained systems.

Also note that the consistently large offset observed between K_R^u and K_R^d is easily understood from the fact that, when $K_R^d \ll K_s$, Eq. (19) implies

$$K_R^u \simeq K_{\text{susp}} + K_R^d, \tag{24}$$

if we neglect terms of order K_R^d/K_s compared to unity. So the observed offset is almost entirely due to the large value of the suspension modulus K_{susp}, although there are some small corrections of order K_R^d/K_s neglected by Eq. (24).

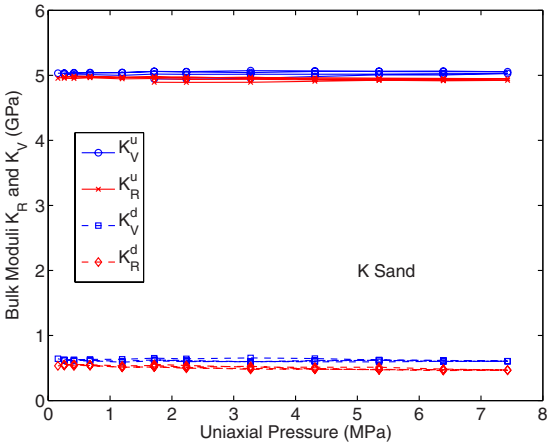


FIG. 7. (Color online) For K sand samples, examples of the undrained and drained bulk modulus estimators K_R and K_V, being the appropriate Reuss and Voigt averages of the compliance and stiffness matrices, respectively. The abscissa values are the uniaxial stresses applied to the anisotropic system. These sand grains had K_s=38.0 GPa, μ_s=44.0 GPa, and density ρ_s =2.65×10^3 kg/m^3. Bulk modulus K_f=2.2 GPa for water. Porosity varied with stress level and was monitored, with values in the range $0.445 \leq \phi \leq 0.480$. Two distinct curves are visible for each stiffness as the stress was first cycled up and then down again, while measurements were made during both phases of the test.

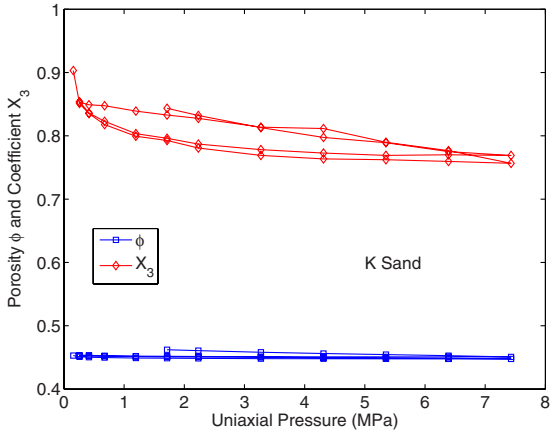


FIG. 8. (Color online) Plots of measured porosity ϕ and computed effective uniaxial stress coefficient X_3 in K sand samples.

IV. DISCUSSION

The analysis presented in this paper has apparently been wholly quasi-static, treating the stiffnesses and bulk moduli as if they were entirely frequency independent so that Gassmann's quasi-static analysis for either isotropic or anisotropic porous media may be used throughout. The authors are, of course, well aware of Biot's theory[2,3] of fluid-saturated porous media. This theory introduces the interesting complication of the Biot slow-wave,[7,30] which is an out-of-phase motion of the fluid and solid frame that produces strongly fluid-dependent attenuation and scattering mechanisms. Since the experiments that motivated the present work were not designed to measure the attenuation of the resulting waves, and since the very confined nature of the experimental chamber essentially prevents excitation of the Biot slow-wave,[31] it seems unnecessary for our present purposes to use the full Biot theory for the analysis. Thus, we have concentrated instead on the Gassmann approach to the problem.

This treatment does not necessarily imply that frequency plays no role in the experiment. For example, there are good

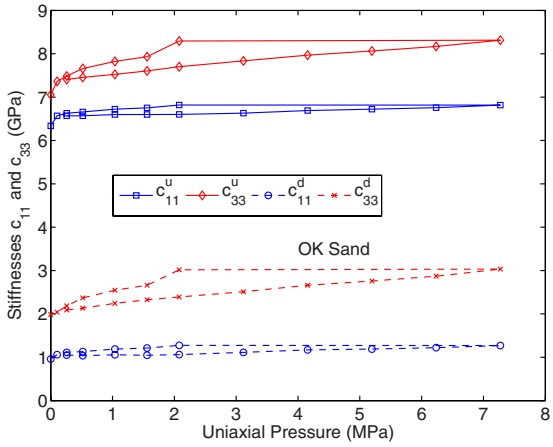


FIG. 9. (Color online) Plots of measured undrained (u) and inverted drained (d) constants for the stiffnesses c_{11} and c_{33} in OK sand samples under uniaxial pressure. Two distinct curves are visible for each stiffness as the stress was first cycled up and then down again, while measurements were made during both phases of the test. The abscissa values are the uniaxial stresses applied to the anisotropic system.

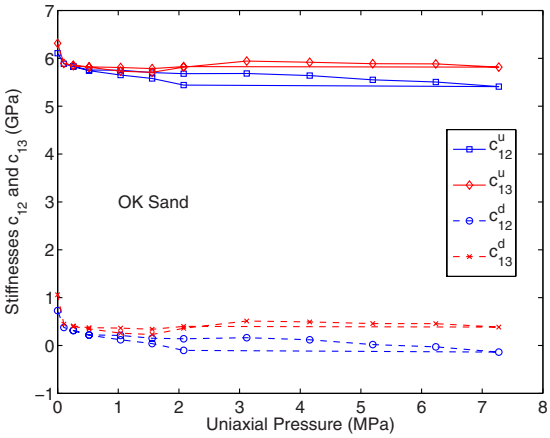


FIG. 10. (Color online) Plots of measured undrained (u) and inverted drained (d) constants for the stiffnesses c_{12} and c_{13} in OK sand samples under uniaxial pressure. The abscissa values are the uniaxial stresses applied to the anisotropic system.

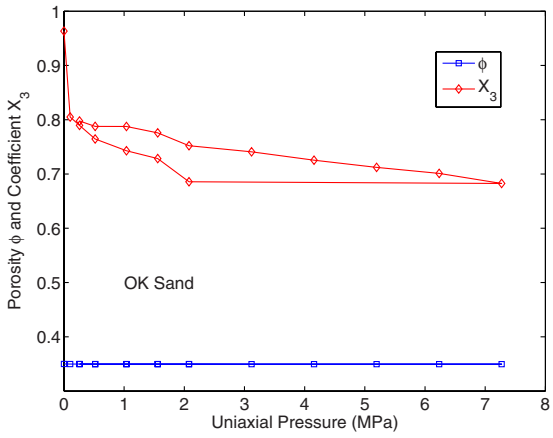


FIG. 12. (Color online) Plots of measured porosity ϕ and computed effective uniaxial stress coefficient X_3 in OK sand samples.

reasons to believe[32] that sometimes an even simpler theory yet, i.e., one treating the frame constants as if they are identically zero, can give a good accounting in some cases of acoustic wave speeds and attenuation in many isotropic granular systems. For present purposes, we view our present work as treating Gassmann's approach as being valid at each fixed frequency—thus making this a viscoelastic theory, rather than just an elastic theory. In the absence of observable slow-waves, this approach seems to give a satisfactory approximation to these data. It leads, as we have shown here in detail, to results that allow us to determine effective frame constants of the drained system, valid in the small range of ultrasonic frequencies pertinent to these specific experiments.

We do not assume that these measured values will necessarily hold true for other measurements in different frequency bands. In literature, it is also commonly assumed that there is a small frequency dependent dissipative component

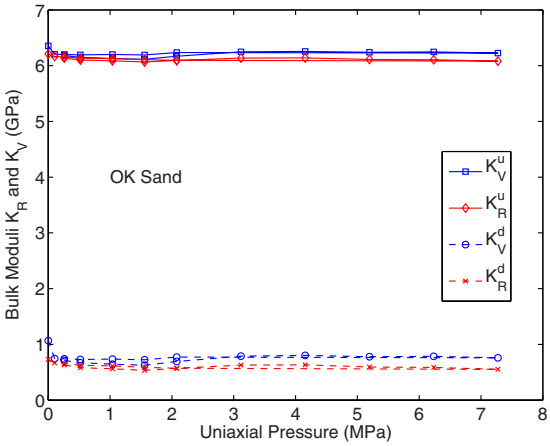


FIG. 11. (Color online) For OK sand samples, examples of the undrained and drained bulk modulus estimators K_R and K_V, being the appropriate Reuss and Voigt averages of the compliance and stiffness matrices, respectively. The abscissa values are the uniaxial stresses applied to the anisotropic system. These sand grains had K_s=38.0 GPa, μ_s=44.0 GPa, and density $\rho_s = 2.65 \times 10^3$ kg/m^3. Bulk modulus K_f=2.2 GPa for water. Porosity varied with stress level and was monitored, with values in the range $0.3496 \leq \phi \leq 0.3502$.

associated with these frame constants (see discussion of these issues by Chotiros[33]), but for our present purposes, we have treated this dissipative component as negligible, and, therefore, set it to zero.

V. CONCLUSIONS AND EXTENSIONS

One important conclusion about the general analysis presented here is that the drained constants are uniquely related to the undrained constants, assuming that porosity, grain bulk modulus, and pore fluid bulk modulus are already known when the porous material is microhomogeneous.[6,21] This fact follows directly from the linearity of the equations being solved. Of course, measurement error in undrained constants and the other constants (porosity, and bulk moduli for grains and pore fluid) is never zero, and so must be taken into account. However, since the equations are linear, the errors also propagate linearly into the computed drained constants. The linearity of this problem also leads to a qualified uniqueness result in the sense that, except for issues already mentioned concerning propagation of measurement errors, the results of the undrained-to-drained inversion process are unique—in principle. This fact has implications for earlier work that has sometimes been done using fitting or iterative methods, the point being that such methods always have an implicit caveat associated with them that the results found might not be unique (and in addition they too are obviously subject to the measurement errors) if the system of poroelastic equations happens to have multiple solutions. All that could be said previously of the iterative or fitting methods is that one solution had been found. The present results improve this situation for all such methods, since we have shown that the solution is actually unique (in principle again, except for the unavoidable measurement errors), and so earlier results of the iterative (or fitting) type may now be assumed to have found good approximations to the (in principle) unique solution of each inversion problem.

One simplification obtained in the chosen range of our analysis is that, since we are specifically considering data on granular materials, we have been able to limit our discussion to ones that happen to be homogeneous mixtures of grains all having the same bulk and shear moduli. This simplification

puts our work in the same class as that of Gassmann,[21] Walton,[34] Johnson *et al.*,[35] and many others, including a large number of the laboratory experimental results, including Domenico,[36] Murphy,[37] and Zimmer *et al.*,[38] on granular bead-packs and some sands that are especially homogeneous. We have not tried to deal specifically with issues related to microscopic heterogeneity (other than the pore space itself), but this complication is actually relatively easy to include as needed for such mixed systems[6,39]—the point being that the pertinent generalizations required here involve only Reuss averages of bulk moduli, as we have shown for the simpler cases presented.

Some very general discussions of anisotropic porous media have been given by Thompson and Willis[40] and Cowin.[41] Both these contributions emphasize the advantages of formulating the anisotropic stress-strain relations for poroelasticity in the compliance form, as we have also chosen to do. In contrast, while expressing the elastic material constants themselves in terms of compliances (i.e., Young's modulus and Poisson's ratio), Loret *et al.*[42] nevertheless emphasized the stiffness formulation of the poroelastic problem; but, in addition, they also gave careful consideration to relationships connecting the undrained and drained constants, as we have again chosen to do. However, these authors focused effort on somewhat different aspects of the problem than the more practical ones considered in the present work. In particular, their studies consider the issue of how to measure and/or estimate the quantities such as K_s and ϕ, which might be considered microscale quantities, especially if the solid material itself is *not* microhomogeneous,[6,21] as we have assumed it is in our measured samples. These issues are indeed important for many applications, but not so important for the data analysis we presented, since the samples we studied were, in fact, composed of very well-sorted materials. Thus, an assumption of microhomogeneity for these laboratory samples is quite reasonable. Such an assumption may not, however, always be equally good for field data in Earth, in ocean sediments, or in other applications of these methods. So the present analysis needs to be—and will be—subsequently generalized to incorporate these considerations pertaining to microheterogeneous frame materials in future work.

ACKNOWLEDGMENTS

Work of both J.G.B. and S.N. was performed under the auspices of the U.S. Department of Energy, at the Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231. We thank Leon Thomsen for helpful comments on this work, and for finding some pertinent references to related and previously published (but nevertheless hard to find) work. Support was provided specifically by the Geosciences Research Program of the DOE Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences.

APPENDIX A: INVERTING A MATRIX FORMED FROM THE SUM OF AN INVERTIBLE MATRIX AND A RANK-1 MATRIX

A result used several times in the main text when we transform between compliance and stiffness matrices in poroelasticity involves the use of a fairly well-known fact about inverting simple matrix sums. The matrices of interest to us are always symmetric, so we use this condition here to simplify and shorten the discussion.

Suppose that we need to invert a matrix of the form

$$\mathbf{M} = \mathbf{M}_0 + a\mathbf{n}\mathbf{n}^T, \tag{A1}$$

where $\mathbf{M}_0$ is a square, invertible, $m \times m$ matrix, and $\mathbf{n}$ is a vector also of length m. The remaining variable a is a nonzero scalar, and for simplicity, we assume a is positive (although this condition is not always required). $\mathbf{M}_0$ is an invertible matrix by assumption, and $a\mathbf{n}\mathbf{n}^T$ is a rank-1 matrix (having only one nonzero eigenvalue), regardless of the size of m, as long as $m \geq 2$ and $a \neq 0$.

It turns out that the form of the matrix inverse of $\mathbf{M}$ must be

$$\mathbf{M}^{-1} = \mathbf{M}_0^{-1} + b\mathbf{M}_0^{-1}\mathbf{n}\mathbf{n}^T\mathbf{M}_0^{-1}, \tag{A2}$$

where b is a scalar to be determined. If form (A2) is correct, then multiplying Eq. (A1) on the right (by symmetry, multiplying on the left produces the same final result) by Eq. (A2) gives

$$\mathbf{I} = \mathbf{I} + [a + b + ab(\mathbf{n}^T\mathbf{M}_0^{-1}\mathbf{n})]\mathbf{n}\mathbf{n}^T\mathbf{M}_0^{-1}, \tag{A3}$$

where $\mathbf{I}$ is an $m \times m$ identity matrix. Equation (A3) shows—since the bracketed expression should vanish—that the previously unknown scalar b must satisfy

$$b = -a/[1 + a(\mathbf{n}^T\mathbf{M}_0^{-1}\mathbf{n})]. \tag{A4}$$

Since, by assumption, $\mathbf{M}_0$ is invertible, and both a and $\mathbf{n}$ are known, this formula gives an explicit result for the scalar b, and therefore justifies the form of the inverse matrix assumed in Eq. (A2).

Also note that this result was used explicitly by Gassmann[21] in the final section (on anisotropic media) of his well-known paper on poroelasticity.

APPENDIX B: UNIAXIAL STRESS CONDITIONS AND THE UNIAXIAL EFFECTIVE STRESS COEFFICIENT X_3

For uniaxial stress conditions, we have $\sigma_{33} = -p_c$, $e_{11} = e_{22} = 0$, and in addition for the undrained case, we have $\zeta = 0$. The transverse stresses $\sigma_{11} = \sigma_{22}$ are uncontrolled, but nevertheless determined by the boundary conditions. The fluid pressure is similarly determined by the system of equations, since the fluid is assumed to be undrained.

The pertinent equations for a transversely isotropic system are

$$(s_{11} + s_{12})\sigma_{11} - s_{13}p_c + \beta_1 p_f = e_{11} = 0, \tag{B1}$$

$$2s_{13}\sigma_{11} - s_{33}p_c + \beta_3 p_f = e_{33}, \tag{B2}$$

and

$$-2\beta_1\sigma_{11}+\beta_3 p_c-\gamma p_f=-\zeta=0, \tag{B3}$$

with $e_{22}=e_{11}$ by symmetry, and similar statements about other symmetric variables and coefficients.

Equations (B1) and (B3) can be turned into a 2×2 system, and solved for σ_{11} and p_f as functions of p_c. The result of this process is

$$\begin{pmatrix} -\sigma_{11} \\ p_f \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} p_c, \tag{B4}$$

where

$$X_1=\frac{-s_{13}\gamma+\beta_1\beta_3}{(s_{11}+s_{12})\gamma-2\beta_1^2} \quad \text{and}$$

$$X_2=\frac{-2s_{13}\beta_1+(s_{11}+s_{12})\beta_3}{(s_{11}+s_{12})\gamma-2\beta_1^2}. \tag{B5}$$

For this uniaxial system, X_2 clearly plays a role similar to that of Skempton's coefficient B in the isotropically confined case, giving the proportionality of pore pressure buildup when confining pressure p_c increases. Also, the transverse stress σ_{11} buildup is entirely analogous physically to $p_f=X_2 p_c$, but the proportionality constant is given instead by $-X_1$ (note that we have chosen pressure p to be positive in compression, while stress σ is positive in tension—which is the reason for some of these minus signs).

We can now define the correct effective stress for the uniaxial problem in the following way:

$$\frac{e_{33}}{-s_{33}} \equiv p_{\text{eff}}=\left[1+\frac{2s_{13}}{s_{33}}X_1-\frac{\beta_3}{s_{33}}X_2\right]p_c=[1-X_3]p_c, \tag{B6}$$

where the pertinent effective stress coefficient is therefore given by

$$X_3 \equiv \frac{2s_{13}^2\gamma-4s_{13}\beta_1\beta_3+(s_{11}+s_{12})\beta_3^2}{s_{33}[(s_{11}+s_{12})\gamma-2\beta_1^2]}. \tag{B7}$$

All the required coefficients in Eq. (B7) except γ are drained constants (containing no factors of K_f), and so X_3 can be computed for the undrained experiments simply by using the earlier analysis in the main text of the present paper.

[1] S. R. Pride and J. G. Berryman, "Connecting theory to experiment in poroelasticity," J. Mech. Phys. Solids **46**, 719–747 (1998).
[2] M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range," J. Acoust. Soc. Am. **28**, 168–178 (1956).
[3] M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range," J. Acoust. Soc. Am. **28**, 179–191 (1956).
[4] M. A. Biot and D. G. Willis, "The elastic coefficients of the theory of consolidation," ASME J. Appl. Mech. **24**, 594–601 (1957).
[5] R. D. Stoll, "Acoustic waves in saturated sediments," *Physics of Sound in Marine Sediments*, edited by L. Hampton (Plenum, New York, 1974), pp. 19–39.
[6] R. J. S. Brown and J. Korringa, "On the dependence of the elastic properties of a porous rock on the compressiblity of the pore fluid," Geophysics **40**, 608–616 (1975).
[7] J. G. Berryman, "Confirmation of Biot's theory," Appl. Phys. Lett. **37**, 382–384 (1980).
[8] G. Mavko and D. Jizba, "Estimating grain-scale fluid effects on velocity dispersion in rocks," Geophysics **56**, 1940–1949 (1991).
[9] D. J. Hart and H. F. Wang, "Laboratory measurements of a complete set of poroelastic moduli for Berea sandstone and Indiana limestone," J. Geophys. Res. **100**, 17741–17751 (1995).
[10] J. G. Berryman, "Transversely isotropic elasticity and poroelasticity arising from thin isotropic layers," *Theoretical and Computational Acoustics 1997*, edited by Y.-C. Teng, E.-C. Shang, Y.-H. Pao, M. H. Schultz, and A. D. Pierce (World Scientific, Newark, NJ, 1997), pp. 457–474.
[11] J. G. Berryman, "Origin of Gassmann's equations," Geophysics **64**, 1627–1629 (1999).
[12] B. T. Hefner and K. L. Williams, "Sound speed and attenuation measurements in unconsolidated glass-bead sediments saturated with viscous pore fluids," J. Acoust. Soc. Am. **120**, 2538–2549 (2006).
[13] J. G. Berryman, "Seismic waves in rocks with fluids and fractures," Geophys. J. Int. **40**, 608–616 (2007).
[14] X. Zhu and G. A. McMechan, "Direct estimation of the bulk modulus of the frame in a fluid-saturated elastic medium by Biot theory," Proceedings of the 60th International Meeting, Society of Exploration Geophysicists, Expanded Abstracts (1990), pp. 787–790.
[15] X. Zhu and G. A. McMechan, "Numerical simulation of seismic responses of poroelastic reservoirs using Biot theory," Geophysics **56**, 328–339 (1991).
[16] K. L. Williams, D. R. Jackson, E. I. Thorsos, D. Tang, and S. G. Schock, "Comparison of sound speed and attenuation measured in a sandy sediment to predictions based on the Biot theory of porous media," IEEE J. Ocean. Eng. **27**, 413–428 (2002).
[17] I. Ishibashi and O. F. Capar, "Anisotropy and its relation to liquefaction resistance of granular material," Soils Found. **43**, 149–159 (2003).
[18] C.-Y Wang, D. S. Dreger, C.-H. Wang, D. Mayeri, and J. G. Berryman, "Field relations among coseismic ground motion, water level change and liquefaction for the 1999 Chi-Chi earthquake," Geophys. Res. Lett. **30**, 1980-1–1980-4 (2003).
[19] S. C. Cowin, "The significance of bone microstructure in mechanotransduction," J. Biomech. **40**, S105–S109 (2007).
[20] Y. J. Yoon and S. C. Cowin, "An estimate of anisotropic poroelastic constants of an osteon," Biomech. Model. Mechanobiol. **7**, 1–11 (2008).
[21] F. Gassmann, "Über die elastizität poröser medien (On elasticity of porous media)," Vierteljahrsschr. Natforsch. Ges. Zur. **96**, 1–23 (1951).
[22] K. Lee, E. Park, and W. Seong, "High frequency measurements of sound speed and attenuation in water-saturated glass-beads of varying size," J. Acoust. Soc. Am. **126**, EL28–EL33 (2009).
[23] A. W. Wood, *A Textbook of Sound* (Bell, London, 1955), p. 360.
[24] A. Reuss, "Berechung der fliessgrenze von mischkristallen (Calculation of the flow limits of mixed crystals from plasticity conditions of monocrystals)," Z. Angew. Math. Mech. **9**, 49–58 (1929).
[25] R. Hill, "The elastic behaviour of a crystalline aggregate," Proc. Phys. Soc., London, Sect. A **65**, 349–354 (1952).
[26] W. Voigt, *Lehrbuch der Kristallphysik* (Teubner, Leipzig, 1928), pp. 962.
[27] A. W. Skempton, "The pore-pressure coefficients *A* and *B*," Geotechnique **4**, 143–147 (1954).
[28] D. Sherlock, K. Nihei, S. Nakagawa, and L. Duranti, "Stress-induced velocity anisotropy measurements in unconsolidated sand using a phased-array uniaxial compaction cell," Proceedings of the 77th International Meeting, Society of Exploration Geophysicists, Expanded Abstracts (2007), Vol. **26**, pp. 1550–1554.
[29] D. E. Miller and C. Spencer, "An exact inversion for anisotropic moduli from phase slowness data," J. Geophys. Res. **99**, 21651–21657 (1994).
[30] T. J. Plona, "Observation of a second bulk compressional wave in a porous medium at ultrasonic frequencies," Appl. Phys. Lett. **36**, 259–261 (1980).
[31] R. C. Y. Chin, J. G. Berryman, and G. W. Hedstrom, "Generalized ray expansion for pulse propagation and attenuation in fluid-saturated porous media," Wave Motion **7**, 43–65 (1985).
[32] K. L. Williams, "An effective density fluid model for acoustic propagation in sediments derived from Biot theory," J. Acoust. Soc. Am. **110**, 2276–2281 (2001).
[33] N. P. Chotiros, "Biot model of sound propagation in water-saturated sand," J. Acoust. Soc. Am. **97**, 199–214 (1995).
[34] K. Walton, "The effective elastic moduli of a random packing of spheres," J. Mech. Phys. Solids **35**, 213–226 (1987).
[35] D. L. Johnson, L. M. Schwartz, D. Elata, J. G. Berryman, B. Hornby, and A. N. Norris, "Linear and nonlinear elasticity of granular media: Stress induced anisotropy of a random sphere pack," ASME J. Appl. Mech. **65**, 380–386 (1998).
[36] S. N. Domenico, "Elastic properties of unconsolidated porous reservoirs," Geophysics **42**, 1339–1368 (1977).

[37]W. F. Murphy III, "Effects of microstructure and pore fluids on the acoustic properties of granular sedimentary materials," Ph.D. thesis, Stanford University, Stanford, CA, 1982.
[38]M. A. Zimmer, M. Prasad, G. Mavko, and A. Nur, "Seismic velocities of unconsolidated sands: Part 1—Pressure trends from 0.1 to 20 MPa," Geophysics **72**, E1–E13 (2007).
[39]B. Gurevich, D. Makarynska, and M. Pervukhina, "Ultrasonic moduli for fluid-saturated rocks: Mavko-Jizba relations rederived and generalized," Geophysics **74**, N25 (2009).
[40]M. Thompson and J. R. Willis, "A reformation of the equations of anisotropic poroelasticity," ASME J. Appl. Mech. **58**, 612–616 (1991).
[41]S. C. Cowin, "Anisotropic poroelasticity: Fabric tensor formulation," Mech. Mater. **36**, 665–677 (2004).
[42]B. Loret, E. Rizzi, and Z. Zerfa, "Relations between drained and undrained moduli in anisotropic poroelasticity," J. Mech. Phys. Solids **49**, 2593–2619 (2001).

Scattering of the fundamental torsional mode at an axial crack in a pipe

M. Ratassepp[a)]
Department of Mechanics, Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086, Estonia

S. Fletcher and M. J. S. Lowe
Department of Mechanical Engineering, Imperial College London, London SW7 2AZ, United Kingdom

(Received 9 May 2009; revised 25 October 2009; accepted 24 November 2009)

A quantitative study of the interaction of the $T(0,1)$ torsional mode with an axial defect in a pipe is presented. The results are obtained from finite element simulations and experiments. The influence of the crack axial extent, depth, excitation frequency, and pipe circumference on the scattering is examined. It is found that the reflection from a defect consists of a series of the wave pulses with gradually decaying amplitudes. Such behavior is caused by the shear waves diffracting from the crack and then repeatedly interacting with the crack due to circumferential propagation. Time-domain reflection coefficient analysis demonstrates that the trend of the reflection strength for different crack lengths, pipe diameters, and frequencies from a through-thickness crack satisfies a simple normalization. The results show that the reflection coefficient initially increases with the crack length at all frequencies but finally reaches an oscillating regime. Also, at a given frequency and crack length the reflection decreases with the increase in pipe circumference. An additional scattering study of the shear wave SH_0 mode at a part-thickness notch in a plate shows that the reflection coefficient, when plotted against depth of the notch, increases with both frequency and notch depth. [DOI: 10.1121/1.3277185]

PACS number(s): 43.35.Cg, 43.20.Fn, 43.20.Mv, 43.35.Zc [YHB] Pages: 730–740

I. INTRODUCTION

The research is motivated by the desire to detect defects in pipelines. Nondestructive testing (NDT) by ultrasonic guided waves has become a well established long range inspection procedure for pipes.[1,2] In this application tens of meters of the structure can be remotely screened for defects from a single test location which is far more efficient than the tedious and time-consuming conventional point-by-point methods. Moreover, with this technique, it is possible to inspect areas which are inaccessible, such as buried or coated regions. The basic idea of the technique is that the presence of the defect in the wall of the pipe will cause the scattering of the guided wave. The measurements of these scattered waves can thus enable defect detection and characterization.

The subject of this paper is to investigate the low frequency $T(0,1)$ mode interaction with an axial crack in a pipe. This axisymmetric torsional mode has already been adopted for guided wave inspection of pipes.[3] Having a simple wave structure and a non-dispersive nature, its incident pulse signals maintain their waveforms and can propagate long distances with small attenuation. Moreover, it is equally sensitive to defects at any depth in the pipe wall. This is because its displacement and stresses are dominated by the in-plane components and are practically uniform through the wall thickness. Also, at lower frequencies it is the only mode in the torsional wave family that can propagate in the pipe when applying axisymmetric torsional excitation, thus excluding multimode generation by the transducers. The $T(0,1)$ mode is particularly attractive for the current investigation because the particle displacement is perpendicular to the plane of propagation and therefore it is expected to be sensitive to cracks that are oriented along the propagation direction.[4] Axisymmetric longitudinal modes can also be used for the detection of cracks in pipes but they have the particle motion entirely in the plane of the propagation and are thus expected to be reflected little by axially aligned cracks.[5] Although some nonaxisymmetric flexural modes can also be sensitive to axial defects, having a significant circumferential stress field, their dispersive nature, and changing wave field along the propagation direction dramatically complicates their exploitation for practical NDT. In principle, it is also possible to use guided circumferential waves for the detection of axial cracks in pipes.[6] However, guided circumferential waves have the limitation that they propagate around the pipe over a small axial extent of the pipe and cannot be used for long range inspection.

The research investigating the potential of using guided waves for the detection of defects in pipes has been intense.[1–14] Quantitative studies[13] of the $T(0,1)$ mode scattering at circumferential notches and cracks have revealed that the reflection coefficient is a roughly linear function of the circumferential extent of the defect. It was shown that when the defect becomes narrower in the circumferential direction of the pipe, its detection possibility decreases. Therefore, defects aligned axially along the pipe have not received as much attention until this time, as they were known to give much smaller reflections than circumferential defects. However, Liu *et al.*[15] demonstrated in their experimental work

a)Author to whom correspondence should be addressed. Electronic mail: madis.ratassepp@ttu.ee

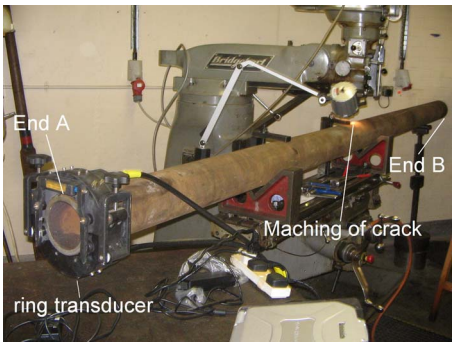


FIG. 1. (Color online) Experimental setup.

with the torsional mode that the reflection coefficient at a through-thickness axial notch was more than 15 times larger than its negligible circumferential width according to previous knowledge. Besides, the reflected signal consisted of a substantial tail whose origin was not clear. Kwun *et al.*[16] proposed that the interaction is accompanied by the circumferential shear-mode resonance which leads to the generation of the regularly spaced decaying signals in the reflection. Such a scattering mechanism can be understood from the paper presented by authors[17] who studied quantitatively the shear horizontal mode, SH_0 mode, interaction in a plate with the crack aligned in the propagation direction of the mode. They found that the scattered field at the crack consisted of contributions of the SH_0 mode, which were reflected in the incident direction and diffracted in the lateral directions, which in the pipe case will propagate circumferentially around the pipe and then interact again with the crack. It was also shown that the scattering strength is influenced by the length of the crack but in a complex way due to the presence of different types of surface waves generated on the faces of the crack.

The aim of this paper is to determine the reflection coefficients from cracks and notches of varying depth and axial length when the $T(0,1)$ mode is traveling in the pipe. The results of this study are important to understand the physical nature of the phenomena and provide some insight into the possibility of detecting and characterizing this kind of defect using ultrasonic guided waves. Finite element (FE) models and experiments on a 5 in. diameter pipe were used for this investigation. The report focuses on the analysis of the reflection as functions of length, depth, thickness and diameter of the pipe, and frequency of the incident signal.

II. EXPERIMENTAL WORK

Experiments were performed on nominal 5 in., schedule 120 steel pipes with an outer diameter of 141.3 mm and wall thickness of 12.5 mm. The pipes were 3.7 m long. Figure 1 shows the experimental set-up. The pipes were supported horizontally on "v" blocks which had a negligible effect on the wave propagation behavior in the pipe. A milling machine was aligned so that an axial slot could be cut along the axis of the pipe, and the axial position of the cut was changed by simply moving the pipe along its axis. This enabled the cutting of the axial notches in an accurate fashion without having to disturb the guided wave instrumentation. The cut was started 1.8 m from the end of the pipe where the transducers were located and extended away from it. Experiments were conducted on two separate pipes, using the following notches: (a) a through thickness notch and (b) 80% deep notch. The notches were 3 mm wide and their length was increased incrementally in steps of 8 mm from 8 up to 120 mm. Due to the shape of the cutting tool, the ends of the notch were rounded and the part-depth notch was flat-bottomed.

The torsional $T(0,1)$ mode was generated by clamping a transducer ring at one end (end A) of the pipe, as shown in Fig. 1. The transducer ring was produced by Guided Ultrasonics Ltd.[18] The ring is made up of two rows of piezoelectric transducer elements,[19] each consisting of 16 elements spaced uniformly around the circumference. The elements impart tangential force and are oriented in such a way as to act in the circumferential direction around the pipe. This means that only the $T(0,1)$ mode is excited.[3] The contact between the transducers and the pipe was dry and no pipe surface cleaning was performed. Although in usual practice it is possible to control the directionality of the generated waves, by controlling the phase difference between the two rows, this was not necessary in this test because the ring was at the end of the pipe, so all the transducers were excited simultaneously.

The equipment used to generate and receive the signals was a Guided Ultrasonics Ltd. Wavemaker G3 instrument. The test signal was 5-cycle tone burst modulated by a Hanning window. For each crack length, the center frequency of excitation was varied from 20 to 65 kHz in increments of 5 kHz. In order to obtain a reference measurement, reflections from end B were recorded before any milling was done. As the signals reflected from the notch were in general quite weak, the quality of each measurement was enhanced by taking 256 averages and applying a band pass filter.

By summing the received signals at all of the transducer elements, it was possible to monitor just the reflected $T(0,1)$ mode[10] and ignore mode converted signals reflected at the crack. The reflection coefficient for each crack length was calculated in the time domain. This was done by calculating the envelope of the time signal, by taking a Hilbert transform, then simply dividing the value of the reflection peak by that of the end-wall reflection peak which had been recorded in the undamaged pipe. Since the signals are fairly narrowband, this approach gives a reflection coefficient which is sensibly representative of the center frequency of the signal.

III. FINITE ELEMENT MODELING

A. Membrane model for through-thickness axial crack

The interaction of the $T(0,1)$ mode with circumferential cracks and notches in pipes has been successfully studied in the past using the FE method.[13] A similar study was performed here to predict the scattering at an axial through-thickness crack using the ABAQUS software with its explicit time stepping procedure.[20] The geometry of the pipe model was discretized using two-dimensional membrane elements, assuming plane stress condition in the pipe wall. This approach has already been shown to work well to model exten-

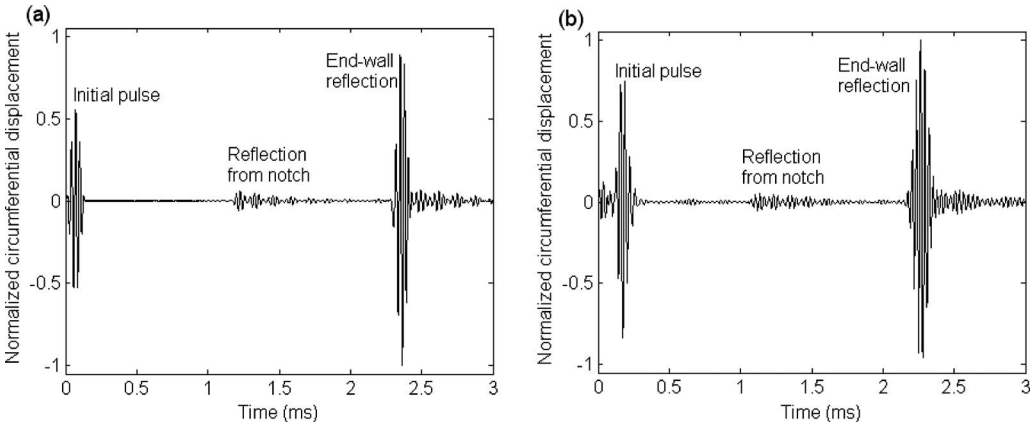


FIG. 2. (a) Predicted FE and (b) experimental time record for 5 in. pipe with 56 mm axial notch and 5-cycle $T(0,1)$ mode incident at 35 kHz.

sional and torsional behavior of the propagating modes around the frequencies at which the experiments were performed. However, it is not possible with this approach to simulate modes in which there is a local bending of the pipe wall, nor to model part-through notches.

A mesh of identically sized linear (four noded) quadrilateral membrane elements was used. The models of the 5 in. pipe represented a 3.7 m length, using 925 elements along the length and 128 elements around the circumference. As a result each element was 4 mm long and 3.16 mm wide around the circumference corresponding to about 12 elements per wavelength at the highest frequency used, respecting the spatial discretization limit required for accurate modeling. Additionally, some larger radius pipes were modeled by using more elements with the same element dimension along the circumference. Zero-width cracks with zero stress on edges were modeled by disconnecting adjacent elements. Notches with nonzero width were introduced by removing elements from the model; the notches were therefore rectangular in shape.

The axisymmetric torsional mode $T(0,1)$ was generated by prescribing identical tangential displacement time histories at all nodes at the end A of the pipe. The central frequency of the tone burst varied from 15 to 65 kHz. The detection of reflected waves was achieved simply by monitoring circumferential displacements at 16 nodes, simulating the receivers, around the circumference at the end A of the pipe. The reflected $T(0,1)$ mode was obtained by summing all the received signals and the reflection coefficient measurements were performed in the same way as in experiments. Some results were also obtained by measuring the nodal displacements around the pipe at a location between the excitation end of the pipe and the defect.

B. 3D model for part-through axial crack

The whole volume of the pipe was discretisized using eight-node "brick" elements. Each node of the element has three degrees of freedom (displacements in x,y,z directions) and such elements permit full three-dimensional (3D) vibration analysis of the pipe. Although this approach enables the modeling of real part-thickness notches, it is computationally expensive due to the large number of elements.

The 3.7 m long pipe was modeled with 925 elements along the length, 128 elements around the circumference, and 5 elements through the wall thickness. Both cracks and finite width notches were modeled 80% deep through the wall thickness. The width of the notch varied through the thickness of the pipe wall, being 3.46 mm wide on the surface of the pipe and 2.96 mm wide at the bottom of the defect. Excitation and monitoring locations were arranged similarly to the membrane model.

IV. RESULTS

A. Through-thickness notch (membrane models)

The results from the experimental data and FE simulations are presented. Typical time records from the experiment and simulation are shown in Figs. 2(a) and 2(b). Both of these records are for a 5 in. pipe with a notch of 56 mm in axial extent, and for a 5-cycle $T(0,1)$ mode incident at 35 kHz. The results are normalized to the maximum absolute displacement value of the end-wall reflection. The figures show clearly the initial tone burst on its way toward the notch and then the reflected $T(0,1)$ mode at the defect and at the pipe end. Interestingly, the reflection consists of a series of pulses with gradually decaying amplitudes. In this example, these pulses are spaced with a period which is approximately equivalent to the duration of the wave pulse passing the circumference of the pipe. Such behavior is caused by the diffraction at the notch, as was shown for the fundamental shear-mode interaction with the crack in a plate.[17] The resulting scattering in that work consisted of the direct reflection and diffraction which was composed of ultrasonic shear waves propagating in both directions perpendicularly to the crack alignment. Similarly, such waves are generated in the present case which propagate circumferentially around the pipe and subsequently interact repeatedly with the crack, giving characteristic echoes after the first reflected pulse. This phenomenon affects the time-domain amplitude of the final reflection if the pipe is sufficiently small diameter and the tone burst is long that overlapping of signals can occur. It can also be seen from the results that there is a small time shift between the two records and that the generated and reflected wave packets in the experimental measurements are slightly wider than those in the FE data.

The reason for this comes from the design of the transducer ring. It has two rows of transducers which are set apart by a specific distance, so one of the rows is slightly away from the end of the pipe. When this ring is excited then there is an additional delayed reflection at the near end of the pipe which distorts the shape of the incident signal. Also, we can see that the back-wall reflection is significantly larger in amplitude than the incident signal. This is due to the doubling of the displacement amplitudes as the measured displacements incorporate both incoming and reflected wave fields at the end A of the pipe.[21]

The variation of the $T(0,1)$ reflection coefficient with axial length l of the notch normalized to incident wavelength λ is shown in Fig. 3(a) at 30, 35, and 40 kHz. In general, the agreement between the FE and experimental results is good. The differences between the results can be accounted by the presence of noise. Very small signals are being examined and the small deviations in experimental set-up can cause a small noise level. The reflection coefficient has similar features to that of the plate case.[17] The predicted FE curve has small characteristic peaks and troughs which are due to interference of reflection components at the tip and at the far end of the notch which are not separable in the first dominant reflection of $T(0,1)$.

Figure 3(b) shows the predicted $T(0,1)$ reflection coefficient versus axial length of the crack normalized to the incident wavelength for a wider range of incident frequencies. If we consider a specific normalized crack length value, we observe that the reflection ratio decreases when the frequency increases. This is not consistent with the previous plate results from where it is known that the reflection coefficient does not depend on the operating frequency when the crack length is similarly scaled to the wavelength of the incident wave for those frequencies. In the plate study the waves were excited and monitored at a single point along the crack alignment and the signals were beamcompensated. However, in the pipe case the reflection from a crack spreads around the pipe and the reflection coefficient is calculated by summing all the amplitudes of received waves measured at the pipe circumference. Recently, an approximate relationship between pipe mode and plate solutions was derived,[22] and it was shown that pipe mode amplitudes are related to plate amplitudes by the law of inverse square root of frequency. The reduction in frequencies can be seen also in the figure.

Another difference between the curves in Fig. 3(b) is the remarkable increase in the reflection coefficient of the 15 kHz result compared to the plateau region exhibited by the higher frequency curves. This is due to a resonance of the different wave packets reflected at the crack. As the excitation signal was a 5-cycle tone burst and the circumferential extent of the pipe is slightly less than 2λ, then there is an overlapping between the prime reflection and the reflections induced later by circumferential waves interacting with the crack. Furthermore, the lateral scattering, and thus the interaction of the circumferential waves with the crack, is strongest at crack dimensions around $l/\lambda=0.7$.[17,23]

Figure 4(a) shows a normalized FE time signal for a 5-cycle $T(0,1)$ wave reflected from a 36 mm long zero-

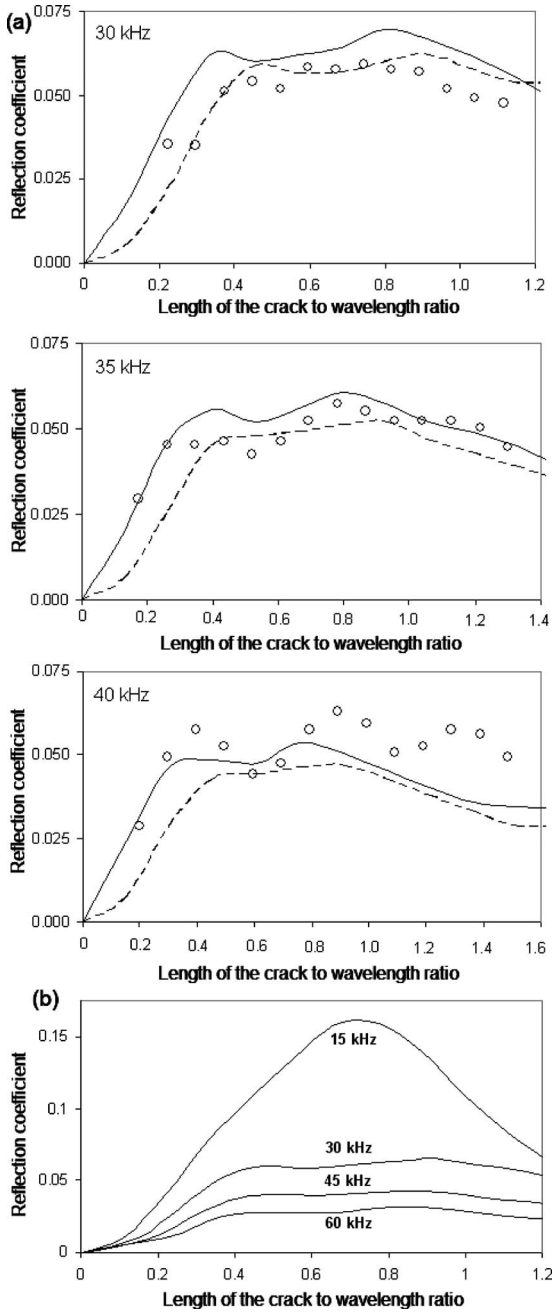


FIG. 3. Results for 5 in. pipe and $T(0,1)$ mode incident: (a) Variation of reflection coefficient with length of the notch normalized to wavelength at 30, 35, and 40 kHz; (○○○)—experiment, (solid line)—FE membrane model with a notch, (dashed line)—FE membrane model with a crack. (b) Variation of reflection coefficient with length of the crack normalized to wavelength at various frequencies.

width notch at 45 kHz. The signals were monitored at 1.4 m from the pipe end A. Its magnitude of reflection coefficient in Fig. 4(b) was calculated by dividing the magnitude of the Fourier transform of the full reflected signal by that of the incident signal. Note that there are two horizontal axes (frequency and normalized circumference) in the figure. Interestingly, it is not a smooth function of the frequency as was found for circumferential through-cracks in pipes,[13] but there are regularly spaced peaks in the spectrum. These peaks are due to resonance of the consecutive wave pulses reflected at the crack. The interpretation of this can be understood by

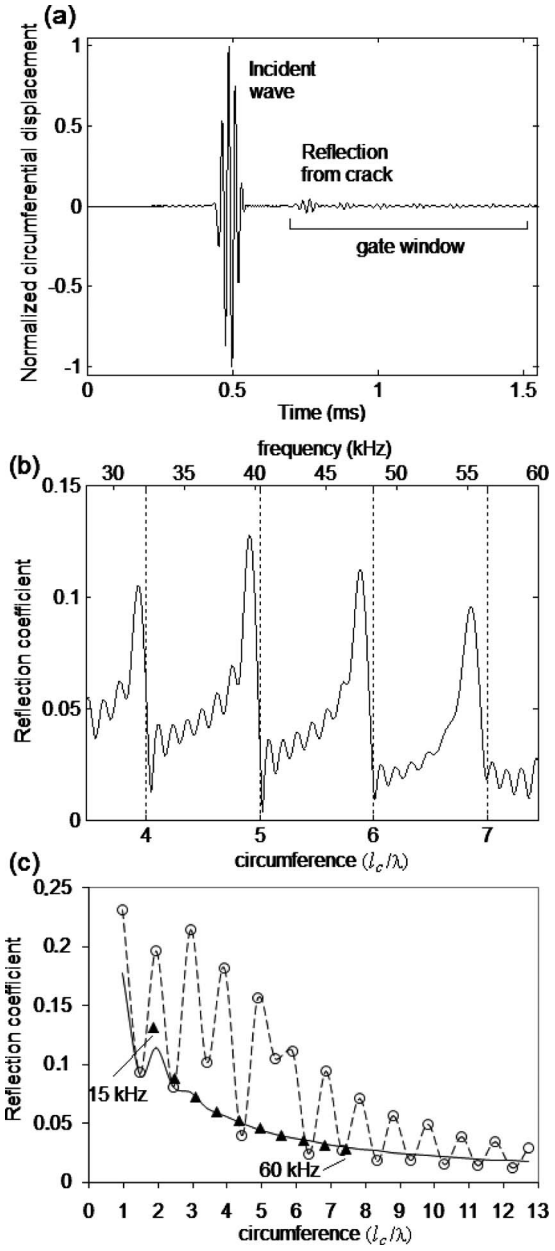


FIG. 4. Results for 5-cycle $T(0,1)$ mode incident at an axial crack 1 =36 mm ($l/\lambda=0.524$): (a) predicted FE time record and (b) reflection coefficient as the function of frequency and pipe circumference for 5 in. pipe at 45 kHz; (c) variation of reflection coefficient with circumference: (solid line)—time-domain values at 45 kHz, (dashed line)—frequency-domain values at 45 kHz, and (solid triangle)—values at 15–60 kHz for 5 in. pipe.

observing the reflection coefficient as a function of the circumference l_c of the pipe measured in wavelengths of the tone. It can be seen that the maxima of the reflection coefficient spectrum occur almost at the frequencies corresponding to integer wavelength values around the pipe circumference. Consequently, constructive interference occurs at those wavelengths when the reflections are an integer number of cycles behind the leading reflection. However, the results indicate that these maxima do not occur exactly at an integer number of wavelength values but at slightly smaller values. The reason for this is not known, but it is believed that this is probably due to a phase delay when the circumferential waves are reflected at the crack, giving the appearance of lengthening the propagation path. The revealed peak-like nature of the signal can be a clear indicator of the presence of an axial crack.

Figure 4(c) shows the prediction for the reflection coefficient versus the circumference of the pipe for a specific case of a 36 mm long crack and a 5-cycle $T(0,1)$ tone burst with center frequency 45 kHz. Both time-domain (solid line) and frequency-domain (dashed line with circles) reflection coefficients are shown. The time-domain reflection coefficient decreases smoothly with the increase in the pipe circumference, except at low pipe circumference values where the curve oscillates. Most of the curve is smooth because the time domain amplitude is given only by the peak of the first arrival wave packet, without interference from later packets. The irregularity at low frequency is because there is an overlapping of the subsequent reflected wave pulses from the crack. Closer examination shows that the amplitude of the reflection is inversely proportional to the pipe circumference which is consistent with the previous knowledge.[22] The overall reduction in the reflection coefficient with increasing circumference means that it is more difficult to detect the crack of the same length in larger radius pipe at the same testing frequency. As expected, the frequency-domain results oscillate with the variation of pipe circumference. The reason for this was explained in Fig. 4(b). The frequency spectrum incorporates the summation of the amplitudes of all the waves reflected at a crack which may at certain pipe circumferences be in phase, having maxima in the frequency spectrum, or being out of phase causing reduction in the amplitude in the final spectrum.

An interesting correlation is observed in the time-domain results if we re-plot the reflection coefficients using different incident frequencies and the same normalized crack length. Figure 4(c) shows that the reflection coefficient values which are obtained for a 5 in. pipe at frequencies from 15 to 60 kHz (solid triangle) agree with the previous 45 kHz results if the circumferences are similarly scaled. This means that, in the studied frequency range for a given normalized circumference and normalized crack length, the reflection has the same amplitude whatever the frequency is. For example, if the pipe diameter is 125 mm, the reflection amplitude from a 62 mm long crack at 30 kHz will be approximately the same as that for the pipe with the diameter 75 mm and a 37 mm long crack at 50 kHz.

B. Part-thickness notch (3D model)

Typical FE time trace showing the interaction of a 35 kHz 5-cycle $T(0,1)$ mode with a 56 mm long and 80% deep notch is presented in Fig. 5(a). The reflection at the notch is hardly noticeable and its amplitude is around half the size of the reflection amplitude of the through-thickness notch with the same length. This means that the variation of the amplitude is not linearly related to the variation of the depth of the notch. This indicates some changes in the wave scattering mechanism at the defect.

The variation of the $T(0,1)$ reflection coefficient with normalized notch length from the 80% deep notch is shown in Fig. 5(b) at 35 and 40 kHz. Although the amplitudes of the

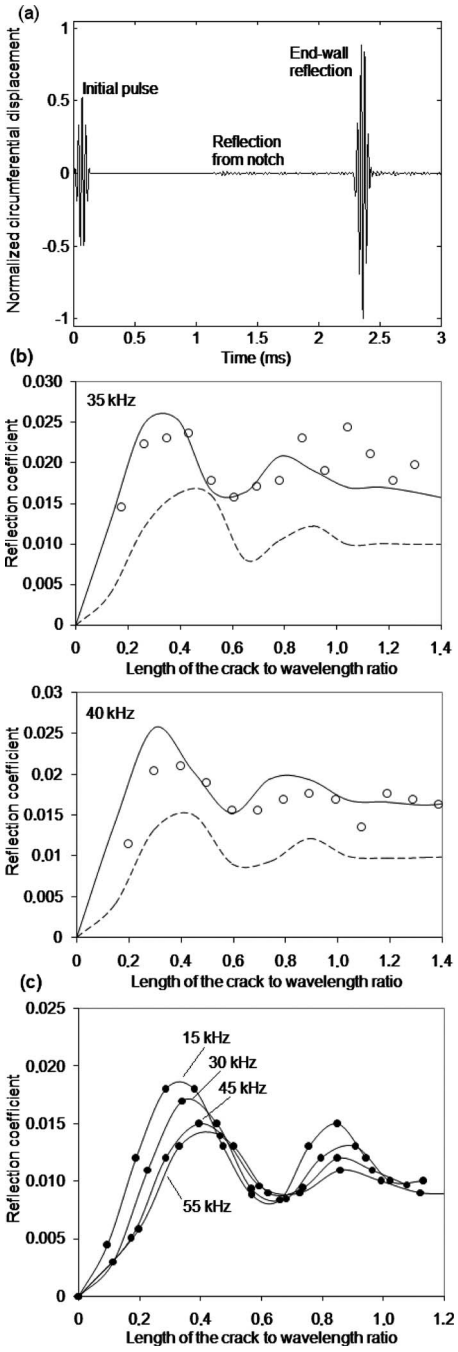


FIG. 5. Results for 5 in. pipe and $T(0,1)$ mode incident: (a) predicted FE time record from 56 mm 80% deep axial notch at 35 kHz; (b) variation of reflection coefficient with length of the 80% deep notch normalized to wavelength at 35 and 40 kHz; (○○○)—experiment, (solid line)—3D FE model with a notch, and (dashed line)—3D FE model with a crack; (c) variation of reflection coefficient with length of the 80% deep crack normalized to wavelength at various frequencies.

reflected signals were very small, the agreement between the experimental and FE notch model results is quite satisfactory. Here it is important to note that the width of the notch affects the reflection strength substantially. It is clearly seen that the reflection from the crack is much smaller than from a finite width notch at all defect lengths.

Figure 5(c) shows FE predictions of the variation of the reflection coefficient with normalized crack length for a wider range of incident frequencies. We can see that the maximum reflection strength is slightly weaker for higher frequencies. Compared with the through-crack results in Fig. 3(b), we can see that the reflection strength has dropped drastically, especially for low frequencies. To understand this, a simple study was performed analyzing the scattering of the SH_0 mode at a part-depth notch in a plate using a 3D FE model. The SH_0 mode in a flat plate is closely related to the $T(0,1)$ mode in a pipe, so understanding the simple plate case is useful and convenient for such studies. Also, the plate model helps to analyze the reflection and diffraction phenomenon separately. A short FE plate study is provided in the Appendix.

Figures 6(a) and 6(b) show the variation of the frequency-domain reflection and diffraction coefficient of the SH_0 mode with normalized crack length for a range of crack depths in a plate. It is clear that when the depth of the crack decreases the amplitudes of scattered waves decrease also. This can be explained by the weakening of mode conversions between the incident shear wave and the surface waves propagating on the crack surfaces[17] due to the reduction in the free surface area of the crack. Furthermore, this has a substantial effect on the amplitude of the scattered waves after primary reflection caused by circumnavigating shear waves interacting repeatedly with the crack in a pipe. The amplitudes of these reflections decrease more quickly compared to the through-crack case because of the reduction in the scattering energy after each interaction.

Further analysis showed that these frequency-domain plate reflection coefficient results can be approximately extended to the time-domain pipe results, confirming again previous results.[22] This is shown in Fig. 7 where the reflection coefficients at 60% and 80% deep cracks are normalized by the through-crack result for varying crack length. We can also see that at varying crack lengths the variation of reflection strength with depth is different. This means that it is not possible to obtain the unique scale factor which links the reflection amplitude from a part-thickness crack by knowing that from a through-crack of the same length because in general the effects of the length and depth of the crack on the scattering process are coupled.[24]

Figures 8(a) and 8(b) show the variation of the reflection and diffraction coefficient with depth of the crack for different frequencies in a plate with constant thickness. The normalized crack length was taken as $l/\lambda=0.7$. If we consider one frequency result, we observe that the scattering strength increases with defect depth in a nonlinear manner. Now, as the frequency decreases, the curves become increasingly concave, the reflectivity and diffraction at low frequencies decreasing markedly. Therefore, this result suggests that it is more difficult to detect shallower defects as the test frequency is decreased. Similar effects were also observed in a case of the $T(0,1)$ mode scattering at full-circumference part-thickness cracks in a pipe.[13] However, in the case of the pipe with an axial crack the strength of the higher frequency scattering also suffers from increasing difference between the phases of received signals reducing the amplitude, which has a more severe effect on the reflection amplitude resulting in smaller reflection coefficients with the increase in frequency as was shown in Fig. 3(b).

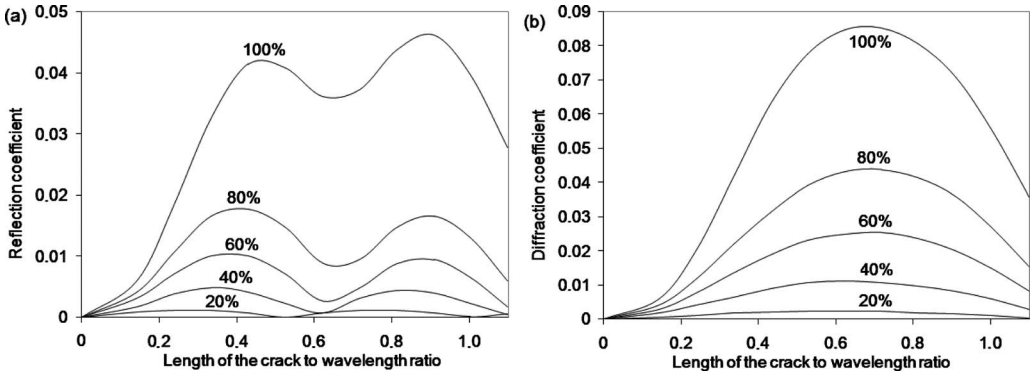


FIG. 6. Predicted variation of (a) reflection and (b) diffraction coefficient with length of the crack normalized to wavelength at 20%, 40%, 60%, 80%, and 100% deep cracks. Results are from 3 mm FE plate model using incident SH_0 mode at 250 kHz.

Figures 9(a) and 9(b) show the predicted variation of the reflection and diffraction coefficient with frequency-thickness in the case of an 80% deep crack. The thickness of the plate was varied from 1, 1.5, 2, 2.5, 3, 3.5 to 4 mm. The incident frequencies were 100, 250, and 350 kHz. Again, it can be seen that the reflection amplitude increases with the frequency-thickness. It is also of interest to observe that when the frequency-thickness product is similar, the reflection amplitude is similar. Thus, the results may be applied equally to other thicknesses of plate and pipe provided the frequency-thickness is retained.

V. DISCUSSION

The motivation for this work is to provide an insight into the possibility of detecting and characterizing axially aligned cracks in the wall of a pipe. As we saw, the interaction of the torsional mode with an axial crack was influenced by several parameters: crack length, crack depth, frequency-thickness, and diameter of the pipe. The results showed two important trends. On the one hand, the study of through-thickness cracks showed that lower frequencies allowed the acquisition of larger reflections, as was seen in Fig. 3(b). On the other hand, investigating reflections at part-depth cracks in plates, as the frequency-thickness value decreased, so did the reflection amplitude, given in Fig. 9. Therefore, it is important to discuss where to make the compromise using different frequencies for different pipes with different crack lengths. In the following, we will find the reflection coefficients for an axial through-crack in a specific defined pipe at all axial extents using FE, and will then introduce approximation formulas to generalize the results for other pipe sizes. The collection of reflection coefficients of the part-depth cracks in a plate helps us to estimate the reflection strength of similar defects in pipes.

A. Effect of the frequency and pipe circumference in case of through-thickness crack

According to Velichko and Wilcox[22] the reflection coefficient from a fixed sized defect is inversely proportional to the pipe radius. Thus, for any pipe circumference the reflection coefficient R at a given crack length and frequency can be approximated with the formula

$$R = R_0 \frac{l_{c0}}{l_c}, \tag{1}$$

where R_0 is the reference reflection coefficient for a pipe with the reference circumference l_{c0} and l_c refers to the circumference of pipe under investigation.

From Eq. (1) we can see another useful relation if the circumference is given in terms of incident wavelength which is related to the frequency. This links reflection coefficients R for different incident frequencies at a given pipe circumference and normalized crack length:

$$R = R_0 \frac{f_0}{f}, \tag{2}$$

where R_0 is the reference reflection coefficient for a pipe with the reference incident frequency f_0 and f is the required test frequency.

The advantage of the relations (1) and (2) is that they enable us to calculate reflection coefficients for all crack lengths at different frequencies and pipe circumferences if the reflection coefficients for all crack lengths from a specific pipe size are known. An example which shows sensitivity limits for practical testing for a clean pipe is given in Fig. 10 where the contour curves for 2% reflection coefficients for different pipe sizes are shown. In this case the reference reflection coefficients are defined for 5 in. schedule 80 pipe. We limit our study by lowest frequency $f = c_{ph}/2\pi r$ (r is the mean radius of the pipe), in which case the wavelength cor-

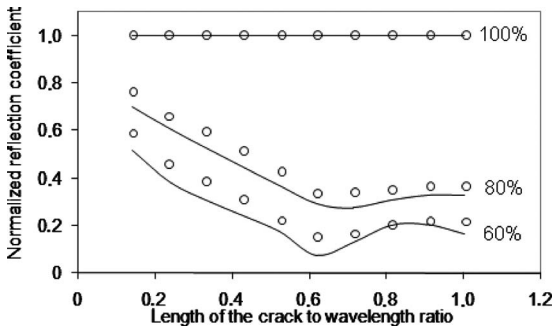


FIG. 7. Predicted variation of normalized reflection coefficient with length of the crack normalized to wavelength for 60% and 80% deep cracks. (solid line)—plate FE results and (○ ○ ○)—pipe FE results; waves were excited at 250 kHz and the thickness of the structure was 3 mm.

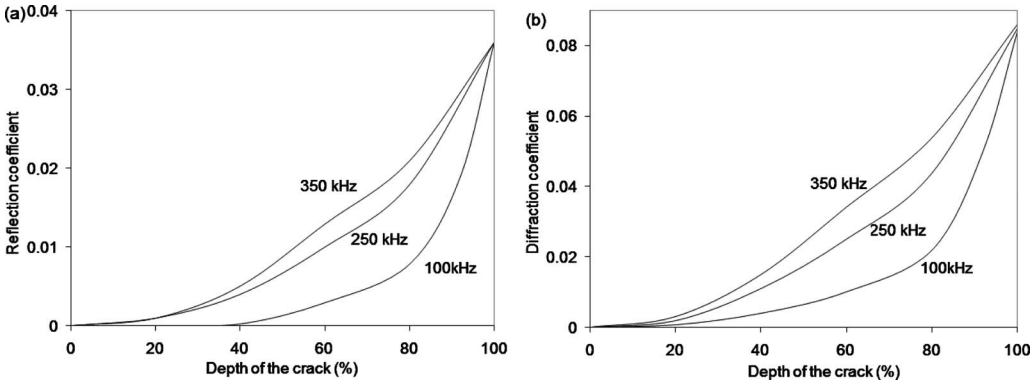


FIG. 8. Predicted variation of (a) reflection and (b) diffraction coefficient as the function of depth of the crack normalized to wavelength at $l/\lambda_{SH_0}=0.7$. The SH_0 mode was excited at 100, 250, and 350 kHz in a 3 mm plate.

responds to the pipe circumference. As expected, it is possible to detect shorter cracks in smaller diameter pipes and also as the pipe size grows the highest frequency that can be used decreases. Another observation from this figure is that the higher frequency waves are more sensitive to shorter cracks but the increase in sensitivity is rather slow. For example, in the case of 5 in. pipe, even if the frequency increases by a factor of 4, it is only possible to detect cracks that are half as long.

B. Estimating the smallest detectable crack depth

The evaluation of the maximum depth of a defect is one of the main targets in pipe inspection. It was demonstrated that the reflection coefficient from an axial crack of a given length increased monotonically with depth at all frequencies but the variation was not monotonic over the observable range of crack length, as was shown in Fig. 5(c). Therefore, the appropriate selection for maximum sensitivity would depend on the crack length. In the studied cases, the highest sensitivity could be established when the crack length to incident wavelength ratio was between 0.2 and 0.5 but in general the exact value depends also on frequency and pipe dimensions.

From the findings in Sec. IV B, we propose that a simplified analysis for the crack depth estimation may be achieved by using combined results from the plate and pipe models. An interesting outcome was that the scattering strength at a part-depth crack in a plate can be scaled to frequency-thickness and the reflection amplitudes at a crack can be approximated with the results from the FE plate model. Therefore, by obtaining the normalized reflection amplitudes for different crack depths and lengths at a specific frequency-thickness value from a plate model, it is possible by using additionally Eqs. (1) and (2) to derive the reflection amplitudes for different pipe sizes having the same frequency-thickness. It is also possible to derive the result by using the theory in Ref. 22.

Such a study was performed by modeling 3 mm thick plates with 60%, 80%, and 100% deep cracks. The SH_0 mode was excited at 60, 100, and 200 kHz and the normalized reflection coefficients at different crack lengths were obtained in the frequency domain. The additional frequency-thickness data points were obtained by making use of the full width of the spectrum of the signal, so that the frequency-thickness range 115–800 kHz mm covered most studied pipe cases. At each frequency-thickness product value, the reflection coefficients for different crack depths were obtained by approximate third degree polynomial fit through the coefficient values at crack depths 60%, 80%, and 100%. Thereafter the reflection maps showing the reflection coefficient with varying crack depth and length for each specific pipe case at a fixed frequency were created.

Figure 11(a) shows a 3D plot of the reflection coefficient for $T(0,1)$ at 35 kHz incident on an axial crack of varying depth and axial length. The results are obtained for cracks with normalized lengths up to $l/\lambda=0.6$ where the highest

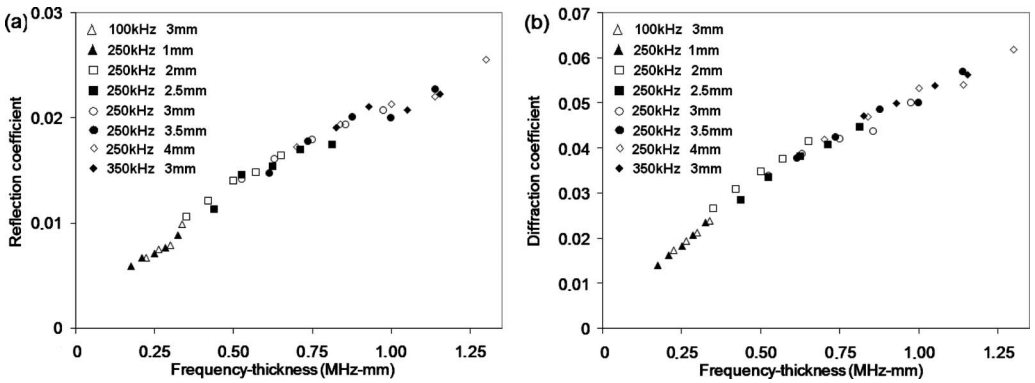


FIG. 9. Predicted (a) reflection and (b) diffraction coefficient versus frequency-thickness for 80% deep crack in a plate at $l/\lambda_{SH_0}=0.7$.

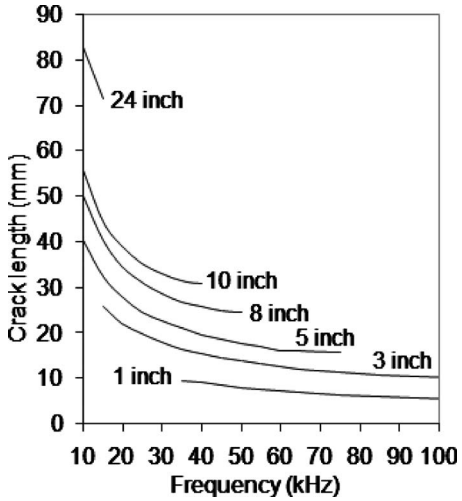


FIG. 10. Variation of the crack length with frequency for reflection amplitude 2% from incident wave amplitude for schedule 80 pipes; results are derived from FE 5 in. pipe membrane model.

sensitivity to crack depth is expected to occur. In Fig. 11(b) the contour curves of this 3D plot show that the satisfactory sensitivity to axial cracks can be achieved only for very deep cracks. Furthermore it can be seen that the sensitivity is similar for various crack depth and length combinations, which means that it is not possible to extract both dimensions from the measurement of a single reflection coefficient; this makes the evaluation of the crack parameters impossible. However, it is possible to suggest the crack length for which the sensitivity is the highest at a given frequency. For example, in this studied case the shallowest crack with the reflection coefficient 2% can be detected when the crack is 40 mm long. Similar analysis was also performed for other frequencies and pipe sizes and the results are shown in Fig. 11(c). There is a clear trend for all pipe sizes showing that the detectable crack depth increases and crack length decreases with the increase in frequency. Therefore, in practice it is useful to test the pipe at several frequencies as the crack length is not known.

VI. CONCLUSIONS

The work in this paper studied the scattering of the fundamental torsional mode by an axially oriented defect in a pipe wall. The FE simulations, confirmed by experimental measurements for selected cases, have provided a useful collection of reflection coefficient results at notches of various lengths and depths. Furthermore, the examination of the time signal has identified an interesting phenomenon of the series of wave pulses with gradually decaying amplitude reflecting from a defect, which leads to a periodic peak-like structure in the frequency spectrum. The phenomenon has been found to be due to the diffraction at a notch which then, propagating in the circumferential direction, interacts repeatedly with the notch, giving additional echoes. Thus the axial notch could be identified by the occurrence of the peaks in the frequency spectrum. It has been shown that the time-domain reflection coefficient increases with the crack length at all frequencies but finally reaches an oscillating regime. The reflection behavior for a through-thickness crack can be normalized to account for the crack length, the pipe circumference, and excitation frequency. Thereby at a given frequency and crack length, the reflection decreases with the increase in pipe circumference. The examination of the reflection coefficient as a function of crack depth demonstrated that the reflection strength decreases substantially with the decrease in the crack depth. The three-dimensional plate study showed that the sensitivity for a given crack length drops with the decreasing frequency-thickness product. However, in the pipe case the higher frequency reflection strength suffers from

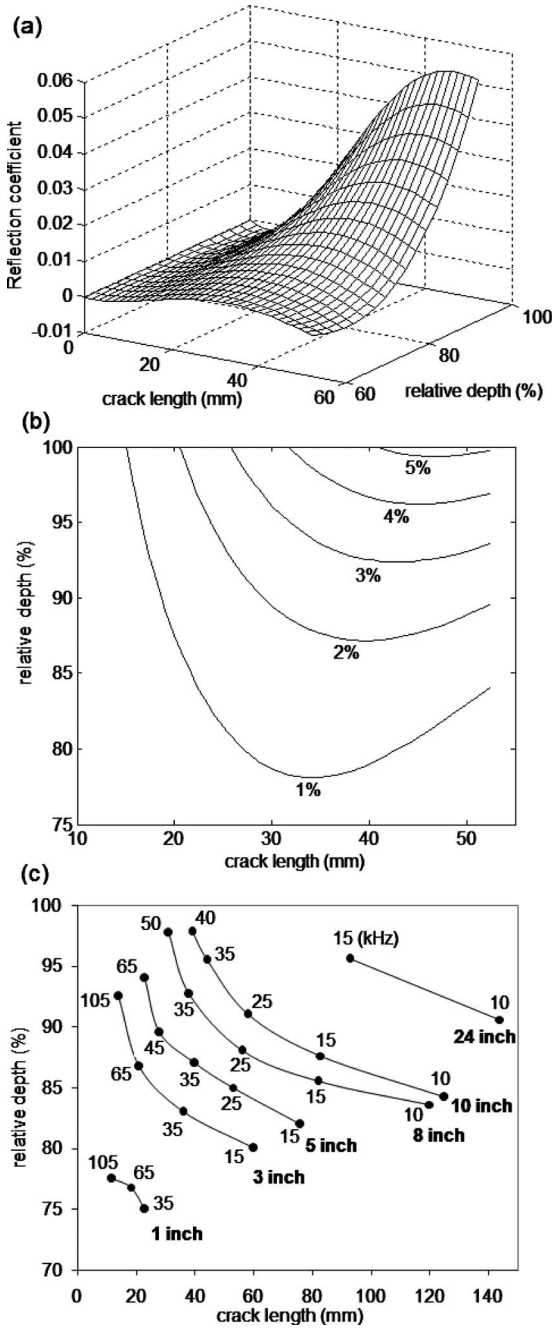


FIG. 11. Results for schedule 80 pipes and $T(0,1)$ mode incident: (a) 3D graph of reflection coefficient from axial cracks with varying depth and length and (b) contour curves of reflection coefficients 1%, 2%, 3%, 4%, and 5% at 35 kHz in a 5 in. pipe; (c) variation of the crack length with depth and frequency for reflection amplitude 2% from incident wave amplitude for various pipe sizes; results are derived from FE 5 in. pipe membrane model.

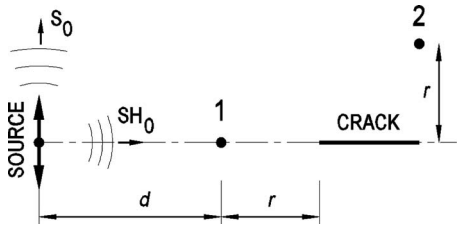


FIG. 12. Geometry for the study of SH_0 mode reflection and diffraction in a plate.

beamspreading of reflected waves to larger area of the pipe. The results demonstrated that the torsional wave is capable of detecting defects which are deeper than 75% of the wall thickness if the noise level is very low. The sensitivity to defects is dependent on frequency, whereby higher frequencies are more sensitive to deeper cracks, though the crack length also influences the detectability. Therefore, in practice it is wise to test at more than one frequency to avoid missing defects due to the unknown length of the crack. The results are shown for various pipe sizes. Further investigations will be carried out with the aim of improving detection sensitivity by using focusing techniques.

ACKNOWLEDGMENTS

M.R. is grateful to the Estonian Ministry of Education and Science Grant No. SF0140072s08 and Estonian Science Foundation Grant No. 7900 for supporting this work. S.F. acknowledges guidance by Dr. C. R. Brett and Mr. J. Trelawny from E-On Engineering.

APPENDIX: FINITE ELEMENT STUDY OF THE SH_0 MODE INTERACTION WITH PART THICKNESS CRACK IN A PLATE

This study largely uses the FE procedures from Ref. 17. Modeling was performed in a three-dimensional domain, using eight noded linear cubic elements in the plate mesh. Five elements were used through the 3 mm thick plate. Cracks of vanishing width were created by disconnecting nodes on elements representing the adjacent faces.

The excitation of SH_0 was achieved by applying point forces uniformly through the plate thickness and parallel to the plate surface in perpendicular direction to the crack axis. This generates circular waves of the SH_0 and S_0 modes, with their principal directions parallel and perpendicular, respectively, to the applied force. The point source excitation consisted of a 5-cycle Hanning windowed tone burst centered at the required frequency. The propagation of the waves was simulated in the time domain.

The surface displacement of the reflected and incident SH_0 signal was monitored in point 1 and diffracted SH_0 signal in point 2, as illustrated in Fig. 12. The far field reflection and diffraction behavior of the SH_0 mode was studied in terms of the frequency-domain ratio of the beam compensated signals, given as follows:

$$\text{Coefficient} = \frac{A\sqrt{r}}{I\sqrt{d}}, \tag{A1}$$

where A is the amplitude of reflected or diffracted signals, I is the amplitude of the incident signal, r is the distance between the tip of the crack and monitoring point, and d is the distance from the source to the monitoring point 1.

[1]D. N. Alleyne and P. Cawley, "Long range propagation of Lamb waves in chemical plant pipework," Mater. Eval. **55**, 504–508 (1997).
[2]P. Cawley, M. J. S. Lowe, D. N. Alleyne, B. Pavlakovic, and P. Wilcox, "Practical long range guided wave testing: Applications to pipe and rail," Mater. Eval. **61**, 66–74 (2003).
[3]D. N. Alleyne, B. Pavlakovic, M. J. S. Lowe, and P. Cawley, "Rapid, long range inspection of chemical plant pipework using guided waves," Insight **43**, 93–96 (2001).
[4]W. Mohr and P. Höller, "Some results on the scattering of guided elastic SH waves from material and geometric waveguide discontinuities," IEEE Trans. Sonics Ultrason. **Su-23**, 369–374 (1976).
[5]Y.-M. Cheong, D.-H. Lee, and H.-K. Jung, "Ultrasonic guided wave parameters for detection of axial cracks in feeder pipes of PHWR nuclear power plants," Ultrasonics **42**, 883–888 (2004).
[6]W. Luo, J. L. Rose, and H. Kwun, "Circumferential shear horizontal wave axial-crack sizing in pipes," Res. Nondestruct. Eval. **15**, 149–171 (2004).
[7]M. Silk and K. Bainton, "The propagation in metal tubing of ultrasonic wave modes equivalent to Lamb waves," Ultrasonics **17**, 11–19 (1979).
[8]C. M. Fortunko and R. E. Schramm, "Evaluation of pipeline girth welds using low-frequency horizontally polarized waves," J. Nondestruct. Eval. **3**, 155–173 (1982).
[9]D. N. Alleyne, M. J. S. Lowe, and P. Cawley, "The reflection of guided waves from circumferential notches in pipes," J. Appl. Mech. **65**, 635–641 (1998).
[10]M. J. S. Lowe, D. N. Alleyne, and P. Cawley, "The mode conversion of a guided wave by a part-circumferential notch in a pipe," J. Appl. Mech. **65**, 649–656 (1998).
[11]J. Li and J. L. Rose, "Excitation and propagation of non-axisymmetric guided waves in a hollow cylinder," J. Acoust. Soc. Am. **109**, 457–464 (2001).
[12]P. Cawley, M. J. S. Lowe, F. Simonetti, C. Chevalier, and A. G. Roosenbrand, "The variation of the reflection coefficient of extensional guided waves in pipes from defects as a function of defect depth, axial extent, circumferential extent and frequency," J. Mech. Eng. Sci. **216**, 1131–1143 (2002).
[13]A. Demma, P. Cawley, and M. J. S. Lowe, "The reflection of the fundamental torsional mode from cracks and notches in pipes," J. Acoust. Soc. Am. **114**, 611–625 (2003).
[14]A. Demma, P. Cawley, M. J. S. Lowe, A. G. Roosenbrand, and B. Pavlakovic, "The reflection of guided waves from notches in pipes: A guide for interpreting corrosion measurements," NDT & E Int. **37**, 167–180 (2004).
[15]Z. Liu, C. He, B. Wu, X. Wang, and S. Yang, "Circumferential and longitudinal defect detection using $T(0,1)$ mode excited by thickness shear mode piezoelectric elements," Ultrasonics **44**, e1135–e1138 (2006).
[16]H. Kwun, S. Y. Kim, H. Matsumot, and S. Vinogradov, "Detection of axial cracks in tube and pipe using torsional guided waves," in *Review of Progress in Quantitative NDE*, edited by D. Thompson and D. Chimenti (Plenum, New York, 2008), Vol. **27**, pp. 193–199.
[17]M. Ratassepp, M. J. S. Lowe, P. Cawley, and A. Klauson, "Scattering of the fundamental shear horizontal mode in a plate when incident at a through-crack aligned in the propagation direction of the mode," J. Acoust. Soc. Am. **124**, 2873–2882 (2008).
[18]Guided Ultrasonics Ltd. http://www.guided-ultrasonics.com. (Last viewed 8/5/2009).
[19]D. N. Alleyne and P. Cawley, "The excitation of lamb waves in pipes using dry-coupled piezoelectric transducers," Nondestr. Test. Eval. **15**, 11–20 (1996).
[20]ABAQUS6.5, Analysis User's Manual, Abaqus, 2004.
[21]H. Bian and J. L. Rose, "A physical interpretation of elastic guided-wave reflection from normal ends of waveguide," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **51**, 838–847 (2004).
[22]A. Velichko and P. D. Wilcox, "Excitation and scattering of guided waves: Relationships between solutions for plates and pipes," J. Acoust. Soc. Am. **125**, 3623–3631 (2009).

[23]P. Rajagopal and M. J. S. Lowe, "Short range scattering of the fundamental shear horizontal guided wave mode normally incident at a through-thickness crack in an isotropic plate," J. Acoust. Soc. Am. **122**, 1527–1538 (2007).

[24]P. Rajagopal and M. J. S. Lowe, "Scattering of the fundamental shear horizontal guided wave by a part-thickness crack in an isotropic plate," J. Acoust. Soc. Am. **124**, 2895–2904 (2008).

Polycrystalline gamma-plutonium's elastic moduli versus temperature

I. Stroe
Department of Physics, Worcester Polytechnic Institute, Worcester, Massachusetts 01609

J. B. Betts, A. Trugman, C. H. Mielke, J. N. Mitchell, M. Ramos, and F. J. Freibert
Los Alamos National Laboratory, Los Alamos, New Mexico 87545

H. Ledbetter
Department of Mechanical Engineering, University of Colorado, Boulder, Colorado 80309

A. Migliori
National High Magnetic Field Laboratory, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 13 May 2009; revised 10 November 2009; accepted 11 November 2009)

Resonant ultrasound spectroscopy was used to measure the elastic properties of pure polycrystalline ^{239}Pu in the γ-phase. Shear and longitudinal elastic moduli were measured simultaneously and the bulk modulus was computed from them. A smooth, linear, and large decrease in all elastic moduli with increasing temperature was observed. The Poisson ratio was calculated and an increase from 0.242 at 519 K to 0.252 at 571 K was found. These measurements on extremely well-characterized pure Pu are in agreement with other reported results where overlap occurs. We calculated an approximate Debye temperature $\Theta_D = 144$ K. Determined from the temperature variation in the bulk modulus, γ-Pu shows the same Grüneisen parameter as copper.
 [DOI: 10.1121/1.3271036]

PACS number(s): 43.35.Cg, 43.20.Ye, 43.25.Gf [RR] Pages: 741–745

I. INTRODUCTION

Plutonium (Pu) is one of the most difficult and interesting elements to study. At ambient pressure, Pu exhibits six distinct solid phases below the melting point. Although extensive theoretical and experimental efforts have been directed toward understanding its properties, many open questions remain.[1] To help resolve some of these, accurate values for the bulk modulus in all the phases is of central importance in calibrating theory and models of Pu because the bulk modulus is an essential thermodynamic quantity, minimally sensitive to impurities, the simplest quantity to compute from an electronic structure model, and nearly invariant between monocrystal and polycrystal specimens. This last property is relevant because it is extremely difficult to produce and measure an untwinned monocrystal of γ-Pu.

The experimental challenges are at least twofold: (i) a pure Pu sample requires considerable effort to produce, and (ii) ^{239}Pu requires special handling techniques because of its radiotoxicity. This limits the experimental investigations to measurements that can be performed in especially designed systems that comply with extensive safety requirements.

Resonant ultrasound spectroscopy (RUS)[2–5] has been used extensively to determine the thermodynamic and elastic properties of materials and to understand phase transitions. The main advantages of RUS are as follows: (i) all the elastic moduli are determined in a single measurement scan, (ii) RUS does not require bonding of the sample to transducers, (iii) it provides exceptional absolute accuracy, and (iv) it exhibits high precision (typically at least 1 in 10^6).

RUS measures directly the adiabatic elastic moduli. They are the second derivative of the energy (not free energy) with respect to strain, are directly related to the atomic bonding of the material, and are very sensitive to phase transitions and their order. Moduli connect to the thermal properties of solids through the Debye model of specific heat. In combination with specific heat and thermal expansion measurements, elastic moduli are used to determine the equation of state and various thermodynamic functions.

We have determined the elastic moduli of pure polycrystaline Pu in the alpha (α), beta (β), gamma (γ), and alloyed delta (δ) phases. Here, we report only the elastic moduli in the γ-phase on the same pure Pu specimen used for α- and β-phase measurements. The elastic moduli of Pu in the α-phase, the details of the experimental setup, the purity and preparation procedures, and a survey of outstanding problems for understanding of Pu are described elsewhere.[1] The elastic moduli of Pu in the β-phase will be reported separately. Because of many experimental problems, the elastic moduli of this specimen of pure Pu in the δ-phase could not be determined.

The X-ray volume of γ-phase Pu (99.85% with Ag as the deliberately introduced main impurity) from 486 to 585 K was described by Zachariasen[6] (analysis of available data by Wallace[7] concludes that the full range of existence of γ-Pu is 488–593 K). To summarize Zachariasen's work, γ-Pu is face-centered orthorhombic (Fddd), with eight equivalent atoms in the unit cell at $\left(0,0,0;0,\frac{1}{2},\frac{1}{2};\frac{1}{2},0,\frac{1}{2};\frac{1}{4},\frac{1}{4},\frac{1}{4};\frac{3}{4},\frac{3}{4},\frac{1}{4};\frac{3}{4},\frac{1}{4},\frac{3}{4};\frac{1}{4},\frac{3}{4},\frac{3}{4}\right)$. Each Pu atom is bonded to ten others at an average distance of 3.157 Å, four neighbors

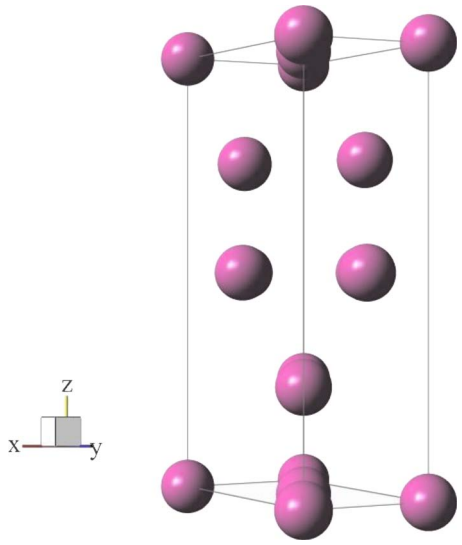


FIG. 1. (Color online) Crystal structure of γ-Pu, viewed perspectively along [110], with origin chosen as [000]. Cell shows Fddd-space-group (centrosymmetric, no. 70) and contains eight atoms with identical Wyckoff symmetry. At ambient pressure, no other metallic element shows this space group, a space group that occurs in many compounds.

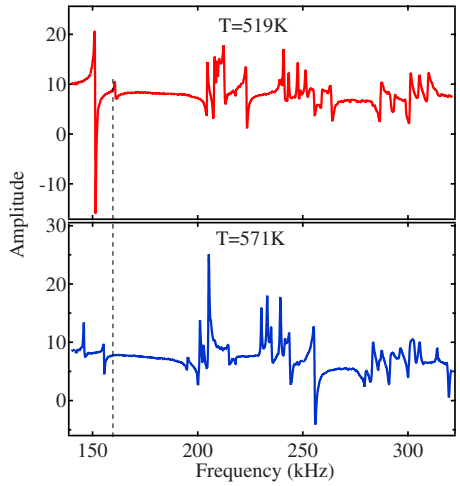


FIG. 2. (Color online) Part of the resonance frequency spectrum of polycrystalline Pu shown near the low- and high-temperature ends of the γ-phase region of existence. The dashed line provides a guide to the eye to emphasize the size of the shift of the resonances with temperature.

are at 3.026 Å, two at 3.159 Å, and four at 3.288 Å. However, in the space group Fddd one has the choice of two origins. By placing the origin at one center of symmetry, Roof[8] pointed out that the structure can be alternatively described with the eight equivalent atoms at $\pm(\frac{1}{8},\frac{1}{8},\frac{1}{8})$ $+(0,0,0;0,\frac{1}{2},\frac{1}{2};\frac{1}{2},0,\frac{1}{2};\frac{1}{2},\frac{1}{2},0)$. This description is consistent with the description of other Pu phases. Along the c-axis, the atoms arrange themselves in nearly close-packed-hexagonal layers, suggesting a clear possible lattice correspondence with face-centered cubic (f.c.c.) δ-Pu. Indeed the γ-Pu crystal structure can be considered as a four-layer distorted hexagonal crystal. The unit cell dimensions in Å, at 535 K, are given by Zachariasen[9] to be $a=3.159\pm0.001$, $b=5.768\pm0.001$, and $c=10.162\pm0.002$. From X-ray experiments,[9] the X-ray density was calculated to be 17.14 ± 0.01 g/cm^3 at 508 K. Zachariasen pointed out that the γ-Pu crystal structure differs from that of any other metal element. The crystal structure of γ-Pu is represented in Fig. 1. Thus, γ-Pu joins α-Pu and β-Pu (both monoclinic) in showing crystal structures typical of compounds (ionic-covalent) than of pure metals. This observation relates to a suggestion by Lawson *et al.* that Pu tends to form "self-intermetallic" compounds.[10]

The only other available data for the elastic moduli of γ-Pu were obtained from measurements of longitudinal, flexural, and torsional resonances,[11] using a Young's modulus and rigidity modulus apparatus. Elastic moduli were obtained after 2 and 3 thermal cycles, a procedure we observed to produce inconsistencies in the phase of the specimen. In our measurements the sample was not cycled, but warmed only once from ambient temperature to just past the γ-δ phase boundary.

II. MEASUREMENT METHOD

Pure electrorefined ^{239}Pu was used for these experiments. Using chemical methods, the composition was determined to be 99.96 wt % Pu, 115 ppm W, 49 ppm Np, 50 ppm O, 53 ppm Si, 32 ppm Am, and less than 25 ppm other impurities. The parallelepiped-shaped sample was cut from a larger button that was arc melted and quenched on a copper hearth several times, until voids and metallurgical imperfections were observed with microscopy to be negligible. The specimen was $0.265\times0.268\times0.270$ cm^3 and the immersion density at 300 K was determined to be 19.55 g/cm$^3\pm0.02\%$.[1] This value is lower than both the density determined from mass and dimensions (19.70 g/cm^3) and the X-ray diffraction density (19.86 g/cm^3 at 294 K).

The RUS system used for this work is described elsewhere.[1] The experiments were carried out in vacuum. The temperature was measured using a RhFe resistance thermometer. At 400 K, we obtained a temperature accuracy of 2%. From the measured resonance frequencies, we computed the elastic moduli using a code that employs the Levenberq–Marquardt algorithm. The code was made available for public use by us.[12]

Figure 2 shows the resonance spectrum from 140 to 320 kHz at two temperatures, i.e., one near the low end, while the other near the high end of the range of existence of γ-Pu. The high signal-to-noise ratio and the high Q (Q was of order 1000) resonances ensure high-accuracy results for this study. The lowest resonance was observed at about 150 kHz. As the temperature was increased, all resonances shifted toward lower values. Using our own algorithms, the center frequencies of 21 resonances were determined at each temperature. Each resonance corresponds to a different vibrational mode. Keeping in mind that in this temperature range the variation in elastic moduli of Pu is featureless, and either the raw resonances or the final computed elastic moduli are to be fitted to a function useful for models, the choice is to either fit the raw data first, and use that smooth fitted to compute for the elastic moduli, or to use the weakly scattered raw results from the fitting codes and then fitted the resulting moduli. Because the fitting codes tend to amplify scatter for the bulk modulus (only a few resonances are strongly depen-

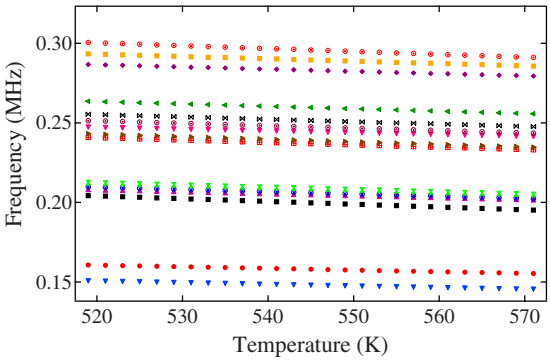


FIG. 3. (Color online) Resonance frequencies versus temperature. All frequencies show a smooth linear decrease with increasing temperature. These fitted values were used in the computation routine to determine the elastic moduli.

dent on it), the best fidelity to the actual measurement is to fit the resonances first, then process the fitted values. Thus the frequencies of each mode versus temperature were fitted to a straight line, a function entirely adequate to capture the temperature dependence of this phase. The linear fit was used to generate input frequencies for the RUS code,[12] which calculates the elastic moduli. A plot of the fitted resonances versus temperature is shown in Fig. 3. In using the fitting codes[12] that determine elastic moduli from measured frequencies, the algorithms can also introduce scatter via small artifacts associated with modes changing relative ordering as temperature changes, introducing corresponding artifacts in what is observed, from the raw resonances, to be smooth behavior with temperature of the moduli. Using linear fits to resonance frequencies as described above is often sufficient to block such artifacts, but smoothness in the mean-square fitting error is a precise test for their presence. Small values of the fitting error are also a useful and quantitative test for isotropy of the polycrystalline specimen. This test is shown in Fig. 4, where we plot root-mean-square (RMS) and Chi-square errors[12] computed by the fitting algorithm versus the moduli, and find no unexpected behavior and strong evidence of isotropy manifested in the small values of the fitting errors. The RMS error has the same value for the shear and bulk modulus.

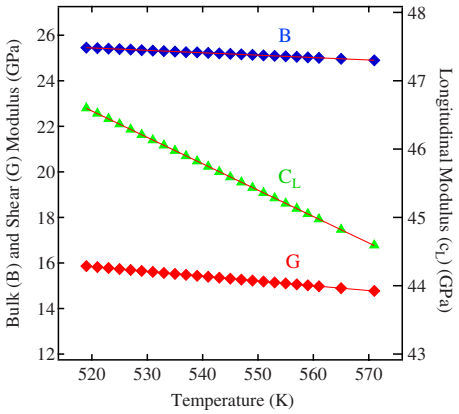


FIG. 4. (Color online) RMS error of longitudinal elastic modulus and Chi-square error of shear and longitudinal elastic moduli versus moduli to test for fitting and isotropy errors. Note: RMS error has the same value for shear and longitudinal moduli.

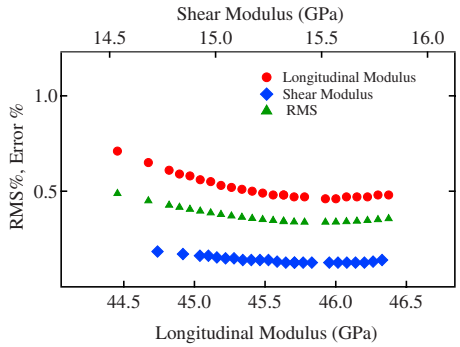


FIG. 5. (Color online) Temperature dependence of bulk (B), shear (G), and longitudinal (C_L) elastic moduli of γ-Pu.

III. RESULTS AND DISCUSSION

The squares of the resonance frequencies, absent small density corrections, are proportional to the elastic moduli, and therefore, the observed decrease in the resonance frequencies with increasing temperature is expected.

The elastic behavior of any isotropic polycrystalline material can be described completely by two independent moduli, for example, the longitudinal (C_L) and the shear (G) moduli. The bulk modulus (B) is related to C_L and G through the following relation:

$$B = C_L - \frac{4G}{3}. \tag{1}$$

Figure 5 shows the smooth, linearly decreasing behavior of C_L, G, and B as a function of temperature. Each elastic modulus was fitted with a straight line and the parameters of the fit are given in Table I.

In Fig. 6 we compare the elastic moduli of pure polycrystaline γ-Pu to those of α-Pu [1] and Ga-stabilized δ-Pu (2.36 at. % Ga)[13] and found, as expected, that, modulus by modulus, the moduli of γ-Pu lie between those of α-Pu (highest) and Ga-stabilized δ-Pu (lowest).

We fitted the higher-temperature values of the elastic moduli of α-Pu and δ-Pu (2.36 at % Ga) with a straight line, and compared the parameters of the linear fit to the ones obtained for γ-Pu in the temperature ranges indicated in the tables. Table II shows the results of the linear fit for the three phases. We found that the slope of the fit for the bulk modulus of γ is much lower than the slope of the bulk modulus of the other two phases. A lower slope dB/dT implies lower anharmonicity. Using a relationship derived by Ledbetter[14]

$$\frac{dB}{dT} = -3k\gamma(\gamma+1)/V_a, \tag{2}$$

TABLE I. Parameters of the linear fits to the measurements of C_L, B, and G of the form $a+bT$.

Modulus	a (GPa)	b (GPa/K)
C_L	66.64	−0.03864
B	30.96	−0.01061
G	26.75	−0.02099

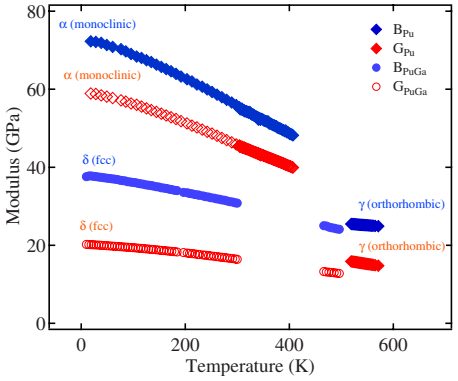


FIG. 6. (Color online) Elastic moduli versus temperature of α-Pu, γ-Pu, and δ-Pu-2.36 at % Ga.

where k is Boltzmann's constant and V_a is the atomic volume, we calculate that the Grüneisen parameter $\gamma=2.0$, which is the same value as for copper ($\gamma=2.0$). Thus, γ-Pu shows normal anharmonicity. α-Pu has $\gamma=3.5$,[15] while δ-Pu has $\gamma=3.7$.[13]

We compared the fitted parameters of the linear fits to the normalized temperature dependence of the shear and bulk modulus in all phases. The results are presented in Table III. The fractional variation in the shear modulus with temperature is similar for α-Pu and γ-Pu, but different for Pu (2.36 at % Ga) alloy. The fractional variation in the bulk modulus with temperature is unexpectedly small for γ-Pu compared to α-Pu and Ga-stabilized δ-Pu.

Poisson's ratio ν for polycrystals is given by

$$\nu = \frac{1}{2}\frac{3B-2G}{3B+G}. \tag{3}$$

The Poisson ratio, as discussed elsewhere for Pu in general,[1] for γ-Pu is unusual in that, unlike the Poisson ratio for the other Pu phases, ν for γ-Pu has a normal value and increases normally with increasing temperature, as shown in Fig. 7. Figure 8 shows a comparison of Poisson ratio versus temperature of aluminum, α-Pu, γ-Pu, Pu (2.36 at % Ga), and Pu (5 at % Ga). The large differences in $d\nu/dT$ strongly suggest large differences in interatomic-bonding changes with temperature.[16] Also, extrapolation to a much lower Poisson ratio at low temperatures suggests strongly the presence of significant covalent bonding.[17]

Although it is unusual to calculate the Debye temperature Θ_D from higher-temperature elastic moduli, it is useful to do so for γ-Pu because Θ_D facilitates useful comparisons with other Pu phases and other materials, and because γ-Pu is only poorly characterized. We calculated Θ_D using the following well-known relationship:

$$\Theta_D = \frac{h}{k}\left(\frac{3}{4\pi V_a}\right)^{1/3} v_m. \tag{4}$$

Here, h denotes Plank's constant, k denotes Boltzmann's constant, V_a denotes atomic volume, and v_m denotes mean sound velocity calculated from mass density, B, and G (extrapolated to $T=0$ K by assuming linear temperature dependence for B and G) in the usual way.[18,19] Our approximate result is $\Theta_D=144$ K, about half way between the Debye

TABLE II. Parameters of the temperature fit for two elastic moduli of γ-Pu and a comparison with those for α-Pu and Ga-stabilized δ-Pu for the temperature ranges indicated.

	Moduli	
Phases	G (GPa)	B (GPa)
α-Pu	$a=62.84\pm0.25$	$a=76.5\pm1.18$
$T=381-407$ K	$b=(-5.6\pm0.06)\times10^{-2}$	$b=(-6.9\pm0.3)\times10^{-2}$
γ-Pu	$a=26.75\pm0.02$	$a=30.96\pm0.02$
$T=519-571$ K	$b=(-2.099\pm0.003)\times10^{-2}$	$b=(-1.061\pm0.004)\times10^{-2}$
δ-PuGa	$a=20.50\pm0.19$	$a=40.1\pm0.96$
$T=466-496$ K	$b=(-1.6\pm0.04)\times10^{-2}$	$b=(-3.2\pm0.2)\times10^{-2}$

TABLE III. Logarithmic temperature derivatives for the elastic moduli of γ-Pu and a comparison with those for α-Pu and Ga-stabilized δ-Pu for the temperature ranges indicated.

	$d\ln \text{Moduli}/dT$	
Phases	$\frac{d\ln G}{dT}$	$\frac{d\ln B}{dT}$
α-Pu		
$T=381-407$ K	$-1.39\times10^{-3}(\pm1.1\%)$	$-1.41\times10^{-3}(\pm0.9\%)$
γ-Pu		
$T=519-571$ K	$-1.37\times10^{-3}(\pm3.6\%)$	$-0.42\times10^{-3}(\pm1.1\%)$
δ-PuGa		
$T=466-496$ K	$-1.19\times10^{-3}(\pm0.2\%)$	$-1.32\times10^{-3}(\pm5.9\%)$

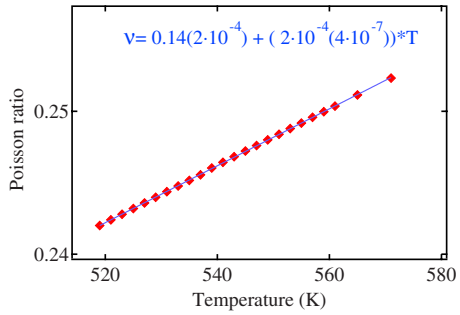


FIG. 7. (Color online) Temperature variation of Poisson's ratio for γ-Pu. The line is a fit of Poisson's ratio versus temperature. The parameters of the fit are given in the figure.

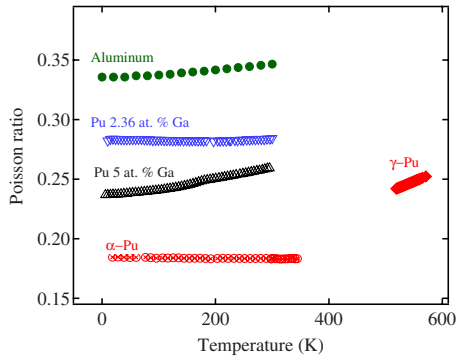


FIG. 8. (Color online) Poisson's ratio versus temperature of aluminum: α-Pu, γ-Pu, Pu-2.36 at % Ga, and δ-Pu-5 at % Ga.

temperatures of α-Pu and δ-Pu, which is 205[10] and 117 K,[18] respectively, a result expected intuitively from simple lattice physics-thermodynamics.

IV. CONCLUSIONS

We report measurement of the full set of elastic moduli of a high purity, isotropic, polycrystalline specimen of ^{239}Pu throughout almost the entire range of the existence of its γ-phase. We found an unusually ordinary behavior of Poisson's ratio (compared with α-Pu and δ-Pu), and unusually low variation in the bulk modulus with temperature. The measured moduli fall, as expected, below those of the lower-temperature and expectedly stiffer α-phase, and above those of the higher-temperature and expectedly softer δ-phase. The measurements pass several internal consistency tests. One previous study approximately agrees with our results.[11]

ACKNOWLEDGMENTS

This work was done under the auspices of the U.S. Department of Energy at Los Alamos National Laboratory in the National High Magnetic Field Laboratory, and was supported by the U.S. National Nuclear Security Administration under Grant No. 20070013DR, the National Science Foundation under Grant No. DMR-0654118, and the State of Florida.

[1]A. Migliori, C. Pantea, H. Ledbetter, I. Stroe, J. B. Betts, J. N. Mitchell, M. Ramos, F. Freibert, D. Dooley, S. Harrington, and C. H. Mielke, "Alpha-plutonium's polycrystalline elastic moduli over its full temperature range," J. Acoust. Soc. Am. **122**, 1994–2001 (2007).
[2]A. Migliori and J. L. Sarrao, *Resonant Ultrasound Spectroscopy* (Wiley-Interscience, New York, 1997).
[3]A. Migliori, J. L. Sarrao, W. M. Visscher, T. M. Bell, M. Lei, Z. Fisk, and R. G. Leisure, "Resonant ultrasound spectroscopic techniques for measurement of the elastic moduli of solids," Physica B **183**, 1–24 (1993).
[4]V. Keppens and A. Migliori, "Elastic properties and thermodynamics," Exp. Methods Phys. Sci. **39**, 221–235 (2000).
[5]A. Migliori, T. W. Darling, and J. P. Baiardo, "Resonant ultrasound spectroscopy (RUS)," in *Handbook of Elastic Properties of Solids, Liquids, and Gases*, edited by M. Levy, H. Bass, and R. Stern (Academic, New York, 2001), Vol. **1**, p. 239.
[6]W. H. Zachariasen, "Crystal-structure studies of plutonium metal," in *The Metal Plutonium*, edited by A. S. Coffinberry and W. N. Miner (University of Chicago Press, Chicago, 1961), pp. 99–107.
[7]D. C. Wallace, "Electronic and phonon properties of six crystalline phases of Pu metal," Phys. Rev. B **58**, 15433–15439 (1998).
[8]R. B. Roof, *Advances in X-Ray Analysis* (Plenum, New York, 1981), p. 221.
[9]W. H. Zachariasen and F. H. Ellinger, "Crystal chemical studies of the 5f-series of elements. XXIV. The crystal structure and thermal expansion of γ-Pu," Acta Crystallogr. **8**, 431–433 (1955).
[10]A. Lawson, J. Goldstone, B. Cort, R. Martinez, F. Vigil, T. Zocco, J. Richardson, and M. Mueller, "Structure of ζ-phase plutonium-uranium," Acta Crystallogr., Sect. B: Struct. Sci. **52**, 32–37 (1996).
[11]A. E. Kay and P. F. Linford, "The elastic constants of plutonium," in *Plutonium 1960*, edited by E. Grison, W. B. H. Lord, and R. D. Flower (Cleaver-Hume, London, 1961), p. 51.
[12]www.magnet.fsu.edu/inhouseresearch/rus/index.html (Last viewed November 9, 2009).
[13]A. Migliori, H. Ledbetter, A. C. Lawson, A. P. Ramirez, D. A. Miller, J. B. Betts, M. Ramos, and J. C. Lashley, "Unexpected elastic softening in δ-Pu," Phys. Rev. B **73**, 052101 (2006).
[14]H. Ledbetter, "Relationship between bulk-modulus temperature-dependence and thermal expansivity," Phys. Status Solidi B **181**, 81–85 (1994).
[15]H. Ledbetter, A. Lawson, and A. Migliori, "Alpha-plutonium's Grüneisen parameter," Los Alamos National Laboratory Report No. LAUR 05-5072, Los Alamos, NM (2009).
[16]W. Koster and H. Franz, "Poisson's ratio for metals and alloys," Metall. Rev. **6**, 1–55 (1961).
[17]H. Ledbetter and A. Migliori, "Elastic-constant systematics in f.c.c. metals, including lanthanides-actinides," Phys. Status Solidi B **245**, 44–49 (2008).
[18]H. Ledbetter, "Estimation of Debye temperatures by averaging elastic coefficients," J. Appl. Phys. **44**, 1451–1454 (1973).
[19]H. Ledbetter and A. Migliori, "Zero-temperature monocrystal elastic coefficients of gallium-stabilized γ-Pu," Los Alamos National Laboratory Report No. LAUR 05-5072, Los Alamos, NM (2009).

Structure of phonon focusing patterns in tetragonal crystals

Litian Wang
Department of Mathematics and Physics, Østfold University College, 1757 Halden, Norway

(Received 3 June 2009; revised 28 September 2009; accepted 23 November 2009)

A set of critical conditions for the characteristic caustic points in the phonon focusing patterns in tetragonal crystals is formulated. A caustic line segment in the focusing pattern is generally associated with either a fold or a cusp on the wave surface. Most of the caustic lines are symmetrical with respect to the principal symmetry plane and can be characterized by the caustic points at the centers of the caustic lines. These characteristic caustic points originated from inflection/parabolic points with zero in-plane/ex-plane curvature, respectively. By employing the Stroh formalism, the inflection/parabolic points on the slowness surface are studied in terms of the so-called zero-curvature transonic states. Since these transonic states are related to extraordinary degeneracies in the Stroh eigenvalue equation, the conditions for the degeneracies can be regarded as critical conditions for the characteristic caustic points. These conditions provide an overview of global structure of the phonon focusing patterns in tetragonal crystals. A set of caustic lines in vicinity of (001) plane is also investigated and exemplified. © *2010 Acoustical Society of America.*
[DOI: 10.1121/1.3273901]

PACS number(s): 43.35.Gk, 43.20.Bi [RW] Pages: 746–753

I. INTRODUCTION

Acoustic wave propagation in anisotropic media is governed by the Christoffel equation, and the phase velocity is described by the slowness surface.[1,2] Among three slowness sheets, two outer ones can be concave locally, and the so-called parabolic line, along which the Gaussian curvature is zero, divides the slowness sheet into concave, saddle, and convex regions. Because of the concavities, the wave surface, which describes the group velocity, will have some folds, and they result in various types of caustic lines in the phonon focusing pattern. The main characteristics of the phonon focusing pattern are therefore directly related to the geometry of the parabolic lines.[3]

The phonon focusing pattern was first studied extensively in cubic crystals by examining the wave surface, and a complete set of critical existence conditions was formulated.[4] Many numerical and experimental investigations were done in documenting the various patterns and their features.[5–7] The understanding of the focusing patterns in tetragonal crystals, however, has been mainly based on numerical simulations,[8–10] and the focusing patterns were classified into 11 categories by comparing their features near the [001] direction to those in cubic crystals.[11] By calculating the Gaussian curvature in the vicinity of two/fourfold symmetry axes, four criteria for negative Gaussian curvature, together with two parameters, were established for categorizing the focusing patterns in detail.[12–14]

Recently, the present author developed an analytical scheme for finding locations of various caustic points in the symmetry planes in cubic crystals.[15,16] The scheme is based on the so-called Stroh formalism for two-dimensional elastodynamics, where the zero-curvature points on the slowness surface are treated in terms of highly degenerated Stroh eigenvalues. The main idea for the scheme will be adopted in this investigation.

Generally, the phonon focusing patterns in tetragonal crystals consist of two types of caustic lines originated from fold/cusp on the wave surface. The existence of these caustic lines can be studied by concentrating on the central points of fold/cusp, which appear usually at principal symmetry planes. Let $(\pm,\pm)$ denote the principal in-plane and ex-plane (transverse) curvature, respectively, with respect to a symmetry plane. Whenever a concave region appears on the slowness surface, the local geometry along the symmetry plane would evolve in following fashions:

(a) $(++)\rightarrow(+-)\rightarrow(--)$,
(b) $(++)\rightarrow(-+)\rightarrow(--)$, and
(c) $(++)\rightarrow(--)$.

The transition $(+,)\rightarrow(-,)$ will take place at an inflection point with vanishing in-plane curvature, while the transition $(,+)\rightarrow(,-)$ will occur at a parabolic point with vanishing transverse curvature. Case (c) refers to a rare case where an inflection point coalesces with a parabolic point, which is a monkey saddle point often found in the cubic crystals.[4,16] Figure 1 illustrates two segments of parabolic lines crossing a symmetry plane at these points and their corresponding caustic lines: The inflection point yields the central point of a parabola-formed caustic line associated with a fold on the wave surface [Fig. 1(a)], while the parabolic point produces the tip of two semicubical parabola-formed (swallowtail) caustic lines associated with a cusp on the wave surface [Fig. 1(b)]. We denote therefore these two caustic points at the symmetry plane as the cuspidal point and the swallowtail point, respectively.

In tetragonal crystals, there will be three sets of cuspidal (C)/swallowtail (S) points along the principal symmetry planes ({100}, {110}, and {001}). Figure 2 illustrates the focusing pattern of InBi where the caustic points along the (010) plane (C_1-S_1, C_2-S_2) and the (110) plane (C_3-C_4-S_6) can be clearly identified. The central region of the pattern

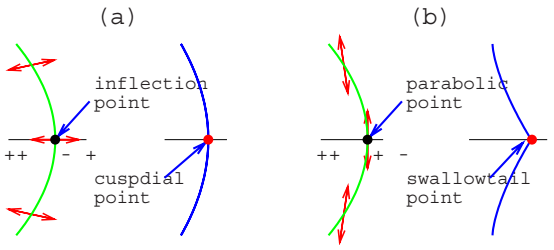


FIG. 1. (Color online) Correspondence between the parabolic line and the caustic line in the neighborhood of a symmetry plane. The small arrows mark directions with vanishing principal curvature. A parabolic line can yield (a) a fold centered at the cuspidal point or (b) a cusp centered at the swallowtail point.

resembles those in cubic crystals because of the presence of a tangential acoustic axis along [001] direction.

Since the concavities on the slowness surface usually originated along the principal symmetry axes as one perturbs elastic constants, we can define a series of critical existence conditions for the emergence of inflection/parabolic points along the symmetry axes, which also serve as critical conditions for the existence of the cuspidal/swallowtail points. In this paper, we will reveal that the common feature among the inflection and parabolic points is that they are all associated with extraordinary degeneracies in the Stroh eigenvalue equation, and their existence conditions can be formulated in terms of the so-called zero-curvature transonic states.[17]

Another objective for this study is to explore the focusing pattern in the vicinity of the (001) plane. The focusing pattern here is different from those along [001] direction due to the absence of tangential acoustic axes along [100]/[010] directions, and it is also typical for other crystals with lower crystal symmetry. We will show that the absence of the tangential acoustic axis leads to new phenomena concerning the coexistence of different caustic lines.

In Sec. II, we will briefly introduce the Stroh formalism and the zero-curvature transonic states. Critical existence conditions for the transonic states are derived by studying degeneracies of the Stroh eigenvalue equation in Sec. III. Finally, the critical existence conditions for the cuspidal/swallowtail points are summarized, and the structure of focusing patterns in tetragonal crystals is discussed and exemplified in Sec. V.

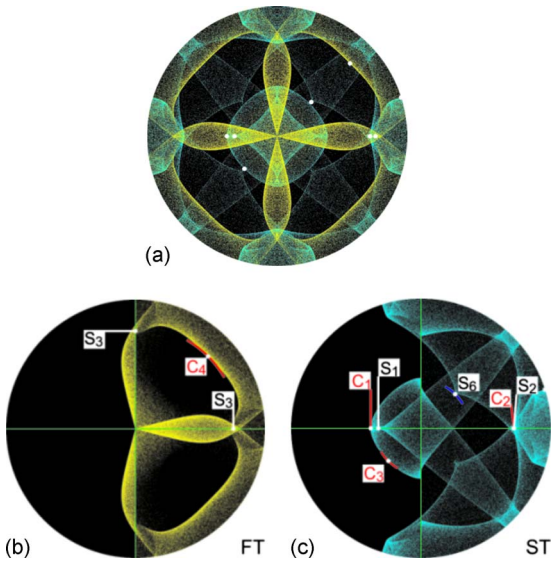


FIG. 2. (Color online) (a) Phonon focusing pattern of InBi. The cuspidal points (C_i) and swallowtail points S_i in the (010) and (110) planes are marked with filled circles. (b) and (c) Partial patterns generated from a half of slowness surface.

II. STROH FORMALISM AND DEGENERACY

The Stroh formalism is an effective approach in two-dimensional elastodynamics in dealing with the wave propagation in anisotropic elastic media. Consider infinite homogeneous elastic media. The equation of motion is given by

$$c_{ijkl}\frac{\partial^2 u_k}{\partial x_j\,\partial x_l} = \rho\frac{\partial^2 u_i}{\partial t^2}, \tag{1}$$

where c_{ijkl} is the elastic stiffness tensor, ρ is the density, and u is the displacement. We define a reference plane $R=(\mathbf{m},\mathbf{n})$, where $\mathbf{m}$ and $\mathbf{n}$ are two unit vectors normal to each other, and consider steady state solutions described by $\mathbf{u}(\mathbf{x})=\mathbf{a}\exp\{ik[(\mathbf{m}+p\mathbf{n})\cdot\mathbf{x}-vt]\}$ with the wave vector given by $\mathbf{k}=k(\mathbf{m}+p\mathbf{n})$, (the value p represents then the direction tangent $p=\tan\phi$ for the wave vector $\mathbf{k}$, where ϕ is the angle between $\mathbf{k}$ and $\mathbf{m}$). The polarization vector $\mathbf{a}$ will obey

$$\Gamma\mathbf{a} = \{(mm) - \rho v^2 I + p[(mn)+(nm)] + p^2(nn)\}\mathbf{a} = 0, \tag{2}$$

where the term (mn), etc., are given by $(mn)_{ik}=m_j c_{ijkl} n_l$, and the repeated subscripts denote summation. The traction $\mathbf{b}$ exerted on the planes parallel to $\mathbf{n}\cdot\mathbf{x}=0$ is given by $\mathbf{b}=-[(nm)+p(nn)]\mathbf{a}$. The Stroh formalism[18–20] is based on the construction of a six-dimensional vector $\xi=[\mathbf{a},\mathbf{b}]^T$ so that Eq. (2) can be transformed into a six-dimensional eigenvalue equation

$$N\xi_i = p_i\xi_i, \tag{3}$$

where

$$N = -\begin{pmatrix} (nn)^{-1}(nm) & (nn)^{-1} \\ (mn)(nn)^{-1}(nm) - (mm) + \rho v^2 I & (mn)(nn)^{-1} \end{pmatrix},$$

and I is a 3×3 identity matrix and $i=1,\ldots,6$. The characteristic equation for Eqs. (2) and (3) will be a sextic equation in p,

$$\sum_{n=0}^{6} a_n(v)p^n = 0. \tag{4}$$

Note that in the Stroh eigenvalue equation, the velocity v becomes a parameter.

The Stroh formalism provides a simple way to determine the group velocity in terms of the transonic state (see Fig. 3). Given a reference plane $R=(\mathbf{m},\mathbf{n})$, let the parameter $v^{-1}=v_0^{-1}$ be such that the Stroh equation yields two real eigenvalues, $p_1=\tan\phi_1$ and $p_2=\tan\phi_2$, representing two bulk-waves pertaining to the outer slowness curve. This solution can be represented by the dashed vertical line L_0 and its two intersection points on the slowness curve. With the increasing slowness v^{-1} (or moving L to v_1^{-1}), we will reach the so-called *transonic state* where two eigenvalues (p_1,p_2) coalesce and leave $p_1=p_2=\tan\phi$, representing one bulkwave, as

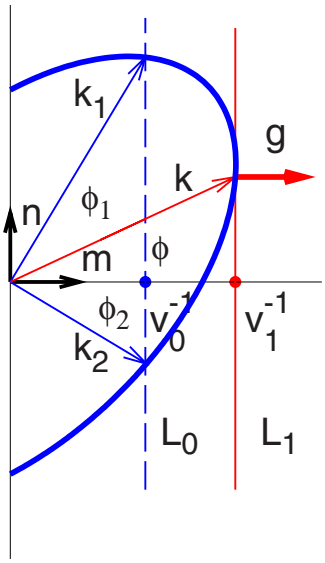


FIG. 3. (Color online) Geometrical interpretation of two solutions of the Stroh equation associated with $v^{-1}=v_0^{-1}$ and $v^{-1}=v_1^{-1}$, represented by the vertical lines L_0 and L_1, respectively. The solid curve (slowness plot) represents the outermost slowness curve. With increasing v^{-1}, two bulkwave solutions with $\mathbf{k}_1$ and $\mathbf{k}_2$ at $v^{-1}=v_0^{-1}$ coalesce into a single bulkwave $\mathbf{k}$ at v_1^{-1}.

shown by line L_1. Since line L_1 tangentially touches the slowness curve, (i) the basis vector $\mathbf{m}$ would now define the curve normal $\mathbf{g}$; (ii) v_1^{-1}, marking the location of L_1, defines the group velocity v_g^{-1}; and (iii) the degenerate eigenvalue $p=\tan\phi$ defines the angle between $\mathbf{k}$ and $\mathbf{g}$. For this bulkwave, the phase velocity and group velocity can be clearly read from the figure: $v_p^{-1}=v_1^{-1}/\cos\phi$ and $v_g^{-1}=v_1^{-1}$.

There are six ordinary transonic states (Types 1–6) and four extraordinary zero-curvature transonic states (Types E1–E4).[17,19] At an ordinary transonic state, the Stroh eigenvalues become degenerate pairwise. For example, we have $p_1=p_2$ and $p_3=p_4$ at a Type 4 transonic state and $p_1=p_2$ and $p_5=p_6$ at a Type 2 transonic state [see Figs. 4(a) and 4(c)]. At a zero-curvature transonic state, there are at least four Stroh eigenvalues becoming degenerate. At an E1 transonic state, four eigenvalues become degenerate, while at an E3 transonic state, all six eigenvalues coalesce [see Figs. 4(b) and 4(d)]. In other words, a zero-curvature point on the slowness surface is always associated with degeneracies in the Stroh eigenvalue equation.

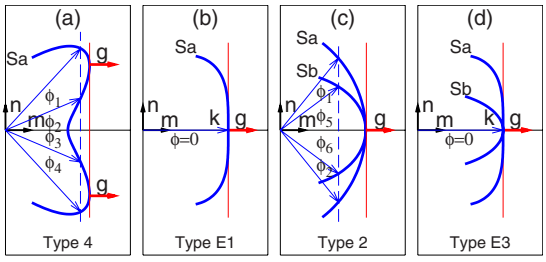


FIG. 4. (Color online) Slowness curves and transition of symmetric transonic states. (a) and (b) Transition from Type 4 to Type E1 when $\phi_{1,\ldots,4}\rightarrow 0$. (c) and (d) Transition from Type 2 to Type E3 when $\phi_{1,\ldots,6}\rightarrow 0$. The solid vertical lines represent the transonic states, and the vector $\mathbf{g}$ represents the curve normal.

Letting $\mathbf{n}$ be the normal of a symmetry plane, the slowness curves are then symmetric with respect to the plane. We will consider now two scenarios. (i) When we trace the transonic state [such as the solid vertical lines in Figs. 4(a) and 4(b)] by rotating $\mathbf{m}$ about $\mathbf{n}$, we can observe transition from a Type 4 transonic state to an E1 transonic state. The E1 transonic state marks then a transverse zero-curvature point on S_a, and hence a parabolic point on the slowness surface. (ii) Since there is always a tangential acoustic axis along [001] in tetragonal crystals, by perturbing elastic constants, a convex outer slowness curve can bulge out and become concave via an E3 transonic state [see Figs. 4(c) and 4(d)]. Such an E3 transonic state marks then a limiting case for the presence of the inflection point in the vicinity of [001] axis.

Note that the degenerate eigenvalues are equal to zero for both E1 and E3 transonic states because $\mathbf{m}$ is defined in the symmetry plane. This leads to two simple conditions for these zero-curvature transonic states,

$$\text{Type E1:} p_i=\tan\phi_i=0, \quad i=1,\ldots,4, \tag{5}$$

$$\text{Type E3:} p_i=\tan\phi_i=0, \quad i=1,\ldots,6. \tag{6}$$

The constraints on the Stroh eigenvalues imposed by Eqs. (5) and (6) can be transformed into simple algebraic conditions. Whenever $\mathbf{m}$ is defined in a symmetry plane, the characteristic equation of Eq. (2), $\|\Gamma\|=0$, can always be expressed in terms of a bicubic equation in p,

$$a_6p^6+a_4p^4+a_2p^2+a_0=0. \tag{7}$$

This suggests that conditions (5) and (6) can be rewritten as

$$\text{Type E1:} a_0=a_2=0,$$

$$\text{Type E3:} a_0=a_2=a_4=0, \tag{8}$$

which define the critical existence conditions for the parabolic/inflection points with respect to a symmetry plane containing $\mathbf{m}$, respectively.

Note that the equation $a_0=\|(mm)-\rho v^2 I\|=0$ is actually the characteristic equation for the Christoffel equation for the bulkwave propagation in the symmetry plane, and its solutions represent three transonic state branches (S_a,S_b,S_c), which are identical to the three slowness sections in the symmetry plane. For an E3 transonic state, $a_2=0$ is constantly zero because of the presence of the tangential acoustic axis.

III. EXISTENCE CONDITIONS FOR ZERO-CURVATURE TRANSONIC STATES

In this section, we will examine the critical existence conditions for the zero-curvature transonic states in three principal symmetry planes in tetragonal crystals: (010), $(\bar{1}10)$, and (001). Let us first outline the procedure as follows: (i) Define the reference plane $R=(\mathbf{m},\mathbf{n})$ with $\mathbf{m}$ lying in a symmetry plane and $\mathbf{n}$ parallel to its normal; (ii) formulate the matrix Γ and corresponding characteristic equation and identify all coefficients $a_n(v,\theta)$; (iii) define transonic state branches (S_a,S_b) with (v_a,v_b) by solving $a_0(v,\theta)=0$; and (iv) formulate explicitly conditions for the E1/E3 tran-

sonic states from $a_2(v,\theta)=0/a_4(v,\theta)=0$ on each transonic state branch along symmetry directions. For the sake of brevity, we will introduce the following notions:

$$\Delta_{kl}=c_{kk}-c_{ll},\quad \lambda_{kl}=c_{kk}\cos^2\theta+c_{ll}\sin^2\theta,$$

$$d_{12}=c_{12}+c_{66},\quad \Delta=c_{11}-c_{12}-2c_{66},$$

$$d_{13}=c_{13}+c_{44},\quad \delta_{46}=c_{44}-(c_{11}-c_{12})/2.$$

A. Inflection and parabolic points in (010) plane

Construct a reference plane R with $\mathbf{n}=\mathbf{e}_y$ and $\mathbf{m}=\cos\theta\mathbf{e}_z+\sin\theta\mathbf{e}_x$. The matrix Γ in Eq. (2) is given by

$$\Gamma_{11}=p^2c_{11}-\rho v^2+\lambda_{46},\quad \Gamma_{12}=d_{12}\sin\theta,$$

$$\Gamma_{22}=p^2c_{66}-\rho v^2+\lambda_{41},\quad \Gamma_{13}=pd_{13}\cos\theta,$$

$$\Gamma_{33}=p^2c_{44}-\rho v^2+\lambda_{34},\quad \Gamma_{23}=pd_{13}\sin\theta\cos\theta. \tag{9}$$

The coefficients a_0,a_2,a_4,a_6 of the characteristic equation are functions of the velocity v and the direction θ. $a_0(v,\theta)=0$ will produce three transonic state branches, where the two slowest ones (S_a,S_b) are given by

$$\rho v_a^2=\tfrac{1}{4}\{c_{33}+2c_{44}+c_{33}\cos 2\theta+2c_{11}\sin^2\theta+2[4c_{33}c_{44}\cos^4\theta+4c_{11}c_{44}\sin^4\theta-(2\Delta_{14}-2\Delta_{13}\cos^2\theta)^2/4+(d_{13}^2-c_{11}c_{33}-c_{44}^2)\sin^2 2\theta]^{-1/2}\}, \tag{10a}$$

$$\rho v_b^2=c_{44}\cos^2\theta+c_{66}\sin^2\theta, \tag{10b}$$

respectively. $a_2(\theta)|_{v_a}=0$ and $a_2(\theta)|_{v_b}=0$ will then determine the locations of the E1 transonic states on S_a and S_b, respectively, and $a_2(\theta)|_{v_a}=a_4(\theta)|_{v_a}=0$ will define the existence condition for the E3 transonic state on S_a. Since most of the concave regions on the slowness surface emerge in the vicinity of symmetry directions, we can examine when the E1/E3 transonic state appears or approaches the symmetry directions.

Along [001] direction $\rho v_a^2=\rho v_b^2=c_{44}$, we have a tangential acoustic axis with $a_0(0)|_{v_a}=a_2(0)|_{v_a}=0$ constantly. The E3 transonic state will occur when $a_4(0)|_{v_a}=0$, which gives

$$a_4\propto(c_{11}\Delta_{34}-d_{13}^2)=0\leftrightarrow \mathrm{A}. \tag{11}$$

Here we denote the equation as condition A. This condition hence refers to the transverse zero-curvature with respect to (010) plane [see Fig. 5(a)], and at the same time, it refers also to the in-plane zero-curvature with respect to (100) plane because of the fourfold symmetry. Therefore, condition A can be regarded as a critical existence condition for the inflection point in the (010) plane.

Along off-symmetry directions, only E1 transonic states can exist on S_a and S_b branches due to the absence of the tangential acoustic axes. The E1 transonic states refer then to the parabolic points, and their locations can be determined by $a_2=0$. We will study the limiting behavior of these parabolic points as they approach the symmetry directions [001] and [100], respectively.

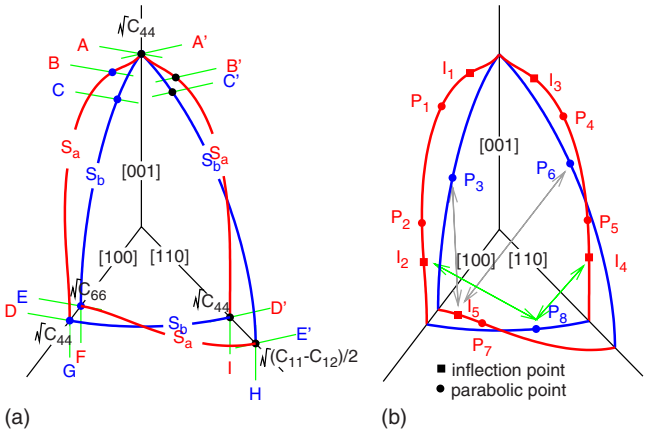


FIG. 5. (Color online) Schematic drawing of the slowness curves S_a and S_b in three principal symmetry planes with (a) locations of transverse zero-curvature transonic states (short lines) and (b) locations of possible inflection (I_i) and parabolic (P_i) points. c_{ij} indicate relevant elastic constants which determine the slowness along the symmetry directions. The innermost slowness sheet S_c is omitted.

In the vicinity of [001] direction, $a_2(\theta)|_{v_a}$ related to the S_a branch, from Eqs. (9) and (10a), can be expressed by

$$a_2(\theta)|_{v_a}=f(\theta)\sin^2\theta. \tag{12}$$

We define $a_{2p}=f(\theta)$ by stripping off the term $\sin^2\theta$ from $a_2(\theta)|_{v_a}$ so that the condition $a_{2p}(0)|_{v_a}=0$ defines critical condition for a parabolic point reaching [001] direction. By setting $\theta=0$, we get

$$a_{2p}\propto-[d_{13}^4-d_{13}^2\Delta_{34}(c_{66}+2d_{12})+\Delta_{34}^2(d_{12}^2+c_{66}\Delta_{16})]/\Delta_{34}=0\leftrightarrow \mathrm{B}. \tag{13}$$

Along [100] direction $(\theta=\pi/2)$ with $\rho v_a^2=c_{44}$, we have

$$a_2\propto-c_{44}\Delta_{14}\Delta_{46}=0\leftrightarrow \mathrm{D}. \tag{14}$$

On the transonic state branch S_b, from Eqs. (9) and (10b), $a_2(\theta)|_{v_b}$ is simpler,

$$a_2(\theta)|_{v_b}=\{\Delta_{46}(c_{11}\Delta_{16}-d_{12}^2)\sin^2\theta+[\Delta_{34}(c_{11}\Delta_{16}-d_{12}^2)-d_{13}^2(2\Delta+c_{66})]\cos^2\theta\}\sin^2\theta. \tag{15}$$

The parabolic point along off-symmetry direction is then given by $a_2(\theta)|_{v_b}=0$, i.e.,

$$\tan^2\theta=\frac{d_{13}^2(2\Delta+c_{66})-\Delta_{34}(c_{11}\Delta_{16}-d_{12}^2)}{\Delta_{46}(c_{11}\Delta_{16}-d_{12}^2)}, \tag{16}$$

and it will approach [001] or [100] when $a_2|_{\theta\to 0,\pi/2}=0$, namely,

$$a_2\propto\Delta_{34}(c_{11}\Delta_{16}-d_{12}^2)-d_{13}^2(2\Delta+c_{66})=0\leftrightarrow \mathrm{C}, \tag{17}$$

and

$$a_2\propto\Delta_{46}(c_{11}\Delta_{16}-d_{12}^2)=0\leftrightarrow \mathrm{E}, \tag{18}$$

respectively.

Because S_b branch is elliptic (10b), direction for the surface normal at the parabolic point φ [e.g., the swallowtail point S_3 in Fig. 2(b)] can be derived directly from Eq. (16), namely,

$$\tan\varphi = \left(\frac{c_{66}}{c_{44}}\right)\tan\theta. \tag{19}$$

B. Inflection and parabolic points in $(\bar{1}10)$ plane

Being a symmetry plane equivalent to the (010) plane, $(\bar{1}10)$ plane can be treated straightforwardly. By rotating the coordinate system about [001] with $\pi/4$, all the results for the (010) plane can be directly converted to the case for the $(\bar{1}10)$ plane. Under the new coordinate system, the elastic tensor remains tetragonal and the transformed elastic constants c'_{IJ} are given by

$$\begin{aligned}
&c'_{11} = c_{11} - \Delta/2, \quad d'_{13} = d_{13},\\
&c'_{12} = c_{12} + \Delta/2, \quad d'_{12} = \Delta_{16},\\
&c'_{66} = c_{66} + \Delta/2, \quad \Delta'_{16} = d_{12},\\
&c'_{13} = c'_{23} = c_{13}, \quad \Delta'_{46} = \delta_{46},\\
&c'_{44} = c'_{55} = c_{44}, \quad \Delta' = -\Delta.
\end{aligned} \tag{20}$$

Since the analysis is exactly parallel to that for the (010) plane, we will just transform conditions A–E by replacing c_{IJ} with the transformed c'_{IJ} above and rewrite them in terms of c_{IJ} using Eq. (20).

On the transonic state branch S_a, the condition for the inflection point along the [001] direction, from Eq. (11), becomes

$$\Delta_{34}(c_{11} + c_{12} + 2c_{66}) - 2d_{13}^2 = 0 \leftrightarrow \mathrm{A}'. \tag{21}$$

The parabolic points on S_a will appear along [001] and [110], from Eqs. (13) and (14), when

$$\begin{aligned}
&-[2d_{13}^4 - d_{13}^2\Delta_{34}(4\Delta_{16} + c_{11} - c_{12}) + \Delta_{34}^2((c_{11} + c_{66})d_{12}\\
&\quad + 2\Delta_{16}^2)]/\Delta_{34} = 0 \leftrightarrow \mathrm{B}',\\
&-c_{44}\delta_{46}(d_{12} - \delta_{46}) = 0 \leftrightarrow \mathrm{D}',
\end{aligned} \tag{22}$$

respectively.

The parabolic point on S_b will reach [001] and [110], from Eqs. (17) and (18), when

$$\begin{aligned}
&d_{13}^2(3\Delta - 2c_{66}) + \Delta_{34}[d_{12}(2c_{11} - \Delta) - 2\Delta_{16}^2] = 0 \leftrightarrow \mathrm{C}',\\
&\delta_{46}[(2c_{11} - \Delta)d_{12} - 2\Delta_{16}^2] = 0 \leftrightarrow \mathrm{E}',
\end{aligned} \tag{23}$$

respectively. This parabolic point and the corresponding swallowtail point will be determined by transformed Eqs. (16) and (19).

C. Inflection and parabolic points in (001) plane

The zero-curvature transonic states with respect to the (001) plane can be examined by defining the reference plane $(\mathbf{m},\mathbf{n})$ with $\mathbf{m}$ lying in (001) plane and $\mathbf{n}$ normal to the plane, $\mathbf{n}=\mathbf{e}_z$ and $\mathbf{m}=\cos\theta\,\mathbf{e}_x+\sin\theta\mathbf{e}_y$. The matrix Γ becomes

$$\begin{aligned}
&\Gamma_{11} = p^2c_{44} - \rho v^2 + \lambda_{16}, \quad \Gamma_{12} = d_{12}\sin\theta\cos\theta,\\
&\Gamma_{22} = p^2c_{44} - \rho v^2 + \lambda_{61}, \quad \Gamma_{13} = pd_{13}\cos\theta,\\
&\Gamma_{33} = p^2c_{33} - \rho v^2 + c_{44}, \quad \Gamma_{23} = pd_{23}\sin\theta.
\end{aligned} \tag{24}$$

The characteristic equation $\|\Gamma\|=0$ remains a bicubic equation in p, and $a_0(v,\theta)=0$ yields the transonic state branches S_a and S_b as follows:

$$\begin{aligned}
\rho v_a^2 = \tfrac{1}{4}\{&2c_{11} + 2c_{66}\\
&- \sqrt{2[\Delta_{16}^2 + d_{12}^2 + (c_{11} + c_{12})\Delta\cos 4\theta]}\},
\end{aligned} \tag{25a}$$

$$\rho v_b^2 = c_{44}. \tag{25b}$$

Since there is no tangential acoustic axis in this plane, we can focus on the existence of E1 transonic states. On transonic state branch S_a, the coefficient $a_2(\theta)|_{v_a}$ is given by

$$\begin{aligned}
a_2(\theta)|_{v_a} = \tfrac{1}{4}\{&d_{13}^2(d_{12} + \Delta_{16}) + c_{44}(d_{12}^2 + \Delta_{16}^2) + (c_{11} - c_{12})\\
&\times(d_{13}^2 + c_{44}(d_{12} + \Delta_{16}))\cos 4\theta - [d_{13}^2 + c_{44}(c_{11}\\
&- 2c_{44} + c_{66})] \times [2\Delta_{16}^2 + 2d_{12}^2 + 2(c_{11}\\
&+ c_{12})\Delta\cos 4\theta]^{1/2}\}.
\end{aligned} \tag{26}$$

TABLE I. Summary of critical conditions for the inflection/parabolic points (I_i/P_i) on the slowness surface (S) and their corresponding cuspidal/swallowtail points (C_i/S_i) on the wave surface (W).

Condition			S	W
[001]		$\rho\hat{v}^2=c_{44}$		
	A	Inflection/(010)	I_1	C_1
	A′	Inflection/$(\bar{1}10)$	I_3	C_3
	B	Parabolic/(010)/S_a	P_1	S_1
	C	Parabolic/(010)/S_b	P_3	S_3
	B′	Parabolic/$(\bar{1}10)$/S_a	P_4	S_4
	C′	Parabolic/$(\bar{1}10)$/S_b	P_6	S_6
[100]		$\rho\hat{v}^2=c_{44}$		
	D	Parabolic/(010)/S_a	P_2	S_2
	G	Parabolic/(001)/S_b	P_8	S_8
	G	Inflection/(010)/S_a	I_2	C_2
[100]		$\rho\hat{v}^2=c_{66}$		
	F	Parabolic/(001)/S_a	P_7	S_7
	E	Parabolic/(010)/S_b	P_3	S_3
	E	Inflection/(001)/S_a	I_5	C_5
[110]		$\rho\hat{v}^2=c_{44}$		
	D′	Parabolic/$(\bar{1}10)$/S_a	P_5	S_5
	I	Parabolic/(001)/S_b	P_8	S_8
	I	Inflection/$(\bar{1}10)$/S_a	I_4	C_4
[110]		$\rho\hat{v}^2=(c_{11}-c_{12})/2$		
	H	Parabolic/(001)/S_a	P_7	S_7
	E′	Parabolic/$(\bar{1}10)$/S_b	P_6	S_6
	E′	Inflection/(001)/S_a	I_5	C_5

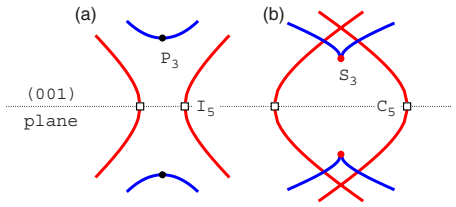


FIG. 6. (Color online) Correspondence between (a) the parabolic lines and (b) the caustic lines associated with the coexistence P_3-I_5 near [100] direction.

The condition $a_2(\theta)|_{v_a}=0$ will result in at most one parabolic point on S_a in the interval $\theta=[0,\pi/4]$, and it approaches [100] ($\theta=0$) and [110] ($\theta=\pi/4$) directions when the following conditions:

$$a_2 \propto c_{44}\Delta_{16}\Delta_{46}=0 \leftrightarrow \mathrm{F},$$

$$a_2 \propto c_{44}d_{12}\delta_{46}=0 \leftrightarrow \mathrm{H}, \tag{27}$$

are satisfied, respectively.

On transonic state branch S_b, $a_2(\theta)|_{v_b}$ becomes

$$a_2(\theta)|_{v_b}=\delta_{46}[d_{13}^2-c_{33}(d_{12}-\delta_{46})]+\Delta[2d_{13}^2-c_{33}(c_{11}+c_{12})]\cos^2 2\theta/4, \tag{28}$$

and $a_2(\theta)|_{v_b}=0$ defines a parabolic point along θ with

$$\tan 2\theta^2=\frac{\Delta_{46}(c_{33}\Delta_{14}-d_{13}^2)}{\delta_{46}[c_{33}(\delta_{46}-d_{12})+d_{13}^2]}, \tag{29}$$

consistent with Eq. (15) in Ref. 14. This parabolic point will appear along [100] ($\theta=0$) and [110] ($\theta=\pi/4$) when

$$a_2 \propto \Delta_{46}(d_{13}^2-c_{33}\Delta_{14})=0 \leftrightarrow \mathrm{G},$$

$$a_2 \propto \delta_{46}[d_{13}^2-c_{33}(d_{12}-\delta_{46})]=0 \leftrightarrow \mathrm{I}, \tag{30}$$

respectively.

Because the S_b branch is circular (25b), the direction for surface normal at this parabolic point φ [e.g., the swallowtail point S_8 shown in Fig. 8(b)] is identically given by Eq. (29).

IV. STRUCTURE OF PHONON FOCUSING PATTERNS

A. Critical existence conditions for caustic points

The 14 conditions above (A–I) constitute the critical existence conditions for the inflection/parabolic points in the principal symmetry directions [001], [100], and [110]. Note that the concave region on the slowness surface can also appear near the (001) plane, and the conditions for the parabolic points (E, E′, G, and I) can, in fact, be regarded as the conditions for the inflection points along the [100] and [110] directions as well. Consider condition E [see Fig. 5(a)]. It refers to the parabolic point along [100] with respect to the (010) plane. At the same time, it implies also an inflection point with respect to the (001) plane. The same holds for condition E′. The existence of such an inflection point in the (001) plane will then be determined by condition $\mathrm{E}\cdot\mathrm{E}'<0$. With the same token, condition G suggests an inflection point along [100] with respect to the (010) plane, and condition I implies another inflection point along [110] with respect to the $(\bar{1}10)$ plane.

In all, we can conclude that, in the symmetry planes $(010)/(\bar{1}10)$, there are up to two inflection points and two parabolic points on S_a and one parabolic point on S_b, and in the (001) plane there is at most one inflection point and one parabolic point on S_a, and one parabolic point on S_b. Figure 5(b) illustrates their topographical positions (marked with I_i and P_i), and their critical conditions are summarized in Table I. Since each inflection/parabolic (I_i/P_i) point produces a cuspidal/swallowtail (C_i/S_i) caustic point in the phonon focusing pattern, the conditions above can be regarded as the critical existence conditions for the cuspidal/swallowtail points.

The phonon focusing pattern along the [001] direction will be dominated by the cuspidal points (C_i) and swallowtail points (S_i) originated from the inflection/parabolic points I_1, P_1, I_3, P_4, P_3, and P_6, and it is comparable to those in cubic crystals. In the (001) plane, however, there is a new

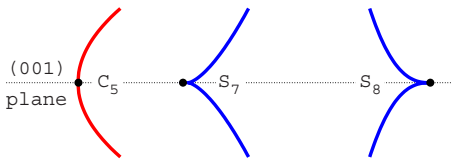


FIG. 7. (Color online) Main feature of caustic lines in the vicinity of the (001) plane. C_5 and S_7 belong to S_a branch, and S_8 belongs to S_b branch.

TABLE II. Locations of the inflection/parabolic points (I_i,P_i) on the slowness surface and the cuspidal/swallowtail points (C_i,S_i) in the focusing pattern for four crystals. The angles (in degrees) are with respect to the symmetry axes.

Crystal	I_i and P_i	C_i and S_i
	InBi	
(010)	$I_1(13.87), I_2(69.65)$	$C_1(34.83), C_2(64.21)$
	$P_1(24.56), P_2(71.98)$	$S_1(29.53), S_2(64.44)$
	$P_3(72.13)$	$S_3(68.13)$
(110)	$I_3(12.55), I_4(69.56)$	$C_3(31.99), C_4(71.40)$
	$P_6(61.75)$	$S_6(33.53)$
(001)	$I_5(17.00)$	$C_5(17.36)$
	CoF_2	
(010)	$P_2(63.93), P_3(43.74)$	$S_2(67.99), S_3(67.73)$
(110)	$I_3(31.57), I_4(56.18)$	$C_3(39.12), C_4(50.00)$
(001)	$I_5(5.57)$	$C_5(38.97)$
	$P_7(31.28), P_8(12.70)$	$S_7(26.98), S_8(12.70)$
	ZnS_2	
(010)	$P_3(43.44)$	$S_3(34.00)$
(110)	$P_6(39.27)$	$S_6(46.93)$
(001)	$I_5(31.43)$	$C_5(32.48)$
	$P_7(19.89), P_8(33.47)$	$S_7(39.82), S_8(33.47)$
	$AgGaS_2$	
(010)	$I_1(16.12), I_2(72.51)$	$C_1(17.63), C_2(75.75)$
(110)	$I_3(13.37), I_4(76.20)$	$C_3(10.21), C_4(85.88)$
	$P_6(64.03)$	$S_6(51.49)$
(001)	$I_5(16.02)$	$C_5(15.43)$
	$P_7(24.24), P_8(26.26)$	$S_7(11.67), S_8(26.26)$

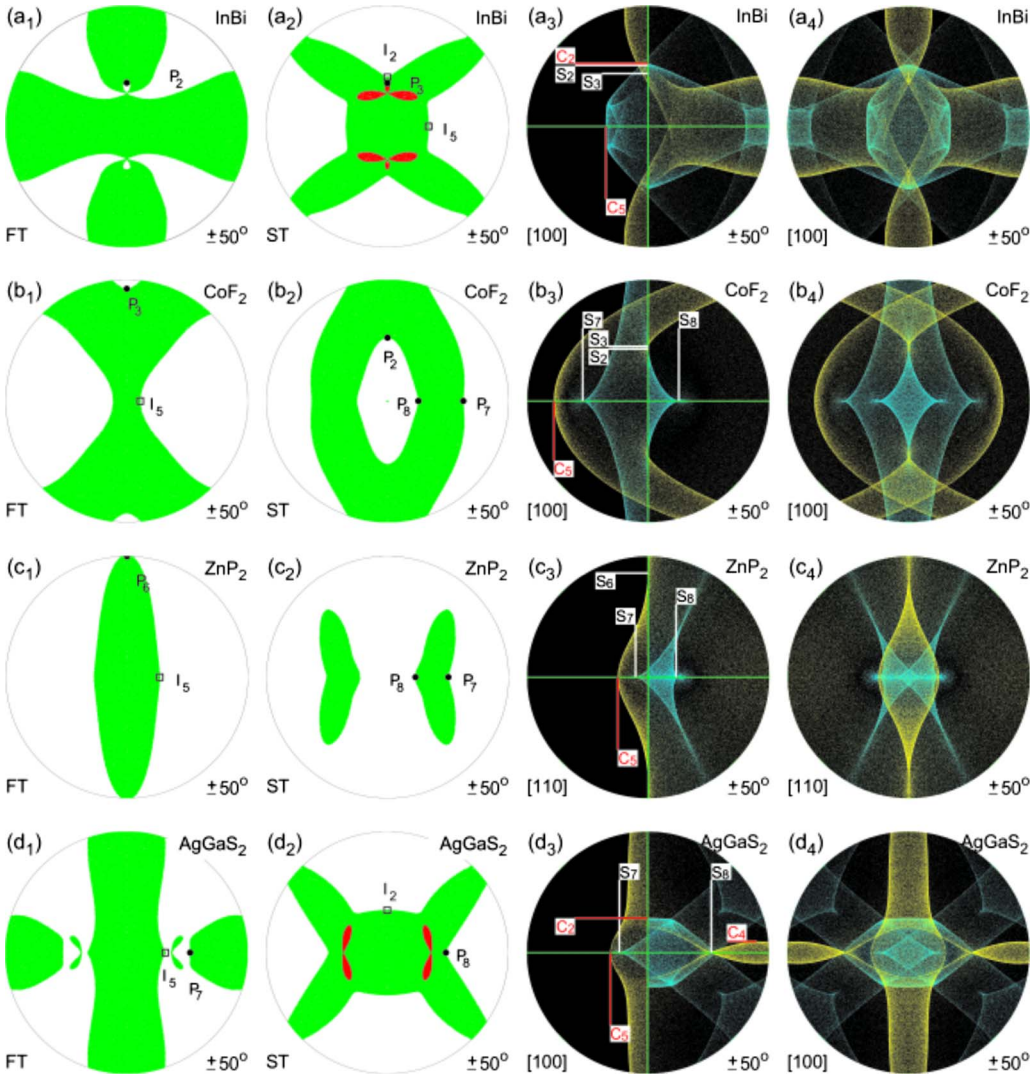


FIG. 8. (Color online) Parabolic lines and simulated phonon focusing pattern of (a) InBi, (b) CoF_2, (c) ZnP_2, and (d) $AgGaS_2$. First two columns: Gray and black regions depict saddle and concave regions on the slowness surface, respectively, and their boundaries represent parabolic lines. The boxes mark inflection points and the filled circles mark parabolic points.

feature—coexistence of the cuspidal and swallowtail points—due to the duality of conditions E, E′, G, and I. These conditions indicate that, whenever a parabolic point appears with respect to a symmetry plane, there will be an inflection point in the symmetry plane normal to it (see Table I). Therefore, the parabolic and inflection points can appear in pair, such as $(P_3 \leftrightarrow I_5)$, $(P_6 \leftrightarrow I_5)$, $(P_8 \leftrightarrow I_2)$, and $(P_8 \leftrightarrow I_4)$, as shown in Fig. 5(b). Such coexistence can be found in many tetragonal crystals. Figure 6 illustrates an example for the coexistence $(P_3 \leftrightarrow I_5)$ and the corresponding caustic points $(S_3 \leftrightarrow P_5)$.

One special situation, which is not covered in this analysis, is the case where a concavity appears along an oblique direction in the symmetry planes. Its critical existence condition refers to vanishing in-plane curvature along an off-symmetry direction, and it has to be addressed in relation to monoclinic crystals.

Analytically, the derivation of the swallowtail points (S_3, S_6, S_8) associated with the S_b branch, such as Eqs. (19) and (29), is simple because S_b is either elliptical or circular. Calculation of other caustic points associated with the S_a branch can be done by employing the analytical scheme established in Refs. 15 and 16.

B. Examples

Generally, the phonon focusing patterns along [001] mainly originated from two inflection points (I_1, I_3) and four parabolic points (P_1, P_3, P_4, P_6). By simulating focusing pattern originated from a half of the slowness surface [e.g., Figs. 2(b) and 2(c)], we can identify the characteristic points C_1, C_3, S_1, S_3, and S_6 systematically.

Compared to cubic crystals, a new feature of the focusing pattern in tetragonal crystals is the caustic points (C_5, S_7, S_8) in the (001) plane and the coexistence phenomena. The existence of C_5, S_7, and S_8 can be deduced by examining the signs of a_2 in Sec. III using *Regula Falsi* method:

$$C_5 \rightarrow \mathrm{E} \cdot \mathrm{E}' < 0,$$

$$S_7 \rightarrow \mathrm{F} \cdot \mathrm{H} < 0,$$

$$S_8 \rightarrow \mathrm{G} \cdot \mathrm{I} < 0. \quad (31)$$

These conditions suggest that the caustic points on the (001) plane may emerge independently with each other. The occurrence of the cuspidal point C_5 need not have the swallowtail point S_7 as its precursor, and the swallowtail point S_7 exists when there is an acoustic axis in the (001) plane (if $d_{12} > 0$). Figure 7 illustrates a possible case for the caustic lines associated with these three points.

In this investigation, we calculated parabolic lines and phonon focusing patterns for all tetragonal crystals listed in Ref. 21. We present here four examples: InBi, CoF_2, ZnS_2, and $AgGaS_2$. Table II lists numerical results using the analytical scheme outlined in Refs. 15 and 16. In order to demonstrate the main features around the (001) plane, the graphical results illustrated in Fig. 8 are centered along [100] or [110] direction and the caustic points C_5, S_7, and S_8 are delineated in detail by simulating the focusing pattern caused by a half of the slowness surface. The coexistence ($C_5 \leftrightarrow S_3/S_6$) and/or ($S_8 \leftrightarrow C_2/C_4$) are observed in all four examples.

Another new feature that one often observes in tetragonal crystals is two small spike-like caustic lines in the neighborhood of [100]/[110] directions, such as the cases in InBi and $AgGaS_2$ [Figs. 8(a_3) and 8(d_3)]. The origin for these lines is the presence of parabolic lines [see elongated loops in Figs. 8(a_2) and 8(d_2)] marking the boundary of concave regions. These parabolic lines anchored at the acoustic axes, and their extremities are associated with the inflection points, which result in the tips of two spike-like caustic lines in the oblique directions.

V. CONCLUDING REMARKS

The zero Gaussian curvature associated with the inflection/parabolic points on the slowness surface is addressed in terms of the zero-curvature transonic state pertaining to the principal symmetry planes. Because the zero-curvature transonic states are related to highly degenerated Stroh eigenvalues, we are able to derive explicitly the critical existence conditions for the inflection/parabolic points with respect to the symmetry planes.

The critical existence conditions provide an overview for the locations of the cuspidal/swallowtail points in the symmetry planes, and thus the main features of the phonon focusing patterns. It is also shown that the phonon focusing pattern near the (001) plane in tetragonal crystals differs from that in cubic crystals due to the absence of the tangential acoustics axis, and it is characterized by three new caustic lines, the coexistence phenomenon, and the spike-like caustic lines.

ACKNOWLEDGMENT

Comments from the reviewers are gratefully acknowledged.

[1] M. J. P. Musgrave, *Crystal Acoustics* (Holden-Day, San Francisco, CA, 1970).
[2] B. A. Auld, *Acoustic Waves and Fields in Solids* (Wiley, New York, 1973), Vol. **1**.
[3] J. P. Wolfe, *Imaging Phonons* (Cambridge University Press, Cambridge, 1998).
[4] A. G. Every, "Ballistic phonons and the shape of the ray surface in cubic crystals," Phys. Rev. B **24**, 3456–3467 (1981).
[5] G. A. Northrop and J. P. Wolfe, "Ballistic phonon imaging in germanium," Phys. Rev. B **22**, 6196–6212 (1980).
[6] D. C. Hurley and J. P. Wolfe, "Phonon focusing in cubic crystals," Phys. Rev. B **32**, 2568–2587 (1985).
[7] R. L. Weaver, M. R. Hauser, and J. P. Wolfe, "Acoustic flux imaging in anisotropic media," Z. Phys. B: Condens. Matter **90**, 27–46 (1993).
[8] C. G. Winternheimer and A. K. McCurdy, "Phonon focusing and phonon conduction in orthorhombic and tetragonal crystals in the boundary-scattering regime," Phys. Rev. B **18**, 6576–6605 (1978).
[9] D. C. Hurley, J. P. Wolfe, and K. A. McCarthy, "Phonon focusing in tellurium dioxide," Phys. Rev. B **33**, 4189–4195 (1986).
[10] A. G. Every and A. K. McCurdy, "Phonon focusing in piezoelectric crystals," Phys. Rev. B **36**, 1432–1447 (1987).
[11] A. G. Every, "Classification of the phonon-focusing patterns of tetragonal crystals," Phys. Rev. B **37**, 9964–9977 (1988).
[12] A. L. Shuvalov and A. G. Every, "Curvature of acoustic slowness surface of anisotropic solids near symmetry axes," Phys. Rev. B **53**, 14906–14916 (1996).
[13] A. L. Shuvalov and A. G. Every, "Evolution of phonon-focusing caustics under linear spatial dispersion," J. Phys. A **33**, 5105–5116 (2000).
[14] A. L. Shuvalov and A. G. Every, "Transverse curvature of the acoustic slowness surface in crystal symmetry planes and associated phonon focusing cusps," J. Acoust. Soc. Am. **108**, 2107–2113 (2000).
[15] L. Wang, "Caustic and anticaustic points in the phonon focusing patterns of cubic crystals," J. Acoust. Soc. Am. **123**, 4140–4146 (2008).
[16] L. Wang, "Analytical determination of parabolic points on slowness surface and swallowtail points on wave surface of cubic crystals," Proc. R. Soc. London, Ser. A **465**, 1231–1247 (2009).
[17] D. M. Barnett, J. Lothe, and S. A. Gundersen, "Zero curvature transonic states and free surface waves in anisotropic elastic media," Wave Motion **12**, 341–360 (1990).
[18] A. N. Stroh, "Steady state problems in anisotropic elasticity," J. Math. Phys. **41**, 77–103 (1962).
[19] P. Chadwick and G. D. Smith, "Foundations of the theory of surface waves in anisotropic elastic materials," *Advances in Applied Mechanics* (Academic, New York, 1977), Vol. **17**, pp. 303–376.
[20] T. C. T. Ting, *Anisotropic Elasticity: Theory and Applications* (Oxford University Press, Oxford, 1996).
[21] R. F. S. Hearmon, *Numerical Data and Functional Relationships in Science and Technology*, Landolt-Börnstein Vols. **11** and **18**, edited by K. H. Hellwege (Springer-Verlag, New York, 1979).

Mindlin plate theory for damage detection: Imaging of flexural inhomogeneities

L. R. Francis Rose[a)]
Air Vehicles Division, Defence Science and Technology Organisation, Melbourne, Victoria 3207, Bundoora, Vicotria 3083, Australia

Chun H. Wang
Sir Lawrence Wackett Aerospace Centre, School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University, Australia

(Received 2 October 2009; revised 24 November 2009; accepted 1 December 2009)

The scattering of plate waves by localized damage or defects that can be modeled as flexural inhomogeneities is examined within the framework of Mindlin plate theory. These inhomogeneities are characterized by variations in one or more of the four plate-theory parameters: the bending stiffness, shear stiffness, rotary inertia, and transverse inertia. It is shown that the Born approximation for the scattered field leads to a plate-theory analog of the Fourier diffraction theorem, which relates the far-field scattering amplitude to the spatial Fourier transform of the inhomogeneity variations. The application of this result is illustrated by using synthetic data derived for an idealized model of a delamination as a flexural inhomogeneity, ignoring mode coupling effects. A computationally efficient implementation of the filtered back-propagation algorithm, based on the eigensystem of the scattering operator, is employed for image reconstruction. The implications for *in-situ* imaging of structural damage in plate-like structures are briefly discussed, and some directions for further work are indicated.
 [DOI: 10.1121/1.3277217]

PACS number(s): 43.35.Zc, 43.20.Gp, 43.35.Bf, 43.40.Dx [TDM] Pages: 754–763

I. INTRODUCTION

This paper continues and completes a theoretical analysis of wave scattering by flexural inhomogeneities within the framework of Mindlin plate theory.[1] The objective is to provide a tractable approach for *in-situ* imaging of certain types of structural damage that can be adequately modeled as flexural inhomogeneities (i.e., regions that are characterized by variations in one or more of the four plate-theory parameters: the bending stiffness, shear stiffness, rotary inertia, and transverse inertia). Hinders and co-workers[2–6] attempted to adapt conventional forms of tomographic imaging for this purpose, but concluded that a more promising approach should be based on diffraction tomography. This paper presents the requisite theoretical framework for diffraction tomographic imaging of flexural inhomogeneities.

It should be emphasized from the outset that the concept of a flexural inhomogeneity is a theoretical idealization. In practice, scattering from realistic types of structural damage will entail a coupling involving all three families of plate-wave modes.[7] An instructive example of this coupling has recently been studied by Grahn,[8] who considered the scattering of an incident symmetric S_0 wave by a step-change in thickness over a circular region. He used an exact representation for the scattered field as a sum over the three families of wave modes, suitably truncated for the purposes of deriving numerical solutions, and he also showed that the results obtained by using the lowest-order plate theories are in good agreement with the exact results for sufficiently low frequencies. It seems reasonable to expect that using higher-order plate theories would lead to good agreement over a wider frequency range, approaching the cut-off frequency for the second antisymmetric mode (A_1), while still retaining analytical simplicity. The present work is restricted to the idealized case of flexural inhomogeneities, which means that there is no coupling to symmetric plate-wave modes. However, this idealization represents a valuable first step before attempting a more realistic analysis using coupled plate theories. An important contribution to this topic has been made by Norris and Vemula,[9] who have derived the exact solution for scattering by a circular inhomogeneity, as well as the optical theorem, within the framework of classical (Kirchhoff) and Mindlin[10] plate theories.

The present work is organized as follows. The configuration that is envisaged is shown in Fig. 1. It consists of a plate containing a region of damage, indicated by the shaded area, which is assumed to be adequately modeled as a flexural inhomogeneity. The plate is equipped with actuators and sensors that can generate and detect flexural waves. The choice of appropriate parameters for characterizing a flexural inhomogeneity is discussed in Sec. II, and a representation for the scattered wave due to an inhomogeneity is derived in Sec. III. A simplified representation based on a linearized approximation is presented next in Sec. IV for weak inhomogeneities, which is the case of greatest interest for nondestructive evaluation and structural health monitoring. It is shown that this approximation leads to a relationship be-

[a)]Author to whom correspondence should be addressed. Electronic mail: francis.rose@dsto.defence.gov.au

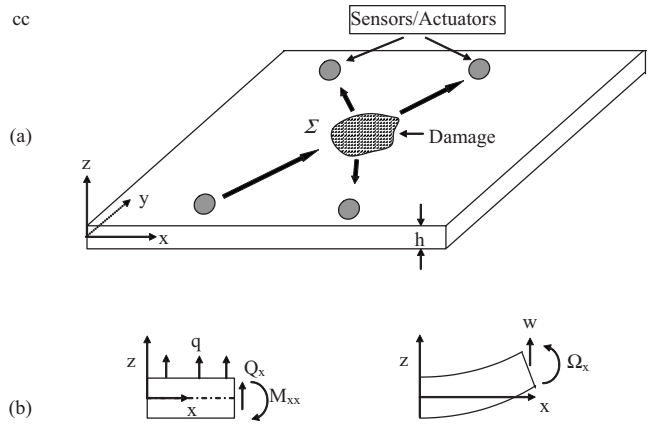


FIG. 1. Coordinate axes and sign conventions for Mindlin plate theory. (a) A plate with built-in actuators and sensors for detection of delamination damage. (b) Sign conventions for positive bending moment M_{xx}, shear force Q_x, rotation Ω_x, and deflection w.

tween the scattered amplitude and the two dimensional (2D) Fourier transform of the inhomogeneity variation that can be used as a basis for diffraction tomography. A procedure for imaging a flexural inhomogeneity on that basis is presented in Sec. V, adapting a recent approach developed for the acoustic case.[11,12] That approach relies on determining the eigensystem of an operator related to the far-field scattering operator,[13] instead of the more conventional approach based on inverting a 2D Fourier transform,[14] which has been discussed elsewhere.[15,16] The procedure is illustrated here by using synthetic data for the scattered field obtained from the solution of Vemula and Norris.[10] A discussion of the present results and some general considerations for *in-situ* imaging of structural damage in plates are presented in Sec. VI.

II. CHARACTERIZING A FLEXURAL INHOMOGENEITY

In an earlier work,[1] hereafter referred to as Part I, the equations of Mindlin plate theory were derived in a form that was designed to highlight their similarity with the equations of three dimensional (3D) elasticity.[17–22] These governing equations are recalled here for later reference and to discuss the characterization of a flexural inhomogeneity within the framework of Mindlin plate theory.

There are two sets of equations. First, the *equations of motion* are given by

$$M_{\beta\alpha,\beta} - Q_\alpha + \rho I \ddot{\Omega}_\alpha = m_\alpha, \tag{1a}$$

$$Q_{\alpha,\alpha} - \rho h \ddot{w} = -q, \tag{1b}$$

where $M_{\alpha\beta}$ denotes the bending moments, Q_α the components of the shear force, ρ the mass density, h the plate thickness, I the moment of inertia $(=h^3/12)$, Ω_α the components of the rotation vector, and w the deflection. The terms on the right hand sides of these equations denote distributed bending moments m_α and lateral pressure q that are assumed to be prescribed quantities. The sign conventions for the forces and displacements are shown in Fig. 1(b), and the usual conventions for derivatives and for summation over repeated subscripts apply.

Second, the *constitutive equations*, relating the bending moments and the shear force to the rotation vector Ω_α and the deflection w, are given by

$$M_{\alpha\beta} = -D\Gamma_{\alpha\beta} = -D\left\{\nu\Omega_{\alpha,\alpha}\delta_{\alpha\beta} + \frac{1-\nu}{2}(\Omega_{\alpha,\beta} + \Omega_{\beta,\alpha})\right\}, \tag{2a}$$

$$Q_\alpha = \bar{\mu}h(w_{,\alpha} - \Omega_\alpha), \tag{2b}$$

where D denotes the bending stiffness $[D=EI/(1-\nu)^2$, where E and ν denote Young's modulus and Poisson ratio, respectively], and $\bar{\mu}h$ the shear stiffness.

It can be seen from these equations that there are four basic parameters in Mindlin plate theory: two stiffnesses (the bending D and shear $\bar{\mu}h$ stiffnesses) and two inertia parameters (the rotary ρI and transverse ρh inertias). A flexural inhomogeneity is therefore characterized by a variation in one or more of those four parameters. It will be convenient to represent these variations in the following form, where a subscript zero is used to identify parameters pertaining to the host plate outside the region of inhomogeneity Σ, and x denotes an arbitrary field point,

$$D(x) = D_0[1 + \delta_1(x)],$$

$$\bar{\mu}h(x) = \bar{\mu}h_0[1 + \delta_2(x)],$$

$$\rho I(x) = \rho I_0[1 + \delta_3(x)],$$

$$\rho h(x) = \rho h_0[1 + \delta_4(x)]. \tag{3}$$

Note that ρI_0 is used to denote the value of the plate-theory parameter ρI for the host plate, and similarly for $\bar{\mu}h_0$ and ρh_0. Thus, a flexural inhomogeneity is in general characterized by a set of four functions $\delta_n(x)$ that are nonzero only within the region of inhomogeneity, i.e., Σ in Fig. 1(a). Reconstructing these four functions from wave-scattering measurements would be a considerably more difficult undertaking than that envisaged in conventional tomography, which provides a reconstruction for only one scalar unknown. That unknown is interpreted as a variation in the wave speed or, equivalently, in the refractive index.[14] For an acoustic medium, there are, in fact, two material properties that affect the wave speed, viz., the compressibility and the density. The conventional assumption for the application of diffraction tomography is to ignore variations in density and therefore to attribute the scattered field purely to variations in compressibility. There have been strategies proposed for imaging the variations in both compressibility and density by using scattered field data for two distinct frequencies,[23–25] but those strategies have not yet been implemented in practice. Similarly for the plate-wave case, it would be possible to image separately the four functions $\delta_n(x)$ by using scattered field data for four distinct frequencies, but it will prove more instructive to focus instead on two special cases for which there is a relationship between these four functions, reducing them effectively to a single unknown function.

Accordingly, we note first that the effect of a delamination can be modeled approximately as a variation in the moment of inertia I, which affects both the bending stiffness D

and the rotary inertia ρI, but leaves the shear stiffness and transverse inertia unchanged, as discussed in previous work.[26,27] Thus, for that case, one can set

$$\delta_1 = \delta_3 = \delta_d, \quad \delta_2 = \delta_4 = 0. \tag{4}$$

Second, a variation δh in the plate thickness h affects all four functions δ_n. If it is assumed that this variation is symmetrical with respect to the plate's mid-plane, so as to avoid a coupling with the extensional modes, and that $\delta h/h \ll 1$, then one can set

$$\delta_1 = \delta_3 = 3\delta_h, \quad \delta_2 = \delta_4 = \delta_h. \tag{5}$$

Thus, for either of those two cases, the inhomogeneity can be characterized by only one unknown function, denoted by δ_d for *delaminations* and by δ_h for *thickness variations*.

Strictly speaking, there are five parameters in Mindlin plate theory: the (in-plane) Poisson ratio ν could also vary within Σ. However, for simplicity, this variation will be ignored in the present work. The following analysis could be modified to accommodate variations in the Poisson ratio should this prove necessary in practice.

III. A REPRESENTATION FOR THE SCATTERED FIELD

Equations (1) and (2) are differential equations that hold locally at every point, i.e., they apply equally for an inhomogeneous plate, provided that the plate-theory parameters take on their local values, as given in Eq. (3). Substituting Eq. (3) into Eqs. (1) and (2) and re-assigning all terms involving the inhomogeneity functions δ_n to the right hand side, one can recast the equations of motion for an inhomogeneous plate as follows:

$$D_0\left\{\nabla\nabla\cdot\boldsymbol{\Omega} - \frac{1-\nu}{2}\nabla\times\nabla\times\boldsymbol{\Omega}\right\} + \bar{\mu}h_0(\nabla w - \boldsymbol{\Omega}) - \rho I_0\ddot{\boldsymbol{\Omega}} = \boldsymbol{m} + \boldsymbol{m}^{\Delta}, \tag{6a}$$

$$\bar{\mu}h_0\,\nabla\cdot(\nabla w - \boldsymbol{\Omega}) - \rho h_0\ddot{w} = -(q + q^{\Delta}), \tag{6b}$$

$$m_{\alpha}{}^{\Delta} = -\{(D_0\delta_1\Gamma_{\beta\alpha})_{\beta} + \bar{\mu}h_0\delta_2(w_{,\alpha} - \Omega_{\alpha}) - \rho I_0\delta_3\ddot{\Omega}_{\alpha}\}, \tag{6c}$$

$$q^{\Delta} = -\{\nabla\cdot[\bar{\mu}h_0\delta_2(\nabla w - \boldsymbol{\Omega})] - \rho h_0\delta_4\ddot{w}\}. \tag{6d}$$

In these equations, $\boldsymbol{m}$ and q are assumed to be prescribed bending moments and lateral force, whereas the effect of the inhomogeneity is modeled by distributed moments $\boldsymbol{m}^{\Delta}$ and lateral force q^{Δ} that depend on the inhomogeneity functions $\delta_n(x)$, as shown in Eqs. (6c) and (6d). This provides the basis for an integral representation of the scattered field in terms of the Green's function introduced in Part I. An analogous representation would be obtained in 3D elasticity. To highlight this analogy, we recall the notation introduced in Part I, whereby $\boldsymbol{u}$ is used to denote a vector of plate-theory displacements, and $\boldsymbol{f}$ is a vector of plate-theory body force, with the following understanding:

$$\boldsymbol{u} \equiv u_i = (-\Omega_x, -\Omega_y, w), \tag{7a}$$

$$\boldsymbol{f} \equiv f_i = (m_x, m_y, q). \tag{7b}$$

Consider now the total displacement field $\boldsymbol{u}$ due to an incident field $\boldsymbol{u}^I$, which is assumed to be a prescribed field. Because Eqs. (6a)–(6d) are linear, $\boldsymbol{u}$ can be expressed as a superposition of two fields,

$$\boldsymbol{u} = \boldsymbol{u}^I + \boldsymbol{u}^S, \tag{8a}$$

where $\boldsymbol{u}^S$ designates the scattered field. Equations (6a)–(6d) can now be written in the operator form

$$\boldsymbol{L}_0[\boldsymbol{u}] = \boldsymbol{L}_0[\boldsymbol{u}^I + \boldsymbol{u}^S] = -(\boldsymbol{f} + \boldsymbol{f}^{\Delta}), \tag{8b}$$

where $\boldsymbol{L}_0[]$ denotes the differential operator on the left hand sides of Eqs. (6a) and (6b) pertaining to a *homogeneous plate*. The incident field $\boldsymbol{u}^I$ is the field that would prevail in the absence of an inhomogeneity, i.e., if the inhomogeneity variations $\delta_n(x)$ were all set equal to zero. Thus,

$$\boldsymbol{L}_0[\boldsymbol{u}^I] = -\boldsymbol{f}, \tag{8c}$$

where $\boldsymbol{f}$ denotes the prescribed moments $\boldsymbol{m}$ and force q in accordance with the notation in Eq. (7b), and

$$\boldsymbol{L}_0[\boldsymbol{u}^S] = -\boldsymbol{f}^{\Delta}, \tag{8d}$$

with $\boldsymbol{f}^{\Delta}$ given by Eqs. (6c) and (6d). (In Secs. IV and V, the incident field will be taken to be a plane wave, i.e., the sources of the incident field are thereby assumed to lie outside the domain of interest, so that $\boldsymbol{f}=\boldsymbol{0}$.) The scattered field can therefore be represented as follows:

$$u_i^S(\boldsymbol{x},t) = \int\int_{\Sigma} f_j^{\Delta}(\boldsymbol{\xi},\tau)\,g_{ij}(\boldsymbol{x}-\boldsymbol{\xi},t-\tau)\,d\xi\,d\eta\,d\tau, \tag{9}$$

where g_{ij} denotes the dynamic Green's function for a homogeneous plate, giving the ith component of displacement due to the jth component of the applied force, in accordance with the notation in Eqs. (7a) and (7b), $\boldsymbol{x}$ denotes an arbitrary field point, $\boldsymbol{\xi}$ denotes an arbitrary point (ξ,η) within the source region Σ, and t denotes the time variable.

Analytical expressions for the Green's function have been derived in Part I for the time harmonic case, using the notation

$$\hat{g}_{ij}(\boldsymbol{x}-\boldsymbol{\xi};\omega) = \int_{-\infty}^{\infty} g_{ij}(\boldsymbol{x}-\boldsymbol{\xi};t)\,e^{i\omega t}dt. \tag{10a}$$

Accordingly, Eq. (9) can be written as follows for the time harmonic case:

$$\hat{u}_i^S(\boldsymbol{x},\omega) = \int\int_{\Sigma} \hat{f}_j^{\Delta}(\boldsymbol{\xi})\,\hat{g}_{ij}(\boldsymbol{x}-\boldsymbol{\xi})\,d\xi\,d\eta. \tag{10b}$$

Equation (10b) constitutes an integral representation for the scattered field $\boldsymbol{u}^S$ in terms of the induced source density $\boldsymbol{f}^{\Delta}$ given by Eqs. (6c) and (6d) and the infinite plate Green's function $\hat{g}_{ij}$ derived in Part I.

It can be seen from Eqs. (6c) and (6d) that this representation involves the spatial derivatives of two of the inhomogeneity variations, viz., δ_1 and δ_2. A more convenient representation can be obtained through an integration by parts, using Gauss's theorem, provided that δ_1 and δ_2 are

assumed to have continuous first derivatives (which implies, in particular, that $\delta_1=\delta_2=0$ on the boundary $\partial\Sigma$ of the region of inhomogeneity Σ). This leads to

$$\hat{u}_i^S(\boldsymbol{x},\omega)=\int\int_\Sigma \{D_0\delta_1\hat{\Gamma}_{\beta\alpha}\hat{g}_{i\alpha,\beta}+[\bar{\mu}h_0\delta_2(\hat{w}_{,\alpha}-\hat{\Omega}_\alpha) + \omega^2\rho I_0\delta_3\hat{\Omega}_\alpha]\hat{g}_{i\alpha}+\bar{\mu}h_0\delta_2(\hat{w}_{,\alpha}-\hat{\Omega}_\alpha)\hat{g}_{i3,\alpha} + \omega^2\rho h_0\delta_4\hat{g}_{i3}\}d\xi d\eta, \tag{11}$$

where $\hat{\Gamma}_{\alpha\beta}$ is given by Eq. (2a).

In Eq. (11), terms such as $\hat{\Omega}_\alpha$ and $\hat{w}_\alpha$ on the right hand side refer to the *total displacement* $\hat{\boldsymbol{u}}=\hat{\boldsymbol{u}}^I+\hat{\boldsymbol{u}}^S$. However, $\hat{\boldsymbol{u}}^S$ is unknown, so that Eq. (11) is, in fact, an integral equation for $\hat{\boldsymbol{u}}^S$, rather than an explicit representation. An explicit representation can be obtained by using the Born approximation,[28] as discussed in Sec. IV.

Before proceeding, there are two points worth noting. First, the integration by parts leading to Eq. (11) requires $\delta_1(x)$ and $\delta_2(x)$ to vanish on the boundary $\partial\Sigma$. By contrast, the exact solutions derived by Vemula and Norris[10] and Grahn[8] involve inhomogeneity variations $\delta_n(x)$ that are constant within a circular region, but exhibit a step discontinuity at the boundary. For such cases, the first and third terms on the right hand side of Eq. (11) should be replaced by

$$D_0\delta_1\int_{\partial\Sigma}\hat{\Gamma}_{\beta\alpha}\hat{g}_{i\alpha}n_\beta ds,\quad \bar{\mu}h_0\delta_2\int_{\partial\Sigma}(\hat{w}_{,\alpha}-\hat{\Omega}_\alpha)\hat{g}_{i3}n_\alpha ds. \tag{12}$$

In the most general case, of course, $\delta_1(x)$ and $\delta_2(x)$ could be non-uniform within Σ as well as having a step discontinuity across the boundary $\partial\Sigma$, in which case the additional terms such as in Eq. (12) should be added to Eq. (11).

Second, it has been implicitly assumed so far that both the host plate and the inhomogeneity are non-dissipative. However, a region of structural damage can be expected to entail some dissipative mechanisms. These can conveniently be modeled with the present formalism by allowing the stiffness variations to be complex-valued functions depending on the frequency ω, i.e., by modeling the damaged region as a linearly viscoelastic inhomogeneity characterized by complex-valued stiffness variations $D_0\hat{\delta}_1(\boldsymbol{x},\omega)$ and $\bar{\mu}h_0\hat{\delta}_2(\boldsymbol{x},\omega)$.

IV. BORN APPROXIMATION

The Born approximation, which is appropriate for *weak inhomogeneities*, consists in approximating the total displacement $\hat{\boldsymbol{u}}=\hat{\boldsymbol{u}}^I+\hat{\boldsymbol{u}}^S$ by $\hat{\boldsymbol{u}}\approx\hat{\boldsymbol{u}}^I$ on the right hand side of Eq. (11), leading to the following approximate representation for $\hat{\boldsymbol{u}}^S$:

$$\hat{\boldsymbol{u}}^S(x,\omega)\approx\hat{\boldsymbol{u}}^B(x,\omega), \tag{13a}$$

$$\hat{u}_i^B(x,\omega)=\int\int_\Sigma \{D_0\delta_1\hat{\Gamma}^I_{\beta\alpha}\hat{g}_{i\alpha,\beta}+[\bar{\mu}h_0\delta_2(\hat{w}^I_{,\alpha}-\hat{\Omega}^I_\alpha) + \omega^2\rho I_0\delta_3\hat{\Omega}^I_\alpha]\hat{g}_{i\alpha}+\bar{\mu}h_0\delta_2(\hat{w}^I_{,\alpha}-\hat{\Omega}^I_\alpha)\hat{g}_{i3,\alpha} + \omega^2\rho h_0\delta_4\hat{g}_{i3}\}d\xi d\eta. \tag{13b}$$

This approximation leads to a relation between the far-field scattered amplitude and the 2D Fourier transforms of the inhomogeneity variations $\delta_n(x)$. To establish that relation, let $\hat{\boldsymbol{u}}^I$ be a plane wave specified by

$$\hat{w}^I(x,\omega)=e^{i\boldsymbol{k}_1{}^I\cdot\boldsymbol{x}}, \tag{14a}$$

$$\hat{\boldsymbol{\Omega}}^I=i\gamma_1\boldsymbol{k}_1\hat{w}^I, \tag{14b}$$

$$\gamma_1=1-\rho\omega^2/\bar{\mu}h, \tag{14c}$$

where the subscript 1 on the wave vector $\boldsymbol{k}_1{}^I$ indicates that the incident wave corresponds to the first of the three possible wave modes for Mindlin plate theory (cf. Part I, Sec. IIC). The scattered field $\hat{\boldsymbol{u}}^S$ will, in general, contain contributions from all three modes. However, for the purposes of imaging inhomogeneities, it will be assumed hereafter that the frequency is restricted to be below the cut-off frequency ω_c for the second and third modes [cf. Part I, Eq. (15f)] so that the first mode is the only propagating mode. Accordingly, it will be sufficient to retain only the terms corresponding to that first mode in the far-field expansions for the Green's function g_{ij} and its derivatives appearing in Eq. (11). Using this simplification and the far-field approximation for the Hankel function appearing in g_{ij} (cf. Part I, Sec. IVC), one can verify that the far-field scattered wave behaves like a diverging cylindrical wave:

$$\hat{w}^S(r,\theta;\omega)=-\frac{i}{4}\sqrt{\frac{2}{\pi k_1 r}}e^{i(k_1r-\pi/4)}A(\theta,\alpha)+O(1/r),\quad k_1r\gg 1, \tag{15a}$$

$$\hat{\Omega}^S_r(\boldsymbol{x},\omega)=ik_1\gamma_1\hat{w}^S(\boldsymbol{x},\omega), \tag{15b}$$

$$\hat{\Omega}^S_\theta(\boldsymbol{x},\omega)=O(1/r), \tag{15c}$$

$$\nabla\cdot\hat{\boldsymbol{\Omega}}^S(\boldsymbol{x},\omega)=-k_1^2\gamma_1\hat{w}^S(\boldsymbol{x},\omega). \tag{15d}$$

Equation (15a) serves to define the *plane-wave scattering amplitude* $A(\theta,\alpha)$ for the scattered wave in the direction θ due to an incident wave in the direction α, i.e.,

$$A(\theta,\alpha)=\lim_{k_1r\to\infty}\left\{\hat{w}^S(r,\theta;\omega)\Big/\left[-\frac{i}{4}\sqrt{\frac{2}{\pi k_1 r}}e^{i(k_1r-\pi/4)}\right]\right\}. \tag{15e}$$

By inserting the Born approximation (13b) for the scattered field into Eq. (15e), one can now derive the following expression for the scattering amplitude:

$$A(\theta,\alpha) = k_1^2 \sum_{n=1}^{4} q_n(\theta-\alpha)\Delta_n(\boldsymbol{k}_1^S - \boldsymbol{k}_1^I), \tag{16a}$$

$$q_1(\theta) = \frac{\gamma_1 k_1^2}{(k_1^2 - k_2^2)}[\cos^2\theta + \nu\sin^2\theta], \tag{16b}$$

$$q_2(\theta) = \frac{\bar{\mu}h_0(1-\gamma_1)^2}{\gamma_1 D_0(k_1^2 - k_2^2)}\cos\theta, \tag{16c}$$

$$q_3(\theta) = -\frac{\gamma_1 \rho I_0 \omega^2}{D_0(k_1^2 - k_2^2)}\cos\theta, \tag{16d}$$

$$q_4(\theta) = -\frac{\rho h_0 \omega^2}{\gamma_1 D_0 k_1^2(k_1^2 - k_2^2)}, \tag{16e}$$

where $\boldsymbol{k}_1{}^I = k_1\boldsymbol{\alpha} = k_1(\cos\alpha, \sin\alpha)$ and $\boldsymbol{k}_1{}^S = k_1\boldsymbol{\theta} = k_1(\cos\theta, \sin\theta)$ denote the incident and scattered wave vectors, and $\Delta_n(\boldsymbol{k})$ denotes the 2D spatial Fourier transform of the inhomogeneity variation $\delta_n(x)$,

$$\Delta_n(\boldsymbol{k}) = \int\int_{\Sigma} \delta_n(x,y)e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}dxdy. \tag{16f}$$

Equation (16a) is the plate-theory analog of the Fourier diffraction theorem,[14] which provides the basis for imaging in diffraction tomography. In general, there are four independent variations $\delta_n(x)$ to be reconstructed, which would require observations of the scattered field at four separate frequencies (all below the cut-off frequency ω_c). However, as noted earlier in Sec. II, we shall focus here on two special cases for which there is a known relation between the inhomogeneity variations so that only one unknown function needs to be reconstructed from the scattered field data.

Consider first the relation specified by Eq. (4), which can serve as an approximate model for delaminations. Then, Eqs. (16a)–(16d) reduce to

$$A(\theta,\alpha) = k_1^2 q_d(\theta-\alpha)\Delta_d(\boldsymbol{k}_1^S - \boldsymbol{k}_1^I), \tag{17a}$$

$$q_d(\theta) = \frac{\gamma_1 k_1^2}{k_1^2 - k_2^2}\left[\cos^2\theta + \nu\sin^2\theta - \frac{\rho I_0 \omega^2}{D_0 k_1^2}\cos\theta\right]. \tag{17b}$$

Equation (17b) gives the scattering pattern due to a small variation in the moment of inertia I (and hence in the bending stiffness D as well as the rotary inertia ρI) over an infinitesimal area. This pattern is significantly more complicated than the corresponding patterns for small parameter perturbations in elasticity[29] or acoustics.[30] Recalling the expression given in Part I, Eq. (15f) for ω_c, and using the value $\pi^2/12$ for the shear coefficient (cf. Part I, Sec. VID), Eq. (17b) can be recast as follows:

$$q_d(\theta) = \frac{\gamma_1 k_1^2}{k_1^2 - k_2^2}\left[\cos^2\theta + \nu\sin^2\theta - \frac{1-\nu}{2}\left(\frac{\pi}{k_1 h}\frac{\omega}{\omega_c}\right)^2\cos\theta\right]. \tag{17c}$$

This shows that the scattering pattern depends on two nondimensional parameters $k_1 h$ and ω/ω_c and on the Poisson ratio ν. However, the third term in Eq. (17c) is generally much smaller than unity so that the scattering pattern is relatively insensitive to the values of $k_1 h$ and ω/ω_c. An example is shown in Fig. 2 for $\omega/\omega_c = 0.5$, $k_1 h = 2.3$, and $\nu = 0.3$.

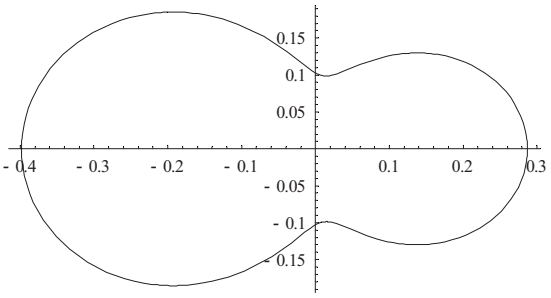


FIG. 2. Shape function q_d for a model material representative of a quasi-isotropic composite laminate at $\omega = 0.5\omega_c$ and $k_1 h = 2.3$; the Poisson ratio $\nu = 0.3$.

Nevertheless, because the pattern is known, it can be factored from the measurements of the scattering amplitude to provide the 2D Fourier transform of the parameter perturbation δ_d in accordance with Eq. (17a), as will be further discussed in Sec. V.

Second, for the relation specified by Eq. (5), corresponding to a thickness variation resulting from, for example, corrosion damage, one finds

$$A(\theta,\alpha) = k_1^2 q_h(\theta-\alpha)\Delta_h(\boldsymbol{k}_1^S - \boldsymbol{k}_1^I), \tag{18a}$$

$$q_h(\theta) = \{3\gamma_1 k_1^2 D_0(\cos^2\theta + \nu\sin^2\theta) + (\bar{\mu}h_0(1-\gamma_1)^2/\gamma_1 - 3\gamma_1\rho I_0\omega^2)\cos\theta - \rho h_0\omega^2/\gamma_1 k_1^2\}/D_0(k_1^2 - k_2^2). \tag{18b}$$

V. IMAGING FLEXURAL INHOMOGENEITIES

To discuss the procedure for imaging flexural inhomogeneities, consider again the configuration in Fig. 1. For simplicity, it will be assumed that the sources and receivers are sufficiently far removed from the region of inhomogeneity that the incident wave can be considered to be a plane wave, and the scattered wave can be considered to have reached the far field, thereby providing directly values of the plane-wave scattering amplitude $A(\theta,\alpha)$ in accordance with the definition in Eq. (15e). However, it is noted that it is possible to derive an explicit analytical relation between the cylindrical-wave scattering amplitude for point-like sources and receivers, and the plane-wave scattering amplitude,[25,31] so that the following analysis is readily adapted to the case of point-like source/receivers. Furthermore, it will also be assumed that the field variable that is measured is the plate deflection w. It is recalled from the discussion in Part I, Secs. III and IV, that circular piezoelectric sensors will usually measure instead the combination $\nabla\cdot\boldsymbol{\Omega}$, which corresponds to the Gaussian curvature of the plate. However, there is a simple linear relation between w and $\nabla\cdot\boldsymbol{\Omega}$, given by Eq. (15d), so that there is no loss of generality in dealing only with w in the following.

To proceed, consider the case where the inhomogeneity is a delamination that is modeled in accordance with Eq. (4),

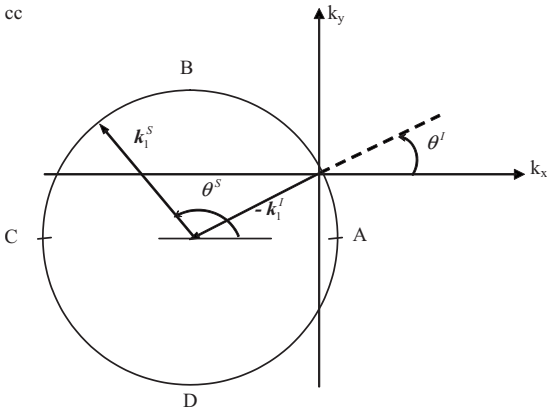


FIG. 3. The point in the transform-parameter plane corresponding to the scattering amplitude in the direction θ^S due to an incident wave in the direction θ^I. For a fixed θ^I, data for various scattering angles θ^S provide the values of the Fourier transform of the inhomogeneity variation δ_d over a circle ABCD in $\boldsymbol{k}$-space.

thereby leading to the relation given by Eq. (17) between the measured values of $A(\theta,\alpha)$ and the 2D Fourier transform of the inhomogeneity variation $\delta_d(x)$. For a fixed angle of incidence α, the values of $A(\theta,\alpha)$ corresponding to various scattering angles θ provide values of the Fourier transform $\Delta_d(\boldsymbol{k})$ on a circle in $\boldsymbol{k}$-space with center $-\boldsymbol{k}_1^I$ and radius k_1, as indicated in Fig. 3. Thus, by varying α, one can collect data that lead to values $\Delta_d(\boldsymbol{k})$ throughout a circular region of radius $2k_1$ in the transform-parameter plane, but not outside that region. The Fourier inversion formula applied to these data now leads to

$$\tilde{\delta}_d(x) = \frac{1}{(2\pi)^2}\int\int_{|\boldsymbol{k}|\le 2k_1} \Delta_d(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}} dk_x dk_y, \tag{19}$$

where the tilde on $\tilde{\delta}_d(x)$ serves to indicate that the function obtained from this Fourier inversion corresponds to a low-pass filtered reconstruction of the inhomogeneity variation $\delta_d(x)$.

Equation (19) constitutes the image reconstruction formula of diffraction tomography,[32,14–16] which has hereby been shown to apply to plate-wave imaging within the framework of Mindlin plate theory. By changing the variables of integration from $k_x = k_1(\cos\theta - \cos\alpha), k_y = k_1(\sin\theta - \sin\alpha)$ to α,θ, one can derive the following plate-theory analog of the filtered back-propagation formula:[33]

$$\tilde{\delta}_d(x) = \frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{2\pi} \frac{A(\theta,\alpha)}{q_d(\theta-\alpha)} \times |\sin(\theta-\alpha)| e^{ik_1(\boldsymbol{\theta}-\boldsymbol{\alpha})\cdot\boldsymbol{x}} d\theta d\alpha. \tag{20}$$

Computational aspects of the inversion formulas (19) and (20) have been discussed by Kak and Slaney.[14]

More recently, Mast *et al.*[11] derived an alternative formulation of diffraction tomography, as a by-product of a more general solution of the non-linear inverse scattering problem. With a subsequent refinement,[12,34] this formulation provides a computationally efficient approach for imaging. Accordingly, we shall now show how their formulation can be adapted for plate-wave imaging.

First, define an integral operator B, whose kernel is suggested by the integrand in Eq. (20), as follows:

$$B\cdot f(\theta) = \int_0^{2\pi} \frac{A(\theta,\alpha)}{q_d(\theta-\alpha)} |\sin(\theta-\alpha)| f(\alpha) d\alpha. \tag{21a}$$

This operator maps functions $f(\alpha)$ defined on the interval $(0,2\pi)$ into functions $g(\theta)$ defined on that same interval. By analogy with the acoustic case,[11,12] B can be shown to be a normal operator. Let λ_n denote the eigenvalues of B, and $f_n(\theta)$ the associated eigenfunctions, i.e.,

$$B\cdot f_n(\theta) = \lambda_n f_n(\theta). \tag{21b}$$

These eigenfunctions will be assumed to be normalized so that

$$\langle f_m, f_n\rangle = \int_0^{2\pi} f_m(\theta)\overline{f_n}(\theta) d\theta = \delta_{mn}, \tag{21c}$$

using standard notation for the inner product, and with $\overline{f_n}$ denoting the complex conjugate of f_n and δ_{mn} denoting the Kronecker delta. Next, let $E_n(\boldsymbol{x})$ denote a superposition of incident plane waves defined by

$$E_n(\boldsymbol{x}) = \int_0^{2\pi} f_n(\alpha) e^{ik_1\boldsymbol{\alpha}\cdot\boldsymbol{x}} d\alpha. \tag{22a}$$

Then

$$\overline{E_n(\boldsymbol{x})} = \langle e^{-ik_1\boldsymbol{\alpha}\cdot\boldsymbol{x}}, f_n\rangle. \tag{22b}$$

By virtue of the orthogonality of the eigenfunctions f_n, Eq. (22b) leads to the following representation for a plane wave:

$$e^{-ik_1\boldsymbol{\alpha}\cdot\boldsymbol{x}} = \sum\nolimits_n f_n(\alpha)\overline{E_n(\boldsymbol{x})}. \tag{22c}$$

Substituting Eq. (22c) into Eq. (20) and using Eqs. (21b) and (22a), one can now derive the final result,

$$\tilde{\delta}_d(x,y) = \frac{1}{8\pi^2}\sum\nolimits_n \lambda_n |E_n(x,y)|^2. \tag{23}$$

In a practical implementation of this approach, the scattering amplitude $A(\theta,\alpha)$ will only be available for a finite number of incident and scattering angles. Consequently, the operator B defined by Eq. (21a) reduces effectively to a matrix operator, the integral in Eq. (21a) being replaced by a matrix multiplication. If it is assumed that there are N angles of incidence α_j and an equal number of scattering angles θ_i, the B operator reduces to an $N\times N$ matrix whose elements are given by

$$B_{ij} = B(\theta_i,\alpha_j) = A(\theta_i,\alpha_j)|\sin(\theta_i-\alpha_j)|/q_d(\theta_i-\alpha_j). \tag{24}$$

Thus, the infinite sum in Eq. (23) is thereby reduced to a finite sum, with λ_n and f_n denoting the eigenvalues and eigenvectors of B_{ij}, and $E_n(x,y)$ denoting the basis functions associated with f_n through a discretized form of Eq. (22a).

A. Resolution and sampling

Before proceeding to numerical examples, we note some general considerations for imaging by diffraction tomography. First, the resolution is limited to approximately half the

wavelength λ, as is characteristic for diffraction-limited imaging. This can be seen from the point spread function, which is the reconstructed image for a point scatterer, i.e., for an inhomogeneity variation described by a Dirac delta function, which corresponds to setting $\Delta_d=1$ in Eq. (19). Evaluating the integral then leads to the following expression for the point spread function:

$$\tilde{\delta}_{\mathrm{PSF}}(x,y)=\frac{k_1}{\pi r}J_1(2k_1r). \tag{25}$$

This function peaks at $r=0$ and falls to zero for $r=0.3\lambda$. Thus, two closely spaced point scatterers can be considered to be adequately resolved if they are separated by twice that distance, i.e., the limit of resolution according to this criterion is 0.6λ. In the present context, the resolution is thereby limited to be approximately equal to the plate thickness h because Mindlin plate theory can only be expected to be appropriate for wavelengths greater than approximately twice the plate thickness.

In practice, this theoretical resolution can only be achieved with adequate sampling and with a sufficiently high signal-to-noise ratio. As a guide for adequate sampling, consider the case of a circular inhomogeneity of radius a, and a source/receiver array with the active elements situated at N equally spaced angular locations, so that the angular spacing between elements is given by $\Delta\theta=2\pi/N$. Then, the Nyquist criterion for an acceptable spatial sampling rate leads to[12]

$$a\cdot\Delta\theta\leq\lambda/2, \tag{26a}$$

$$N\geq 2k_1a. \tag{26b}$$

This criterion will be used in the numerical examples below.

As a practical guide for the validity of the Born approximation used in Sec. IV, the phase difference that is accrued on traversing an inhomogeneity, relative to propagation in a homogeneous medium, should be less than $\pi/2$.[14] This criterion can be expressed as

$$k_1an_d\leq 1, \tag{27a}$$

where n_d denotes the change in refractive index within the inhomogeneity

$$n_d=|k_{1d}-k_1|/k_1. \tag{27b}$$

This variation in refractive index is related to the inhomogeneity variation δ_d, but not in a simple manner, because δ_d affects both the stiffness D and the inertia ρI, as can be seen from Eqs. (3) and (4). However, numerical calculations confirm that this criterion is satisfied in the following numerical examples.

B. Numerical examples

To illustrate the imaging procedure, consider first the case of a circular delamination modeled in accordance with Eqs. (3) and (4). The theoretical solution for the scattered field[10] will be used as synthetic data for the purposes of implementing the image reconstruction formula (23).

Specifically, the plate is assumed to be a quasi-isotropic composite laminate (i.e., one that can be adequately modeled as being isotropic) of thickness $h=2.0$ mm, Young's modulus $E=50$ GPa, density $\rho=1600$ kg/m^3, and Poisson ratio $\nu=0.3$. With this choice of parameters, the cut-off frequency is equal to 866.7 kHz. The frequency of the incident and scattered waves is taken to be 173 kHz so that $\omega/\omega_c=0.2$ and the wavelength $\lambda=10.0$ mm. The delamination is assumed to have a radius $a=15$ mm and to be centered at the origin, and the inhomogeneity variation is specified to be $\delta_d(\boldsymbol{x})=-0.2$. These choices lead to $k_1a=9.4$, and hence the minimum sampling rate for the scattering amplitude, according to Eq. (26b), is $N\geq 19$. In the following examples, we have used $N=32$, unless otherwise specified. The imaging domain is chosen to be -60 mm$<(x,y)<60$ mm.

The first step is to calculate the values $A_{ij}=A(\theta_i,\alpha_j)$ for N equi-spaced view angles θ_i and incident angles α_j. This constitutes the raw data. The matrix elements B_{ij} are then obtained from Eq. (24), and hence, by determining the eigensystem λ_n and f_n, and the associated basis functions $E_n(x,y)$, a reconstructed image can be computed in accordance with Eq. (23). However, as noted by Lin *et al.*,[12] the term $|\sin(\theta_i-\alpha_j)|$ has a large bandwidth and therefore requires a large sampling rate for adequate sampling accuracy. Accordingly, a substantial improvement in image quality can be obtained as follows. The $N\times N$ matrix of data values A_{ij} is used in conjunction with a suitable interpolation scheme to generate in effect a continuous function $A'(\theta,\alpha)$. This interpolated function is now re-sampled at a higher rate to produce a $4N\times 4N$ matrix A'_{ij}, and those values are used in Eq. (24) to obtain a $4N\times 4N$ matrix B_{ij}. The eigensystem and associated basis functions of this enlarged B matrix are then used for the image reconstruction in accordance with Eq. (23). The resulting image is shown as a 3D plot of $\tilde{\delta}_d(x,y)$ in Fig. 4(b) and as a density plot in Fig. 4(c). This reconstructed image is the spatial convolution of the exact object function [which is shown, for reference, in Fig. 4(a)] and the point spread function given by Eq. (25). It can be seen that sharp edges become blurred, and the image exhibits characteristic modulations corresponding to decaying oscillations of the point spread function for $r>0.3\lambda$. Figure 4(d) shows a profile of the reconstructed image along the x-axis, i.e., $\tilde{\delta}_d(x,y=0)$.

Figures 5(a) and 5(b) show the reconstructed image for the same inhomogeneity, but at two different frequencies, $\omega/\omega_c=0.1$ and 0.5. For the higher frequency, a higher sampling rate is required so as to comply with the Nyquist criterion (26b), and a value $N=64$ was used for that case. It can be seen from those examples that an excellent image quality can be achieved within the limitations of diffraction-limited imaging.

Finally, consider next the case where the center of the inhomogeneity is displaced from the origin to a location (x_0,y_0), as indicated in Fig. 6(a), so that the scattering pattern is no longer radially symmetric, i.e., $A(\theta,\alpha)$ is not simply a function of $(\theta-\alpha)$. In the present case, $x_0=y_0=10$. The scattered field can be determined by a translation of the solution for a centered circular inhomogeneity derived by Vemula and Norris.[10] The reconstructed image for that case is shown in Figs. 6(b)–6(d). It can be seen that the image quality is significantly degraded relative to Figs. 4(b)–4(d). This high-

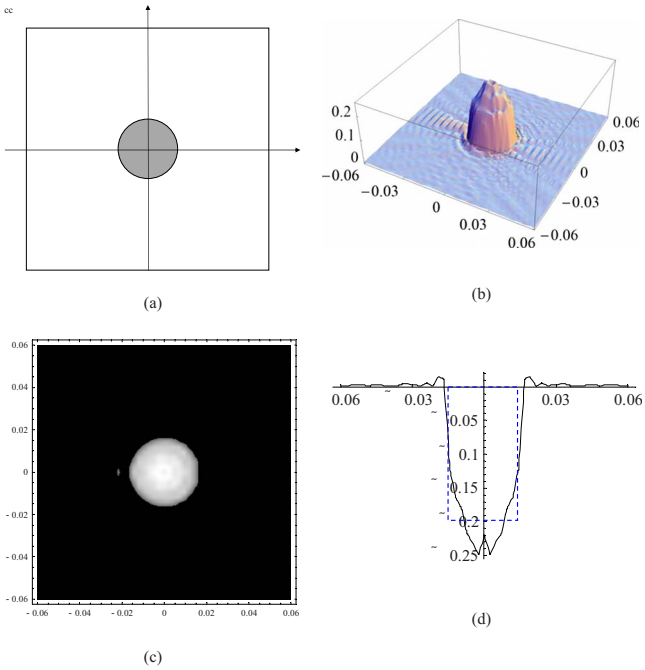


FIG. 4. (Color online) Reconstructed image of a delamination damage located at the center of the imaging domain. (a) The configuration used for generating the synthetic data. (b) 3D plot of the reconstructed inhomogeneity variation $|\tilde{\delta}_d(x,y)|$. (c) Density plot of $|\tilde{\delta}_d(x,y)|$. (d) Profile plot of $\tilde{\delta}_d(x,y)$ along $y=0$; the dashed line shows the actual variation.

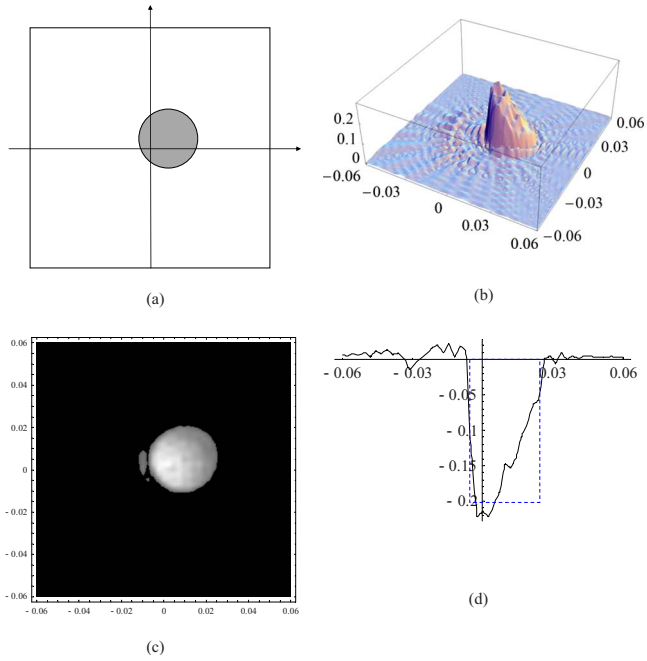


FIG. 6. (Color online) Reconstructed image of a delamination damage located off the center of the imaging domain. (a) The configuration used for generating the synthetic data. (b) 3D plot of the reconstructed inhomogeneity variation $|\tilde{\delta}_d(x,y)|$. (c) Density plot of $|\tilde{\delta}_d(x,y)|$. (d) Profile plot of $\tilde{\delta}_d(x,y)$ along $y=0$; the dashed line shows the actual variation.

lights a limitation of the present imaging algorithm relative to a numerical implementation of the inverse Fourier transform in Eq. (19), as discussed by Rohde *et al.*[15,16] However, the advantage of the present approach from the viewpoint of computational efficiency can be retained by using a two-pass approach to imaging, in which the objective in the first pass is to obtain an approximate location for the center (or, more generally, the centroid) of the inhomogeneity, and translating the origin of coordinates for the imaging domain to the location of that centroid. For the first pass, the location of the inhomogeneity can be determined using alternative algorithms, such as the synthetic time reversal method developed by Wang *et al.*[35] The two-pass process should ensure an imaging quality similar to that shown in Fig. 4 for the centered circular inhomogeneity.

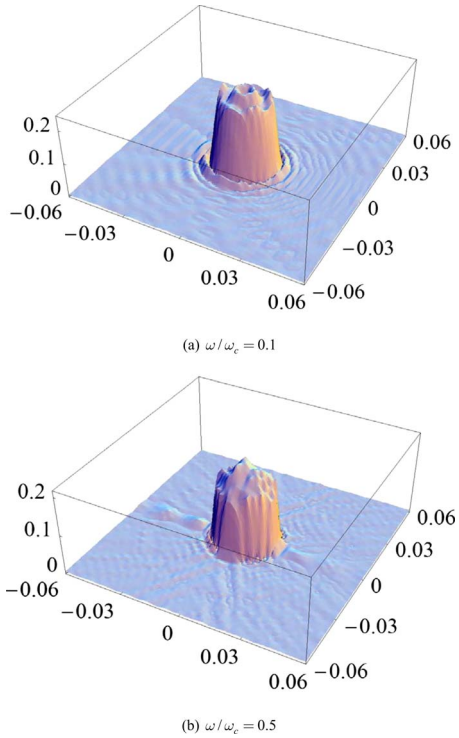


FIG. 5. (Color online) Influence of incident wave frequency on imaging quality at two different frequencies: (a) $\omega/\omega_c=0.1$ and (b) $\omega/\omega_c=0.5$.

VI. DISCUSSION

The motivation for the present work has been to explore theoretically the possibility of *in-situ* imaging of certain forms of structural damage in plate-like structures. This constitutes a significantly more ambitious objective than simply detecting and locating the damage. It has been shown that this objective can be achieved subject to certain simplifications and idealizations. The most important of these is that the mode coupling that would be induced by realistic forms of damage is ignored. In that case, the interaction between an incident wave and the damaged region can be adequately captured within the framework of a plate theory by modeling the damage as an inhomogeneity. In particular, the present work has focused on forms of laminar damage that can be modeled as flexural inhomogeneities within the framework of Mindlin plate theory. A further idealization has been to characterize the inhomogeneity in terms of a single unknown function, as given in Eqs. (4) and (5), for the case of delaminations and of thickness variations. It was then shown that a plate-theory analog of diffraction tomography can be derived, and the imaging procedure was illustrated in Sec. V.

Once this connection with diffraction tomography has been established, further work could proceed in two different directions. First, one can examine more closely the practical implementation of diffraction tomography, accepting the various idealizations mentioned above. From that perspective, we note, in particular, the recent work of Devaney and Dennison[36,37] which provides an alternative formulation of the linearized inverse scattering problem that appears to be more versatile in dealing with arbitrary source-receiver arrays. This formulation can be readily adapted to the plate-theory case, as briefly discussed by Rohde *et al.*[38] This new approach appears to be well suited for assessing the trade off between resolution and noise tolerance. This is an important practical consideration that has been ignored in the present work, but that warrants further study. It can be anticipated that an increased number of active elements in a source-receiver array will lead to increased resolution, but this will also entail an increased sensitivity to noise. Thus, the resolution that can be achieved in practice is likely to be noise limited, rather than diffraction limited.

Second, one can attempt to extend the present formulation to include mode conversion between extensional and flexural plate-wave modes,[26] using either the classical (Kirchhoff) plate theory or the more refined Mindlin–Kane theory for the extensional modes. A solution for scattering by a circular inhomogeneity within Mindlin–Kane theory has recently been derived.[27] It would seem worthwhile to combine this solution with that for the flexural case to study mode coupling. This can be expected to provide a significant improvement relative to the solution using only the lowest-order plate theories for the extensional and flexural waves.[8,39]

Finally, it is recalled that diffraction tomography relies on using the Born approximation for the scattered field, which is only appropriate for weak inhomogeneities. However, that is the case of greatest interest from the viewpoint of non-destructive evaluation and structural health monitoring. It is also worth noting that despite a considerable research effort on non-linear inverse scattering over the past several decades, no practical alternative to imaging based on a linearized inverse has emerged that offers consistently better imaging quality.[40]

VII. CONCLUSION

A plate-theory analog of imaging by diffraction tomography has been presented within the framework of Mindlin plate theory. This fulfills the original objective from Part I of developing a theoretical framework for imaging certain forms of structural damage in plate-like structures. The approach relies on several simplifications and idealizations, most notably ignoring mode coupling effects. However, the approach can be extended to include such effects, and this would appear to be a worthwhile extension to pursue. The eigensystem of the scattering operator is shown to offer a computationally efficient approach to plate-wave diffraction tomography. This imaging algorithm enables high resolution of flexural damage to be determined using a two-pass process in which the first pass aims to provide an approximate location for the centroid of the damaged region.

ACKNOWLEDGMENTS

The authors are grateful to Associate Professor Martin Veidt and Dr. Andrew H. Rohde for several stimulating discussions during the course of the present work.

[1]L. R. F. Rose and C. H. Wang, "Mindlin plate theory for damage detection: Source solutions," J. Acoust. Soc. Am. **116**, 154–171 (2004).
[2]J. C. P. McKeon and M. K. Hinders, "Parallel projection and crosshole Lamb wave contact scanning tomography," J. Acoust. Soc. Am. **106**, 2568–2577 (1999).
[3]E. V. Malyarenko and M. K. Hinders, "Fan beam and double crosshole Lamb wave tomography for mapping flaws in aging aircraft structures," J. Acoust. Soc. Am. **108**, 1631–1639 (2000).
[4]E. V. Malyarenko and M. K. Hinders, "Ultrasonic Lamb wave diffraction tomography," Ultrasonics **39**, 269–281 (2001).
[5]K. R. Leonard, E. V. Malyarenko, and M. K. Hinders, "Ultrasonic Lamb wave tomography," Inverse Probl. **18**, 1795–1808 (2002).
[6]K. R. Leonard and M. K. Hinders, "Multi-mode Lamb wave tomography with arrival time sorting," J. Acoust. Soc. Am. **117**, 2028–2038 (2005).
[7]M. Lemistre and D. Balageas, "Structural health monitoring system based on diffracted Lamb wave analysis by multiresolution processing," Smart Mater. Struct. **10**, 504–511 (2001).
[8]T. Grahn, "Lamb wave scattering from a circular partly through-thickness hole in a plate," Wave Motion **37**, 63–80 (2003).
[9]A. N. Norris and C. Vemula, "Scattering of flexural waves on thin plates," J. Sound Vib. **181**, 115–125 (1995).
[10]C. Vemula and A. N. Norris, "Flexural wave propagation and scattering on thin plates using Mindlin theory," Wave Motion **26**, 1–12 (1997).
[11]T. D. Mast, A. I. Nachman, and R. C. Waag, "Focusing and imaging using eigenfunctions of the scattering operator," J. Acoust. Soc. Am. **102**, 715–725 (1997).
[12]F. Lin, A. I. Nachman, and R. C. Waag, "Quantitative imaging using a time-domain eigenfunctions method," J. Acoust. Soc. Am. **108**, 899–912 (2000).
[13]A. N. Norris, "A direct inverse scattering method for imaging obstacles with unknown surface conditions," IMA J. Appl. Math. **61**, 267–290 (1998).
[14]A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (IEEE, New York, 1988), Chap. 6, free download at http://www.slaney.org/pct/pct-toc.html (Last viewed 11/17/09).
[15]A. H. Rohde, "Lamb wave diffraction tomography for imaging material inhomogeneities in isotropic plates," Ph.D. thesis, University of Queensland, Brisbane, Australia (2007).
[16]A. H. Rohde, M. Veidt, L. R. F. Rose, and J. Homer, "A computer simulation study of imaging flexural inhomogeneities using plate-wave diffraction tomography," Ultrasonics **48**, 6–15 (2008).
[17]J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, Amsterdam, 1973).
[18]K. Aki and P. G. Richards, *Quantitative Seismology* (Freeman, San Francisco, CA, 1980), Vol. **1**.
[19]R. D. Mindlin and J. Yang, *An Introduction to the Mathematical Theory of Vibrations of Elastic Plates* (World Scientific, Singapore, 2007).
[20]*The Collected Papers of Raymond D. Mindlin*, edited by H. Deresiewicz, M. P. Bieniek, and F. L. DiMaggio (Springer-Verlag, New York, 1989).
[21]J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge University Press, New York, 1999).
[22]J. L. Rose, "A baseline and vision of ultrasonic guided wave inspection potential," J. Pressure Vessel Technol. **124**, 273–282 (2002).
[23]S. J. Norton, "Generation of separate density and compressibility images in tissue," Ultrason. Imaging **5**, 240–252 (1983).
[24]A. J. Devaney, "Variable density acoustics tomography," J. Acoust. Soc. Am. **78**, 120–130 (1985).
[25]A. Witten, J. Tuggle, and R. C. Waag, "A practical approach to ultrasonic imaging using diffraction tomography," J. Acoust. Soc. Am. **83**, 1645–1652 (1988).
[26]C. H. Wang and L. R. F. Rose, "Wave reflection and transmission in beams containing delamination and inhomogeneity," J. Sound Vib. **264**, 851–872 (2003).
[27]C. H. Wang and F. K. Chang, "Scattering of plate waves by a cylindrical

inhomogeneity," J. Sound Vib. **282**, 429–451 (2005).
[28]P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1968), Chap. 8.
[29]R. Wu and K. Aki, "Scattering characteristics of elastic waves by an elastic heterogeneity," Geophysics **50**, 582–595 (1985).
[30]J. P. Lefebvre, "A linearised inverse problem: Acoustical impedance tomography of biological media," in *Electromagnetic and Acoustic Scattering: Detection and Inverse Problem*, edited by C. Bourrely, P. Chiappetta, and B. Torresani (World Scientific, Singapore, 1988).
[31]A. J. Devaney and G. Beylkin, "Diffraction tomography using arbitrary transmitter and receiver surfaces," Ultrason. Imaging **6**, 181–193 (1984).
[32]R. K. Mueller, M. Kaveh, and G. Wade, "Reconstructive tomography and applications to ultrasonics," Proc. IEEE **67**, 567–587 (1979).
[33]A. J. Devaney, "A filtered backpropagation algorithm for diffraction tomography," Ultrason. Imaging **4**, 336–350 (1982).
[34]T. D. Mast, A. I. Nachman, D. L. Liu, and R. C. Waag, "Quantitative imaging with eigenfunctions of the scattering operator," in 1997 IEEE International Ultrasonics Symposium Proceedings, Toronto, Canada (1997), IEEE Catalog No. 97CH36118-2, pp. 1507–1510.
[35]C. H. Wang, J. T. Rose, and F.-K. Chang, "A synthetic time-reversal imaging method for structural health monitoring," Smart Mater. Struct. **13**, 415–423 (2004).
[36]A. J. Devaney and M. Dennison, "Inverse scattering in inhomogeneous background media," Inverse Probl. **19**, 855–870 (2003).
[37]M. L. Dennison and A. J. Devaney, "Inverse scattering in inhomogeneous background media: II. Multi-frequency case and SVD formulation," Inverse Probl. **20**, 1307–1324 (2004).
[38]A. H. Rohde, L. R. F. Rose, M. Veidt, and C. H. Wang, "Two inversion strategies for plate wave diffraction tomography," Mater. Forum **33**, 489–495 (2009).
[39]M. J. S. Lowe, P. Cawley, J.-Y. Kao, and O. Diligent, "The low frequency reflection characteristics of the fundamental antisymmetric Lamb wave a_0 from a rectangular notch in a plate," J. Acoust. Soc. Am. **112**, 2612–2622 (2002).
[40]M. Brandfass, A. D. Lanterman, and K. F. Warnick, "A comparison of the Colton-Kirsch inverse scattering methods with linearised tomographic inverse scattering," Inverse Probl. **17**, 1797–1816 (2001).

Analytical method for the ultrasonic characterization of homogeneous rigid porous materials from transmitted and reflected coefficients

J.-P. Groby[a)]
Département de Recherche en Électromagnétisme-Laboratoire de Signaux et Systèmes, UMR8506 CNRS/Supélec/Univ. Paris-Sud, F-91192 Gif-sur-Yvette, France

E. Ogam
Laboratoire de Mécanique et d'Acoustique, UPR7051 CNRS, F-13402 Marseille Cedex 20, France

L. De Ryck,[b)] N. Sebaa, and W. Lauriks
Laboratory of Acoustics and Thermal Physics, K.U.Leuven, B-3001 Heverlee, Belgium

(Received 8 December 2008; revised 9 December 2009; accepted 11 December 2009)

A frequency domain method dedicated to the analytic recovery of the four relevant parameters of macroscopically homogeneous rigid frame porous materials, e.g., plastic foams, at the high frequency range of the Johnson–Champoux–Allard model is developed and presented. The reconstructions appeal to experimental data concerning time domain measurements of the ultrasonic fields reflected and transmitted by a plate of the material at normal incidence. The effective density and bulk modulus of the material are first reconstructed from the frequency domain reflection and transmission coefficients. From the latter, the porosity, tortuosity, and thermal and viscous characteristic lengths are recovered. In a sense, the method presented herein is quite similar in the ultrasonic range, but also quite complementary, to the method developed by Panneton and Olny [J. Acoust. Soc. Am. **119**, 2027–2040 (2006); **123**, 814–824 (2008)] at low frequency, which appeal to experimental data measured in an impedance tube.
 [DOI: 10.1121/1.3283043]

PACS number(s): 43.35.Zc, 43.20.Jr, 43.20.Ye, 43.40.Sk [KVH] Pages: 764–772

I. INTRODUCTION

A rigid frame porous material is a porous material whose frame is immobile. This assumption is not only conditioned by the saturating fluid usually a light fluid such as air, but also by the frequency of the acoustic solicitation. The frequency band suitable for the rigid frame approximation is bounded at high frequency by the diffusion limit, when the wavelength is of the order of, or smaller than the pore size, and at low frequency by the fluid-solid decoupling frequency below which the skeleton may vibrate. These bounds depend on the material properties, and so on its characteristics.

The equations of motion in a rigid frame porous material, derived from Biot's theory[1,2] and in later publications,[3–8] reduce to those of an equivalent fluid, with complex frequency-dependent effective density and bulk modulus. These later publications present semiphenomenological models that describe the viscous and thermal dissipations over specific frequency ranges, assuming known the characteristics of the saturating fluid, i.e., the viscosity η, the saturating pressure P_0, the specific heat ratio γ, and density ρ_f. Concerning the viscous dissipation, the Johnson *et al.* model[3] was made to fit the exact high- and low- (imaginary) frequency limits of the effective density, while the Wilson[7] model was contrived to match the mid frequency range of the latter. Concerning the thermal dissipation, the Champoux–Allard model[4] was developed to fit the exact high frequency limits of the effective bulk modulus, while the Wilson[7] model was contrived to match the mid frequency range of the latter. Inspired by the fact that the low frequency development of the Champoux–Allard model is not exact, Lafarge *et al.*[6] proposed an alternative expression of the effective bulk modulus and introduce the static thermal permeability. However, this parameter seems to be difficult to determine, even at low frequency.[6,9,10]

Herein, the Johnson–Champoux–Allard model[3,4] (JCAM) is considered, because high frequency measurements are performed. The parameters involved in this model are the porosity ϕ, which is the ratio of the fluid volume to the total sample volume; the tortuosity α_∞, which describes the change in magnitude and direction of the fluid microvelocity due to curliness of the pores; the characteristic viscous and thermal lengths Λ and Λ', which relate to the geometry of the pores through the viscous and thermal losses; and the flow resistivity σ, which is the ratio of the fluid viscosity η to the fluid permeability κ_f. The inverse problem consists in recovering some or all of these five parameters from measurements of the scattered fields, i.e., the reflected and transmitted fields in the case of a rigid frame porous plate.

Subsequent to the experimental apparatus of Beranek[11] to determine the porosity (porosimeter), or the flowmeter of Brown and Bolt[12] to determine the flow resistivity, several methods have been developed to characterize macroscopi-

[a)]Author to whom correspondence should be addressed. Electronic mail: jean-philippe.groby@univ-lemans.fr. Present address: Laboratoire d'Acoustique de l'Université du Maine, UMR6613 CNRS/Univ. du Maine, F-72085 Le Mans Cedex 9, France.
[b)]Present address: LMS international NV, B-3001 Leuven, Belgium.

cally homogeneous porous samples, or layered porous samples. From ultrasonic measurements, methods developed either in the frequency or in the time domain allow the determination of the porosity,[13] the tortuosity,[14] the porosity and tortuosity simultaneously,[15] the viscous and thermal characteristic lengths[16] (using the $Q\delta$ method), or the porosity, tortuosity, and viscous characteristic length (the thermal characteristic length being set to three times the viscous one) simultaneously.[17,18] Specific methods have been developed for the characterization of highly absorbent porous materials, based on the variation of the static pressure of the saturating fluid,[19] mainly for the determination of the viscous characteristic length and the tortuosity. Recently, the simultaneous reconstruction of the three profiles of porosity, tortuosity, and viscous characteristic length of a macroscopically inhomogeneous rigid frame porous plate from simulated reflection coefficients for various angles of incidence was carried out[20] through an optimization process.

A characteristic of these many studies and methods, particularly in the frequency domain, is the difference in sensitivity of each parameter on the physical quantities. For example, both the density and bulk modulus do not depend on σ in the asymptotic high frequency range. Moreover, most of the methods for the simultaneous reconstruction of the parameters appeal to minimization techniques.[18,19] On the other hand, two analytic methods were developed[9,10] appealing to measurements in an impedance tube.[21] The idea is first to reconstruct the equivalent density and bulk modulus for various frequencies from the measurement of the impedance and the acoustic index of refraction and then to recover the parameters involved in several dissipation models mainly through analytic inversion, but also after extrapolation if needed. This leads to four equations at each frequency formed by the real and imaginary parts of the effective density and bulk modulus. Concerning the JCAM, these methods enable the recovery of the tortuosity, viscous and thermal characteristic lengths, and flow resistivity assuming the porosity to be known. This assumption seems to be due to the necessary use of the complex form of the complete JCAM, which involves five parameters, for the correct modeling of the phenomena in the frequency range considered therein. For the number of to-be-reconstructed parameters to be lower than or equal to the number of equations, a simplified model should be used, for example, the asymptotic high frequency one. In this ultrasonic frequency range, only four parameters are relevant. Complete analytical recovery of the latter is also enabled from the knowledge of the *complex* density and bulk modulus. The low frequency bound of this asymptotic regime is not clear and depends on the material, but it can be assumed that for frequencies higher than 100 kHz, this regime is reached for most of the porous foams. The high frequency bound is the diffusion limit, which is related to some of the parameters to be reconstructed, particularly the characteristic length.[22]

The aim of this article is to present a simple, efficient, and analytic frequency domain method for the simultaneous recovery of the four parameters appealing to the high frequency approximation of the JCAM, i.e., ϕ, α_∞, Λ, and Λ', for a macroscopically homogeneous rigid frame porous plate saturated by air, i.e., $\eta=1.839\times10^{-5}$ m^2 s^{-1}, $P_0=1.013\,25\times10^5$ Pa, $\rho_f=1.213$ kg m^{-3}, and $\gamma=1.4$, by means of ultrasonic measurements at normal incidence. This method can be considered as complementary to those proposed in Refs. 9 and 10 for a complete analytical characterization of macroscopically homogeneous porous materials.

II. ANALYTICAL CHARACTERIZATION OF A MACROSCOPICALLY HOMOGENEOUS RIGID FRAME POROUS PLATE

In what follows, the pressure $p(x,t)$ is related to its Fourier transform $P(x,\omega)$ through

$$p(x,t)=\int_{-\infty}^{\infty}P(x,\omega)e^{-i\omega t}d\omega. \tag{1}$$

A. Basic theoretical retrieval equations

When solicited by a normally incident plane wave, the reflection $\mathcal{R}(\omega)$ and transmission $\mathcal{T}(\omega)$ coefficients of a homogeneous plate of thickness L occupied by the equivalent fluid medium $M^{[1]}$ and surrounded by the same fluid, denoted as medium $M^{[0]}$, take the following forms:[23]

$$\mathcal{R}=\frac{R(1-e^{2ik^{[1]}L})}{1-R^2e^{2ik^{[1]}L}},$$

$$\mathcal{T}=\frac{(1-R^2)e^{i(k^{[1]}-k^{[0]})L}}{1-R^2e^{2ik^{[1]}L}}, \tag{2}$$

wherein $R=(Z^{[1]}-Z^{[0]})/(Z^{[1]}+Z^{[0]})$ is the reflection coefficient at the interface between two semi-infinite media $M^{[1]}$ and $M^{[0]}$, with $Z^{[j]}=\rho^{[j]}c^{[j]}=\sqrt{K^{[j]}\rho^{[j]}}$ the impedance of $M^{[j]}$, and $\rho^{[j]}$, $c^{[j]}$, and $K^{[j]}$ being the density, sound speed, and bulk modulus, respectively, $j=0,1$. Let us now introduce $S_{11}(\omega)$ and $S_{21}(\omega)$, the coefficients of the scattering matrix $\mathbf{S}$, related to $\mathcal{R}$ and $\mathcal{T}$ through

$$S_{11}=\mathcal{R}=\frac{R(1-e^{2ik^{[0]}\zeta L})}{1-R^2e^{2ik^{[0]}\zeta L}},$$

$$S_{21}=\mathcal{T}e^{ik^{[0]}L}=\frac{(1-R^2)e^{ik^{[0]}\zeta L}}{1-R^2e^{2ik^{[0]}\zeta L}}, \tag{3}$$

wherein $\zeta(\omega)=k^{[1]}/k^{[0]}=c^{[0]}/c^{[1]}=\sqrt{K^{[0]}\rho^{[1]}/K^{[1]}\rho^{[0]}}$ is the acoustic index of refraction, i.e., the ratio between the wavenumbers of media $M^{[1]}$ and $M^{[0]}$. The acoustic index of refraction and the impedance ratio $z(\omega)=Z^{[1]}/Z^{[0]}=\sqrt{K^{[1]}\rho^{[1]}/K^{[0]}\rho^{[0]}}$ can be obtained analytically by inverting Eq. (3), yielding[24–26]

$$z=\pm\sqrt{\frac{(1+S_{11})^2-S_{21}^2}{(1-S_{11})^2-S_{21}^2}}, \tag{4a}$$

$$\zeta=\frac{-i}{k^{[0]}L}(\ln(X\pm i\sqrt{1-X^2}))+\frac{2m\pi}{k^{[0]}L}, \tag{4b}$$

wherein $X=(1-S_{11}^2+S_{21}^2)/2S_{21}$ and m is an integer.

For a rigid frame porous plate, the signs in Eqs. (4a) and (4b) can be determined by the requirements $\mathrm{Re}(z)\geq0$ and

$\mathrm{Im}(z) \geq 0$ for Eq. (4a), and $\mathrm{Im}(\zeta) \geq 0$ for Eq. (4b). Both requirements result from the outgoing wave condition together with the Fourier transform convention. In fact, z and ζ are related and their relationship can be used to determine the sign in Eqs. (4a) and (4b) as reported in Ref. 24. Effectively, a small perturbation of S_{11} and S_{21} easily produced in experimental measurements may change the sign of $\mathrm{Re}(z)$ and $\mathrm{Im}(\zeta)$, making it impossible to satisfy the previous requirements, as discussed in Ref. 27. The derived method is particularly adapted when $\mathrm{Re}(z)$ and $\mathrm{Im}(\zeta)$ are close to zero and is also of impractical interest for rigid frame porous materials because $\mathrm{Re}(z)$ is always greater than unity. Nevertheless, once the value of z is obtained, ζ can be determined without sign ambiguity through

$$\zeta = \frac{-i}{k^{[0]}L}\left(\ln\left(\frac{S_{21}(z+1)}{z+1-S_{11}(z-1)}\right)\right) + \frac{2m\pi}{k^{[0]}L}. \tag{5}$$

The integer m related to the branch index of $\mathrm{Re}(\zeta)$ can be determined by the requirement $\mathrm{Re}(\zeta) \geq \sqrt{\alpha_\infty}$ [because $\mathrm{Re}(c^{[1]}(\omega)) \leq c^{[0]}/\sqrt{\alpha_\infty}$]. Its determination is not straightforward, because the requirement depends on one of the parameters to reconstruct.

The condition $\mathrm{Re}(\zeta) \geq 1$ should normally be sufficient for its correct determination, but problems occur for large α_∞ and/or for high frequency solicitation because of large $k^{[0]}L$. This usually leads authors to use small thickness sample. This problem can be solved by use of an iterative scheme. The latter is initialized with m as determined through the condition $\mathrm{Re}(\zeta) \geq 1$ and consists in adding 1 to the previously calculated m as long as the slope of $\mathrm{Re}(\zeta(\omega))$ is not negative. This condition is a mathematical translation of the fact that the phase velocity $c^{[1]}(\omega)$ associated with a rigid frame porous material is an increasing function of the frequency toward its high frequency limit $c^{[0]}/\sqrt{\alpha_\infty}$; i.e., $\mathrm{Re}(\zeta) = \mathrm{Re}(c^{[0]}/c^{[1]}(\omega))$ is a decreasing function of ω. The latter iterative scheme is chosen because the condition $\mathrm{Re}(\zeta) \geq 1$ can lead to an underevaluated m for large $k^{[0]}L$ and so an underevaluated ζ. The reconstructed value of α_∞ is then underevaluated and can reach a nonphysical value less than unity. An iterative scheme based on the reconstructed value of α_∞ is also of impractical interest.

Another condition for the correct determination of the integer m is connected to the fact that both effective density and bulk modulus should be continuous functions of frequency. This problem was not encountered during the experiments presented here, but is also related to large values of $k^{[0]}L$. The reader can also refer to Ref. 24, keeping in mind that the high frequency limit clearly defines the initialization of the continuity condition that must be applied.

B. Reconstruction of ϕ, α_∞, Λ, and Λ' from the high frequency approximation of the JCAM

Let us assume in what follows that z and ζ can be evaluated for frequencies ν ($\omega = 2\pi\nu$) in the interval $[\nu_1, \nu_2]$ over which the high frequency approximation of the JCAM is valid. The equivalent density and bulk modulus ratios reduce to

$$\tilde{\rho} = \frac{\rho^{[1]}}{\rho^{[0]}} = \frac{\alpha_\infty}{\phi}\left(1 + \frac{2}{\Lambda}\sqrt{\frac{i\eta}{\omega\rho_f}}\right),$$

$$\tilde{K} = \frac{K^{[1]}}{K^{[0]}} = \frac{1}{\phi}\left(1 + \frac{2(1-\gamma)}{\Lambda'}\sqrt{\frac{i\eta}{\omega\ \mathrm{Pr}\ \rho_f}}\right), \tag{6}$$

wherein $\mathrm{Pr} = 0.71$ is the Prandtl number.

Noting that $\tilde{K}(\omega) = z/\zeta$ and $\tilde{\rho}(\omega) = z\zeta$, several reconstruction algorithms can be developed for the recovery of the four parameters appealing to the high frequency approximation of the JCAM. The algorithm that was found to give the most accurate results is based on a phase and amplitude analysis of both the equivalent density and bulk modulus. At each frequency, a value $\bar{x}(\omega)$ of the parameter x is recovered analytically, $x = \phi$, α_∞, Λ, or Λ'. The final value of x is taken to be the average value of $\bar{x}(\omega)$ over $[\omega_1, \omega_2]$; i.e., $x = \mathrm{mean}[\bar{x}(\omega)]$. The latter operation regularizes the inverse problem.

From $\mathrm{Im}(\tilde{K})/\mathrm{Re}(\tilde{K})$ and $\mathrm{Im}(\tilde{\rho})/\mathrm{Re}(\tilde{\rho})$, which are both independent of α_∞ and ϕ, the two characteristic lengths are first recovered from

$$\bar{\Lambda} = \sqrt{\frac{2\eta}{\omega\rho_f}}\left(\frac{\mathrm{Re}(\tilde{\rho}) - \mathrm{Im}(\tilde{\rho})}{\mathrm{Im}(\tilde{\rho})}\right),$$

$$\bar{\Lambda}' = (1-\gamma)\sqrt{\frac{2\eta}{\omega\ \mathrm{Pr}\ \rho_f}}\left(\frac{\mathrm{Re}(\tilde{K}) - \mathrm{Im}(\tilde{K})}{\mathrm{Im}(\tilde{K})}\right). \tag{7}$$

Then, ϕ is recovered from $\|\tilde{K}\|$ via

$$\bar{\phi} = \frac{1}{\|\tilde{K}\|}\left(1 + \frac{2(1-\gamma)}{\Lambda'}\sqrt{\frac{2\eta}{\omega\ \mathrm{Pr}\ \rho_f}} + \frac{4(1-\gamma)^2\eta}{\Lambda'^2\omega\ \mathrm{Pr}\ \rho_f}\right)^{1/2}, \tag{8}$$

and finally α_∞ is recovered from $\|\tilde{\rho}\|$ by means of

$$\overline{\alpha_\infty} = \phi\|\tilde{\rho}\|\left(1 + \frac{2}{\Lambda}\sqrt{\frac{2\eta}{\omega\rho_f}} + \frac{4\eta}{\Lambda^2\omega\rho_f}\right)^{-1/2}. \tag{9}$$

III. EXAMPLE OF A RECONSTRUCTION FROM EXPERIMENTAL DATA

The method has been applied to efficiently characterize several homogeneous rigid frame porous samples (polyurethane foam, melamine foam, etc.) with low and medium flow resistivities.

In what follows, a $L = 27.5$ mm thick medium resistivity rigid frame porous material sample is considered. This material has already been characterized at various frequencies (100 and 200 kHz) in the ultrasonic range by some of the authors of the present paper,[18] wherein a smaller sample of 7 mm thick was employed. Its characteristic parameters were also previously determined by use of different methods[15,16] and are recalled in Table I. Its flow resistivity is $\sigma = 38\,000$ N m s^{-4} (measured with a flowmeter).

A. Experimental setup and procedure

The experimental setup is shown in Fig. 1. Two air-coupled piezoelectric Ultran NCG100-D25 transducers de-

TABLE I. Properties of the studied homogeneous material as recovered in Refs. 15 and 16, as recovered in Ref. 18, and as recovered with our analytic method.

	ϕ	α_∞	Λ (μm)	Λ' (μm)
Use of Refs. 15 and 16	0.7 ± 0.05	1.25 ± 0.05	50 ± 10	$3\times\Lambda=150\pm30$
Ref. 18	0.72–0.74	1.21–1.26	42–66	$3\times\Lambda=126$–198
Present method	0.752	1.246	55.4	178.4

noted Ta and Tb, whose central frequencies are 100 kHz, are placed at x_a and x_b. The airborne wave is generated by Ta and detected by Ta in the pulse-echo mode in the reflection experiments or by Tb in the transmission experiments. All recorded signals are averaged over 512 waveforms.

The method is very sensitive to small variations of the sample alignment during the recording of the reflected $p^r(x_a,t)$ and the transmitted $p^t(x_b,t)$ fields. The sample remains in place for the measurement of both of these fields, its first interface being placed at the origin of the Cartesian coordinate system such that $p^r(x_a,t)$ and $p^t(x_b,t)$ can be recorded over the same time window of the oscilloscope; i.e., the first interface of the sample is placed approximately at equal distance from the two transducers; i.e., $x_a \approx |x_b|$. This procedure avoids problems caused by temporal shift between the windows used to measure reflected or transmitted waves with and without the sample, which is a source of error as discussed in Ref. 28 for the reconstruction of ζ and z.

The main difficulty with ultrasonic measurements is the evaluation from the time domain responses of the frequency domain reflection and transmission coefficients $\mathcal{R}$ and $\mathcal{T}$, with sufficient accuracy for the inversion to be carried out, and over a sufficiently large interval of frequency for the inverse problem to be regularized. This operation requires particular care, because it is often considered to be highly uncertain and/or ill-posed. This has led some authors to prefer a time domain inversion algorithm based on fractional derivatives[17,18] or to use time domain deconvolution.

The Fourier transforms of $p^r(x_a,t)$ and $p^t(x_b,t)$ are related to $\mathcal{R}$ and $\mathcal{T}$ through

$$P^r(x_a) = \mathcal{R}A^i e^{ik^{[0]}x_a} = \mathcal{T}P^{ri}(x_a),$$

$$P^t(x_b) = \mathcal{T}A^i e^{-ik^{[0]}x_b} = \mathcal{T}P^i(x_b), \tag{10}$$

wherein $A^i(\omega)$ is the spectrum of the incident field, $P^i(x_b)$ is the incident field as recorded with T_b in absence of the sample, and $P^{ri}(x_a)$ corresponds to time delay of the reflected field between the first interface of the sample and the location of T_a. The latter field, abusively called here "incident field," can be recorded in time domain with T_a in pulse-echo mode when a perfectly reflecting plate is placed such that its first interface is located at the origin. Any time delay between $p^{ri}(x_a,t)$ and $p^r(x_a,t)$ is expected if the first interface of the sample and the perfectly reflecting plate are located at the same position. Nevertheless, to avoid such source of error, $p^r(x_a,t)$ is translated in time in such a way that its maximum of amplitude occurs at the same time as the one of $p^{ri}(x_a,t)$. The frequency domain reflection and transmission coefficients are then evaluated by means of $\mathcal{R}=P^r(x_a,\omega)/P^{ri}(x_a,\omega)$ and $\mathcal{T}=P^t(x_b,\omega)/P^i(x_b,\omega)$ over the intersection of the two −3 dB bandwidths as calculated for $P^{ri}(x_a,\omega)$ and $P^i(x_b,\omega)$. These bandwidths are effectively different because the latter incorporate, to some extent, transfer functions of the transducers T_a and T_b, which are close but not identical.

In addition, the amplitude of $p^r(x_a,t)$ is often very small with a very poor signal to noise ratio (SNR) for relatively low and medium flow resistivity porous materials. The noise in a signal is usually assumed to be a stationary random process; i.e., its mean value is independent of the time origin. In the experiments, the acquisitions are done using a triggering process, which ensures the time origin of the signals. Recorded signals are averaged. The averaging operation induces a large decrease in the noise that is incoherent with respect to the trigger signal. Therefore, the main part of the remaining noise is coherent[29] with the trigger signal, which is the same for all the experiments. The remaining noise is mainly electronic from the measurement devices. In the reflection experiments only one transducer is used. To get rid of this remaining noise component, which is coherent, reproducible, and somewhat synchronized with the trigger, the latter, called $p^n(x_a,t)$, is recorded without the sample and perfectly reflecting plate, the emission being still on (the experimental setup is identical). The incident and reflected fields used for the inversion are $\widetilde{p^r}(x_a,t)=p^r(x_a,t)-p^n(x_a,t)$ and $\widetilde{p^{ri}}(x_a,t)=p^{ri}(x_a,t)-p^n(x_a,t)$, as shown Fig. 2 for $\widetilde{p^r}(x_a,t)$.

For highly absorbent foam and/or large flow resistivity porous material, a similar procedure can be employed to determine $\widetilde{p^t}(x_b,t)$ and $\widetilde{p^i}(x_b,t)$, by a prior recording of $p^n(x_b,t)$ with T_b, without the sample and by turning off the excitation delivered by T_a. This procedure is certainly less efficient than the one described for the experiment in reflection because two transducers are needed in transmission experiments, this imposing the excitation to be switched off.

In summary, the incident fields are recorded, without the sample, as follows: (1) Record $p^i(x_b,t)$ using Tb, (2) record the noise $p^n(x_a,t)$ using Ta in pulse-echo mode, (3) place an infinitely rigid plate (here an aluminum plate is used) whose first interface is located at the origin, and (4) record $p^{ri}(x_a,t)$

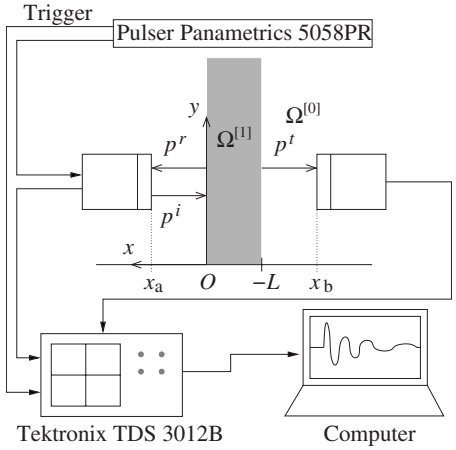


FIG. 1. Experimental setup and description of the configuration. The transducers are Ultran NCG100-D25 air-coupled piezoelectric transducers.

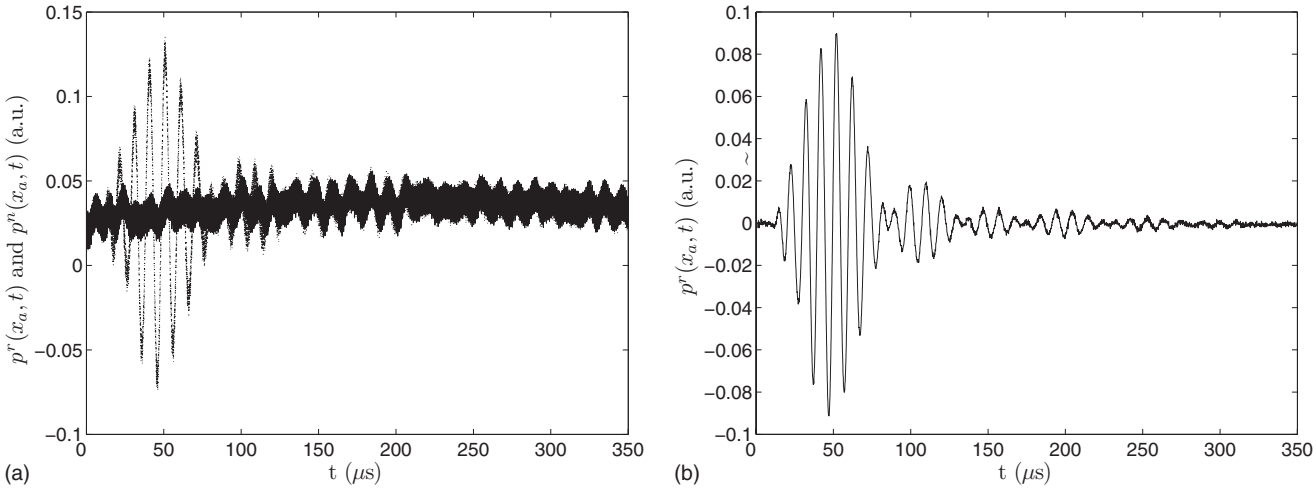


FIG. 2. (a) $p^r(x_a,t)$ (dotted curve) and $p^n(x_a,t)$ (solid curve) as recorded on the oscilloscope and (b) $\widetilde{p^r}(x_a,t)$ as used for the inversion.

using Ta in pulse-echo mode. The infinitely rigid plate is then replaced with the sample, carefully placed at the origin, and $p^r(x_a,t)$ is recorded in the pulse-echo mode using Ta while $p^t(x_b,t)$ is recorded employing Tb.

B. Numerical reconstruction from experimental data

Once m is correctly determined, the following parameters are recovered from $\bar{\phi}(\nu)$, $\overline{\alpha_\infty}(\nu)$, $\bar{\Lambda}(\nu)$, and $\overline{\Lambda'}(\nu)$,

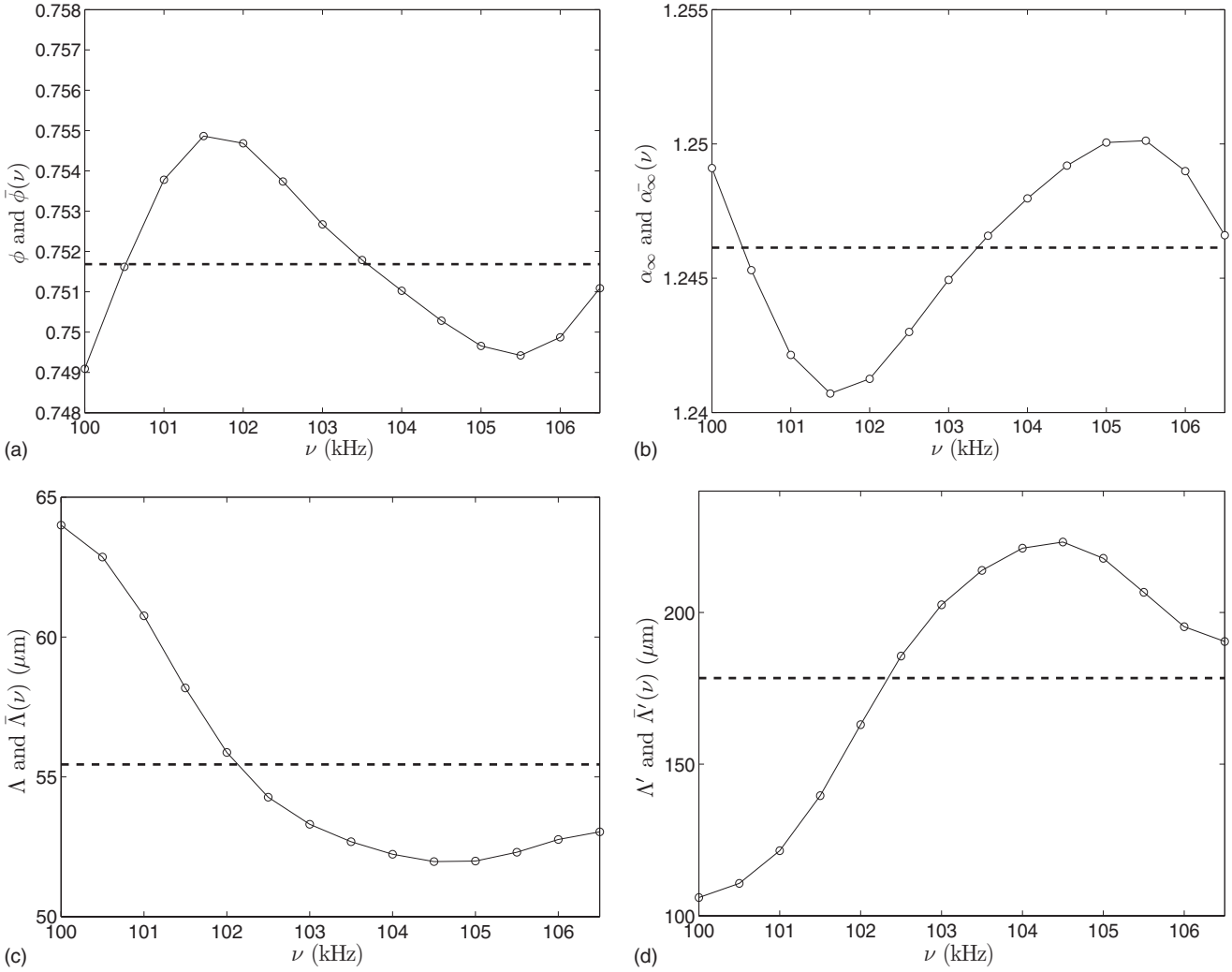


FIG. 3. (a) $\bar{\phi}(\nu)$ (solid curve) and its average value ϕ (dashed line), (b) $\bar{\alpha}_\infty(\nu)$ (solid curve) and its average value α_∞ (dashed line), (c) $\bar{\Lambda}(\nu)$ (solid curve) and its average value Λ (dashed line), and (d) $\overline{\Lambda'}(\nu)$ (solid curve) and its average value Λ' (dashed line). The frequency range corresponds to the intersection of the two −3 dB bandwidths as calculated for $\widetilde{P^{ri}}(x_a,\omega)$ and $P^i(x_b,\omega)$.

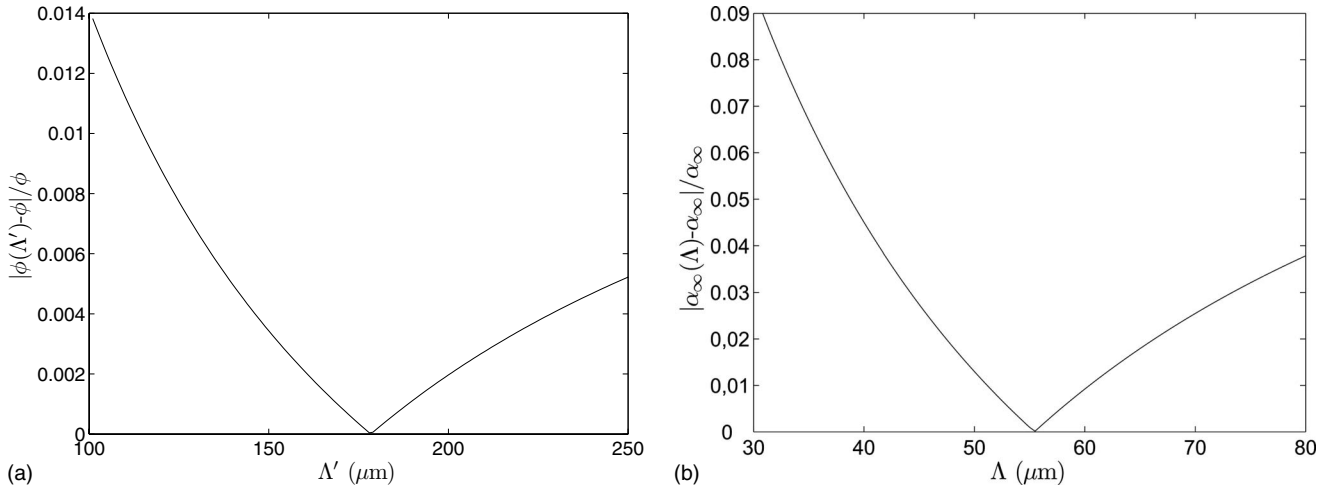


FIG. 4. Absolute value of the relative error of the recovered parameter ϕ as a function of Λ' (a) and α_∞ as a function of Λ (b), over the frequency bandwidth used for the reconstruction.

which are plotted in Fig. 3: $\phi=0.752$, $\alpha_\infty=1.246$, $\Lambda=55.4$ μm, and $\Lambda'=178.4$ μm. These parameters agree with those found in Ref. 18 via a time domain method and other ultrasonic methods[15,16] (see Table I). The unbiased estimators of the standard deviation, as calculated for the vector x as $s(x)=\sqrt{1/(n-1)\Sigma_{i=1}^{n}(x_i-\text{mean}(x))^2}$, are $s(\phi)=2\times10^{-3}$, $s(\alpha_\infty)=3.3\times10^{-3}$, $s(\Lambda)=4.25$ μm, and $s(\Lambda')=42.3$ μm. The latter estimator for Λ' seems large at first glance, but the constraint $\Lambda'=3\Lambda$ used in Ref. 18 leads to a value $\Lambda'=150\pm30$ μm, whose error is also of the same order as $s(\Lambda')$. The latter constraint, i.e., $\Lambda'=3\Lambda$, is debatable because the value of Λ' is usually between 2Λ and 3Λ for most of the porous foams.[22]

A specific feature of the reconstruction of the parameters appealing to the high frequency approximation of the JCAM is the fact that for given S_{11} and S_{21}, the reconstruction of ϕ through Eq. (8) is, as expected, rather insensitive to a variation of Λ', Fig. 4. Likewise, the influence of $s(\Lambda')$ is not significant on the reconstruction of the other parameters. The reconstruction of α_∞ is strongly sensitive on Λ, Fig. 4.

The reconstructed normalized density and bulk modulus, as calculated by introducing the reconstructed parameters in Eq. (6), over the frequency range used for the reconstruction are in good agreement with the measured ones, Fig. 5. It is clear that the developed analytic method for the recovery of the characteristic parameters realizes some kind of least squares fitting of both density and bulk modulus. This fit is much more efficient than a direct fitting of the curve. For example, use of an algorithm based on a basic fit of $\text{Im}(\rho^{[1]}/\rho^{[0]})$, which is theoretically equal to $1/\Lambda\sqrt{2\eta/\omega\rho_f}$, leads to a negative value of Λ because the experimental slope is clearly positive. Use of such an algorithm would require reconstructions over a larger frequency range.

The time domain reconstructed reflected and transmitted fields, as calculated from the reconstructed parameters introduced in Eqs. (6) and (2) via inverse Fourier transform of Eq. (10), $P^{ri}(x_a)$ and $P^{i}(x_b)$ being the measured incident signals, agree quite well with the experimentally recorded fields, Fig. 6. The Bravais–Pearson linear correlation coefficients between the latter fields are 0.9811 in reflection and 0.9819 in transmission.

The method was applied to characterize various samples with other piezoelectric Ultran transducers, whose central frequency is 200 kHz. Table II presents the values of the parameters as recovered with the present method and as recovered with other ultrasonic methods.[15–17] The flow

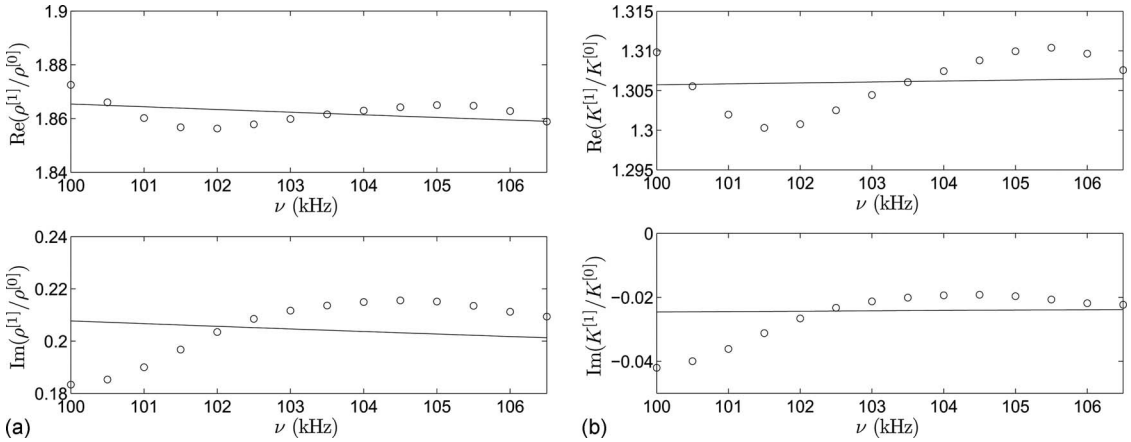


FIG. 5. Comparison between experimental (o) and reconstructed (—) real and imaginary parts of (a) the normalized density $\tilde{\rho}$ and (b) the bulk modulus $\tilde{K}$.

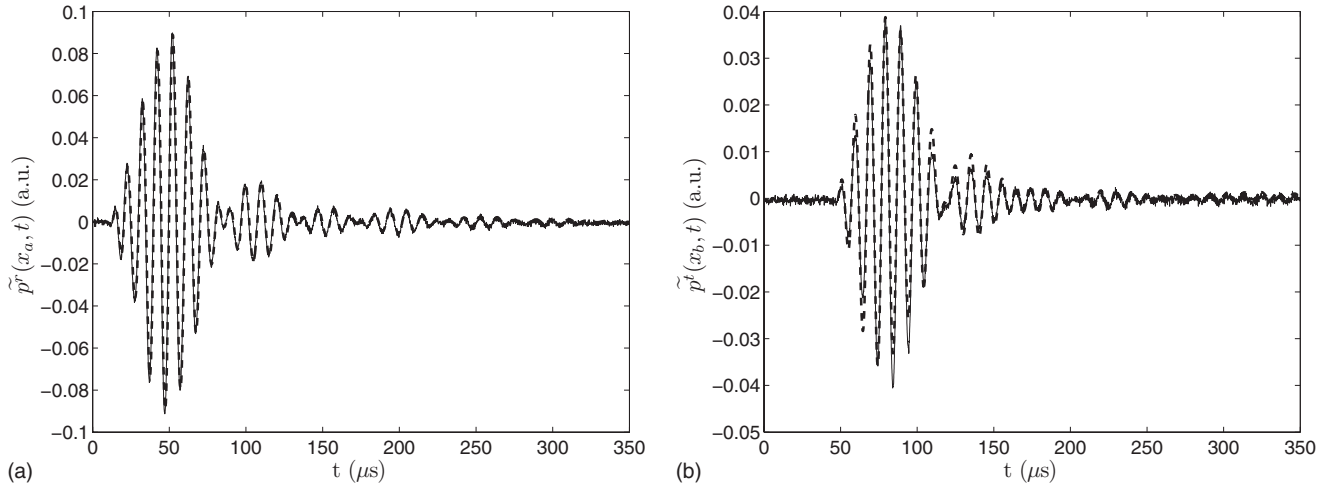


FIG. 6. Comparison between experimental (—) and simulated (– – –) signals, with recovered $\phi=0.752$, $\alpha_\infty=1.246$, $\Lambda=55.4$ μm, and $\Lambda'=178.4$ μm: (a) reflected signal and (b) transmitted signal.

resistivity of the polyurethane foam is low, i.e., $\sigma=2830$ N m s^{-4}, while the one of the melamine is $\sigma=12\,000$ N m s^{-4}.

C. Discussion

The method is sensitive to the approximation of the continuous Fourier transform by a discrete one and to the temporal shift, especially for the evaluation of S_{11}, as was pointed out in Ref. 28. Nevertheless, assuming that the Fourier transformed is correctly performed, and that the proposed process does not lead to temporal shift, other sources of error arise.

From

$$\frac{\partial z^2}{\partial S_{21}}=\frac{8S_{21}S_{11}}{((1-S_{11})^2-S_{21}^2)^2},$$

$$\frac{\partial z^2}{\partial S_{11}}=\frac{4(1-S_{11}^2-S_{21}^2)}{((1-S_{11})^2-S_{21}^2)^2}, \tag{11}$$

it is obvious that a weak transmission, i.e., large flow resistivity materials, has little influence on the retrieval of z. For the low and medium flow resistivity foams tested, S_{11} is close to zero. The reconstruction is also highly sensitive when S_{21} is close to one, which is in contradiction with the use of thin layer sample, for which reconstruction is much more efficient. On the other hand, from

$$\frac{\partial \zeta}{\partial S_{11}}=\frac{-1}{ik^{[0]}L}\frac{z-1}{z+1-S_{11}(z-1)},$$

$$\frac{\partial \zeta}{\partial S_{21}}=\frac{-1}{S_{21}ik^{[0]}L}, \tag{12}$$

it is obvious that ζ, for S_{11} close to zero, is rather insensitive to S_{11}, while it is very sensitive to S_{21} when the latter is close to zero. This also imposes conditions not only on the absorption of the sample to characterize, but also on its thickness: both should be small. This method seems particularly adapted to low and medium resistive materials, but less to highly resistive ones, for which S_{21} vanishes.

On the other hand, it is obvious from Eq. (7) that Λ and Λ' are, respectively, highly sensitive to $\mathrm{Im}(\tilde{\rho})$ and $\mathrm{Im}(\tilde{K})$, when the latter is small. Particularly, the reconstructed characteristic lengths are singular when the imaginary part of the density or of the bulk modulus vanishes. Both of these quantities are close to zero in the asymptotic high frequency regime of the JCAM. The regularization process partly avoids such problems, but decreasing the frequency of the solicitation is the best way to increase the value of both imaginary parts. From Eqs. (8) and (9), it is clear that reconstructed porosity is highly sensitive to $\|\tilde{K}\|$ when the latter is small, while the reconstructed tortuosity is less sensitive to variation of $\|\tilde{\rho}\|$.

One of the biggest limitations of the present method is closely linked to experiments, and particularly to the SNR for the evaluation of the reflection coefficient. The flow resistivity of the sample should not be very small. Effectively, very low resistivity foams exhibit quasi-null reflected fields. Their characterization by the present method is also impossible. A similar remark arises for highly resistive porous material, such as rock wool, which exhibits quasi-null transmitted fields.

TABLE II. Properties of other homogeneous materials as recovered with Refs. 15–17 and as recovered by our analytic method.

	L (mm)	ϕ	α_∞	Λ (μm)	Λ' (μm)
Melamine foam					
Use of Refs. 15–17	10	0.99	1.001	150	250
Present method		0.99	1.011	159.2	259.2
Polyurethane foam					
Use of Refs. 15–17	27	0.97	1.07	270	670
Present method		0.975	1.057	319.4	621

Another limitation is strongly linked to ultrasonic measurements, for which $k^{[0]}L$ is high. This induces possible difficulties in the evaluation of the correct branch of $\mathrm{Re}(\zeta)$. The frequency band used to perform the inversion, chosen to be the intersection of the -3 dB bandwidth of both Fourier transform of the recorded fields by T_a and T_b without sample, is thin. Use of larger band transducers should help in the regularization of the inverse problem and in its resolution. In this case, particular care should be paid to the frequency range used for the reconstruction, in a sense the model should be valid. Check of the constancy and to some extent to the continuity of the recovered parameters over the frequency range, as proposed in Refs. 9 and 10, should avoid this possible problem.

One of the main advantages of the present method is that the reconstructed parameters are not constrained, as it is usually the case when minimization techniques are applied. On the other hand, this is also a disadvantage because nonphysical values, such as negative characteristic lengths, porosity larger than 1, or tortuosity smaller than 1, can be reconstructed if experiments are not performed with sufficient care. Moreover, among all the experiments performed, it was found that whatever the reconstructed parameters are, even nonphysical ones, the reconstructed time domain signals are always in agreement with the experimental ones. An a priori correct ratio between the reconstructed Λ' and Λ seems to be indicative of the quality of the experiments, and leads to the accurate reconstructions of the other parameters. This a priori is connected to the type of material tested. For most of the industrial foams, a ratio Λ'/Λ that lies between 2 and 3 seems to be a good indicator of the reconstruction. For other industrial porous materials, other ratios can be used, for example, a ratio close to 2 for fibrous materials.

It is then important to check the consistency of the recovered parameters with regard to the frequency of solicitation in order to validate the use of the asymptotic high frequency JCAM. This can be simply achieved at high frequency (diffusion limit) by checking if the wavelength of the central frequency of the incident field is larger than the thermal characteristic length. Effectively, the latter characteristic length represents a measure of the average pore size (although the "pore" is not always straightforwardly defined) while the viscous characteristic length corresponds to the average size of the "constrictions" in the porous medium,[22] i.e., the average distance between pore walls in the narrower areas of the pore volume. For example, the ratio of the wavelength in air at 100 kHz over the recovered value of Λ' for the first medium resistivity porous sample studied in this article is around 19. As stated in the Introduction, the low frequency bound of the asymptotic high frequency JCAM being not clear, it is then quite difficult to validate the use of the dissipation model with regard to the latter, particularly because the flow resistivity is not recovered through our method and despite the fact that a frequency higher than 100 kHz usually ensures this validation. Nevertheless, this problem can be partly solved by performing ultrasonic experiments for various central frequencies of the transducers, for example, at 100 and 200 kHz, keeping in mind that the dissipation model should be valid, that additional problems related to large $k^{[0]}L$ could occur, and that a correct SNR is required. If the reconstructed parameters are quite similar, the constancy of the dissipation model and so of the reconstructed parameters is achieved. This operation is similar to use of larger band transducers, in the sense it would help the regularization of the problem, when the dissipation model is valid.

IV. CONCLUSION

A frequency domain method has been developed, which allows the analytic reconstruction of the four parameters appealing to the high frequency approximation of the JCAM (porosity, tortuosity, and thermal and viscous characteristic lengths) for macroscopically homogeneous rigid frame porous materials. The reconstruction is achieved by first retrieving the complex and frequency-dependent acoustic index of refraction and impedance ratio, from which the four latter parameters are reconstructed. The method and algorithm were tested on relatively low and medium flow resistivity porous samples, which mainly cover the range of resistivity values of plastic foams, and its accuracy has been demonstrated in the ultrasonic domain on three materials. The experiments must be very precise even if the reconstructed time traces are very close to the experimental ones. This is to avoid reconstructed parameters that are meaningless. This situation is often avoided when minimization techniques are used to solve inverse problems. It was found empirically that when the ratio of the recovered thermal and viscous characteristic lengths lies between 2 and 3, as it is the case for most industrial foams, the experiments are performed with sufficient efficiency and the recovered parameters turn out to be accurate.

For the reconstruction to be performed, time domain reflected and transmitted signals should be recorded with a high SNR. Some methods have been employed here to increase the SNR, especially for the reflected field, but it also imposes some constraints on the materials that can be tested with the present method. The flow resistivity should not be very small for the record of the reflected field, while it should not be very large for the record of the transmitted field.

This algorithm being based on analytic developments, further investigations should lead to the development of a more complex algorithm for the complete analytic characterization of homogeneous foam samples, i.e., the recovery of additional parameters such as the flow resistivity from additional lower frequency measurements or parameters of other models relevant for mid frequency range, for example. This method can be viewed as a complementary to the low frequency method developed in Refs. 9 and 10 in the sense that it allows the recovery of the porosity, which should be known in Refs. 9 and 10. The other recovered parameters in the two frequency ranges may then be compared.

ACKNOWLEDGMENT

The authors would like to thank A. Wirgin for his useful help during the experiments and the writing of this paper.

[1]M. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid—I. Low frequency range," J. Acoust. Soc. Am. **28**, 168–178 (1956); M. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid—II. High frequency range," J. Acoust. Soc. Am. **28**, 179–191 (1956).
[2]M. Biot, "Mechanics of deformation and acoustic propagation in porous media," J. Appl. Phys. **33**, 1482–1498 (1962).
[3]D. Johnson, J. Koplik, and R. Dashen, "Theory of dynamic permeability and tortuosity in fluid-saturated porous media," J. Fluid Mech. **176**, 379–402 (1987).
[4]Y. Champoux and J. Allard, "Dynamic tortuosity and bulk modulus in air-saturated porous media," J. Appl. Phys. **70**, 1975–1979 (1991).
[5]K. Attenborough, "Acoustical characteristics of porous materials," Phys. Rep. **82**, 179–227 (1982).
[6]D. Lafarge, P. Lemarinier, J.-F. Allard, and V. Tarnow, "Dynamic compressibility of air porous structures at audible frequencies," J. Acoust. Soc. Am. **102**, 1995–2006 (1997).
[7]D. Wilson, "Relaxation-matched modeling of propagation through porous media, including fractal pore structure," J. Acoust. Soc. Am. **94**, 1136–1145 (1993).
[8]L. De Ryck, J.-P. Groby, P. Leclaire, W. Lauriks, A. Wirgin, Z. Fellah, and C. Depollier, "Acoustic wave propagation in a macroscopically inhomogeneous porous medium saturated by a fluid," Appl. Phys. Lett. **90**, 181901 (2007).
[9]R. Panneton and X. Olny, "Acoustic determination of the parameters governing viscous dissipation in porous media," J. Acoust. Soc. Am. **119**, 2027–2040 (2006).
[10]X. Olny and R. Panneton, "Acoustic determination of the parameters governing thermal dissipation in porous media," J. Acoust. Soc. Am. **123**, 814–824 (2008).
[11]L. L. Beranek, "Acoustic impedance of porous materials," J. Acoust. Soc. Am. **13**, 248–260 (1942).
[12]R. Brown and R. Bolt, "The measurement of flow resistance of porous acoustic materials," J. Acoust. Soc. Am. **13**, 337–344 (1942).
[13]Y. Champoux, M. Stinson, and G. Daighle, "Air-based system for the measurement of porosity," J. Acoust. Soc. Am. **89**, 910–916 (1991).
[14]J. Allard, B. Castagnede, M. Henry, and W. Lauriks, "Evaluation of tortuosity in acoustic porous materials saturated by air," C. R. Acad. Sci. **322**, 754–755 (1994).
[15]Z. Fellah, S. Berger, W. Lauriks, C. Depollier, C. Aristégui, and J.-Y. Chapelon, "Measuring the porosity and the tortuosity of porous materials via reflected waves at oblique incidence," J. Acoust. Soc. Am. **113**, 2424–2433 (2003).
[16]P. Leclaire, L. Kelders, W. Lauriks, N. Brown, M. Melon, and B. Castagnède, "Determination of the viscous and thermal characteristic lengths of plastic foams by ultrasonic measurements in helium and air," J. Appl. Phys. **80**, 2009–2012 (1996).
[17]Z. Fellah, C. Depollier, S. Berger, W. Lauriks, P. Trompette, and J.-Y. Chapelon, "Determination of transport parameters in air-saturated porous materials via reflected ultrasonic waves," J. Acoust. Soc. Am. **114**, 2561–2569 (2003).
[18]Z. Fellah, F. Mitri, M. Fellah, E. Ogam, and C. Depollier, "Ultrasonic characterization of porous materials: Inverse problem," J. Sound Vib. **302**, 746–759 (2007).
[19]C. Ayrault and S. D. Griffiths, "Separation of viscothermal losses and scattering in ultrasonic characterization of porous media," Ultrasonics **45**, 40–49 (2006).
[20]L. De Ryck, W. Lauriks, P. Leclaire, J.-P. Groby, A. Wirgin, and C. Depollier, "Reconstruction of material properties profiles in one-dimensional macroscopically inhomogeneous rigid frame porous media in the frequency domain," J. Acoust. Soc. Am. **124**, 1591–1606 (2008).
[21]H. Utsuno, T. Tanaka, and T. Fujikawa, "Transfer function method for measuring characteristic impedance and propagation constant of porous materials," J. Acoust. Soc. Am. **86**, 637–643 (1989).
[22]J.-F. Allard, *Propagation of Sound in Porous Media: Modeling Sound Absorbing Materials* (Chapman and Hall, London, 1993).
[23]J.-P. Groby, L. De Ryck, P. Leclaire, A. Wirgin, W. Lauriks, R. P. Gilbert, and Y. S. Xu, "Use of specific Green's functions for solving direct problems involving a heterogeneous rigid frame porous medium slab solicited by acoustic waves," Math. Methods Appl. Sci. **30**, 91–122 (2007).
[24]X. Chen, T. Grzegorczyk, B. Wu, J. Pacheco, and J. Kong, "Robust method to retrieve the constitutive effective parameters of metamaterials," Phys. Rev. E **70**, 016608 (2004).
[25]P. Markos and C. Soukoulis, "Transmission properties and effective electromagnetic parameters of double negative metamaterials," Opt. Express **11**, 649–661 (2003).
[26]R. Smith, S. Schultz, P. Markos, and C. Soukoulis, "Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients," Phys. Rev. B **65**, 195104 (2002).
[27]R. Ziolkowski, "Design, fabrication, and testing of double negative metamaterials," IEEE Trans. Antennas Propag. **51**, 1516–1529 (2003).
[28]A. Nicolson, "Measurement of the intrinsic properties of materials by time domain techniques," IEEE Trans. Instrum. Meas. **19**, 377–382 (1970).
[29]J. S. Bendat and A. G. Piersol, *Engineering Application of Correlation and Spectral Analysis* (Wiley, New York, 1993).

Vector intensity reconstructions in a volume surrounding a rigid spherical microphone array[a)]

Earl G. Williams[b)]
Acoustics Division, Naval Research Laboratory, Washington DC 20375

Kazuhiro Takashima
Nittobo Acoustic Engineering Co., Ltd., Tokyo, Japan

(Received 26 May 2009; revised 4 December 2009; accepted 4 December 2009)

A technique is described to image the vector intensity in the near field of a spherical array of microphones flush mounted in a rigid sphere. The spatially measured pressure is decomposed into Fourier harmonics in order to reconstruct the volumetric vector intensity outside the sphere. The theory for this reconstruction is developed in this paper. The resulting intensity images are very successful at locating and quantifying unknown exterior acoustic sources, ideal for application in noise control problems in interior spaces such as automobiles and airplanes. Arrays of varying numbers of microphones and radii are considered and compared and errors are computed for both theory and experiment. It is demonstrated that this is an ill-posed problem below a cutoff frequency depending on array design, requiring Tikhonov regularization below cutoff. There is no low frequency limit on operation, although the signal-to-noise ratio is the determining factor for high-spatial resolution at low frequencies. It is shown that the upper frequency limit is set by the number of microphones in the array and is independent of noise. The accuracy of the approach is assessed by considering the exact solution for the scattering of a point source by a rigid sphere. Several field experiments are presented to demonstrate the utility of the technique. In these experiments, the partial field decomposition technique is used and holograms of multiple exterior sources are separated and their individual volumetric intensity fields imaged. In this manner, the intensity fields of two uncorrelated tube sources in an anechoic chamber are isolated from one another and separated intensity maps are obtained from over a broad frequency range. In a practical application, the vector intensity field in the interior of an automobile cabin is mapped at the fundamental of the engine vibration using the rigid sphere positioned at the driver's head. The source regions contributing to the interior cabin noise are identified. [DOI: 10.1121/1.3278591]

PACS number(s): 43.40.Sk, 43.60.Pt, 43.60.Jn, 43.60.Fg [EJS] Pages: 773–783

I. INTRODUCTION

Measurement of the acoustic intensity has long been a popular subject due to the fact that it directly measures the power per unit area radiated by sources and thus is an ideal source quantification tool in noise analysis. The popular two microphone intensity probe and more recently p-u probes measure the vector intensity very accurately at a single point over a wide frequency range. Multiple-microphone probes have been designed to measure the intensity vector at a point. Our objective in this research is to use a multiple microphone probe to project the intensity throughout a volume as opposed to a single point.[1,2] This projection or imaging is not a direct field measurement but is a near-field holographic reconstruction of the field away from the measurement surface and as such provides a spatially smoothed rendition of the intensity. This smoothing effect, true of all inverse problems of near-field acoustical holography (NAH), still provides an accurate prediction of the acoustic intensity although the objective is no longer a subpercent error, but instead an accurate image of the flow of the intensity in a volume is desired. In previous work, a transparent open sphere comprised of 50 microphones was used to measure the volumetric intensity field inside an air-craft cabin.[3] Much of the relevant literature was reviewed in that paper, so it is not repeated here. Our interest in this paper is the closed sphere, a rigid spherical surface impregnated with a set of microphones. The rigid sphere with surface mounted microphones has received a lot of interest lately [4–8] in which beamforming is the main objective. This work makes use of the Green's function for a rigid spherical surface, the same as the one used in the present paper. Other literature has concentrated on the multiple-microphone open sphere[9,10] again in beamforming research from which the directions of sound sources can be obtained.

We extend the previous analysis[3] by developing in Sec. II a near-field holographic formulation using a rigid sphere of microphones to predict the intensity vector in a volume outside the sphere. A theoretical discussion of the expected errors is given in Sec. III. In Sec. IV, we develop an approach to determine the all-important Fourier coefficients of the measured pressure field on a sphere. The developed formulation is tested in a numerical experiment consisting of a

[a)]This work is presented at Acoustics'08 Paris.
[b)]Author to whom correspondence should be addressed. Electronic mail: earl.williams@nrl.navy.mil

point source scattering off a rigid sphere and the resulting errors are delineated as a function of frequency for spherical arrays consisting of 31, 64, and 100 sensors at two different radii in Secs. V and VI. Stabilization of the reconstructed field is necessary at low frequencies and a mathematical technique for this is discussed in Sec. VII. In Sec. VIII, we present real experiments with a commercially available 31 element sphere probe applied to multiple sources in an anechoic chamber. Partial reference NAH is used to separate the two holograms for the sources. Finally, in a practical demonstration of the utility of the approach for noise identification inside an automobile, the intensity field is imaged inside an automobile driven on a dynamometer.

II. RIGID SPHERE ACOUSTIC INTENSITY PROBE: RECONSTRUCTION EQUATIONS

Consider a spherical reconstruction volume $\mathcal{V}$ represented by spherical coordinates $\mathbf{r}=(r,\theta,\phi)\in\mathcal{V}$ of extent defined by $a\leq r\leq r_{\max}$, which is source free (homogeneous wave equation applies) with a rigid sphere located at the interior boundary $r=a$. Any pressure field on a sphere of radius a can be decomposed into a set of orthonormal modes called spherical harmonics, $Y_n^m(\theta,\phi)$,

$$p(\mathbf{a},\omega)\equiv\sum_{n=0}^{\infty}\sum_{m=-n}^{n}P_{mn}(a,\omega)Y_n^m(\theta,\phi), \tag{1}$$

with $\mathbf{a}=(a,\theta,\phi)$ and where the unknowns P_{mn} in this equation are called the Fourier coefficients given by

$$P_{nm}(a,\omega)=\int\int p(\mathbf{a},\omega)Y_n^m(\theta,\phi)^*d\Omega, \tag{2}$$

with $d\Omega\equiv\sin\theta d\theta d\phi$.

Given these Fourier coefficients, the acoustic pressure $p(\mathbf{r},\omega)$ may be represented anywhere in $\mathcal{V}$ by an expansion in terms of the spherical harmonics and a radial propagator, $G_n(r,a,\omega)$,

$$p(\mathbf{r},\omega)=\sum_{n=0}^{\infty}G_n(r,a,\omega)\sum_{m=-n}^{n}P_{mn}(a,\omega)Y_n^m(\theta,\phi). \tag{3}$$

The propagator must be composed of radial functions that satisfy the wave equation in spherical coordinates, that is, $j_n(kr)$ and $y_n(kr)$, the spherical Bessel functions[11] with $k=\omega/c$. It must also satisfy the boundary condition at the rigid sphere, $\partial G_n(r,a,\omega)/\partial r=0$, and must equal unity at $r=a$ which leads to

$$G_n(r,a,\omega)=(ka)^2(j_n(kr)y_n'(ka)-j_n'(ka)y_n(kr)). \tag{4}$$

In the modeling of a spherical measurement array, we are forced to limit the number of terms in the expansion of Eq. (3) and use a regularization filter $0\leq F_n\leq 1$ (discussed in Sec. VII). Thus we rewrite Eq. (3) as

$$p_N(\mathbf{r},\omega)\equiv\sum_{n=0}^{N}F_n(\omega)G_n(r,a,\omega)\sum_{m=-n}^{n}P_{mn}(a,\omega)Y_n^m(\theta,\phi), \tag{5}$$

where p_N has the important interpretation as a spatially bandlimited representation of the exact pressure p.

The components of the velocity vector satisfying Euler's equation $i\rho ck\vec{v}(\mathbf{r})=\vec{\nabla}p(\mathbf{r})$, where

$$\vec{\nabla}p=\hat{e}_r\frac{\partial p}{\partial r}+\hat{e}_\theta\frac{1}{r}\frac{\partial p}{\partial\theta}+\hat{e}_\phi\frac{1}{r\sin\theta}\frac{\partial p}{\partial\phi}, \tag{6}$$

are given in terms of these unknown Fourier coefficients.[11]

$$v_\theta(\mathbf{r},\omega)=\frac{1}{i\omega\rho}\sum_{n=0}^{N}F_n(\omega)\frac{G_n(r,a,\omega)}{r}\sum_{m=-n}^{n}P_{mn}(a,\omega)\frac{\partial Y_n^m(\theta,\phi)}{\partial\theta},$$

$$v_\phi(\mathbf{r},\omega)=\frac{1}{i\omega\rho}\sum_{n=0}^{N}F_n(\omega)\frac{G_n(r,a,\omega)}{r}\sum_{m=-n}^{n}P_{mn}(a,\omega)\frac{imY_n^m(\theta,\phi)}{\sin\theta},$$

$$v_R(\mathbf{r},\omega)=\frac{1}{i\rho c}\sum_{n=0}^{N}F_n(\omega)G_n'(r,a,\omega)\sum_{m=-n}^{n}P_{mn}(a,\omega)Y_n^m(\theta,\phi), \tag{7}$$

with $G_n'\equiv\partial G_n/\partial kr$ and the equalities hold strictly only in the limit as $N\to\infty$ and $F_n=1$, $\forall n$ (no filter). Finally, the active intensity vector $\vec{I}$ in spherical coordinates is then determined by the usual expression using unit vectors $\hat{e}$:

$$\vec{I}(\mathbf{r},\omega)=\tfrac{1}{2}\Re[p^*(v_\theta\hat{e}_\theta+v_\phi\hat{e}_\phi+v_R\hat{e}_R)]. \tag{8}$$

The radial dependence in the reconstruction equations is borne by the propagator in Eq. (4) and its derivative with respect to kr in Eq. (7). Its behavior forms the underpinnings of the success or failure of the projection of the acoustic intensity vector. A representative behavior is displayed in Fig. 1 in which the logarithms of G_n (left plot) and G_n' (right plot) are plotted as a function of ka under the condition that $r=2a$; that is, the reconstruction is carried out at a radius two times the radius of the rigid sphere for $n=4,7,9$. The figure shows that the propagators exhibit a low frequency region in which the propagator increases with n located before the first null in each curve. This behavior with n is borne out by the small argument expansions $[n\gg(kr)^2]$ for the Bessel functions that can be shown to yield

$$G_n(r,a,\omega)\to\frac{n+1}{2n+1}\left(\frac{kr}{ka}\right)^n. \tag{9}$$

This formula also indicates that as r increases (in this low frequency region) the propagators also increase. These increases are important because they imply an ill-posed problem; that is, small variations in P_{mn} due to noise when multiplied by the propagator lead to large variations in the reconstructed field. This effect leads to the need for regularization, discussed later in this paper.

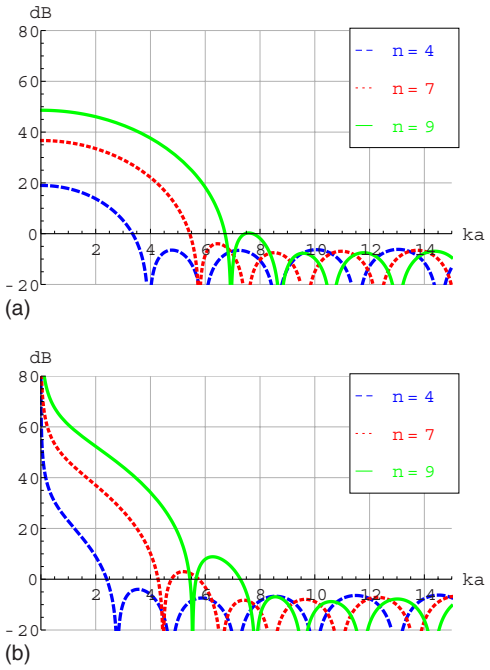


FIG. 1. (Color online) The pressure propagator $20\log_{10}|G_n|$ (left) and the normal velocity propagator $20\log_{10}|G_n'|$ (right) versus ka with $r=2a$ for the three values of n indicated in the key.

III. ERRORS ANALYSIS

References 11 (Chap. 7) and 3 provided an extensive error analysis for an open spherical array which can be repeated for the closed (rigid) array, with only minor modification of the equations. We provide just a summary here for the sake of brevity. To formulate the error, we assume that the measured pressure $\tilde{p}$ is contaminated by Gaussian, spatially uncorrelated noise ϵ that adds to the exact pressure p so that $\tilde{p}=p+\epsilon$. The variance of the noise is σ^2 independent of position and its mean is zero. There is an important error that does not depend on noise but arises from the use of a finite number of basis functions to represent the field, indicated by the limit N in Eq. (5). We call this the "base" error and define it by (setting $F_n=1$ in p_N)

$$\mathcal{E}_b(N,\mathbf{r},\omega) \equiv \left(\frac{\langle|p(\mathbf{r})-p_N(\mathbf{r})|^2\rangle}{\langle|p(\mathbf{r})|^2\rangle}\right)^{1/2},$$

where we have defined the brackets (an average over a sphere) as

$$\langle|p_N|^2\rangle \equiv \frac{1}{4\pi}\int\int |p_N|^2 d\Omega.$$

The base error provides a measure of how well a sphere with a fixed number of measurement points will do in the best of circumstances, a noiseless world. Using Eqs. (3) and (5), the base error becomes

$$\mathcal{E}_b(N,\mathbf{r},\omega) = \left(\frac{\sum_{n=N+1}^{\infty} G_n(r,a,\omega)^2 \sum_{m=-n}^{n} |P_{mn}|^2}{\sum_{n=0}^{\infty} G_n(r,a,\omega)^2 \sum_{m=-n}^{n} |P_{mn}|^2}\right)^{1/2}, \quad (10)$$

where P_{mn} are computed for a noise-free pressure. We can see that as $N\to\infty$, the base error vanishes. Note the base error is computed without any filtering.

The error due to noise can be formulated by adding noise to the band-limited term at the measurement radius, $\tilde{p}_N(\mathbf{a})=p_N(\mathbf{a})+\epsilon$, computing the Fourier coefficients with the added noise $\tilde{P}_{mn}(a,\omega)$, projecting to r using Eq. (5), and writing the noise error as

$$\mathcal{E}_\sigma(N,\mathbf{r},\omega) = \left(\frac{\langle E[|\tilde{p}_N(\mathbf{r})-p_N(\mathbf{r})|^2\rangle]}{\langle|p(\mathbf{r})|^2\rangle}\right)^{1/2},$$

where $E[\cdot]$ represent the expected value. The total error is $\mathcal{E}=\sqrt{\mathcal{E}_b^2+\mathcal{E}_\sigma^2}$. With a bit of algebra, we find that the "noise" error reduces to

$$\mathcal{E}_\sigma(N,\mathbf{r},\omega) = \sqrt{\frac{\sigma^2}{M\langle|p(\mathbf{r})|^2\rangle}\sum_{n=0}^{N}(2n+1)F_n(\omega)^2 G_n(r,a,\omega)^2}, \quad (11)$$

where $M\geq(N+1)^2$ is the number of microphones in the array. The condition $M\geq(N+1)^2$ provides a Nyquist condition that the number of microphones must be equal too or greater than the total number of spherical harmonics, that is, the number of basis functions used in the expansion if Eq. (1) must not be larger than the number of points. In the noise error equation, $\sqrt{\langle|p(\mathbf{r})|^2\rangle}/\sigma$ looks like the signal-to-noise ratio (SNR) at the projected radius r for the exact field. One can also interpret the term $\sigma^2\Sigma_{n=0}^{N}(2n+1)F_n^2G_n^2/M$ as the amplification of the noise variance due to projection to the radius r especially when $F_n=1\ \forall n$. This amplification arises at low frequencies, as Fig. 1 indicates, but not at higher frequencies (to the right of the first null). This leads us to the following important conclusion: when $r>a$ and for unity filters the noise term *increases* with N at low frequencies, opposite to the base error term which diminishes with N. At higher frequencies the noise error no longer increases with N, whereas the base error does. This is identical to the conclusions about the open array made in the Refs. 3 and 11. Another important conclusion arises from this. The low frequency behavior implies that the reconstruction problem is ill-posed only at low frequencies (for the radial range considered here) and leads to the requirement for the regularization filter $F_n(\omega)$ in Eq. (11) that can limit the value n to be less than the Nyquist value N and thus limit the noise amplification. We discuss this in more detail in Sec. VII.

Of course the errors above are only for the reconstructed pressure. It is impossible to derive simple formulas for the intensity vector errors and they must be computed numerically. We do this in Sec. VI below for five different cases to illustrate the dependency with the radius of the rigid sphere as well as with the number of microphones in the array.

IV. CALCULATION OF FOURIER COEFFICIENTS

All the field calculations in Eqs. (5) and Eq. (7) depend on the Fourier coefficients, P_{mn} evaluated at the surface of the sphere as implied in Eq. (1). These coefficients can be computed by a discretization of either Eq. (2) or Eq. (1). We use the latter equation since Eq. (2) requires an area be assigned to each of the measurement nodes, which may be difficult for non-uniform distributions. We now describe this method.

Define a set of M pressure measurement locations on a rigid sphere of radius a given by $p(a,\theta_i,\phi_i)$, $1\leq i\leq M$ represented by a column vector $\mathbf{p}\in\mathbb{C}^{M\times 1}$. We approximate Eq. (1) by limiting the sum over n to $n\leq N$ and write it as a matrix equation, treating P_{mn} as a column vector with $(N+1)^2$ components $\mathbf{P}=[P_{00},P_{-11},P_{01},P_{11},\ldots]^t$, where t represents transpose, and $Y_n^m(\theta_i,\phi_i)$ a matrix $\mathbf{Y}\in\mathbb{C}^{M\times(N+1)^2}$ ordered such that $\mathbf{p}=\mathbf{YP}$. A singular value decomposition (SVD) of $\mathbf{Y}$ yields $\mathbf{Y}=\mathbf{U\Sigma V^H}$. $\mathbf{\Sigma}$ is a diagonal matrix of singular values and since $\mathbf{U}$ and $\mathbf{V}$ are unitary we can solve for the Fourier coefficients $\mathbf{P}$ (that is, P_{mn}),

$$\mathbf{P}=\mathbf{V\Sigma}^{-1}\mathbf{U^H p}. \tag{12}$$

The condition number of the matrix $\mathbf{Y}$ is given by the ratio of the largest singular value to the smallest and provides a measure for how stable the inversion is. For a 31 microphone array, the condition number with $N=4$ (25 singular values) was 3.3. Condition numbers for a 64 ($N=7$) and 100 ($N=9$) element arrays (these arrays are described below) were 6.3 and 15.6, respectively. These condition numbers are excellent and all the inverted singular values can be used as a result. Thus Eq. (12) could be computed using the pseudo-inverse instead of a SVD. This method proved to be very accurate in calculating the Fourier coefficients.

In the paper on the open sphere,[3] the microphone locations were selected so that they fell at the integration points for a 50 point exact quadrature algorithm. The SVD method employed here eliminates this requirement and thus is much more general and can be used for any distribution of microphones on the sphere.

V. POINT SOURCE ANALYSIS

The reconstruction equations presented in Sec. IV can be evaluated by considering a point source located at (r_0,θ_0,ϕ_0) outside the rigid sphere. The point source provides a tractable mathematical formalism that can be studied to help understand application in a more complex source environment such as one may find inside an aircraft or automobile. In fact, we can compute the exact field by solving the rigid scattering problem, as has been done in Ref. 12. The incident field from this source is given by $p_i=e^{ikR}/kR$, where R is centered at (r_0,θ_0,ϕ_0) and the center of the rigid sphere is located at the origin. The total field (scattered plus incident fields) outside the rigid sphere is given by[12] [Eq. (10.67)]

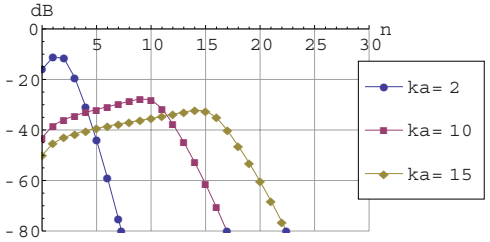


FIG. 2. (Color online) $20\log_{10}|P_{mn}|$ from Eq. (16) with $r_0=5a$, $\theta_0=\phi_0=0$ (point source on the polar axis) for three values of ka.

$$p(\mathbf{r})=4\pi i\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left[j_n(kr_<)-\frac{j_n'(ka)}{h_n'(ka)}h_n(kr_<)\right]\times h_n(kr_>)Y_n^m(\theta,\phi)Y_n^m(\theta_0,\phi_0)^*, \tag{13}$$

where $r_>$ is the greater of r,r_0 and $r_<$ the lesser of the two, and $h_n\equiv h_n^{(1)}$ is the spherical Hankel function. We recognize the first term of this expansion $j_n(kr_<)h_n(kr_>)$ as due to the incident field only, and the second term the scattered field only. Since we reconstruct the field in the source free region given by $a\leq r\leq r_0$, then Eq. (13) becomes

$$p(\mathbf{r})=4\pi i\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left[\frac{j_n(kr)h_n'(ka)-j_n'(ka)h_n(kr)}{h_n'(ka)}\right]\times h_n(kr_0)Y_n^m(\theta,\phi)Y_n^m(\theta_0,\phi_0)^*,$$

so that

$$p(\mathbf{r})=-4\pi\sum_{n=0}^{\infty}G_n(r,a,\omega)\frac{h_n(kr_0)}{(ka)^2h_n'(ka)}\sum_{m=-n}^{n}\times Y_n^m(\theta,\phi)Y_n^m(\theta_0,\phi_0)^*, \tag{14}$$

where $G_n(r,a,\omega)$ was given in Eq. (4). Based on Eq. (5), it is clear that the Fourier coefficients of the total pressure field are

$$P_{mn}=-4\pi\frac{h_n(kr_0)}{(ka)^2h_n'(ka)}Y_n^m(\theta_0,\phi_0)^*. \tag{15}$$

The velocity vector of the total field is easily determined by inserting Eq. (15) into Eq. (7). The intensity then follows from Eq. (8) with $N\to\infty$ and $F_n=1$.

The nature of the Fourier coefficients of Eq. (15) is illustrated in Fig. 2 for three different values of ka, plotted on a decibel scale. If we choose the point source to be on the polar axis, all the $m\neq 0$ harmonics are zero and we have

$$P_{0n}=-\sqrt{4\pi(2n+1)}\frac{h_n(kr_0)}{(ka)^2h_n'(ka)}. \tag{16}$$

This figure demonstrates the important fact that the Fourier coefficients decay rapidly when $n>ka$. Although we choose $r_0=5a$ in this example, larger values for r_0 exhibit the same break points at $n=ka$. This fact mimics the well know separation of plane waves into evanescent (exponentially decaying) and non-evanescent plane waves, the evanescent region corresponding to $n>ka$. In a fashion similar to planar NAH, these mimicked evanescent waves are amplified by the

growth of the propagator, with respect to r and n, that multiplies the Fourier coefficients.

VI. INTENSITY RECONSTRUCTIONS: NUMERICAL SIMULATIONS

We consider three spherical arrays consisting of 31, 64, and 100 microphones. We can derive the ideal locations of the M microphones on the array by using the solution to the so called Thomson problem discussed at the web site.[13] This site provides a construction algorithm by treating the microphones as equally charged particles with $1/r$ repulsive forces, separated so that the total energy of the system is minimized. The microphone positions appear as equally spaced on a sphere as a result of the algorithm. We derived the arrays for the three cases presented here in this manner. We also adjusted the radius of the computed arrays so that two cases are considered $a=0.13$ m and $a=0.20$ m.

In the following simulation, a point source is placed at 1.0 m from the origin and the exact total intensity field [$\vec{I}_{ex}$ is computed by inserting Eq. (16) into the reconstruction equations, Eqs. (5) and (7) leading to Eq. (8), where N was chosen to be sufficiently large for convergence of the summation is achieved]. For the simulation of the measured field we assume that the coefficients, Eq. (16), are not known and must be calculated using the approach given in Eq. (12). Furthermore, spatially incoherent, random Gaussian noise with a variance of σ^2 is added to the simulated measurement of the pressure. We construct an equally spaced cubical grid outside the sphere covering the region from $r=a$ to $r=0.4$ m for evaluation of the intensity vector. We then reconstruct the intensity vector using the equations of Sec. II on this grid for each of the three microphone arrays introduced, as well as the two radii.

It is instructive first to compute a base error by setting $\sigma=0$ and computing the resulting reconstructed intensity $\vec{I}$ using the equations in Sec. II, with no filtering. Critical in these equations is the value of N. N is determined by the number of microphones in the array which limits the number of Fourier coefficients one can determine via the Nyquist condition $(N+1)^2 \leq M$, discussed above. For $M=31,64,100$, this results in $N=4,7,9$, respectively. With these values of N, we compute the resulting errors for arrays of radius $a=0.13$ m and $a=0.2$ m using

$$\%\text{error} = 100\left(\frac{\sum (\vec{I}-\vec{I}_{ex})\cdot(\vec{I}-\vec{I}_{ex})}{\sum \vec{I}_{ex}\cdot\vec{I}_{ex}}\right)^{1/2} \tag{17}$$

where the sum is over all the lattice points (cubical lattice spaced 0.08 m) between $r=a$ and $r=0.4$ m. The base error is important because it presents the "best we can do" scenario.

Five cases are presented in Fig. 3 for various combinations of radius and microphone density, as indicated in the caption. We see that the error increases with frequency for all the cases. It appears that the 25% error points on the curves line up closely with the equation $kr_{max}=N$ (equivalent to $f=136N$), where $r_{max}=0.4$. This is not surprising given the discussion of Fig. 2. Thus increasing the number of microphones increases the frequency limit of the system. However,

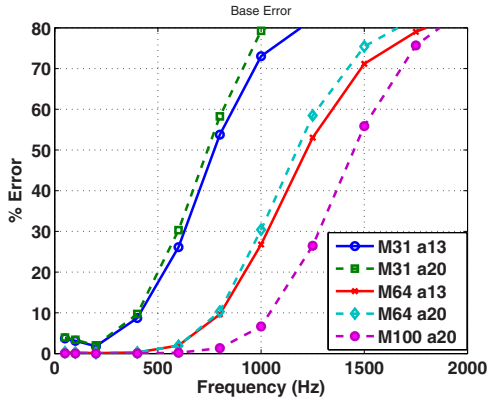


FIG. 3. (Color online) Error in the intensity vector averaged over space from $r=a$ to $r=0.4$ m for the "base" error (independent of noise). The number following M in the key is the number of microphones and the number following a represents the corresponding radius of the rigid sphere, that is, 0.13 and 0.20 m.

Fig. 3 shows that increasing the radius of the measurement sphere has little effect on the base error. We also conclude from Fig. 3 that a 31 element array can be expected to be fairly accurate (within 50%) up to about 800 Hz, a 64 element array to 1200 Hz and a 100 element array to about 1500 Hz. Final conclusions about this accuracy, however, must be made after a consideration of the influence of noise in the measurement.

The addition of spatially incoherent random noise ($\sigma \neq 0$) into the simulation poses a significant new challenge. If we look at the noise error for the reconstruction of the pressure field at a radius r given in Eq. (11) coupled with the fact that the propagator G_n behaves as indicated in Fig. 1, we conclude that as the limit of the summation N (or r for that matter) increases for a given σ the error also increases at the low frequencies before the first null. The maximum value of $N \equiv N_m$ is set by the Nyquist condition. However, if we choose to use a smaller value of $N<N_m$ below this first null the noise error will diminish, albeit at the sacrifice of an increase in the base error. Due to this increase in the base error, there is value of N for which the total error will begin to increase as N is reduced further. This value of N in which the total error is minimized is the optimum value, and it depends on frequency and r/a. Although this conclusion was based on the pressure reconstruction, the same holds true for the intensity reconstruction. We illustrate this in the following.

We set σ so that the SNR is 30 dB and reconstruct the total intensity using the equations of Sec. II for the same point source location and lattice used above. However, in this process we run through a set of all possible values of N, that is, $0 \leq N \leq N_m$, reconstructing the intensity at each value of N and choose the result with the least error, where the error is computed with Eq. (17) above. The results are shown in Fig. 4 with the total error on the left and the value of N corresponding to the minimum on the right. Notice that the increased errors in the high frequency regime reflect the behavior discussed for the base error in Fig. 3, and thus we attribute this increase to the base system error. We also can see clearly that the 100 element array has the lowest error

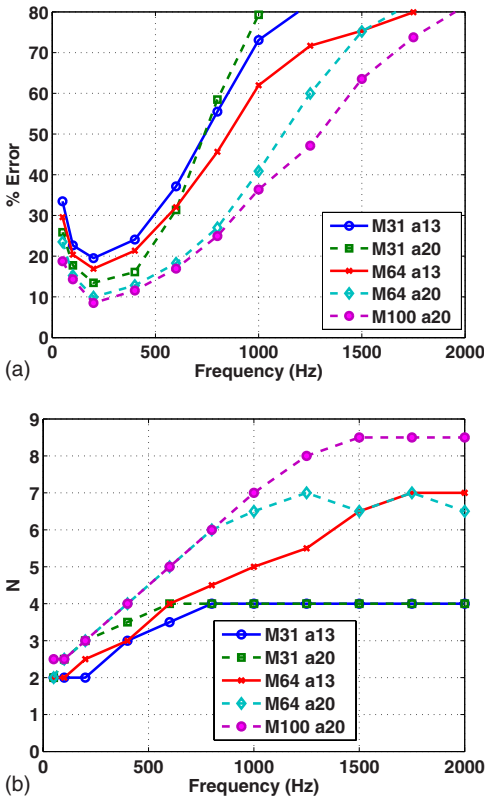


FIG. 4. (Color online) Total error in the intensity vector averaged over space from $r=a$ to $r=0.4$ m for the total error ("base" plus "noise" error terms) with a SNR of 30 dB. Key is the same as in Fig. 3. The plot on the right shows the value of N that reduced the total intensity error to a minimum.

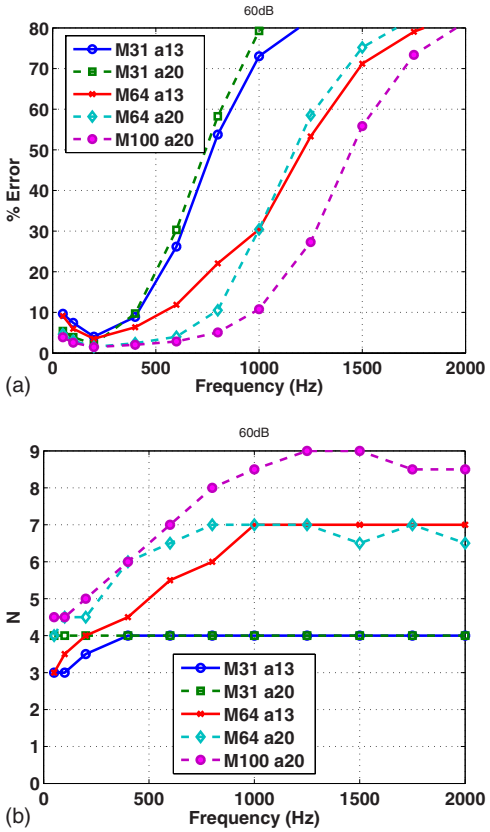


FIG. 5. (Color online) Same as Fig. 4 but with the SNR increased from 30 to 60 dB.

across all frequencies. Furthermore, it is evident that increasing the size of the 31 and 64 element arrays almost always improves the performance. The increase in error from 200 to 50 Hz for all the cases is somewhat of a surprise and we believe due to the fact that the divergence of the point source field is not well characterized by so few harmonics ($N=2$, that is, nine harmonics, for the 31 and 64 element cases, as the right-side plot indicates). We verified this by moving the source out to 10.0 m in which case the errors were much lower; the errors for $M=31$, $a=0.13$ m dropped to 9% at 50 Hz and increased monotonically above that.

In Fig. 5, the point source simulation is repeated with a 60 dB SNR and the minimum errors are displayed on the left plot with the value of N at the minimum error in the right plot. Clearly the errors at the low frequencies are much reduced, and the high frequency response now looks very much like the base error result of Fig. 3, again bearing out the fact that this region is controlled by the base error. It is also evident that the maximum values of $N=N_m$ occur at much lower frequencies, for example, the 31 microphone array of radius $a=0.13$ reached the maximum of $N=4$ at 400 Hz. Note that this implies that no filtering or regularization is needed above this frequency.

The careful reader may notice that some N values in the right plot of the last two figures are half-integers like $N=8.5$. This corresponds to a taper with respect to n in the regularization filter $F_n(\omega)$ of Eqs. (5) and (7), that is, the value of $F_n=1/2$ is used as a taper between $F_{n-1}=1$ and $F_{n+1}=0$. Thus, for example, the value of $N=2.5$ implies that the filter is set to one for $n\leq 2$, equal to 0.5 for $n=3$ and 0 for $n=4$. This process may seem odd at first, but we have borrowed from past experience with Tikhonov regularization filters used extensively in NAH. In this case the taper of the filter is specified allowing only values of 0, 1/2, or 1, but in Tikhonov regularization a more complex filter taper is computed automatically. We proceed to discuss regularization in a bit more detail in Sec. VII.

VII. REGULARIZATION

As we have discussed above, the pressure and velocity reconstructions given in Eqs. (5) and (7) present ill-posed problems at low frequencies when $r>a$ (as we expand the reconstruction sphere toward the sources) due to the amplification of the noise as was demonstrated for the pressure reconstruction in Eq. (11). Study of the propagators in Fig. 1 indicated that regularization is needed below the first null where the propagator increases with n and r although it may not be necessary at higher frequencies (for the range limit considered $r\leq 2a$). In Sec. VI, we presented a simple, *ad hoc*, regularization scheme in which we allowed the filter F_n to have three values, 0, 1/2, or 1, that depended on knowing the exact field. When the exact field is not known, an automated procedure is provided by Tikhonov regularization yielding a filter function depending on the propagator. We summarize it very briefly here and point the interested reader to references for more details.

The filter produced by Tikhonov regularization has the following form, now depending on a parameter α:

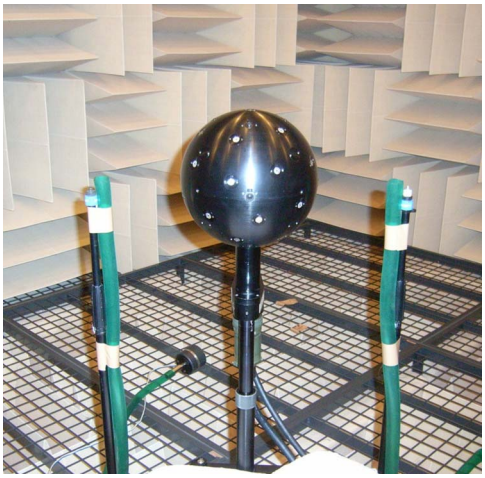


FIG. 6. (Color online) Rigid sphere of radius 0.13 m with 31 microphones of interest in this paper. The photograph also shows two tubes (green) acting as point sources inclined at a right angle relative to the center of the sphere, discussed below.

$$F_n^\alpha = \frac{1}{1 + g_n^\alpha (g_n^\alpha/(1+g_n^\alpha))^2}, \tag{18}$$

where $g_n^\alpha \equiv \alpha G_n'(r,a)/\rho c$, G_n' is the velocity propagator, and α is the regularization parameter to be determined from the Morozov discrepancy principle.[14] Note that as the velocity propagator becomes very large compared with $\rho c/\alpha$, the filter is close to zero, giving the desired filtering effect. To determine the value of α, an estimate of the noise in the measurement is required. We do this at low frequencies by assuming that the Fourier components for $n=N_m$ are below the noise. For example, for the 31 element array we would use $N_m=n=4$ below 500 Hz. As Hansen[15] pointed out, this estimate of the standard deviation σ of the noise is then (given $2n+1$ Fourier components in the noise) by

$$\sigma = \|P_{mn}\|/\sqrt{2N_m+1}, \quad n=N_m, \quad m=-N_m, \ldots, N_m. \tag{19}$$

As Ref. 14 describes, the Morozov discrepancy principle then returns a value of α to be used in the regularization filter.

VIII. EXPERIMENTAL RESULTS

To validate the feasibility of a 31 element array of radius 0.13 m for reconstructing the intensity vector in a volume of radius 0.4 m, we generate holograms using the point source formulation above [Eq. (14) with $r=a$], add noise, reconstruct the intensity, and compute the errors in the reconstructed versus exact intensity fields. The array of interest is shown in Fig. 6. We present examples of reconstructions at 200 and 600 Hz in which noise has been added to the hologram to achieve a SNR of 30 dB. The noisy holograms are then used in Eq. (12) to predict the Fourier coefficients with $N_m=4$. For the 200 Hz case, the SNR is then estimated (actual value is not used) using the $n=4$ harmonics and the filter of Eq. (18) determined by Eq. (19). For the 600 Hz case, $F_n^\alpha=1$ and no filter is used, since the problem is no longer ill-posed as discussed above. The filtered pressure and velocity vectors are computed over a fixed lattice cubical lattice (lattice spacing of 0.06 m) using Eqs. (5) and (7) and the intensity vector computed on this lattice from Eq. (8). On the left panels of Figs. 7 and 8 are the results at 200 and 600 Hz with the exact field shown on the right panel of each. The length of the vector is proportional to the intensity (W/m^2) at the base of the cone, and the cones are color coded also to show the magnitude of the intensity level. The point source is located outside the display window at $x_0=1$ m and $y_0=z_0=0$. At the bottom center of each figure is the Tikhonov filter values, F_n^α, for $n=0,1,\ldots,4$. For the 200 Hz case the errors are less than 30%, and the location of the point source can be easily inferred from the directions of the intensity vectors. It is clear that the errors have increased in the 600 Hz case (about 62% at $r=0.4$ m) although a curious effect has occurred. The reconstructed intensity drops off more rapidly circumferentially away from the source axis (the x axis) than the exact result. This error actually improves the ability to locate the source; the direction of the vectors still

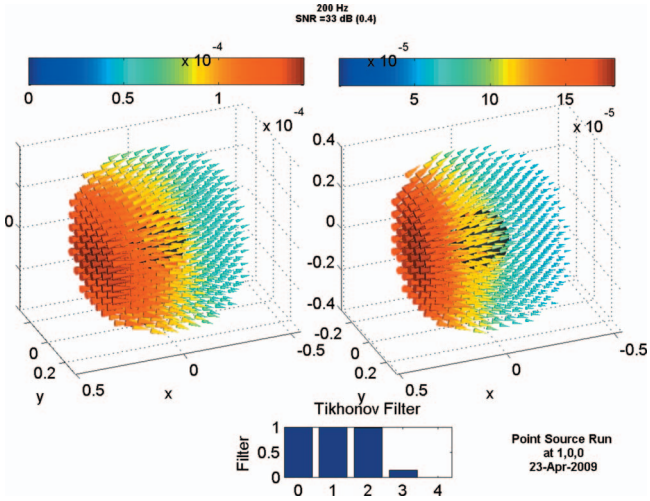


FIG. 7. Volumetric intensity reconstruction of the total field of a point source located on the x axis at $x=1$ m (to the left) from the sphere center, for 200 Hz with SNR=30 dB. The plot on the left is the filtered reconstruction and on the right the exact field. The Tikhonov filter is shown at bottom center.

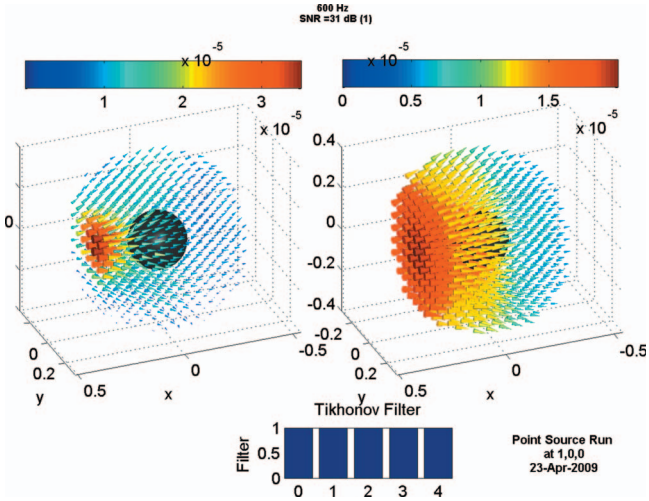


FIG. 8. Same as Fig. 7 except at 600 Hz. Exact field is on the right.

point away from the point source. We have no explanation for this effect, but it is triggered when there are not enough Fourier components (we are limited to $n \leq 4$) to accurately reconstruct the field. This corresponds to the frequency regime where the base error dominates. This effect occurs at all of the higher frequencies as well. The Tikhonov filter in Fig. 7 indicates that only the $n=0,1,2$ components were unfiltered for the 200 Hz case, whereas in the 600 Hz case no filter was used. We now discuss laboratory experiments to demonstrate the utility of the vector intensity reconstructions.

A. Two open tube sources

Two point sources generated by two long tubes at the end of separate vibration generators were arranged as shown in Fig. 6 above. The measurements were carried out at in the anechoic chamber at Nittobo Acoustic Engineering Corporation using a 31 element sphere of radius 0.13 m. The intensity vectors were computed on a rectangular lattice of 1331 points with equal spacing of 0.06 m extending over a cubical volume 0.6 m on a side. The two sources were at right angles to the center of the sphere and were located a distance of 0.28 m from the sphere center. They were driven independently (uncorrelated) with random noise that was recorded to use as references. The cross spectral densities between the 31 microphones and the two reference signals were computed using ensemble averaging using a 1024 point fast Fourier transform (FFT) and 58 overlapping ensembles. With a sample rate of 12 kHz, each ensemble consisted of 0.085 s of data. Using the partial field decomposition technique,[3,16,17] holograms were produced correlated to each source at selected frequencies. The two holograms were then processed separately and the vector intensity fields displayed. We show results at two frequencies below, 200 and 600 Hz.

Figure 9 shows the computed intensity field from the partial field holograms of the left and right tube sources and Fig. 10 shows the resulting Tikhonov filters based on noise calculated from the nine $n=4$ harmonics ($m=-4,-3,\ldots,4$) (using the Morozov discrepancy principle to determine the value of α in the filter). In Fig. 6, the orientation of the photograph is the same as the orientation of the intensity displays. The filters for the left and right sources are slightly different due to a difference in the signal-to-noise ratios for each due to unequal drive levels. We can see that the approach has separated the two sources, and the contributions and intensity maps of each determined. The ability to separate multiple sources is a hallmark of the partial field decomposition technique discussed in the references.

In Fig. 11, the reconstructions at 600 Hz are shown. In this case no Tikhonov filter is necessary, as discussed above. Two holograms were generated, each correlated to the reference driver at the end of each tube. Again the intensity field reconstructed from these holograms emanates from the direction of each source. We observe the same effect noted for the point source of Fig. 8 in that the intensity level drops rapidly in the circumferential direction away from the point of maximum intensity. We would expect that with more harmonics, our images would appear more like the one on the right panel

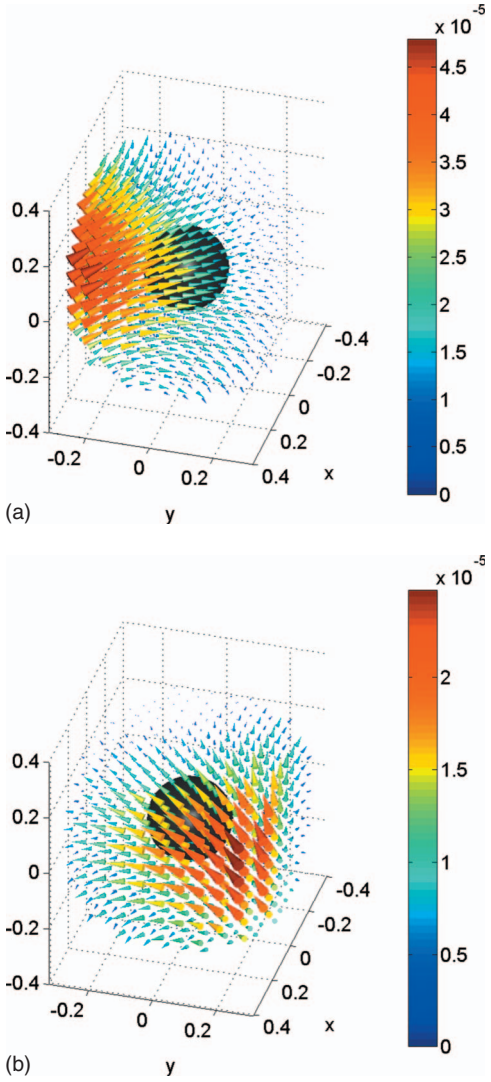


FIG. 9. Volumetric intensity from partial field hologram related to the reference signal of the left tube driver (left panel) and of the right tube driver (right panel). Consistent with the vector displays, the orientation of the sphere and sources are shown in the photograph of the experiment in Fig. 6.

of Fig. 8. However, to measure more harmonics, the number of microphones in the array must be increased as we have discussed above. Again this anomaly appears to help with the determination of the direction of the actual source, even though some accuracy with respect to source level has been sacrificed. The same anomaly is apparent up to 1000 Hz.

B. Application inside an automobile body

One of the most promising applications is for interior noise diagnostics inside automobiles, aircraft, etc. The 31 element rigid sphere discussed in Sec. VIII A was placed at the operator location inside an automobile driven on a dynamometer at 3000 rpm and a reference accelerometer was placed on the engine block. Partial reference holography was used to extract the hologram related to the fundamental of the engine tonal, at 105 Hz, and the measured pressure hologram correlated very highly to the engine block accelerometer. The volumetric acoustic intensity was imaged from this hologram, as described above on a cubic grid with lattice spacing 0.06 m. Figure 12 shows two views of the resulting

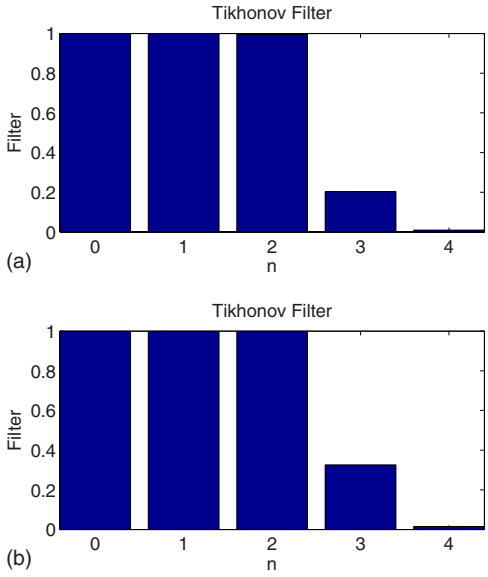


FIG. 10. (Color online) The Tikhonov filter coefficients for the left and right reconstructions of Fig. 9, respectively.

intensity field with an outline of the quarter section of the automobile including driver-side window and windshield framing the display (looking from the back seat on the left, and into the driver door on the right). The mean sound pressure level over the sphere microphones was about 89 dBA at 105 Hz. In the figure, it can be seen that the intensity originates from the rear, coming from the left side of the back seat. The hemisphere plots in the figure indicate the radial intensity on an imaginary spherical surface of radius 0.3 m, red entering the sphere and blue leaving the sphere in units $\mu W/cm^2$. The seat-side hemisphere is shown on the left and the windshield side hemisphere is shown on the right oriented aligned with the photo on the extreme right as well as the top left vector display. On the left hemisphere, one can see the large influx of power (red area) from the left of the back seat of the cabin. The right hemisphere also indicates an inward power flow at a lower level coming from the windshield (orange area). It also shows a fairly large outflow of power (deep blue) in an area near the left under-dash and floor area. It is very difficult to create comprehensive yet intelligible displays of the vector fields. Rotating the graphics display is very helpful and ultimately a stereographic display would be ideal. One final note on the resolution is obtained with the vector display. At 105 Hz, the acoustic wavelength is over 3.2 m, and the vector display covers a cube 0.6 m on a side which corresponds to about $1/5$ of a wavelength. Clearly the displayed field presents a spatial resolution much better than half-wavelength. The Tikhonov filter values F_n^α are displayed at the bottom of the plot.

As a final demonstration of the potential of the theory, we open up the Tikhonov filter beyond its "correct" settings to increase the spherical harmonic content (adding most of the $n=3$ component) for the extracted hologram and repeat the intensity calculation for 105 Hz. The results in Fig. 13 show that increasing the harmonic contributions provides a more concentrated intensity yielding a more precise indication of the direction of the interior noise. The source now appears to be near the back door as opposed to the rear seat as conjectured above.

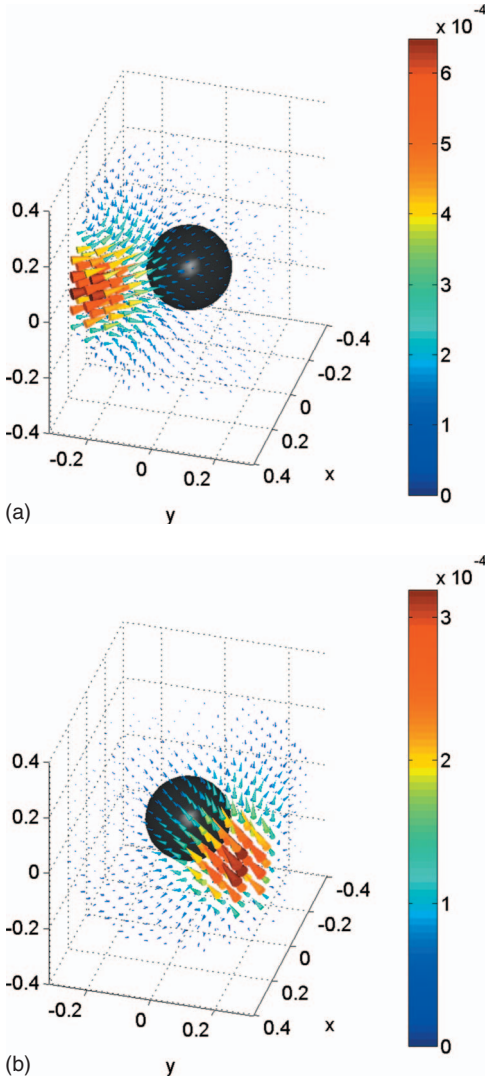


FIG. 11. Volumetric intensity from partial field holograms at 600 Hz, related to the left tube reference signal of the left and right tube reference on the right panel.

IX. CONCLUSIONS

This paper considered both the theory and experimental implementation of a vector intensity mapping device based on the reconstruction techniques of NAH. A rather detailed error analysis was provided so that the results could be interpreted. Intended for use in interior spaces such as an aircraft, we exercised the theory by using numerical holograms generated from a point source intended to be a very simple model of an interior boundary source. Arrays of two different sizes and three different microphone densities were considered. The locations of the microphones embedded in the sphere were "equally" spaced by treating the microphones as equally charged particles with $1/r$ repulsive forces. We found that increasing the number of microphones in the array increased the upper frequency limit for successful reconstructions, essentially doubling the range from 1 to 2 kHz

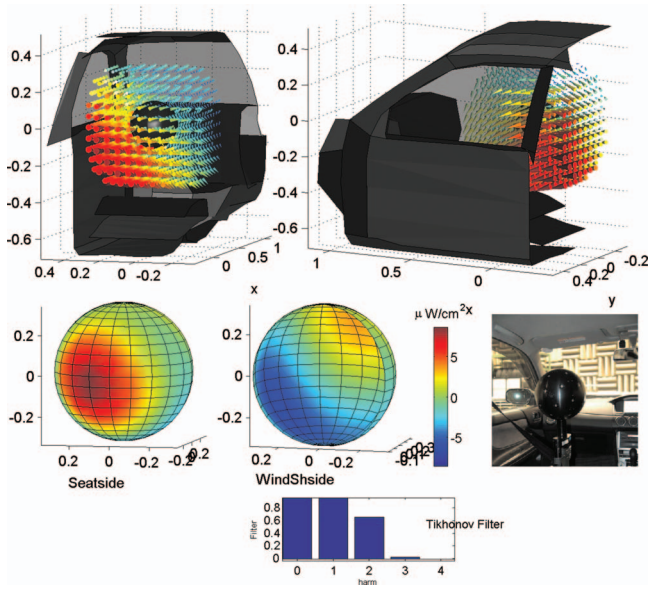


FIG. 12. Volumetric intensity correlated with an engine block accelerometer at the fundamental frequency (105 Hz) of the engine tonals.

when going from 31 to 100 microphones. The reconstruction space was limited to a sphere of radius 0.4 m which was three times the smaller array and twice the larger array considered in this study. We do not feel that this restriction is significant since the device is intended to work in a closed space, and the theory is only valid in the spherical volume that does not cross any boundary. A theoretical analysis of the propagators indicated that reconstruction in this volume is an ill-posed problem requiring regularization below a certain frequency, the latter increases as the number of microphones in the array increases. With this regularization, a high density array was shown to be more accurate at low frequencies at a typical SNR level than the less dense array. This regularization may seem formidable at first glance, but as the

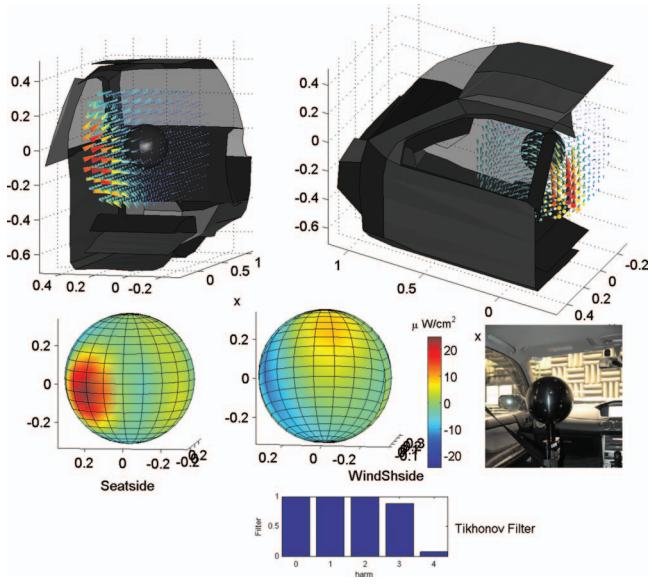


FIG. 13. Same as Fig. 12 except the estimate of the SNR has been increased by 6 dB to increase the number of harmonics used in the reconstruction of the volumetric intensity.

ad hoc filter used in Sec. VI indicated one can obtain excellent results with a predefined filter. Although we did not show the result, when too many harmonics are included, the directions of the intensity vector in the volume are nonsensical.

An earlier paper[3] considered an open sphere consisting of 50 microphones, and the conclusions drawn here are very consistent with those of the open sphere. (That work also used a point source at 1.0 m range). The open sphere also used a predefined filter set, although it required an estimate of the SNR to choose it. Both open and rigid spheres represent ill-posed problems at low frequencies, but not at higher frequencies. However, the rigid sphere does not have the interior eigenvalue problem of the open sphere due to the difference in propagators which should lead to slightly more accurate results. [In the open sphere, the pressure and velocity propagators contained the term $j_n(ka)$ in the denominator which has a set of zeros causing the propagator to blow up at these eigenvalues]. Furthermore, the sound diffraction around the rigid sphere is accounted for in the formulation, whereas diffraction due to the structure of the open sphere cannot be removed in the formulation. Certainly the interior of the rigid sphere can be used to house all the cabling and much of the electronics without effecting the operation, not possible in the open sphere.

One drawback is the fact that a rigid sphere placed near a boundary might change the impedance of the boundary and thus alter the intensity flow pattern, especially since the normal component of the intensity must vanish at the sphere surface. This does not appear to be a question that is easy to answer and should be the subject of future research.

When linked with commercially available hardware, our method to reconstruct the vector intensity provides real-time imaging of the acoustic intensity vector in a volume surrounding the array. Since the spatial map of the intensity vector points away from the sources, it can be used to identify the location and strength of the sources. This approach is ideally suited for noise control identification during operation in interior spaces such as automotive cabins as we have demonstrated and aircraft interiors, the latter studied in the open sphere investigations.

ACKNOWLEDGMENT

This work was supported by the U.S. Office of Naval Research.

[1]E. G. Williams and K. Takashima, "Vector intensity reconstructions in a volume surrounding a rigid spherical measurement array," in Acoustics'08 Paris (2008), Vol. **123**, p. 3309.
[2]K. Takashima, H. Nakagawa, and E. G. Williams, "Vector intensity measurement with a rigid spherical microphone array in a vehicle cabin," in Acoustics'08 Paris (2008), Vol. **123**, p. 3312.
[3]E. G. Williams, N. P. Valdivia, P. C. Herdic, and J. Kloss, "Volumetric acoustic vector intensity imager," J. Acoust. Soc. Am. **120**, 1887–1897 (2006).
[4]J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in IEEE ICASSP'02 (2002), Vol. **2**, pp. 1781–1784.
[5]Z. Li and R. Duraiswami, "A robust and self-reconfigurable design of spherical microhone array for multi-resolution beamforming," in IEEE ICASSP'05 (2005), Vol. **4**, pp. 1137–1140.
[6]M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposi-

tion using spherical microphone array," J. Acoust. Soc. Am. **118**, 3094–3103 (2005).
[7]Boaz Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," J. Acoust. Soc. Am. **116**, 2149–2157 (2004).
[8]Z. Li and R. Duraiswami, "Flexible and optimal design of spherical microphone arrays for beamforming," IEEE Trans. Audio, Speech, Lang. Process. **15**, 702–714 (2007).
[9]B. Rafaely, "Analysis and design of spherical microphone arrays," IEEE Trans. Audio, Speech, Lang. Process. **13**, 135–143 (2005).
[10]A. Meyer and D. Döbler, "Noise source localization within a car interior using 3d-microphone arrays," in Berlin Beamforming Conference (BeBeC 2006) (2006), pp. 1–7.
[11]E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, London, 1999).
[12]*Electromagnetic and Acoustic Scattering by Simple Shapes*, edited by J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi (Hemisphere, New York, 1987).
[13]http://physics.syr.edu/condensedmatter/thomson/thomsonapplet.htm (Last viewed 11/22/09).
[14]E. G. Williams, "Regularization methods for near-field acoustical holography," J. Acoust. Soc. Am. **110**, 1976–1988 (2001).
[15]P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems* (Siam, Philadelphia, PA, 1998).
[16]J. Hald, "STSF—A unique technique for scan-based nearfield acoustical holography without restriction on coherence," Technical Report No. 1, from Bruel & Kjaer, Naerum, Denmark, 1989.
[17]D. L. Hallman and J. S. Bolton, "Multi-reference nearfield acoustical holography," in Proceedings Inter-Noise '92, (1992), pp. 1165–1170.

Prediction of the acoustical performance of enclosures using a hybrid statistical energy analysis: Image source model

Franck Sgard[a] and Hugues Nelisse
Institut de Recherche Robert-Sauvé en Santé et Sécurité au Travail (IRSST), 505, Boulevard de Maisonneuve Ouest, Montréal Québec H3A 3C2, Canada

Noureddine Atalla, Celse Kafui Amedin, and Remy Oddo
Department of Mechanical Engineering, GAUS, University of Sherbrooke, 2500 Boulevard de l'Université, Sherbrooke QC J1K2R1, Canada

(Received 3 March 2009; revised 11 November 2009; accepted 19 November 2009)

Enclosures are commonly used to reduce the sound exposure of workers to the noise radiated by machinery. Some acoustic predictive tools ranging from simple analytical tools to sophisticated numerical deterministic models are available to estimate the enclosure acoustical performance. However, simple analytical models are usually valid in limited frequency ranges because of underlying assumptions whereas numerical models are commonly limited to low frequencies. This paper presents a general and simple model for predicting the acoustic performance of large free-standing enclosures which is capable of taking into account the complexity of the enclosure configuration and covering a large frequency range. It is based on the statistical energy analysis (SEA) framework. The sound field inside the enclosure is calculated using the method of image sources. Sound transmission across the various elements of the enclosure is considered in the SEA formalism. The model is evaluated by comparison with existing methods and experimental results. The effect of several parameters such as enclosure geometry, panel materials, presence of noise control treatments, location of the source inside the enclosure, and presence of an opening has been investigated. The comparisons between the model and the experimental results show a good agreement for most of the tested configurations.
 [DOI: 10.1121/1.3273892]

PACS number(s): 43.50.Gf, 43.40.At [KVH] Pages: 784–795

I. INTRODUCTION

Enclosures are a common noise control solution to reduce the airborne sound radiated by a noisy machine. They usually consist of a box surrounding the noise source that is made up of an assembly of multi-layered structures [metal or wood skin which can be stiffened, a sound absorbing material (e.g., mineral wool) and a perforated metal plate or an impervious screen]. Panels may contain openings which might be designed purposely (e.g., conveyor belts) or leaks which are the results of bad assembly and/or mounting conditions. The enclosure can be placed very close to the source in which case they are denoted as "close-fitting" enclosure. The acoustic performance of an enclosure is generally quantified by its insertion loss which is the difference between the sound power of the source in free field condition and the sound power transmitted by the enclosure surrounding it. The insertion loss depends on multiple factors such as sound transmission loss of each component, absorbing materials inside the enclosure, presence of leaks and apertures, coupling between the machine, the ground and the enclosure walls, and position and extent of the machine.

Extensive work has been done on the modeling of acoustic enclosures. Through the years, several authors have developed simplified analytical models for predicting the performance of enclosures for an excitation by an internal sound source.[1–5] These works mostly focus on close-fitting enclosures by considering the problem of a vibrating source panel coupled to a radiating panel (enclosure wall) through an air gap. The limitations of these models lie either in the infinite plate model,[1–3] unrealistic modeling of the source,[1–4] or restriction to low frequencies.[5] Besides, Tweed and Tree[6] reviewed Jackson's and Junger's work[1,2,4] and concluded that their models were not reliable for prediction of the enclosure efficiency. Byrne *et al.*[7] proposed an analytical approach for the case where the enclosure is directly attached to the machine by connectors (machine-mounted type). More recently, Hasheminejad and Mehdizadeh [8] investigated convex-shaped panels and developed a closed form solution for the insertion loss of an airtight close-fitting multi-layered hemispherical enclosure containing porous and viscoelastic materials based on appropriate field expansions. Other models based on modal expansions of the fields in the panels and the cavity or Fourier transform methods have been proposed to investigate the vibroacoustic response of a vibrating panel subjected to an external acoustic excitation and coupled to a cavity.[5,9–15] Numerical methods such as finite element method (FEM), mixed FEM/boundary element method (BEM) or integrogeometrical methods have also been used to predict the low frequency insertion loss of close-fitting enclosures.[16] The vibroacoustic response of a plate backed cavity subjected to mechanical and acoustical excitation (external or internal) with[17–19] or without[20] sound absorbing treatment or the vibroacoustic response of a flexible box treated[21–23] or untreated[24,25] have also been investigated with

[a]Author to whom correspondence should be addressed. Electronic mail: frasga@irsst.qc.ca

these numerical techniques. These works relying on deterministic methods are, however, limited to low frequencies, are usually time consuming for day-to-day design purposes, and may require meshing expertise and capabilities which restrain their domain of application. For higher frequencies, energy based methods are commonly utilized. Simplified models based on the calculation of the sound power radiated from the enclosure through the various transmission paths (structural, leaks, etc.) have been developed to design enclosures.[26–28] These models rely generally either on the measured panel sound transmission losses or the calculated sound transmission losses, e.g., using textbook analytical or empirical formulas or transfer matrix method (TMM).[29] Comparisons between the theory using diffuse field calculated sound transmission loss and experimental results indicate that overall insertion loss trends can be correctly captured especially in the case of air-tight enclosures made up of high critical frequency single walls with or without acoustical treatment but may lead to errors in some frequency bands for walls with critical frequency lying in the calculation spectrum and for double walls.[27] Other energy based techniques such as ray tracing, mixed BEM-ray tracing, and energy flow method allow for the calculation of the sound transmission through flexible panels and can be used to predict the insertion loss of machinery enclosures. Ray-tracing method is classically used in architectural acoustics where only the room internal sound field is of interest. Walls are then assumed not to vibrate. Barbry and Trompette[30] currently extended the ray-tracing technique by considering the rays transmitted through the enclosure panels in order to calculate its sound insertion loss. Jean[23] combined the ray-tracing technique to assess the Green's function and the BEM to predict the internal and external sound fields in cavities with flexible surfaces. Le Bot[31] and Cotoni *et al.*[32] developed an integral energy approach to predict the sound field inside an enclosure and for fluid-structure interaction problems, respectively. Several authors also proposed statistical energy analysis (SEA) approaches for assessing the insertion loss of enclosures for internal sound excitation[5,33,34] or placed into an external diffuse field.[35–37] Ver,[33] Cole *et al.*,[37] and Ming and Pan[34] included both resonant and non-resonant transmission paths in their analysis whereas Lyon,[35] Eichler,[5,36] Oldham and Hillarby[5,36] only considered resonant transmission. In the most recent work, Ming and Pan[34] proposed actually two different SEA based models accounting for the structural coupling between the walls and depending on the importance of the cavity modal density. These authors show that comparisons between the models and experiments are encouraging. They insist on the fact that absorbing acoustical materials increase the damping in the cavity but also modify the non-resonant transmission through enclosure panels. However, in their work, acoustic treatments are supposed to have only a mass effect and no rigorous model of their effect on the non-resonant transmission is considered. Note that the discrepancies between the simulation and the experimental results for the treated enclosure in Ming and Pan's work[34] could also be partly explained by the fact that the internal sound field is not really diffuse. In a real enclosure, the diffuse field assumption for the sound field is indeed likely not to hold because of the sound absorbing treatment. Additionally, the SEA model proposed by Ming and Pan[34] does not account for the source location inside the enclosure which may have an impact on the insertion loss of the system especially in the presence of apertures.

This work presents some of the results of a recent research project whose objective was to develop a general, simple, and efficient acoustical SEA-based design tool for machinery enclosures.[38] The model is deemed simple in the sense that it does not require extensive pre- and post-processing work as in standard numerical techniques (FEM). Close-fitting enclosures are not considered in this study. Rather, one considers large free-standing enclosures for which the volume of the source is much smaller than the volume of the enclosure. One specific improvement is to relax some of the limitations associated with classical SEA. It consists of using a hybrid SEA-image sources approach. First, the SEA cavity subsystem is partitioned into subcavities. The image sources method is then used to calculate the internal sound pressure field in each subcavity which is then considered as an equivalent source of pressure in the SEA equations. There are several originalities to the present approach: (i) use of the latest developments for the vibroacoustic behavior of multi-layered acoustical materials based on transfer matrix approach to avoid measurement of the associated insertion loss, sound transmission loss of panels, and absorption; (ii) a new model for diffused field sound transmission loss through apertures is considered; (iii) the image source method is combined with SEA to improve the quality of results in the whole frequency range. To evaluate the model, two types of enclosures (parallelepipedic and L-shaped geometries) have been built. The effects of several parameters on the acoustical performance of the enclosures, such as the noise source location, the presence of an opening, the panel nature and the presence of acoustical treatment, are investigated. The insertion losses of 26 different configurations were measured using the intensity scan method in a semi-anechoic room. The sound field inside the enclosures was also assessed using a set of microphones. The model predictions for the internal sound field are then compared to other existing models. Experimental results are discussed and compared to the predictions.

This paper is organized as follows. First, the theoretical model is expounded. Second, the experimental set-up together with the different configurations of interest are described. Third, the image source model prediction of the internal sound field is evaluated by comparisons with several other existing numerical, semi-analytical, or geometrical approaches. Fourth, the SEA-based models are evaluated by comparisons between the calculations and the experiments. Finally, the main findings and prospects of this work are summarized in the Conclusion.

II. THEORY

The SEA general framework is used to solve the global problem. In the SEA formalism, geometrical (panels, acoustic cavities, openings/apertures, sound treatments, etc.) and physical complexities (multi-layer subsystems, structural couplings between panels, structure borne excitation, etc.)

can be added in a natural and relatively simple way (assuming that most SEA assumptions are met). In the present work, two types of enclosures are considered: They consist of box or L-shaped assemblies of elastic homogeneous panels of uniform thickness treated with sound absorbing materials protected by a perforated plate. Each panel may contain one or more openings. Only airborne excitations are considered so the structural connections at panel boundaries are neglected. It is considered that each sound source has dimensions smaller than the acoustic wavelength (point sources). Each source is considered to generate acoustic power in the enclosure and the transmitted power through each part of the enclosure is then predicted to obtain the insertion loss.

A. Sea model

The enclosure is modeled as an assembly of separate structural SEA subsystems (panels) enclosing one (or more) air cavity containing the sound source(s). To calculate the exterior power radiated by the enclosure, a large cavity surrounding the structure is used. The dimensions and damping of this acoustic space are set in order to simulate sound radiation in open space so there is negligible power flowing from this external cavity to the enclosure panels.

1. General equations

The SEA equations are obtained by writing the power flow balances for all the subsystems which leads to

$$\omega[\Upsilon]\{E\} = \{\Pi\}, \tag{1}$$

where ω is the circular frequency, $\{E\}$ is the vector of total energies stored in each subsystem, $\{\Pi\}$ is the vector of injected powers in each subsystem, and $[\Upsilon]$ is the loss factor matrix. All quantities are averaged over both space and frequency (1/3 octave bands).

The matrix of loss factors coefficients is given by

$$\Upsilon_{ii} = \eta_{ii} + \sum_{\substack{j=1 \\ j\neq 1}}^{N} \eta_{ij},$$

$$\Upsilon_{ij} = -\eta_{ji} \quad \text{if } i \neq j, \tag{2}$$

where η_{ii} is the damping loss factor (DLF) of subsystem i and η_{ij} denotes the coupling loss factor (CLF) between subsystems i and j. CLF η_{ji} is calculated in terms of η_{ij} using the reciprocity equation $n_i\eta_{ij}=n_j\eta_{ji}$, where n_i is the modal density of subsystem i.

Since the enclosure is made of flat rectangular panels and regular cavities, classical relations are used to compute the modal densities.[39] For the sake of simplicity, only flexural waves are retained for structural subsystems in the analysis, a fairly valid assumption for thin plates and airborne transmission. If needed, adding structural connections and/or additional wave type can be done by using the appropriate CLF and DLF and modal densities relations (see, for example, Ref. 39).

Once loss factors are calculated [Eq. (2)] and substituted into Eq. (1), the system of equations is solved to get subsystem total energies. These energies are related to the physical variables by

$$E = \begin{cases} M\langle v^2\rangle & \text{for a structural subsystem (panel)} \\ \dfrac{\langle p^2\rangle V}{\rho_0 c_0^2} & \text{for an acoustical subsystem (cavity),} \end{cases} \tag{3}$$

where $\langle v^2\rangle$ denotes the mean square velocity of the panel, M its mass, and $\langle p^2\rangle$ the mean square pressure inside the cavity of volume V filled with a fluid of speed of sound c_0 and density ρ_0. Once the linear system has been solved, the energy vector is used to compute the input power in the exterior cavity and thus the total acoustic power radiated by the enclosure.

2. Damping loss factors

The damping loss factor of flexural waves for the structural subsystems (panels) can be measured by classical techniques. Bolduc[40] provided an extensive review of the available techniques. The damping loss factor of the cavities can be assessed by

$$\eta_{ii} = \frac{c_0 S_i \bar{\alpha}_i}{4\omega V_i}, \tag{4}$$

where V_i and S_i are, respectively, the volume of cavity i and the area of the surface enclosing cavity i, and $\bar{\alpha}_i$ is the average diffuse field sound absorption coefficient of cavity i. The coefficient $\bar{\alpha}_i$ is calculated using

$$\bar{\alpha}_i = \frac{\sum \alpha_{i,j}^d S_{i,j}}{\sum S_{i,j}}, \tag{5}$$

where $S_i=\Sigma_j S_{i,j}$ and $S_{i,j}$ and $\alpha_{i,j}^d$ are, respectively, the area and the diffuse field sound absorption coefficient of the jth wall in the ith cavity. This coefficient can be calculated using the following equation:

$$\alpha_{i,j}^d = \frac{\int_0^{\theta_{\lim}} \alpha_{i,j}(\theta_i)\sin\theta_i\cos\theta_i d\theta_i}{\pi \sin^2\theta_{\lim}}, \tag{6}$$

where $\alpha_{i,j}(\theta_i)$ is the plane wave sound absorption coefficient for oblique incidence θ_i. If $\alpha_{i,j}(\theta_i)$ can be assessed experimentally or numerically, $\alpha_{i,j}^d$ is then called statistical sound absorption coefficient. In practice, this coefficient is assessed in reverberant rooms using reverberation time measurements with and without material sample. One then refers to Sabine's sound absorption coefficient if Sabine's formula is used to assess it. Sabine's sound absorption coefficient can also be approximated using transfer matrix approach accounting for the finite lateral extent of the material.[41] In Eq. (6), limit angle $\theta_{\lim}$ is commonly chosen equal to 78° rather than 90°. Sound absorption in cavity is modeled as a combination of wall absorption as well as air absorption. The coefficient $\bar{\alpha}_i$ is then written as $\bar{\alpha}_{wi}+(4mV_i/S_i)$, where $\bar{\alpha}_{wi}$ is the walls average sound absorption coefficient of the ith cavity and m is an energy attenuation constant of sound waves propagating in air. This value is a function of frequency, temperature, and relative humidity. Tabulated values exist[33] or an analytical formula[42] is available to calculate this value.

3. Coupling loss factors

The power exchange between resonant modes of a panel i and those of a cavity j, called a resonant path, is characterized by the resonant coupling loss factor η_{ij}^{res} given by

$$\eta_{ij}^{\text{res}} = \frac{\rho_0 c_0 \sigma_p}{\rho_s \omega}\left(\frac{S_{\text{coup}}}{S_p}\right), \tag{7}$$

in which ρ_0 and c_0 are the air density and sound speed in the cavity, respectively, σ_p is the radiation efficiency of a panel of area S_p, ρ_s is the panel surface mass density, and S_{coup} is the coupling surface between panel i and cavity j. In this work, the radiation efficiency σ_p is calculated using the formulations developed by Leppington *et al.*[43] which are valid for a simply supported baffled panel. At low frequencies σ_p can be corrected[39,44] to take into account the fact that the panel is not embedded into a plane rigid baffle and that the boundary conditions may be different from simply supported (say, e.g., clamped). Indeed, the real boundary conditions of the panel can be expected to be somewhere between simply supported and clamped. In addition, panels in the enclosure radiate into a space delimited by plates at right angle (90° on the source side and 270° on the external side) which can affect significantly the radiation efficiency at low frequencies. In order to model both boundary conditions and baffle effects, Craik[39] introduced a correction factor under the critical frequency given by

$$C_\sigma = m_0 - (m_0 - 1)\left[\frac{f}{f_c}\right]^4, \tag{8}$$

where m_0 is the desired correction (for example, $m_0=2$ for perfect clamping boundary conditions). Radiation efficiency is thus corrected with factor C_σ ($\sigma_p \rightarrow C_\sigma \sigma_p$) under the critical frequency, this correction factor being equal to 1 at (and above) the critical frequency of the panel.

The direct acoustic path between two cavities i and j is usually (and sometimes erroneously) called the non-resonant path and is taken into account using the coupling loss factor calculated from

$$\eta_{ij}^{\text{nres}} = \frac{c_0 S_{\text{coup}} \tau_p}{4\omega V_i}, \tag{9}$$

where S_{coup} is the coupling area between the two cavities and τ_p is the diffuse field sound transmission coefficient of the infinite panel (or aperture) separating the two cavities.

$$\tau_p = \frac{\int_0^{\theta_{\text{lim}}} \tau(\theta_i)\sin\theta_i \cos\theta_i d\theta_i}{\pi \sin^2\theta_{\text{lim}}}, \tag{10}$$

where $\tau(\theta_i)$ is the oblique incidence plane (θ_i is the angle between the normal to the panel and the direction of propagation of the plane wave) wave sound transmission coefficient (mass law part) of the panel. In the general case, Eq. (10) can be calculated numerically using Gauss–Kronrod quadrature algorithm. For a homogeneous isotropic panel, $\tau(\theta_i)$ is given by

$$\tau(\theta_i) = \left(\frac{m\omega\cos\theta_i}{2\rho_0 c_0}\right)^2, \tag{11}$$

where m denotes panel surface density. Then

$$\tau_p = -\frac{1}{16\pi}\left(\frac{m\omega}{\rho_0 c_0 \sin\theta_{\text{lim}}}\right)^2 (\cos^4\theta_{\text{lim}} - 1). \tag{12}$$

When a panel is covered with a sound absorbing treatment, four different quantities can be affected: (i) damping (absorption) in the cavity in which the treatment appears, (ii) sound transmission through the panel, (iii) damping of the panel, and (iv) added mass and stiffness on the panel.

The first aspect is dealt with by using the absorption coefficient of the treatment in the DLF calculation of the cavity [see Eq. (4)]. The additional panel damping brought by the treatment can be assessed using either damping measurements or TMM predictions. In the present study, measurements with unassembled panels were carried out to obtain the panel damping. The potential modification of the dynamic properties of the panels (mass and stiffness) due to the presence of the sound treatment can also in theory be assessed experimentally or using the TMM. However, in the present work, it is assumed that the sound absorbing materials that were used are lightweight and soft and have negligible effect on the panel mass and stiffness.

For the transmission part, the sound treatment can affect both the resonant and the non-resonant paths. Therefore, corrections are made on the corresponding CLF to take into account these effects. For the resonant path, the correction is made by modifying the panel radiation efficiency σ_p appearing in Eq. (7) using

$$\sigma_{p+\text{treat}} = \sigma_p \times 10^{-\text{IL}/10}, \tag{13}$$

where IL is the sound treatment insertion loss defined as the difference between the transmission loss of the panel with and without treatment. The insertion loss is calculated using TMM.

For the non-resonant path, the correction is made by modifying the transmission coefficient referred to in Eq. (9) using

$$\tau_{p+\text{treat}} = \tau_p \times 10^{-\text{IL}/10}. \tag{14}$$

It is worth mentioning that the same insertion loss is used to evaluate the correction for both the resonant and non-resonant paths. It has been shown that this is a reasonable assumption.[45]

Finally, an aperture (opening or leak) is modeled as an additional non-resonant path between the two acoustic spaces on both sides of the aperture. The associated coupling loss factor is given by Eq. (9), where S_{coup} is the aperture area and τ_d corresponds now to the diffuse field transmission coefficient of the opening. The latter is predicted using the modal approach recently proposed by Sgard *et al.*[46] Note that the total coupling non-resonant loss factor of the panel with the opening is simply the sum of the two paths where the surface of the panel has to be diminished by the aperture area.

B. Modeling of the internal sound field

In order to better describe spatial variations of the exciting sound field inside the enclosure, a strategy based on the image source method is employed. The interior sound field is calculated at a given number of points inside the enclosure using the image source method. This information is thereafter used as the excitation field on the different panels (structural SEA subsystems) of the enclosure. To integrate this excitation in the SEA model, the interior cavity is first partitioned into subcavities. These cavities are then used as SEA acoustic subsystems with no power flow between them (no acoustic coupling between the cavities) but coupled to the panels. The sound pressure level (SPL) predictions obtained with the image source method are thus spatially averaged and then used as SPL input data (pressure constraint) in the corresponding interior subcavities in the SEA model. It corresponds to assigning input energies to each corresponding subcavity using Eq. (3) and imposing these values in Eq. (1). The consequence is that the energy of a subcavity is directly proportional to its volume and therefore, the power flow from a subcavity to a panel and/or the exterior cavity is independent of the volume [see, for example, Eq. (9) for the non-resonant path]. In addition, the exterior space is chosen to be large enough so that the CLF relating this space to a subcavity (calculated using the reciprocity relations) is very small regardless of the volume/modal density of the subcavity. Consequently, this justifies that a partitioning scheme in which arbitrary subcavity volumes can be chosen (see also Ref. 47 for qualitative discussion on this topic). This approach allows for a more realistic technique to predict the interior sound field while using the SEA framework to predict the transmitted noise through the enclosure panels and apertures.

In the image source method, the sound pressure field at a point M in the enclosure due to a source located at a point M_0 is considered as an infinite sum of the contributions of acoustic fields created by virtual sources which are images at multiple orders of the real source through the different walls of the enclosure together with the direct acoustic field. In practice, the series is truncated since only a finite number of virtual sources contribute to the sound field at a point M. Assuming a temporal dependency $e^{j\omega t}$, the coherent sound pressure at a given point M in the enclosure reads

$$\hat{p}(M) = \hat{A}\frac{e^{-jk_0 r}}{4\pi r} + \sum_{i=1}^{N} \hat{A}\hat{\Re}_i \frac{e^{-jk_0 r_i}}{4\pi r_i} + \Xi_{N+1}, \tag{15}$$

where $\hat{A}$ is the amplitude of the monopole sound source; r is the distance between the receiver M and the source M_0; N is the number of contributing sources for a specified order of reflection; r_i is the distance between the receiver M and the image source of order N_i; k_0 is the wavenumber. A dissipative part can be included to take into account the atmospheric attenuation. This should be considered when the enclosure is large and reverberant. $\hat{\Re}_i$ is the global reflection coefficient corresponding to the walls which are intersected by the segment linking the virtual source of index i and order N_i and receiver M. Ξ_{N+1} is the residual contribution of all the remaining virtual sources. It tends to zero when N increases.

The convergence of the series depends on many parameters like the geometry of the enclosure, the position of the source compared to the receiver, and the sound absorbing coefficient of the walls. The more absorbing the walls, the greater the convergence speed and the quicker $\Xi_{N+1} \to 0$. The algorithm to generate the contributing virtual sources consists in calculating the position of the multiple virtual sources, discarding those which do not pass certain geometric tests (validity and visibility) as described by Borish[48] and adding up the contribution.

If a specular reflection model is assumed, reflection coefficient $\hat{\Re}_i$ is equal to $\hat{\Re}_i = \Pi_{j=1}^{N_i} \hat{\Re}_{ij}(\theta_{ij})$, where $\hat{\Re}_{ij}(\theta_{ij})$ is the reflection coefficient of the wall which depends in principle on the incidence angle θ_{ij} between the vector linking the virtual source of index i and receiver M and the normal to the jth intersected wall. The complex valued reflection coefficient at oblique incidence $\hat{\Re}_{ij}(\theta_{ij})$ can be assessed assuming a locally reacting behavior using a plane wave reflection coefficient $\hat{\Re}_{pw}(\theta)$ $(\hat{\Re}_{pw}(\theta) = (Z\cos\theta - Z_0)/(Z\cos\theta + Z_0)$, where Z is the normal surface impedance of the wall and $Z_0 = \rho_0 c_0$) or a spherical wave reflection coefficient $\hat{\Re}_{sp}(\theta)$ $[\hat{\Re}_{sp}(\theta) = \hat{\Re}_{op}(\theta) + (1 - \hat{\Re}_{op}(\theta))F(u)$ with $F(u) = 1 + j\sqrt{\pi}u\,\mathrm{erf}(u)$ and $u = (1+j)/2\sqrt{k_0 r_i}(\beta + |\cos\theta|)$, where $\beta = \rho_0 c_0 / Z$ is the normalized admittance] depending on the desired accuracy.[29] Alternatively, based on the assumption that it is a real number, the reflection coefficient can also be calculated by the relationship $\hat{\Re} = \sqrt{1-\alpha}$, where α denotes the normal incidence sound absorption coefficient, the statistical sound absorption coefficient, or the Sabine sound absorption coefficient which is usually the data provided by the manufacturer depending on the data available. The reflection coefficient does not depend then on the incidence angle anymore. If the incidence angle is to be accounted for in the calculation, the locally reacting assumption can be used again to derive a normal surface impedance using the relationship $\hat{Z} = Z_0[(1+\sqrt{1-\alpha})/(1-\sqrt{1-\alpha})]$, where α is one of the aforementioned absorption coefficients which can then be substituted back in the oblique incidence plane wave or spherical wave reflection coefficient. Some architectural acoustics commercial softwares interested in predicting sound pressure levels in rooms use incoherent summation of image sources contributions. In such cases, Eq. (15) then becomes

$$|\hat{p}(M)|^2 \approx \frac{|\hat{A}|^2}{16\pi^2 r^2} + \sum_{i=1}^{N} |\hat{\Re}_i|^2 \frac{|\hat{A}|^2}{16\pi^2 r_i^2} + \Theta_{N+1}, \tag{16}$$

where Θ_{N+1} is a residual term taking into account the incoherent contribution of the remaining virtual sources. This term can be obtained by assuming that the residual field is reverberant and that its power is equal to the power absorbed by the room after the Nth reflection. The residual term simply then reads

$$\Theta_{N+1} = \frac{4(1-\bar{\alpha})^{N+1}}{S\bar{\alpha}}, \tag{17}$$

where $\bar{\alpha}$ is the average sound absorption coefficient of the enclosure and S is the total area of the enclosed cavity.

TABLE I. External dimensions of the enclosures.

	L_x (m)	L_{y1} (m)	L_{z1} (m)	L_{y2} (m)	L_{z2} (m)
Box	1.52	1.3	2.0	⋯	⋯
L-shape	1.52	1.3	1.0	1.0	1.0

For the energy reflection coefficient, one can classically use $|\hat{\mathfrak{R}}_i|^2=\Pi_{j=1}^{N_i}|\hat{\mathfrak{R}}_{ij}|^2=\Pi_{j=1}^{N_i}(1-\alpha_{ij})$ or the other ways of calculating the amplitude reflection coefficient presented previously.

C. Calculation of the insertion loss

The enclosure insertion loss IL is calculated as the difference between the sound power radiated by the sound source L_{w0} in free field condition (expressed in dB: ref 10^{-12} W) and radiated by the enclosure L_w given by:

$$L_w = 10\log\left(\sum_{i=1}^{N}\omega\eta_{i\text{ext}}E_i\right)+120, \tag{18}$$

where E_i is the energy of subsystem i and $\eta_{i\text{ext}}$ is the coupling loss factor between subsystem i and the external cavity.

III. EXPERIMENTAL SET-UP

In order to evaluate the performance of the model and to investigate experimentally the effect of different parameters (materials constituting the panels, presence of absorbing treatments, presence of an opening, and source location) on the insertion loss, an enclosure has been built in the Groupe d'Acoustique of University of Sherbooke (GAUS) laboratories. The frame consisted of vertical and horizontal tubular steel beams filled with urethane foam. The box-shaped structure could be easily modified to add a lateral module which converted the original geometry into an L-shaped enclosure. The enclosures external dimensions are provided in Table I. A layer of neoprene was taped onto the frame to ensure the air-tightness between the metallic skeleton and the panels which were attached using bolts. The walls of the enclosure consisted of either 1.16 mm thick metal sheets or 1.61 cm thick presswood panels. The sound treatment was made up of 7 cm thick rockwool (Roxul) covered with a 1 mm thick steel perforated plate (perforation rate equal to 33% and radius of perforation equal to 1.5 mm). Seal material was used to avoid acoustic leaks especially around the junctions between the panels and the door and the panels and the floor. The interior of the enclosure was accessible through a sandwich door consisting of two 1 mm thick steel plates with a core of rockwool. One of the enclosure panels could comprise an opening consisting of a rectangular steel duct of cross-section 400×200 mm^2. When the sound absorbing treatment was present, the duct was 174 mm deep whereas in the case of the bare steel enclosure, the opening depth corresponded to the panel thickness, namely, 1.16 mm for steel and 1.615 cm for presswood. Photographs of some of the tested configurations are displayed in Fig. 1.

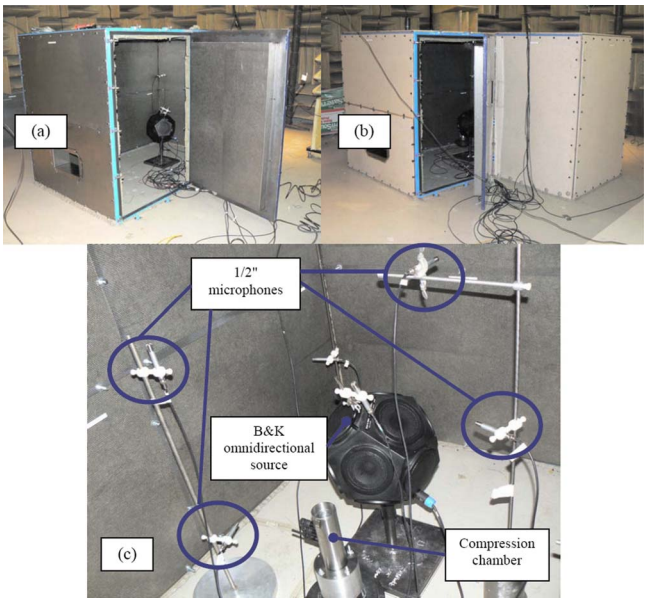


FIG. 1. (Color online) Experimental set-up: enclosure surrounding a noise source and radiating inside a semi-anechoic room. (a) Box-shaped enclosure (steel panels). (b) L-shaped enclosure (presswood panels). (c) Interior of the enclosure, B&K omnidirectional sound source and compression chamber, and microphones.

In order to get an adequate signal to noise ratio outside the enclosure, two sources were combined in order to deliver sufficient power in the frequency band of interest [80–6300 Hz]. The sound field inside the enclosure was generated with both a B&K omnidirectional sound source and a compression chamber. The first source radiated more energy in the low frequency range (below 1 kHz) whereas the second one radiated more energy above 1 kHz. The sound powers of the individual two sources have been measured in a semi-anechoic room using five 0.5 in. B&K microphones attached along a rotating semi-circumferential antenna. The internal sound field inside the box-shaped and the L-shaped enclosure have been spatially sampled with, respectively, 8 and 11.5 in. B&K microphones. The sound power radiated by the enclosure was measured by intensity scanning. A sound intensity probe was used to scan all the faces one by one and to assess the total radiated sound power by summing the sound power radiated from each face. In the cases where there was an opening, a special attention was paid when scanning the opening. Each time the structure of the enclosure was modified, preliminary checks were carried out to ensure that the enclosure was correctly sealed. The presence of unwanted acoustic leaks was diagnosed through intensity measurements. When identified, additional seal was added in the corresponding zone in order to minimize their importance. A B&K pulse system was used to control both the generation of the signal (white noise) and the acquisition of sound pressure and sound intensity signals.

The experiment has been designed to investigate the effect of (i) the geometry of the enclosure (rectangular box vs L-shaped box), (ii) the panels material (steel vs presswood),

TABLE II. List of tested configurations.

Configuration	Material	Geometry	Acoustic material	Source location 1 (○)	Source location 2 (◎)	Source location 3 (●)	Opening 200mmx400mm
1	Steel	Box	☒	☒			
2	Steel	Box	☒		☒		
3	Steel	Box	☒	☒			☒
4	Steel	Box	☒		☒		☒
5	Steel	L	☒			☒	☒
6	Steel	L	☒	☒			☒
7	Steel	L	☒		☒		☒
8	Steel	L	☒			☒	
9	Steel	L	☒	☒			
10	Steel	L	☒		☒		
11	Presswood	L	☒			☒	
12	Presswood	L	☒	☒			
13	Presswood	L	☒		☒		
14	Presswood	L	☒			☒	☒
15	Presswood	L	☒	☒			☒
16	Presswood	L	☒		☒		☒
17	Presswood	Box	☒	☒			☒
18	Presswood	Box	☒		☒		☒
19	Presswood	Box	☒	☒			
20	Presswood	Box	☒		☒		
21	Presswood	Box		☒			
22	Presswood	Box			☒		
23	Steel	Box		☒			
24	Steel	Box			☒		
25	Steel	Box		☒			☒
26	Steel	Box			☒		☒

(iii) the presence of a sound absorbing treatment, (iv) the location of the source inside the enclosure, and (v) the presence of an opening. A total of 26 configurations were tested. They are described in Table II. In the table, the source position refers symbolically to the approximate position of the geometrical centre of the B&K omnidirectional source and the compression chamber. For each configuration, both sound pressure levels inside the enclosure and sound intensity radiated by each face were measured.

IV. RESULTS

A. Internal sound field

First, the image sources model for the internal sound field is evaluated by comparisons with existing models and experimental data. Consider a box-shaped enclosure corresponding to configuration 1 in Table II with a point source of unit amplitude [$\hat{A}=1$ in Eq. (15)] located at ($x0=0.43$ m, $y0=0.55$ m, $z0=1.25$ m). All walls are covered with the sound absorbing treatment described in Sec. IV B while the floor is made up of concrete. A specular spherical reflection model is used. The sound absorbing material is supposed locally reacting and its complex surface impedance at normal incidence has been calculated using the TMM. In the TMM, the acoustic treatment is defined from the material thickness and physical properties of each layer (Roxul+perforated plate) on rigid backing, properties which were measured at GAUS material characterization laboratory. Geometrical parameters (porosity, flow resistivity, tortuosity, and characteristic lengths), density, and viscoelastic parameters of Roxul have also been assessed in the same laboratory using direct or indirect measurement techniques.

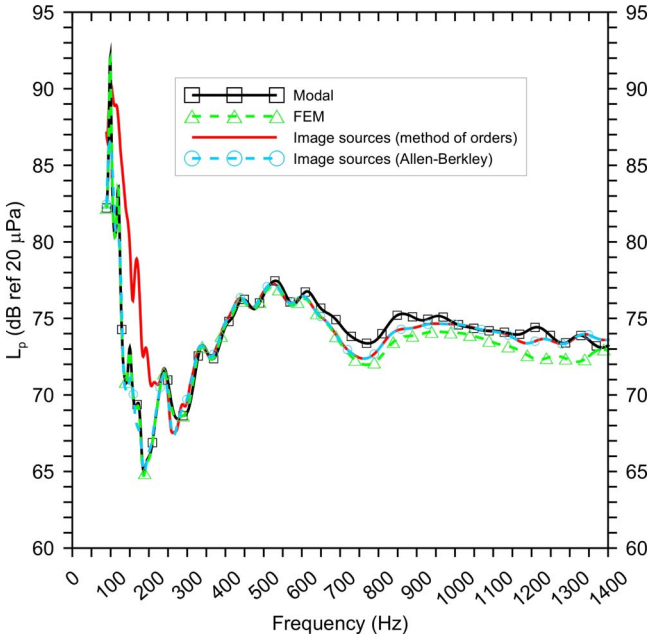


FIG. 2. (Color online) Comparison between narrow band coherent sound pressure levels averaged over eight microphone positions calculated using modal approach (modal), FEM, image source model, and image source model based on Allen–Berkley's algorithm. Case of a box-shaped air-tight enclosure with sound absorbing treatment.

Figure 2 shows the narrow band SPL averaged over eight points located throughout the cavity and obtained with the image sources model using a coherent summation [Eq. (15)] and several other approaches (modal approach and finite element method). In the modal approach, the sound pressure field inside the enclosure is expanded in terms of eigenmodes of the cavity with rigid walls. The maximum frequency of interest (1400 Hz) has been selected in order to have reasonable computation time. A related modal truncation frequency of 2100 Hz which corresponds to 3423 kept modes and 1.5 times has been chosen. In the finite element model, the internal air volume is discretized using $39\times30\times38$ HEXA8 acoustic elements. Again this mesh has been adopted to ensure convergence up to 1400 Hz. For the image source method, two algorithms are compared. The first is the method of orders described by Eq. (15). The sound pressure at a point M due to a point source located at a point M_0 is considered as the sum of all contributing image sources at a given reflection order. It requires determining if a virtual source contributes indeed to the total sound pressure field using a variety of geometrical tests as mentioned previously. It is valid for any polyhedron geometry with arbitrary distribution of sound absorbing material (one face can be covered with multiple materials). The second is the Allen–Berkley approach.[49] In this method, the sound pressure is obtained through a summation of contributions of sources located in replicas of the original room through symmetries with respect to each wall. No geometrical test is required. Note that this approach is only valid for parallelepipedic geometry

where each wall is covered with a uniform material. In the following, unless otherwise specified, an order equal to 12 has been selected for the method of orders. A distance $d_{max}=9L_z$ representing the maximum distance used to truncate the infinite series in Allen–Berkley's paper[49] has been chosen for Allen–Berkley technique. It represents a fair compromise between accuracy of the solution and computation time.

Figure 2 indicates a good agreement between modal approach, finite element method, and image source method (Ref. 49) up to 600 Hz. Above this frequency, small differences of the order of 1 dB are observed due to the modal truncation frequency, which is too low, and to the finite element mesh, which is too coarse. Image source results based on the method of orders and Allen–Berkley's approach[49] are identical above 200 Hz. Differences are observed below 200 Hz between the two techniques. This occurs when using spherical reflection coefficients with complex-valued wall surface impedance with a reflection order which is too low. Above 200 Hz, a reflection order equal to 8 gives the same results as an order equal to 12. An order equal to 4 yields to results less than 1 dB from those obtained with an order 8.

Calculations based on incoherent summation of virtual source contributions with or without residual [see Eq. (17)] have also been carried out but are not shown here for the sake of conciseness. It was found that third-octave band comparisons of SPL predicted with (i) analytical formulas based on the summation of the direct field and reverberant field (room constant) or Sabine's model (reverberant diffuse field only),[42] (ii) ray tracing,[50] and (iii) radiative transfer[31] show very good correlation between the different approaches. Sabine's formula provides the poorer agreement since the direct field is not included in the calculation and the cavity is very absorbing (diffuse field sound absorption coefficient of 0.5 at 250 Hz and larger than 0.8 above 500 Hz). At very low frequencies (100 Hz), the image sources based on the method of orders without residual shows convergence problems but this can be remedied if the residual is taken into account. If the residual is included in the summation, the reflection order can be decreased to 6 for this specific case with a loss of accuracy smaller than 1 dB.

Comparisons of SPL predictions and measurements are shown in Fig. 3. Sound pressure levels averaged over eight microphone positions are used. Prediction results are based on the image sources (method of orders) for a box-shaped air-tight enclosure with sound absorbing treatment and an opening for two source location (configurations 3 and 4 in Table II). In this configuration, the sound power of the two acoustic sources has been used as an input to the image source software. A fairly good agreement between the image source approach and measurements is observed (discrepancies less than 3 dB). Note that comparisons have also been performed in the case of L-shaped enclosure where the interior cavity has been naturally divided into three zones. The image source model captures the right trend but differences are observed in the zone which is not visible from the source (shadow zone) since the present technique does not account for diffraction effects.

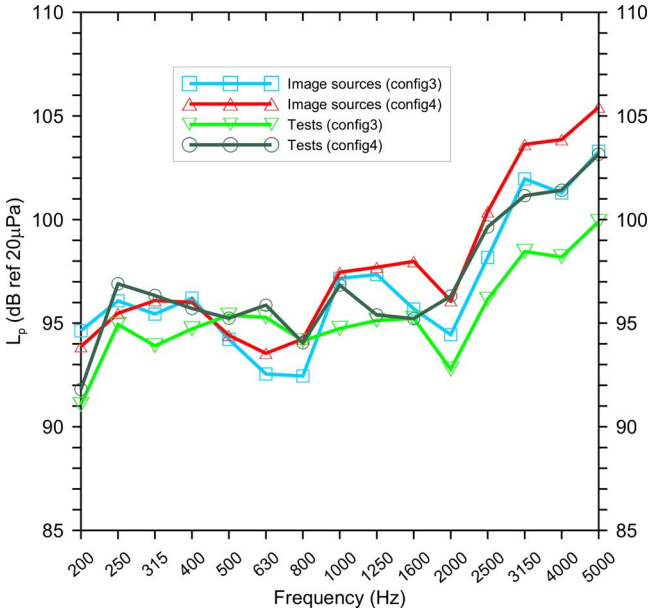


FIG. 3. (Color online) Comparison between one-third octave sound pressure levels averaged over eight microphone positions, measured and calculated using the image source model. Case of a box-shaped air-tight enclosure with sound absorbing treatment and an opening for two source location (configurations 3 and 4).

B. Insertion loss

The quality of prediction of the insertion loss depends on the knowledge of the input parameters. Elastic properties (density, Young's modulus, and Poisson's ratio) and loss factor of the panels (steel and presswood) have been taken from the literature or estimated from measurements with impact hammer on the panels. The consistency of these values has been checked by comparing transmission loss measurements in reverberant rooms and results predicted by TMM. These properties are provided in Table III and used in the coupling loss factors, modal densities, and acoustical treatment insertion loss calculations. The latter indicators have been assessed using the TMM assuming that the panels have lateral infinite extent. The steel plate has been modeled as a thin plate (Kirchhoff theory), whereas presswood panel has been modeled as an isotropic solid. The insertion losses of the treatment have been calculated for each support panel (steel and presswood plate). The field incidence sound absorption coefficient of the absorbing treatment Roxul has also been predicted using the transfer matrix method.

As far as the calculation of interior and exterior resonant coupling loss factors are concerned, values of $m=3$ and $m=1.5$, respectively, have been chosen in Eq. (8). Factor $m=1.5$ corresponds to a correction of boundary conditions which are in-between simply supported and clamped (correction of approximately 2 dB at low frequencies). Factor m

TABLE III. Physical properties of steel and wood used in the calculations.

	Steel	Presswood
Young's modulus E (GPa)	210	2.15
Poisson's ratio	0.33	0.25
Density (kg m^{-3})	7850	601

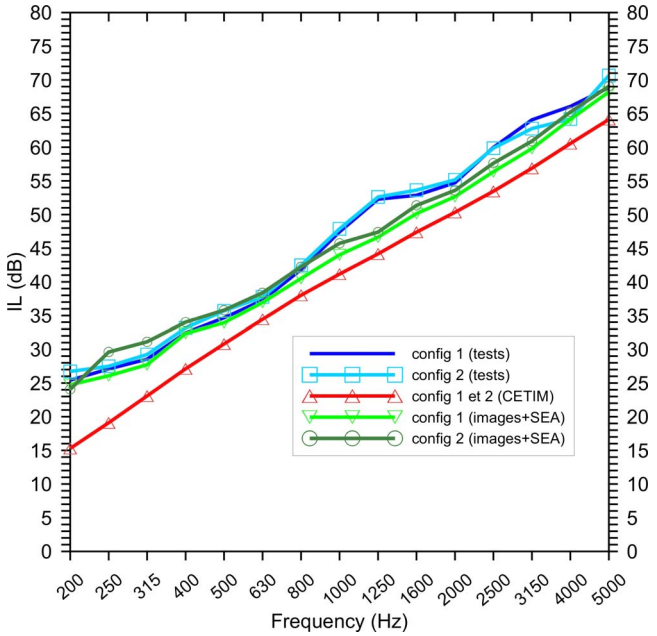


FIG. 4. (Color online) Comparison between measured and calculated insertion loss with CETIM's analytical model (Ref. 26) and hybrid image sources+SEA model. Case of a box-shaped air-tight enclosure with steel panels and sound absorbing treatment for two source location (configurations 1 and 2).

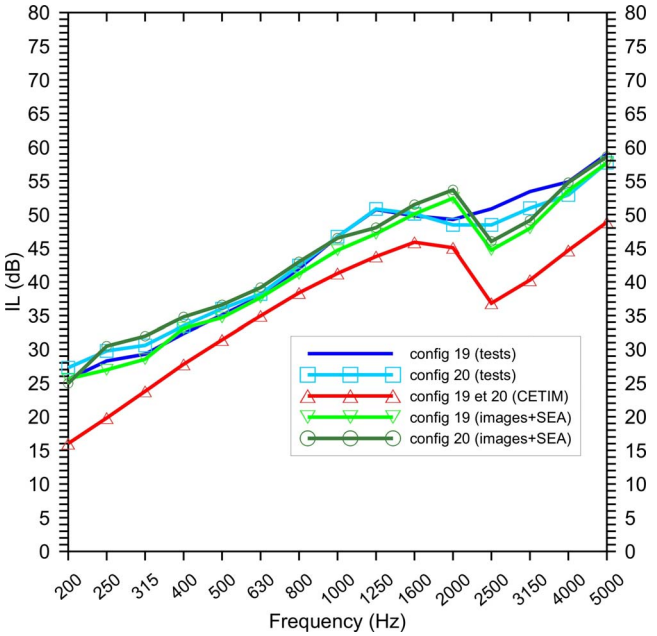


FIG. 5. (Color online) Comparison between measured and calculated insertion loss with CETIM's analytical model (Ref. 26) and hybrid image sources+SEA model. Case of a box-shaped air-tight enclosure with presswood panels and sound absorbing treatment for two source location (configurations 19 and 20).

=3 combines the fact that the panel radiates inside a confined box (right angle baffles) and the aforementioned boundary condition effect.

In the image source technique, a coherent summation based on a real reflection coefficient independent of the incidence angle and calculated with the formula $\hat{\Re}=\sqrt{1-\alpha}$ has been used. For the box-shaped enclosure, the internal cavity is arbitrarily subdivided into two subcavities unless otherwise specified. For L-shaped enclosures, it is subdivided into three subcavities. In each of the subcavity, the sound pressure level averaged over ten microphone positions randomly distributed in each subcavity is calculated using the image source method in order to obtain a better representation of the spatial variation in the sound field. This substructuration allows one to simulate the effect of the source position inside the enclosure. As will be shown later, the subdivision of the cavity is more important in configurations with openings.

Figures 4 and 5 show a comparison between insertion loss measurements and predictions using an analytical model [CEntre Technique des Industries Mécaniques (CETIM) (Ref. 26)] and the proposed hybrid image sources+SEA model for a box-shaped air-tight enclosure with, respectively, steel and presswood panels and sound absorbing treatment for two source locations [configurations 1 and 2/configurations 19 and 20 (see Table II)]. The CETIM model[26] is based on the calculation of the power radiated by the enclosure through various transmission paths (structural, leaks, mechanical links, etc.) using the transmission loss of each component. This model has been chosen as it constitutes a good and easy to implement representation of most of simple energy based approaches available in the literature. In this paper, the sound transmission loss of treated panels together with field incidence sound absorption coefficient calculated with the TMM has been used as input data into the CETIM's model. Recall that in the tested configurations, mechanical sound transmission paths are neglected. These figures show that the results given by the hybrid image sources-SEA approach are in very good agreement for both the steel and presswood boxes. However, the CETIM model underestimates the insertion loss from 5 to 10 dB since the acoustic absorption area is calculated from Sabine's formula. Reductions in the insertion loss (see Fig. 5) around the critical frequency of the presswood panel are observed as expected. The source location has no significant influence on the insertion loss of such an air-tight enclosure (as long as the source is not too close from one of the panel).

Figure 6 depicts the comparison between measured and calculated insertion loss with CETIM's analytical model, hybrid image sources+2 cavity subsystem SEA model and hybrid image sources+1 cavity subsystem SEA model for a box-shaped enclosure with steel panels and sound absorbing treatment. An opening of 200×400 mm^2 is placed on one the panel and two source locations are considered [configurations 3 and 4 (see Table II)]. Figure 6 indicates that using two acoustic subsystems allows for better simulation of the effect of source location, contrarily to the model with only one acoustic subsystem. Moreover, it is observed that when the source is closer to the opening, the insertion loss is lower than when the source is farther from it particularly at high frequencies. Again, the image source method with a coherent summation provides a better agreement with measurements compared to the analytical model.

Figure 7 presents a comparison of SEA-based prediction models with different degrees of sophistication with the measurements for configuration 3 (see Table II). In the legend, SEA standard (1 cavity) refers to classical SEA where the internal sound field is supposed to be reverberant and diffuse,

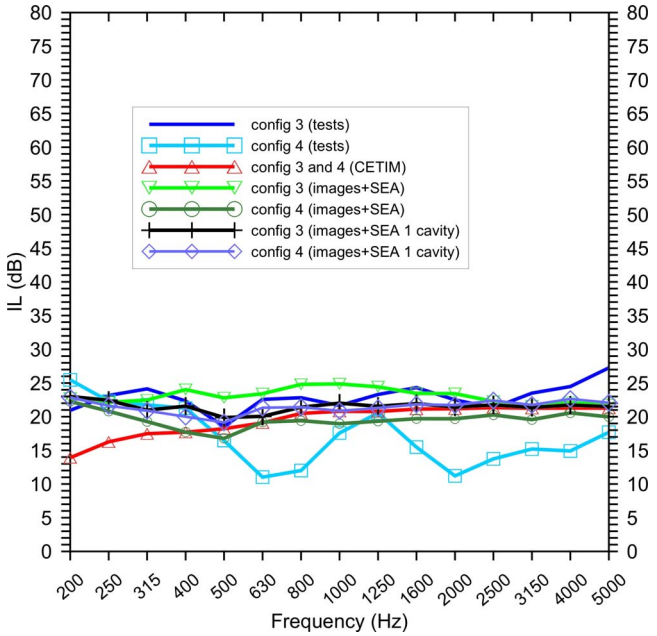


FIG. 6. (Color online) Comparison between measured and calculated insertion loss with CETIM's analytical model (Ref. 26), hybrid image sources +2 cavities subsystems SEA model, and hybrid image sources+1 cavity subsystem SEA model. Case of a box-shaped enclosure with steel panels and sound absorbing treatment with an opening for two source location (configurations 3 and 4).

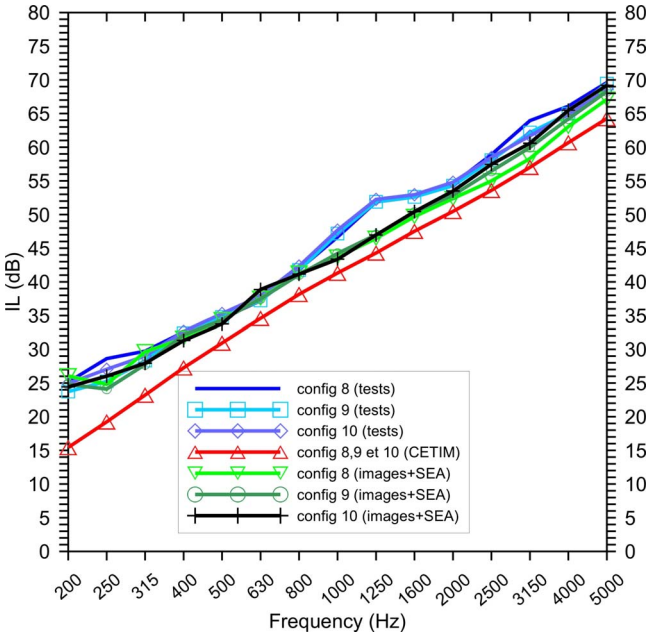


FIG. 8. (Color online) Comparison between measured and calculated insertion loss with CETIM's model (Ref. 26) and hybrid image sources+2 cavity subsystem SEA model. Case of an L-shaped air-tight enclosure with steel panels and sound absorbing treatment for three source location (configurations 8–10).

and SEA standard (2 cavities) refers to a classical SEA in which the enclosure volume has been divided into two subcavities in which the internal sound field is supposed to be reverberant and diffuse. SEA standard+direct field (2 cavities) corresponds to the previous model in which the average contribution of the direct field in each subcavity has been obtained from an incoherent summation. As expected, it is seen that SEA standard (1 cavity) fails to predict correctly the insertion loss. Things become even worse when the enclosure cavity is split into two subcavities and only the reverberant field is retained. Indeed, the reverberant SPL in the subcavity close to the opening is decreased compared with the case of a single cavity leading to an increase in the insertion loss. On the other hand, adding the contribution of the direct field to the reverberant field improves the results. This simplified model proves to be appropriate for a convex-shaped enclosure and a simple point source but is expected to fail for more complicated geometries like L-shaped where shadow zones and more complex reflection patterns may exist when the source is located on one end of the enclosure and for extended sources. Finally, the best results in Fig. 7 are observed when the hybrid image source-SEA approach is used with two subcavities.

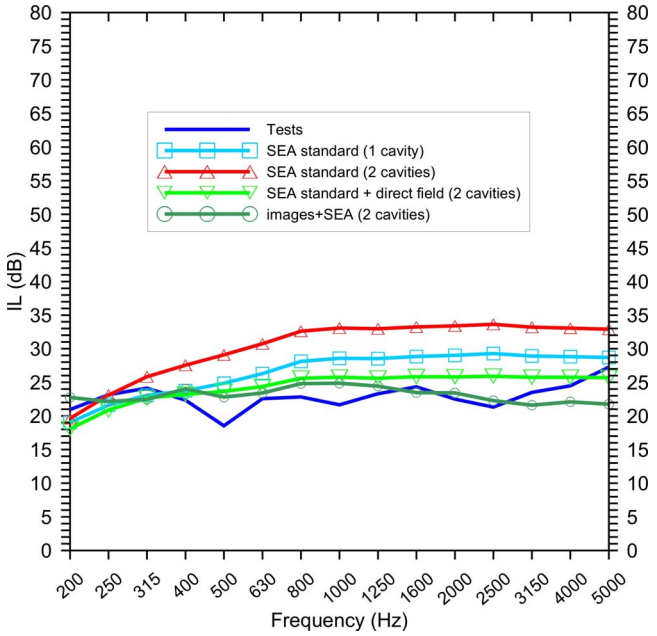


FIG. 7. (Color online) Comparison between measured and calculated insertion loss with classical SEA model (1 cavity), classical SEA model (2 cavities), classical SEA model (2 cavities)+direct field, and hybrid image sources+2 cavity subsystem SEA model. Case of a box-shaped enclosure with steel panels and sound absorbing treatment with an opening for source location number 1 (configuration 3).

Figures 8 and 9 show the comparisons between the measured and calculated insertion loss for configurations 8–10 and 14–16, respectively (see Table II). The configurations in Fig. 8 correspond to the L-shaped enclosure with steel walls with an acoustical treatment and three source locations. The configurations in Fig. 9 corresponds to the L-shaped enclosure with presswood walls and an opening of dimensions 200×400 mm^2 on one of the walls and three source locations. The label images+SEA refers to hybrid images method-SEA with three subcavities. For the configuration without opening, it is seen that the effect of the source location on the acoustical performance of the enclosure is not important. Conclusions are identical to those of Fig. 4 for the rectangular enclosure. Apart from the analytical CETIM's approach, all the hybrid models provide acceptable predictions with regard to the effect of the source location. In the case of an opening (Fig. 9), the source position has a high influence on the enclosure insertion loss. The hybrid tech-

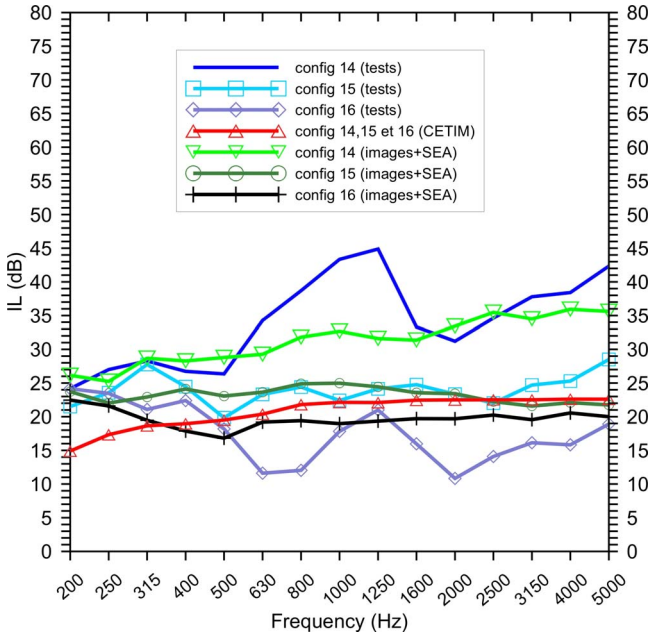


FIG. 9. (Color online) Comparison between measured and calculated insertion loss with CETIM's model and hybrid image sources+3 cavity subsystems SEA model. Case of an L-shaped enclosure with presswood panels and sound absorbing treatment with an opening for three source location (configurations 14–16).

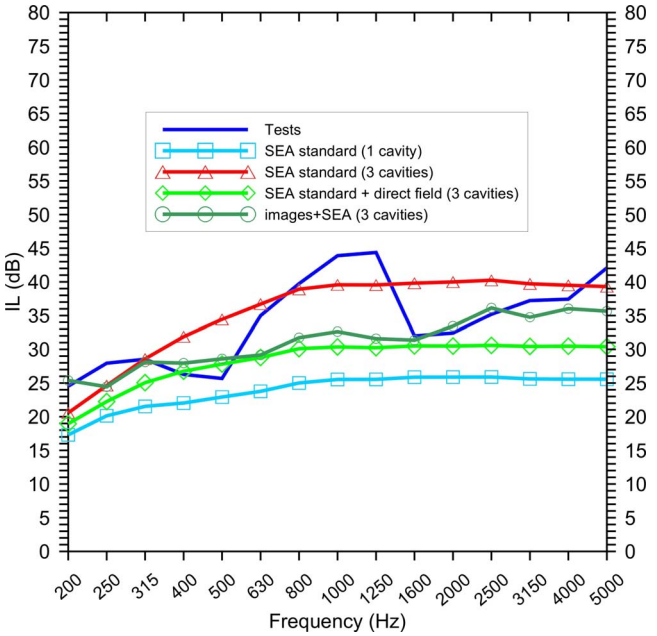


FIG. 10. (Color online) Comparison between measured and calculated insertion loss with classical SEA model (1 cavity subsystem), classical SEA model (3 cavity subsystems), classical SEA model (3 cavities)+direct field and hybrid image sources+3 cavity subsystem SEA model. Case of an L-shaped enclosure with steel panels and sound absorbing treatment with an opening for source location number 3 (configuration 5).

nique is able to capture the effect of the source position, namely, a reduction in the insertion loss as the source is closer to the opening. Note, however, that large discrepancies (up to 10 dB) are observed in some frequency bands especially in the case of configurations 14 and 16. Clearly, some of the physics of the problem is missing in the proposed approach. This could be due to the diffraction effect around the protuberant right angle corner, the nature of the incident field on the large opening (which cannot be considered as diffuse in the calculation of the sound transmission coefficient of the aperture when the source is close to the opening), and diffraction effects around the aperture edges. Investigation is underway to identify the cause of these discrepancies. CETIM's model once again provides an average curve of the enclosure performance, independent of the source location, but discrepancies with experimental results may become important (case of configurations 14, for instance).

Figure 10 compares different degrees of sophistication of prediction models based on SEA with measurement for configuration 5 [L-shaped enclosure with steel walls, acoustical treatment and an opening for source position number 3 (see Table II)]. The best results are again obtained with the hybrid model. SEA standard+direct field (3 cavities) gives less accurate results compared to the ones presented previously in Fig. 7. This is mainly due to the L-shaped geometry but the trends are nonetheless correctly captured. Again, accurate predictions cannot be obtained by using only one cavity. Finally, as observed before, using only the reverberant field in three subcavities without inclusion of the direct field grossly overestimates the insertion loss.

V. CONCLUSION

This paper presented a general and simple model to predict the sound insertion loss of large free-standing enclosures. It is based on a hybrid method combining the SEA for the sound transmission across the various elements of the enclosure and the method of image sources for the sound field inside the enclosure. Compared to available techniques, SEA is relevant to deal with this problem, given the large size of enclosures, the large frequency range in which the response is sought, and the uncertainties related to enclosure walls properties and their mounting conditions. SEA also gives a general framework to solve the global problem and its matrix form makes it easy for the enclosure to be much more complicated geometrically [as many subsystems (panels and apertures) as desired] and physically (e.g., additional structural couplings between panels and structure borne excitation possible).

Compared to previous works published on the subject, the present approach has several features:

- The latest advances in material modeling (use of transfer matrix approach rather than use of measured impedance data) offers more flexibility to calculate coupling loss factors for various sound absorbing materials.
- The use of image sources technique allows for a better description of the internal sound field and to account for the location of the acoustic source in the enclosure.
- Leaks and opening are accounted for through a recent model of sound transmission through apertures[46] which has been validated using various experimental configurations.

This hybrid approach has been evaluated by comparing, on one hand, with other existing models and, on the other hand, with experimental results using rectangular and L-shaped enclosures with several interior source locations, materials, and

openings. The comparisons between the models and the experimental results showed a good agreement for most of the 26 tested configurations. The coherent image source model coupled with SEA has been found a reliable tool to predict the acoustic at low frequencies compared to the "classical" simplified model.

[1]R. S. Jackson, "The performance of acoustic hoods at low frequencies," Acustica **12**, 139–152 (1962).
[2]R. S. Jackson, "Some aspects of the performance of acoustic hoods," J. Sound Vib. **3**, 82–94 (1966).
[3]F. J. Fahy, *Sound and Structural Vibration* (Academic, London, England, 1989).
[4]M. C. Junger, "Sound transmission through an elastic enclosure acoustically coupled to a noise source," ASME Paper No. 70-WA/DE-12.
[5]D. J. Oldham and S. N. Hillarby, "The acoustical performance of small close fitting enclosures, Part 1: Theoretical models," J. Sound Vib. **150**, 261–281 (1991).
[6]L. W. Tweed and D. R. Tree, "Three methods for predicting the insertion loss of close-fitting acoustical enclosures," Noise Control Eng. **10**, 74–79 (1978).
[7]K. P. Byrne, H. M. Fischer, and H. V. Fuchs, "Sealed, close-fitting, machine-mounted acoustic enclosures with predictable performance," Noise Control Eng. **31**, 7–15 (1998).
[8]S. M. Hasheminejad and S. Mehdizadeh, "Acoustic performance of a multi-layer close-fitting hemispherical enclosure," Noise Control Eng. **54**, 86–100 (2006).
[9]A. J. Pretlove, "Forced vibrations of a rectangular panel backed by a closed rectangular cavity," J. Sound Vib. **3**, 252–261 (1966).
[10]R. W. Guy and M. C. Bhattacharya, "The transmission of sound through a cavity-backed finite plate," J. Sound Vib. **27**, 207–223 (1973).
[11]E. H. Dowell, G. F. I. Gorman, and D. A. Smith, "Acoustoelasticity: General theory, acoustic natural modes and forced response to sinusoidal excitation, including comparisons with experiment," J. Sound Vib. **52**, 519–542 (1977).
[12]R. W. Guy, "The response of a cavity backed panel to external airborne excitation: A general analysis," J. Acoust. Soc. Am. **65**, 719–731 (1979).
[13]S. Narayanan and R. L. Shanbhag, "Sound transmission through elastically supported sandwich panels into a rectangular enclosure," J. Sound Vib. **77**, 251–270 (1981).
[14]L. Cheng and C. Lesueur, "Influence of damping on the vibro-acoustic response: Theoretical study of a plate coupled to a cavity and acoustically excited (in French)," J. Acoust. **2**, 105–118 (1989).
[15]Y. Y. Lee and C. F. Ng, "The effects of coupled source/cavity modes on the acoustic insertion loss of close-fitting enclosures," Journal of Building Acoustics **2**, 549–567 (1997).
[16]P. Agahi, U. P. Singh, and J. O. Hetherington, "Numerical prediction of the insertion loss for small rectangular enclosures," Noise Control Eng. **47**, 201–208 (1999).
[17]N. Atalla and R. Panneton, "The effects of multilayer sound-absorbing treatments on the noise field inside a plate backed cavity," Noise Control Eng. **44**, 235–243 (1996).
[18]F. X. Bécot and F. Sgard, "On the use of poroelastic materials for the control of the sound radiated by a cavity backed plate," J. Acoust. Soc. Am. **120**, 2055–2066 (2006).
[19]F. Sgard, N. Atalla, and C. K. Amedin, "Vibro-acoustic behavior of a cavity backed by a plate coated with a meso-heterogeneous porous material," Acta. Acust. Acust. **93**, 106–114 (2007).
[20]F. Sgard and N. Atalla, "Mean flow effects on a plate-backed cavity, Part 1: Theory," Acta Acust. **83**, 243–251 (1997).
[21]F. Polonio, T. Loyau, J. M. Parot, and G. Gogu, "Acoustic radiation of an open structure: Modeling and experiments," Acta Acust. **90**, 496–511 (2004).
[22]A. F. Seybert, C. Y. R. Cheng, and T. W. Wu, "The solution of coupled interior/exterior acoustic problems using the boundary element method," J. Acoust. Soc. Am. **88**, 1612–1618 (1990).
[23]P. Jean, "Coupling geometrical and integral methods for indoor and outdoor sound propagation—Validation examples," Acta Acust. **87**, 236–246 (2001).
[24]Y. Y. Lee and C. F. Ng, "Sound insertion loss of stiffened enclosure plates using the finite element method and the classical approach," J. Sound Vib. **217**, 239–260 (1998).
[25]M. Tournour and N. Atalla, "Vibroacoustic behavior of an elastic box using state-of-the-art FEM-BEM approach," Noise Control Eng. **46**, 83–90 (1998).
[26]CEntre Technique des Industries Mécaniques (CETIM), *Sound Insulation of Machinery Using Enclosures Proceedings* (Publications Centre Technique des Industries Mécaniques, Senlis, France, 1995) (in French).
[27]N. Trompette, T. Loyau, and G. Lovat, "Machinery enclosures—Design: Basic rules and experimental implementation," Cahiers de notes documentaires Institut National de Recherche en Sécurité—Hygiène et sécurite du travail **182**, 49–72 (2001) (in French).
[28]E. O'Keefe, "A computer simulation for determining far-field noise levels radiated from a noise source within a rectangular enclosure," in Inter-Noise (1980).
[29]J. F. Allard and N. Atalla, *Propagation of Sound in Porous Media: Modelling Sound Absorbing Materials*, 2nd ed. (Wiley-Blackwell, New York, 2009).
[30]J. L. Barbry and N. Trompette, "Improvement of the prediction of panels and machinery enclosures acoustical performances," Département Ingénierie des équipements de travail de l'Institut National de Recherche en Sécurité, Nancy, France, 2007 (in French).
[31]A. Le Bot and A. Bocquillet, "Comparison of an integral equation on energy and the ray-tracing technique in room acoustics," J. Acoust. Soc. Am. **108**, 1732–1740 (2000).
[32]V. Cotoni, A. Le Bot, and L. Jezequel, "Sound transmission through a plate by an energy flow approach," Acta. Acust. Acust. **88**, 827–836 (2002).
[33]I. L. Ver, "Reduction of noise by acoustic enclosures," in Proceedings of ASME Design Engineering Technical Conference (1973).
[34]R. Ming and J. Pan, "Insertion loss of an acoustic enclosure," J. Acoust. Soc. Am. **116**, 3453–3459 (2004).
[35]R. H. Lyon, "Noise reduction of rectangular enclosures with one flexible wall," J. Acoust. Soc. Am. **35**, 1791–1797 (1963).
[36]E. Eichler, "Thermal circuit approach to vibrations in coupled systems and the noise reduction of a rectangular box," J. Acoust. Soc. Am. **37**, 995–1007 (1965).
[37]V. Cole, M. J. Crocker, and P. K. Raju, "Theoretical and experimental studies of the noise reduction of an idealized cabin enclosure," Noise Control Eng. **20**, 122–132 (1983).
[38]F. Sgard, H. Nélisse, N. Atalla, C. K.Amedin, and R. Oddo, "Development of a software to design machinery acoustic enclosures (in French)," Technical Report No. R-619, 2009, Institut Robert Sauvé en Sécurité et Santé du Travail, Montréal, Canada.
[39]R. Craik, *Sound Transmission Through Buildings Using SEA* (Gower, Brookfield, Vermont, USA, 1996).
[40]M. Bolduc, "Acquiring statistical energy analysis damping loss factor for complex structures with low to high damping characteristics," Ph.D. thesis, Université de Sherbrooke, Sherbrooke, Quebec, Canada, 2007.
[41]Mecanum Inc, Nova 2008 User's Guide, Mecanum Inc. and University of Sherbrooke, Sherbrooke, Quebec, Canada, 2008.
[42]A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981).
[43]F. G. Leppington, E. G. Broadbent, and K. H. Heron, "The acoustic radiation efficiency of rectangular panels," Proc. R. Soc. London, Ser. A **382**, 245–271 (1982).
[44]K. H. Heron, "The radiation efficiency of non-isotropic plates with various edge boundary conditions and baffle geometries," DERA/MSS/MSTR2/TR990674.1999 DERA.
[45]H. Nélisse, T. Onsay, and N. Atalla, "Structure-borne insertion loss of sound package components," in SAE Meeting, 2003.
[46]F. Sgard, H. Nélisse, and N. Atalla, "On the modeling of diffuse field sound transmission loss of finite thickness apertures," J. Acoust. Soc. Am. **122**, 302–313 (2007).
[47]F. J. Fahy, "A note on the subdivision of a volume of air in a vehicle enclosure into sea subsystems," J. Sound Vib. **271**, 1170–1174 (2004).
[48]J. Borish, "Extension of the image model to arbitrary polyhedra," J. Acoust. Soc. Am. **75**, 1827–1836 (1984).
[49]J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am. **65**, 943–950 (1979).
[50]D. Van Maercke, "Simulation of sound fields in time and frequency domain using a geometrical method," in 12th ICA (1986).

Efficient numerical modeling of traffic noise

Erik M. Salomons,[a] Han Zhou, and Walter J. A. Lohman
TNO Built Environment and Geosciences, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands

(Received 18 May 2009; revised 28 August 2009; accepted 19 November 2009)

An optimized method is presented for the numerical evaluation of the sound field generated by an incoherent line source, which is commonly used to model road and rail traffic noise. Two different solutions for the numerical integration over the line source are distinguished, a point source solution and a line source solution. With proper segmentation of the line source, both solutions yield accurate results. Special attention is paid to receiver positions close to the (infinite) line through the (finite) line source. At these positions, conventional methods give numerical errors, which occur frequently in calculations of large-scale noise maps of cities, employing automatically generated geographical input data. The problems are avoided by using the optimized method presented here. The method is based on a combination of angular segmentation and linear segmentation of the line source and can be used to minimize the number of point-to-point calculations for noise mapping.
 [DOI: 10.1121/1.3273890]

PACS number(s): 43.50.Lj, 43.50.Rq, 43.28.Js, 43.50.Sr [KVH] Pages: 796–803

I. INTRODUCTION

To regulate and control road and rail traffic noise in the European Union, the European Commission requires that member states periodically calculate noise maps of major cities, highways, and railways.[1] Consequently, there is currently great interest in efficient noise-mapping algorithms based on practical engineering models for traffic noise.

Engineering models for traffic noise are commonly based on the concept of *incoherent line source*.[2–7] In the case of road traffic noise this means that the sound level at a receiver near a road is calculated by incoherent summation of contributions from vehicles moving along the road. For numerical calculations the road is divided into straight sections, and each section is represented by a line source at a fixed height above the road section. Each line source is divided into small segments, and the sound level is usually expressed as a sum over angular sectors corresponding to the segments.[3] This formulation fails if the receiver is located on the line through the line source, as illustrated in Fig. 1. In practice, a numerical problem occurs when the receiver is *near* the line through the line source, for example, at a distance of 1 m or less. The problem is not serious for receiver positions such as position A in Fig. 1 since these are not relevant receiver positions in practice, but the problem is serious for receiver positions such as positions B and C. The end points of the line source shown in Fig. 1 may represent intersections where several road sections converge, or they may represent points where an underground tunnel starts.

The problem is of considerable practical interest, as noise-mapping calculations for a city usually employ automatically generated geographical input data, and it frequently happens that a receiver is near the line through a line source. Furthermore, a traffic noise algorithm should also be applicable to test situations in which the position of a receiver is chosen exactly on the line through a line source. The problem is basically a division by zero, which can be avoided with an approach described in this article.

The results presented here may be useful for a harmonized European traffic noise model, which is expected to be developed for future noise-mapping rounds for the European Commission. Currently the noise maps are calculated with various national traffic noise models or with so-called *interim* models: the French road traffic noise model[8] and the Dutch rail traffic noise model[9,10] (the interim models are used by member states that have no national model). The two interim models employ different solutions for the numerical integration over the line source, a point source solution and a line source solution, respectively (these solutions are described in this article). The line source solution has the numerical problem described above, while the point source solution avoids the problem. On the other hand, the line source solution is more accurate than the point source solution for many receiver positions, unless an optimized segmentation algorithm is used. Such a segmentation algorithm is presented in this article. The algorithm can be used to minimize the number of segments used to represent a line source and consequently to minimize the number of point-to-point calculations for noise mapping.

In Sec. II the modeling of a traffic stream by an incoherent line source is described, and an integral equation is derived for the equivalent sound level generated by an incoherent line source. In Sec. III the numerical evaluation of the integral equation is described. In Sec. IV an optimized algorithm for the numerical evaluation is described.

II. EQUIVALENT SOUND LEVEL GENERATED BY A TRAFFIC STREAM

Vehicles are modeled as point sources moving with constant speed along a horizontal line segment at height z_S above a straight road section on a horizontal ground surface. An expression for the sound level at a receiver at height z_R is derived in this section.

[a]Author to whom correspondence should be addressed. Electronic mail: erik.salomons@tno.nl

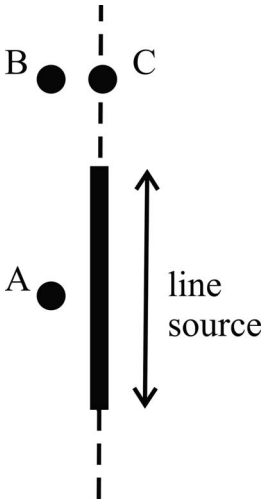


FIG. 1. Line source (thick line), line through the line source (dashed line), and three receiver positions A, B, and C.

Figure 2 shows the geometry in the horizontal xy plane for the special case that the source height z_S is equal to the receiver height z_R. The line source is located on the y axis between points $(0,y_1)$ and $(0,y_2)$. A point source at position $(0,y)$ on the line source is indicated in the figure. The receiver is located at position (x_R,y_R) at normal distance $d=|x_R|$ from the line source. Also indicated in the figure are angle ϕ and source-receiver distance r. Angle ϕ runs from $\phi_1=\arctan[(y_1-y_R)/d]$ to $\phi_2=\arctan[(y_2-y_R)/d]$. Distance r is given by $r=[d^2+(y_R-y)^2]^{1/2}$.

The analysis presented in this article also applies to the general three-dimensional case with different source height z_S and receiver height z_R. In this case the horizontal xy plane is replaced with the tilted plane through the line source and the receiver, and the formulas derived in the following remain valid if we replace $d=|x_R|$ with $d=[x_R^2+(z_R-z_S)^2]^{1/2}$.

We first consider a single point source moving with constant speed v along the y axis from $y=y_1$ to $y=y_2$. The source position as a function of time t is given by $y=y_1+vt$. The travel time is $T=(y_2-y_1)/v$.

The sound pressure level L_p at the receiver at time t is

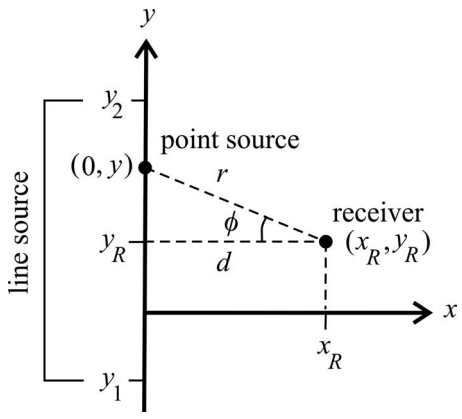


FIG. 2. Geometry with a point source at position $(0,y)$ on a line source on the y axis between points $(0,y_1)$ and $(0,y_2)$, and a receiver at position (x_R,y_R), at normal distance $d=|x_R|$ from the line source.

$$L_p = L_W - 10\log 4\pi r^2 - A_{\text{excess}}, \tag{1}$$

where L_W is the sound power level of the source, r is the source-receiver distance at time t (we neglect the small difference between emission time and immission time), and A_{excess} is the excess attenuation, which accounts for the effects of ground reflections, screening by noise barriers, atmospheric refraction, and atmospheric absorption.[11] The quantities L_W and A_{excess} are in general functions of frequency, and the broadband sound level is calculated by logarithmic summation of the spectral levels resulting from Eq. (1). The A-weighting is applied to L_W to account for the sensitivity of the human ear. Further, the quantities L_W and A_{excess} are in general functions of y or, equivalently, functions of the emission angle ϕ shown in Fig. 2. The Doppler effect for a moving source is assumed to be included (approximately) in the directivity function $L_W(\phi)$. The equivalent sound pressure level over the travel time T is

$$L_{\text{eq},1} = 10\log\left(\frac{1}{T}\int_0^T 10^{(L_W-10\log 4\pi r^2-A_{\text{excess}})/10}dt\right), \tag{2}$$

where index 1 indicates that we consider a single point source. A change in integration variable from t to y yields

$$L_{\text{eq},1} = 10\log\left(\frac{1}{y_2-y_1}\int_{y_1}^{y_2} 10^{(L_W-10\log 4\pi r^2-A_{\text{excess}})/10}dy\right). \tag{3}$$

This expression corresponds to the case of a single vehicle moving in time T from $y=y_1$ to $y=y_2$. In general, there are several vehicles simultaneously on the line segment from $y=y_1$ to $y=y_2$. We write the average number of vehicles per unit length as Q/v, where Q is the number of vehicles passing per unit time. The average number of vehicles on the line segment from $y=y_1$ to $y=y_2$ is then $(y_2-y_1)Q/v$. The equivalent sound pressure level becomes

$$L_{\text{eq}} = 10\log\left(\int_{y_1}^{y_2} 10^{(L_W-10\log 4\pi r^2-A_{\text{excess}})/10}dy\right) + 10\log\left(\frac{Q}{v}\right). \tag{4}$$

The integral in Eq. (4) can also be written as an integral over the angle ϕ given by $\phi=\arctan[(y-y_R)/d]$, using the relation $d\phi/dy=d/[d^2+(y-y_R)^2]$:

$$L_{\text{eq}} = 10\log\left(\int_{\phi_1}^{\phi_2} 10^{(L_W-10\log 4\pi d-A_{\text{excess}})/10}d\phi\right) + 10\log\left(\frac{Q}{v}\right). \tag{5}$$

Equations (4) and (5) represent a uniform distribution of incoherent point sources over the line segment from $y=y_1$ to $y=y_2$. The distribution is sometimes called an *incoherent line source*. We have neglected any coherence between different points of the line segment. This is based on the fact that the coherence length along the line segment is small due to the random nature of the noise emitted by cars and trains. In addition, atmospheric turbulence causes loss of coherence of

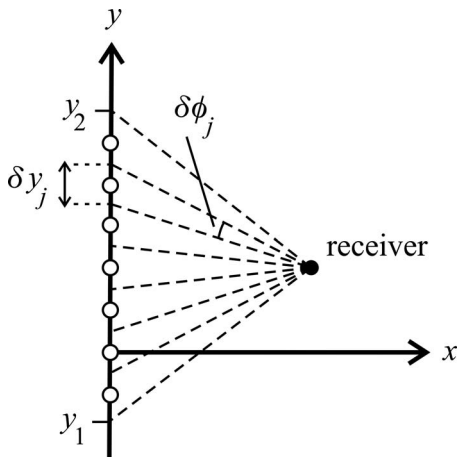


FIG. 3. Segmentation of the line source on the y axis, between $y=y_1$ and $y=y_2$, with point sources (open circles) at the centers of the segments. Indicated is length δy_j and angular sector $\delta\phi_j$ for segment j.

sound waves emitted at different points of the line segment.[4,11]

Road and rail traffic noise can be modeled with incoherent line sources. If a single car or a train is represented by more than one point source (e.g., to represent tire-road noise and engine noise separately), each point source corresponds to an incoherent line source. Different traffic lanes of a road can be represented by different incoherent line sources.

III. NUMERICAL FORMULATION

A. Discretization

For the numerical calculation of the level L_{eq} one can use Eq. (4) or Eq. (5). The discrete representation of Eq. (4) is

$$L_{\mathrm{eq}} = 10\log\left(\sum_j 10^{(L_W - 10\log 4\pi r_j^2 - A_{\mathrm{excess},j})/10}\delta y_j\right) + 10\log\left(\frac{Q}{v}\right), \tag{6}$$

where index j runs over a set of small contiguous segments δy_j between $y=y_1$ and $y=y_2$ (see Fig. 3). We have omitted index j for level L_W for simplicity. The segment lengths δy_j may be constant or they may vary, depending on the segmentation algorithm (see Sec. III C). Each segment is represented by a point source at the center of the segment. The quantity $A_{\mathrm{excess},j}$ is the excess attenuation of the field of the point source and r_j is the distance from the point source to the receiver. The excess attenuations $A_{\mathrm{excess},j}$ for all point sources j can be calculated with practical engineering formulas, as described in Sec. III B.

The discrete representation of Eq. (5) is

$$L_{\mathrm{eq}} = 10\log\left(\sum_j 10^{(L_W - 10\log 4\pi d - A_{\mathrm{excess},j})/10}\delta\phi_j\right) + 10\log\left(\frac{Q}{v}\right), \tag{7}$$

where $\delta\phi_j$ is the angular sector of segment j with length δy_j (see Fig. 3). Equations (6) and (7) can be written as follows:

$$L_{\mathrm{eq}} = 10\log\left(\sum_j 10^{L_j/10}\right), \tag{8}$$

where

$$L_j = L_W + 10\log\left(\frac{Q}{v}\right) - A_{\mathrm{geo},j} - A_{\mathrm{excess},j} \tag{9}$$

is the contribution from segment j, with geometrical attenuation $A_{\mathrm{geo},j}$ given by

$$A_{\mathrm{geo},j} = -10\log\left(\frac{\delta y_j}{4\pi r_j^2}\right) \tag{10}$$

in the case of Eq. (6) and

$$A_{\mathrm{geo},j} = -10\log\left(\frac{\delta\phi_j}{4\pi d}\right) \tag{11}$$

in the case of Eq. (7). The numerical solution based on Eq. (10) will be referred to as *point source solution* and the solution based on Eq. (11) will be referred to as *line source solution*.

The point source solution is employed by the French road traffic noise model, which is the current interim model for road traffic noise mapping in Europe.[8] The line source solution is employed by the Dutch models for road and rail traffic noise (with a monopole directivity for cars and a dipole directivity for trains).[9,10] The Dutch rail traffic noise model is the current interim model for rail traffic noise mapping in Europe. A review of national rail traffic noise models in Europe[12] showed that about half of the models employ the point source solution while the other half employ the line source solution.

In the case of free-field conditions ($A_{\mathrm{excess}}=0$) and omnidirectional point sources (L_W independent of ϕ), Eq. (5) can be integrated analytically and we find the following exact solution:

$$L_{\mathrm{eq}} = L_W + 10\log\left(\frac{Q}{v}\right) + 10\log\left(\frac{\Phi}{4\pi d}\right), \tag{12}$$

where $\Phi=\Sigma_j\delta\phi_j$ is the total angular sector of the line source. For $d\to 0$ and $y_R>y_2$ or $y_R<y_1$ this solution approaches the following expression:

$$L_{\mathrm{eq}} = L_W + 10\log\left(\frac{Q}{v}\right) + 10\log\left(\frac{(y_R-y_2)^{-1}-(y_R-y_1)^{-1}}{4\pi}\right), \tag{13}$$

as follows from the expression $\Phi=\arctan[d/(y_R-y_2)] - \arctan[d/(y_R-y_1)]$ for $d>0$; here we have used the trigonometric relations $\arctan(x)=\pi/2-\arctan(1/x)$ and $\arctan(x)\approx x$ for small x. Equation (13) can also be derived by analytical integration of Eq. (4) with $A_{\mathrm{excess}}=0$, L_W =constant, and $d=0$.

In practice, $A_{\mathrm{excess},j}$ is usually nonzero, but for the development of an optimized numerical solution, the free-field situation is useful since for this case we have the above exact solution as a reference for different numerical approximations.

B. Excess attenuation

To describe the calculation of the excess attenuation $A_{\text{excess},j}$ for segment j, we consider a point source at position (x_S, y_S, z_S) and a receiver at position (x_R, y_R, z_R), where we omit subscript j of the point source for simplicity. The excess attenuation is written as

$$A_{\text{excess}} = A_{\text{ground}} + A_{\text{air}}, \tag{14}$$

where A_{ground} is the ground attenuation and A_{air} is the attenuation due to atmospheric absorption. For the calculation of the excess attenuation we consider two different models:

(i) the empirical Dutch standard model for road or rail traffic noise,[9,10] which is similar to the ISO 9613-2 model[13] and
(ii) the analytical solution for a monopole point source above a rigid ground surface.

It should be noted that the analytical solution (ii) is employed in recent European traffic noise models,[14,15] which may be considered as a basis for a harmonized European traffic noise model for future noise-mapping rounds. The remainder of this section presents a description of the models (i) and (ii) as used for numerical examples presented in Secs. III D and IV C

The analytical solution (ii) follows from the expression $\exp(ikR_1)/R_1 + \exp(ikR_2)/R_2$ for the sound pressure generated by a monopole source above a rigid ground surface:

$$A_{\text{ground}} = -20 \log \left| 1 + \frac{R_1}{R_2} \exp(ikR_2 - ikR_1) \right|, \tag{15}$$

where $R_1 = [(x_S - x_R)^2 + (y_S - y_R)^2 + (z_S - z_R)^2]^{1/2}$ is the direct source-receiver distance, $R_2 = [(x_S - x_R)^2 + (y_S - y_R)^2 + (z_S + z_R)^2]^{1/2}$ is the distance between the image source and the receiver, and k is the wavenumber. The atmospheric absorption attenuation is given by $A_{\text{air}} = \alpha R_1$, where α is the absorption coefficient, which we calculate with the ISO model[16] for a temperature of 15 °C and a relative humidity of 70% (we use $R_1 \approx R_2$ for the atmospheric absorption contribution, so equal atmospheric absorption values are used for the direct and reflected paths). Octave-band spectra of the excess attenuation are calculated from narrow-band spectra with 12 frequencies per octave band (the 1/12-octave band center frequencies). The analytical solution (15) is valid for a non-refracting atmosphere. For practical predictions of traffic noise levels, effects of refraction and turbulent scattering cannot be neglected, but for the purpose of this article, it is useful to consider an analytical solution for a non-refracting atmosphere.

With the Dutch standard traffic noise model, the ground attenuation is given by $A_{\text{ground}} = -3\gamma_0 - 6$ for 63 Hz and $A_{\text{ground}} = -3\gamma_0 - 2$ for 125 Hz–8 kHz, with $\gamma_0 = \max[0, 1 - 30(z_S + z_R)/R_0]$, where $R_0 = [(x_S - x_R)^2 + (y_S - y_R)^2]^{1/2}$ is the horizontal source-receiver distance. These formulas are empirical formulas that are valid for downwind conditions (similar formulas are employed with the ISO model[13]). The atmospheric absorption attenuation is given by $A_{\text{air}} = \alpha R_1$, and the model employs the values 0, 0, 1, 2, 4, 10, 23, and 58 dB/km for the atmospheric absorption coefficient α for the octave bands 63 Hz–8 kHz, respectively.

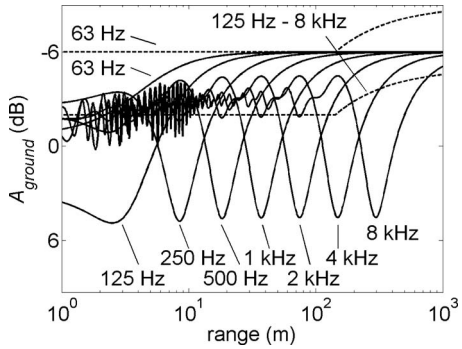


FIG. 4. Ground attenuation as a function of range, for a source height of 0.75 m, a receiver height of 4 m, and a rigid ground surface, calculated with the Dutch standard model for traffic noise (Refs. 9 and 10) (dashed lines) and the analytical solution (15) (solid lines) for the eight octave bands 63 Hz–8 kHz.

Figure 4 shows the ground attenuation A_{ground} for the eight octave bands as a function of horizontal source-receiver distance (range) for a point source at height of 0.75 m and a receiver at height of 4 m. Dashed lines represent the Dutch standard model and solid lines represent the analytical solution (15). The analytical solution shows marked interference minima and maxima, which in practical situations are partly reduced by coherence loss due to atmospheric turbulence.[11] The Dutch standard model solution is a smooth function of range (for absorbing ground the solution shows larger variations with range).

C. Segmentation

In the case of free-field conditions and omni-directional point sources, the most efficient segmentation of the line source is with constant angles $\delta\phi_j$ (unless the receiver is very close to the line through the line source). This follows from the exact solution (12), which implies that equal angular intervals have equal contributions to the sound level. One may use angles $\delta\phi_j$ of 5° for example.

In practice the excess attenuation is always nonzero, so the most efficient segmentation differs from the constant angle segmentation. In general, however, the excess attenuation can be considered as a correction to the geometrical attenuation, so the constant angle segmentation can be used in general as a practical approach. A complete avoidance of all possible numerical errors due to finite segment lengths would require an advanced segmentation algorithm that would adjust segment lengths to variations in the excess attenuation and source emission. In this article we develop a more practical segmentation algorithm that can be used for large-scale noise-mapping calculations, assuming more or less smooth functions for the excess attenuation (such as shown in Fig. 4) and source emission directivity.

If the receiver approaches positions at the line through the line source, such as position C in Fig. 1, the total angular sector Φ of the line source approaches zero, and the constant angle segmentation becomes unpractical. In this case one may turn to constant segment lengths δy_j. This is illustrated in Sec. III D.

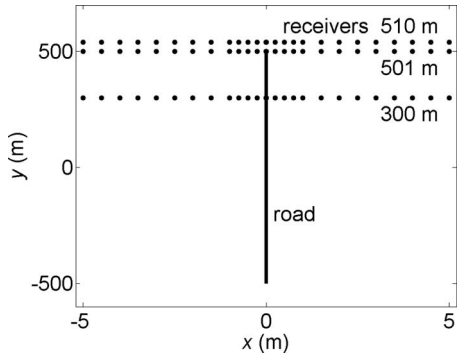


FIG. 5. Geometry with a 1 km long road and three lines of receivers at $y=300$, 501, and 510 m.

D. Numerical examples

We consider the geometry shown in Fig. 5, with a 1 km long road, and three lines of receivers at $y=300$ m, $y=501$ m, and $y=510$ m. We use the values $v=100$ km/h and $Q=1000$ h^{-1} (1000 passenger cars per hour) and calculate the A-weighted emission spectrum L_W with the Dutch standard traffic noise model[10] (there is no emission directivity with this model):

$$L_W = 74.5, 86.9, 92.6, 96.5, 103.7, 101.2, 93.0, 82.1 \text{ dB}$$

for the eight octave bands with center frequencies of 63 Hz–8 kHz, respectively.

For the segmentation of the line source we used

(1) constant angles $\delta\phi_j=5°$ for receivers at $y=300$ m and
(2) constant segment lengths $\delta y_j=25$ m for receivers at $y=501$ m and $y=510$ m.

This segmentation leads to considerable numerical errors, as shown by the figures presented below. In Sec. IV an improved segmentation algorithm is presented.

Figure 6 shows results for the free-field case ($A_{\text{excess}}=0$), with equal source and receiver height. The graph shows two solutions: (i) the line source solution based on Eq. (11) and (ii) the point source solution based on Eq. (10). The line source solution is identical to the exact solution in this free-field case, and therefore can be considered as a reference solution for assessing the accuracy of the point source solution. The point source solution deviates considerably from the reference solution. The deviation can be reduced by using smaller segment lengths δy_j, but clearly the line source solution is more efficient near the line through the road in this free-field case.

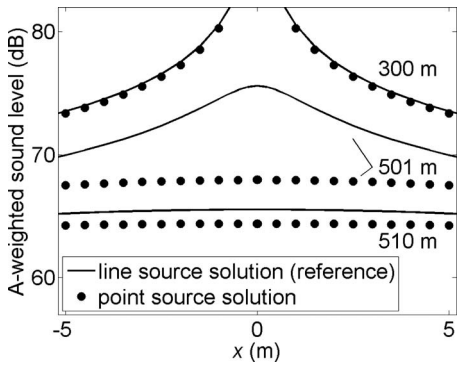


FIG. 6. A-weighted sound level calculated for the receivers shown in Fig. 5 for the free-field case ($A_{\text{excess}}=0$) with the line source solution based on Eq. (11) and the point source solution based on Eq. (10). The line source solution is identical to the exact solution in this free-field case.

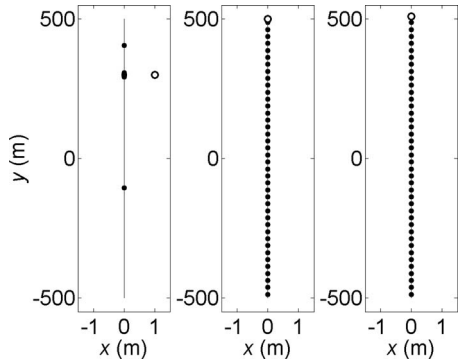


FIG. 7. Illustration of the segmentation used for Fig. 6 for three receiver positions, $(x,y)=(1,300)$ (left), (0,501) (middle), and (0,510) (right). The open circles represent the receivers and the dots on the line source represent point sources at the centers of the segments. The numbers of segments (and point sources) are 36, 40, and 40, respectively.

Figure 7 illustrates the segmentation used for Fig. 6 for three receiver positions, $(x,y)=(1,300)$, (0,501), and (0,510). The dots on the line source represent point sources at the centers of the segments. For the receiver at position (1,300) there are 36 segments of 5°, concentrated near the receiver due to the angular segmentation. For the receivers at positions (0,501) and (0,510) there are 40 segments of 25 m distributed uniformly over the road of 1 km.

Figure 8 is similar to Fig. 6, but now the excess attenuation based on the analytical solution (15) for a rigid ground surface has been used, with a source height $z_S=0.75$ m and a receiver height $z_R=4$ m. The line source solution is not an exact solution in this case, so we have included an accurate reference solution in the graph, calculated with small segments $\delta y_j=0.5$ m (for this segmentation the line source solution and the point source solution agree). Both the line source solution and the point source solution deviate from the reference solution.

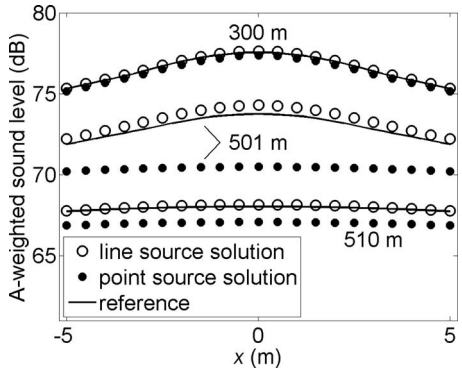


FIG. 8. As Fig. 6, with the excess attenuation based on the analytical solution (15) for a rigid ground surface for a source height $z_S=0.75$ m and a receiver height $z_R=4$ m. The line source solution is not an exact solution in this case, so we have included an accurate reference solution in the graph calculated with small segments $\delta y_j=0.5$ m.

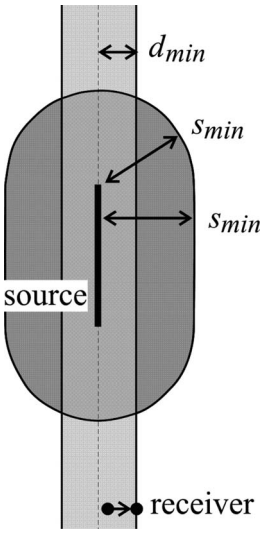


FIG. 9. Line source (thick solid line) with region $s<s_{\min}$ where receivers are excluded (dark gray area) and region $d<d_{\min}$ near the line through the line source where the receiver is shifted to the nearest point with $d=d_{\min}$ for calculating the geometrical attenuation (light gray area). The shift of the receiver is shown schematically at the bottom of the figure.

For receivers at $y=300$ m, Fig. 6 illustrates that the sound level diverges for $x=0$. This is not a problem since receivers *on* the road, or very close to the road, are not relevant in practice. We adopt a practical approach: we consider only receivers that are at least a minimum distance of 1 m away from the line source. In mathematical terms, we exclude receivers in the region $s<s_{\min}$, where s is the shortest distance from the receiver to the line source, and we use $s_{\min}=1$ m (see Fig. 9 described in Sec. IV A).

IV. OPTIMIZED ALGORITHM

In this section an improved algorithm is presented to calculate the sound level of an incoherent line source. Section IV A describes how a numerical problem near the line through the line source can be avoided. Section IV B presents an optimized segmentation algorithm. In Sec. IV C numerical examples are presented.

A. Numerical problem near the line through the line source

Near the line through the line source, both d and $\delta\phi_j$ approach zero, so the numerical evaluation of the line source geometrical attenuation (11) is a problem.

An *incorrect* approach to solve this problem would be to simply apply a lower limit to the denominator d in Eq. (11) because the numerator $\delta\phi_j$ also approaches zero if d approaches zero and $y_R>y_2$ or $y_R<y_1$. Although it may seem obvious that this approach is incorrect, we found deviations of more than 10 dB for $d=0$ and $y_R>y_2$ with commercial software based on the standard Dutch traffic noise model (the software does not specify how situations with small values of d are handled).

A *correct* approach is to apply consistent lower limits to both d and $\delta\phi_j$. To achieve this, we simply avoid receiver positions at a normal distance d smaller than $d_{\min}$ from the line through the line source. If d is smaller than $d_{\min}$, then we shift the receiver to normal distance $d_{\min}$, as illustrated in Fig. 9. Equation (11) for the geometrical attenuation is evaluated for the shifted receiver position, so with $d=d_{\min}$ and corresponding angular sectors $\delta\phi_j$. In addition, we exclude receivers in the region $s<s_{\min}$, with $s_{\min}=1$ m, as described in Sec. III D and also illustrated in Fig. 9.

It should be noted that Fig. 9 shows a single plane through the line source, for example, the horizontal plane. In three dimensions, the region $d<d_{\min}$ is a cylindrical volume around the line source and the region $s<s_{\min}$ is a cylindrical volume with rounded ends around the line source. A receiver on a viaduct over a road, for example, may be located outside these volumes.

B. Optimized segmentation algorithm

With the shift $d\rightarrow d_{\min}$ described in Sec. IV A, the total angular sector $\phi_2-\phi_1$ remains always finite, so the segmentation can be based on constant angles $\delta\phi_j$ for all receiver positions. To reduce numerical segmentation errors, the segmentation is refined, as described in this section.

Taking into account the shift $d\rightarrow d_{\min}$, the angles ϕ_2 and ϕ_1 are given by $\phi_1=\arctan[(y_1-y_R)/d_x]$ and $\phi_2=\arctan[(y_2-y_R)/d_x]$, with $d_x=\max(d,d_{\min})$. First, the angular interval $[\phi_1,\phi_2]$ is divided into N_{seg} angular sectors $\delta\phi_j=(\phi_2-\phi_1)/N_{\text{seg}}$ with $N_{\text{seg}}=\max[N_{\text{seg,min}},R((\phi_2-\phi_1)/\delta\phi_{\max})]$. Here function $R(x)$ rounds x to the nearest larger integer, and $N_{\text{seg,min}}$ and $\delta\phi_{\max}$ are numerical parameters.

Next, some segments are divided into smaller segments. A maximum segment length is calculated as

$$\delta y_{\max}=\min\left(\frac{r}{3},\delta y_{\max,0}\right), \tag{16}$$

where $\delta y_{\max,0}$ is a numerical parameter. Each segment with length δy larger than $\delta y_{\max}$ is broken down into a minimum number n of (equally long) segments shorter than $\delta y_{\max}$, with $n=R(\delta y/\delta y_{\max})$.

The term $r/3$ in Eq. (16) ensures that the spherical nearfields of the point sources on the line source are avoided. This is based on an analysis of the field of a row of point sources,[2] which shows that a transition from spherical to cylindrical spreading occurs at a distance of $r/\pi\approx r/3$. The parameter $\delta y_{\max,0}$ in Eq. (16) may be used for refined sampling in situations with large variations in the excess attenuation, for example, due to varying ground properties or varying screening near a vertical edge of a noise barrier.

C. Numerical examples

In this section we present numerical examples of calculations with the optimized algorithm described in Secs. IV A and IV B. We used the following numerical parameters: $\delta\phi_{\max}=5°$, $N_{\text{seg,min}}=10$, $\delta y_{\max,0}=100$ m, and $d_{\min}=0.1$ m. The examples are for the same geometry and source parameters as considered in Sec. III D.

Figure 10 is analogous to Fig. 6 presented in Sec. III for a free-field situation. We recall that the line source solution is identical to the exact solution in this free-field case and therefore can be considered as a reference solution for assessing the accuracy of the point source solution. While Fig.

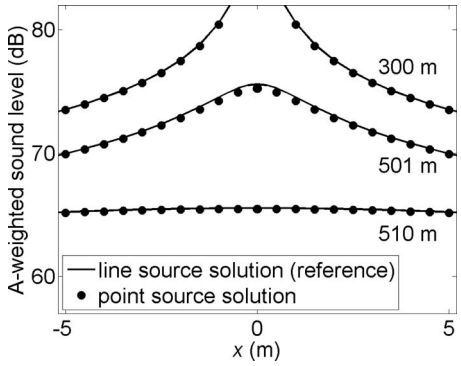


FIG. 10. As Fig. 6, calculated with the optimized algorithm described in Secs. IV A and IV B.

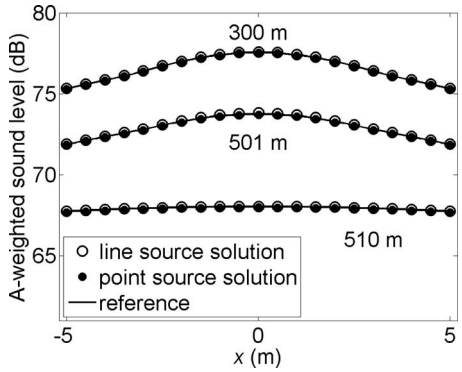


FIG. 12. As Fig. 8, calculated with the optimized algorithm described in Secs. IV A and IV B.

6 shows that considerable deviations between the reference solution and the point source solution occur if the basic segmentation algorithm described in Sec. III is used, Fig. 10 shows that the deviations are negligible if the optimized algorithm described in Secs. IV A and IV B is used.

Figure 11 illustrates the segmentation used for Fig. 10 for three receiver positions, $(x,y)=(1,300)$, (0,501), and (0,510), analogous to the segmentation shown in Fig. 7 (for Fig. 6). For the receiver at position (1,300), the number of segments is 52, which is a bit larger than the number of 36 segments in Fig. 7. For the receivers at positions (0,501) and (0,510), the number of segments is 21, which is even smaller than the number of 40 in Fig. 7, while the agreement between the point source solution and the reference solution is better in Fig. 10 than in Fig. 6.

Figure 12 is analogous to Fig. 8 presented in Sec. III for the excess attenuation based on the analytical solution for a rigid ground surface, with a source height $z_S=0.75$ m and a receiver height $z_R=4$ m. Also in this case the optimized algorithm yields more accurate results, both with the line source solution and with the point source solution.

The distribution of segments over the line source for Fig. 12 is very similar to the distribution shown in Fig. 11 for the case of equal source and receiver heights. The numbers of segments for the three receiver positions shown in Fig. 11 are 51, 26, and 21, respectively.

Figure 13 is similar to Fig. 12, but now the Dutch standard model has been used for the excess attenuation. The agreement with the reference solution is similar to the agreement in Fig. 12. The A-weighted sound levels in Fig. 13 are slightly lower than the levels in Fig. 12.

V. CONCLUSIONS

The numerical modeling of traffic noise has been studied. Two different solutions for the sound field of an incoherent line source have been compared, a point source solution and a line source solution. It has been shown that conventional methods for calculating traffic noise may yield considerable numerical errors, in particular, for receiver positions near the infinite line through the line source. These errors occur frequently with large-scale noise mapping of cities, employing automatically generated geographical input data. The errors originate from improper segmentation of the line source and from improper evaluation of the geometrical attenuation for the line source solution.

An optimized method has been presented that avoids the errors. The problem with the geometrical attenuation is avoided by shifting receivers that are closer to the line through the line source than a minimum distance of typically 0.1 m. An optimized general segmentation algorithm of the line source is used to avoid segmentation errors for receivers close to the line through the source, at least for situations with more or less smooth functions for the excess attenuation and source directivity, as commonly employed by traffic

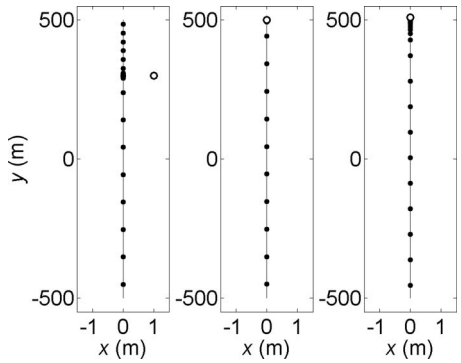


FIG. 11. Illustration of the segmentation used for Fig. 10 for three receiver positions. The numbers of segments (and point sources) are 52, 21, and 21, respectively.

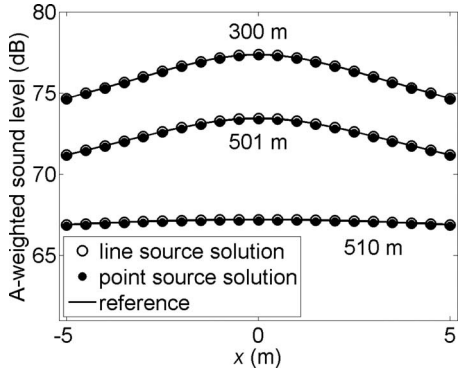


FIG. 13. As Fig. 12, but now calculated with the excess attenuation according to the Dutch standard model.

noise engineering models. The segmentation algorithm can be used to minimize the number of point-to-point calculations for large-scale noise mapping of cities.

With the optimized method presented in this article, both the point source solution and the line source solution for the sound field of an incoherent line source yield accurate results and therefore can be used for application in practical engineering models for road and rail traffic noise.

[1]European Directive on Environmental Noise, 2002/49/EC, URL: http://ec.europa.eu/environment/noise/home.htm (website of the European Commission, last viewed 3/28 2009).

[2]E. J. Rathe, "Note on two common problems of sound propagation," J. Sound Vib. **10**, 472–479 (1969).

[3]J. D. Van der Toorn, "Measurements of sound emission by single vehicles," Noise Control Eng. **11**, 110–115 (1978).

[4]E. M. Salomons, "A coherent line source in a turbulent atmosphere," J. Acoust. Soc. Am. **105**, 652–657 (1999).

[5]E. M. Salomons, A. C. Geerlings, and D. Duhamel, "Comparison of a ray model and a Fourier—Boundary element method for traffic noise situations with multiple diffractions and reflections," Acta Acust. Acust. **83**, 35–47 (1997).

[6]R. Makarewicz and P. Kokowski, "Prediction of noise changes due to traffic speed control," J. Acoust. Soc. Am. **122**, 2074–2081 (2007).

[7]R. Makarewicz and M. Zoltowski, "Variations or road traffic noise in residential areas," J. Acoust. Soc. Am. **124**, 3568–3575 (2008).

[8]NMPB-Routes-96, Noise from road infrastructures, calculation method including meteorological effects," document prepared by the French Institutes CERTU, SETRA, LCPC, and CSTB, 1997 (in French).See also G. Dutilleux, J. Defrance, B. Gavreau, and F. Besnard, "The revision of the French method for road traffic noise prediction," J. Acoust. Soc. Am. **123**, 3150-3150 (2008). The full paper is available on the CD-ROM with the Proceedings of the Joint ASA-EAA Meeting in Paris, July 2008 (Acoustics 08).

[9]Dutch standard model for rail traffic noise, Appendix of Regulation LMV 2006 332519 of Dutch Ministry of Environment (in Dutch). The model provides a non-spectral method SRM1 and a spectral method SRM2; for this study the method SRM2 has been used. The model is in many ways similar to the Dutch standard model for road traffic noise (Ref. 10). The description of the model can be downloaded from URL http://www.stillerverkeer.nl. (maintained by the Dutch public authorities, last viewed 3/28/2009).

[10]See supplementary material at http://dx.doi.org/10.1121/1.3273890 E-JASMAN-126-051911, for Dutch standard model for road traffic noise ("Calculation and measurement method for road traffic noise 2002"), regulation LMV 2002 025825 of Dutch Ministry of Environment (in Dutch). The model provides a non-spectral method SRM1 and a spectral methof SRM2; for this study the method SRM2 has bee used, referred to as DMS (Dutch standard model). The description of the model can be downloaded from URL http://www.stillerverkeer.nl (Last viewed March 28, 2009). An English translation of the model description is available through the Electronic Physics Auxiliary Publication Service (EPAPS) of the American Institute of Physics.

[11]E. M. Salomons, *Computational Atmospheric Acoustics* (Kluwer, Dordrecht, 2001).

[12]H. J. A. van Leeuwen, "Railway noise prediction models: A comparison," J. Sound Vib. **231**, 975–987 (2000).

[13]ISO 9613-2:1996 Acoustics—Attenuation of sound during propagation outdoors—Part 2: General method of calculation (International Organization for Standardization, Geneva, Switzerland, 1996).

[14]H. G. Jonasson, M. Dittrich, D. van Maercke, J. Defrance, E. Salomons, I. Noordhoek, D. Heimann, B. Plovsing, G. Watts, X. Zhang, E. Premat, I. Schmich, F. Aballea, M. Baulac, F. de Roo, M. Bakermans, D. Kühner, B. De Coensel, D. Botteldooren, F. Vanhove, S. Logghe, and R. Bütikofer, "Building Europe's future harmonized noise mapping methods," Acta Acust. Acust. **93**, 173–262 (2007) (special issue about the European projects Harmonoise and Imagine).

[15]G. B. Jónsson and F. Jacobsen, "A comparison of two engineering models for outdoor sound propagation: Harmonoise and Nord2000," Acta Acust. Acust. **94**, 282–289 (2008).

[16]ISO 9613-1:1993, Acoustics—Attenuation of sound during propagation outdoors—Part 1: Calculation of the absorption of sound by the atmosphere (International Organization for Standardization, Geneva, Switzerland, 1993).

Noise-induced annoyance from transportation noise: Short-term responses to a single noise source in a laboratory

Jaehwan Kim,[a] Changwoo Lim,[b] Jiyoung Hong, and Soogab Lee
Center for Environmental Noise and Vibration Research, School of Mechanical and Aerospace Engineering, Seoul National University, Room 205, Building 44, Seoul 151-744, Republic of Korea

(Received 4 June 2009; revised 13 October 2009; accepted 22 November 2009)

An experimental study was performed to compare the annoyances from civil-aircraft noise, military-aircraft noise, railway noise, and road-traffic noise. Two-way within-subjects designs were applied in this research. Fifty-two subjects, who were naïve listeners, were given various stimuli with varying levels through a headphone in an anechoic chamber. Regardless of the frequency weighting network, even under the same average energy level, civil-aircraft noise was the most annoying, followed by military-aircraft noise, railway noise, and road-traffic noise. In particular, penalties in the time-averaged, A-weighted sound level (TAL) of about 8, 5, and 5 dB, respectively, were found in the civil-aircraft, military-aircraft, and railway noises. The reason could be clarified through the high-frequency component and the variability in the level. When people were exposed to sounds with the same maximum A-weighted level, a railway bonus of about 3 dB was found. However, transportation noise has been evaluated by the time-averaged A-weighted level in most countries. Therefore, in the present situation, the railway bonus is not acceptable for railway vehicles with diesel-electric engines.
 [DOI: 10.1121/1.3273896]

PACS number(s): 43.50.Qp, 43.50.Lj [BSF] Pages: 804–814

I. INTRODUCTION

Most research on the effects of transportation noise has been performed in advanced western countries or the EU and US. The majority of papers have reported that aircraft noise is the most annoying and railway noise is the least annoying.[1–3] The "railway bonus" refers to the fact that railway noise is less annoying than other transportation noise. However, several studies have revealed somewhat contradictory results.[4–6]

Recent field studies in Korea[7–9] have yielded different results with regard to dose-response relationships and the railway bonus in comparison with those in western countries.[3,5] Annoyance responses to aircraft and railway noise in Korea were higher than in western countries.[7,8] Road-traffic noise in Korea causes less sleep disturbance than railway noise and the self-reported sleep disturbance from road-traffic noise in Korea is similar to that in European countries.[9] Bonus effects were found not for railway noise but for road-traffic noise. The results of Japanese studies were similar to the Korean results.[7] Railway noise in Japan caused annoyance that was much higher than in European countries. The response to railway noise in Japan is reportedly about 10 dB higher than that to road-traffic noise.[10] Hui and Yano[11] suggested the possibility of a railway bonus in the case of auditory tasks and the impossibility of a railway bonus in the case of non-auditory tasks. In their experiments, road-traffic noise and railway noise were presented as maskers of speech signals. It was reported that in the context of auditory tasks, the cause of a railway bonus was masking effects.

The majority of investigations on the impacts on transportation noise have found that aircraft noise causes the greatest annoyance. There has been a great diversity of opinion about the impacts of ground vehicles (especially with regard to railway noise). Although a bonus of 5 dB has been applied to railway noise in many EU countries, it seems that railway bonuses are controversial and not justified in every situation. In Ref. 12, railway bonuses have not been applied to all types of railway vehicle but restricted to electric railway vehicles with 12–20 cars and speeds of less than 250 km/h.

This is a very complex problem. The annoyance response is affected by not only acoustical components but also non-acoustical factors that are social, environmental, psychological, and economic in nature.[13,14] Even with similar noise levels and sources, the annoyance response differs across countries because the annoyance response to noise is affected by several external factors including the source's spectral difference, cultural differences, language differences, variations in survey questions, and differences in climatic conditions.[13,14] More detailed studies for various types of vehicle should be carried out in diverse environments.

In the present experimental research, annoyance from four kinds of transportation noise was investigated in order to compare the responses for the respective noise sources. The cause of the differential response to noises was explored through an analysis of the factors that influence transportation noise annoyance. The possibility of railway bonuses is also discussed.

[a]Author to whom correspondence should be addressed. Electronic mail: kjh03@snu.ac.kr
[b]Present address: Department of Aircraft Noise Policy, Republic of Korea Air Force, Republic of Korea.

II. CHOICE IN SOUND REPRODUCTION

When research on the evaluation of subjective responses to noise is carried out using a headphone and loudspeaker in a laboratory, recorded sounds should be played to the subjects. Researchers have adopted various methods for sound recording and playback. With regard to sound recording, the choice is between a conventional microphone and an artificial dummy head. In general, a conventional microphone (1/2 in. condenser microphone) is used for the objective evaluation of environmental noise. Sounds at the entrance of the outer ear are different from those that are recorded by a conventional microphone because of head-related transfer functions (HRTFs). The discrepancy between the two sounds is caused by natural acoustic phenomena around the human head. This implies that sound recordings using an artificial dummy head ensure a better laboratory environment. Empirically, Kim *et al.*[15] compared the responses to traffic sounds recorded by a conventional microphone with those to traffic sounds that were recorded by a dummy head. The subjective responses to transportation noise showed remarkable variations according to the manner of sound recording and reproduction. Transportation noise annoyance also increased faster, i.e., by about 1%, in the presence of the binaural effect.

The second choice concerns devices for sound reproduction, i.e., headphones vs loudspeakers. Loudspeakers usually have been used in many laboratory studies. For binaural synthesis, two loudspeakers are required and the trans-aural cross-talk should be removed by a robust cross-talk cancellation algorithm. Though trans-aural cross-talk is a serious problem in laboratory studies on the impacts of environmental noise, the requirement for binaural synthesis has not been mentioned in most research papers on the present topic. Loudspeakers offer the advantage of natural hearing situations and the disadvantages of acoustic quality control, such as binaural synthesis, and spatial limitations, e.g., hearing rooms. Headphones offer the advantage of acoustic quality control and the disadvantage of unnatural hearing situations.

III. METHOD AND MATERIALS

A. Measurement and stimuli

1. Binaural recording and noise measurement

In this study, four kinds of noise (civil- and military-aircraft noise, railway noise, and road-traffic noise) were recorded by head and torso simulator (HATS) (Brüel & Kjær Type 4100) in order that the binaural sounds (HRTF-filtered sounds) could be heard by subjects. At the same time, noise measurement (HRTF-unfiltered sounds) was carried out using a free-field microphone (Brüel & Kjær Type 4190) for calculating the corresponding noise metrics, which are predictor variables for explaining the noise-induced annoyance.

It is practically impossible for two devices to be located at the same position simultaneously. The close (side-by-side) arrangement of a microphone and a dummy head causes acoustical interference in the sound field around the microphone by the dummy head and vice versa. It is also inevitable that the dummy head obstructs the sound from reaching the microphone. However, the noises that reach the two measuring devices are almost the same because the measurements are performed in an open environment with no obstacles, and the noise source is almost equidistant from each measuring device. Further, in the present study, the free-field microphone was located about 5 m away from the dummy head parallel to the forward direction of the noise sources in order to decrease the acoustical interference in each sound field.[16]

The two devices were placed about 15 m away from the side of the road and railway and about 100 m away from the runway of military and civil aircraft. They were located at about 1.7 m above the ground. Their output terminals were directly connected to a B&K's data acquisition and analysis system platform named as PULSE (Brüel & Kjær Type 3560C), which provides the functions of sound recording and real-time monitoring, including an analysis of signals in the time and frequency domains. A single microphone and two microphones that were inserted in both ears of the dummy head were calibrated before the noise measurement and binaural recording. Each transportation noise was simultaneously sampled by a dummy head and a microphone every 1/2 s[16].

Civil-aircraft, military-aircraft, and railroad-vehicle sounds were recorded through one pass-by. The civil-aircraft, military-aircraft, and railway vehicles were operated by a turbo fan, turbo jet, and diesel-electric engine, respectively. The railway vehicle had eight cars (one locomotive and seven passenger cars). The velocities of the aircraft and the railway vehicle were about 250 and 80–100 km/h, respectively. Road-traffic sound was recorded from a continuous stream of vehicles on the highway, whose velocities ranged from 80 to 120 km/h. This was done for considering general noise generation/exposure patterns that included both intermittent (aircraft and railway noises) and continuous (road-traffic noise) noises.

2. Test stimuli

The process for the construction of the test stimuli was as follows. An A-weighting network filter was constructed for calculating the time-averaged A-weighted sound level (TAL) for 15 s, $L_{Aeq,15\ s}$. The original signals (HRTF-unfiltered signals) were measured by a microphone pass through the filter. After this process, $L_{Aeq,15\ s}$ was calculated. The original values of the acoustic pressure (HRTF-unfiltered signals) were multiplied in order to change them into signals with the specific TAL value. These coefficients were acquired through MATLAB version 6.5 using the mathematical expression for the TAL for T s, $L_{Aeq,T}$. Then, the coefficients were also multiplied with the signals (HRTF-filtered acoustic pressure) that were recorded by the two microphones in the ears of the dummy head. A total of 36 stimuli (nine stimuli per noise source) with various noise levels were generated and converted into 36 wave files. The stimuli of civil- and military-aircraft, railway, and road-traffic noises were generated in increments of 5 dB in TAL from 50 to 85 dB. Additionally, a stimulus with a TAL of 40 dB was generated for each noise source.

The durations of stimuli might be contentious with regard to the reliability of annoyance ratings because of the

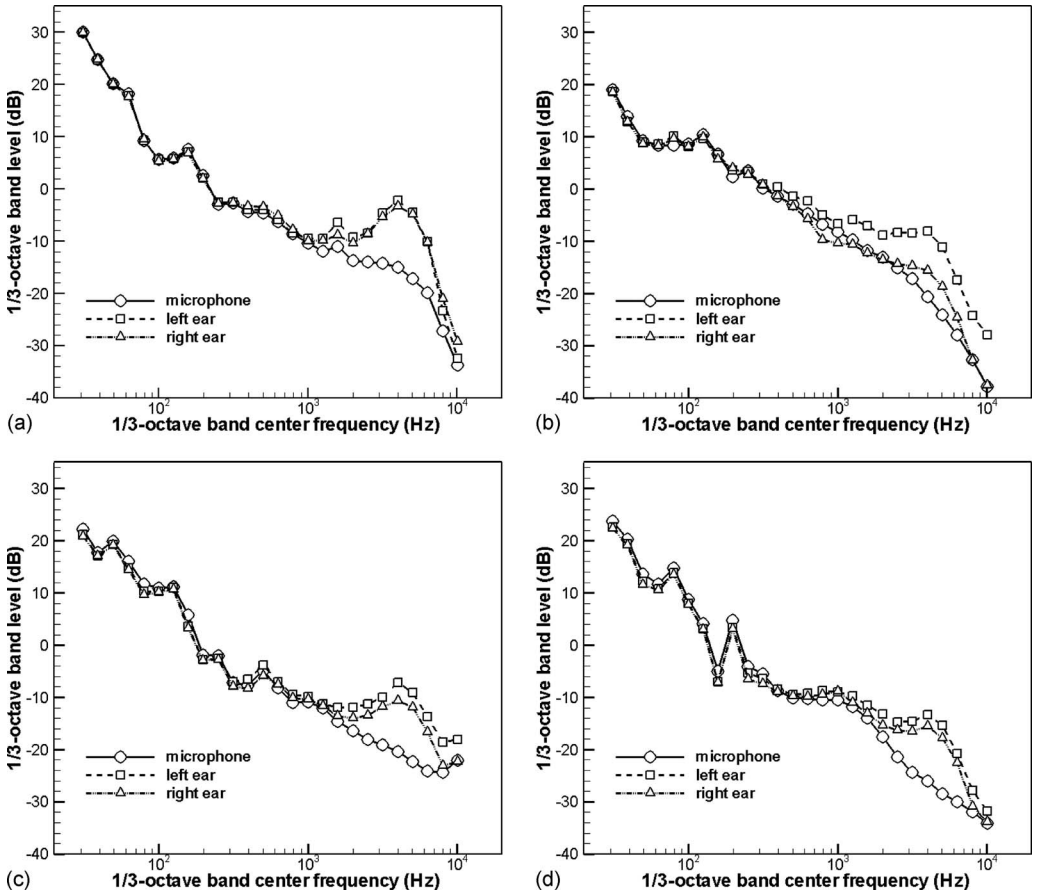


FIG. 1. Spectral difference between signals of a conventional microphone and both ears of HATS for (a) spectra of civil-aircraft noise, (b) spectra of military-aircraft noise, (c) spectra of railway noise, and (d) spectra of road-traffic noise. For all panels, 1/3-octave band levels were normalized to their time average A-weighted sound levels for 15 s.

varying exposure times that are encountered in real situations. In this study, the exposure time is considered by just one pass-by intermittent source in a general operating condition. Poulsen[16] reported that the exposure time did not have a significant effect on annoyance ratings. The choice of duration in the present study could be justified.

As shown in Fig. 1, the spectra of the stimuli differed from the signals that were measured by a conventional microphone. In particular, an acoustical boost by the pinna was found in the high-frequency bands. The TAL of stimuli was a little higher than that of the signals that were measured by the microphone. The level difference was less than 3 dB. The A-weighted statistical levels were found by means of Fig. 2. All the levels were acquired by exponential averaging every 1/8 s. For civil-aircraft, military-aircraft, railway, and road-traffic noises, the differences in levels between L10 and L50 were 12.8, 13.5, 11.7, and 1.3 dB, respectively, and between L10 and L90 were 24.5, 28.9, 25.3, and 3.6 dB, respectively.

B. Apparatus

1. Test section

An anechoic chamber was used as the test room for not only the audiometric screening test but also the subjective test in order to keep subjects from being disturbed by unwanted external sounds. The size of the anechoic chamber was $4.1 \times 4.1 \times 3.2$ m^3 and that of the test section was $3.2 \times 3.2 \times 2.1$ m^3. The absorbent material of the chamber was urethane foam. The cut-off frequency was approximately 200 Hz. Background noise in the anechoic chamber was less than 20 dB.

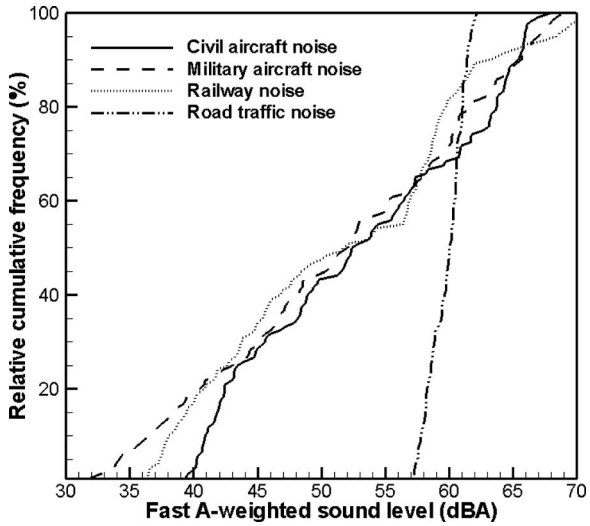


FIG. 2. Relative cumulative frequency of each transportation noise with time average A-weighted sound level 60 dB. Time integration was set to be fast (1/8 s exponential average). They were measured by a microphone, not an artificial dummy head (HATS).

2. Sound reproduction system

A system that could control the sound through a desktop personal computer (PC) (Pentium IV) with a RME's DIGI 96 Pro sound card was constructed in order to test the subjective responses to transportation noise. The output terminal of the sound card in the PC was connected to the front end of a power amplifier. The output stage of the power amplifier was interfaced with a headphone (Sennheiser HD25).

Each piece of the equipment for sound reproduction distorted the original input sound due to its frequency response; the distorted sound might have had a bad effect on the laboratory test. The sound field between the pinnae and ear couplers of the headphone distorted the sound. Therefore, acoustical correction had to be performed for the reliability of this research. A headphone was put on HATS and the output terminal of HATS was connected to the input terminal of PULSE (Brüel & Kjær Type 3560 C) for monitoring the output sound in terms of both the time and frequency domains. Sound monitoring was conducted using white noise as the input sound. The level differences between the input and output sounds in each frequency component were found to range from 20 Hz to 2 kHz. They were added to the test stimuli using a software (COOL EDIT PRO version 2.0) prior to the subjective test.

There was only one headphone in the anechoic chamber. Other devices were put out of the chamber so that subjects would not be disturbed by the noise from the sound reproduction equipment, such as fan noise in the computer, noise by operators, etc.

C. Participants

Fifty-two naïve subjects (32 male and 20 female) participated in the laboratory study. Their ages ranged from 20 to 35 years. The mean age was 26 and the standard deviation of the age was 3.2. Males were between 20 and 33 years of age (mean of 26.3 and standard deviation of 2.8). Females were between 21 and 35 years of age (mean of 25.6 and standard deviation of 3.8). All the participants had normal hearing [i.e., the hearing level (HL) was smaller than 15 dB of the reference equivalent threshold sound pressure level[17] in this research]. They were paid fees for their participation.

D. Laboratory test

1. Procedure of the laboratory test

All the subjects were screened audiometrically in order to filter participants who had abnormal hearing. The audiometric screening test was performed in the octave band center frequency between 20 Hz and 20 kHz according to the ascending method in Ref. 18. After the audiometric screening test, subjects were instructed on the procedure and experimental method. The main instructions were as follows.

- Whatever makes a noise or disturbs the experiment should not be permitted in the test section.
- Annoyance is different from the loudness of each noise event. Thus, the level or loudness of each presented sound must not be considered as the reference.

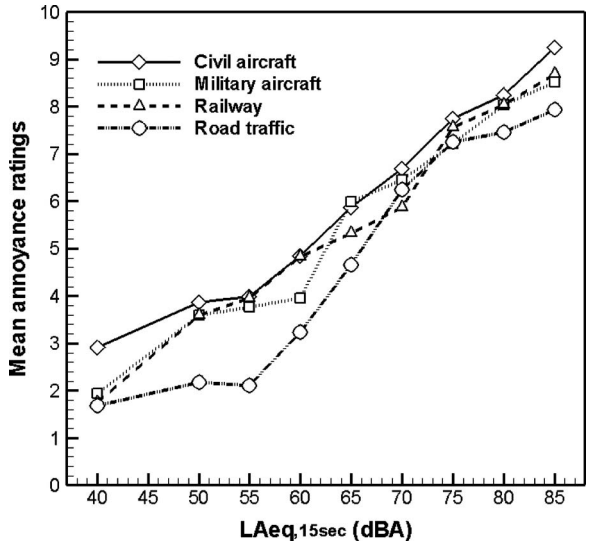


FIG. 3. Mean annoyance ratings for civil-aircraft noise, military-aircraft noise, railway noise, and road-traffic noise as a function of the time average A-weighted sound level of each noise.

- Annoyance is defined as a feeling of displeasure that is evoked by a noise or any feeling of resentment, displeasure, discomfort, and irritation when a noise intrudes into someone's thoughts and moods or interferes with their activity.
- Subjects may imagine reading a book, watching TV, or any similar activity.
- The scoring of annoyance on the questionnaire should be performed when no sound is heard following the respective noise event.

Integers between 1 and 36, inclusive, were assigned to all 36 stimuli. Random permutations of the integers between 1 and 36, inclusive, were generated using MATLAB. Each subject was presented with 36 stimuli in a random order. Each stimulus lasted for 15 s. Subjects evaluated the annoyance from each stimulus on the questionnaire for 15 s. Each experiment lasted about 90 min, including hearing tests and rest breaks.

2. Questionnaire and rating scale

The questionnaire included questions about the participant's age, gender, Weinstein's[19] noise sensitivity, and annoyance. Subjects were asked to answer the question in Korean, which translated to "What extent of annoyance would you feel if you had heard the noise in your common environment?" in English. They marked the extent of annoyance in the questionnaire after hearing each stimulus. In their study, Fields *et al.*[20] reported that a 0–10 numerical scale is likely to be easily understood by people of all countries and cultures who are familiar with currencies in a base-10 monetary system and other familiar counting situations. Team 6 of the International Commission on the Biological Effects of Noise (ICBEN) recommended a five-point (Not at all–Slightly–Moderately–Very–Extremely) verbal and a 0–10 (0 means "Not at all" and 10 means "Extremely") numerical scale for community annoyance ratings. Therefore, a 0–10

TABLE I. Mauchly's tests of sphericity. Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

Within-subject effects	Mauchly's W	Approx. chi-square	df	Sig.	Huynh–Feldt's epsilon
Noise level	0.186	80.915	35	<0.001	0.785
Vehicle type	0.841	8.614	5	0.126	0.943
Noise level * vehicle type	<0.001	608.129	299	<0.001	0.532

numerical scale was used as the annoyance rating scale in this research. The end points of this scale were labeled with the Korean equivalents of "not annoyed at all" and "very annoyed," respectively. The choice for the 11-point numerical scale is based on the assumption that respondents are more cognitively familiar with 0–10 scaling than with the shorter seven- or nine-point numeric scales.[21]

Weinstein's noise sensitivity scale (0–105) consists of 21 items, most of which express attitudes toward noise in general and emotional reactions to a variety of environmental sounds that are encountered in the everyday lives of students, who are the target population of the questionnaire.[22] For every statement, six response options ranging from strong disagreement to strong agreement are presented.[22] In 14 of the 21 items, agreement with the item indicates greater noise sensitivity of the respondent.[23]

IV. ANALYSIS AND RESULTS

A. Comparison across vehicle types

For civil-aircraft, military-aircraft, railway, and road-traffic noises, the mean annoyance ratings of subjects, as a function of the TAL of each noise, are shown in Fig. 3. Civil aircraft was more annoying than the other sources, while road traffic was the least annoying of all. Main (noise level and vehicle type) and interaction effects were tested in a repeated measure Analysis of Variance (ANOVA) (53 subjects × 4 vehicle types × 9 noise levels).

The assumptions in tests of within-subject effects are that the variances of the sets of data are equal (i.e., the errors are homoscedastic) and that the errors are uncorrelated with each other. Mauchly's test of sphericity is used to verify whether the assumptions of univariate models are met. If Mauchly's test shows violation of sphericity, this may be compensated by an epsilon adjustment. The correction is made by multiplying the degrees of freedom for the effect by the epsilon value and by multiplying the degrees of freedom for the error term by the epsilon value. The significance of the F-value is then determined using the corrected degrees of freedom. In this analysis, Huynh–Feldt's epsilon was used.

Tables I and II show the results of Mauchly's test of sphericity and within-subject effect tests, respectively. As shown, Mauchly's W for the vehicle type was not statistically significant ($p>0.05$). The noise level and interaction (noise level * vehicle type) violated the sphericity assumption ($p>0.05$). Huynh–Feldt's epsilon adjustment was required for the noise level and interaction. The F-test or adjusted F-test for two main effects and one interaction effect were statistically significant. The differences in the mean annoyance ratings across vehicle types were significant [$F(3,153)=32.460$, $p<0.001$]. The ratings increased with the time-averaged A-weighted sound levels [$F(6.728,320.155)=319.179$, $p<0.001$]. *Post hoc* pair-wise multiple comparison tests did not justify these conclusions at all paired conditions. The difference between military-aircraft noise and railway noise was not significant ($p>0.05$). The mean annoyance ratings at the TAL of 50 dB were not significantly different from those at the TAL of 55 dB ($p>0.05$). The mean annoyance ratings at the TAL of 75 dB were not significantly different from those at the TAL of 80 dB ($p>0.05$). The mean annoyance ratings at the TAL of 80 dB were not significantly different from those at the TAL of 85 dB ($p>0.05$). These results might be caused by nonlinearity, which asymptotically approaches either 0 (not annoyed at all) or 10 (very annoyed) at the lower and higher levels.

In Table III, it is seen that the slopes and intercepts in the linear annoyance models depend on the vehicle type. In

TABLE II. Repeated measure two-way ANOVA. Noise level and noise level * vehicle type violated sphericity assumption. F-tests adjusted by Huynh–Felt's epsilon for them were conducted. Vehicle type satisfied sphericity assumption.

Source		Sum of square	df	Mean square	F	Sig.
Noise level	Sphericity assumed	8818.779	8	1102.347	319.179	<0.001
	Huynh–Feldt	8818.779	6.278	1404.814	319.179	<0.001
Error (noise level)	Sphericity assumed	1409.110	408	3.454		
	Huynh–Feldt	1409.110	320.155	4.401		
Vehicle type	Sphericity assumed	340.312	3	113.437	32.460	<0.001
	Huynh–Feldt	340.312	2.829	120.301	32.460	<0.001
Error (vehicle type)	Sphericity assumed	534.668	153	3.459		
	Huynh–Feldt	534.668	144.270	3.706		
Noise level * vehicle type	Sphericity assumed	171.832	24	7.160	2.195	0.001
	Huynh–Feldt	171.832	12.778	13.447	2.195	0.009
Error (noise level * vehicle type)	Sphericity assumed	3393.168	1224	3.262		
	Huynh–Feldt	3393.168	651.697	6.127		

TABLE III. Linear least-squares fit for source-specific and overall data. Dependent variable is mean annoyance ratings, and independent variable is L_{Aeq} (mean annoyance ratings$=b_0+b_1*L_{Aeq}$). Mean and σ_e are annoyance rating averaged across 52 subjects and standard error of the estimated annoyance ratings in each vehicle or total, respectively. DW is Dubin–Watson's D statistic.

Vehicle type	b_1	b_0	r	Mean	σ_e	DW
Civil aircraft	0.148	−3.599	0.698	5.934	2.101	1.871
Military aircraft	0.153	−4.337	0.735	5.496	1.949	1.704
Railway	0.154	−4.420	0.751	5.513	1.877	1.785
Road traffic	0.166	−5.972	0.765	4.750	1.943	1.656
Overall	0.155	−4.582	0.730	5.423	2.014	1.677

the road-traffic noise model, the slope was the steepest and the intercept was the smallest. In the civil-aircraft noise model, the slope was the gentlest and the intercept was the greatest. The railway and military-aircraft noises were moderate in both. The differences in the annoyance responses to noise sources increased as the corresponding TALs decreased. In view of psychometrics, the TAL of the equally annoying road-traffic noise is larger than those of the civil-aircraft, military-aircraft, and railway noises. Penalties should be added to the TALs of civil-aircraft, military-aircraft, and railway noises in relation to the differences in levels. The penalties for the civil-aircraft, military-aircraft, and railway noises are $14.295-0.108L_{Aeq}$, $9.849-0.078L_{Aeq}$, and $9.349-0.072L_{Aeq}$, respectively. The overall mean annoyance rating was about 5 (more precisely, the expected mean annoyance was 5.423, as shown in Table III). In the model of civil-aircraft, military-aircraft, and railway noises, each of the TALs of 58.1, 61.0, and 61.2 dB caused an expected annoyance of 5. The penalty for the civil-aircraft noise compared with community responses to road-traffic noise was a TAL of 8 dB on the average. The penalty for the military-aircraft and railway noises compared with the subjective responses to road-traffic noise was a TAL of 5 dB on the average.

B. Comparison in terms of other noise rating indices

1. Other A-weighted noise ratings

Various A-weighted noise metrics, such as MXAL, modified TAL, NPL, and TNI, were investigated. The time weighting was set to "fast" (125 ms) for their calculation. In this research, the difference between ASEL and TAL was a constant of 11.76 dB; thus, the two indices were perfectly correlated. Therefore, ASEL was excluded from the analysis. The modified TAL, NPL, and TNI are computed from the following formula. The noise metric for each stimulus is summarized in Table IV.

$$L_{Aeq}{}' = L_{Aeq} + f(\sigma'), \tag{1}$$

where

$$\sigma' = \left[\frac{1}{T}\int_0^T \left(\frac{dL}{dt}\right)^2 dt\right]^{0.5}. \tag{2}$$

Further,

$$f(\sigma') = 10\log_{10}(1+15\sigma') \tag{3}$$

and

$$\mathrm{NPL} = L_{Aeq} + k\sigma. \tag{4}$$

In the above, k is an empirical constant, which is set equal to 2.56, and σ is the standard deviation.

$$\mathrm{TNI} = 4(L_{10}-L_{90}) + (L_{90}-30). \tag{5}$$

Regression analyses regarding the four independent variables, except for ASEL, were performed for each noise source and for all noise sources taken together. In Table V, the correlation between annoyance for the overall noise source and TNI was extremely poor: TNI explained only 16.6% of the variance in annoyance, which means that TNI is not a good noise metric and that further analysis is meaningless. The correlation between the modified TAL and annoyance ratings was significantly different from the correlation between the NPL and annoyance ratings (Fisher's $Z=3.770$, $p<0.001$). The correlation between TAL (or ASEL) and the annoyance ratings was significantly different from the correlation between the NPL and annoyance ratings (Fisher's $Z=3.440$, $p<0.001$). The differences in the goodness of fit between ASEL, TAL, MXAL, and the modified TAL were not statistically significant ($Z=0.133-0.463$, $p>0.05$).

Consequently, TAL, ASEL, MXAL, and the modified TAL might be relatively good acoustic measures for annoyance responses to short-term exposure to community noise. Thus, TNI and NPL were no longer taken into account. Annoyance responses to each noise were compared regarding the two noise metrics.

a. Maximum A-weighted sound level ($L_{A\,\max}$) The penalties for civil-aircraft, military-aircraft, and railway noises were $9.367-0.108L_{A\,\max}$, $4.090-0.078L_{A\,\max}$, and $1.988-0.072L_{A\,\max}$, respectively. In the model of civil-aircraft, military-aircraft, and railway noises, each of the maximum A-weighted sound levels of 66.1, 69.7, and 71.5 dB caused an expected annoyance of 5. The civil-aircraft penalty was a maximum A-weighted sound level of 2 dB on the average. The railway bonus was a maximum A-weighted sound level of 3 dB on the average. No bonus or penalty was found for military aircraft. Even with the same maximum A-weighted level, civil-aircraft noise was the most annoying, followed by road-traffic, military-aircraft, and railway noises, in that order.

TABLE IV. Summary of the levels in each noise metric for every stimuli.

Type of stimulus	L_{Aeq}	$L_{A\,max}$	$L_{Aeq}{}'$	L_{NP}	TNI	L_{Beq}	L_{Ceq}	L_{Deq}	L_{Zeq}
Civil aircraft	40	48	60.7	63.4	89	41.6	43.3	46	43.8
	50	58	70.7	73.4	99	51.6	53.3	56	53.8
	55	63	75.7	78.4	104	56.6	58.3	61	58.8
	60	68	80.7	83.4	109	61.6	63.3	66	63.8
	65	73	85.7	88.4	114	66.6	68.3	71	68.8
	70	78	90.7	93.4	119	71.6	73.3	76	73.8
	75	83	95.7	98.4	124	76.6	78.3	81	78.8
	80	88	100.7	103.4	129	81.6	83.3	86	83.8
	85	93	105.7	108.4	134	86.6	88.3	91	88.8
Military aircraft	40	48.9	63.7	66.7	102.4	42.3	43.4	44.7	43.5
	50	58.9	73.7	76.7	112.4	52.3	53.4	54.7	53.5
	55	63.9	78.7	81.7	117.4	57.3	58.4	59.7	58.5
	60	68.9	83.7	86.7	122.4	62.3	63.4	64.7	63.5
	65	73.9	88.7	91.7	127.4	67.3	68.4	69.7	68.5
	70	78.9	93.7	96.7	132.4	72.3	73.4	74.7	73.5
	75	83.9	98.7	101.7	137.4	77.3	78.4	79.7	78.5
	80	88.9	103.7	106.7	142.4	82.3	83.4	84.7	83.5
	85	93.9	108.7	111.7	147.4	87.3	88.4	89.7	88.5
Railway	40	50.3	62	65	89.5	42.7	44.9	45.1	45.2
	50	60.3	72	75	99.5	52.7	54.9	55.1	55.2
	55	65.3	77	80	104.5	57.7	59.9	60.1	60.2
	60	70.3	82	85	109.5	62.7	64.9	65.1	65.2
	65	75.3	87	90	114.5	67.7	69.9	70.1	70.2
	70	80.3	92	95	119.5	72.7	74.9	75.1	75.2
	75	85.3	97	100	124.5	77.7	79.9	80.1	80.2
	80	90.3	102	105	129.5	82.7	84.9	85.1	85.2
	85	95.3	107	110	134.5	87.7	89.9	90.1	90.2
Road traffic	40	42.2	57.5	43.4	22.2	42.1	43.8	44	44.2
	50	52.2	67.5	53.4	32.2	52.1	53.8	54	54.2
	55	57.2	72.5	58.4	37.2	57.1	58.8	59	59.2
	60	62.2	77.5	63.4	42.2	62.1	63.8	64	64.2
	65	67.2	82.5	68.4	47.2	67.1	68.8	69	69.2
	70	72.2	87.5	73.4	52.2	72.1	73.8	74	74.2
	75	77.2	92.5	78.4	57.2	77.1	78.8	79	79.2
	80	82.2	97.5	83.4	62.2	82.1	83.8	84	84.2
	85	87.2	102.5	88.4	67.2	87.1	88.8	89	89.2

b. Modified time-averaged A*-weighted sound level* ($L_{Aeq}{}'$)*: The modified TAL* The penalties for civil-aircraft, military-aircraft, and railway noises were $13.349-0.108L_{Aeq}{}'$, $5.590-0.078L_{Aeq}{}'$, and $6.434-0.072L_{Aeq}{}'$, respectively. In the model of civil-aircraft, military-aircraft, and railway noises, each of the modified TALs of 78.8, 84.7, and 83.2 dB caused an expected annoyance of 5. The civil-aircraft penalty was a modified TAL of 5 dB on the average. No bonus or penalty was found for either military-aircraft or railway noise (less than 2 dB).

2. Other frequency weighting networks

The time-averaged B-, C-, D-, and Z-weighted sound levels were computed for all stimuli (see Table IV). The least-squares fit was performed between the annoyance ratings and each of the sound levels. The variances of the annoyance explained by L_{Beq}, L_{Ceq}, L_{Deq}, and L_{Zeq} were 0.530 ($r=0.728$), 0.530 ($r=0.728$), 0.542 ($r=0.736$), and 0.530 ($r=0.728$), respectively. No significant difference was found in the goodness of fit across the sound levels ($p>0.05$). They were not significantly different from the product-moment correlations between relatively good acoustic measures (L_{Aeq}, L_{AE}, $L_{A\,max}$, and $L_{Aeq}{}'$) and the annoyance response ($p>0.05$).

The penalties or bonuses of civil-aircraft, military-aircraft, and railway noises are shown in Table VI. For the noise metrics, the civil-aircraft penalties were $15.127-0.108L_{Zeq}$, $14.958-0.108L_{Beq}$, $15.175-0.108L_{Ceq}$, and $12.880-0.108L_{Deq}$, respectively. The military-aircraft penalties were $10.807-0.078L_{Zeq}$, $9.837-0.078L_{Beq}$, $10.494-0.078L_{Ceq}$, and $9.500-0.078L_{Deq}$, respectively. The railway penalties were $8.792-0.072L_{Zeq}$, $8.952-0.072L_{Beq}$, $8.590-0.072L_{Ceq}$, and $8.560-0.072L_{Deq}$, respectively. On average, the civil-aircraft penalties were about 8 dB for the B-, C-, and Z-weighted L_{eq} and about 6 dB for the D-weighted L_{eq}. On average, the military-aircraft penalties were about 5 dB

TABLE V. Linear least-squares fit for source-specific and overall data. Dependent variable is mean annoyance ratings, and independent variable is each A-weighted noise metric (mean annoyance ratings=b_0+b_1*noise metric). r is product-moment correlation, r^2 is the ratio of variance accounted for by regression to total variance, and σ_e is standard error of the estimated annoyance ratings in each vehicle or total.

Vehicle type	Noise metric	b_1	b_0	r	r^2	σ_e
Overall	$L_{A\,\max}$	0.153	−5.568	0.737	0.543	1.990
Civil aircraft		0.148	−4.783	0.698	0.488	2.101
Military aircraft		0.153	−5.659	0.735	0.541	1.949
Railway		0.154	−6.008	0.751	0.564	1.877
Road traffic		0.166	−6.338	0.765	0.585	1.943
Overall	$L_{Aeq}{}'$	0.154	−7.761	0.735	0.540	1.996
Civil aircraft		0.148	−6.664	0.698	0.488	2.101
Military aircraft		0.153	−7.952	0.735	0.541	1.949
Railway		0.154	−7.812	0.751	0.564	1.877
Road traffic		0.166	−8.880	0.765	0.585	1.943
Overall	L_{NP}	0.118	−4.515	0.673	0.453	2.179
Overall	TNI	0.035	1.915	0.408	0.166	2.689

for the B- and C-weighted L_{eq}, about 4 dB for the D-weighted L_{eq}, and about 6 dB for the Z-weighted L_{eq}. On average, the railway penalty was about 4 dB for the B-, C-, D-, and Z-weighted L_{eq}. These results are similar to the results when the A-weighted L_{eq} was controlled. The civil-aircraft noise was the most annoying, the military-aircraft and railway noises were moderate, and the road-traffic noise was the least annoying.

C. Effects of factors on annoyance

1. Acoustical factors

The influence of several acoustical factors on annoyance was investigated using principal component factor and multiple regression analyses. A total of seven factors, such as TAL (L_{Aeq}), MXAL ($L_{A\,\max}$), $L_{A\,\max-Aeq}$, σ, σ', L_{C-A}, and L_{D-B}, were selected. The differences between the C- and A-weighted equivalent continuous levels (L_{C-A}) represented the low-frequency content in the respective transportation noises. The differences between the D-weighted and B-weighted equivalent continuous levels (L_{D-B}) represented the high-frequency content in the respective transportation noises.

A principal component factor analysis was performed for the seven factors in order to classify them into uncorrelated components. Three components accounted for 90% of

TABLE VI. Linear least-squares fit for source-specific and overall data. Dependent variable is mean annoyance ratings, and independent variable is each time average noise metric with regard to frequency weightings (mean annoyance ratings=b_0+b_1*noise metric). r is product-moment correlation, r^2 is the ratio of variance accounted for by regression to total variance, and σ_e is standard error of the estimated annoyance ratings in each vehicle or total.

Vehicle type	Noise metric	b_1	b_0	r	r^2	σ_e
Overall	L_{Beq}	0.155	−4.897	0.728	0.530	2.018
Civil aircraft		0.148	−3.839	0.698	0.488	2.101
Military aircraft		0.153	−4.689	0.735	0.541	1.949
Railway		0.154	−4.836	0.751	0.564	1.877
Road traffic		0.166	−6.322	0.765	0.585	1.943
Overall	L_{Ceq}	0.155	−5.138	0.728	0.530	2.020
Civil aircraft		0.148	−4.081	0.698	0.488	2.101
Military aircraft		0.153	−4.858	0.735	0.541	1.949
Railway		0.154	−5.174	0.751	0.564	1.877
Road traffic		0.166	−6.600	0.765	0.585	1.943
Overall	L_{Deq}	0.156	−5.431	0.736	0.542	1.994
Civil aircraft		0.148	−4.494	0.698	0.488	2.101
Military aircraft		0.153	−5.055	0.735	0.541	1.949
Railway		0.154	−5.211	0.751	0.564	1.877
Road traffic		0.166	−6.632	0.765	0.585	1.943
Overall	L_{Zeq}	0.155	−5.190	0.728	0.530	2.020
Civil aircraft		0.148	−4.156	0.698	0.488	2.101
Military aircraft		0.153	−4.873	0.735	0.541	1.949
Railway		0.154	−5.218	0.751	0.564	1.877
Road traffic		0.166	−6.667	0.765	0.585	1.943

TABLE VII. Rotated component matrix of principal component factor analysis. Three latent factors were extracted by Varimax rotation. They accounted for 90.0% (first factor: 40.7%, second factor: 28.3%, and third factor: 21.0%) of the trace of the matrix analyzed.

Acoustical factor	Components		
	1	2	3
L_{Aeq}	−0.053	0.998	0.002
$L_{A\,max}$	0.162	0.987	−0.004
$L_{A\,max-Aeq}$	0.983	0.057	−0.027
σ	0.984	0.052	0.172
σ'	0.871	0.032	−0.050
L_{C-A}	0.203	0.026	−0.857
L_{D-B}	0.293	0.027	0.840

the total variance. The first component represented the level variability, the second component represented the energy exposed, and the third component represented spectra (see Table VII). These findings were similar to Cermak and Cornillon's results.[23] L_{C-A} was negatively correlated with L_{D-B}. To avoid multicollinearity, multiple regression analyses were performed with independent variables, such as one factor of the exposed energy (L_{Aeq}), one factor of the level-variability property, and one factor of the spectral property (L_{C-A} or L_{D-B}). In Table VIII, it is seen that the rate of the explained variance increased by as much as 1.5% due to the level-variability property, 1.4% by the spectral property, and 2.0% by both. The regression coefficient of L_{C-A} was not significant ($p>0.05$).

The L_{D-B} of civil-aircraft, military-aircraft, and railway noises were 2.56, 0.53, and 0.56 dB, respectively, larger than those of road-traffic noise. σ of civil-aircraft, military-aircraft, and railway noises were 7.79, 9.11, and 8.44 dB, respectively, larger than those of road-traffic noise. From the third row in Table VIII, it can be seen that aircraft noise annoyance was 0.637 (2.56×0.249) larger than road-traffic noise annoyance in terms of ΔL_{D-B} and 0.545 (7.79×0.07) larger than road-traffic annoyance in terms of $\Delta\sigma$. The total increase in the civil-aircraft noise annoyance was equivalent to a TAL of 7.6 dBA ($\fallingdotseq$8 dBA). The constitution and the proportion of penalty that is accounted for by each acoustical factor with regard to the noise source are summarized in Table IX. For other noise metrics, the cause of differential responses to noise sources could be analyzed similarly.

2. Individual noise sensitivity

Individual noise sensitivities for all subjects were investigated using Weinstein's method. Their median was 63, the lower quartile was 31.5, and the upper quartile was 84. Least-squares fits were performed for L_{Aeq} and noise sensitivity and undertaken for L_{Aeq}, L_{D-B}, σ, and noise sensitivity. The assumptions, such as linearity, independence, multicollinearity, etc., were not violated in both cases. The multiple correlations were 0.761 ($r^2=0.579$) and 0.775 ($r^2=0.600$). In both cases, the increase in multiple correlations was statistically significant ($Z=2.242$, $p<0.05$; $Z=2.136$, $p<0.05$). An increase of 4.6% was obtained in the explained variance. From the present and prior results, it was confirmed that the effect of individual noise sensitivity on annoyance was more than those of the two acoustical factors or their combination. However, individual noise sensitivity did not account for the difference in annoyance with regard to noise sources because a within-subject design was applied in this experiment.

TABLE VIII. Linear relationships between annoyance and indices for acoustical properties. r is multiple correlation, r^2 is the ratio of variance accounted for by regression to total variance, and σ_e is standard error of the estimate.

Linear relationships	r	r^2	σ_e
Annoyance$=-5.601+0.155L_{Aeq}+0.366L_{D-B}$	0.740	0.547	1.982
Annoyance$=-5.343+0.155L_{Aeq}+0.099\sigma$	0.740	0.548	1.981
Annoyance$=-5.811+0.155L_{Aeq}+0.070\sigma+0.249L_{D-B}$	0.744	0.553	1.970

TABLE IX. Constitution and proportion of penalty by each acoustical factor with regard to noise source. Second column shows penalty caused by the difference between high-frequency components of each noise source and road traffic, and third column shows penalty caused by the difference between level variability of each noise source and road traffic.

Noise source	ΔL_{D-B} [dBA (%)]	$\Delta\sigma$ [dBA (%)]	Total penalty [dBA (%)]
Civil aircraft	4.1(53.9)	3.5(46.1)	7.6(100)
Military aircraft	0.9(18.0)	4.1(82.0)	5.0(100)
Railway	0.9(19.1)	3.8(80.9)	4.7(100)

V. DISCUSSION

Many investigations have been performed on comparing the annoyances from various types of vehicle. All of them affirm that vehicle noises cause differential levels of annoyance. However, there have been inconclusive and conflicting findings regarding the impact of railway noise.

Through a cluster of six different field surveys, Fields and Walker[1] compared railway noise annoyance with aircraft noise and road-traffic noise annoyance. In order to identify a systematic difference, the results of the six surveys were analyzed using identical noise and annoyance ratings. Railway noise annoyance was less than road-traffic and aircraft noise annoyances and increased less rapidly with the noise level. A railway bonus of about 10 dB was found. Miedema and Vos[3] synthesized 55 data sets that were acquired from different surveys. In spite of large variations in each data set for the same type of vehicle, the proposed dose-response relationships showed that railway noise caused the least annoyance and aircraft noise caused the greatest annoyance. The gap between road-traffic and railway noises was about 5 dB with respect to the road-traffic DNL of 60 dB. Finegold *et al.*[5] performed meta analysis for data based on Field *et al.* and developed source-specific dose-response relationships using a logistic model. The result of Finegold *et al.* was somewhat different from the two results referred above. At the higher level (DNL>75 dB), railway noise caused more annoyance than road-traffic noise. At the lower level, the annoyances caused by both were not significantly different.

In a laboratory study by Öhström *et al.*,[4] annoyances from two types of road-traffic noise (lorry and moped noises), railway noise, and aircraft noise were tested. Under

the condition of controlled TAL (L_{Aeq}), lorry noise was the least annoying, followed by railway, aircraft, and moped noises, in that order. Under the condition of controlled MXAL ($L_{A\,max}$), lorry noise was the least annoying, followed by moped, railway, and aircraft noises, in that order. The differences were statistically significant. The two road-traffic vehicles caused different levels of annoyance. The reason might be the differing time patterns and spectral characteristics that are caused by engine types and operating conditions. Versfeld and Vos[24] demonstrated that driving conditions and the spectral difference significantly influenced the annoyance. From this point of view, it is necessary that the referred field surveys should be reanalyzed in more detail.

In this research, noise annoyances from four kinds of vehicle with specific types of engine were compared. Different results were revealed with regard to the noise metrics. When transportation noise was evaluated as the maximum A-weighted level, civil-aircraft noise caused the most annoyance, followed by road-traffic noise, military-aircraft noise, and railway noise, in that order. The railway bonus was about 3 dBA with regard to the maximum A-weighted level. However, the railway bonus was not found for other noise metrics. Despite the same time-averaged (or exposure) level, it was found that aircraft noise was the most annoying and road-traffic noise was the least annoying. The civil-aircraft penalty was a TAL of about 8 dBA that was caused by the high-frequency component (53.9%) and level variability (46.1%). The military-aircraft penalty was a TAL of about 5 dBA that was caused by the high-frequency component (18.0%) and level variability (82.0%). The railway penalty was a TAL of about 5 dBA that was caused by the high-frequency component (19.1%) and level variability (80.9%) (see Table IX). Not only the exposure (or average) energy level but also the peak energy level might be important in the situation of short-term exposure of noise because people hear the peak level instantaneously. Two standard noise regulations with regard to the maximum noise level and the average noise level will be necessary. However, transportation noise has been evaluated by only the TAL in most countries, including Korea. Therefore, the railway bonus will not be acceptable for railway vehicles with diesel-electric engines under an evaluation system that is based on the average noise level.

Recent field studies[7–9] justified these findings in Korea. In their researches, the difference in annoyance from each noise source (civil aircraft, railway, and road traffic) was not different from that in this research. Of course, it is difficult to compare the present research with the field researches directly. As known, there are various differences between the both in noise exposure and annoyance rating conditions, process of annoyance perceived, interference of other noise, and so on. In other words, outdoor instantaneous annoyance is rated, and subjects have not suffered from specific noise, and other background noises are not considered in this experiment, whereas indoor cumulative annoyance is rated by mental integrations, and respondents have suffered from specific noise, and interference of other background noises has an influence on annoyance ratings in the field studies. Therefore, relative difference in annoyance responses according to vehicle types between a laboratory study and field studies was compared. In the field studies, the engine type of each noise source was similar to (or the same as) that in the present laboratory study. Civil-aircraft noise caused more annoyance than railway noise,[7,8] and road-traffic noise was the least annoying (exposure-response relationship for road-traffic noise was not reported). Self-reported sleep disturbance from railway noise was higher than that from road-traffic noise.[9]

VI. CONCLUDING REMARKS

The main purpose of the present experimental study was to compare the annoyance responses to civil-aircraft noise, military-aircraft noise, railway noise, and road-traffic noise. It should be noted that the annoyances for specific types and operating conditions of four vehicles were measured in simulated outdoor conditions. The following conclusions were drawn from the results.

(1) Civil-aircraft noise caused the most annoyance and road-traffic noise caused the least annoyance for the same TAL (L_{Aeq}). Almost equal annoyance was rated at the same TAL (L_{Aeq}) for the military-aircraft and railway noises.
(2) To obtain equally annoying levels under road-traffic noise, a penalty of 8 dB [TAL, viz., (L_{Aeq})] should be applied to civil-aircraft noise; likewise, a penalty of 5 dB (L_{Aeq}) should be applied to the military-aircraft and railway noises.
(3) The differential responses to the noises can be explained by the high-frequency component and level variability. The greater these factors are, the greater the annoyance that can be expected.
(4) A railway bonus would not be acceptable for railway vehicles with diesel-electric engines under the noise-evaluation system that is based on metrics of average energy.

ACKNOWLEDGMENTS

This work was supported by Engineering Research Institute at Seoul National University and the "Core Environmental Technology Development Project for Next Generation" in Korea Institute of Environmental Science and Technology.

[1]J. M. Fields and J. G. Walker, "Comparing the relationships between noise level and annoyance in different surveys: A railway noise vs. aircraft and road traffic comparison," J. Sound Vib. **81**, 51–80 (1982).
[2]K. D. Kryter, "Community annoyance from aircraft and ground vehicle noise," J. Acoust. Soc. Am. **72**, 1222–1242 (1982).
[3]H. M. E. Miedema and H. Vos, "Exposure-response relationships for transportation noise," J. Acoust. Soc. Am. **104**, 3432–3445 (1998).
[4]E. Öhström, M. Björkman, and R. Rylander, "Laboratory annoyance and different traffic noise sources," J. Sound Vib. **70**, 334–341 (1980).
[5]L. S. Finegold, C. S. Harris, and H. E. von Gierke, "Community annoyance and sleep disturbance: Updated criteria for assessing the impacts of general transportation noise on people," Noise Control Eng. J. **42**, 25–30 (1994).
[6]B. Griefahn, A. Marks, and S. Robens, "Noise emitted from road, rail and air traffic and their effects on sleep," J. Sound Vib. **295**, 129–140 (2006).
[7]C. Lim, J. Kim, J. Hong, and S. Lee, "The relationship between railway noise and community annoyance in Korea," J. Acoust. Soc. Am. **120**,

2037–2042 (2006).
[8]C. Lim, J. Kim, J. Hong, S. Lee, and S. Lee, "The relationship between civil aircraft noise and community annoyance in Korea," J. Sound Vib. **299**, 575–586 (2007).
[9]J. Hong, J. Kim, C. Lim, K. Kim, and S. Lee, "Comparative study on sleep disturbance from railway and road traffic noise," in Proceedings of the Inter-Noise 2007, Istanbul, Turkey (2007), pp. 181–181.
[10]J. Kaku and I. Yamada, "The possibility of a bonus for evaluating railway noise in Japan," J. Sound Vib. **193**, 445–450 (1996).
[11]M. Hui and T. Yano, "An experiment on auditory and non-auditory disturbances caused by railway and road traffic noises in outdoor conditions," J. Sound Vib. **277**, 501–509 (2004).
[12]International Organization for Standardization ISO 1996-1:2003, Acoustics—Description, measurement and assessment of environmental noise—Part 1: Basic quantities and assessment procedures (2003).
[13]J. M. Fields, "Effects of personal and situational variables on noise annoyance in residential areas," J. Acoust. Soc. Am. **93**, 2753–2763 (1993).
[14]H. M. E. Miedema and H. Vos, "Demographic and attitudinal factors that modify annoyance from transportation noise," J. Acoust. Soc. Am. **105**, 3336–3344 (1999).
[15]J. Kim, C. Lim, J. Hong, and S. Lee, "The influence of binaural effects on annoyance for transportation noise," Noise Control Eng. J. **55**, 204–216 (2007).
[16]T. Poulsen, "Influence of session length on judged annoyance," J. Sound Vib. **145**, 217–224 (1991).
[17]International Organization for Standardization ISO 389-1, Acoustics—Reference zero for the calibration of audiometric equipment—Part 1: Reference equivalent threshold sound pressure level for pure tones and supra-aural earphone (International Organization For Standardization, Geneva, 1998).
[18]International Organization for Standardization ISO 8253-1, Acoustics—Audiometric test methods—Part 1: Basic pure tone air and bone conduction threshold audiometry (International Organization For Standardization, 1989).
[19]N. D. Weinstein, "Individual differences in reactions to noise: A longitudinal study in a college dormitory," J. Appl. Psychol. **63**, 458–466 (1978).
[20]J. M. Fields, R. G. De Jong, T. Gjestland, I. H. Flindell, R. F. S. Job, S. Kurra, P. Lercher, M. Vallet, T. Yano, R. Guski, U. Felscher-suhr, and R. Schumer, "Standardized general-purpose noise reaction questions for community noise surveys: Research and a recommendation," J. Sound Vib. **242**, 641–679 (2001).
[21]International Standard Organization ISO/TS 15666, Acoustics—Assessment of noise annoyance by means of social and socio-acoustic surveys (2003).
[22]K. Zimmer and W. Ellermeier, "Psychometric properties of four measures of noise sensitivity: A comparison," J. Environ. Psychol. **19**, 295–302 (1999).
[23]G. W. Cermak and P. C. Cornillon, "Multidimensional analyses of judgments about traffic noise," J. Acoust. Soc. Am. **59**, 1412–1420 (1976).
[24]N. J. Versfeld and J. Vos, "Annoyance caused by sounds of wheeled and tracked vehicle," J. Acoust. Soc. Am. **101**, 2677–2685 (1997).

Speech levels in meeting rooms and the probability of speech privacy problems

J. S. Bradley[a] and B. N. Gover
National Research Council, 1200 Montreal Road, Ottawa, K1A 0R6, Canada

(Received 19 August 2009; revised 24 November 2009; accepted 3 December 2009)

Speech levels were measured in a large number of meetings and meeting rooms to better understand their influence on the speech privacy of closed meeting rooms. The effects of room size and number of occupants on average speech levels, for meetings with and without sound amplification, were investigated. The characteristics of the statistical variations of speech levels were determined in terms of speech levels measured over 10 s intervals at locations inside, but near the periphery of the meeting rooms. A procedure for predicting the probability of speech being audible or intelligible at points outside meeting rooms is proposed. It is based on the statistics of meeting room speech levels, in combination with the sound insulation characteristics of the room and the ambient noise levels at locations outside the room. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277220]

PACS number(s): 43.55.Hy [NX] Pages: 815–822

I. INTRODUCTION

It is sometimes required that speech from a meeting room be unintelligible or even inaudible to an eavesdropper outside the room. When this is achieved we can describe the room as having a certain degree of speech privacy, and when the degree of privacy is adequate, we can say that the room is speech secure. When speech from the room is intelligible or audible outside the room, we can say there has been a speech privacy lapse. Whether the speech from the meeting room is intelligible or audible to a particular eavesdropper outside the room will depend on the level of the speech in the room, the sound insulation characteristics of the room boundary, and the level of ambient noise at the location of the eavesdropper outside the room. Of course, speech levels in meeting rooms are not constant but vary from moment to moment. The likelihood of a speech privacy lapse will depend on the probability of louder speech levels occurring.

The average speech levels of children, male, and female adult talkers have been determined for varied vocal effort by Pearsons *et al.*[1] However, the statistical distribution of speech levels over time in meeting rooms has not been previously investigated. In this new work, the statistical properties of speech levels incident on the boundaries of meeting rooms were determined by measurements in a large number of meetings and meeting rooms.

This paper first investigates the relationships of average meeting room speech levels with meeting room parameters such as room size and number of occupants. Following this, the statistical distributions of speech levels in meeting rooms are presented. The new results form the basis for a new procedure for estimating the likelihood of speech privacy lapses for meeting rooms from the statistical properties of speech levels in meeting rooms. This new procedure is presented in Secs. IV–VI.

[a]Author to whom correspondence should be addressed. Electronic mail: john.bradley@nrc-cnrc.gc.ca

II. PROCEDURE

Sound levels in meeting rooms were measured over 24-h periods using data logging sound level meters located near the periphery of the room. The integrating sound level meters (Brüel and Kjær type 2236) were fitted with external batteries so that they could be left operating unattended for 24-h periods. The sound level meters stored an energy average *A*-weighted sound level (L_{eq}) for every 10 s interval over the 24-h period. Four meters were placed in each room and were typically located approximately 1 m from the room boundaries with the microphone 1.2 m above the floor. These locations were chosen so that the recorded speech levels were representative of the sound incident on the walls of the meeting room.

Figure 1(a) illustrates an example of measured 10 s L_{eq} values over a 24-h period. The levels associated with each meeting were identified from records like this one and the known schedule of meetings in the room. Figure 1(b) illustrates the 10 s L_{eq} values versus time for one meeting from the 24-h record in Fig. 1(a). During the meeting the sequence of 10 s L_{eq} values is seen to vary somewhat randomly. The stored 10 s L_{eq} values for the 4 meters in each meeting were first used to calculate an average L_{eq} for each meeting. The L_{eq} values from the intervals between meetings were used to analyze ambient noise levels in the meeting rooms. These could be compared with L_{90} values (the *A*-weighted noise level exceeded 90% of the time) that were also measured for each 10 s interval.

Data were obtained for a total of 79 meetings in 32 different meeting rooms. Twenty-nine of the meetings included the use of sound amplification systems. The sound amplification systems were quite simple and most had one or two loudspeakers located at the front of the room. Two of the 32 rooms were repeats to include measurements of meetings with and without sound amplification in the same rooms. The range of numbers of people present, room volumes, and floor areas are listed in Table I.

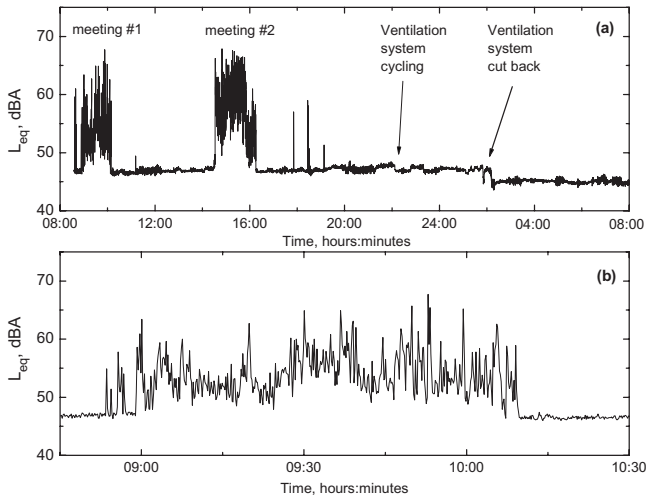


FIG. 1. Recorded time history of 10 s L_{eq} values from one data logger, (a) complete 24-h period, and (b) enlarged portion for one meeting.

III. AVERAGE SPEECH LEVELS IN MEETING ROOMS

A. Meeting-average results

Meeting room-average L_{eq} values were calculated for all 10 s L_{eq} values measured in each of the 79 meetings. These meeting-average L_{eq} values are summarized in Table II. As well as the average L_{eq} values, the standard deviations of the 10 s L_{eq} values are given and the number of meetings included in each of the categories of rooms listed. There are small differences of typically 1 or 2 dB between the averages for some categories. These will be discussed in Secs. III B–III D.

B. Effects of room size

The results in Table II indicate that sound levels were on average 2.8 dBA higher in smaller rooms (less than 100 m^3) than in larger rooms (greater than 100 m^3) for rooms without amplified speech sound. There were no small rooms with sound amplification.

The effect of room size on speech L_{eq} values was investigated by plotting measured room-average L_{eq} values versus various parameters describing the rooms and numbers of people present. The variables considered are listed in Table III. There were very few significant systematic effects of these variables on the room-average speech L_{eq} values. However, when the room-average L_{eq} values for non-amplified rooms were plotted versus meeting room volume, as shown in Fig. 2, there was a statistically significant ($p<0.001$) decrease in L_{eq} values with increasing room volume. A similar but less significant effect ($p<0.01$) was found when the data were plotted versus meeting room floor area. Significant variations of L_{eq} values with room size were not found for rooms using sound amplification systems. There were also no significant effects of the number or the density of people in the rooms.

Although the regression line in Fig. 2 indicates changes of L_{eq} values by as much as 5 dB with room volume are possible, effects this large are probably not likely for conditions in most meeting rooms. This is partly because the larger rooms included in Fig. 2, with volumes of approximately 500 m^3, would usually have sound amplification systems which were found to eliminate the changes in speech levels with room volume.

Speech in smaller rooms (i.e., less than 100 m^3) not only had a little higher sound levels but also the sound levels seemed a little more scattered between rooms of similar room volume. This is probably due to the unavoidable problem of some talkers being quite close to the microphones in the smaller rooms, which might influence the slope of the regression line in Fig. 2. This also suggests that for speech privacy concerns, there may also be higher speech sound levels incident on the room boundaries due to the presence of talkers close to the room boundaries in smaller meeting rooms and hence an increased risk of speech privacy problems at locations outside the room.

C. Effects of amplification

The results in Table II indicate that on average, speech levels were 2 dB higher in rooms with sound amplification

TABLE I. Summary of meeting rooms measured.

No. of meeting room cases measured[a]	32
No. of meetings measured	79
No. of people in each meeting	2–300 people
Range of room volumes	39–16 000 m^3
Range of room floor areas	15–570 m^2

[a]Includes 30 different rooms, two of which were measured with and without sound amplification systems.

TABLE II. Meeting-average sound levels (L_{eq}, dBA) for various amplified and non-amplified conditions with the standard deviation of each group of N meetings.[a]

Condition	Meeting-average speech level, L_{eq} (dBA)	Standard deviation (dBA)	No. of meetings (N)
All meetings	60.7	4.1	79
All amplified meetings	62.0	4.5	29
All Non-amplified meetings	60.0	3.4	50
Non-amplified small rooms	61.1	2.6	26
Non-amplified large rooms	58.3	3.4	24
Amplified: single loudspeaker system	62.1	4.4	14
Amplified: multiple loudspeaker system	61.8	4.8	15

[a]Standard deviation of the 10 s L_{eq} values.

TABLE III. Variables considered for possible systematic effects on the measured speech levels.

Quantity	Units
Meeting room volume	m^3
Meeting room floor area	m^2
Volume density of people	No. of people/m^3
Area density of people	No. of people/m^2
No. of people	Number

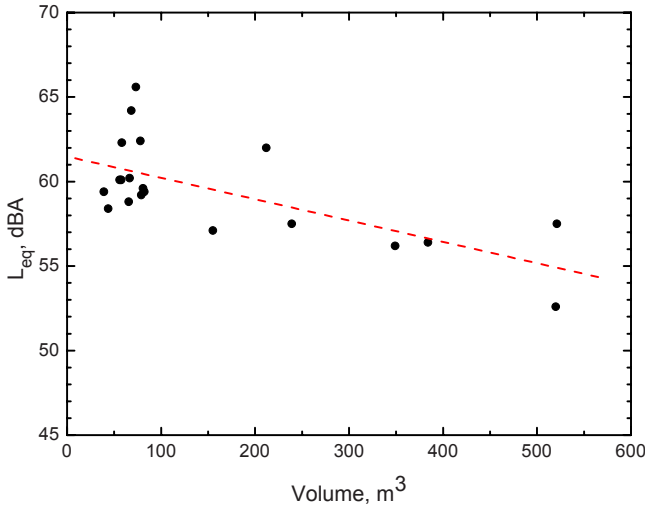


FIG. 2. (Color online) Plot of meeting-average speech L_{eq} values versus meeting room volume and best-fit regression line for meetings without sound amplification systems. (R^2=0.453, $p<0.001$).

systems operating. At first this may seem a rather small difference because amplification systems would be expected to have a larger effect. In two of the rooms, meetings were measured in the same room with and without sound amplification. The room-average speech L_{eq} values for these two rooms, shown in Table IV, indicate that the amplification systems increased speech levels by an average of 10.8 dB.

The smaller difference between the averages for all amplified and all non-amplified conditions in Table II is probably due to the locations of the measurement microphones. They were located around the periphery of the rooms to obtain speech levels representative of the sound incident on the room boundaries. The results in Table II seem to indicate that the amplification systems were in general set up to obtain speech levels, at more distant listening positions, similar to the speech levels found in smaller and non-amplified rooms. That is, sound amplification systems do not normally create much higher sound levels than are found in good non-amplified conditions. (This may not be true for much larger rooms such as theatres and auditoria).

D. Effects of background noise

Estimates of the background noise levels in the meeting rooms were obtained using two different approaches. One approach was to use measured L_{eq} values, obtained in each meeting room when they were unoccupied. The second approach used the L_{90} values obtained during the meetings in each room. Both L_{eq} and L_{90} values were energy averages of values for all 10 s intervals during meetings. The average results for all meetings in Table V show that the two approaches led to nearly identical results. This confirmed that the L_{90} values measured during the meetings were a good indication of the actual ambient noise levels in the rooms.

TABLE IV. Meeting-average speech levels measured with and without sound amplification systems in two rooms.

Room	Non-amplified meeting (L_{eq}) (dBA)	Amplified meeting (L_{eq}) (dBA)	Difference (L_{eq}) (dBA)
Annex (521 m^3)	52.8	64.9	12.1
Centennial (520 m^3)	52.6	61.4	8.8
Average			10.8

TABLE V. Comparison of average ambient noise measurements from daytime L_{90} values (during meetings) and L_{eq} values (between meetings, i.e., no meeting sounds).

Measure	Mean value	Standard deviation	No. of measurements
L_{90} during meetings	49.4	3.3	149
L_{eq} of ambient noise between meetings	49.3	4.5	57

The relationship between meeting-average speech levels and ambient noise levels is shown in Fig. 3. This figure shows systematically increasing speech levels with increasing ambient noise levels. For unamplified speech, this is a well-known effect (the Lombard effect[2]) whereby people naturally talk louder to maintain an acceptable signal-to-noise (S/N) ratio. The results in Fig. 3 indicate that this type of effect occurs for all meetings including both amplified and unamplified speech. The diagonal lines on Fig. 3 indicate conditions with signal-to-noise ratios of +10 and +15 dB. A +10 dB signal-to-noise ratio corresponds to reasonably good conditions for speech and a +15 dB signal-to-noise ratio to very good conditions for adults with unimpaired hearing.[3]

From the results in Fig. 3, it is evident that it is important to control ambient noise levels in meeting rooms to improve speech privacy. This is true because when ambient noise levels are allowed to increase, speech levels will also increase correspondingly, and the higher speech levels are more likely to be audible or intelligible outside the meeting room. Of course, for good speech intelligibility inside the meeting room, it is also important to have low background noise levels (ideally 35 dBA in a classroom sized rooms[4,5]).

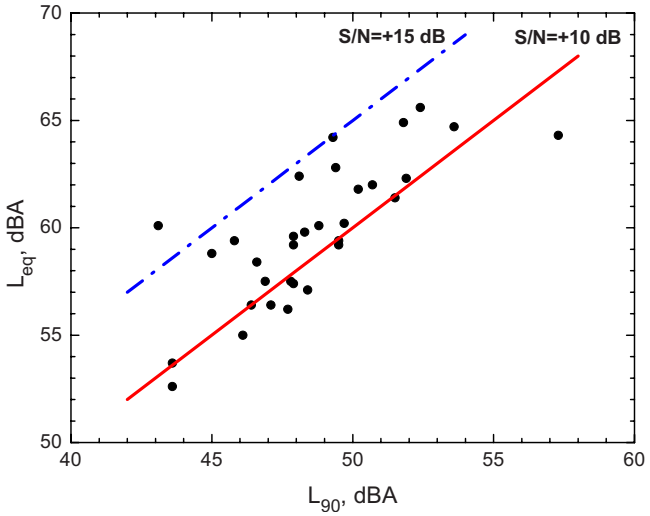


FIG. 3. (Color online) Meeting-average speech levels (L_{eq}) versus ambient noise levels in the meeting rooms (L_{90}). The solid line shows situations with a +10 dB (S/N) ratio and the dash-dotted line shows the more ideal case of a +15 dB speech-to-noise ratio.

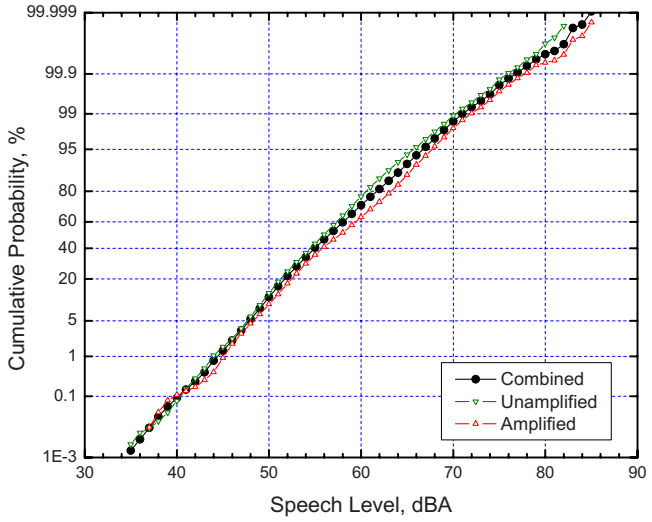


FIG. 4. (Color online) Cumulative probability distributions of 10 s speech L_{eq} values for unamplified speech cases, amplified speech cases, and the combined data.

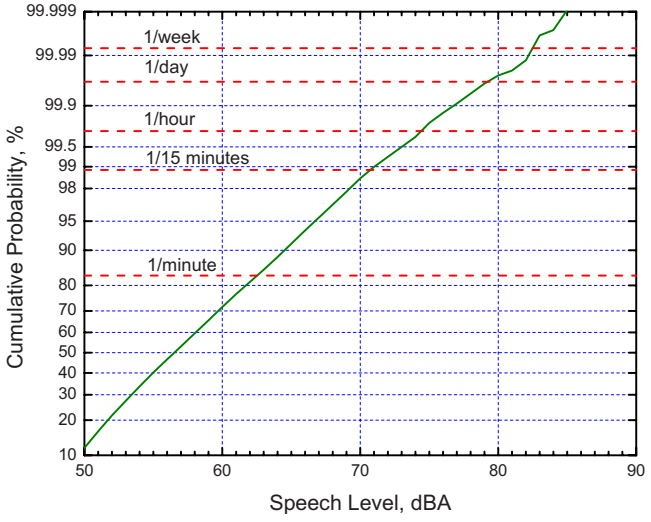


FIG. 5. (Color online) Cumulative probability distributions of 10 s speech L_{eq} values for the combined data with expanded scales to better show the data for higher speech levels. The labels on the horizontal dashed lines (1/min to 1/week) indicate the equivalent intervals between occurrences of these 10 s speech L_{eq} values.

That is, reduced ambient noise levels in meeting rooms would improve speech intelligibility in the meeting room and increase speech privacy at locations outside the meeting room.

IV. STATISTICAL DISTRIBUTION OF SPEECH LEVELS

Higher speech levels in meeting rooms are more likely to lead to speech privacy problems at positions outside the room. It is therefore important to determine the probability of various higher speech levels occurring in typical meeting rooms. This was done by examining the statistical distribution of the 10 s speech L_{eq} values for all meetings. A total of 110 773 speech L_{eq} values were included. Because the average values in Table II indicate a small difference between the amplified and unamplified speech cases, the distributions of 10 s speech L_{eq} values were first considered separately for the two cases as well as for the combined data.

The distributions of 10 s L_{eq} values are shown in Fig. 4 plotted as cumulative probability distributions. There are small differences as expected between the amplified and unamplified cases and the differences vary a little with speech level. However, it was decided that because the values of the combined data distribution were usually within 1 dB of the separate amplified and unamplified speech distributions, the combined data could be used to closely approximate conditions in all meeting rooms. The combined distribution was then used to determine the likelihood of speech privacy lapses for all meeting rooms. Figure 5 plots the cumulative probability distribution of the combined data with expanded scales to make it possible to read off probabilities of the higher speech levels occurring.

From the probabilities of the occurrence of speech levels in Figs. 4 and 5, one can calculate the corresponding average time interval between occurrences of particular higher sound levels taking into account the 10 s duration of each L_{eq} measurement of speech levels. Each probability indicates the frequency of occurrence of all speech levels up to and including the corresponding speech level on the x-axis. For example, a 90% probability corresponds to a speech level of 64.5 dBA, indicating that 90% of the time 10 s speech L_{eq} values would be no higher than 64.5 dBA. Hence, 10% of the time this speech level would be exceeded. There are 360 intervals of 10 s duration in 1 h and this would correspond to speech levels exceeding 64.5 dB in 36 of them. On average there would be a 1.67 min interval between times when this speech level is exceeded. Table VI shows the results of the calculations of the average interval between occurrences of a range of speech levels from the frequency of occurrence of each speech level. Figure 5 also includes horizontal dashed lines to indicate the speech levels corresponding to several intervals (1/min to 1/week).

V. PROBABILITY OF SPEECH PRIVACY PROBLEMS

The audibility and intelligibility of transmitted speech sounds are related to the signal-to-noise ratio of the transmitted speech and the ambient noise at the listener's position outside the room. Previous research has identified the uniformly weighted frequency-averaged signal-to-noise ratio to be an accurate predictor of the degree to which speech is intelligible or audible.[6] It is calculated from the spectra of the transmitted speech and of the background noise signals at the listener's position as given by the following:

$$\mathrm{SNR}_{\mathrm{UNI32}} = \frac{1}{16}\sum_{f=160}^{5000}[L_{\mathrm{ts}}(f) - L_n(f)]_{-32},\ \mathrm{dB}, \qquad (1)$$

where in each of the 16 $\frac{1}{3}$-octave bands centered at frequency f from 160 to 5000 Hz, $L_{\mathrm{ts}}(f)$ is the level of the transmitted speech at the listener's position outside the room, $L_n(f)$ is the level of the background noise at the same position, and the subscript "−32" indicates that the quantity in square brackets (i.e., the signal-to-noise ratio in each band) is to be limited to a minimum of −32 dB.

A condition where 50% of attentive listeners could just understand some speech was defined as the threshold of

TABLE VI. Example results for the calculation of the average time intervals between occurrences of particular speech levels in the meeting room speech data. "Prob" is the probability of occurrence in percent of a speech level equal to or less than the corresponding speech level in column 1. The next two columns indicate the number of times per hour and per 8 h work day that the speech level would be exceeded (i.e., the number of 10 s L_{eq} intervals). The last two columns give the average time interval between 10 s L_{eq} values exceeding these levels.

Speech (dBA)	Prob, % ⩽	N/hour exceed	N/day exceed	Minutes between	Hours between
55	40.265	215.05	1720.4	0.28	0.0047
60	71.622	102.16	817.3	0.59	0.0098
65	91.388	31.00	248.0	1.94	0.0323
70	98.540	5.25	42.0	11.42	0.1903
75	99.798	0.73	5.8	82.43	1.3738
80	99.974	0.09	0.8	636.67	10.6112

intelligibility.[6] This corresponds to a particular value of SNR_{UNI32}. In laboratory experiments this value was −16 dB, but in moderately reverberant conditions with ideally diffuse noise was −11 dB, varying about ±1 dB with variations in reverberation time above 0.5 s.[7] At a lower signal-to-noise condition, there is a point where, for 50% of attentive listeners, speech sounds were just audible. This was defined as the threshold of audibility and corresponds to a higher degree of privacy. The threshold of audibility corresponds to a SNR_{UNI32} value of −22 dB (in both laboratory and real room conditions).[6,7]

To apply SNR_{UNI32} for the assessment of closed room speech privacy, it is necessary to determine the transmitted speech levels and the background noise levels at the listener position. For a given speech level inside the room, the level of transmitted speech depends on the sound insulation provided by the building structure. The relevant measure of sound insulation is the difference in sound level between the average level of a uniform test sound field inside the room and the received level at a spot listener location outside the room.[8] If the average level of the uniform field in the room is $L_s(f)$ and the corresponding received level outside the room is $L_r(f)$, then the level difference in each frequency band is $LD(f)=L_s(f)-L_r(f)$. A uniform field is measured inside the room to represent the average of talkers, who could be located anywhere in the room. The spot receiver locations outside the room are usually chosen to be 0.25 m from the boundaries of the room to minimize the effect of the receiving space, to more realistically represent the locations of potential eavesdroppers, and to allow evaluation of variations of sound insulation, such as due to doors or ducts.

Since the level of speech varies from moment to moment, the speech levels in the room can be assessed statistically. Sections II–IV reported new measurements of speech levels $L_{sp}(f)$ in a large number of meetings and the probability of occurrence of particular levels. The level of the background noise $L_n(f)$ can be measured at the receiver positions outside the meeting room.

In terms of measurable quantities, SNR_{UNI32} outside the room is given by

$$SNR_{UNI32} = \frac{1}{16}\sum_{f=160}^{5000}[L_{sp}(f) - LD(f) - L_n(f)]_{-32}, \quad (2)$$

where $L_{sp}(f)$ is the speech level inside the room, $LD(f)$ is the measured level difference between the average level inside the room and the level at a listener position, and $L_n(f)$, is the background noise at the listener position. Frequently the −32 dB limitation has minimal effect, and Eq. (2) can be simplified to

$$SNR_{UNI32} = L_{sp}(\text{avg}) - LD(\text{avg}) - L_n(\text{avg}), \quad (3)$$

where (avg) indicates the arithmetic average of $L_{sp}(f)$, $LD(f)$, and $L_n(f)$ over the 16 $\frac{1}{3}$-octave bands from 160 to 5000 Hz.

A particular speech privacy criterion such as the threshold of intelligibility corresponds to a particular value of SNR_{UNI32}, such as $SNR_{UNI32,0}$. Rewriting Eq. (3) as an inequality, when the following is true of the speech level,

$$L_{sp}(\text{avg}) \leqslant SNR_{UNI32,0} + LD(\text{avg}) + L_n(\text{avg}), \quad (4)$$

then the condition at the listening point is at least as private as the criterion condition, $SNR_{UNI32,0}$. For example, if $SNR_{UNI32,0}=-16$ dB, conditions would be equal to or below the threshold of intelligibility. The quantities $LD(\text{avg})$ and $L_n(\text{avg})$ are properties of the closed room, and Eq. (4) dictates the maximum source room speech level $L_{sp}(\text{avg})$ for which the conditions are adequately private relative to the selected criterion. From the statistics of speech levels, this can be used to determine the interval of time between expected privacy "lapses," which would correspond to instances for which the speech level is larger than that allowed by Eq. (4).

The privacy criterion is usually chosen as the threshold of intelligibility, and consequently the criterion value of −16 dB is used. Although higher intelligibility threshold values were found in more reverberant and highly diffuse environments, this lower value was judged to be a better estimate of worst case conditions in typical office buildings. Using this value Eq. (4) yields the following relationship:

$$LD(\text{avg}) + L_n(\text{avg}) \geqslant L_{sp}(\text{avg}) + 16, \quad (5)$$

which determines the maximum speech level for which conditions at the listening point remain at or below the threshold

TABLE VII. Summary of expected average time intervals between intelligibility and audibility lapses for SPC values from 60 to 90.

SPC	Time between intelligibility lapses	Time between audibility lapses
60	0.32 min	⋯
65	0.76 min	⋯
70	2.87 min	0.62 min
75	18.03 min	2.09 min
80	2.28 h	12.54 min
85	15.30 h	1.53 h
90	⋯	11.22 h

of intelligibility. The likelihood of this speech level being exceeded is the likelihood that the conditions at the listening point will be above the threshold of intelligibility.

VI. SPEECH PRIVACY CLASS

As previously noted, the quantities LD(avg) and L_n(avg) are properties of the building containing the closed room. The sum of these two terms governs the speech privacy rating of a room, and is called the speech privacy class (SPC)

$$\mathrm{SPC} = LD(\mathrm{avg}) + L_n(\mathrm{avg}). \tag{6}$$

For the threshold of intelligibility criterion, Eq. (5) indicates that the probability of intelligible transmitted speech occurring for a particular SPC value can be determined from the probability of a particular related speech level occurring. By replacing the criterion $\mathrm{SNR}_{\mathrm{UNI32,0}}$ value with a value of −22 dB, the results would indicate the probability of the threshold of audibility being exceeded.

The probabilities of various speech levels are given in Figs. 4 and 5 but are in terms of A-weighted speech levels. To convert the speech levels to frequency averages in the form of L_{sp}(avg) values, Pearsons'[1] speech spectra were used. Pearsons' "raised voice" level spectra for male and female talkers were averaged to obtain a representative spectrum shape for meeting room speech. Pearsons' raised voice level spectra were selected because the overall A-weighted levels of these spectra are similar to the average speech level recorded in the 79 meetings that were measured (see Table II). For this raised speech spectrum, the frequency average speech level L_{sp}(avg) was 12.2 dB less than the A-weighted level of the same spectrum. This was used to convert A-weighted speech levels to frequency averaged speech levels, L_{sp}(avg). It was then possible to determine the probabilities of the occurrence of L_{sp}(avg) values required to determine the probabilities of the related SPC values occurring.

For SPC values in five point steps, the related L_{sp}(avg) values were used to determine the probability of speech privacy lapses. This was done in terms of the threshold of intelligibility being exceeded and for the threshold of audibility being exceeded. The results of these calculations are given in terms of the average time intervals between speech privacy lapses in Table VII.

Using Eq. (6), the SPC values can be related to combinations of ambient noise level at the listener position, L_n(avg), and the measured level difference, LD(avg). Some examples are included in Table VIII for the SPC values used in Table VII. In this case, approximately equivalent A-weighted noise levels, $L_n(A)$, to the average noise levels, L_n(avg), were determined to provide comparable more familiar values. This was done by determining the difference between L_n(avg) and $L_n(A)$ values for a −5 dB per octave spectrum shape, which is representative of indoor ambient noises.[9,10] This resulted in an adjustment of +10.7 dB to L_n(avg) values to approximate equivalent A-weighted levels.

TABLE VIII. SPC values for combinations of background noise, L_n(avg), and related LD(avg) values.

	Very quiet	Quiet	Moderate noise
$L_n(A)$, dBA	25.0	35.0	45.0
L_n(avg), dB	14.3	24.3	34.3
SPC	***LD*(avg)**	***LD*(avg)**	***LD*(avg)**
60	45.7	35.7	25.7
65	50.7	40.7	30.7
70	55.7	45.7	35.7
75	60.7	50.7	40.7
80	65.7	55.7	45.7
85	70.7	60.7	50.7
90	75.7	65.7	55.7

The combination of the information in Tables VII and VIII can be used to estimate the probability of speech privacy lapses for meeting rooms where the ambient noise levels outside the room, L_n(avg), and the transmission characteristics from the room, LD(avg), have been determined.

VII. EFFECTS OF OTHER VARIABLES AND SOURCES OF ERROR

A number of other factors could influence the speech privacy of meeting rooms or be potential sources of error in the prediction of the expected speech privacy. In particular situations, it may be desirable to include some estimate of their effects.

A. Other variables influencing speech privacy

The criterion value $\mathrm{SNR}_{\mathrm{UNI32,0}}$ for the threshold of intelligibility was determined to be −16 dB in approximately free field conditions with spatially separated speech and noise sources,[6] and −11 dB in conditions with ideally diffuse noise and reverberation times of approximately 0.8 s.[7] Differences in meeting room reverberation time and the spatial separation of speech and noise sources could influence the choice of $\mathrm{SNR}_{\mathrm{UNI32,0}}$ value for the threshold of intelligibility. However, there is no evidence that these effects influence the threshold of audibility.

Speech levels at positions outside the meeting room are determined from measured level differences between the source room-average level and the level at each receiver position. Although the source room-average level usefully represents the average of all possible locations of talkers in the meeting room, and provides more repeatable results, some locations might lead to higher transmitted speech levels and

hence to reduced speech privacy. This is most likely for talker positions close to the meeting room boundary and this is most likely to be a problem in smaller rooms. (The four points to the upper left of Fig. 2 may illustrate this phenomenon).

Although a talker located close to the room boundary in the meeting room might be expected to lead to higher incident speech levels, such a talker would most likely be directing their voice toward other occupants in the middle of the room and not toward the room boundary. The effects of the directionality of the talker's voice might lead to a larger reduction in incident levels than the increased incident levels due to the proximity of the talker to the room boundaries. It is not clear how much such a talker location would modify the speech privacy of the room, but the effects would also depend on how reverberant conditions in the meeting room were.

Some listeners are better able to understand speech in noise than others. Their listening abilities would be influenced by their hearing sensitivity and their familiarity with the language being spoken. The listening test results used to determine the thresholds of audibility and intelligibility were based on the condition for which 50% of the listeners with better than average hearing sensitivity could just hear or just understand the transmitted speech. A significant portion of listeners would be able to hear or understand speech at transmitted speech levels several decibels lower. Such more skilled than average listeners would pose a little higher risk of speech privacy lapses.

To account for such better than average listeners, the threshold of intelligibility and audibility could be set at a lower percentage of listeners being able to hear or understand transmitted speech than 50%. This would provide a reduced risk of speech privacy problems but would probably lead to more expensive constructions.

The calculations assumed knowledge of an average noise level at the listener position. Typical ambient noises can be expected to vary in level over time. Intermittent increases in ambient noise levels due to various transient sources are likely but would not be a problem because they would only result in increases in speech privacy. Transient reductions in ambient noise levels seem much less likely to occur and hence fluctuating noise levels are not expected to be a significant problem.

B. Possible sources of error

There are always possible errors associated with measurements of sound levels. There could be errors associated with the measured speech levels used to describe the statistical properties of speech levels in meeting rooms. However, with such a large number of measurements with many different independent calibrations of equipment, it is expected that the possible small errors would tend to average out. Measurements of ambient noise levels and level differences for particular meeting rooms could also be a source of errors. The ASTM E2638 standard[8] discusses the accuracy of LD(avg) measurements.

There are also possible errors in approximate conversions between measures. For example, the measured speech levels were in terms of A-weighted levels, $L_{sp}(A)$, which were converted to corresponding L_{sp}(avg) values. Pearsons' speech spectra used to determine the appropriate conversion are for speech levels close to the talker and would not exactly represent speech spectra at more distant locations in each meeting room. The possible errors would be expected to vary from room to room and are not easily estimated accurately.

Similar conversions from L_n(avg) to $L_n(A)$ values were also made and would include similar possible errors. However, these were not a part of the calculations and were only to help readers understand the significance of the L_n(avg) values.

VIII. CONCLUSIONS

In meetings using sound amplification systems, average speech levels near the periphery of the room were only 2.0 dBA higher than in rooms without sound amplification. There was a small effect of room size on average speech levels in meeting rooms without sound amplification. However, this effect was not found in meeting rooms using sound amplification systems. Sound amplification systems were found to increase speech levels at more distant locations to more closely equal those in smaller rooms without sound amplification. The decrease in levels with room size for meetings without sound amplification was not concluded to be an important effect, because in the larger meeting rooms, sound amplification systems are normally used. In smaller rooms (<100 m^3) average speech levels were more varied because of the higher probability of talkers being closer to the measurement microphones.

Speech levels in rooms, both with and without sound amplification, increased with increasing background noise levels. To maximize speech intelligibility in the meeting room, and speech privacy at points outside the room, it is important to reduce ambient noise levels inside meeting rooms. Averages of 10 s L_{90} values during meetings were found to be a good indication of background noise levels in meeting rooms and indicated similar values to L_{eq} values measured in the unoccupied rooms.

The measurements of 10 s L_{eq} values of sounds in meeting rooms provided a large sample of meeting room speech levels for estimating the probability of particular speech levels being exceeded and the average interval between occurrences of particular speech levels in meeting rooms. Because meetings with and without sound amplification had similar speech levels and room size effects were not usually important, the combined speech data from all meetings were used to estimate speech level statistics in all meeting rooms.

Using the uniformly weighted frequency-averaged signal-to-noise ratio concept and values of this measure for the thresholds of intelligibility and audibility of transmitted speech, the probability of meeting room speech privacy lapses can be determined. Speech privacy was related to the SPC, which is the sum of the background noise level at a position outside the room, L_n(avg), and the level reduction

from room-average levels within the meeting room to this same measurement point outside the room, LD(avg). For any combination of LD(avg) and L_n(avg), the probability of speech being audible or intelligible at points just outside the room can be estimated. The procedure is applicable for all levels of speech privacy from modest confidentiality to high speech security.

There are a number of special cases that may be less well predicted from these results. These would include smaller rooms with talkers close to the room boundaries and especially so if they are near sound isolation weak spots such as doors. They may also include rooms with teleconferences where participants may, on average, talk louder than in other rooms. These cases were not included in the current study but could be considered in future investigations.

ACKNOWLEDGMENTS

Funding for this work was provided by Public Works and Government Services Canada (PWGSC) and the Royal Canadian Mounted Police (RCMP).

[1]W. O. Olsen, "Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearsons, Bennett, and Fidell (1977) report," J. Audiol. **7**, 1–5 (1998).
[2]J. C. Junqua, "The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex," Speech Commun. **20**, 13–22 (1996).
[3]J. S. Bradley, "Predictors of speech intelligibility in rooms," J. Acoust. Soc. Am. **80**, 837–845 (1986).
[4]American National Standard, "Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools," ANSI S12.60 (American National Standards Institute, New York, 2002).
[5]J. S. Bradley and H. Sato, "The Intelligibility of speech in elementary school classrooms," J. Acoust. Soc. Am. **123**, 2078–2086 (2008).
[6]B. N. Gover and J. S. Bradley, "Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms," J. Acoust. Soc. Am. **116**, 3480–3490 (2004).
[7]J. S. Bradley, M. Apfel, and B. N. Gover, "Some spatial and temporal effects on the speech privacy of meeting rooms," J. Acoust. Soc. Am. **125**, 3038–3051 (2009).
[8]ASTM E2638-08, "Standard test method for objective measurement of the speech privacy provided by a closed room," American Society for Testing and Materials, Philadelphia (2008).
[9]D. F. Hoth, "Room noise spectra at subscribers' telephone locations," J. Acoust. Soc. Am. **12**, 449–504 (1941).
[10]W. E. Blazier, "Revised noise criteria for application in the acoustical design and rating of HVAC systems," Noise Control Eng. J. **16**, 64–73 (1981).

Interaural cross correlation in a sound field represented by spherical harmonics

Boaz Rafaely[a] and Amir Avni
Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

(Received 16 May 2009; revised 3 December 2009; accepted 5 December 2009)

Spatial impression is an important acoustic quality of concert halls. An accepted objective measure for spatial impression is the interaural cross-correlation (IACC) coefficient. Recently, spherical microphone arrays have been studied for room acoustics analysis and music recordings. This study presents a theoretical formulation for the computation of IACC using spherical-harmonics representations of the sound field, as measured by spherical microphone arrays, and spherical-harmonics representation of head-related transfer functions (HRTFs), taken from HRTF databases. As spherical microphone arrays typically use a finite number of microphones, they may not be able to capture the complete spatial information in a sound field. Therefore, the effect of limited spherical-harmonics order on the accuracy of IACC approximation using the proposed method is studied using simulated and measured data. The method presented in this paper can be further used to study the effect of limited spatial information on the spatial perception of sound fields. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3278605]

PACS number(s): 43.55.Hy, 43.55.Mc [NX] Pages: 823–828

I. INTRODUCTION

Spatial impression is an important attribute in concert halls. In particular, apparent source width and listener envelopment were found to contribute significantly to spatial impression.[1] Although these are subjective measures, several objective measures have been proposed for these qualities. Among these, the interaural cross-correlation (IACC) coefficient seems to be one of the most important measures.[2] IACC measures the correlation over time between the sound pressure at the two ears of a listener, and is influenced by the sound field around the listener's head, and the acoustic effect of the head. The latter is typically described by the head-related transfer functions (HRTFs), which are the response functions between a distant source and the pressure at the ears and are functions of frequency and source direction. IACC of a concert hall can be calculated from *in-situ* measurements made on a human listener or using a dummy head.[3]

Recently, spherical microphone arrays have been developed and applied to the analysis of sound fields in rooms and auditoria. Decomposition of the sound fields into plane waves was shown useful in the identification of early reflections in an auditorium,[4,5] while beamforming applied to directional sound-field analysis was shown useful in the calculation of a range of spatial sound-field measures.[6,7] As data from spherical arrays may become more available, it may be useful to calculate IACC using such data, without the need for additional measurements using human or dummy heads.

This paper presents a theoretical formulation of the sound field and of HRTFs using spherical harmonics, showing that a product of the spherical-harmonics coefficients of the two can be used in the calculation of IACC. In addition to the ability to calculate IACC using sound-field and HRTF data, this formulation allows investigation of the effect of sound-field order, in the spherical-harmonics domain, on the values of the IACC. Changes in IACC values may suggest changes in sound perception and so this may be used as a simple objective way to investigate the effect of sound-field order on its spatial perception.

To summarize, this paper offers several contributions.

(1) Theoretical formulation for the calculation of IACC using sound-field data and HRTF data, both in the spherical-harmonics domain.
(2) The theoretical formulation allows calculation of IACC using sound-field data, as measured by a spherical microphone array or generated by simulation, for example, and HRTF data, as measured on humans or dummy heads.
(3) Calculation of IACC using sound-field data represented in the spherical-harmonics domain facilitates a study of the effect of sound-field order on IACC and therefore on spatial impression. The results presented in this study may therefore be useful in areas such as sound-field reproduction and synthesis, and auralization.
(4) A secondary contribution is the analysis of the effect of HRTF directions set on the modeling of HRTF data using spherical harmonics, where the advantage of direction sets that cover the entire directional range is highlighted.

This paper is presented as follows. After a review on sound-field representation by spherical harmonics, the theoretical formulation of IACC calculation from sound-field and HRTF data is presented. Then, HRTF modeling by spherical harmonics is studied, as this is important for the IACC calculation. This paper is concluded with an experimental study

[a]Author to whom correspondence should be addressed. Electronic mail: br@ee.bgu.ac.il

based on both measured and simulated data of the sound field in a concert hall and measured HRTF data from the CIPIC (Center for Image Processing Integrated Computing) database.[8]

II. SOUND-FIELD REPRESENTATION BY SPHERICAL HARMONICS

Spherical-harmonics decomposition is employed in this paper to describe the sound field[9] and is briefly introduced. The standard spherical coordinate system, (r,θ,ϕ),[10] is used throughout the paper. Consider a sound pressure function $p(k,r,\theta,\phi)$, which is square integrable over $\Omega \equiv (\theta,\phi)$, with k the wavenumber, then its spherical Fourier transform, denoted by $p_{nm}(k,r)$, and the inverse transform are given by[11]

$$p_{nm}(k,r) = \int_{\Omega \in S^2} p(k,r,\Omega) Y_n^{m*}(\Omega) d\Omega, \tag{1}$$

$$p(k,r,\Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} p_{nm}(k,r) Y_n^m(\Omega), \tag{2}$$

where $\int_{\Omega \in S^2} d\Omega \equiv \int_0^{2\pi} \int_0^{\pi} \sin\theta d\theta d\phi$. The spherical harmonics $Y_n^m(\cdot)$ are defined by

$$Y_n^m(\Omega) \equiv Y_n^m(\theta,\phi) \equiv \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}, \tag{3}$$

where n is the order of the spherical harmonics, $P_n^m(\cdot)$ is the associated Legendre function, and $i=\sqrt{-1}$. When the sound field $p(k,r,\Omega)$ is produced by a single plane wave with amplitude $a(k)$, and with an incident direction Ω_0, then $p_{nm}(k,r)$ is given by[9]

$$p_{nm}(k,r) = a(k) b_n(kr) Y_n^{m*}(\Omega_0), \tag{4}$$

where $b_n(kr)=4\pi i^n j_n(kr)$ for free-field,[9] or an open pressure measurement sphere of radius r, and $b_n(kr)=4\pi i^n (j_n(kr) - [j_n'(kr_0)/h_n'(kr_0)] h_n(kr))$ for a measurement sphere of radius r around a rigid sphere of radius $r_0 \le r$. Other configurations and corresponding expressions for $b_n(kr)$ have also been presented.[12] Here $j_n(\cdot)$ is the spherical Bessel function, $h_n(\cdot)$ is the spherical Hankel function, and j_n', h_n' represent derivatives.

In the case of a more general sound field, a plane-wave decomposition representation can be employed by assuming that the sound field is composed of an infinite number of plane-waves arriving from all directions with spatial amplitude density of $a(k,\Omega_0)$. The coefficients p_{nm} can be calculated in this case by taking an integral of Eq. (4) over all directions Ω_0:[13]

$$\begin{aligned} p_{nm}(k,r) &= \int_{\Omega_0 \in S^2} a(k,\Omega_0) b_n(kr) Y_n^{m*}(\Omega_0) d\Omega_0 \\ &= a_{nm}(k) b_n(kr), \end{aligned} \tag{5}$$

with $a_{nm}(k)$ the spherical Fourier transform of $a(k,\Omega)$. Now, having approximated p_{nm} from measurements, plane-wave decomposition or computation of a_{nm} can simply be calculated using

$$a_{nm}(k) = \frac{p_{nm}(k,r)}{b_n(kr)}. \tag{6}$$

Methods to design arrays that avoid zero or low values of b_n,[12,14] and theoretical[13] and experimental[4,5] investigations of plane-wave decomposition have been previously presented.

III. IACC CALCULATION FROM SPHERICAL-HARMONICS MODELS OF HRTF AND THE SOUND FIELD

HRTFs represent the frequency response from a typically far source to the left and right ear canals of a human head or a dummy head. Left-ear and right-ear HRTFs are denoted by $H^l(k,\Omega)$ and $H^r(k,\Omega)$, respectively, where $k=2\pi f/c$, with f the frequency in Hz, c the speed of sound, and Ω the direction of the source used in the measurement of the HRTF. Since HRTFs are a function of the solid angle Ω, we can apply the spherical Fourier transform as in Eq. (1) to compute $H_{nm}^l(k)$ and $H_{nm}^r(k)$.

The frequency response from a source in a room or an auditorium to a listener's ears will be a function of the HRTFs but also the sound field in the auditorium. Assuming that the sound field in the room is composed of plane waves arriving from directions Ω_0 with amplitudes $a(k,\Omega_0)$ relative to the origin of the coordinate system, then the contribution of a single plane wave to the pressure at the left and right ears of the listener is given by $a(k,\Omega_0)H^l(k,\Omega_0)$ and $a(k,\Omega_0)H^r(k,\Omega_0)$, respectively. Now, with a continuum of plane waves composing the sound field in the room, the pressure at the left ear can be written as[15]

$$p^l(k) = \int_{\Omega_0 \in S^2} a(k,\Omega_0) H^l(k,\Omega_0) d\Omega_0. \tag{7}$$

Denoting $\tilde{a}(k,\Omega) \equiv a^*(k,\Omega)$ where "*" denotes complex conjugate and substituting the spherical Fourier transform representation of $\tilde{a}$ and H^l as in Eq. (2), we can write

$$\begin{aligned} p^l(k) = \int_{\Omega_0 \in S^2} & \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \tilde{a}^*{}_{nm}(k) Y_n^{m*}(\Omega_0) \\ & \times \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} H_{n'm'}^l(k) Y_{n'}^{m'}(\Omega_0) d\Omega_0. \end{aligned} \tag{8}$$

Exchanging terms and using the orthogonality property of the spherical harmonics,[9] we get

$$p^l(k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \tilde{a}^*{}_{nm}(k) H_{nm}^l(k). \tag{9}$$

Equation (8) shows how the spherical-harmonics representation of the sound field and the spherical-harmonics representation of the HRTFs are combined to represent the pressure at the ear. A similar expression can be developed for representing the pressure at the right ear:

$$p^r(k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \tilde{a}^*{}_{nm}(k) H_{nm}^r(k). \tag{10}$$

The normalized interaural cross-correlation function is given by[3]

$$\rho_{t_1,t_2}(\tau) = \frac{\int_{t_1}^{t_2} p^l(t)p^r(t+\tau)dt}{\sqrt{\int_{t_1}^{t_2}[p^l(t)]^2dt \cdot \int_{t_1}^{t_2}[p^r(t)]^2dt}}. \quad (11)$$

The interaural cross-correlation coefficient is then given by taking the maximum value over τ,

$$\mathrm{IACC}_{t_1,t_2} = \max_{\tau}|\rho_{t_1,t_2}(\tau)|, \quad \tau \in (-1,1)\ \mathrm{ms}. \quad (12)$$

Taking the inverse Fourier transforms of $p^l(k)$ and $p^r(k)$ to compute $p^l(t)$ and $p^r(t)$, respectively, we can substitute the latter in Eqs. (11) and (12) to compute the IACC.

To summarize, the following procedure is followed for the calculation of the IACC using spherical-harmonics models.

(1) Spherical-harmonics model of the HRTFs is prepared using a HRTF database, and H^l_{nm}, H^r_{nm} are calculated.
(2) A sound field is measured using a spherical microphone array, and its spherical-harmonics representation $\tilde{a}_{nm}$ is extracted.
(3) Impulse responses to the left and right ears are calculated using Eqs. (9) and (10) and the inverse Fourier transform.
(4) IACC is calculated using Eqs. (11) and (12).

Although the use of the spherical-harmonics model is not necessary for the calculation of the IACC, it does facilitate the study of the effect of spherical-harmonics order on the spatial properties of a sound field as quantified by the IACC.

IV. HRTF MODELING BY SPHERICAL HARMONICS

Analysis of HRTF using spherical harmonics has been previously presented.[16,17] Nevertheless, in this work we aim to study the overall modeling error as a function of order N and frequency, and the effect of the direction sampling set. Given Q left-ear or right-ear head-related transfer functions denoted by $H(k,\Omega_q)$, $q=1,\ldots,Q$, they can be represented by a spherical-harmonics expansion as in Eq. (2) using coefficients $H_{nm}(k)$ written in a matrix form for a specified wavenumber as

$$\mathbf{H} = \mathbf{Y}\mathbf{H_{nm}}, \quad (13)$$

where $\mathbf{H}$ is a $Q\times 1$ vector of HRTFs given by

$$\mathbf{H} = [H(k,\Omega_1), H(k,\Omega_2), \ldots, H(k,\Omega_Q)]^T, \quad (14)$$

and $\mathbf{H_{nm}}$ is the corresponding $(N+1)^2\times 1$ vector of spherical-harmonics coefficients given by

$$\mathbf{H_{nm}} = [H_{00}(k), H_{1(-1)}(k), H_{10}(k), H_{11}(k), \ldots, H_{NN}(k)]^T, \quad (15)$$

where it is assumed that coefficients of orders $n \le N$ are used in this representation. Finally, the $Q\times(N+1)^2$ matrix $\mathbf{Y}$ of spherical harmonics is given by

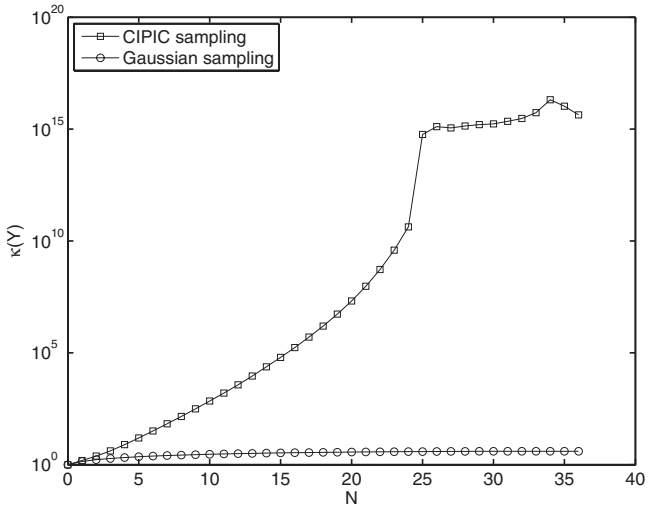


FIG. 1. Condition number of matrix $\mathbf{Y}$ with the CIPIC direction sampling set and the Gaussian direction sampling set.

$$\mathbf{Y} = \begin{bmatrix} Y_0^0(\Omega_1) & Y_1^{-1}(\Omega_1) & Y_1^0(\Omega_1) & Y_1^1(\Omega_1) & \ldots & Y_N^N(\Omega_1) \\ Y_0^0(\Omega_2) & Y_1^{-1}(\Omega_2) & Y_1^0(\Omega_2) & Y_1^1(\Omega_2) & \ldots & Y_N^N(\Omega_2) \\ \vdots & \vdots & \vdots & \vdots & & \\ Y_0^0(\Omega_Q) & Y_1^{-1}(\Omega_Q) & Y_1^0(\Omega_Q) & Y_1^1(\Omega_Q) & \ldots & Y_N^N(\Omega_Q) \end{bmatrix}. \quad (16)$$

In practice, $\mathbf{H}$ can be taken from a HRTF database, while $\mathbf{Y}$ can be generated knowing the Q HRTF directions in the database. Now, the coefficient vector $\mathbf{H_{nm}}$ can be derived numerically by solving Eq. (13), with a possible solution given by

$$\mathbf{H_{nm}} = \mathbf{Y}^{\dagger}\mathbf{H}, \quad (17)$$

where $\mathbf{Y}^{\dagger}$ is the pseudo-inverse of $\mathbf{Y}$, solving Eq. (13) in a least-squares sense.

It is clear from Eq. (17) that the condition number of $\mathbf{Y}$ needs to be sufficiently low in order to ensure numerical robustness of the computation. Furthermore, it was shown that low condition number also ensures robustness against uncertainty in the computation, such as inaccuracy in the direction angles Ω_q.[14]

The CIPIC HRTF database[8] has been used in this work for the experimental study. In the CIPIC database, HRTFs over a dense grid of directions have been measured for a large number of human subjects. The condition number of matrix $\mathbf{Y}$ with directions Ω_q taken from the CIPIC database was plotted against the order N and compared to another distribution of points in Ω, referred to as the Gaussian sampling.[18] Both sets have a total of 1250 directions. Figure 1 shows that the condition number using the Gaussian sampling is reasonably small at the entire range of orders, while the condition number using the CIPIC sampling degrades with increasing order N. The poor conditioning of the CIPIC sampling scheme can be attributed to the missing samples around elevation directions $\Omega=(\theta,\phi)=(180°,\phi)$, i.e., the south-pole directions. This missing direction range is probably of less importance to most applications involving HRTFs, but is required to achieve numerically robust decom-

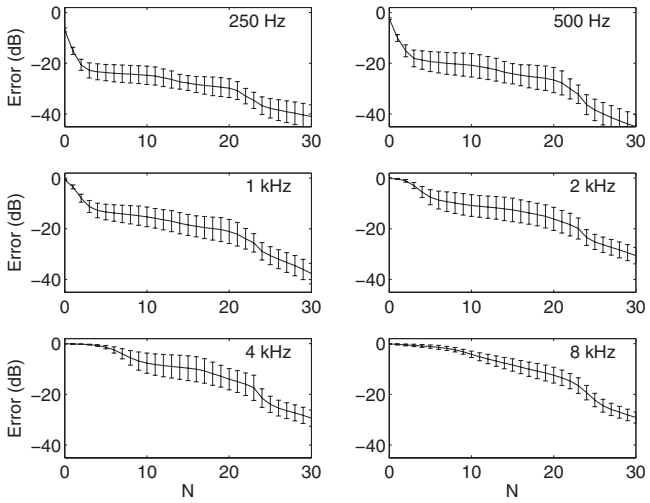


FIG. 2. Approximation error ϵ_{250} at 250 Hz through ϵ_{8000} at 8 kHz, as a function of order N, averaged over ten subjects (20 ears) from the CIPIC database. Also shown are standard deviation bars for each curve.

position of the HRTFs into spherical harmonics. In the Gaussian sampling scheme, the samples are distributed more evenly along Ω, keeping the condition number small. This shows that decompositions of HRTFs into spherical harmonics of high orders require directions which are sampled more uniformly, covering the entire range of directions. Nevertheless, the CIPIC database has been used here even with the poor conditioning; useful decomposition can still be achieved due to the large dynamic range in numerical programming tools such as MATLAB.

In the next study, the CIPIC database was used to investigate HRTF representation using spherical harmonics. The left-ear and right-ear HRTFs of ten subjects from the CIPIC database (overall 20 HRTF sets) have been used in a simulation study aiming to investigate the accuracy of modeling the measured HRTFs by spherical harmonics as a function of spherical-harmonics order N. Equation (17) was used to compute the spherical-harmonics representation $\mathbf{H_{nm}}$, using matrix $\mathbf{Y}$ computed using the known HRTF directions and matrix $\mathbf{H}$ holding the HRTF data set. Equation (13) was then used to reconstruct an approximated HRTF matrix $\hat{\mathbf{H}}$. An error metric for this approximation was then defined as

$$\epsilon(f) = \frac{\|\mathbf{H} - \hat{\mathbf{H}}\|}{\|\mathbf{H}\|}, \tag{18}$$

where $\|\cdot\|$ denotes the 2-norm. In practice, the analysis was performed in octave bands by initially filtering each head-related impulse response with an octave band-pass filter and then integrating the resulting error $\epsilon(f)$ over frequency. This process generated six error metrics for each octave band: ϵ_{250}, ϵ_{500}, ϵ_{1000}, ϵ_{2000}, ϵ_{4000}, and ϵ_{8000}, corresponding to center frequencies of 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz, respectively. Figure 2 shows the error results for each octave band as a function of order N. It is clear that at low frequencies, where the HRTFs vary less as a function of direction, good approximation is achieved using relatively low order N, while at high frequencies, higher spherical-harmonics orders are required to achieve good approximation. Error values of −3 and −10 dB were used as cut-off values to represent two levels of approximation accuracies. The minimal order N required to achieve a level of error not higher than the cut-off value is shown in Table I for each of the octave bands. A maximum error of −3 dB is achieved for orders ranging from $N=0$ at the low frequency end to $N=6$ at the high frequency end. To achieve a lower error of −10 dB, orders ranging from 1 to 12 along the frequency range are required.

TABLE I. The first two lines show the minimal spherical-harmonics order N required to achieve an approximation error less than −10 and −3 dB for each octave band. The third line shows the value of kr calculated using the center frequency of each frequency band, assuming that $r=0.09$ m and $c=343$ m/s.

	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
−10 dB error	1	1	2	6	8	12
−3 dB error	0	1	1	3	5	6
kr	0.41	0.82	1.65	3.30	6.60	13.19

The third row in the table shows values of kr for each of the center frequencies, assuming that $c=343$ m/s and $r=0.09$ m, approximating an average head. The values in this row are very similar to the −3 dB error row. This can be explained by modeling the human head as a rigid sphere, in which case a plane-wave sound field around the sphere has a spherical-harmonics expansion with significant coefficients only in the range $kr \leq N$.[9,13] As the rigid-sphere model is not an accurate model for a head, in particular, at high frequencies, the third row in the table may be used to explain the general behavior of the first two rows in the table, but not necessarily to derive explicit spherical-harmonics order bounds for a given HRTF data set.

V. STUDY I: SIMULATED SOUND FIELD

A study aiming to investigate the accuracy of reproducing IACC values using finite-order spherical-harmonics representation has been performed using simulated and measured data. First, the sound field in the Elmia concert hall has been simulated using ODEON (AUDITORIUM version 9.2). The hall has 1033 seats and a volume of 12 800 m^3. Data from the hall simulation, such as the direction-of-arrival, amplitude, and delay of the direct sound and room reflections, have been used in a MATLAB (version 7.7.0)-based simulation to compute the pressure sound field at various positions in the room volume, including representation of the sound field using spherical harmonics. Six measurement positions and two source positions have been examined, producing a total of 12 spatial arrangements. For each arrangement, the $IACC_E$ value has been calculated using early room reflections (first 100 ms) and measured HRTF data of nine human listeners taken from the CIPIC database and Eq. (12). Table II shows the mean and standard deviation of the $IACC_E$ values across all 12 arrangements and nine listeners. The table shows that $IACC_E$ values are high at low frequencies and low at high frequencies, with relatively low variance.

Now, $IACC_E$ for all arrangements and listeners has been computed using spherical-harmonics representation, as de-

TABLE II. $IACC_E$ values at various frequencies and $E3$ (averaging IACC values for 500 Hz–2 kHz) for the Elmia concert hall simulation. Mean and STD are the average and standard deviation over 12 positions in the room and nine subjects for each position.

	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	$E3$
Mean	0.727	0.634	0.588	0.384	0.395	0.232	0.536
STD	0.08	0.09	0.065	0.094	0.08	0.055	0.064

tailed in Sec. III for various orders N. The error between the direct calculation and the spherical-harmonics-based calculation was averaged over all arrangements and listeners, producing root-mean-squared error plots shown in Fig. 3. Each sub-figure shows results for $IACC_E$ calculation performed at a single octave band, and E3 averaging bands 500–2000 Hz. Similar to the behavior demonstrated in Table I, the errors in Fig. 3 tend to increase with frequency, with a cut-off corresponding to $N \approx kr$ as in Table I, producing $IACC_E$ errors of around 0.1–0.2 (in IACC value scale). Also, an error of around 0.1 in IACC-E3 is achieved with order $N=3$ and a lower error of 0.055 is achieved with order $N=6$.

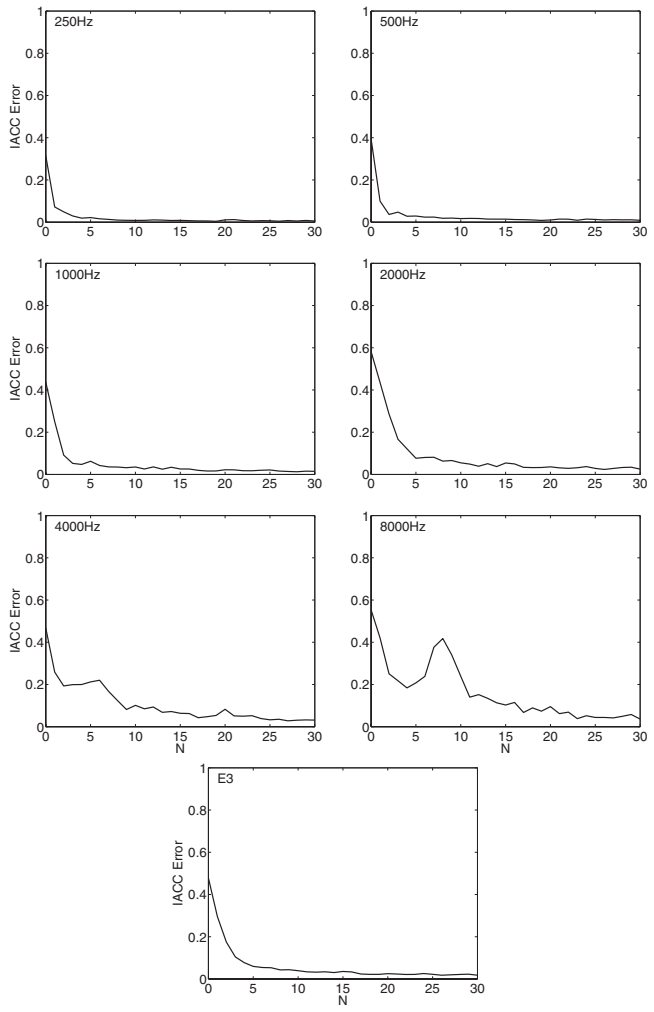


FIG. 3. The root-mean-squared error between $IACC_E$ values calculated directly using Eq. (12) and calculated using spherical harmonics of order N averaged over all 12 hall positions and nine listeners. Frequency ranges of calculation in terms of octave bands are marked on each sub-figure.

Another aspect under study was the value of τ when computing the IACC according to Eq. (12). There were very small fluctuations of τ using different spherical-harmonics orders and the averaged error between τ values in the direct calculation and the spherical-harmonics-based calculation was negligible, e.g., less than 0.05 ms, regardless of the spherical harmonics order or the octave band.

To summarize, this experimental study demonstrated the use of spherical-harmonics representation in the computation of $IACC_E$, presenting diminishing errors with increasing spherical-harmonics order. As an example, $IACC_{E3}$ was computed with reasonable accuracy using spherical-harmonics orders in the range $N=3-6$.

VI. STUDY II: MEASURED SOUND FIELD

An experimental study has been preformed aiming to verify the theory presented in this paper, using a KEMAR (Knowles Electronics Manikin for Acoustics Research) dummy head and a spherical microphone array. The experiment has been performed at the Kreitman building, a 250 seat lecture hall in Ben-Gurion University having a reverberation time of around 1 s. A Brüel&Kjær omni-source loudspeaker was placed around the center of the stage and was driven by a swept-sine signal using MATLAB, a sound card type ESI U24XL, and a Brüel&Kjær amplifier type 2716c. The sound field was then recorded using GRAS KEMAR dummy head positioned on a seat around the central area of the hall. The KEMAR included head and torso and was fitted with large ears, and a set of GRAS 40AO 0.5 in. prepolarized pressure ear-microphones, connected to a Brüel&Kjær NEXUS 2690-OS2 amplifier. The impulse response from the loudspeaker to each of the ear-microphones was calculated, from which $IACC_E$ was computed according to Eq. (12). Then, the KEMAR was removed and a spherical array, mhacoustics em32 Eigenmike®, was positioned on the same seat. The spherical array consists of 32 microphones around a rigid sphere, having a radius of 9 cm. The impulse response from the loudspeaker to the microphone was computed as described above. The spherical array data, together with the HRTF data from the CIPIC database, were used to compute the $IACC_E$, as described in Secs. II and III. The procedure described above was repeated for 20 different seats in the hall, all around the central seating area. The mean-squared error between the KEMAR $IACC_E$ and the em32 $IACC_E$ was calculated as an average over all seats. In this experiment a sound-field representation using spherical harmonics of order up to $N=3$ was used due to the limited number of microphones in the em32 array.

Table III presents the average error between the $IACC_E$ values computed by the KEMAR and the em32 measurements. The error is presented for each of the midoctave bands and their average $E3$. Spherical harmonics up to order of $N=3$ should provide a sufficient number of coefficients for sound-field analysis up to the 2 kHz octave band, as shown in Table III. An error of 0.07–0.12 for different octave bands was observed in the experiment. This is comparable to the errors presented in Fig. 3 for $N=3$ in study I. Note that

TABLE III. The first two rows present the $IACC_E$ mean and STD values using the measurements of the KEMAR dummy. The values are averaged over 20 seats around the hall and computed at various frequencies and $E3$. The last row presents the average error between the two $IACC_E$ values of the KEMAR and the em32.

	500 Hz	1 kHz	2 kHz	E_3
Mean	0.51	0.52	0.33	0.45
STD	0.08	0.08	0.13	0.06
Error	0.07	0.12	0.09	0.09

$IACC_{E3}$ was computed using spherical-harmonics order of $N=3$ and produced $IACC_{E3}$ average error of 0.09 compared to 0.1 in the simulated sound field, Fig. 3.

In conclusion, this experiment demonstrated the feasibility of computing IACC values using a spherical microphone array having a limited spherical-harmonics sound-field representation and a HRTF database.

VII. CONCLUSION

This paper presented a mathematical framework for computing IACC from spherical-harmonics representation of the sound-field and of HRTF data. The framework is useful for computing IACC when sound-field data have been measured by spherical microphone arrays or for analyzing the effect of spherical-harmonics order on spatial impression.

ACKNOWLEDGMENT

This research was supported by the Israel Science Foundation (Grant No. 155/06).

[1]L. L. Berank, "Concert hall acoustics—2008," J. Audio Eng. Soc. **56**, 532–544 (2008).
[2]T. Okano, L. L. Berank, and T. Hidaka, "Relations among interaural cross-correlation coefficient ($IACC_E$), lateral fraction (LF_E), and apparent source width (ASW) in concert halls," J. Acoust. Soc. Am. **104**, 255–265 (1998).
[3]ISO 3382:1997, International Organization for Standardization, Geneva, Switzerland, Acoustics—Measurement of the reverberation time of rooms with reference to other acoustical parameters (1997).
[4]M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposition using spherical microphone array," J. Acoust. Soc. Am. **118**, 3094–3103 (2005).
[5]B. Rafaely, I. Balmages, and L. Eger, "High-resolution plane-wave decomposition in an auditorium using a dual-radius scanning spherical microphone array," J. Acoust. Soc. Am. **122**, 2661–2668 (2007).
[6]B. N. Gover, J. G. Ryan, and M. R. Stinson, "Microphone array measurement system for analysis of directional and spatial variations of sound fields," J. Acoust. Soc. Am. **112**, 1980–1991 (2002).
[7]B. N. Gover, J. G. Ryan, and M. R. Stinson, "Measurement of directional properties of reverberant sound fields in rooms using a spherical microphone array," J. Acoust. Soc. Am. **116**, 2138–2148 (2004).
[8]V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2001), New Paltz, NY (2001), pp. 99–102.
[9]E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, New York, 1999).
[10]G. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 5 ed. (Academic, San Diego, CA, 2001).
[11]J. R. Driscoll and D. M. Healy, Jr., "Computing Fourier transforms and convolutions on the 2-sphere," Adv. Appl. Math. **15**, 202–250 (1994).
[12]I. Balmages and B. Rafaely, "Open-sphere designs for spherical microphone arrays," IEEE Trans. Audio, Speech, Lang. Process. **15**, 727–732 (2007).
[13]B. Rafaely, "Plane-wave decomposition of the pressure on a sphere by spherical convolution," J. Acoust. Soc. Am. **116**, 2149–2157 (2004).
[14]B. Rafaely, "The spherical-shell microphone array," IEEE Trans. Audio, Speech, Lang. Process. **16**, 740–747 (2008).
[15]D. Menzies and M. Al-Akaidi, "Nearfield binaural synthesis and ambisonics," J. Acoust. Soc. Am. **121**, 1559–1563 (2007).
[16]M. J. Evans, J. A. S. Angus, and A. I. Tew, "Analyzing head-related transfer function measurements using surface spherical harmonics," J. Acoust. Soc. Am. **104**, 2400–2411 (1998).
[17]Y. Tao, A. I. Tew, and S. J. Porter, "A study on head-shape simplification using spherical harmonics for HRTF computation at low frequencies," J. Audio Eng. Soc. **51**, 799–805 (2003).
[18]V. I. Krylov, *Approximate Calculation of Integrals* (Macmillan, New York, 1962).

Boundary element analyses for sound transmission loss of panels

Ran Zhou[a] and Malcolm J. Crocker
Department of Mechanical Engineering, Auburn University, Auburn, Alabama 36849

(Received 12 March 2009; revised 9 November 2009; accepted 16 November 2009)

The sound transmission characteristics of an aluminum panel and two composite sandwich panels were investigated by using two boundary element analyses. The effect of air loading on the structural behavior of the panels is included in one boundary element analysis, by using a light-fluid approximation for the eigenmode series to evaluate the structural response. In the other boundary element analysis, the air loading is treated as an added mass. The effect of the modal energy loss factor on the sound transmission loss of the panels was investigated. Both boundary element analyses were used to study the sound transmission loss of symmetric sandwich panels excited by a random incidence acoustic field. A classical wave impedance analysis was also used to make sound transmission loss predictions for the two foam-filled honeycomb sandwich panels. Comparisons between predictions of sound transmission loss for the two foam-filled honeycomb sandwich panels excited by a random incidence acoustic field obtained from the wave impedance analysis, the two boundary element analyses, and experimental measurements are presented.
 [DOI: 10.1121/1.3273885]

PACS number(s): 43.55.Nd, 43.20.Rz, 43.20.Tb, 43.40.At [NX] Pages: 829–840

I. INTRODUCTION

Composite sandwich structures have been used increasingly in recent years instead of metal structures for the construction of aircraft, spacecraft, and ships because of their high stiffness-to-weight ratios. This trend is driven by demands for higher load capacity and reduced fuel consumption for cars, trucks, and aerospace structures.[1–4] Noise and vibration control in machinery and vehicles is often a matter of concern.[5–7] Once noise sources have been identified, the noise of machinery can be reduced by the use of machine enclosures or by cabin enclosures used to protect passengers in the case of aircraft or vehicles. The sound transmission loss of sandwich structure enclosure walls is mostly determined by the mass and the dynamic stiffness of the structures. Because of the presence of the core, the dynamic stiffness of sandwich structures is strongly dependent on frequency and decreases with increasing frequency. Thus, the sound transmission loss of sandwich panels can be much different from that of metal panels used with metal enclosures.

Wave impedance analysis is the most straightforward approach for the calculation of the sound transmission loss of panels. The wave impedance of a panel is derived from the governing equations for the forced vibration of the panel. Since the acoustic particle velocity must match the transverse velocity of the panel at the fluid-structure interfaces, the pressures in the incident, reflected, and radiated waves at the interfaces can be related to the wave impedance of the panel. Wave impedance analysis assumes that the panel is infinite, so that only the non-resonant forced motion is considered. Most previous work has used wave impedance analysis to predict the sound transmission loss of composite sandwich panels.[8–11]

In many cases, machines and/or ground and aerospace vehicles have machine enclosures, vehicle cabin enclosures, or other structural systems comprised of beams or panels which have geometries approximating those of rectangular shapes.[1,2,12,13] Thus the forced response and radiation of sound from metal and composite sandwich rectangular panels and the transmission of sound through such panels and its suppression are of interest in several practical problems.

The introduction of computers has permitted increasing use of numerical simulation analyses, including boundary element method (BEM) and finite element method. For a baffled planar structure, boundary element method has the advantage over finite element method of avoiding the discretization of the fluid domain. Thus it is efficient for solving exterior interactions, where the fluid occupies an unbounded domain. Three-domain, fluid-structure-fluid systems have been modeled as coupled systems[14–16] and uncoupled systems.[17–19] The uncoupled approach assumes that the structure is rigid, and that the external load on the structure is only associated with the pressures in the incident and reflected waves at the fluid-structure interface. The coupled approach considers the effect of the pressures in the radiated waves at the fluid-structure interfaces.

Mariem and Hamdi[14] presented a boundary element analysis to compute the sound transmission loss of a baffled panel. The elastic potential energy, the kinetic energy, and the work done by the displacement of the panel are described in their approach. The radiated sound pressure field is associated with a modified Green's function and their analysis uses the classical formula of Rayleigh. They calculated the sound transmission loss for a baffled clamped thin circular panel excited by normal incidence plane sound waves. Their

[a]Author to whom correspondence should be addressed. Electronic mail: zhouran@auburn.edu

numerical results predict that the radiated energy can be greater than the incident energy near to the first few resonance frequencies. However, there is an error in their expression for the total load on the panel.

Barisciano[17] studied the sound transmission loss of honeycomb sandwich panels using boundary element and finite element models. The finite elements of honeycomb sandwich panels were constructed using PATRAN. The computed velocities of the panel excited by the fluid forces were imported to a boundary element analysis software as the boundary conditions of the fluid domain. Barisciano treated the fluid-structure-fluid system as an uncoupled system and used an incorrect finite element model for the honeycomb sandwich panels.

Filippi *et al.*[15] studied the response of a thin elastic baffled rectangular panel in a light fluid excited by an incident acoustic field. They only predicted the noise reduction (difference in sound pressure levels) across the panel.

Thamburaj and Sun[18] examined the effects of material and geometrical properties on the theoretical sound transmission loss of a sandwich beam. The governing equations for the sandwich beam were developed by applying Lagrange's principle. They assumed that the external load on the beams is due to the incident and reflected pressures only.

The two composite sandwich panels in our study have plain weave fabric-reinforced graphite composites as the face sheet materials and polyurethane (PUR) foam-filled paper honeycomb as the core material. The combination of the PUR foam and honeycomb material gives the core the advantage of possessing both foam and honeycomb properties, a high shear modulus, and a large bonding area.

In this paper, first, a boundary element model is presented for a fluid-light-panel-fluid system. Approximations for the resonance eigenvalues and eigenmodes of the light-fluid-loaded panel are presented. Then a boundary element model is proposed for the study of the sound transmission loss of symmetric sandwich panels. The computed transmission loss values were obtained for an aluminum panel and the two composite sandwich panels from two boundary element analyses: (1) our BEM computer program and (2) a transmission loss (TL) model created in a boundary element analysis commercial software, LMS SYSNOISE. The effects of angle of rotation and modal energy loss factor on the sound transmission loss of the three panels are discussed. Finally, predictions of the sound transmission loss of the two foam-filled honeycomb sandwich panels from the two boundary element analyses and the wave impedance analysis are compared with experimental values.

II. BOUNDARY ELEMENT MODEL FOR A FLUID-STRUCTURE-FLUID SYSTEM

For conventional multi-domain systems, the acoustical domains are first divided into several subdomains, and the Helmholtz integral equation is applied to each subdomain. Two interface conditions, the continuity of normal particle velocity and the continuity of sound pressure, are then enforced at the interface between two neighboring subdomains. For the fluid-structure-fluid system, the structure involved is assumed to be elastic, and the sound pressure to have a step

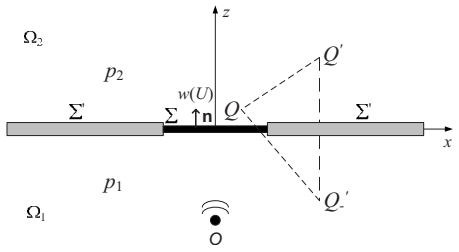


FIG. 1. Sound fields Ω_1 and Ω_2 created by a baffled planar vibrating structure.

across the structure. Then the fluid-structure-fluid interactions require slight modifications to the interface conditions. The first condition is the continuity of normal particle velocity, which is the same as for fluid-fluid systems. The second condition becomes the continuity of the normal stress, which relates the normal displacement and the stiffness and mass matrices of the structure to the pressure step across the structure.

A. Governing equation for the fluid-loaded panel

Consider a fluid-structure-fluid system, as illustrated in Fig. 1. An elastic panel occupies the domain Σ on the plane $z=0$ in a three-dimensional space. The baffle occupies the region Σ' and is perfectly rigid. The two half spaces, $\Omega_1(z<0)$ and $\Omega_2(z>0)$, contain a fluid. The system is excited by a simple harmonic sound source, $O(Q)e^{i\omega t}$, located in Ω_1 at point Q, (x,y,z).

Let $w(U)=w(x,y)$ be the panel normal displacement of node U, located at $(x,y,0)$; $p_2(Q)=p_2(x,y,z)$ and $p_1(Q)=p_1(x,y,z)$ denote the sound pressure fields in the Ω_2 and Ω_1 fields, respectively. The sound pressure step $q(Q)=q(x,y)$ across the panel is defined by

$$q(x,y)=\lim_{l\to 0}|p_1(x,y,-l)-p_2(x,y,l)|, \quad l>0. \tag{1}$$

Then the sound pressure functions $p_2(Q)$ and $p_1(Q)$ satisfy the Helmholtz equation:

$$(\nabla^2+k^2)p_2(Q)=0, \quad Q\in\Omega_2, \tag{2a}$$

$$(\nabla^2+k^2)p_1(Q)=O(Q), \quad Q\in\Omega_1. \tag{2b}$$

At the interface, the continuity of normal stress produces

$$(K-M\omega^2)w(U)=q(U)A, \quad U\in\Sigma, \tag{3}$$

where K, M, and A are the stiffness, the mass, and the coverage area of node U, respectively. The continuity of normal velocity gives

$$\frac{\partial p_2(U)}{\partial n}=-\frac{\partial p_1(U)}{\partial n}=\rho_f\omega^2 w(U), \quad U\in\Sigma, \tag{4}$$

where ρ_f is the mass density of the fluid, and $\mathbf{n}$ is the unit normal vector to the surface Σ, outward to the sound pressure field Ω_2. On the baffle plane, we have

$$\frac{\partial p_2(U)}{\partial n}=\frac{\partial p_1(U)}{\partial n}=0 \quad \text{and} \quad w(U)=0, \quad U\in\Sigma'. \tag{5}$$

The half-space Green's function is expressed as

$$G(Q,Q') = -\frac{e^{-ikr(Q,Q')}}{4\pi r(Q,Q')} - \frac{e^{-ikr(Q,Q'_-)}}{4\pi r(Q,Q'_-)}, \tag{6}$$

where the coordinates of the points Q' and Q'_- are (x',y',z') and $(x',y',-z')$, respectively.

The structure and the baffle are coplanar. Thus, the normal derivative of the modified Green's function is zero along Σ'. Then the sound pressure fields can be written as

$$p_2(Q) = \rho_f \omega^2 \int_\Sigma w(U')G(Q,U')dS(U'), \quad Q \in \Omega_2,$$

$$p_1(Q) = p_0(Q) - \rho_f \omega^2 \int_\Sigma w(U')G(Q,U')dS(U'), \tag{7}$$

$$Q \in \Omega_1, \tag{8}$$

where $p_0(Q)$ is the sound pressure generated by the source and its image in the presence of the baffle.

Then the pressure step across the panel is

$$q(U) = p_1(U) - p_2(U), \quad U \in \Sigma,$$

$$= p_0(U) - 2\rho_f \omega^2 \int_\Sigma w(U')G(U,U')dS(U'). \tag{9}$$

Equation (9) is similar to the equation derived by Mariem and Hamdi.[14] Although their expression for the half-space Green's function is identical to Eq. (6), Mariem and Hamdi, however, used the term $4\rho_f\omega^2\int_\Sigma w(U')G(U,U')dS(U')$ in their expression for the pressure step across the panel.

Hence, we have the equation for the panel displacement:

$$(K - M\omega^2)w + 2\rho_f\omega^2 \int_\Sigma \int_\Sigma w(U')G(U,U')dS(U')dS(U)$$

$$= \int_\Sigma p_0(U)dS(U), \quad U \in \Sigma. \tag{10}$$

The discretization of the panel by finite elements yields the matrix equation

$$[\mathbf{K} - (\mathbf{M} + 2\mathbf{B})\omega^2]\{w\} = \mathbf{F}, \tag{11}$$

where $\mathbf{B}$ is the complex symmetric matrix that represents the effect of the fluid on the structural behavior and depends on the wave number k, and $\mathbf{F}$ is the loading matrix acting on the planar structure.

B. Eigenvalues and eigenmodes of the fluid-loaded panel

Crighton[20] pointed out that if the intrinsic fluid loading parameter, $\bar{\varepsilon}$, is small, a singular perturbation method can be used to solve the fluid-structure-fluid problem. This parameter is given by

$$\bar{\varepsilon} = \frac{\rho_f c_f}{2\pi f_c m}, \tag{12}$$

where m and f_c are the mass per unit area and the critical frequency of the panel, and c_f is the sound speed in the fluid.

Habault[21] developed an approximation for the eigenvalues and eigenmodes of the fluid-loaded panel,

$$\lambda_n = \lambda_n^0 + \varepsilon\lambda_n^1, \quad \psi_n = \psi_n^0 + \varepsilon\psi_n^1, \quad \varepsilon = 2\rho_f/m, \tag{13}$$

where n is the order of the mode; λ_n^0 and ψ_n^0 are the zeroth order eigenvalues and eigenmodes of the panel; i.e., the eigenvalues and eigenmodes *in vacuo*; ε is the actual fluid loading parameter, and

$$\frac{\lambda_n^1}{\lambda_n^0} = \frac{\int_\Sigma \int_\Sigma \psi_n^0(U')G(U,U')\psi_n^0(U)dS(U')dS(U)}{\int_\Sigma \psi_n^0(U)\psi_n^0(U)dS(U)}, \tag{14a}$$

$$\psi_n^1(U) = -\sum_{m=1,\ldots,n-1,n+1}^{\infty} \frac{\lambda_n^0}{\lambda_m^0 - \lambda_n^0}\psi_m^0(U) \times \frac{\int_\Sigma \int_\Sigma \psi_n^0(U')G(U,U')\psi_m^0(U)dS(U')dS(U)}{\int_\Sigma \psi_m^0(U)\psi_m^0(U)dS(U)}. \tag{14b}$$

It is noted that the first order eigenvalue correction caused by the fluid loading, λ_n^1, is independent of the other eigenmodes, and the eigenvalue, λ_n, becomes complex. The absolute value, $|\lambda_n^1/\lambda_n^0|$, decreases with increasing values of n. The first order eigenmode correction, ψ_n^1, is associated with the other zeroth order eigenvalues and the couplings between the zeroth order eigenmodes. Equation (14b) shows that the contributions of those modes with zeroth order eigenvalues, which are not close to λ_n^0, are expected to be small. Furthermore, the coupling between high order eigenmodes (which have large values of n) is also expected to be small.

C. Governing matrix equation for the fluid-loaded panel

The normal displacement frequency response can be expressed as a linear combination of fluid-loaded panel eigenmode responses,[21]

$$\{w\} = \sum_j a_j\{\psi_j\} = [\psi]\{a\}, \tag{15}$$

where $\{a\}$ contains the modal participation factors and $[\psi]$ is a matrix, whose columns are the fluid-loaded panel eigenmodes. Then Eq. (11) can be expressed as

$$([\psi]^T[\mathbf{K} - (\mathbf{M} + 2\mathbf{B})\omega^2][\psi])\{a\} = [\psi]^T\mathbf{F}. \tag{16}$$

When the eigenmodes are generalized with respect to mass, the following relations are obtained:

$$\{\psi_j\}^T[K]\{\psi_i\} = \begin{cases} \lambda_i, & i = j \\ 0, & i \neq j, \end{cases}$$

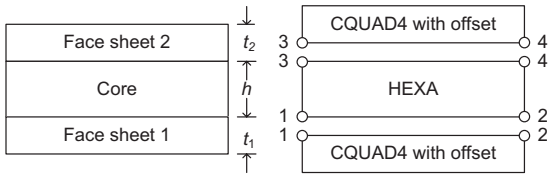


FIG. 2. Finite element model for sandwich structures using MSC NASTRAN.

$$\{\psi_j\}^T[M+2B]\{\psi_i\} = \begin{cases} 1-\varepsilon\beta, & i=j \\ 0, & i\neq j, \end{cases} \tag{17a}$$

with

$$\beta = \frac{\int_\Sigma \int_\Sigma \psi_n(U')G(U,U')\psi_n(U)dS(U')dS(U)}{\int_\Sigma \psi_n(U)\psi_n(U)dS(U)}. \tag{17b}$$

Hence Eq. (16) is associated only with the eigenvalues λ_i and the generalized eigenmodes. The modal energy loss factor can be accounted for by including a fraction array $\{\eta_n\}$,

$$\lambda_n^* = \lambda_n(1+i\eta_n). \tag{18}$$

In many papers on the sound transmission loss of panels or beams,[14,17,18] the mass density of air is considered to be small compared with the panel surface mass density. Thus the eigenvalues and eigenmodes of the fluid-loaded panels, λ_n and ψ_n, are assumed to be the eigenvalues and eigenmodes *in vacuo*, λ_n^0 and ψ_n^0.

In some commercial softwares, the fluid loading matrix, **B** in Eq. (11), is treated as an added mass matrix, which is evaluated at each frequency. Thus the governing matrix equation of the panel displacement is expressed as[22]

$$([\psi^0]^T[\mathbf{K}-(\mathbf{M}+2\mathbf{M}_a(\omega))\omega^2][\psi^0])\{a\} = [\psi^0]^T\mathbf{F}, \tag{19}$$

where $\mathbf{M}_a$ is the added mass matrix.

D. Symmetric sandwich panels

An analytical expression for the generalized eigenmodes of the normal displacements of sandwich panels is not available. The generalized modal eigenmodes must be obtained from finite element analysis.

Johnson and Kienholz[23] presented a finite element model for sandwich structures using MSC NASTRAN. In this model, the face sheets are modeled with plate elements, such as CQUAD and CTRIA, with two rotational and three translational degrees of freedom per node. The core is modeled with solid elements, such as HEXA and PENTA, with three translational degrees of freedom per node. The plate elements are offset to one surface of the plate and are coincident with the nodes of the adjoining solid elements, as shown in Fig. 2.

For a three-layer sandwich panel, the generalized eigenmodes of the normal displacements of the two face sheets can be written as

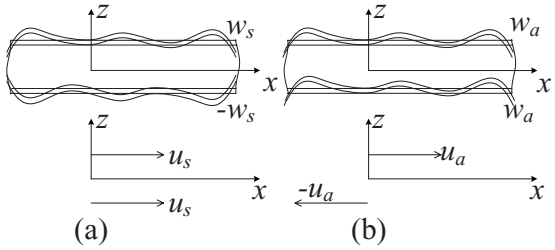


FIG. 3. Face sheet displacements: (a) symmetric motion and (b) antisymmetric motion.

$$\psi_{mn}(x,y,z) = \begin{cases} V_{mn,2}(x,y), & z>0 \\ V_{mn,1}(x,y), & z<0. \end{cases} \tag{20}$$

For sandwich panels, which have identical face sheets and whose vibration is dominated by antisymmetric motion in the frequency range of interest, the normal displacements of the face sheets are the same, $V_{mn,2}=V_{mn,1}$ [see Fig. 3(b)]. Hence these symmetric three-layer sandwich panels can be treated as equivalent single-layer panels whose generalized eigenmodes satisfy

$$\psi_{mn}(x,y) = V_{mn} = V_{mn,2} = V_{mn,1}. \tag{21}$$

In Sec. II C, it was shown that Eq. (16) is only associated with the eigenvalues and the generalized eigenmodes of the fluid-loaded structure. The eigenvalues and the generalized eigenmodes of the fluid-loaded structure can be evaluated approximately by using Eqs. (13) and (14). Then only the generalized eigenmodes for one face sheet of the sandwich panel *in vacuo* are required for this boundary element model.

III. SOUND TRANSMISSION LOSS MODEL

The sound pressure in the incident plane wave generated by the sound source O (see Fig. 1) can be written as

$$p_{\text{inc}} = P_{\text{inc}}\ \exp\{i\omega t - ik(x\sin\theta\cos\phi + y\sin\theta\sin\phi + z\cos\theta)\}, \tag{22}$$

where P_{inc} is the amplitude of the source, θ is the angle of incidence measured from the normal, and ϕ is the angle of rotation. The sound transmission coefficient is defined as

$$\tau = \frac{W_r}{W_i}, \tag{23}$$

where W_i denotes the virtual sound power flow, which would pass through the surface of the panel, if the panel were removed,

$$W_i = \frac{P_{\text{inc}}^2 A_p \cos\theta}{2\rho_f c_f}, \tag{24}$$

where A_p is the area of the panel. W_r is the sound power radiated by the panel in the semi-infinite fluid domain. Normally, it is determined indirectly. A field point mesh, a hemisphere, or some other geometric shape that encloses the receiving side of the structure is defined. After the sound pressure, $p_2(Q)$, obtained from Eq. (7), and the particle velocity are computed at the field points, then the sound power

TABLE I. Eigenfrequencies of a 3.175 mm (0.125 in.) thick 0.84×0.428 m^2 aluminum panel *in vacuo*.

(m,n)	(1, 1)	(2, 1)	(3, 1)	(1, 2)	(2, 2)	(4, 1)	(3, 2)	(5, 1)	(4, 2)	(1, 3)	(2, 3)	(6, 1)	(5, 2)	(3, 3)	(4, 3)	(6, 2)
f (Hz)	52.9	85.5	140	180	212	217	266	317	343	393	425	440	442	479	554	564

radiated is determined by integrating the sound intensity over the area of the field point mesh.

In this work, the sound power radiated, W_r, is determined directly from the panel normal displacement, $w(U)$, and the sound pressure in the receiving side, which is applied on the structure, $p_r(U)$.

$$W_r = \frac{1}{2}\mathrm{Re}\left(-\int_\Sigma i\omega w^*(U)p_r(U)dS(U)\right), \quad U \in \Sigma$$
$$= \frac{1}{2}\mathrm{Re}\left(\int_\Sigma v_n^*(U)p_r(U)dS(U)\right), \tag{25}$$

where $p_r(U)=\omega^2\rho_f\int_\Sigma w(U')G(U,U')dS(U')$ is a part of the equation for the panel normal displacement, Eq. (10). Thus, the sound power radiated, W_r, can be computed after the matrix equation [Eq. (11)] is solved.

Finally, the panel transmission loss is defined by

$$\mathrm{TL} = 10\log_{10}\left(\frac{1}{\tau}\right)\mathrm{dB}. \tag{26}$$

In practice, sound waves are usually incident upon a structure from many angles simultaneously. Therefore, an idealized random incidence model is usually assumed, in which plane waves are incident from all directions with equal probability and which have random phases. The random incidence transmission coefficient, $\bar{\tau}$, is obtained by averaging $\tau(\theta,\phi)$ over all angles of incidence and rotation as follows:

$$\bar{\tau} = \frac{\int_0^{2\pi}\int_0^{\theta_{\lim}}\tau(\theta,\phi)\sin\theta\cos\theta d\theta d\phi}{\int_0^{2\pi}\int_0^{\theta_{\lim}}\sin\theta\cos\theta d\theta d\phi}. \tag{27}$$

Based on field and laboratory measurements, the limiting angle $\theta_{\lim}$ is usually assumed to be equal to 78°.[24]

IV. NUMERICAL RESULTS FOR ALUMINUM PANEL

In this section, the sound transmission characteristics of an aluminum panel are studied. A BEM computer program using MATLAB language, which uses an approximation for the eigenvalues and eigenmodes of the fluid-loaded panel [Eqs. (13) and (14)], was developed to calculate the sound transmission loss of the panel. The predicted sound transmission values are compared with results obtained from a TL model created in a commercial software, SYSNOISE.

The generalized eigenmodes of the normal displacements for a simply supported single-layer panel *in vacuo* can be expressed as

$$\psi_{mn} = V_{mn}\sin\left(\frac{m\pi}{l_x}x\right)\sin\left(\frac{n\pi}{l_y}y\right), \tag{28}$$

where l_x and l_y are the dimensions of the panel, and V_{mn} is the amplitude.

The generalized eigenmodes of a 3.175 mm (0.125 in.) thick aluminum panel *in vacuo* were computed using a finite element analysis commercial software, MSC NASTRAN. The 0.84×0.428 m^2 panel was divided into 40×24 CQUAD4 elements. The eigenfrequencies of the aluminum panel *in vacuo* are given in Table I. The first 16 modes were included in the BEM computer program. The energy loss factor was assumed to be constant for all modes. The frequency increment used was 1 Hz. It was found that the eigenfrequencies of the fluid-loaded aluminum panel are slightly smaller than those of the same aluminum panel *in vacuo*. The differences between these frequencies are 4%, 2%, 1%, and 0.5% for modes, (1, 1), (3, 1), (5, 1), and (1, 3), respectively.

The predicted sound transmission loss values of the aluminum panel for plane sound waves at normal incidence are given in Fig. 4. The sound transmission loss prediction at the first resonance frequency is less sensitive to the value of the modal energy loss factor. It was found that the sound transmission loss predictions for a frequency of 50 Hz, which is near to the first resonance frequency, for energy loss factors η=0.01, 0.001, and 0.0001 are −2.0, −5.1, and −5.4 dB, respectively. The differences between these predictions are much smaller than the corresponding differences for the sound transmission loss predictions at higher order resonance frequencies.

It is seen that the radiated power is greater than the virtual power flow at the first resonance frequency. The spe-

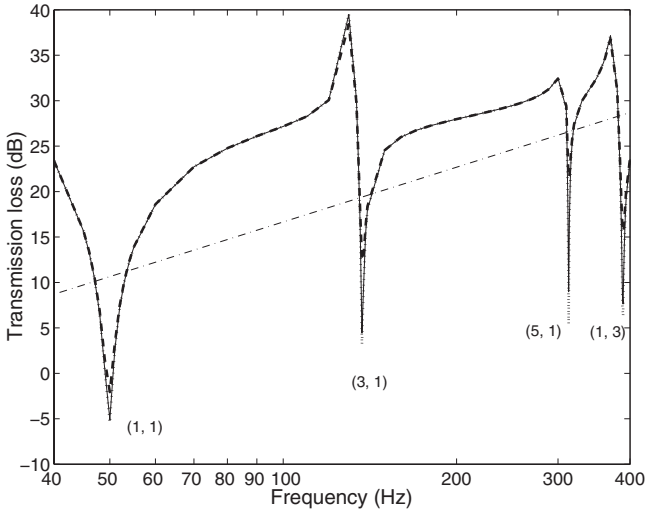


FIG. 4. Predicted sound transmission loss values of a 3.175 mm (0.125 in.) thick aluminum panel made using the BEM computer program for plane sound waves at normal incidence: (– – –) η=0.01; (——) η=0.001; (···) η=0.0001; (–·–·–) the mass law for normal incidence.

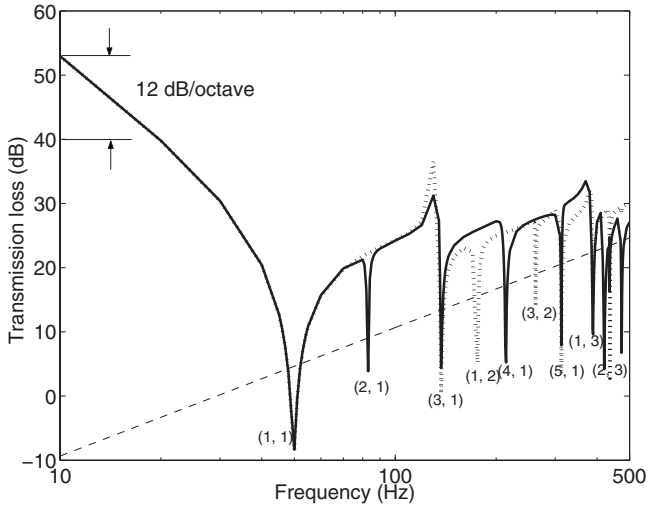


FIG. 5. Predicted sound transmission loss values of 3.175 mm (0.125 in.) thick aluminum panel made using the BEM computer program for plane sound waves at oblique incidence, $\theta=60°$, with $\eta=0.0001$: (——) traveling in the x-axis direction, $\phi=0°$; (···) traveling in the y-axis direction, $\phi=90°$; (– – –) the mass law for oblique incidence.

cific acoustic impedance of the half-space Green's function, Eq. (6), is identical to that for a spherical wave,

$$Z=\frac{i\rho_f c_f kr}{1+ikr}=\rho_f c_f\left(\frac{k^2r^2}{1+k^2r^2}+i\frac{kr}{1+k^2r^2}\right). \tag{29}$$

Unlike the specific acoustic impedance of a plane sound wave, the specific acoustic impedance of a spherical sound wave has both resistive and reactive components. When $kr \gg 1$, the specific acoustic impedance approaches $\rho_f c_f$ the specific acoustic impedance of a plane wave. For boundary element analysis, both the resistive and reactive components are required. The virtual sound power flow, Eq. (24), is defined by using plane wave concepts, and the radiated power, Eq. (25), is associated with the specific acoustic impedance of spherical waves. That may explain why the radiated sound power is greater than the virtual sound power flow at low order resonance frequencies.

When a plane wave is at normal incidence on the panel, the sound pressure generated by the source and its image on the panel, $p_0(U)$, is uniform [see Eq. (22)]. Only odd-odd (volume displacing) panel modes are excited, so only odd-odd modes radiate sound power. When a plane wave is obliquely incident on the panel and is traveling in the x-axis direction, $\phi=0°$, $p_0(U)$ is constant in the y-axis direction. Then the radiated power for odd-even or even-even modes of the panel is quite small. Hence the resonance dips in the predicted transmission loss for incident waves traveling in the x-axis direction ($\phi=0°$) occur at the resonance frequencies whose y-axis mode numbers are odd. Likewise, the resonance dips in the predicted transmission loss for incident waves traveling in the y-axis direction ($\phi=90°$) occur at the resonance frequencies whose x-axis mode numbers are odd. The sound transmission loss values of the aluminum panel for sound waves at oblique incidence, $\theta=60°$, predicted using the BEM computer program are presented in Fig. 5.

At very low frequencies, much below the first resonance frequency, the normal displacement $w(U)$ approaches $p_0(U)A_p/K$, and the half-space Green's function can be expanded in a Taylor series as

$$\begin{aligned} G(U,U') &= -\left\{\frac{1}{2\pi r(U,U')}-i\frac{\omega}{2\pi c_f}-\frac{r(U,U')}{2\pi}\frac{\omega^2}{c_f^2}\right. \\ &\quad \left.+i\frac{r^2(U,U')}{2\pi}\frac{\omega^3}{c_f^3}+\cdots\right\} \\ &\approx -\frac{1}{2\pi r(U,U')}+i\frac{\omega}{2\pi c_f} \\ &\quad \text{for } \omega \ll \omega_{11}, \quad U,U' \in \Sigma. \end{aligned} \tag{30}$$

Then the radiated sound power W_r is proportional to ω^4. This explains why the sound transmission loss of the aluminum panel decreases by 12 dB, not 6 dB, when the frequency is doubled at very low frequency.

The boundary element analysis commercial software, LMS SYSNOISE, uses Eq. (19) to obtain the normal panel displacement. A transmission loss model was built in the commercial software to obtain the sound transmission loss of the aluminum panel. A finite element model database and a boundary element indirect baffled model database were also defined. The element mesh, 40×24 CQUAD4 elements, was imported to the commercial software as the structural meshes for both databases to occupy the plane $z=0$. The computed eigenmodes of the first 16 modes obtained from MSC NASTRAN were also imported to the commercial software as the structural modes in the finite element model database. A plane wave source was defined in the boundary element model database at 5 m below the center of the panel. The two databases were linked to solve the normal displacement of the aluminum panel. LMS SYSNOISE was used to calculate the radiated power of the planar structures indirectly. Then a hemisphere with a radius of 1 m was used for the field point mesh to estimate the sound power radiated.

Comparisons of the sound transmission losses of the aluminum panel for plane sound waves at oblique incidence calculated from our BEM computer program and the TL model using the boundary element analysis commercial software are given in Figs. 6 and 7. In both cases, the energy loss factor was assumed to be constant for all modes; the frequency increment used was 1 Hz. Figures 6 and 7 give plots of transmission loss against frequency on a linear scale so that the resonance dips can be seen more clearly.

It is seen that the contributions of even-even modes can have a substantial effect to the sound transmission loss of the aluminum panel for plane sound waves at oblique incidence, such as for angles $\theta=60°$ and $\phi=45°$ (see Fig. 6). The predictions obtained from our BEM computer program and from the TL model using the boundary element analysis commercial software are similar. For this panel, the TL model using the commercial software provides a good approximation for the resonance frequencies. However, it underestimates the transmission loss near some panel resonance frequencies for very small values of the energy loss factor, such as $\eta=0.0001$. It is noted that this small value of the energy loss factor is not at all appropriate for practical alu-

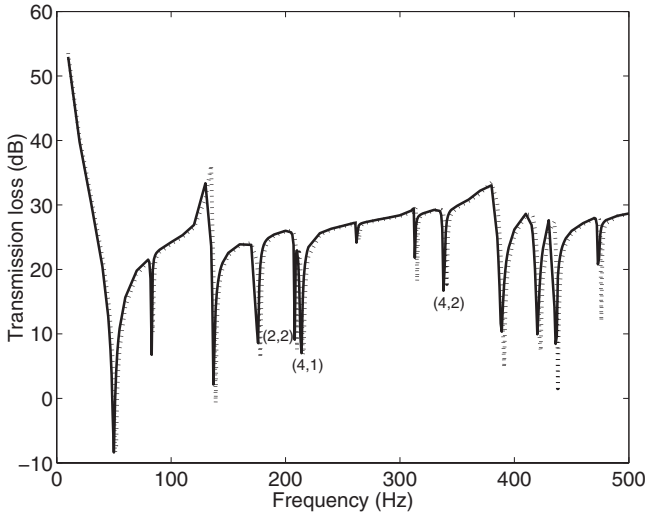


FIG. 6. Predicted sound transmission loss values of a 3.175 mm (0.125 in.) thick aluminum panel with η=0.0001 for plane sound waves at oblique incidence, θ=60° and ϕ=45°: (—) our BEM computer program; (···) the TL model with the boundary element analysis commercial software.

minum structures. Energy loss factors of aluminum panels in practice are known to be dominated by the energy losses at the panel edge constraints. These loss factors are dependent on panel size, with smaller panels of the same boundary conditions generally having larger loss factors (larger perimeter to area ratios). It was found that the difference between the predictions obtained from the two boundary element analyses is negligible when a practical value of the energy loss factor of this aluminum panel, η=0.01, was used in the sound transmission loss calculations (see Fig. 7).

V. SOUND TRANSMISSION LOSS OF COMPOSITE SANDWICH PANELS

Unlike single-layer panels, three-layer sandwich panels undergo not only antisymmetric motion but symmetric motion as well. The boundary element model presented in this paper can be applied to the sandwich panels which have identical normal displacements for both face sheets in the frequency range of interest. To determine the contribution of the antisymmetric and symmetric face sheet displacements to sound transmission loss of composite sandwich panels, a wave impedance analysis of sandwich panels is reviewed in this section. The sound transmission loss values of two foam-filled honeycomb sandwich panels from the wave impedance analysis are compared with experimental results. The sound transmission characteristics of the two composite sandwich panel were investigated by use of the two boundary element analyses. The contribution of the effect of air loading to the sound transmission loss of the two composite sandwich panels is also discussed.

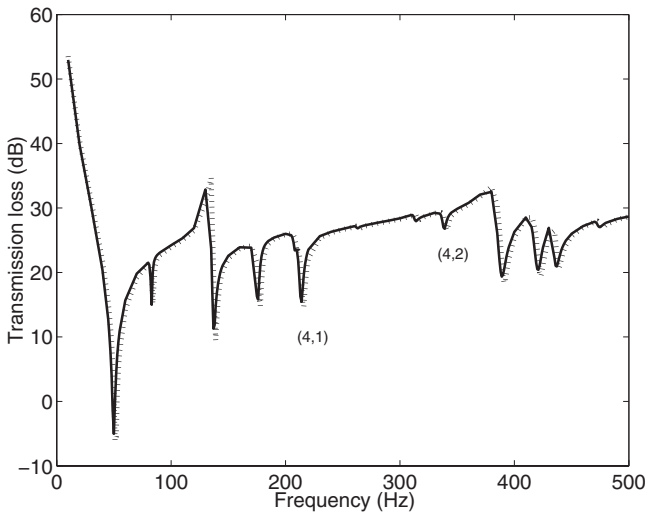


FIG. 7. Predicted sound transmission loss values of a 3.175 mm (0.125 in.) thick aluminum panel with η=0.01 for plane sound waves at oblique incidence, θ=60° and ϕ=45°: (—) our BEM computer program; (···) the TL model with the boundary element analysis commercial software.

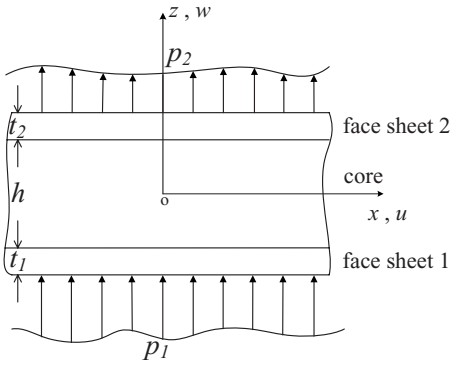


FIG. 8. The geometry and loads of a sandwich panel

A. Wave impedance analysis for symmetric sandwich panels

For symmetric sandwich panels, the symmetric and antisymmetric motions are naturally uncoupled.[8] The governing matrix equations for symmetric and antisymmetric motions of sandwich panels with orthotropic cores can be written as follows:[11]

$$\begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{12} & D_{22} & D_{23} \\ D_{13} & D_{23} & D_{33} \end{bmatrix} \begin{Bmatrix} w_s \\ u_s \\ \zeta \end{Bmatrix} = \begin{Bmatrix} -(p_1+p_2) \\ 0 \\ 0 \end{Bmatrix}, \tag{31a}$$

$$\begin{bmatrix} D_{44} & D_{45} \\ D_{45} & D_{55} \end{bmatrix} \begin{Bmatrix} w_a \\ u_a \end{Bmatrix} = \begin{Bmatrix} p_1-p_2 \\ 0 \end{Bmatrix}, \tag{31b}$$

where u and w are the in-plane and transverse displacements of the face sheet, and p_j is the pressure on the face sheet j, as shown in Fig. 8. The subscripts s and a denote the displacements caused by symmetric and antisymmetric motions, respectively, as shown in Fig. 3. The variable ζ is used to describe the translational motion of the core. The value of D_{ij} can be evaluated with the elastic stiffness constants of the face sheets and core, and geometrical properties of the panels, as shown by Moore and Lyon.[11]

When the three axes of the coordinate system are not completely aligned with the three principal axes of the orthotropic material, the stiffness in the rotated axis system becomes $\overline{C_{ij}}$ instead of C_{ij},

$$\overline{C_{11}} = l^4 C_{11} + 2l^2 m^2 (C_{12} + 2C_{66}) + m^4 C_{22}, \tag{32a}$$

$$\overline{C_{13}} = l^2 C_{13} + m^2 C_{23}, \tag{32b}$$

$$\overline{C_{33}} = C_{33}, \tag{32c}$$

$$\overline{C_{55}} = m^2 C_{44} + l^2 C_{55}, \tag{32d}$$

where C_{ij} is the elastic stiffness constant of the orthotropic material; ϕ denotes the angle of rotation, $l=\cos\phi$ and $m=\sin\phi$. Moore and Lyon[11] neglected the contribution of C_{66} to the rotated axis stiffness $\overline{C_{11}}$.

Damping is incorporated by allowing the stiffness constants in the material to become complex,

$$C_{ij}^{*} = C_{ij}(1 + i\eta), \tag{33}$$

where η is the energy loss factor of the material. The symmetric and antisymmetric wave impedances, z_s and z_a, for transverse panel motion are

$$z_s = \frac{-(p_1 + p_2)}{i\omega w_s} = \frac{\begin{vmatrix} D_{11} & D_{12} & D_{13} \\ D_{12} & D_{22} & D_{23} \\ D_{13} & D_{23} & D_{33} \end{vmatrix}}{i\omega \begin{vmatrix} D_{22} & D_{23} \\ D_{23} & D_{33} \end{vmatrix}}, \tag{34}$$

$$z_a = \frac{(p_1 - p_2)}{i\omega w_a} = \frac{1}{i\omega}\left(D_{44} - \frac{D_{45}^2}{D_{55}}\right). \tag{35}$$

The transmission coefficient due to the antisymmetric and symmetric motions is

$$\tau(\theta,\phi) = \left| \frac{1}{1 + z_s/(2z_f')} - \frac{1}{1 + z_a/(2z_f')} \right|^2 \quad \text{where } z_f' = \rho_f c_f/\cos\theta. \tag{36}$$

If, for particular conditions, $z_s \gg z_a$, the antisymmetric motion is dominant in the frequency band of interest, and the transmission coefficient can be approximated by

$$\tau(\theta,\phi) \approx \tau_a(\theta,\phi) = \left| \frac{1}{1 + z_a/(2z_f')} \right|^2. \tag{37}$$

When only the antisymmetric motion is considered, the transverse displacement of face sheet 1 is the same as that of face sheet 2 [see Fig. 3(b)].

B. Experimental sound transmission loss values of two foam-filled honeycomb sandwich panels

Two composite sandwich panels were fabricated for sound transmission loss measurements in the Sound and Vibration Laboratory at Auburn University. The transmission suite consists of two adjacent 51.2 m^3 reverberation rooms. Each room has two walls made of wood with fiberglass filled in between them. The two rooms are mounted on airbags and they are also separated from each other by fiberglass. The panels were clamped in a frame between the two rooms. The panel edge conditions were intended to be fully fixed. The frame reduced the test dimensions of the panels to 0.84 $\times$0.42 m^2. The sound transmission loss of the two panels was measured according to the standard test method, ASTM E90-99, and determined experimentally by using the conventional two-room method.

TABLE II. Properties of composite sandwich panels B and C.

	Property	Panel B	Panel C
Face sheet	Density (kg/m^3)	1600	1600
	$C_{11}=C_{22}$ (GPa)	50	40
	C_{12} (GPa)	8	7
	C_{66} (GPa)	10	10
	Thickness t (mm)	0.5	0.5
	Energy loss factor η_f	0.01	0.01
Core	Density (kg/m^3)	160	120
	Thickness h (mm)	6.35	12.7
	$C_{11}=C_{22}$ (MPa)	80	60
	C_{33} (MPa)	300	220
	$C_{12}=C_{13}=C_{23}$ (MPa)	14	10
	C_{44} (MPa)	140	60
	C_{55} (MPa)	90	100
	C_{66} (MPa)	36	30
	Energy loss factor η_c	0.01	0.01

The assumed properties of composite sandwich panels B and C are given in Table II. Since the sound transmission loss measurements were conducted in reverberation rooms, for comparison purposes the sound transmission loss was evaluated by averaging $\tau(\theta,\phi)$ over the angle of incidence θ, and the angle of rotation, ϕ, see Eq. (27). The stiffness of the orthotropic material, for both the face sheets and cores, $\overline{C_{ij}}$, was evaluated in the calculation of the sound transmission coefficient, $\tau(\theta,\phi)$, by using Eq. (32). Because the measured transmission loss results for the two composite sandwich panels are presented in one-third octave bands, the predicted transmission loss values were also frequency-averaged.

The predictions of the sound transmission loss for the composite sandwich panels B and C (1) due to both antisymmetric and symmetric motions, using Eq. (36), and (2) only antisymmetric motion, using Eq. (37), were calculated and are shown in Figs. 9 and 10. The first resonance frequencies of panels B and C are in the one-third octave bands with center frequencies of 250 and 315 Hz, respectively. The experimental transmission loss results for composite sandwich panel B are similar to the field incidence mass law curve at frequencies above its first resonance frequency and below its critical coincidence frequency, while those of composite sandwich panel C are significantly lower than the field incidence mass law curve in the corresponding frequency region. At low frequencies, the sound transmission loss curve measurements of composite sandwich panel C show that it is strongly influenced by its resonant behavior. At very low frequencies, below their first resonance frequencies, the experimental results show that panel behavior is dominated by stiffness effects. The coincidence dips of the two composite sandwich panels are not as apparent as those of metal panels. Neither of the coincidence dips returns to the mass law curves and the experimental transmission loss values depart from the mass law curves by about 15 dB.

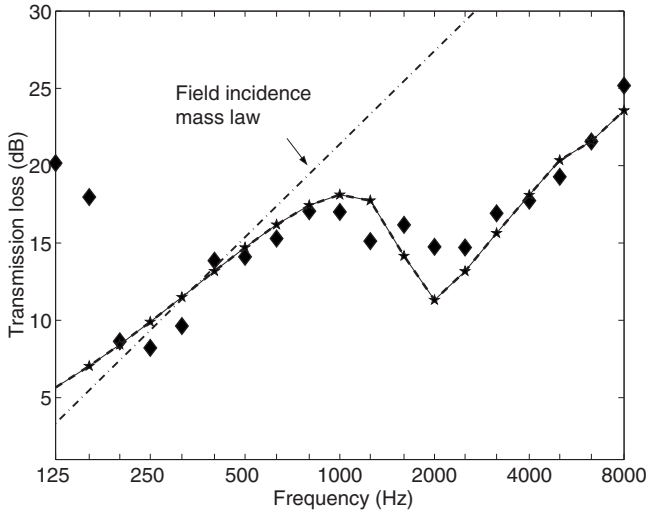


FIG. 9. Predicted sound transmission loss values of composite sandwich panel B made using the wave impedance analysis: (–★–) only caused by antisymmetric motion; (– – –) caused by both motions; (♦) experimental values.

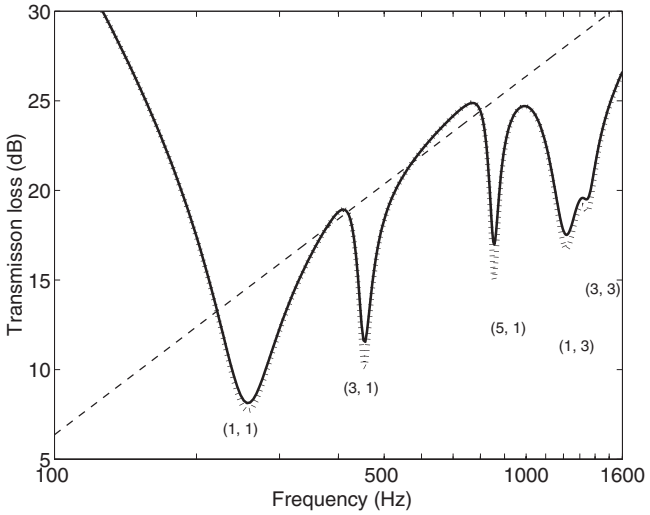


FIG. 11. Predicted sound transmission loss values of composite sandwich panel B made using our BEM computer program for plane sound waves at normal incidence: (—) η=0.01; (···) η=0.0001; (– – –) the mass law for normal incidence.

It is seen that the transmission losses of composite sandwich panels B and C are primarily caused by the antisymmetric motions in the frequency band of interest. The predicted critical coincidence frequencies for composite sandwich panels B and C associated with the antisymmetric motion are in the one-third octave bands with center frequencies of 2000 and 1250 Hz, respectively. It was found that the difference between the predicted transmission loss values (1) due to both motions and (2) only antisymmetric motion is less than 0.02 dB up to 8 kHz. The predictions of the sound transmission loss for composite sandwich panel B are in good agreement with the experimental values for frequencies except below its first resonance frequency and near coincidence. The wave impedance analysis provides a reasonable prediction for the sound transmission loss of composite sandwich panel C well above its critical coincidence frequency.

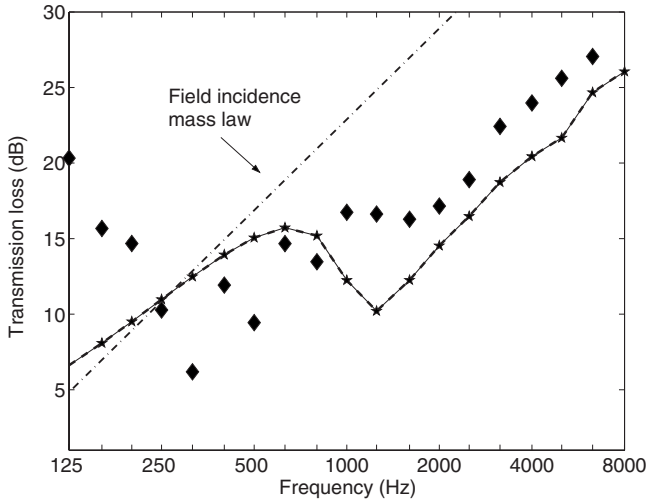


FIG. 10. Predicted sound transmission loss values of composite sandwich panel C made using the wave impedance analysis: (–★–) only caused by antisymmetric motion; (– – –) caused by both motions; (♦) experimental values.

C. Predicted sound transmission loss values of composite sandwich panels B and C

The sound transmission characteristics of the two composite sandwich panels were also computed using a BEM computer program which incorporates the boundary element model for symmetric sandwich panels proposed in Sec. II D. The finite element mesh in the finite element analysis for the sandwich panel was generated in MSC NASTRAN, and consisted of 60×36 CQUAD4 elements for each face sheet and 60×36 HEXA elements for the core. Hence, the finite element mesh of the equivalent panel used in the BEM computer program consists of 60×36 CQUAD4 elements. The first 20 modes were included in the BEM computer program. The energy loss factor was assumed to be constant for all modes. The frequency increment used was 4 Hz. It was found that the differences between the eigenfrequencies of the fluid-loaded sandwich panels and those of the panels *in vacuo* are 11%, 7%, and 3% for panel B, and 6%, 6%, and 1% for panel C, for modes, (1, 1), (3, 1), and (1, 3), respectively.

The predicted transmission loss values of composite sandwich panels B and C for plane sound waves at normal incidence, θ=0°, are shown in Figs. 11 and 12. All of the modal energy loss factors for the composite sandwich panels were assumed to be independent of frequency. The calculated eigenfrequencies *in vacuo* obtained from the finite element analysis are given in Table III. It is seen that only odd-odd modes radiate sound power. Since modal overlap exists, the contribution of some odd-odd modes to the sound transmission loss predictions is negligible, such as mode (3, 3) for panel B and mode (5, 1) for panel C. Because of the larger value of ε, then the sound transmission loss predictions for the two sandwich panels are much less sensitive to the modal energy loss factor than are those of the aluminum panel. The difference between the sound transmission loss predictions for panel C with η=0.01 and 0.0001 at the resonance frequencies is less than that for panel B. The sound

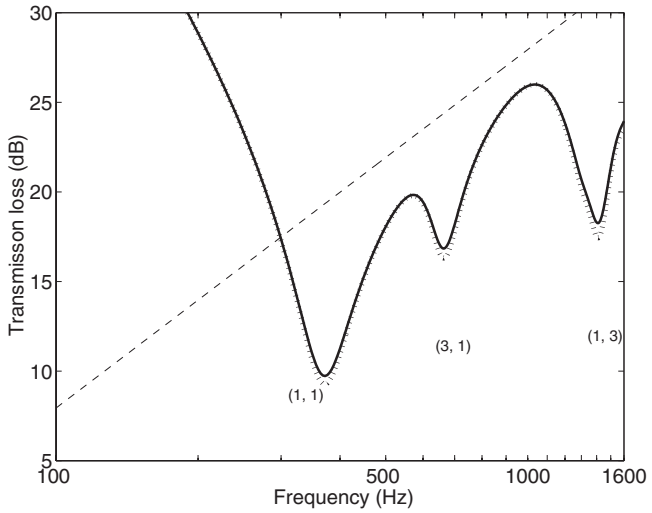


FIG. 12. Predicted sound transmission loss values of composite sandwich panel C made using our BEM computer program for plane sound waves at normal incidence: (—) η=0.01; (···) η=0.0001; (– – –) the mass law for normal incidence.

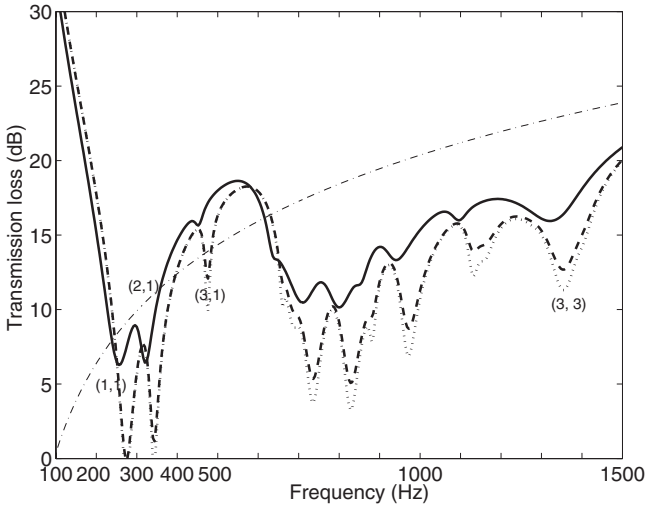


FIG. 13. Predicted sound transmission loss values of composite sandwich panel C made using the two boundary element analyses for plane sound waves at oblique incidence, θ=60° and ϕ=45°: (—) our BEM computer program, η=0.01; (– – –) the TL model using the commercial software, η=0.01; (···) the TL model using the commercial software, η=0.0001; (–·–·–) the mass law for oblique incidence.

transmission loss predictions for panel C are much smaller than the mass law values for normal incidence.

To investigate the contribution of the effect of air loading to the sound transmission loss of the foam-filled honeycomb sandwich panels, a corresponding TL model was built with the commercial software. A finite element model database and a boundary element indirect baffled model database were created in the boundary element analysis commercial software, LMS SYSNOISE. The element mesh for one face sheet was imported into the commercial software as the structural meshes for both databases and used to occupy the plane z=0. The computed eigenmodes of the first 20 modes of the face sheet were also imported into the commercial software as the structural modes in the finite element model database. A plane wave source was defined in the boundary element model database at 5 m below the center of the panel. The two databases were linked to solve the displacement of the sandwich panel. A 1 m radius hemisphere field point mesh that was assumed to cover the receiver side of the panel was used to obtain the radiated sound power of the panel. The frequency increment used was 4 Hz.

The predicted transmission loss values of the two composite sandwich panels for plane sound waves at oblique incidence, θ=60°, and ϕ=45°, are shown in Figs. 13 and 14. Figures 13 and 14 give plots of the transmission loss against frequency on a linear scale so that the resonance dips can be seen more clearly. Unlike the aluminum panel, the difference between the sound transmission loss predictions obtained from (1) our BEM computer program and (2) the TL model using the commercial software is apparent at each resonance frequency. For the two sandwich panels, the TL model using the commercial software considerably underestimates the

TABLE III. Eigenfrequencies of composite sandwich panels B and C *in vacuo* from Johnson and Kienholz's (Ref. 23) finite element model.

Panel B								
Clamped-clamped	(m,n)	(1,1)	(2,1)	(3,1)	(4,1)	(1,2)	(2,2)	(3,2)
	f (Hz)	289	360	494	681	706	757	853
	(m,n)	(5,1)	(4,2)	(6,1)	(5,2)	(1,3)	(2,3)	(3,3)
	f (Hz)	909	998	1166	1187	1259	1301	1377
Simply-supported	(m,n)	(1,1)	(2,1)	(3,1)	(1,2)	(2,2)	(4,1)	(3,2)
	f (Hz)	145	220	360	501	554	554	657
Panel C								
Clamped-clamped	(m,n)	(1,1)	(2,1)	(3,1)	(1,2)	(2,2)	(4,1)	(3,2)
	f (Hz)	395	508	711	864	947	986	1099
	(m,n)	(5,1)	(4,2)	(1,3)	(2,3)	(5,2)	(3,3)	(6,1)
	f (Hz)	1311	1320	1426	1495	1599	1616	1671
Simply-supported	(m,n)	(1,1)	(2,1)	(3,1)	(1,2)	(2,2)	(4,1)	(3,2)
	f (Hz)	222	338	550	695	775	837	929

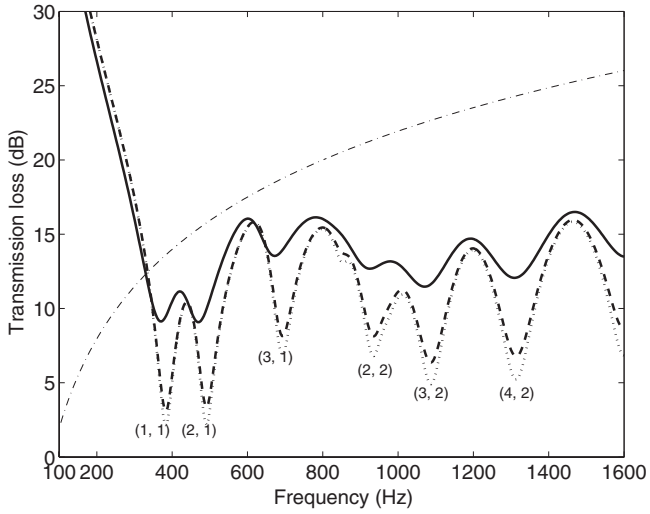


FIG. 14. Predicted sound transmission loss values of composite sandwich panel C made using the two boundary element analyses for plane sound waves at oblique incidence, θ=60° and ϕ=45°: (—) our BEM computer program, η=0.01; (– – –) the TL model using the commercial software, η =0.01; (···) the TL model using the commercial software, η=0.0001; (–·–·–) the mass law for oblique incidence.

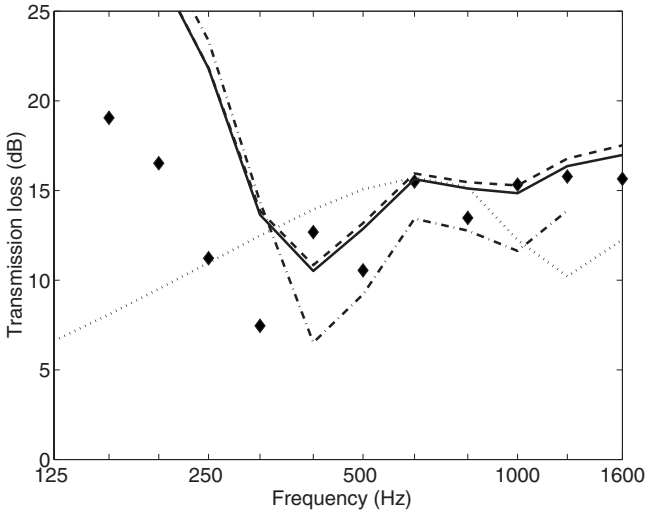


FIG. 16. Predictions of sound transmission loss for composite sandwich panel C: (– – –) our BEM computer program, η=0.01; (——) our BEM computer program, η=0.0001; (–·–·–) the TL model using the commercial software, η=0.01; (···) the wave impedance analysis; (♦) experimental values.

sound transmission loss values near the panel resonance frequencies. It is seen that the effect of air loading on the eigenfrequencies of the air-loaded panel C is smaller than on those of panel B. The increase in the modal energy loss factor from 0.0001 to 0.01 is seen, and makes less than a 3 dB difference in the predictions obtained from the TL model using the commercial software for both sandwich panels.

Figures 15 and 16 show comparisons between the sound transmission loss values for composite sandwich panels B and C measured experimentally and those calculated from our BEM computer program and the TL model using the commercial software. The predicted sound transmission loss values were evaluated by averaging $\tau(\theta,\phi)$ over the angles of incidence and rotation, and were also frequency-averaged.

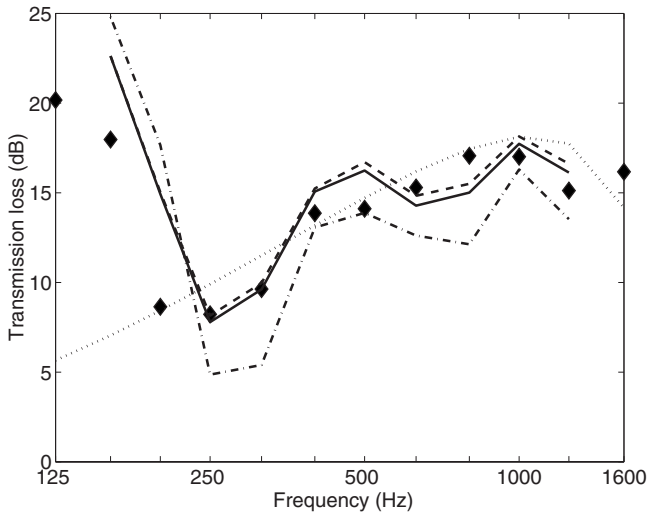


FIG. 15. Predictions of sound transmission loss for composite sandwich panel B: (– – –) our BEM computer program, η=0.01; (——) our BEM computer program, η=0.0001; (–·–·–) the TL model using the commercial software, η=0.01; (···) the wave impedance analysis; (♦) experimental values.

The predicted transmission loss values in Figs. 15 and 16 increase with increasing values of the modal energy loss factor. The increase is small, less than 1 dB. We attempted to produce clamped-clamped boundaries in our experiments. However, clamped-clamped boundary conditions are difficult to achieve in practice. The disagreement at low frequencies is thought to be caused by the ideal clamped-clamped boundary conditions assumed in the finite element analysis. The non-ideal clamped-clamped boundary conditions will shift the first several resonance frequencies to lower frequencies and modify the transmission loss values in low one-third octave bands. The predictions for composite sandwich panel B agree with the experimental values and are near to the values made using the wave impedance analysis. The predictions for composite sandwich panel C are in better agreement with the experimental values than those made using the wave impedance analysis. The discrepancy in the 800 Hz one-third octave band may be caused by frequency-dependent material properties.

VI. CONCLUSIONS

For the type of composite sandwich panels with graphite fiber face sheets and foam-filled honeycomb cores studied in this work, the sound transmission loss is primarily caused by the antisymmetric motion of the panels in the frequency band of interest. The wave impedance analysis provides accurate predictions of the sound transmission loss for the thin composite sandwich panel B at frequencies above its first resonance frequency.

For finite single-layer isotropic rectangular panels, the sound transmission loss depends not only on the angle of incidence θ but also on the angle of rotation ϕ. The contributions of other than odd-odd modes to the transmission loss of single-layer panels were found to be negligible for plane sound waves at normal incidence, although these can be substantial for plane sound waves at oblique incidence. The TL

model used in the commercial boundary element analysis software provides accurate estimates for the sound transmission loss of isotropic panels when the value of the fluid loading parameter ε is small.

The boundary element model for the sound transmission loss of symmetric sandwich panels proposed herein requires a minimum increase in computation effort relative to the effort needed for single-layer panels. The main reason for this fact is that the eigenmodes of only one of the face sheets are required in the boundary element model for symmetric sandwich panels. It was found that the effect of changes in the modal energy loss factor on the sound transmission loss predictions is negligible for the two sandwich panels studied. The predictions from our BEM computer program agree well with experimental transmission loss values of the two composite sandwich panels B and C above the first panel resonance frequency of each panel.

[1]Z. Li and M. J. Crocker, "A review of vibration damping in sandwich composite structures," Int. J. Acoust. Vib. **10**, 159–169 (2005).

[2]K. H. Hornig and G. T. Flowers, "Performance of heuristic optimisation methods in the characterisation of the dynamic properties of sandwich composite materials," Int. J. Acoust. Vib. **12**, 60–68 (2007).

[3]T. Chavan and D. N. Manik, "Optimum design of vibro-acoustic systems using SEA," Int. J. Acoust. Vib. **13**, 67–81 (2008).

[4]Z. Li and M. J. Crocker, "Effects of thickness and delamination on the damping in honeycomb-foam sandwich beams," J. Sound Vib. **294**, 473–485 (2006).

[5]B. Roozen, J. van der Oetelaar, A. Geerlings, and T. Vliegenthart, "Source identification and noise reduction of a reciprocating compressor: A case history," Int. J. Acoust. Vib. **14**, 90–98 (2009).

[6]R. B. Randall, "The application of fault simulation to machine diagnostics and prognostics," Int. J. Acoust. Vib. **14**, 81–89 (2009).

[7]J. Tuma, "Gearbox noise and vibration prediction and control," Int. J. Acoust. Vib. **14**, 99–108 (2009).

[8]C. L. Dym and M. A. Lang, "Transmission of sound through sandwich panels," J. Acoust. Soc. Am. **56**, 1523–1532 (1974).

[9]C. L. Dym, C. S. Ventres, and M. A. Lang, "Transmission of sound through sandwich panels: A reconsideration," J. Acoust. Soc. Am. **59**, 364–367 (1976).

[10]C. L. Dym and D. C. Lang, "Transmission loss of damped asymmetric sandwich panels with orthotropic cores," J. Sound Vib. **88**, 299–319 (1983).

[11]J. A. Moore and R. H. Lyon, "Sound transmission loss characteristics of sandwich panel constructions," J. Acoust. Soc. Am. **89**, 777–791 (1991).

[12]L. Gelman, P. Jenkins, I. Petrunin, and M. J. Crocker, "Vibroacoustical damping diagnostics: Complex frequency response function versus its magnitude," Int. J. Acoust. Vib. **11**, 120–124 (2006).

[13]L. Xu and X. Jia, "Electromechanical coupled vibration for double coupled micro beams," Int. J. Acoust. Vib. **12**, 15–24 (2007).

[14]J. B. Mariem and M. A. Hamdi, "A new boundary finite method for fluid-structure interaction problems," Int. J. Numer. Methods Eng. **24**, 1251–1267 (1987).

[15]P. J. T. Filippi, P. O. Mattei, and C. Maury, "Sound transmission through a thin baffled plate: Validation of a light fluid approximation with numerical and experimental results," J. Sound Vib. **229**, 1157–1169 (2000).

[16]T. W. Wu and A. Dandapani, "A boundary element solution for sound transmission through thin panels," J. Sound Vib. **171**, 145–157 (1994).

[17]L. P. Barisciano, "Broadband transmission loss due to reverberant excitation," NASA/CR Report No. 1999-209687, NASA Langley Research Center, Hampton, VA.

[18]P. Thamburaj and J. Q. Sun, "Effect of material and geometry on the sound and vibration transmission across a sandwich beam," J. Vibr. Acoust. **123**, 205–212 (2001).

[19]L. A. Roussos, "Noise transmission loss of a rectangular plate in an infinite baffle," NASA Technical Paper No. NAS 1.6 2398, 1985.

[20]D. G. Crighton, "The 1988 Rayleigh medal lecture: Fluid loading—The interaction between sound and vibration," J. Sound Vib. **133**, 1–27 (1989).

[21]D. Habault and P. J. T. Filippi, "Light fluid approximation for sound radiation and diffraction by thin elastic plates," J. Sound Vib. **213**, 333–374 (1998).

[22]"Calculation procedures," in LMS SYSNOISE Rev 5.6 User's Manual.

[23]C. D. Johnson and D. A. Kienholz, "Finite element prediction of damping in structures with constrained viscoelastic layers," AIAA J. **20**, 1284–1290 (1982).

[24]I. Ver and C. Holmer, "Interaction of sound waves with solid structures," *Noise and Vibration Control* (McGraw-Hill, New York, 1971).

The improvement of a simple theoretical model for the prediction of the sound insulation of double leaf walls

John L. Davy[a)]
School of Applied Sciences, RMIT University, GPO Box 2476V, Melbourne, Victoria 3001, Australia

(Received 1 April 2009; revised 15 November 2009; accepted 18 November 2009)

This paper presents a revised theory for predicting the sound insulation of double leaf cavity walls that removes an approximation, which is usually made when deriving the sound insulation of a double leaf cavity wall above the critical frequencies of the wall leaves due to the airborne transmission across the wall cavity. This revised theory is also used as a correction below the critical frequencies of the wall leaves instead of a correction due to Sewell [(1970). J. Sound Vib. **12**, 21–32]. It is found necessary to include the "stud" borne transmission of the window frames when modeling wide air gap double glazed windows. A minimum value of stud transmission is introduced for use with resilient connections such as steel studs. Empirical equations are derived for predicting the effective sound absorption coefficient of wall cavities without sound absorbing material. The theory is compared with experimental results for double glazed windows and gypsum plasterboard cavity walls with and without sound absorbing material in their cavities. The overall mean, standard deviation, maximum, and minimum of the differences between experiment and theory are −0.6 dB, 3.1 dB, 10.9 dB at 1250 Hz, and −14.9 dB at 160 Hz, respectively.
 [DOI: 10.1121/1.3273889]

PACS number(s): 43.55.Rg, 43.55.Ti, 43.40.Rj, 43.20.Rz [LMW] Pages: 841–849

I. INTRODUCTION

There are still no really satisfactory theoretical models for predicting the sound insulation of walls. This means that acoustical consultants have to use measured values of sound insulation when designing buildings. Measurements of sound insulation are expensive and time consuming to make. This paper improves the accuracy of a simple theoretical model of sound insulation.

The author (Davy, 1990, 1991, 2009a) derives a theoretical model for the sound transmission of double leaf cavity walls due to the airborne transmission across the wall cavity above the critical frequencies of the cavity wall leaves. In this derivation, the integral over angles of incidence from 0° to 90° is approximated by extending the range of integration to $-\infty$ to ∞ following the approach of Cremer (1942) for single leaf walls. This extension of range approximation is used to make the integral easier to evaluate analytically. However, it is possible to evaluate the integral analytically without extending the range of integration. Unfortunately, although the principles of integration involved are straightforward, the algebra becomes very messy. In this paper, the integral is evaluated analytically with the assistance of the MAPLE 11 software package.

The model, developed for use above the critical frequencies of the wall leaves, is also used as a correction below the critical frequencies by assuming that the non-existent critical angle is 90°. This correction replaces Sewell's (1970) correction in the region just below the critical frequencies. This new correction has the advantage that it does not become ill-defined at the critical frequencies of the cavity wall leaves.

When the theory is compared with measurements on double glazed windows it is found that it is necessary to include "stud" borne transmission due to the window frames in order to be able to correctly predict the sound transmission for the wider air gaps. To model steel studs correctly, the author's theory (Davy, 1993, 2009a) is extended to include a minimum stud transmission, and a value for this quantity is derived by comparison with experimental data.

When there is no sound absorbing material in the wall cavity, this paper derives two empirical linear regression equations for the "effective" sound absorption coefficient of the cavity. One of these equations is for double glazing and the other is for gypsum plaster board cavity walls. Because of approximations made in the author's theory, the effective sound absorption coefficient of the cavity is expected to be larger than the actual physical sound absorption coefficient of the cavity.

The theory is compared with experimental results for narrow and wide air gap double glazing and gypsum plaster board cavity walls, both with and without sound absorbing material in the wall cavity. The gypsum plaster board comparison walls include the cases of no structural connection between the wall leaves and steel and wooden studs.

II. ABOVE THE CRITICAL FREQUENCIES

Above the lowest of the critical frequencies of the two wall leaves the method given by Eqs. (1)–(18) of Davy, 2009c is followed. Because these equations are applied to both wall leaves, the subscript $i=1$ or 2 is applied to the variables τ (transmission coefficient), Z (bending wave im-

[a)]Electronic mail: john.davy@rmit.edu.au. Also at CSIRO Materials Science and Engineering, P.O. Box 56, Highett, Victoria 3190, Australia.

pedance), m (mass per unit area), η (total damping loss factor), ω_c (angular critical frequency), f_c (critical frequency), η_{int} (internal damping loss factor), η_{rad} (radiation damping loss factor), a (ratio of mass impedance to twice the characteristic impedance of air), and θ_c (coincidence angle). To avoid confusion with a symbol used later on, the variable r (ratio of frequency to critical frequency) is replaced with the variable ξ, and the subscript i is also applied to it. Equation (2) of Davy, 2009c can also be written as [Cremer (1942), Eq. (8.19)]

$$\tau = 2\int_0^1 \tau(\theta)\sin\,\theta d(\sin\,\theta). \tag{1}$$

Equation (18) of Davy, 2009c can be written as

$$\tau_i(\theta) = \frac{1}{s_i^2}\frac{1}{q_i^2 + (x - p_i)^2}, \tag{2}$$

where

$$p_i = 1 - \frac{1}{\xi_i}, \tag{3}$$

$$s_i = \frac{2a_i\xi_i}{\sigma(\theta_{ci})}, \tag{4}$$

$$q_i = \frac{1 + \dfrac{a_i\eta_i}{\sigma(\theta_{ci})}}{s_i}, \tag{5}$$

and $\sigma(\theta_{ci})$ is the single sided forced radiation efficiency at the coincidence angle of the ith wall leaf. Following the approach of Davy (1990, 1991, 2009c), above the lower of the critical frequencies of each leaf, the sound transmission coefficient $\tau(\theta)$ of a double leaf cavity wall is approximated as

$$\tau(\theta) = \frac{\tau_1(\theta)\tau_2(\theta)}{\alpha^2}, \tag{6}$$

where α is the sound absorption coefficient of the wall cavity. The reason for adoption of this equation is explained at the end of Sec. III.

Putting Eq. (6) into Eq. (2) of Davy, 2009c gives the diffuse field sound transmission coefficient as

$$\tau = \frac{I}{s_1^2 s_2^2 \alpha^2}, \tag{7}$$

where α is the sound absorption coefficient of the wall cavity of the double leaf wall and

$$I = \int_0^1 \frac{dx}{[q_1^2 + (x - p_1)^2][q_2^2 + (x - p_2)^2]}. \tag{8}$$

The integral in Eq. (8) can be evaluated using the methods outlined in Secs. 2.101–2.103 of Gradshteyn and Ryzhik, 1965. However, the algebra is rather complicated. Thus the integral was evaluated with the assistance of the MAPLE 11 software package.

$$I = \frac{A + B + C}{D}, \tag{9}$$

where

$$A = q_1 q_2 (p_2 - p_1)\ln\left\{\frac{[q_1^2 + (p_1 - 1)^2](q_2^2 + p_2^2)}{[q_2^2 + (p_2 - 1)^2](q_1^2 + p_1^2)}\right\}, \tag{10}$$

$$B = q_1[(p_1 - p_2)^2 + q_1^2 - q_2^2]\left[\arctan\left(\frac{p_2}{q_2}\right) - \arctan\left(\frac{p_2 - 1}{q_2}\right)\right], \tag{11}$$

$$C = q_2[(p_2 - p_1)^2 + q_2^2 - q_1^2]\left[\arctan\left(\frac{p_1}{q_1}\right) - \arctan\left(\frac{p_1 - 1}{q_1}\right)\right], \tag{12}$$

$$D = q_1 q_2[(p_2 - p_1)^2 + (q_2 + q_1)^2][(p_2 - p_1)^2 + (q_2 - q_1)^2]. \tag{13}$$

If the two wall leaves have the same properties, $q_1 = q_2 = q$ and $p_1 = p_2 = p$. This makes Eq. (9) indeterminate. In this situation, Eq. (8) becomes

$$I = \int_0^1 \frac{dx}{[q^2 + (x - p)^2]^2}. \tag{14}$$

Evaluating Eq. (14) using the MAPLE 11 software package gives

$$I = \frac{q^2 - p(p - 1)}{2q^2(q^2 + p^2)[q^2 + (p - 1)^2]} + \frac{\arctan\left(\dfrac{p}{q}\right) - \arctan\left(\dfrac{p - 1}{q}\right)}{2q^3}, \tag{15}$$

where use has been made of the fact that $q > 0$ and $p \geq 0$.

The integrand in Eq. (8) has local maxima when $x = p_i$. If $q_i \ll 1$, which is usually the case, and if $|p_1 - p_2| \gg q_i$, the integrand is half its local maximum value when $|x - p_i| = q_i$. Since q_i is usually very much less than 1, and p_i is between 0 and 1 if the frequency is greater than or equal to f_{ci}, the values of x where the integrand is significantly different from zero usually lie well inside the integral limits from 0 to 1. Because of this Davy (1990, 1991, 2009a) approximated the integral in Eq. (8) by extending the limits of integration from minus infinity to plus infinity.

$$I = \int_{-\infty}^{\infty} \frac{dx}{[q_1^2 + (x - p_2)^2][q_1^2 + (x - p_2)^2]}. \tag{16}$$

This integral can be evaluated using the calculus of residues. This evaluation was carried out using the MAPLE 11 software package and gave

$$I = \frac{\pi(q_1 + q_2)}{q_1 q_2[(q_1 + q_2)^2 + (p_1 - p_2)^2]}. \tag{17}$$

If

$$\frac{a_i \eta_i}{\sigma(\theta_{ci})} \gg 1, \tag{18}$$

which is usually the case, then

$$q_i = \frac{\eta_i}{2\xi_i}. \tag{19}$$

Substituting Eqs. (17) and (19) into Eq. (7) and evaluating give

$$\tau = \frac{\pi\sigma^2(\theta_{c1})\sigma^2(\theta_{c2})(\eta_1\xi_2 + \eta_2\xi_1)}{2a_1^2a_2^2\eta_1\eta_2\alpha^2[(\eta_1\xi_2 + \eta_2\xi_1)^2 + 4(\xi_1 - \xi_2)^2]}. \tag{20}$$

If an infinite wall leaf is assumed, the forced radiation efficiency is given by

$$\sigma(\theta_{ci}) = \frac{1}{\sqrt{1 - \frac{1}{\xi_i}}} = \frac{1}{\sqrt{1 - \frac{\omega_{ci}}{\omega}}} = \frac{1}{\sqrt{1 - \frac{f_{ci}}{f}}} = \frac{1}{\cos\ \theta_{ci}}. \tag{21}$$

Because Eq. (21) gives an infinite result at the critical frequency, a common approximation is to assume that the forced radiation efficiency is unity above the critical frequency. If this assumption is made, Eq. (20) becomes

$$\tau = \frac{\pi(\eta_1\xi_2 + \eta_2\xi_1)}{2a_1^2a_2^2\eta_1\eta_2\alpha^2[(\eta_1\xi_2 + \eta_2\xi_1)^2 + 4(\xi_1 - \xi_2)^2]}. \tag{22}$$

Equation (22) should agree with Eqs. (30)–(32) of Davy, 1990, Eqs. (9)–(13) of Davy, 1991, and Eqs. (40)–(44) of Davy, 2009a. It does not do so for two reasons. The first reason is errors in the older equations. Equation (33) of Davy, 1990 and Eq. (12) of Davy, 1991 should read $v = 4(\xi_1 - \xi_2)$ rather than $v = 4(\eta_1 - \eta_2)$. Also the factor α^2 is missing from the denominator of Eq. (34) of Davy, 1990. The second reason is that the approximations made are slightly different. However, if the critical frequencies of the two wall leaves are the same, then $\xi_1 = \xi_2 = \xi$ and Eqs. (22) and (34) of Davy, 1990 (after correction), Eq. (13) of Davy, 1991 (after correction), and Eq. (44) of Davy, 2009a all reduce to

$$\tau = \frac{\pi}{2a_1^2a_2^2\eta_1\eta_2\xi\alpha^2(\eta_1 + \eta_2)}. \tag{23}$$

It should be noted that the ξ_i in Davy, 1990, 1991, 2009a is the square root of the ξ_i used in this paper. If the masses per unity area and the damping loss factors of both wall leaves are the same, then $a_1 = a_2 = a$ and $\eta_1 = \eta_2 = \eta$, and all four equations reduce further to

$$\tau = \frac{\pi}{4a^4\eta^3\xi\alpha^2}. \tag{24}$$

At the critical frequency of the ith wall leaf, $\xi_i = 1$ and $p_i = 0$. This means that one of the maxima of the integrand in Eq. (8) occurs at zero, which is the lower limit of integration in Eq. (8). Thus the extension of the limits of integration, which occurs in Eq. (16), means that this maximum contributes about double to the integral compared to what it did before the extension of the limits of integration. Thus, in this paper, Eqs. (5), (7), (9)–(13), and (15) will be used to calculate the sound transmission coefficient instead of the further approximations given by Eqs. (17), (19), (20), and (22)–(24).

Because Eq. (21) gives an infinite value for the forced radiation efficiency at the critical frequency, this paper uses Davy's (2009b) theory to calculate the forced radiation efficiency. This theory is an updated version of the theory in Davy, 2004. First the cosine of the coincidence angle is calculated. Since Eqs. (5), (7), (9)–(13), and (15) are going to be used as a correction term below the critical frequency, the cosine of the coincidence angle is set to zero for frequencies below the critical frequency. The actual equations used are Eqs. (33)–(39) of Davy, 2009c.

To obtain better agreement with experimental results, $\sigma(\theta_{ci})$ is set equal to 1 if the frequency is greater than or equal to the lower of the two critical frequencies. If the frequency is between 0.9 times and 1 times the lower of the two critical frequencies, $\sigma(\theta_{ci})$ is linearly interpolated in the frequency domain between the value of $\sigma(\theta_{ci})$ at 0.9 times the lower of the two critical frequencies and the value of 1 at the lower of the two critical frequencies.

An alternative approximation is used by Davy (1990, 1991, 2009a). It produces a different result from the approximation used in this paper if the critical frequencies of the two wall leaves are different. This other approximation follows Cremer (1942) and uses Eq. (1) rather than Eq. (2) of Davy, 2009c. The $\sin\ \theta$ term in Eq. (1) is approximated by the average of $\sin\ \theta_{c1}$ and $\sin\ \theta_{c2}$. This is the approximation that causes the difference if the critical frequencies of the wall leaves are different. It is not needed if Eq. (2) of Davy, 2009c is used but works satisfactorily if the critical frequencies are fairly similar. The second $\sin^4\ \theta$ term in Eq. (12) of Davy, 2009c is approximated as

$$\sin^4\ \theta \approx \sin^4\ \theta_{ci}\left(1 + 4\frac{\sin\ \theta - \sin\ \theta_{ci}}{\sin\ \theta_{ci}}\right). \tag{25}$$

Then the limits of integration are extended from minus infinity to plus infinity as in Eq. (16). Finally radiation efficiency is approximated to be unity.

III. BELOW THE CRITICAL FREQUENCIES

Below the lowest of the critical frequencies of the two wall leaves the method given by Eqs. (18)–(32) of Davy, 2009a is followed. An important difference is that a correction factor (Sewell, 1970) for when the critical frequency is approached is omitted from Eq. (24) of Davy, 2009a. That is, Eq. (24) of Davy, 2009a becomes

$$a_i = \frac{\omega m_i}{2\rho_0 c}, \tag{26}$$

where $\rho_0 c$ is the characteristic impedance of air. Another difference is that the square root of the panel area $\sqrt{A}$ in Eq. (34) of Davy, 2009a is replaced with the length $2a$ of the side of an equivalent square panel defined by Eq. (34) of Davy,

2009c. Note that this a is different from the a used earlier in this paper. To avoid confusion with symbols used earlier in this paper p is replaced with P and q is replaced with Q.

Davy (2009a) used a variable limiting angle of integration (Sewell, 1970) and showed that the use of this angle for single leaf walls is equivalent to the use of the forced radiation efficiency of the finite size of the wall averaged over all angles of incidence. Thus the use of the variable limiting angle of integration is equivalent to the spatial windowing technique of Villot *et al.* (2001). For double leaf walls, Eq. (34) of Davy, 2009a, which is used in this paper, limits the maximum value of the limiting angle of integration to 61° (Davy, 1998) in order to make the theory agree better with experimental results. Similarly, Schoenwald *et al.* (2008) found that the spatial windowing technique was not sufficient on its own for double leaf walls and combined it with Kang's (2000) Gaussian distribution of directional incident sound energy.

To compensate for the removal of Sewell's (1970) correction term in Eq. (26), the sound transmission coefficient, above the normal incidence mass-air-mass resonance frequency and below the upper of the two critical frequencies, is calculated as the sum of Eq. (33) of Davy, 2009a and Eq. (7) where Eq. (9) or Eq. (15) has been used to evaluate Eq. (7). At and above the upper of the two critical frequencies, only Eq. (7) is used. Equation (33) of Davy, 2009a can be included immediately above the upper of the two critical frequencies, but it causes problems if used a long way above both critical frequencies.

Q is half the sum of the ratio of the surface density of one wall leaf to that of the other with the inverse of this ratio. Thus if the wall leaves are the same $Q=1$. Px is the product of the ratio of the mass impedance of one wall leaf to twice the specific acoustic impedance of air with the same ratio for the other wall leaf and with the sound absorption coefficient of the cavity and with the square of the cosine of the angle of incidence. Because Q is usually very much less than Px, Eq. (28) of Davy, 2009a can be approximated as follows:

$$\tau(\theta) = \frac{1}{a_1^2 \cos^2(\theta) a_2^2 \cos^2(\theta) \alpha^2}. \tag{27}$$

The first term in the denominator of Eq. (12) of Davy, 2009c is usually very much smaller than the second term. Ignoring this first term, assuming that ξ_i is very small, and using the value of the radiation efficiency for an infinite panel, which is given by Eq. (21), change Eq. (12) of Davy, 2009c to

$$\tau_i(\theta) = \frac{1}{a_i^2 \cos^2(\theta)}. \tag{28}$$

Thus Eq. (27) can be written as Eq. (6). This is the reason for the adoption of Eq. (6).

IV. STUD BORNE TRANSMISSION

This paper is primarily concerned with airborne transmission across the cavity. However, to the author's surprise, it was found that it was necessary to include "stud borne" transmission between the edges of the glass panes due to the window frames in order to correctly predict the sound insulation of double glazing with larger air gaps. Thus this section gives the formulas for stud borne transmission across the cavity (Davy, 1993, 2009a). The actual equations used are Eqs. (47)–(50) of Davy, 2009a where the symbols D (ratio of total to non-resonant radiation) and g (sum of the product of the mass per unit area of each wall leaf with the square root of the angular critical frequency of the other wall leaf) of Davy, 2009a have been replaced with the symbols H and G to avoid confusion with the use of the symbols D and g earlier in this paper. Equation (46) of Davy, 2009a is split into the following two equations:

$$\tau = \frac{32\rho_0^2 c^3 HJ}{G^2 b \omega^2}, \tag{29}$$

where the stud transmission ratio is

$$J = \frac{2}{1 + \left(1 - \frac{4\omega^{3/2} m_1 m_2 c C_M}{G}\right)^2}. \tag{30}$$

The studs have a mechanical compliance of C_M where $C_M=0$ gives the rigid stud case. The mechanical compliance of a stud has dimensions of length per (force per length of the stud). The stud transmission J is the ratio of the stud borne sound transmission coefficient for the wall to the stud borne sound transmission coefficient for the same wall with rigid studs ($C_M=0$). The stud transmission J is restricted to be greater than or equal to the set minimum stud transmission K. C_M and K are selected to give the best agreement with experiment. This is different from Davy, 2009a, which gives the formula for resilient studs but recommends setting the stud compliance to zero and using a set constant stud transmission J for resilient studs instead of the stud transmission J given by Eq. (30). In Davy, 2009a, J is selected to give the best agreement with experiment.

Research by Poblet-Puig *et al.* (2006, 2009), Guigou-Carter *et al.* (1998), and Guigou-Carter and Villot (2006) has shown that a steel stud can be modeled as a translational spring with a translational stiffness, which varies with frequency in the range from 10^5 to 10^8 Pa. The constant value of mechanical compliance used in Sec. V corresponds to a translational stiffness of 6×10^5 Pa, which lies toward the bottom end of the above range. The value of the minimum stud transmission used in Sec. V is −23 dB. This also lies in the 0 to −40 dB stud transmission range determined by Poblet-Puig *et al.* (2006, 2009) for a standard steel stud. It would be possible to use the frequency dependent translational stiffness values determined by Poblet-Puig *et al.* (2006, 2009) with the theory developed in this paper.

σ_i, which appears in Eqs. (48) and (49) of Davy, 2009a, is the single sided radiation efficiency of a free reverberant bending wave vibration field of the ith wall leaf. Note that, below the critical frequency of the ith wall leaf, this is different from the single sided radiation efficiency of a forced bending wave on the ith wall leaf, which is given by Eqs. (33)–(39) of Davy, 2009c. The corrected versions of Maidanik's formulas for the single sided radiation efficiency of a free reverberant bending wave vibration field given by Vér and Holmer (1971) are used in this paper. However, the

TABLE I. This table shows the mean, the standard deviation, the maximum, and the minimum in decibels of the differences between experiment and theory for each of the seven cases presented in the accompanying figures. Also shown are the frequencies in hertz at which the maximum and minimum differences occur. The overall column shows the average of the seven means, the root mean square of the seven standard deviations, the maximum of the seven maxima, and the minimum of the seven minima. In Figs. 2 and 3, only the combined theoretical results are used for the calculations whose results are shown in this table.

	Figure 1	Figure 2	Figure 3	Figure 4	Figure 5	Figure 6	Figure 7	Overall
Mean (dB)	−0.6	1.1	1.2	−2.1	0.9	−3.3	−1.5	−0.6
Standard deviation (dB)	2.4	2.7	2.7	3.1	3.9	4.3	2.3	3.1
Maximum (dB)	3.0	7.6	8.4	4.2	10.9	8.5	2.1	10.9
Minimum (dB)	−9.3	−2.0	−2.6	−12.5	−5.4	−14.9	−5.6	−14.9
Frequency of maximum (Hz)	200	80	160	50	1250	80	4000	1250
Frequency of minimum (Hz)	2000	400	80	2000	315	160	160	160

maximum value of the radiation efficiency is limited to the value of 1. Previous research (Cremer (1942); Davy, 1993, 2009a) has shown that this assumption works well for predicting the sound transmission of third octave bands of noise above the critical frequency. It is also consistent with the restrictions placed on the result of Eqs. (33)–(39) of Davy, 2009c above the critical frequency in Sec. II.

The stud borne transmission is not included in the combined transmission below the mass-air-mass resonance frequency. In this frequency range, the wall leaves are already effectively coupled by the air cavity.

V. COMPARISON WITH EXPERIMENT

Quantitative measures of the differences between experiment and theory for each of the seven cases shown in the following figures are given in Table I. This table shows the mean, the standard deviation, the maximum, and the minimum in decibels. Also shown are the frequencies in hertz at which the maximum and minimum differences occur. The overall column shows the average of the seven means, the root mean square of the seven standard deviations, the maximum of the seven maxima, and the minimum of the seven minima. In Figs. 2 and 3, only the combined theoretical results are used for the calculations whose results are shown in Table I.

The first comparison is for the case where there is no vibration connection between the two leaves of the wall (except possibly at the edges) and hence only airborne cavity wall transmission is involved. In this first comparison there is sound absorbing material in the wall cavity. The value of 1 is used for the cavity sound absorption coefficient for cavity walls with sound absorbing material in the cavity in this paper, since the experimental results show little dependence of sound insulation on the type or the thickness of the sound absorbing material in the cavity, providing that the sound absorbing material is not too thin or lightweight.

Figure 1 compares theory with experimental results for five 40 mm double steel stud, 16 mm gypsum plaster board cavity walls with cavity absorption. There is a 10 mm gap between the separate studs. The double studs are spaced on 610 mm centers. The experimental results in Fig. 1 were measured by the National Research Council of Canada (NRCC). The last of the walls measured (TL-92-975) had no studs. The gypsum plaster board is assumed to have a density of 770 $\mathrm{kg/m^3}$, Young's modulus of 1.85×10^9 Pa, Poisson's ratio of 0.3, and an internal damping loss factor of 0.03. The agreement between theory and experiment is good. Figure 1 should be compared with Fig. 3 of Davy, 2009a where the same experimental data are presented. Slight differences in the two theories can be seen in the frequency region around the peak below the critical frequency.

To determine the appropriate value of the "sound absorption coefficient" to be used to predict the sound insulation of cavity walls without sound absorbing material in the cavity, the NRCC measurements on double glazed windows (Quirt, 1981, 1982) were analyzed. It soon became apparent that it was necessary to include stud borne transmission between the edges of the glass panes due to the window frames in order to correctly predict the sound insulation of double glazing with larger air gaps. This is the reason for the inclusion of Sec. IV in this paper. The window size is 2.02×1.8 $\mathrm{m^2}$. The frame has two vertical dividers. Thus each layer of glass consists of three separate panes of glass. This gives a stud spacing of $b=2.02/3=0.67$ m. The theoretical calculations assume that the density is 2500 $\mathrm{kg/m^3}$, Young's modulus is 6.5×10^{10} Pa, and Poisson's ratio is 0.22. The stud compliance C_M of the window frames is assumed to be 0 $\mathrm{Pa^{-1}}$.

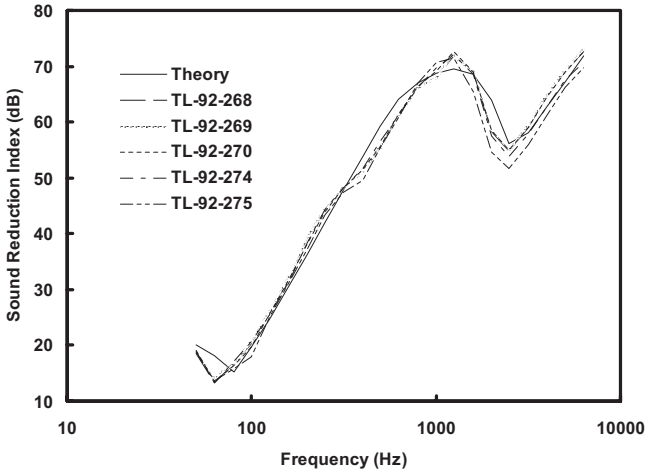


FIG. 1. Comparison of theory and experiment (measured by National Research Council of Canada) for the sound insulation of five double steel stud 16 mm gypsum plaster board cavity walls with cavity absorption. The cavity width is 90 mm.

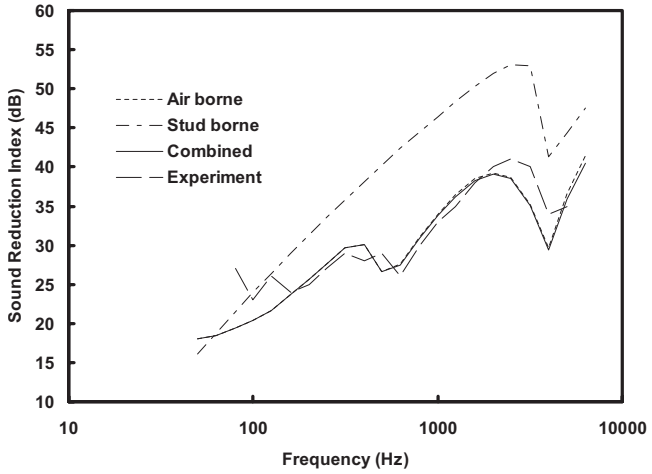


FIG. 2. Comparison of the airborne, stud borne, and combined theoretical results with the experimental result (Quirt, 1981, 1982) for the sound insulation of a sealed double glazed unit consisting of two 3 mm glass panes separated by an air cavity of 3 mm width.

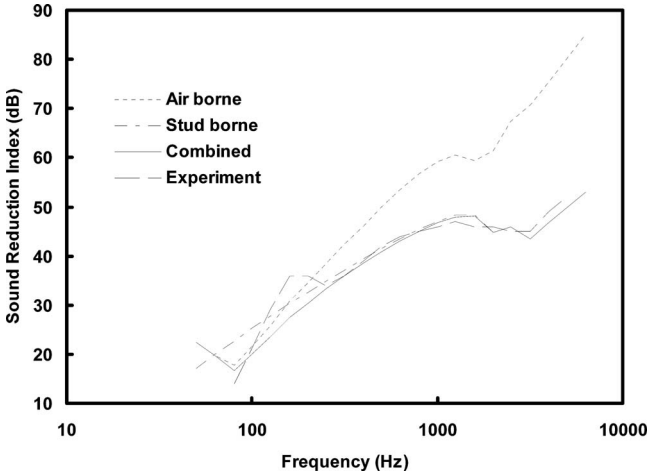


FIG. 3. Comparison of the airborne, stud borne, and combined theoretical results with the experimental result (Quirt, 1981, 1982) for the sound insulation of a double glazed window consisting of 4 and 6 mm glass panes separated by an air cavity of 100 mm width.

The sound absorption coefficient of the cavity and the internal damping loss factor of the glass panes were assumed to be constant as functions of frequency. They were adjusted to provide the best fit between theory and experiment for each of the double glazing combinations that were tested. The sound absorption coefficient α of the cavity increased with increasing cavity width d. A linear regression of the sound absorption coefficient as a function of cavity width produced the following equation:

$$\alpha = 0.027 + 2.4d. \tag{31}$$

Note that the value of the sound absorption coefficient is limited by Eq. (35) of Davy, 2009a. Because the second term on the right hand side of Eq. (27) of Davy, 2009a has been ignored, the effective sound absorption coefficient α of the cavity is expected to be considerably larger than the actual physical sound absorption coefficient of the cavity. Equation (31) is less accurate at larger cavity widths because, as will be seen, the stud borne transmission across the window air cavity is dominant in the theoretical values for larger cavity widths.

The average value of the internal damping loss factor of the glass panes was 0.064. Although this internal damping loss factor seems high for glass, it should be noted that Cremer (1942) assumed a damping loss factor of 0.1 for glass. This internal damping loss factor actually includes the loss of vibrational energy at the edges of the glass panes.

The narrowest window air cavity gap measured by NRCC was 3 mm. Figure 2 shows a comparison of the airborne, stud borne, and combined theoretical results with the experimental result for a double glazed unit consisting of a 3 mm glass pane, a 3 mm air gap, and another 3 mm glass pane. Figure 2 shows that the theoretical sound insulation is controlled by the airborne transmission across the cavity. Remember that the theoretical stud borne transmission is not used below the mass-air-mass resonance, which occurs at about 550 Hz in this situation. The theory underestimates the experimental sound reduction index below 160 Hz and from 2500 to 4000 Hz. The agreement is good in the other frequency ranges. The underestimation in the 2500–4000 Hz range can be removed by increasing the apparent sound absorption coefficient in this frequency range.

The next comparison shows the case where the two "wall leaves" are different and where the window air cavity gap is large. Figure 3 compares the airborne, stud borne, and combined theoretical results with the experimental result for a double glazed window consisting of 4 and 6 mm glass panes separated by an air cavity of 100 mm width. Figure 3 shows that the theoretical sound insulation is controlled by the stud borne transmission across the cavity above 160 Hz and by the airborne transmission across the cavity below 160 Hz. The combined theoretical results agree reasonably well with the experimental results, although the theory fails to predict the local maxima at 160 and 200 Hz.

To estimate the stud mechanical compliance C_M and the minimum stud transmission K for use in the case of steel studs, test data were taken for five walls from Halliwell *et al.*, 1998. All five walls consisted of two layers of 16 mm gypsum plaster board mounted on each side of 90 mm steel studs at 406 mm centers. All five walls had 90 mm of porous sound absorbing material in the wall cavity. The first wall had sprayed cellulose fiber, while the second and third walls had glass fiber, and the last two walls had mineral fiber. The best fit to the experimental data is obtained by setting the stud mechanical compliance C_M equal to 1.6×10^{-6} Pa^{-1} and the minimum stud transmission K equal to 0.005.[1]

The gypsum plaster board is assumed to have a density of 770 $\mathrm{kg/m}^3$ and Poisson's ratio of 0.3. To ensure that the two layers of 16 mm gypsum plaster board have the same critical frequency as a single layer of 16 mm gypsum plaster board, Young's modulus is set equal to one-quarter of the 1.85×10^9 Pa used for a single 16 mm layer of gypsum plaster board. To obtain good agreement above the critical frequency, the internal damping loss factor is set to 0.02. The comparison of theory with experiment is shown in Fig. 4. The theory overestimates at 100 Hz and in the 1000–2000 Hz range. The agreement is good at other frequencies.

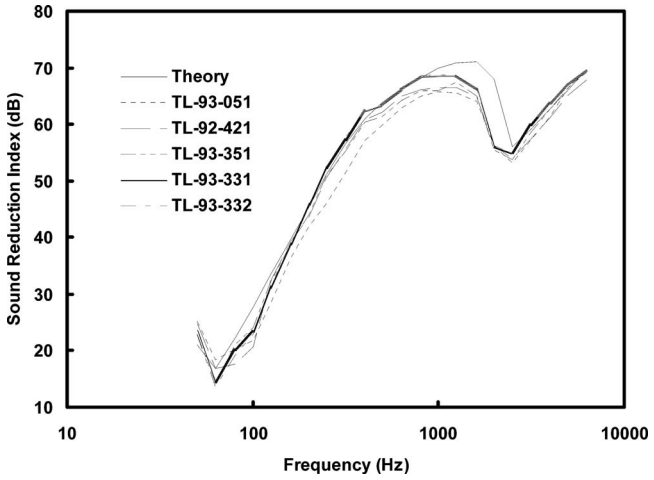


FIG. 4. Comparison of theory and experiment (Halliwell *et al.*, 1998) for the sound insulation of five 90 mm steel stud cavity walls with cavity absorption. Each side of the steel studs has two layers of 16 mm gypsum plaster board attached.

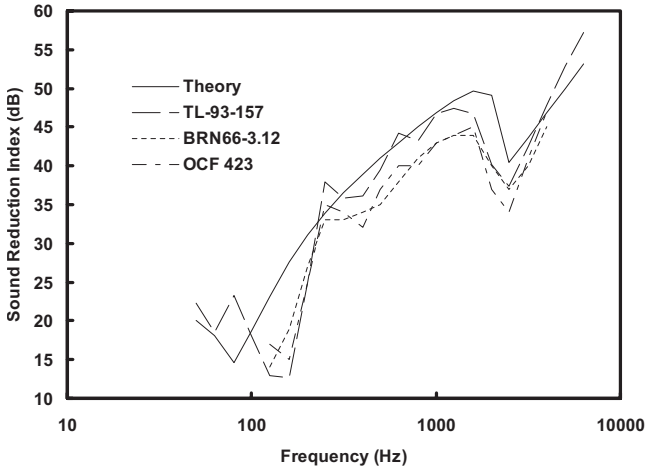


FIG. 6. Comparison of theory and experiment (Halliwell *et al.*, 1998; Northwood, 1968; Dupree, 1981) for the sound insulation of three 90 mm wooden stud, 16 mm gypsum plaster board cavity walls with cavity absorption.

The next comparison is for gypsum plaster board cavity walls without cavity absorption. The last three experimental results in Fig. 5 (TL-92-265 to TL-92-267) are for the same construction as the results in Fig. 1, except that there is no sound absorbing material in the cavity. In the first three experimental results (TL-92-262 to TL-92-264), the double 40 mm steel studs with a 10 mm gap are replaced with 90 mm steel studs on 813 mm centers. These 90 mm steel stud results have been included, because surprisingly they are as good as or better than the double 40 mm steel stud results. The theoretical curve includes stud borne transmission, but the theoretical stud borne transmission only has a small effect on the total theoretical sound insulation in the region of 100 Hz. Another experimental result for the case with no studs, which is not included here, produced lower results. Presumably this is because the studs help inhibit the oblique propagation of sound in the cavity in this case without sound absorption in the cavity. The experimental measurements in Fig. 5 were also measured by the NRCC.

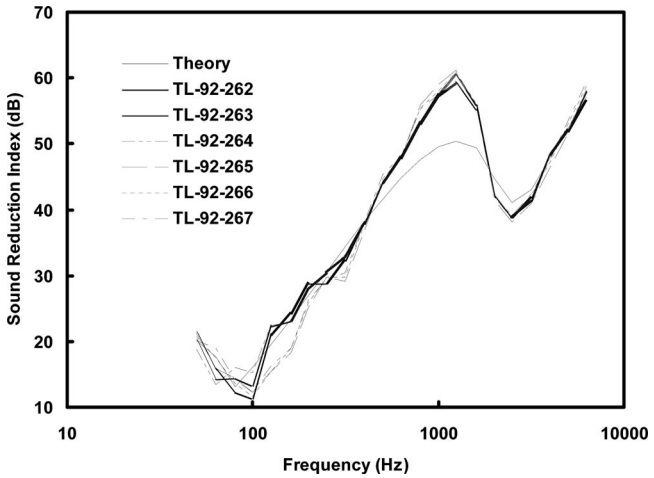


FIG. 5. Comparison of theory and experiment (measured by National Research Council of Canada) for the sound insulation of three steel stud and three double steel stud 16 mm gypsum plaster board cavity walls without cavity absorption. The cavity width is 90 mm.

The original intention was to use Eq. (31) to predict the cavity absorption coefficient. However, comparison with the experimental results shown in this figure, and the results without cavity absorption from Halliwell *et al.*, 1998 and from NAHB, 1971, produced the following equation:

$$\alpha = 0.043 + 0.73d. \tag{32}$$

The mean value of the damping loss factor was 0.044. This value of the damping loss factor and the value of sound absorption coefficient given by Eq. (32) were used for the theoretical calculations in Figs. 5 and 7. The theory in Fig. 5 underestimates the experimental results in the range from 500 to 1600 Hz. Use of Eq. (31) instead of Eq. (32) does a better job of predicting the peak in this range at the expense of overestimating in the range from 160 to 630 Hz. Figure 5 should be compared to Fig. 4 of Davy, 2009a where the same experimental data are plotted. The theory of Davy (2009a) is closer to the experimental results in the frequency range from 500 to 1600 Hz. The theory of this paper is closer to the experimental results in the 250–400 Hz range and at 2000 Hz.

Figure 6 compares theory and experiment for the sound insulation of three 90 mm wooden stud, 16 mm gypsum plaster board cavity walls with cavity absorption. The stud spacing is 406 mm. The first two experimental results are a new and an old measurement by the National Research Council of Canada (Halliwell *et al.*, 1998; Northwood, 1968). The third experimental measurement is by Owens/Corning Fiberglas (Dupree, 1981). The second and third experimental results also appear as an average in NAHB, 1971. In order to obtain reasonable agreement above the critical frequency, the damping loss factor used to obtain the theoretical result had to be increased to 0.1.

The theory does not do a good job of predicting the experiment results below 250 Hz. Between 315 and 1250 Hz it is in reasonable agreement with the most recent measurement, but overpredicts the two older measurements. It overpredicts between 1600 and 2500 Hz. It is in very rough

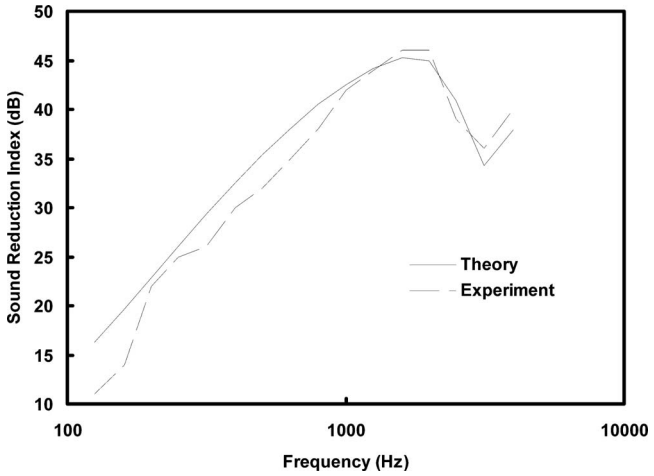


FIG. 7. Comparison of theory and experiment (NAHB, 1971) for the sound insulation of a 13 mm gypsum plaster board, 90 mm cavity wall without cavity absorption.

agreement with the experimental results above 2500 Hz, but only because of the adoption of a theoretical damping loss factor of 0.1.

Note that the predicted mass-air-mass resonance frequency of about 80 Hz is significantly less than the measured mass-air-mass resonance frequencies of 125 or 160 Hz. This may be due to a structural resonance, which is not included in the theory described in this paper. Bradley and Birta (2001) showed that the sound insulation of wood stud exterior walls can be significantly degraded by a structural resonance if the two wall leaves are rigidly coupled by the wooden studs. They explained this structural resonance in terms of the analysis conducted by Lin and Garrelick (1977). The effects of this resonance can be reduced by structurally isolating the two wall leaves with resilient mounts, thin steel studs, staggered studs, or double studs. The frequency of the resonance is about double the calculated mass-air-mass resonance, and it reduces in frequency as the rigid stud spacing is increased and as the depth of the rigid studs is increased.

Bradley and Birta (2000) reported the results of laboratory sound insulation measurements on typical Canadian building facades. These measurements showed the structural resonance at 125 Hz. However, field measurements (Bradley *et al.*, 2001; Bradley, 2002) with actual aircraft noise showed little effect due to this structural resonance.

Figure 7 compares theory and experiment for 13 mm gypsum plaster board on each side of 90 mm thick wooden studs with no sound absorbing material in the wall cavity. The studs are spaced at 400 mm centers. The experimental results are the average results of two separate experimental measurements (NAHB, 1971). Although the theory overestimates the experimental results below 1000 Hz, it still does a reasonable job of predicting the general trend of the experimental data. The predicted results would compare better with the experimental ones if the value of the cavity absorption coefficient was modified. However, in order to have a prediction method, the cavity absorption given by linear regression equation (32) has been used.

VI. CONCLUSIONS

A simple theory for predicting the sound insulation of double leaf cavity walls has been revised. An approximation to the range of integration over angle of incidence for the theory above the critical frequency is removed. This enables the theory for above the critical frequency to be used as a correction to theory for below the critical frequency instead of Sewell's (1970) correction and thus gives continuous sound insulation values. The stud borne transmission via the window frames is included when modeling the sound insulation of double glazed windows. This stud borne transmission via the window frames is particularly important for windows with wide air gaps. Linear regression equations for the effective sound absorption coefficient of a wall cavity without sound absorbing material in the wall cavity as a function of cavity width are derived for both double glazed windows and gypsum plaster board cavity walls. This effective sound absorption increases with increasing wall cavity width.

Comparison with experiment shows that the theory does a reasonably good job of predicting the general trend of the experimental values. The overall mean, standard deviation, maximum, and minimum of the differences between experiment and theory are −0.6 dB, 3.1 dB, 10.9 dB at 1250 Hz, and −14.9 dB at 160 Hz, respectively. The theory struggles most when attempting to predict the sound insulation of double leaf gypsum plasterboard walls without sound absorbing material in the wall cavity. This indicates that the assumption that the cavity sound absorption coefficient is constant with frequency, except at low frequencies, probably needs to be revised.

[1]It should be noted that the value of the stud mechanical compliance C_M is similar to the value of 1×10^{-6} Pa^{-1} used by Davy (1990). The minimum stud transmission K is not needed in Davy, 1990 because the theoretical airborne transmission across the cavity is greater in that paper and because the limiting angle θ_l is not limited to a maximum value of 61°.

Bradley, J. S. (**2002**). "IBANA-Calc validation studies," Institute for Research in Construction Research Report No. IRC RR-125, National Research Council of Canada, Ottawa, Canada.

Bradley, J. S., and Birta, J. A. (**2000**). "Laboratory measurements of the sound insulation of building façade elements," Institute for Research in Construction Internal Report No. IRC IR-818, National Research Council of Canada, Ottawa, Canada.

Bradley, J. S., and Birta, J. A. (**2001**). "On the sound insulation of wood stud exterior walls," J. Acoust. Soc. Am. **110**, 3086–3096.

Bradley, J. S., Lay, K., and Norcross, S. G. (**2001**). "Measurements of the sound insulation of a wood framed house exposed to aircraft noise," Institute for Research in Construction Internal Report No. IRC IR-831, National Research Council of Canada, Ottawa, Canada.

Cremer, L. (**1942**). "Theorie der schalldämmung wände bei schrägem einfall (Theory of the sound attenuation of thin walls with oblique incidence)," Akust. Z. **7**, 81–104; [Most of this article has been republished with an English language summary in Northwood, T. D. (**1977**). "Theory of the sound attenuation of thin walls with oblique incidence," in *Architectural Acoustics*, Benchmark Papers in Acoustics Vol. **10**, edited by T. D. Northwood (Dowden, Hutchinson and Ross, Stroudsburg, PA), pp. 367–399].

Davy, J. L. (**1990**). "A model for predicting the sound transmission loss of walls," *The Australian Vibration and Noise Conference 1990* (The Institution of Engineers, Australia, Canberra, Australia), National Conference Publication No. 90/9, pp. 23–27.

Davy, J. L. (**1991**). "Predicting the sound insulation of stud walls," *The Costs of Noise—Proceedings of Inter-Noise 91 Conference*, edited by A. Lawrence (Australian Acoustical Society, Sydney, Australia), Vol. **1**, pp. 251–254.

Davy, J. L. (**1993**). "The sound transmission of cavity walls due to studs," in *People Versus Noise—Proceedings of the Inter-Noise 93 Conference*, edited by P. Chapelle and G. Vermeir (Belgium Acoustical Association, Leuven, Belgium), Vol. **2**, pp. 975–978.

Davy, J. L. (**1998**). "Problems in the theoretical prediction of sound insulation," *Sound and Silence: Setting the Balance, Proceedings of the International Conference on Noise Control Engineering*, edited by V. C. Goodwin and D. C. Stevenson (New Zealand Acoustical Society Inc., Auckland, New Zealand), Paper No. 44.

Davy, J. L. (**2004**). "The radiation efficiency of finite size flat panels," *Acoustics 2004, Transportation Noise and Vibration—The New Millennium, Proceedings of the Annual Conference of the Australian Acoustical Society*, Gold Coast, Australia, 3–5 November, edited by M. J. Mee, R. J. Hooker, and I. D. M. Hillock (Australian Acoustical Society, Castlemaine, Victoria, Australia), pp. 555–560.

Davy, J. L. (**2009a**). "Predicting the sound insulation of walls," Build. Acoust. **16**, 1–20.

Davy, J. L. (**2009b**). "The forced radiation efficiency of finite size flat panels which are excited by incident sound," J. Acoust. Soc. Am. **126**, 694–702.

Davy, J. L. (**2009c**). "Predicting the sound insulation of single leaf walls—Extension of Cremer's model," J. Acoust. Soc. Am. **126**, 1871–1877.

Dupree, R. B. (**1981**). *Catalog of STC and IIC Ratings for Wall and Floor/Ceiling Assemblies—With TL and ISPL Data Plots* (Office of Noise Control, California Department of Health Services, Berkeley, CA).

Gradshteyn, I. S., and Ryzhik, I. M. (**1965**). *Table of Integrals, Series, and Products*, 4th ed., prepared by Yu. V. Geronimus and M. Yu. Tseytlin, translated from Russian by Scripta Technica Inc., edited by A. Jeffrey (Academic, New York).

Guigou-Carter, C., and Villot, M. (**2006**). "Analytical and experimental study of single frame double wall," Euronoise 2006, Tampere, Finland, 30 May–1 June.

Guigou-Carter, C., Villot, M., and Vernois, L. (**1998**). "Study of cavity ties for improving efficiency of acoustical lining," Euronoise, Munich, Germany.

Halliwell, R. E., Nightingale, T. R. T., Warnock, A. C. C., and Birta, J. A. (**1998**). "Gypsum board walls: Transmission loss data," Internal Report No. IRC-IR-761, Institute for Research in Construction, National Research Council of Canada, Ottawa, Canada.

Kang, H.-J. (**2000**). "Prediction of sound transmission loss through multilayered panels by using Gaussian distribution of directional incident energy," J. Acoust. Soc. Am. **107**, 1413–1420.

Lin, G.-F., and Garrelick, J. M. (**1977**). "Sound transmission through periodically framed parallel plates," J. Acoust. Soc. Am. **61**, 1014–1018.

NAHB (**1971**). "Acoustical manual—Apartment and home construction," NAHB Study 210-21, National Association of Home Builders Research Foundation Inc., Rockville, MD.

Northwood, T. D. (**1968**). "Transmission loss of plasterboard walls," Building Research Note BRN-66, Division of Building Research, National Research Council of Canada, Ottawa, Canada.

Poblet-Puig, J., Rodriguez-Ferran, A., Cuigou-Carter, C., and Villot, M. (**2006**). "Experimental and numerical characterization of metallic studs," Euronoise 2006, Tampere, Finland, 30 May–1 June.

Poblet-Puig, J., Rodriguez-Ferran, A., Cuigou-Carter, C., and Villot, M. (**2009**). "The role of studs in the sound transmission of double walls," Acta Acust. Acust. **95**, 555–567.

Quirt, J. D. (**1981**). "Measurements of the sound transmission loss of windows," Building Research Note BRN-172, Division of Building Research, National Research Council of Canada, Ottawa, Canada.

Quirt, J. D. (**1982**). "Sound transmission through windows. I. Single and double glazing," J. Acoust. Soc. Am. **72**, 834–844.

Schoenwald, S., Gerretsen, E., and Martin, H. J. (**2008**). "Prediction of the sound transmission loss of multilayered small sized elements," Proceedings of the Acoustics '08, Paris, France, 29 June–4 July.

Sewell, E. C. (**1970**). "Transmission of reverberant sound through a single-leaf partition surrounded by an infinite rigid baffle," J. Sound Vib. **12**, 21–32.

Vér, I. L., and Holmer, C. I. (**1971**). "Interaction of sound waves with solid structures," in *Noise and Vibration Control*, edited by L. L. Beranek (McGraw-Hill, New York), Chap. 11, pp. 270–361.

Villot, M., Guigou, C., and Gagliardini, L. (**2001**). "Predicting the acoustical radiation of finite size multi-layered structures by applying spatial windowing on infinite structures," J. Sound Vib. **245**, 433–455.

A mesh-free approach to acoustic scattering from multiple spheres nested inside a large sphere by using diagonal translation operators

Andrew J. Hesford[a] and Jeffrey P. Astheimer
Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14627

Leslie F. Greengard
Courant Institute of Mathematical Sciences, New York University, New York, New York 10012

Robert C. Waag
Department of Electrical and Computer Engineering and Department of Imaging Sciences, University of Rochester, Rochester, New York 14627

(Received 15 January 2009; revised 2 December 2009; accepted 3 December 2009)

A multiple-scattering approach is presented to compute the solution of the Helmholtz equation when a number of spherical scatterers are nested in the interior of an acoustically large enclosing sphere. The solution is represented in terms of partial-wave expansions, and a linear system of equations is derived to enforce continuity of pressure and normal particle velocity across all material interfaces. This approach yields high-order accuracy and avoids some of the difficulties encountered when using integral equations that apply to surfaces of arbitrary shape. Calculations are accelerated by using diagonal translation operators to compute the interactions between spheres when the operators are numerically stable. Numerical results are presented to demonstrate the accuracy and efficiency of the method. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277219]

PACS number(s): 43.58.Ta, 43.20.Fn, 43.80.Qf, 43.80.Ev [TDM] Pages: 850–861

I. INTRODUCTION

The efficient calculation of scattering from inhomogeneities embedded in an acoustically large domain is of interest in a variety of applications. This paper is restricted to the case of spherical inclusions contained within a large sphere, because this defines a rich class of simply described model problems. That is, the geometry and material properties require only the specification of the center locations, radii, sound speeds, densities, and attenuation parameters. Such geometries serve, for example, as useful models for inverse scattering algorithms.[1–4]

An application of particular interest is the estimation and correction of aberration through human tissue.[5–7] Tissue-mimicking phantoms are often used to study this problem under controlled circumstances that are clinically relevant. Spheres are attractive components of laboratory phantoms because the shapes are easy to make and have been shown to mimic the aberration behavior observed in human tissue.[6] A single enclosing medium keeps the sphere locations fixed relative to each other and mimics the properties of background tissue. A frequency-domain algorithm that computes scattering from spherical objects in the presence of an enclosing medium allows for the systematic evaluation of the performance of inverse scattering methods, with straightforward experimental validation of the forward model as well.

To be concrete, the modeling of ultrasound wave scattering by a spherical phantom 48 mm in diameter is considered. Inside this phantom are 12 smaller spheres designed to mimic the properties of human fat, skin, and muscle tissue. Laboratory measurements are conducted by immersing the phantom in water and transmitting temporal pulses with a center frequency of 2.5 MHz. At this frequency, the diameter of the outer sphere is approximately $80\lambda_0$, where λ_0 is the wavelength in water. Calculation of scattering by this large outer sphere with several smaller inclusions is a challenging problem that demands efficient methods.

An important approach to the scattering problem for general geometries in two or three dimensions is based on integral-equation formulations accelerated by the fast multipole method (FMM).[8–12] In full, hierarchical implementations, the FMM reduces the cost of application of the integral operator from $O(N^2)$ to $O(N \log N)$ or $O(N)$, where N is the number of discrete elements used to model the scatterer surfaces.[11] However, robust solvers that are high-order accurate require substantial meshing efforts, quadratures for singular integrals, and care to ensure that the integral equation is numerically well-conditioned. While FMM-based solvers have been applied to general scattering problems involving tens of millions of unknowns,[12,13] specialized techniques that take advantage of the scattering geometry to reduce the number of unknowns can improve efficiency and performance. In the case of spheres, partial-wave expansions[14,15] obviate many of the issues that complicate general-purpose solvers.

In the computational literature, the use of partial-wave expansions is often referred to as the T-matrix approach,[16–18] and is based simply on defining the linear relationship between incoming and outgoing waves, centered on each scatterer, that satisfy the physical interface conditions. These re-

[a]Author to whom correspondence should be addressed. Electronic mail: hestford@ece.rochester.edu

lationships are diagonal when fields are expressed in terms of spherical harmonic expansions. In an iterative solution process, one basically updates the outgoing waves by computing the incoming fields due to all of the other scattering expansions based on the current iteration. Since the "self-interaction" on each sphere is in diagonal form, the bulk of the work consists of converting the outgoing expansions emanating from each scatterer into incoming expansions centered on all of the others.

When large numbers of small scatterers are involved, FMM-based schemes are natural candidates. Previous research has produced effective schemes that borrow at least some FMM concepts, such as *T*-matrix approaches, multipole expansions, and efficient translation operators.[19–21] The methods implemented in these papers employ rotation operators and recurrence relations to efficiently evaluate the required translation operators. This "rotate-shift-rotate" approach is substantially more efficient than direct translation of spherical harmonics, but is less effective in the presence of a large enclosing sphere.

This paper presents a *T*-matrix method for solving the scattering problem using a formulation similar to that of Ref. 19. Boundary conditions are imposed on the surface of each scatterer, and a linear system is solved to calculate the scattered field, expressed through a partial-wave expansion centered on each sphere. In the spherical harmonic basis, the modes on a single sphere are uncoupled, so that the self-interaction of each sphere is simply diagonal. Unlike general-purpose integral-equation methods, no geometric meshing or singular integrals are encountered, and high-order accuracy is straightforward to achieve. Thus, the number of degrees of freedom required on each sphere is modest, dependent only on its acoustic size and, more weakly, on its separation distance from other scatterers.

Despite the broad similarities, this work differs from Ref. 19 in three ways. First, the translation of scattered fields between spheres is accomplished using diagonal forms[22] rather than the rotate-shift-rotate approach, reducing the cost of each translation from $O(p^3)$ to $O(p^2 \log p)$, where p is the degree of the partial-wave expansion.[23] Second, this paper examines the more complicated problem of scattering from small inclusions embedded in an acoustically large phantom which may be tens or hundreds of wavelengths in diameter. Third, the problem formulation uses interface conditions that match both the pressure and normal particle velocity across the inclusion boundaries. Ensuring the continuity of these physical parameters results in wave behavior that replicates the interaction of acoustic waves with soft tissue, parametrized in terms of sound speed, absorption, and density. The analysis of Refs. 19 and 20 relied on impedance boundary conditions, which provide a convenient mathematical formalism for describing the fast algorithm, but are not adequate as a physical model in the present setting.

Of these three, the introduction of an enclosing medium is the most interesting new feature, allowing the calculation of scattering due to actual tissue-mimicking phantoms realizable in the laboratory. The addition of the enclosing sphere requires augmentation of the system of equations with additional unknowns to account for the additional material interface. The number of unknowns is proportional to the surface area of the enclosing sphere measured in wavelengths and is typically much larger than the number of unknowns associated with each small scatterer.

As a technical matter, the diagonal translation operators employed in this method suffer from numerical instability if the distance over which the translation is carried out is too small. Under such conditions, the rotate-shift-rotate approach of Ref. 19 is prescribed in place of diagonal translations. This happens infrequently, when spheres are within a wavelength of each other or so. A detailed analysis of the issue is presented in this paper.

Finally, note that the existence of the diagonal form lies at the heart of the FMM for high-frequency scattering, as explained in detail in Ref. 23. However, the model problems of interest did not require a full implementation of the FMM. Computational costs of this method are dominated by interaction between the enclosing phantom and the small scatterers rather than interactions between the small scatterers themselves. If the number of inclusions in the model were increased by an order of magnitude or more, then the recommended approach would be to couple a scheme like that of Ref. 20 with the analytic tools discussed here.

II. THEORY

The problem of acoustic scattering from an arbitrary medium characterized by a spatially varying complex sound speed $c(\boldsymbol{r})$ is considered. The imaginary part of the sound speed describes the loss in the medium. If all pressure fields have frequency ω and an $e^{-i\omega t}$ time dependence, the total pressure field $\varphi(\boldsymbol{r})$ satisfies the Helmholtz equation

$$[\nabla^2 + k^2(\boldsymbol{r})]\varphi(\boldsymbol{r}) = 0, \tag{1}$$

where the wavenumber $k=\omega/c$. The total pressure is

$$\varphi(\boldsymbol{r}) = \xi(\boldsymbol{r}) + \varphi_s(\boldsymbol{r}), \tag{2}$$

where ξ is a known incident field and φ_s is the field scattered by the medium. The scattered field satisfies the Sommerfeld radiation condition,

$$\lim_{r\to\infty}[\partial_r\varphi_s - ik\varphi_s] = 0, \tag{3}$$

where $r=|\boldsymbol{r}|$. The boundary conditions between two distinct homogeneous regions require continuity of total pressure and normal particle velocity across the interface Γ separating the two regions. Hence,

$$\varphi_1(\boldsymbol{r}) = \varphi_2(\boldsymbol{r}), \quad \boldsymbol{r} \in \Gamma, \tag{4a}$$

$$\partial_n\varphi_1(\boldsymbol{r}) = \rho_r\partial_n\varphi_2(\boldsymbol{r}), \quad \boldsymbol{r} \in \Gamma, \tag{4b}$$

where $\rho_r=\rho_1/\rho_2$. In the i-th medium, the pressure is $\varphi_i(\boldsymbol{r})$ and the density is ρ_i. The direction n of differentiation in Eq. (4b) is that of the outward normal to Γ at the point $\boldsymbol{r}$.

A. Scattering from a single sphere

Scattering from a single homogeneous sphere of radius a embedded in a homogeneous background material is most simply described when the center of the sphere is placed at

the origin of a spherical coordinate system with coordinate vectors of the form $\boldsymbol{r}=(r,\theta,\phi)$. The boundary of the sphere is described by the set of points

$$\Gamma=\{\boldsymbol{r}:r=a\}. \tag{5}$$

With this description, the normal n coincides with the unit radial vector $\boldsymbol{e}_r$. The interior of the sphere supports a regular pressure wave $\varphi_2(\boldsymbol{r})$, while the exterior of the sphere supports a pressure $\varphi_1(\boldsymbol{r})=\xi(\boldsymbol{r})+\varphi_s(\boldsymbol{r})$ consisting of a regular incident wave and an outgoing, singular scattered field.

The use of spherical harmonic expansions for all field components makes evaluation of the boundary conditions straightforward. Because the incident and internal waves are both regular, they have regular harmonic expansions. The expansion of the incident field is limited to the outside of the sphere, while the expansion of the internal wave is valid only inside the sphere. The scattered field is defined outside of the sphere by a harmonic expansion that is singular at the origin. Taking into account these constraints, the harmonic expansions of the field components are

$$\xi(\boldsymbol{r})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\chi_{lm}j_l(k_1r)Y_{lm}(\theta,\phi),\quad r>a, \tag{6a}$$

$$\varphi_s(\boldsymbol{r})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}B_{lm}h_l(k_1r)Y_{lm}(\theta,\phi),\quad r>a, \tag{6b}$$

$$\varphi_2(\boldsymbol{r})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}C_{lm}j_l(k_2r)Y_{lm}(\theta,\phi),\quad r<a, \tag{6c}$$

where χ_{lm}, B_{lm}, and C_{lm} are coefficients of their respective harmonic expansions. The function $Y_{lm}(\theta,\phi)$ represents the spherical harmonic of degree l and order m, h_l is a spherical Hankel function of the first kind of order l, j_l is a spherical Bessel function of order l, and k_i is the wavenumber in the i-th medium. The boundary conditions (4) can be reduced to a series of well-known scattering quotients[14,15] Q_l mapping coefficients of the incident field (6a) to those of the scattered field (6b) without concern for the internal field (6c),

$$B_{lm}=\frac{\gamma j_l(k_1a)j_l'(k_2a)-j_l'(k_1a)j_l(k_2a)}{h_l'(k_1a)j_l(k_2a)-\gamma h_l(k_1a)j_l'(k_2a)}\chi_{lm}=Q_l\chi_{lm}, \tag{7}$$

where $\gamma=\rho_1k_2/\rho_2k_1$ and the prime indicates differentiation of the Bessel functions with respect to their whole arguments. Because each scattering quotient Q_l in Eq. (7) maps an incoming spherical harmonic coefficient to exactly one outgoing coefficient of the same degree and order, the scattering operator is diagonal. Calculation of the scattering of incoming waves is, therefore, as efficient as possible.

B. Multiple spheres

When N disjoint scattering spheres are considered, scattering quotients such as those in Eq. (7) can be computed with slight modifications. The field outside of all spheres can be represented as the superposition of a known incident field ξ and scattered fields radiating from each object,

$$\varphi_0(\boldsymbol{r})=\xi(\boldsymbol{r})+\sum_{i=1}^{N}\varphi_{s,i}(\boldsymbol{r}). \tag{8}$$

Let S_i denote the i-th sphere, which has center $\boldsymbol{c}_i$, wavenumber k_i, density ρ_i, and radius a_i. The regular wave inside the sphere S_i will be denoted φ_i. The scattered fields $\varphi_{s,i}$ are singular at the centers of the spheres. The incident, scattered, and internal fields have the respective expansions

$$\xi(\boldsymbol{r})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\chi_{lm}^{i}j_l(k_0r_i)Y_{lm}(\theta_i,\phi_i),\quad r_i>a_i, \tag{9a}$$

$$\varphi_{s,i}(\boldsymbol{r})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}B_{lm}^{i}h_l(k_0r_i)Y_{lm}(\theta_i,\phi_i),\quad r_i>a_i, \tag{9b}$$

$$\varphi_i(\boldsymbol{r})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}C_{lm}^{i}j_l(k_ir_i)Y_{lm}(\theta_i,\phi_i),\quad r_i<a_i, \tag{9c}$$

where (r_i,θ_i,ϕ_i) represents the spherical coordinates of the vector $\boldsymbol{r}_i=\boldsymbol{r}-\boldsymbol{c}_i$.

The scattered fields $\varphi_{s,j}$ are regular within sphere S_i whenever $i\neq j$. Therefore, within S_i, the field $\varphi_{s,j}$ has a spherical harmonic expansion that can be expressed as

$$\varphi_{s,j}(\boldsymbol{r})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}D_{lm}^{ij}j_l(k_ir_i)Y_{lm}(\theta_i,\phi_i),\quad r_i<a_i, \tag{10}$$

in which the regular wave coefficients D_{lm}^{ij} are translations of the singular wave coefficients B_{lm}^{j} from sphere S_j to S_i. The boundary conditions (4) can be used with the expansions (9) and (10) to relate the scattered-field coefficients B_{lm}^{i} to the translated coefficients D_{lm}^{ij} and the coefficients χ_{lm}^{i} of the incident field. Because spherical harmonics are orthogonal, the boundary conditions split into equations of the form

$$I_{lm}^{i}j_l(ka_i)+B_{lm}^{i}h_l(ka_i)=C_{lm}^{i}j_l(k_ia_i), \tag{11a}$$

$$I_{lm}^{i}j_l'(ka_i)+B_{lm}^{i}h_l'(ka_i)=\gamma_iC_{lm}^{i}j_l'(k_ia_i), \tag{11b}$$

where k is the wavenumber of the background medium and $\gamma_i=\rho_0k_i/\rho_ik$. The regular wave coefficients

$$I_{lm}^{i}=\chi_{lm}^{i}+\sum_{j\neq i}D_{lm}^{ij} \tag{12}$$

describe the total wave impinging on sphere S_i. This wave is the superposition of the incident field and the scattered fields from all spheres S_j distinct from S_i. The solution of the system (11) provides scattering quotients for the sphere S_i,

$$B_{lm}^{i}=\frac{\gamma_ij_l(ka_i)j_l'(k_ia_i)-j_l'(ka_i)j_l(k_ia_i)}{h_l'(ka_i)j_l(k_ia_i)-\gamma_ih_l(ka_i)j_l'(k_ia_i)}I_{lm}^{i}=Q_l^{i}I_{lm}^{i}. \tag{13}$$

The coefficients Q_l^{i} in Eq. (13) are equivalent to those in Eq. (7), except that the incident field coefficients χ_{lm}^{i} are replaced by the regular wave coefficients I_{lm}^{i} representing all fields incident upon S_i. In an iterative approach based on the T-matrix method, each iteration of the solution uses coefficients I_{lm}^{i} derived from the fields computed in the previous iteration to produce new scattered-field coefficients B_{lm}^{i}.

C. Translations of spherical harmonics

The keystone of the spherical scattering operator with coefficients defined in Eq. (13) is a translator that converts and recenters singular harmonic expansions of the form (9b) to regular wave expansions of the form (10). The translator is developed from the addition theorem for spherical harmonics.[24,25] This theorem relates a singular spherical harmonic to a linear combination of regular harmonics,

$$h_l(kr)Y_{lm}(\theta,\phi)=\sum_{l'=0}^{\infty}\sum_{m'=-l'}^{l'}\alpha_{l'm'lm}(k,\boldsymbol{r}'')j_{l'}(kr')Y_{l'm'}(\theta',\phi'), \tag{14}$$

where $\boldsymbol{r}=\boldsymbol{r}'+\boldsymbol{r}''$. The translator coefficients $\alpha_{l'm'lm}$ are defined by the expression[24,25]

$$\alpha_{l'm'lm}(k,\boldsymbol{r}'')=4\pi\sum_{l''=0}^{\infty}i^{l'+l''-l}\sum_{m''=-l''}^{l''}h_{l''}(kr'')Y_{l''m''}(\theta'',\phi'')\times\int_{S_0}d\hat{\boldsymbol{s}}Y_{lm}(\hat{\boldsymbol{s}})Y^*_{l'm'}(\hat{\boldsymbol{s}})Y^*_{l''m''}(\hat{\boldsymbol{s}}), \tag{15}$$

where S_0 is the unit sphere. However, direct evaluation of the integral in Eq. (15) results in an inefficient, dense operation.

An efficient diagonal form of the translator (15) is the basis for the high-frequency FMM.[22,25,26] This diagonal form may be derived from the addition theorem for the three-dimensional, free-space Green's function[27] expressing the field at $\boldsymbol{c}_j$ due to a point source located at $\boldsymbol{d}$ relative to the coordinate system centered on $\boldsymbol{c}_i$,

$$\frac{e^{ik|\boldsymbol{r}_{ji}+\boldsymbol{d}|}}{|\boldsymbol{r}_{ji}+\boldsymbol{d}|}=ik\sum_{l=0}^{\infty}(-1)^l(2l+1)j_l(kd)h_l(kr_{ji})P_l(\hat{\boldsymbol{d}}\cdot\hat{\boldsymbol{r}}_{ji}), \tag{16}$$

where $\boldsymbol{r}_{ji}=\boldsymbol{c}_j-\boldsymbol{c}_i$ and P_l is the Legendre polynomial of degree l. A general vector $\boldsymbol{x}$ has magnitude $x=|\boldsymbol{x}|$ and direction $\hat{\boldsymbol{x}}=\boldsymbol{x}/x$. Taking the Fourier transform of Eq. (16) with respect to $\boldsymbol{d}$ yields

$$\frac{e^{ik|\boldsymbol{r}_{ji}+\boldsymbol{d}|}}{|\boldsymbol{r}_{ji}+\boldsymbol{d}|}\approx\frac{ik}{4\pi}\int_{S_0}d\hat{\boldsymbol{s}}e^{ik\hat{\boldsymbol{s}}\cdot\boldsymbol{d}}\alpha_{ji}(\hat{\boldsymbol{s}}), \tag{17}$$

in which

$$\alpha_{ji}(\hat{\boldsymbol{s}})=\sum_{l=0}^{L}i^l(2l+1)P_l(\hat{\boldsymbol{r}}_{ji}\cdot\hat{\boldsymbol{s}})h_l(kr_{ji}) \tag{18}$$

are the coefficients of the diagonal translation operator.

The addition theorem (16) converges as $l\to\infty$ despite the tendency for $h_l(kr_{ji})$ to approach infinity because the Bessel functions $j_l(kd)$ vanish more rapidly. However, through the Fourier transform operation and the interchange of the integration and summation, this balancing effect is lost in the diagonal form (18). Therefore, the translator sum must be truncated to a maximum number of terms, L, large enough to cause the addition theorem (16) to converge within a desired tolerance, but not large enough to cause numerical inaccuracy in Eq. (18) due to a discrepancy between the magnitudes of h_0 and h_L that exceeds the dynamic range of the chosen finite-precision arithmetic. When the translation distance r_{ji} is large compared to the radii of the source and target spheres, the excess bandwidth formula prescribes a truncation point

$$L\approx ka+1.8d_0^{2/3}(ka)^{1/3}, \tag{19}$$

in which d_0 represents the desired number of digits of accuracy and a is the maximum of the source and target sphere radii.[11] Expression (19) has been found to give accurate results for electromagnetic Helmholtz problems.[28] The parameter L can also be computed numerically with essentially no cost[23] by analyzing the convergence of the addition theorem (16). This is necessary when the translation distance is not large compared to the radius a and the excess bandwidth formula (19) gives incorrect results.

The diagonal form of the translation operator (18) can be applied to general far-field signatures of wave fields that express the fields as coefficients of outgoing plane waves. A field $\varphi(\boldsymbol{r})$ has a far-field signature centered about an arbitrary point $\boldsymbol{a}$ that is defined by the relationship[11,26]

$$\hat{\varphi}_{\boldsymbol{a}}(\hat{\boldsymbol{s}})=\lim_{r\to\infty}k\tau e^{-ik\tau}\varphi(\boldsymbol{a}+\tau\hat{\boldsymbol{s}}). \tag{20}$$

Using the large-argument approximations for Hankel functions, the scattered field $\varphi_{s,i}$ expanded according to Eq. (9b) has a far-field signature given by[23]

$$\hat{\varphi}_{s,i}(\hat{\boldsymbol{s}})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}(-i)^{l+1}B^i_{lm}Y_{lm}(\hat{\boldsymbol{s}}). \tag{21}$$

The outgoing far-field signature (21) represents the radiation pattern, centered on $\boldsymbol{a}=\boldsymbol{c}_i$, of the sphere S_i induced by the incident field.

The translation operator (18) converts an outgoing far-field signature $\hat{\varphi}_{s,i}(\hat{\boldsymbol{s}})$ for a source sphere S_i into an incident plane-wave expansion that represents the field $\varphi_{s,i}$ everywhere within a target sphere S_j. The spheres S_i and S_j must be disjoint for the translation to be valid. The coefficients $\hat{\varphi}_{ji}(\hat{\boldsymbol{s}})$ of the incident plane-wave expansion are valid inside sphere S_j and are related to the outgoing far-field signature of sphere S_i by the expression

$$\hat{\varphi}_{ji}(\hat{\boldsymbol{s}})=\alpha_{ji}(\hat{\boldsymbol{s}})\hat{\varphi}_{s,i}(\hat{\boldsymbol{s}}). \tag{22}$$

The incoming plane-wave expansion is related to the translated spherical harmonic expansion coefficients D^{ji}_{lm} in Eq. (10) by the expression[23]

$$\hat{\varphi}_{ji}(\hat{\boldsymbol{s}})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}(-i)^{l+1}D^{ji}_{lm}Y_{lm}(\hat{\boldsymbol{s}}). \tag{23}$$

Hence, given the coefficients B^i_{lm} that describe a scattered field radiating outward from S_i, the translated regular wave coefficients D^{ji}_{lm} in expansion (10) can be found by converting the scattered field $\varphi_{s,i}$ to a far-field signature $\hat{\varphi}_{s,i}$, pointwise multiplying by the translator α_{ji}, and recovering the spherical harmonic expansion coefficients of the resulting incoming plane-wave expansion.

D. Spherical enclosing medium

The addition of an enclosing sphere in the description of scattering is straightforward when using partial-wave expansions. For interactions *between* interior spheres, no change to the system of equations is necessary. However, in the scattering quotients (13), the coefficients χ^i_{lm} of the known incident field must be replaced by the coefficients A_{lm} of a general spherical harmonic expansion

$$\varphi_e(\boldsymbol{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} j_l(kr) Y_{lm}(\theta,\phi), \quad r < a, \tag{24}$$

that is valid within the sphere S, which has radius a, density ρ, wavenumber k, and a center $\boldsymbol{c}$ that coincides with the origin of the coordinate system. The coefficients A^i_{lm} represent a re-expansion of the field φ_e about the center of the sphere S_i. The recentered field is given by

$$\varphi_e(\boldsymbol{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A^i_{lm} j_l(kr_i) Y_{lm}(\theta_i,\phi_i), \quad r_i < a_i, \tag{25}$$

where the coordinate variables follow the notation in Eq. (9). Since expansion (24) is valid everywhere inside the enclosing sphere S, and the internal sphere S_i is contained entirely within S, expansion (25) is valid everywhere inside S_i.

The addition of an enclosing sphere introduces two additional fields: the previously described internal field φ_e and an external scattered field $\varphi_{e,s}$. Accompanying these additional fields are two additional boundary conditions that govern the behavior of the fields at the boundary of the enclosing sphere. Those boundary conditions have the same form as boundary conditions (4) for the inner spheres. However, in this case, the external and internal fields are given by

$$\varphi_1(\boldsymbol{r}) = \xi(\boldsymbol{r}) + \varphi_{e,s}(\boldsymbol{r}), \tag{26a}$$

$$\varphi_2(\boldsymbol{r}) = \varphi_e(\boldsymbol{r}) + \sum_{i=1}^{N} \varphi_{s,i}(\boldsymbol{r}). \tag{26b}$$

The scattered fields $\varphi_{s,i}$ radiate outward from the spheres S_i contained within S. Hence, each field can be represented as a spherical harmonic expansion

$$\varphi_{s,i}(\boldsymbol{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B^{ei}_{lm} h_l(kr) Y_{lm}(\theta,\phi), \tag{27}$$

where the expansion is relative to the center of the enclosing sphere S. Since S contains S_i, expansion (27) is valid everywhere outside of S.

If the incident field has spherical harmonic expansion coefficients χ_{lm} and the external scattered field $\varphi_{e,s}$ has expansion coefficients B_{lm}, relation (26) can be employed in boundary conditions (4) with expansions (24) and (27) to provide two expressions:

$$A_{lm} = T_l \chi_{lm} + R_l \sum_{i=1}^{N} B^{ei}_{lm}, \tag{28a}$$

$$B_{lm} = T^e_l \sum_{i=1}^{N} B^{ei}_{lm} + R^e_l A_{lm}. \tag{28b}$$

The relation in Eq. (28a) describes the internal field φ_e as a transmission of the incident field superimposed with a reflection of the scattered fields produced by all internal spheres. Likewise, the expansion of the external field $\varphi_{e,s}$ is defined in Eq. (28b) to be the superposition of the transmission of the scattered fields produced by the internal spheres and the reflected internal field φ_e. The transmission and reflection operators appearing in Eq. (28) are given by

$$T_l = \frac{j_l'(k_0a)h_l(k_0a) + j_l(k_0a)h_l'(k_0a)}{\gamma j_l'(ka)h_l(k_0a) - j_l(ka)h_l'(k_0a)}, \tag{29a}$$

$$R_l = \frac{\gamma h_l(k_0a)h_l'(ka) - h_l(ka)h_l'(k_0a)}{\gamma j_l'(ka)h_l(k_0a) - j_l(ka)h_l'(k_0a)}, \tag{29b}$$

$$T^e_l = \frac{\gamma j_l(k_0a)h_l'(ka) - j_l'(k_0a)h_l(ka)}{j_l(k_0a)h_l'(k_0a) - j_l'(k_0a)h_l(k_0a)}, \tag{29c}$$

$$R^e_l = \frac{\gamma j_l'(ka)j_l(k_0a) - j_l(ka)j_l'(k_0a)}{j_l(k_0a)h_l'(k_0a) - j_l'(k_0a)h_l(k_0a)}, \tag{29d}$$

where $\gamma = \rho_0 k / \rho k_0$, with k_0 and ρ_0, respectively, equal to the wavenumber and density of the background medium.

The field $\varphi_{e,s}$ influences neither the incident field ξ nor the internal field φ_e. The scattered fields $\varphi_{s,i}$, in turn, are influenced only by each other and the internal field φ_e. Therefore, the solution of the scattering problem can proceed without regard to the external scattered field $\varphi_{e,s}$. Once the fields φ_e and $\varphi_{s,i}$ have been computed, the external scattered field $\varphi_{e,s}$ can be computed in a single pass if desired. This reduces the computational demand of the method.

The solution of Eq. (13) requires knowledge of the shifted coefficients A^i_{lm} given the known coefficients A_{lm}. Likewise, the solution of Eq. (28a) requires shifted coefficients B^{ei}_{lm} given the known coefficients B^i_{lm}. The diagonal translation operators provide an efficient mechanism for determining these shifted coefficients.[26] Given the incident plane-wave expansion $\hat{\varphi}_e$ centered at the origin, the shifted incident expansion $\hat{\varphi}'_e$ centered about the point $\boldsymbol{c}_i$ is given by

$$\hat{\varphi}'_e(\hat{\boldsymbol{s}}) = e^{ik\hat{\boldsymbol{s}}\cdot\boldsymbol{c}_i}\hat{\varphi}_e(\hat{\boldsymbol{s}}). \tag{30}$$

This expansion is valid everywhere inside of S_i, which is wholly contained within the enclosing sphere S. Similarly, the outgoing far-field pattern $\hat{\varphi}_{s,i}$ centered at $\boldsymbol{c}_i$ is related to an outgoing pattern $\hat{\varphi}'_{s,i}$ centered at the origin by

$$\hat{\varphi}'_{s,i}(\hat{\boldsymbol{s}}) = e^{-ik\hat{\boldsymbol{s}}\cdot\boldsymbol{c}_i}\hat{\varphi}_{s,i}(\hat{\boldsymbol{s}}). \tag{31}$$

The shifted far-field signature is valid outside of the enclosing sphere. The spherical harmonic expansion coefficients corresponding to the shifted incoming plane-wave expansion (30) and the outgoing far-field pattern (31) are the coefficients A^i_{lm} and B^{ei}_{lm}, respectively.

The spherical harmonic expansions of each field component involve infinitely many terms. However, for practical computation, the expansions must be truncated. The trunca-

tion point is determined according to an excess bandwidth formula (19). This ensures that the fields radiating from or incident upon each field are sampled with sufficient density to accurately represent all relevant harmonic content.

For simplicity, the expansions for all internal spheres can be truncated to the same degree. However, the large outer sphere typically requires expansions of much higher degrees for accurate representation of the fields. Since the internal scattered fields $\varphi_{s,i}$ have lower degrees, their far-field patterns $\hat{\varphi}_{s,i}$ can be sampled more coarsely than the far-field patterns $\hat{\varphi}_e$ and $\hat{\xi}$. Consequently, when the patterns $\hat{\varphi}_{s,i}$ are translated to the center of the enclosing sphere, they must be more finely sampled before solving Eq. (28a). In addition, when the internal field φ_e is translated to each internal sphere for use in Eq. (13), the far-field pattern $\hat{\varphi}_e$ must be decimated after translation to match the coarser sampling grid used to define $\hat{\varphi}_{s,i}$. The methods used for decimation and interpolation are detailed in Refs. 11 and 23.

E. Iterative solution

In the absence of an enclosing sphere, equations of the form (13) for each sphere S_i, $i=1,\ldots,N$, form a system of equations that describe the relationships between all field quantities. However, these relationships relate unknown coefficients B_{lm}^i to the known incident field coefficients χ_{lm}^i and the unknown coefficients D_{lm}^{ij} derived from unknown coefficients B_{lm}^j. To produce a system of equations suitable for iterative inversions, Eq. (13) must be rearranged to form

$$B_{lm}^i - Q_l^i \sum_{j\neq i} D_{lm}^{ij} = Q_l^i \chi_{lm}^i. \tag{32}$$

The resulting modified equations define a scattering map that relates the unknown coefficients B_{lm}^i for each sphere to the known coefficients of the incident field at each sphere. The scattering map is composed of both the scattering quotients Q_{lm}^i and the diagonal operators that convert outgoing coefficients B_{lm}^j into incoming coefficients D_{lm}^{ij}. The system of equations that describe the scattering map are invertible using iterative methods such as GMRES.[29] Convergence of the iterative scheme is monitored by tracking the residual norm associated with the solution and the scattering map.

When an enclosing sphere surrounds all inner scatterers, the equations governing the fields at the boundary of each inner scatterer are reordered to form

$$B_{lm}^i - Q_l^i \left[A_{lm}^i - \sum_{j\neq i} D_{lm}^{ij} \right] = 0. \tag{33}$$

This relationship is again derived from Eq. (13) with the previously noted exception that the coefficients χ_{lm}^i of the incident field are replaced with the coefficients A_{lm}^i of the regular field inside the enclosing sphere. The coefficients A_{lm}^i centered on S_i are related to the unknown regular field coefficients A_{lm} centered on the enclosing sphere through the diagonal translation operators (30). These unknown coefficients are related to the known incident field coefficients χ_{lm} by rearranging Eq. (28a) to yield

$$A_{lm} - R_l \sum_{i=1}^{N} B_{lm}^{ei} = T_l \chi_{lm}, \tag{34}$$

where the coefficients B_{lm}^{ei} centered on the enclosing sphere are related to the coefficients B_{lm}^i centered on S_i by Eq. (31). The entire system of equations consisting of Eqs. (33) and (34) form a scattering map that relates all unknown quantities to the known incident field coefficients using the three diagonal translation operators and the scattering quotients Q_{lm}. This map may be inverted using GMRES in a manner similar to that for the case involving no enclosing sphere.

F. Comparison to the fast multipole method

The fundamental aspects of the fast multipole method are efficient diagonal forms for translation operators and aggregation and distribution procedures. The aggregation procedure collects far-field signatures of individual scattering elements into a source group signature that is translated using diagonal forms to a target group of scattering elements. The translated signature is then distributed to individual elements within the target group. Modern, hierarchical implementations further reduce computational demand.[11,23] The name "fast multipole method" for high-frequency scattering reflects the use of diagonal translation operators, which are Fourier transforms of translation operators for multipole (spherical harmonic) expansions.

The method presented to compute scattering from spheres uses the same diagonal translation operators employed in the FMM. However, because the number of spheres considered in the numerical studies was not large, aggregation and distribution of fields for groups of scattering elements were unnecessary. Consequently, application of the name "fast multipole method" to the presented approach is not appropriate. Instead, the described method is called the "sphere multipole method" to emphasize the use of spheres as scattering elements and the use of spherical harmonic (multipole) expansions to represent fields.

III. NUMERICAL RESULTS

Three examples are presented to illustrate the accuracy and efficiency of the method. The first example consists of simple problems involving two spheres at various distances that demonstrate the accuracy of translation even for short distances. The second example is a larger problem involving 12 spheres with differing material properties that exhibits increased efficiency over an algorithm using dense translations. In the third example, the 12 spheres are embedded in an enclosing sphere to form a geometry resembling a tissue-mimicking phantom that is used for aberration estimation.

A. Two spheres

Two tests each considered scattering from two spheres immersed in water. The sound speed in the water was 1509 m/s and the density was 997 kg/m^3. The sound speed in each sphere was 1613 m/s with an attenuation slope of 1.61 dB/cm MHz and a density equal to that of water. The two spheres were separated along the y axis and the incident

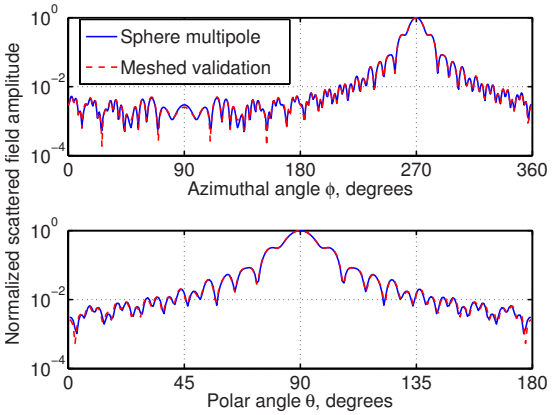


FIG. 1. (Color online) Azimuthal and polar scattering patterns for two spheres, each with a 2 mm radius, separated by 10 mm center to center. The excitation frequency was 2.5 MHz.

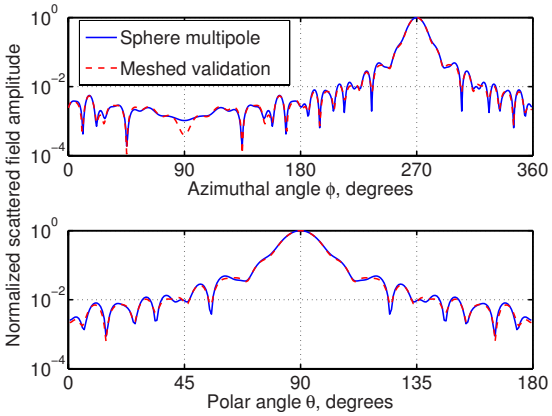


FIG. 2. (Color online) Azimuthal and polar scattering patterns for two spheres, with respective radii 1 and 2 mm, separated by 3.2 mm center to center. The excitation frequency was 2.5 MHz.

plane wave was traveling in the $-y$ direction at a frequency of 2.5 MHz. The results were compared to simulations using a volume integral-equation method that requires a meshed representation of the scattering geometry.

In the first calculation, two spheres with 2 mm radii were separated by 10 mm. The scattering patterns of the two spheres are shown in Fig. 1 along with the results from the reference simulation. The large separation between the spheres ensures that the scattering pattern is accurately computed. The mean-squared error (MSE) between the fields computed by the fast multiple-scattering method and the integral-equation method is defined as

$$\mathrm{MSE} = \left[\int_{\Omega} d\hat{s} |\varphi_s(\hat{s}) - \varphi_{s,r}(\hat{s})|^2 \Big/ \int_{\Omega} d\hat{s} |\varphi_{s,r}(\hat{s})|^2 \right]^{1/2}, \tag{35}$$

where Ω is the unit sphere, φ_s is the scattered field computed with the fast multiple-scattering method, and $\varphi_{s,r}$ is the reference scattered field computed with the integral-equation method. The MSE for fields shown in Fig. 1 is 2.4%. This error reflects inaccuracies due to truncation of the field components and diagonal translation operators in the fast method in addition to discretization and numerical integration inaccuracies in the integral-equation method. If the fast method is used to compute a solution neglecting interactions between spheres (by forcing the translators to vanish in all directions), the MSE increases to 66.6%. This large increase suggests that the interactions between spheres are significant to the overall solution. The low error of the complete solution therefore provides evidence that the translators are accurate.

Two spheres with different radii and a shorter translation distance were also considered. One sphere had a radius of 2 mm, while the radius of the other sphere was 1 mm. These two spheres were separated by 3.2 mm along the y axis. Strictly speaking, this is a violation of the constraints on the use of the diagonal form that is inaccurate for scatterers in proximity. The excess bandwidth formula (19) does not correctly predict the truncation point L required for accurate evaluation of the diagonal translator, and proper selection of L according to the convergence of the addition theorem (16) results in an unstable translator. The scattered field computed when the diagonal translator is truncated according to the excess bandwidth formula is shown in Fig. 2 to reasonably agree with the integral-equation solution. The MSE of the solution is 5.0%. When interactions between spheres are disregarded, the MSE increases to 38.1%. While this example is not as strongly influenced by interactions between spheres as the previous example, these interactions are still a significant part of the overall solution.

B. Twelve spheres

A more challenging test was posed by the geometry shown in Fig. 3. Twelve spheres, each composed of one of three materials that mimic the properties of human tissues, were embedded in a larger, contrasting sphere. The tissues and their properties are listed in Table I. The entire phantom was immersed in a background with the acoustic properties of water. The characteristics of each of the inner spheres are listed in Table II.

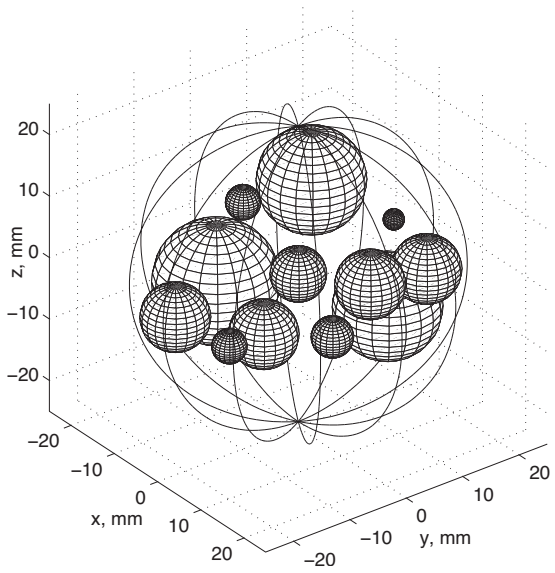


FIG. 3. An arrangement of 12 tissue-mimicking spheres. A spherical enclosing boundary is shown, although some tests did not consider the enclosing sphere.

TABLE I. Material properties of spheres designed to mimic human tissue.

Tissue	Sound speed (m/s)	Density (kg/m^3)	Absorption (dB/cm MHz)
Water	1509.0	997.0	0.00
Fat	1478.0	950.0	0.52
Muscle	1547.0	1050.0	0.91
Skin	1613.0	1120.0	1.61

Considering the 12 inner spheres in Fig. 3 without an enclosing sphere significantly reduces the complexity of the scattering problem. However, even without an enclosing sphere, the direct method used to validate the sphere multipole method was incapable of solving a full-scale problem with the desired excitation frequency using the available desktop computing resources. The validation method uses a matrix solver based on spherical harmonic expansions and non-diagonal harmonic translators. To produce validation results, the dimensions of the phantom were scaled uniformly by a factor of 1/3. Hence, for a sphere S_i with radius a_i and center location $\boldsymbol{c}_i$, the scaled sphere S_i' had radius $a_i'=a_i/3$ and center location $\boldsymbol{c}_i'=\boldsymbol{c}_i/3$. The scattering patterns shown in Fig. 4 were produced by the scaled spheres due to an incident plane wave oscillating at 2.5 MHz and traveling in the $+x$ direction. The solution was obtained in 9 s on a four-processor, shared-memory computer system using less than 200 Mbytes of memory. The MSE of the solution shown in Fig. 4 is 3.5%.

C. An enclosing sphere

Addition of the enclosing sphere surrounding the smaller tissue-mimicking spheres shown in Fig. 3 results in an accurate model of a laboratory phantom. Scattering from this structure at a range of frequencies in the neighborhood of 2.5 MHz is of interest. Simulated results are useful for equipment calibration and validation of laboratory measurements. Furthermore, the numerical results are useful for aberration correction algorithms.

When the enclosing sphere is acoustically large, the interior field must be described by a spherical harmonic expansion with a high degree. This increases the computational load of both the sphere multipole and validation methods. While the sphere multipole method is fully capable of solving the scattering problem for incident waves with frequencies in excess of 2.5 MHz, the validation code was incapable of simulating such frequencies with available resources. Thus, for validation purposes, the geometry was again reduced by a uniform scaling factor.

TABLE II. Characteristics of the spheres in the tissue-mimicking phantom.

Radius (mm)	Center (mm)			Tissue
	x	y	z	
4.0	0.0	0.0	0.0	Fat
5.0	14.0	2.0	4.0	Skin
5.0	5.0	−10.0	−4.0	Fat
3.0	17.0	−7.0	0.0	Fat
7.8	−10.0	10.0	7.2	Muscle
7.8	5.0	12.0	−7.2	Muscle
5.0	14.0	12.0	3.0	Muscle
5.0	−5.0	−18.0	−3.0	Skin
2.5	7.5	−18.0	−2.0	Skin
1.5	−4.0	20.0	0.0	Skin
2.5	−18.0	4.0	2.0	Skin
9.1	−12.5	−5.0	−5.2	Muscle

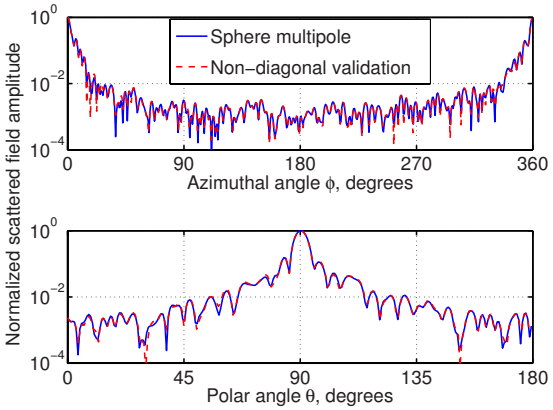


FIG. 4. (Color online) Azimuthal and polar scattering patterns of the 12-sphere acoustic phantom (with dimensions scaled by 1/3). The excitation frequency was 2.5 MHz.

The 12 spheres shown in Fig. 3 were enclosed in a contrasting sphere with a radius of 24 mm. The material of the enclosing sphere had a sound speed of 1570 m/s, an attenuation slope of 0.3 dB/cm MHz, and a density of 970 kg/m^3. The dimensions of the phantom were reduced by a factor of 3 to facilitate computation of validation results. An incident plane wave traveling along the $+x$ direction and a frequency of 2.5 MHz were used to excite the scaled geometry. The scattering patterns produced by the tissue-mimicking phantom in the presence of this incident wave are shown in Fig. 5 with a MSE of 6.5%. Running on four processing cores, the sphere multipole method computed a solution in 35 s using 200 Mbytes of system memory.

The sphere multipole method has also been used to simulate the enclosed, 12-sphere geometry at full scale with a much higher incident frequency. With an incident fre-

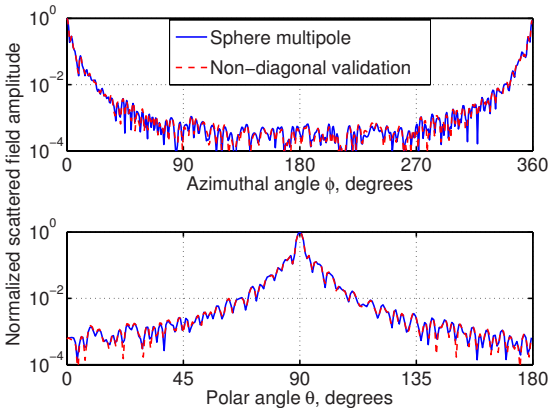


FIG. 5. (Color online) Azimuthal and polar scatter patterns for a 12-sphere acoustic phantom embedded in a contrasting enclosing sphere (with dimensions scaled by 1/3). The excitation frequency was 2.5 MHz.

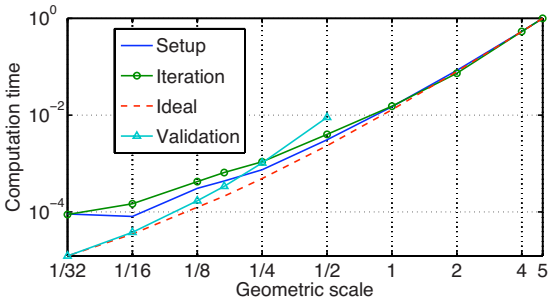


FIG. 6. (Color online) Scaling of the computation times of the sphere multipole method and a direct validation technique as a function of the scaling of the geometry depicted in Fig. 3.

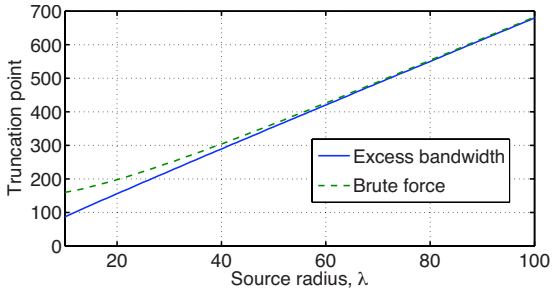


FIG. 7. (Color online) The number of terms required for desired accuracy below 10^{-6} in the addition theorem for a translation $\boldsymbol{r}_{ji}=1.1a\hat{x}$ and a source location $\boldsymbol{d}=a\hat{x}$.

quency of 7.5 MHz, the sphere multipole method was able to compute a solution in 15 min with a memory consumption less than 1 Gbyte. No validation result was available for a problem of this size. These results indicate that the described method will be useful in simulating propagation through the phantom at a range of frequencies that matches those used in laboratory equipment.

D. Scaling studies

The programming environments and strategies of the implementation of the sphere multipole method and the validation codes are different. Direct comparisons between the run times of the two methods are, therefore, not meaningful. Instead, a scaling study was employed to establish the efficiency of the sphere multipole method. To perform such a study, the dimensions of the 12-sphere phantom pictured in Fig. 3 and described in Table II were reduced to various scales. The enclosing sphere was omitted because of the limitations of the validation code. Altering the geometric scale of the problem is similar to scaling the excitation frequency by the reciprocal of the geometry scaling factor. Thus, these results also describe how the algorithm scales with increasing incident frequency. However, because loss in the spheres is specified as an attenuation slope, scaling the excitation frequency alters the attenuation observed in the spheres. While this will not alter the periteration or setup times of the algorithm, different attenuation values can affect the number of iterations required for convergence.

The results of the scaling study are shown in Fig. 6. In the sphere multipole method, setup-related overhead does not depend on the number of iterations required to produce a solution. This overhead is primarily dominated by calculation of the translation operators. The iteration time reported is normalized by the minimum number of iterations required to produce a solution with a backward error less than 10^{-6}. The number of iterations increased slightly as the scale of the problem was increased. The "ideal" line shown in the figure scales as $O(L^3)$, where L is the maximum degree used in spherical harmonic expansions. The number of spheres was fixed in these studies and does not, therefore, affect the asymptotic behavior of the method.

The scaling of the direct validation method is also shown in Fig. 6. The excessively large demands on memory and computer time prevented calculation of validation results at geometric scales greater than 0.5. The validation times have been normalized to ensure that the time observed at the minimum scale $(1/32)$ agrees with the ideal prediction of the fast method. The figure shows that the asymptotic behaviors of the fast method and the direct method diverge rapidly. The validation method solves the same system of boundary equations solved by the sphere multipole method. However, the direct method uses dense spherical harmonic translation matrices to translate $O(L^2)$ singular wave coefficients to $O(L^2)$ regular wave coefficients. Because these matrices contain $O(L^4)$ elements, the direct method can scale no more efficiently than $O(L^4)$.

E. Truncation of the translation operator

Analyses of truncation error in the addition theorem (16) consider general distributions of point sources throughout a spherical source region of radius d. As discussed below, such considerations are not generally applicable to an analysis of translation error for fields scattered by homogeneous spheres because the harmonic bandwidth of the scattered fields is lower than that of the singular fields produced by point sources. Nevertheless, analysis of truncation error in the addition theorem illuminates worst-case concerns associated with the application of diagonal translation operators over short distances.

The addition theorem (16) is always guaranteed to converge when $d<r_{ji}$ because the Bessel term $j_l(kd)$ vanishes faster than $h_l(kr_{ji})$ increases in magnitude as l increases. However, convergence is slowest when $d=a$ and $\boldsymbol{d}$ is parallel or antiparallel to $\boldsymbol{r}_{ji}$, resulting in a Bessel term $j_l(kd)$ that vanishes slowly and a Legendre polynomial $P_l(\hat{\boldsymbol{d}}\cdot\hat{\boldsymbol{r}}_{ji})$ with maximum absolute value. The excess bandwidth formula provides a suitable estimate for the truncation point L when the translation distance r_{ji} is large relative to the source radius a. When this condition is not satisfied, a brute-force approach can be employed to directly evaluate the minimum truncation point L that ensures agreement between both sides of the addition theorem within a desired tolerance. A comparison of the truncation points computed using the excess bandwidth formula and brute-force evaluation for a translation $\boldsymbol{r}_{ji}=1.1a\hat{x}$ and point-source location $\boldsymbol{d}=a\hat{x}$ is shown in Fig. 7 for a desired tolerance of 10^{-6}.

For short translation distances, the improved estimate of L obtained from brute-force analysis may be required to compute accurately the diagonal translation operator (18). However, when r_{ji} is small, the translator may be unstable

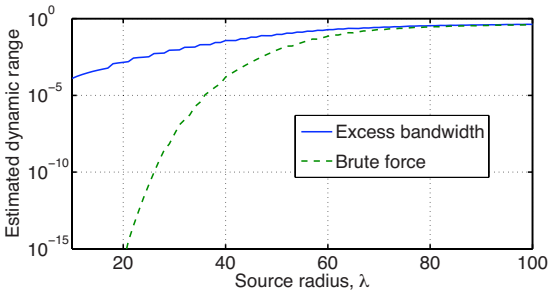


FIG. 8. (Color online) Estimate of the dynamic range required to accurately compute the diagonal translation operator for a translation $\boldsymbol{r}_{ji}=1.1a\hat{x}$ and a source location $\boldsymbol{d}=a\hat{x}$.

with finite-precision arithmetic. This is because the magnitudes of the Hankel functions $h_l(kr_{ji})$ increase rapidly and monotonically when $l>kr_{ji}$. In such cases, the magnitude $|h_L(kr_{ji})|$ of the high-order term may eclipse the magnitude $|h_0(kr_{ji})|$ of the low-order term in finite-precision arithmetic. An estimate of the required dynamic range for the aforementioned translation is shown in Fig. 8 as the ratio $|h_0(kr_{ji})/h_L(kr_{ji})|$ for a desired error below 10^{-6}. When the ratio is smaller than the minimum representable number in finite-precision arithmetic, the translator sum (18) is not computed accurately.

IV. DISCUSSION

Three aspects of the described method merit special comment. The first is the truncation of and error in the diagonal translation operator. Although this topic has been discussed in considerable detail (for example, in Ref. 11), the context here is different. The number of terms required in the expansion is determined by the accuracy of the representation (9) and not by point source considerations in the standard FMM. The second issue is the computational efficiency of this method. The asymptotic complexity of the sphere multipole method depends primarily on the number of scattering spheres and their acoustic sizes. The third aspect of the sphere multipole method is the suitability of the algorithm to parallel computing.

A. Truncation of the diagonal translation operator

The worst-case estimates provided in Sec. III E assume that the scattered field radiating from a source sphere behaves as a point source located on the sphere surface. This contributes terms of high degree to the harmonic expansion of the field, but the excess bandwidth formula (19) assumes that the source and target spheres are well-separated to ensure that the source singularity is far from the target. Because high-degree harmonics are not significant in the far field, the excess bandwidth formula underestimates the harmonic content of singular source fields and prescribes a lower truncation point for the diagonal translator than necessary for near-field translations. Therefore, premature truncation of the translator according to the excess bandwidth formula results in inaccurate computation of the contributions due to high-degree components.

However, scattering from homogeneous spheres of finite radius are approximately band-limited since the scattering coefficients (13) tend to vanish rapidly as $l\to\infty$. Furthermore, the maximum harmonic degrees of all scattered fields are known as a consequence of harmonic expansions being used with diagonal scattering operators at the surface of each sphere. As discussed in Ref. 23, the harmonic bandwidth of a field is the sole determinant of the number of terms required in a diagonal translator for that field. Stability of the diagonal form is likewise dependent on the required number of terms and the translation distance, and is easily discerned from the ratio of the magnitudes of the lowest-order and highest-order terms. Even when worst-case error analysis (based on singular field assumptions) suggests that diagonal forms will be unstable for a particular source size and translation distance, the harmonic bandwidth of the source field may be sufficiently low to allow stable diagonal translation with a reduced number of terms. This phenomenon is shown in Fig. 2, where the excess bandwidth formula was used to determine the translator truncation points for short distances and the results are still accurately computed.

Because the harmonic bandwidths of all scattered fields are known during computation, and because the stability of diagonal forms is easily assessed based on the required numbers of terms and the translation distances, the cost to implement runtime checks to determine whether a given field may be accurately translated using a precomputed diagonal form is insignificant. If a check indicates that the diagonal form is not suitable for a particular translation, a rotate-shift-rotate approach [30–32] is recommended as an alternative. This technique allows on-demand, efficient computation without the stability concerns accompanying diagonal forms. While the rotate-shift-rotate approach is less efficient than diagonal translation, the method is only required for near-neighbor translations and should not dominate the overall computational cost of the method. In the examples presented earlier, instability of the diagonal forms was not an issue, and the rotate-shift-rotate approach was not required.

B. Computational efficiency

The computational cost of the described method is heavily dependent on the maximum number of terms used in spherical harmonic expansions of fields. For simplicity, the expansion degree for each of the scattering spheres is assumed to be L, obtained either numerically or from Eq. (19). For a spherical harmonic expansion of degree L, the total number of terms in the expansion is $O(L^2)$. When the spherical harmonic expansions are transformed to far-field radiation patterns, the patterns must be sampled at $O(L^2)$ points on the unit sphere.

If scattering from N spheres is to be determined, the scattered field from each sphere must be translated to every other sphere. The total number of translations required is thus $O(N^2)$ and each translator sum (18) must be computed for $O(L^2)$ points on the unit sphere. Evaluation of the translator at a single point can be completed in $O(L)$ time. Hence,

in the simplest implementation, the overall cost for constructing the $O(N^2)$ translators is $O(N^2L^3)$. Storage of these translators requires $O(N^2L^2)$ memory.

Once the translators are constructed, their application requires only the pointwise multiplication of the far-field pattern of a field with the translator. These multiplications can be performed in $O(L^2)$ time. For a single iteration in the calculation of scattering, each translation is performed once. Hence, for each iteration, the total cost of translating far-field signatures between spheres is $O(N^2L^2)$. In contrast, each dense translation requires $O(L^4)$ time, for a total cost of $O(N^2L^4)$, while translations using a rotate-shift-rotate approach require $O(L^3)$ work for a total cost of $O(N^2L^3)$.

Application of the diagonal translation operator to interactions between scattering spheres also requires transformations between spherical harmonics and far-field signatures. With $O(L^2)$ coefficients, the far-field signature must be sampled at $O(L^2)$ distinct points on the unit sphere. Therefore, a direct evaluation of the signature (21) requires $O(L^4)$ time. However, using separation of variables,[23] the far-field signature can be evaluated in $O(L^3)$ time. A divide-and-conquer scheme has also been implemented[33] to further reduce the complexity to $O([L \log L]^2)$. A similar acceleration technique can convert an incoming plane-wave expansion (23) into a spherical harmonic expansion in $O([L \log L]^2)$ time. These transformations must be performed twice for each sphere: once to convert the field scattered from the sphere from a spherical harmonic expansion to a far-field signature and once again to convert the incoming plane-wave expansion from other spheres (which are superimposed) into a spherical harmonic expansion for scattering. Hence, the transformations can be evaluated with a complexity of $O(N[L \log L]^2)$ per iteration.

Because spherical scattering operators are diagonal, the scattering from each sphere is performed in $O(L^2)$ computer time for $O(L^2)$ spherical harmonic coefficients. A scattering operator must be applied to incoming fields centered on sphere so a total of N scattering operations must be performed. The overall cost for applying scattering operators defined in Eq. (13) is $O(NL^2)$ per iteration. This is not significant compared to the complexity of the translation operators and the spherical harmonic transformations. Hence, the overall complexity of a single iteration of the sphere multipole method can be reduced to $O(N^2L^2)+O(N[L \log L]^2)$.

If the number N of spheres is increased while the size of each sphere is kept constant, the algorithm as presented scales as $O(N^2)$. Likewise, if N is kept constant while the size of each sphere is increased (for example, by increasing the incident frequency), the algorithm scales as $O([L \log L]^2)$. The N^2 scaling issue can be overcome by a full implementation of the (hierarchical) FMM,[10,11,20,23] reducing the cost to $O(N \log N)$ without the need to adopt an integral-equation formulation. The $[L \log L]^2$ scaling is more problematic and, for sufficiently large spheres, optimal schemes would require full integral equation approaches with FMM acceleration applied to high-order surface meshes. Nevertheless, for the medium-sized problems that are addressed in this study, the scheme of coupling "spectrally accurate" representations in a spherical harmonic basis with efficient translation operators that generate sphere-to-sphere interactions is sufficient. For scattering spheres with very large radii (in excess of several hundred wavelengths), the additional overhead associated with a full FMM implementation is insignificant compared to the cost of the transformation between spherical harmonic expansions and far-field signatures.

The memory requirements of the fast algorithm are dominated by storage of the diagonal translators and all field coefficients. There are $O(NL^2)$ field coefficients associated with fields scattered by N spheres. The diagonal translators require storage of $O(N^2L^2)$ values for $O(N^2)$ translators. Hence, the overall storage complexity of the algorithm scales as $O(N^2L^2)$.

C. Parallel computing

The sphere multipole method can easily be made parallel with little or no synchronization overhead by using multiple threads on shared-memory systems. Both the scattering quotients in Eq. (13) and the efficient transformations between spherical harmonic expansions and outgoing far-field signatures[23] can be applied to fields radiating from each scattering sphere independently with no synchronization. In addition, translations of multipole fields between spheres can be computed in parallel, but some synchronization is required when the translated fields are used to augment the incoming wave expansion for each sphere (12) to prevent conflicting memory access between multiple threads. Distributed-memory implementations of the FMM have been developed,[13,34,35] but care must be taken to prevent the communication demands from overshadowing the advantages of additional processors. When computing scattering at multiple frequencies for pulse excitations, a more straightforward parallel approach evaluates the solution at each desired frequency on a distinct processor. Virtually no communication or synchronization is required in such an approach.

V. CONCLUSION

A method that couples analytic expressions for scattering from a sphere with diagonal translation operators has been presented to solve the acoustic scattering problem for multiple spheres that can be enclosed within a larger sphere. Like the method of Ref. 19, a multiple-scattering formalism is used that avoids the need for discrete representation of scatterers and the evaluation of singular integrals as is typical in general-purpose integral-equation methods (with or without fast multipole acceleration). The use of diagonal translation here reduces the computational complexity involved in mapping the outgoing wave from each source sphere to each target sphere in a spherical harmonic basis. Numerical results establish the accuracy and efficiency of the new method.

ACKNOWLEDGMENTS

Jason C. Tillett, David P. Duncan, and Gheorghe Salahura are thanked for helpful discussions, suggestions, and comments about material in this paper. This research was funded in part by NIH Grant No. EB 00280, U.S. Department of Energy Contract No. DEFG0288ER25053, and the

University of Rochester Diagnostic Ultrasound Research Laboratory Industrial Associates.

[1]W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," IEEE Trans. Med. Imaging **9**, 218–225 (1990).
[2]G. L. Wang, W. C. Chew, T. J. Cui, A. A. Aydiner, D. L. Wright, and D. V. Smith, "3D near-to-surface conductivity reconstruction by inversion of VETEM data using the distorted Born iterative method," Inverse Probl. **20**, S195–S216 (2004).
[3]T. D. Mast, A. I. Nachman, and R. C. Waag, "Focussing and imaging using eigenfunctions of the scattering operator," J. Acoust. Soc. Am. **102**, 715–726 (1997).
[4]F. Lin, T. K. Varslot, J. P. Astheimer, and R. C. Waag, "An eigenfunction method for reconstruction of large-scale and high-contrast objects," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **54**, 1316–1332 (2007).
[5]T. D. Mast, L. M. Hinkelman, L. A. Metlay, M. J. Orr, and R. C. Waag, "Simulation of ultrasonic pulse propagation, distortion, and attenuation in the human chest wall," J. Acoust. Soc. Am. **106**, 3665–3677 (1999).
[6]J. C. Lacefield, W. C. Pilkington, and R. C. Waag, "Distributed aberrators for emulation of pulse distortion by abdominal wall," ARLO **3**, 47–52 (2002).
[7]M. Tabei, T. D. Mast, and R. C. Waag, "Simulation of ultrasonic focus aberration and correction through human tissue," J. Acoust. Soc. Am. **113**, 1166–1176 (2003).
[8]L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations," J. Comput. Phys. **73**, 325–348 (1987).
[9]V. Rokhlin, "Rapid solution of integral equations of scattering theory in two dimensions," J. Comput. Phys. **86**, 414–439 (1990).
[10]C. C. Lu and W. C. Chew, "A fast algorithm for solving hybrid integral equations," IEE Proc., Part H: Microwaves, Antennas Propag. **140**, 455–460 (1993).
[11]*Fast and Efficient Algorithms in Computational Electromagnetics*, edited by W. C. Chew, J. Jin, E. Michielssen, and J. Song (Artech House, Boston, 2001).
[12]L. Gürel and Ö. Ergül, "Fast and accurate solutions of extremely large integral-equation problems discretised with tens of millions of unknowns," Electron. Lett. **43**, 499–500 (2007).
[13]S. Velamparambil and W. C. Chew, "Analysis and performance of a distributed memory multilevel fast multipole algorithm," IEEE Trans. Antennas Propag. **53**, 2719–2727 (2005).
[14]A. Lowan, P. Morse, H. Feshbach, and M. Lax, "Scattering and radiation from circular cylinders and spheres—Tables of amplitudes and phase angles," Technical Report No. 62.1R, U.S. Navy Department, Office of Research and Inventions, Arlington, VA, 1945.
[15]P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
[16]P. C. Waterman and R. Truell, "Multiple scattering of waves," J. Math. Phys. **2**, 512–537 (1961).
[17]V. K. Varadan and V. V. Varadan, *Acoustic, Electromagnetic and Elastic Wave Scattering: Focus on the T-Matrix Approach* (Pergamon, New York, 1980).
[18]S. Koc and W. C. Chew, "Calculation of acoustical scattering from a cluster of scatterers," J. Acoust. Soc. Am. **103**, 721–734 (1998).
[19]N. A. Gumerov and R. Duraiswami, "Computation of scattering from *N* spheres using multipole reexpansion," J. Acoust. Soc. Am. **112**, 2688–2701 (2002).
[20]N. A. Gumerov and R. Duraiswami, "Computation of scattering from clusters of spheres using the fast multipole method," J. Acoust. Soc. Am. **117**, 1744–1761 (2005).
[21]N. A. Gumerov and R. Duraiswami, "A scalar potential formulation and translation theory for the time-harmonic Maxwell equations," J. Comput. Phys. **225**, 206–236 (2007).
[22]V. Rokhlin, "Diagonal forms of translation operators for the Helmholtz equation in three dimensions," Appl. Comput. Harmon. Anal. **1**, 82–93 (1993).
[23]H. Cheng, W. Y. Crutchfield, Z. Gimbutas, L. F. Greengard, J. F. Ethridge, J. Huang, V. Rokhlin, N. Yarvin, and J. Zhao, "A wideband fast multipole method for the Helmholtz equation in three dimensions," J. Comput. Phys. **216**, 300–325 (2006).
[24]W. C. Chew, *Waves and Fields in Inhomogeneous Media* (IEEE, New York, 1995).
[25]W. C. Chew, S. Koc, J. M. Song, C. C. Lu, and E. Michielssen, "A succinct way to diagonalize the translation matrix in three dimensions," Microwave Opt. Technol. Lett. **15**, 144–147 (1997).
[26]M. A. Epton and B. Dembart, "Multipole translation theory for the three-dimensional Laplace and Helmholtz equations," SIAM J. Sci. Comput. (USA) **16**, 865–897 (1995).
[27]L. J. Jiang and W. C. Chew, "A mixed-form fast multipole algorithm," IEEE Trans. Antennas Propag. **53**, 4145–4156 (2005).
[28]J. Song and W. C. Chew, "Error analysis for the truncation of multipole expansion of vector Green's functions," IEEE Microw. Wirel. Compon. Lett. **11**, 311–313 (2001).
[29]Y. Saad and M. H. Schultz, "GMRES—A generalized minimal residual algorithm for solving nonsymmetric linear-systems," SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **7**, 856–869 (1986).
[30]W. C. Chew, "Recurrence relations for the three-dimensional scalar addition theorem," J. Electromagn. Waves Appl. **6**, 133–142 (1992).
[31]C. H. Choi, J. Ivanic, M. S. Gordon, and K. Ruedenberg, "Rapid and stable determination of rotation matrices between spherical harmonics by direct recursion," J. Chem. Phys. **111**, 8825–8831 (1999).
[32]N. A. Gumerov and R. Duraiswami, "Recursions for the computation of multiple translation and rotation coefficients for the 3-D Helmholtz equation," SIAM J. Sci. Comput. (USA) **25**, 1344–1381 (2003).
[33]D. Healy Jr., D. Rockmore, P. Kostelec, and S. Moore, "FFTs for the 2-sphere: Improvements and variations," J. Fourier Anal. Appl. **9**, 341–385 (2003).
[34]L. Ying, G. Biros, D. Zorin, and H. Langston, "A new parallel kernel-independent fast multipole method," in Proceedings of the ACM/IEEE SC2003 Conference on Supercomputing (SC'03) (2003), Vol. **1**, pp. 14–30.
[35]O. Ergül and L. Gürel, "Efficient parallelization of the multilevel fast multipole algorithm for the solution of large-scale scattering problems," IEEE Trans. Antennas Propag. **56**, 2335–2345 (2008).

Prosodic peak estimation under segmental perturbations

Greg Kochanski
Phonetics Laboratory, Oxford University, Oxford OX1 2JF, United Kingdom

(Received 2 May 2008; revised 19 October 2009; accepted 20 October 2009)

Despite the apparent simplicity, measuring the position of peaks in speech fundamental frequency (f_0) can produce unexpected results in a model where f_0 is the superposition of a supersegmental component and a segmental component. In these models, the measured f_0 peak position can be as much as an entire syllable different from the peak of the intonation component. This difference can be large enough so that the measured peak positions could falsely suggest a phonological distinction in the intonation where none really exists. This paper then discusses measurement techniques that are less sensitive to segmental effects than directly measuring the position of the f_0 maximum. A algorithm, called the "bracketed maximum," is presented. The performance of these techniques is compared on a corpus of speech data where the intonation is expected to be in a stable position. The bracketed maximum can reduce the variance of peak position measurements by at least 15%, in the presence of changing segmental structure, thereby presumably yielding a more accurate measurement of the intonation peak position.
 [DOI: 10.1121/1.3268511]

PACS number(s): 43.60.Ac, 43.70.Jt, 43.72.Ar [DOS] Pages: 862–873

I. INTRODUCTION

Many papers on intonation are based on measurements of the timing of peaks in fundamental frequency (f_0) contours (e.g., Chen *et al.*, 2004; House, 2003; Ladd *et al.*, 1999; Arvaniti *et al.*, 1998; Silverman and Pierrehumbert, 1990; Pierrehumbert and Steele, 1989). Peak timing is an easy, objective measurement, but f_0 is generally considered to be jointly determined by the segments of the utterance ("microprosodic perturbations") and suprasegmental prosodic properties. This paper investigates obtaining timing of underlying prosodic peaks from measurements on f_0 curves.

This paper has two objectives: First, to show that separating segmental and suprasegmental effects is not straightforward, even in a simple "toy" model; there are unexpectedly large nonlinear interactions. Second, to compare several different ways of measuring the peak timing to find which is best at estimating the timing of the prosodic peaks.[1] We look for the measurement technique that is least affected by changes in the segmental structure of a sentence.

This paper assumes a superposition model for intonation. Superposition models assume that an observable is made of a sum of two components and that the two components are independent of each other. One can always split an observed time-series into two or more components; in fact, there are an infinite number of ways to do so. However, a superposition model requires more than just the addition of two components; it also asserts that the two components are independent and that they can be combined in any desired way.[2] This provides some constraint on how f_0 can be split into two components.

Another constraint comes about because intentional control of f_0 is not fast enough to reproduce or compensate for segmental effects. The muscles of the larynx cannot produce an increase in f_0 in less than 100 ms (Stevens, 1998, pp. 40–48 and references therein; Xu and Sun, 2000), while segmental effects can come and go in that interval. This means that, to some level of approximation, we can split the segmental effects away from the prosodic part by their characteristic time scale: Short-term effects can be treated as aerodynamic or segmental effects, and changes over time scales longer than 100–200 ms are plausibly prosodic. Such a division is consistent with the linguistic association of prosody with suprasegmental properties (i.e., properties spanning more than one segment).

For f_0, the first component in a superposition model is the prosodic contour, which is normally assumed to be determined by the choice of accents and their positions. The second component is the segmental effect: a change in f_0 that depends on the phone. Low vowels characteristically have a low f_0, whereas high vowels have a somewhat higher f_0 on average (Crandall, 1925; Taylor, 1933; Peterson and Barney, 1952; House and Fairbanks, 1953; Ladefoged, 1964; Whalen and Levitt, 1995). This vowel-dependent shift in f_0 is variously estimated to be from 4 Hz to above 10 Hz, i.e., roughly 0.5 semitones. While this shift is fairly small compared to the (roughly) three semitone standard deviation of f_0 in normal speech (Baken, 1987, Table 5-2), it will be seen that it is not small enough to safely ignore.

Consonants can also affect f_0. For instance, nasality has been observed to have a significant effect (Silverman, 1987). Also, van Santen and Hirschberg (1994) showed that voiceless consonants and voiced obstruents can make short-lived changes of about 20 Hz near the onset of voicing. Thus the size of these consonant effects is comparable to those caused by vowel height.

One example of a superposition intonation model is that of Fujisaki (1983), who constructed f_0 contours from a superposition of short-term and long-term prosodic effects. Other models of intonation that add segmental effects onto a prosodic component are those from Morlec *et al.* (1996), Di

Cristo and Hirst (1986), van Santen and Hirschberg (1994),[3] van Santen and Möbius (1999), and Ross and Ostendorf (1999).

A superposition model for intonation can be written as

$$f(t) = p(t) + s(t), \tag{1}$$

where $f(t)$ is the observed (surface) frequency at time t, $s(t)$ is the segment-related frequency shift, and $p(t)$ is the prosodic (accent-related) part. Superposition implies that for any $p(t)$ which is a possible prosodic contour, and for any $s(t)$ which is a possible segmental frequency shift, then $p(t)+s(t)$ is a valid, physically possible intonation contour.[4]

The independence assumption in superposition models is probably just an approximation. But, if it were not a useful approximation, one would expect to see a substantially different choice of segments in regions of low f_0 vs high f_0, or alternatively, substantially different phonetic implementations of segments in low f_0 regions from high f_0 regions. While there are differences in segmental content of accented vs non-accented regions (e.g., schwa is rare under accented regions or see Greenberg *et al.*, 2001), the correlation with f_0 is presumably not strong, because high f_0 is not strongly correlated with accent/prominence (Kochanski *et al.*, 2005; Kochanski and Orphanidou, 2008).

Also, the fact that segmental effects have been measured suggests that they are not compensated by speakers: If speakers automatically compensated for $s(t)$, then one would not expect an observable frequency shift when changing from a high vowel to a low vowel. This is consistent with the assumption that $p(t)$ is not used to compensate for $s(t)$—i.e., that they are independent.[5]

II. A TOY SUPERPOSITION MODEL

I will illustrate the behavior of the model with simple idealized forms for $s(t)$ and $p(t)$. This section explores the mathematical consequences of superposition models of intonation, and does not propose particular forms or numerical values for any particular sentence.[6]

In this example, the equation for $s(t)$ is chosen to represent an alternating sequence of phones that have intrinsic frequency shifts that differ by 8 Hz in successive syllables. The segmental part is thus taken to be

$$s(t) = 4\,\sin(2\pi t/d) \tag{2}$$

(in hertz), where $d=0.5$ s is the period after which the segmental structure repeats; successive syllables can be imagined to occur every $d/2=0.25$ s.

The intonation part is

$$p(t) = 170 + 29\,\cos(2\pi(t-\tau)/D), \tag{3}$$

which is a wave with period $D=1.3$ s (i.e., about five syllables). Low and high intonation targets are separated here by two or three syllables. In this example, the average $f_0=170$ Hz is chosen to be midway between typical male and female values. The accents cause ± 29 Hz pitch excursions (roughly three semitones) relative to the average. Figure 1 shows the underlying intonation contour $p(t)$, as a dashed curve and the surface f_0 contour, $f(t)$. The segmental shifts are small compared to the assumed intonation.

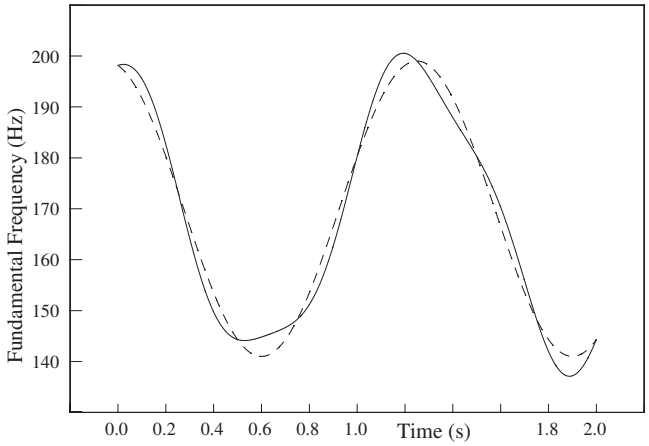


FIG. 1. The underlying intonation contour $p(t)$ (dashed) and the surface intonation $f(t)$ (solid line) for the model utterance. The plots shows fundamental frequency vs time. The difference between the contours is the segmental effect $s(t)$.

The time at which the first peak in $p(t)$ occurs is given by τ in Eq. (3). Figure 2 displays a set of different $p(t)$ curves, each shifted slightly. These shifts correspond to changes in the alignment of the intonation peaks relative to the segmental structure of the sentence.

Figure 2 shows the intonation contours generated by the model. They differ only in the value of τ, which increases from $\tau=-50$ ms in steps of 30 ms. This set of curves corresponds to changes in the alignment of the intonation component $p(t)$ relative to the fixed segmental structure $s(t)$. These curves are simply shifted versions of the dashed curve in Fig. 1.

This example uses frequency shifts caused by vowels; however, consonants can also contribute substantial shifts to $s(t)$. If there were only shifts on vowels, then peaks or valleys of $s(t)$ would be aligned with the syllable centers. But, if one includes segmentally specified f_0 shifts from consonants, peaks and valleys of $s(t)$ can also occur between syllables or at the edges of syllables. In this toy example, I plot curves in terms of a shift in $p(t)$, but equivalent results could be ob-

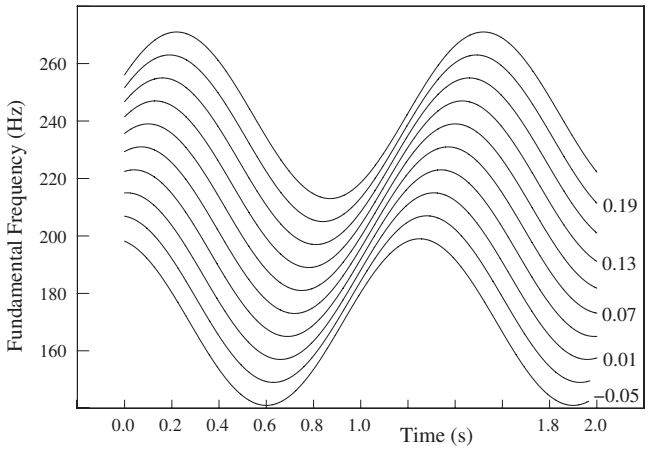


FIG. 2. A set of contours for the intonation component $p(t)$. These form inputs to the intonation model, Eq. (1). The contours are shifted vertically for clarity. Contours are labeled by τ (in s) and the maxima of neighboring curves differ by 30 ms.

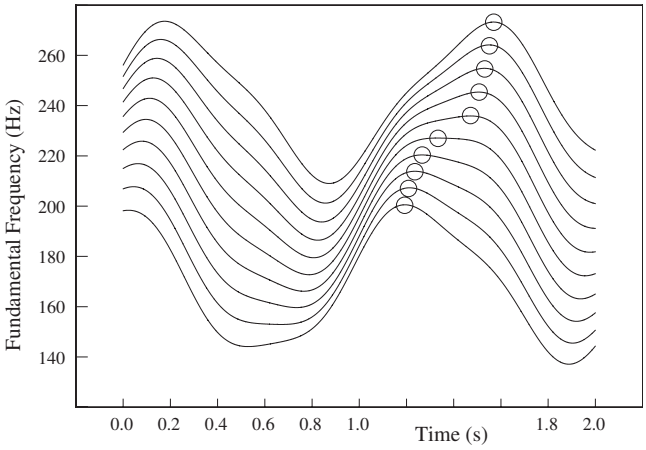


FIG. 3. A set of contours for $f(t)$, derived from Fig. 2 by adding segmental effects [Eq. (2)]. Small circles mark the maxima of the contours. Plotted as per Fig. 2.

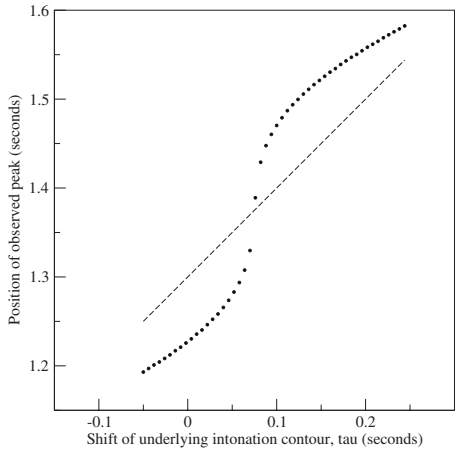


FIG. 4. The relationship between the time shift τ of the underlying intonation contour and the maximum of the observed fundamental frequency, arg max$\{f(t)\}$. The vertical axis indicates the position of the maxima (open circles) in Fig. 3. The diagonal dashed line corresponds to the position of the underlying prosodic peak.

tained by holding $p(t)$ constant and changing the pattern of segmental shifts. This is because the mathematics depends only on the alignment *difference* between segmental and prosodic peaks.

Figure 3 shows the resulting f_0 contours produced from the curves in Fig. 2 by way of Eqs. (2) and (1). As the intonation contour shifts with respect to the segments, a sudden jump in the time of the maximum occurs between $\tau = 70$ ms and $\tau = 100$ ms.

It is perhaps surprising that segmental effects, which shift f_0 up and down, can lead to substantial differences between the timing of the peak of the underlying prosodic contour and the surface f_0 contour. Rather than being on top of peaks in the underlying prosodic contour, peaks in f_0 are pulled toward nearby segments that cause a positive shift in f_0.

This can result in segmental anchoring: Imagine a group of utterances that have the same segmental structure but progressively different underlying prosodic f_0 contours. In such a situation, an f_0 peak can be "pinned" to a particular segment that has a positive segmental effect. Even if the peak of $p(t)$ shifts from utterance to utterance (within some range), the peak of f_0 might always occur within the same segment.

If there are two nearby segments with positive segmental effects with the underlying peak of $p(t)$ in between, the f_0 peak can switch from being pinned on one to being pinned on the other. This can lead to an effect where the f_0 peak jumps a substantial interval (e.g., the spacing between two syllables) as the result of an arbitrarily small change in the alignment of the underlying prosodic contour. Such a jump could easily be misinterpreted as evidence for distinct phonological states of prosodic alignment.

A. Explaining the jump

One can understand the sudden jump by considering two limiting cases of the model. First, suppose that the amplitude of $s(t)$ is very small compared to $p(t)$. Then, the time of the observed f_0 peak would smoothly follow the peak of $p(t)$.

Now, consider the opposite case where $p(t)$ is small compared to $s(t)$. Then, segmental shifts would dominate, and the overall maximum would always be very near one of the maxima of $s(t)$, but the underlying prosody would select which maximum would be the highest. One of the maxima of $s(t)$ would be pushed up by the (tiny but nonzero) $p(t)$. Therefore, if one were to vary τ, the maximum of $f(t)$ would move only occasionally, when it would jump from one maximum of $s(t)$ to the next. (This case describes an extreme form of segmental anchoring, where strong correlations arise between measured f_0 maxima and segment locations.)

For realistic values of $s(t)$ and $p(t)$, the model gives results intermediate between these two extreme cases. The maximum of $f(t)$ is loosely anchored to the maximum of $p(t)$, but it changes non-uniformly as the underlying alignment changes.

Figure 4 plots the relationship between the maxima of $f(t)$ and τ. (It is a more detailed plot of the times of the maxima at the left edge of Fig. 3, near $t=0$.) It is s-shaped and strongly nonlinear. There is a non-trivial relationship between the time of the observed peaks in $f(t)$ and the time-alignment of the accents, which is the underlying quantity of interest to intonational phonologists.

The nonlinearity becomes more pronounced as the magnitude of $s(t)$ increases or as the magnitude of $p(t)$ decreases. In fact, if the curvature (i.e., the second time derivative) of $s(t)$ were to exceed the curvature of $p(t)$, Fig. 4 would have a discontinuous jump so that an infinitesimal change in τ would lead to a substantial jump in the time of the observed f_0 peak. These jumps occur when the top of the maximum is flat (or nearly so); such flat-topped f_0 contours are described in Knight, 2002, Ogden *et al.*, 2000, especially pp. 194–195, Fig. 8, and others. With a small pitch range in $p(t)$ or flat-topped profile, a small shift in alignment or segmental structure can even move the time of peak f_0 from one syllable to the next.

B. Segmental anchoring

Segmental anchoring occurs when a peak in $s(t)$ approximately coincides with a peak in $p(t)$. This should hap-

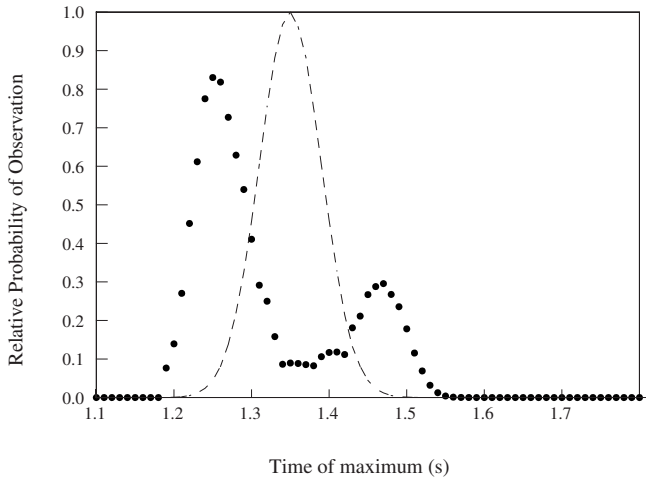


FIG. 5. The distribution of the times of f_0 peaks including segmental shifts (dots). This histogram assumes a Gaussian distribution of τ (the location of the underlying intonation maxima) with a standard deviation of 40 ms. (A shifted version of the distribution of τ is shown by the thin dashed line.) The horizontal axis is the position of the maximum and the vertical axis is the probability of observation, collected into 10 ms bins.

pen when high vowels or certain consonants are near accents that raise f_0. The mechanism is discussed in Sec. II A.

The anchoring effect can be important even in tightly controlled experiments that generate several intonations for the same text. Dilley *et al.* (2005) tested whether accents are anchored to each other or to the segmental structure, and concluded that for their data, the anchoring is segmental. Several other papers can be interpreted as showing anchoring of f_0 contours to the syllable (Arvaniti *et al.*, 1998; van Santen and Hirschberg, 1994; Ladd *et al.*, 2000), possibly because of this effect.

C. False phonological contrasts

The opposite of segmental anchoring occurs when a prosodic peak approximately coincides with a segment that pushes f_0 down. This is the situation shown in Sec. II.

The resulting nonlinearity shown in Fig. 4 is important because it could lead a researcher to falsely assert a discrete (and thus presumably phonological) contrast when none really exists. Suppose that the alignment of the peak of $p(t)$ is somewhat variable and is taken from a single unimodal distribution with no phonological distinctions. For example, take τ from a Gaussian distribution with a mean of 50 ms and a standard deviation of 40 ms. One can think of this as a corpus containing many utterances, each with a slightly different alignment.

To show the effect, the average peak alignment will be placed midway between maxima of the segmental effect. Figure 5 shows that despite the unimodal distribution of alignments, the distribution of observed f_0 maxima is bimodal.[7] This could easily lead to a false belief that there are two underlying phonological categories (e.g., early peak vs late peak), one for each maximum.[8,9] In reality, though, the bimodal distribution is generated by an interaction between the intonation, the segmental shifts, and the measurement procedure.

Thus, taking the maximum of an f_0 contour and interpreting it as an intonation peak are dangerous procedures. Small segmental effects can lead to large changes in the position of the maximum. The root cause of the trap is that the peaks of observed f_0 that one *can* measure are not the same as peaks in the prosodic contour, $p(t)$, that one would *like* to measure. An algorithm is therefore needed to estimate peaks in $p(t)$ from observations of $f(t)$ (see Fig. 5).

D. The bracketed maximum algorithm

It has been shown that the simple maximum[10] of $f(t)$ is not a satisfactory estimate of the maximum of the underlying prosody. But how can one do better? Ideally, one would model the segmental effects $s(t)$ and simply subtract that function from $f(t)$ to directly yield the desired $p(t)$. However, this would not be trivial, as segmental shifts seem to vary from speaker to speaker. Intonation experiments and data analysis would have to become rather larger and more complex to estimate the segmental shifts of each phone for each speaker in all relevant contexts.

A more practical approach is the "bracketed maximum,"[11] which springs from two ideas. First, if one thinks of the segmental effects as "noise," one might be able to get a better measurement of the prosodic peak position if we could take the average of two measurements from locations where the segmental effects were approximately independent.[12] This is possible by measuring in two different segments.

Normally, peaks in the prosodic component of f_0 are expected to be wide enough to cover more than one phoneme and their left and right edges will often be in different phonemes with different, almost uncorrelated segmental effects. As long as the two measurements are close enough to be on the same accent, it is plausible that the errors in the average position may be smaller than the error in either individual measurement.

Second, making a timing measurement on the side of a peak is intrinsically less sensitive to segmental effects than a measurement near the top of a peak. Suppose one is looking for the top of a broad maximum. There might be a region perhaps 50–100 ms long over which the frequency changes by only 4 Hz. Now, if segmental effects were different, so that f_0 changed by a few hertz, then the observed maximum might move anywhere within that large region, controlled by the pattern of segments.

On the other hand, if one makes timing measurements on the side of a peak, they will be less affected by segmental effects because f_0 changes rapidly. To take a plausible example, if $p(t)$ swings through 40 Hz in 100 ms, it will take only 10 ms to swing through the 4 Hz range corresponding to segmental effects. So, if the segmental structure of the sentence were rearranged without changing $p(t)$, one would expect the time at which f_0 crosses a given threshold to change by only 10 ms or so. Changing the segmental effects will thus have a smaller effect on the timing measurements if the measurements are made on the sides of the peak.

The bracketed maximum is a simple technique that implements these ideas and can reduce the effect of segmen-

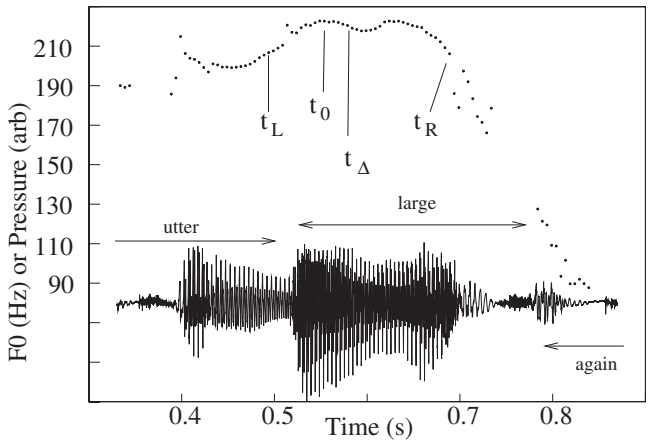


FIG. 6. Sample audio data are plotted (lower curve) with a transcription and an f_0 contour above that. The f_0 contour is annotated with the simple maximum (labeled t_0), intermediate results from the bracketed maximum algorithm with $\Delta=20$ Hz (labeled t_L and t_R), and the final result (labeled t_Δ). These data were chosen to show the operation of the bracketed maximum algorithm on a broad f_0 peak.

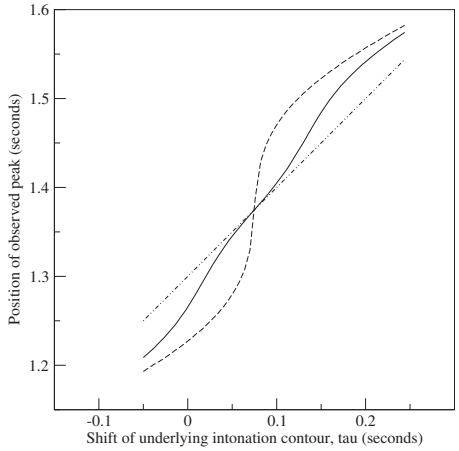


FIG. 7. The bracketed maximum (solid), the simple maximum (dashed) vs τ, and the maximum of the underlying intonation contour.

tally determined frequency shifts. It involves measuring on both sides of the maximum of $f(t)$ and averaging the two measurements. It should work as long as the segmental shifts are small compared to the prosodic f_0 swings. (Edge cases are discussed in Appendix B.)

The technique involves four steps.

- Find the maximum of f_0.
- Go backward from the maximum as long as f_0 is within Δ Hz of the maximum, or until an unvoiced region is encountered. Call this time t_L. Here, Δ is called the measurement offset.
- Go forward from the maximum, as long as f_0 is within Δ Hz of the maximum,[13] or until you encounter an unvoiced region. Call this time t_R.
- Average t_L and t_R to produce an estimate of the time at which $p(t)$ is maximal. This is an estimate of the underlying intonation peak alignment.

Note that for $\Delta=0$, the result is just the simple maximum. The algorithm may be downloaded from http://kochanski.org/gpk/papers/2008/Segmental_Additive/algorithm.py.txt and/or as part of this paper's supplemental materials (EPAPS).

(Figure 6 shows sample data annotated to show the operation of the bracketed maximum algorithm.)

In the toy model of Sec. II, this algorithm yields a more accurate and more linear relationship between the observations and τ, compared to the simple maximum. Figure 7 shows a comparison between τ the result of the bracketed maximum algorithm and the result of the simple maximum.

Another advantage of the bracketed maximum can be seen in Figure 8. This is computed as per Fig. 7, except that it shows the distribution of peak positions for both the bracketed maximum and simple maximum algorithms. The bracketed maximum gives a unimodal distribution and thus does not falsely suggest a phonological distinction between two categories. The full-width at half-maximum is 107 ms, close to the 94 ms full-width at half-maximum for the distribution of τ.

The measurement offset (Δ) that one uses is not particularly critical in this "toy" model. Performance gradually improves as the measurement offset is increased, with a broad optimum when it is half of the peak-to-peak swing of $p(t)$ (e.g., 29 Hz in this example). However, the bulk of the improvement comes by the time the measurement offset reaches the peak-to-peak segmental effect, $s(t)$ (e.g., 8 Hz).

E. Limits on the choice of Δ

One limit on using the bracketed maximum algorithm on real speech is that one does not know the size of the relevant pitch excursions. Since the bracketed maximum will not produce a useful value for f_0 maxima that are smaller than the measurement offset, Δ should not be made too large.

However, very small f_0 maxima are unlikely to be perceptually important or phonologically distinctive. When it is

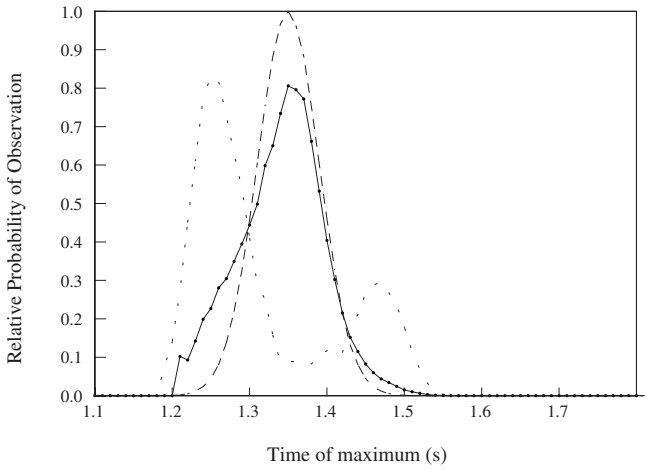


FIG. 8. The distribution of estimated intonation peak positions $(t_L+t_R)/2$ for a simulated experiment using the bracketed maximum technique (solid line with dots). The resulting histogram is unimodal, reflecting the unimodal distribution of τ (alignment). Results from the simple maximum (from Fig. 5) are reproduced as a dotted line for comparison, and the underlying distribution of the prosodic peak position (τ) is shown as a dashed line.

important for the speaker to communicate a distinction between two possible meanings of a sentence, one expects that he or she will produce an easily perceptible peak, large enough to be reliably detected by the listener. Under favorable conditions, frequency differences as small as about 3 Hz may be detectable (Chuang and Wang, 1978), but other experiments show that larger pitch motions are needed (Peng, 2000, Sec. 10.3). Mack and Gold (1984) showed that the minimum detectable pitch shift is a function of the complexity of the stimuli, ranging from 2 Hz for a buzz-tone though 4 Hz for monotone sentences, to more than 6 Hz for sentences with near-natural intonation patterns.

Since the discrimination threshold is typically defined as a 75% correct detection of a difference under quiet laboratory conditions without distractions, one expects that a yet larger shift is necessary for reliable communication in realistic conditions.[14]

Linguistic evidence also points to somewhat larger values. For instance, Holm and Bailly (2002) (Sec. 3.1) noted that different repetitions of the same utterance typically differ by one semitone (about 10 Hz). Braun *et al.* (2006) (Sec. 3.C) obtained similar results, suggesting that phonologically distinct utterances are separated by about three semitones, while changes smaller than about one semitone were unimportant.

Overall, these considerations suggest that any phonological differences that the speaker wishes to be understood by the listener will probably be encoded by pitch shifts of 10 Hz or larger. So, with Δ near 10 Hz, the algorithm should produce reliable results for most f_0 maxima that have any communicative function.

III. TESTS ON REAL SPEECH

A. Introduction

The argument so far rests on an idealized mathematical model (although the parameters used in the model are consistent with observations). One of the predictions of the model is that given a corpus of sentences with the same prosodic contour but different segmental structures, segmental effects would make the timing of the observed f_0 peak vary.

Thus, a measurement technique like the bracketed maximum should reduce the variability of peak timing in such a corpus. We can check this chain of logic experimentally by comparing various measurement algorithms. Whichever procedure produces the smallest variance in peak timing is presumably least affected by segmental effects.

This simple maximum algorithm will now be compared to the following:

- The bracketed maximum, described in Sec. II D.
- Smoothing the data with a median filter, then taking the time at which the median is maximum. This follows Taylor, 1993 and Xu and Xu, 2003. It ignores variations of f_0 on short time scales such as 100 ms or less. These short time scales are where the segmental effects are most dramatic.
- Smoothing the data by averaging over a window centered around each point, and then taking the time when the average is maximal.
- The MOMEL algorithm of Di Cristo and Hirst (1986); it transforms intonation contours into a quadratic spline approximation, producing a smooth representation that bridges over unvoiced regions.

A priori, one expects that these approaches generally will smooth away structures on short time scales (e.g., segmental effects) and will generally preserve longer, suprasegmental f_0 motions.

B. Data

I use a corpus of utterances with fixed sentence patterns where there is reason to believe that the intonational phonology always specifies a peak in the same position. One can then compare algorithms to see which one gives a more stable estimate of the intonation peak.

The database consists of single-syllable words embedded within a frame sentence (Slater and Coleman, 1996); it was previously collected for a different purpose. The corpus contains 4970 utterances from a single speaker of Southern British English (each text is recorded five times). There are four frames used: "Can you utter *X* again, please?" (where *X* is a word in the form *CxC*, beginning and ending with a consonant), "Can you utter *X* today?" (*X* is *CxV*, beginning with a consonant and ending with a vowel), "Have you uttered *X* again?" (*X* is a *VxC* word), and "Have you uttered *X* today?" (*X* is *VxV*). The speaker was phonetically trained and knew he was reading a list containing minimal pairs of English words. In such a database, the frame sentence and the words are semantically neutral, and the syntax is such that any word can be used as *X*. Consequently, the intonation should be identical within groups of utterances that share the same frame, and plausibly among all the utterances.

One expects an accent on *X*, because it is most informative. (General information on focus and accent location can be found in Ladd, 1996, Sec. 5.1 and references therein.) The accented word should thus be louder and longer than its neighbors (e.g., Kochanski *et al.*, 2005 and references therein). A random sample of 100 utterances was checked by the author, and an obvious prominence was heard on the variable word in 98 cases. In 3 of those 98 cases, words in the frame were judged to be as prominent as the variable word. So, in 95 out of the 100 samples, the variable word was the most prominent word in the utterance.

C. Signal processing and overview of data

The acoustic data were processed to extract time-series measurements of quasi-duration, loudness[a], f_0, and aperiodicity, as per Kochanski *et al.* (2005). Time axes were normalized to span the range from 0 to 1 between the beginning of the second syllable of "utter"/"uttered" and the end of the first syllable in "today"/"again." Thus, normalized time 0 to 1 always spans three syllables and the variable syllable is centered near 0.5.

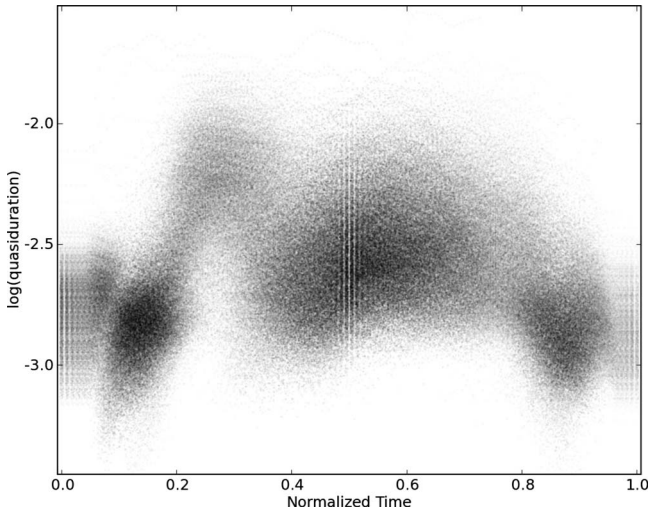


FIG. 9. Running duration scatter-plot for the entire corpus. The quasi-duration is a measure of the stationarity of the speech spectrum; small values imply a rapidly changing spectrum. The plot shows the log of the normalized quasi-duration against normalized time, otherwise plotted as per Fig. 11.

For each utterance, we compute estimates of the prosodic peak position for different values of the measurement offset Δ. This yields values $t_i(\Delta)$ where i indexes the utterance.

We included only f_0 maxima on or near X by restricting the analysis to f_0 data with maxima at normalized time between 0.2 and 0.8. We did this because the f_0 in utter/uttered can sometimes be higher than the peak f_0 within the variable syllable[15] (see Fig. 11); this can lead to unexpected results in the computed peak positions. Utterances where either t_L or t_R was outside that region were dropped.

A scatter-plot of the quasi-duration for the corpus is shown in Fig. 9; this is computed per Kochanski *et al.* (2005). (At each time, the quasi-duration measures how far one can go forward and backward in time before the spectrum changes substantially. The quasi-duration at any time is roughly proportional to the duration of the phone at that time.) The frame syllables are centered near normalized times of 0.17 and 0.82 and the variable syllable is approximately centered. The variable syllable typically has a longer vowel with a more stable formant structure than found in the frame syllables. (The region near normalized time 0.3 corresponds approximately to the boundary between the preceding frame syllable and the variable syllable.) A plot of an estimate of the perceptual loudness appears in Fig. 10. The variable syllable can be seen to be typically longer and louder than its neighbors, and thus should typically be prominent. (Incidentally, one can see a bimodal distribution of loudness from the two possible frames near normalized time 0.25.)

Figure 11 gives an overview of f_0 data from the entire corpus. It shows a three-syllable region centered on X, plotted as per Braun *et al.* (2006). It is created by placing 342 000 dots, 1 per f_0 measurement. In the plot, dot sizes are proportional to an estimate of the perceptual loudness in regions where the speech waveform is approximately periodic and are reduced where the waveform is aperiodic. This emphasizes loud, periodic regions (e.g., the central vowel of each syllable and regions where the f_0 data are most reliable). The image has been smoothed slightly to reduce printing artifacts; this blurs the individual dots slightly. The variable syllable typically also has an f_0 peak, providing some further evidence of its prosodic prominence.

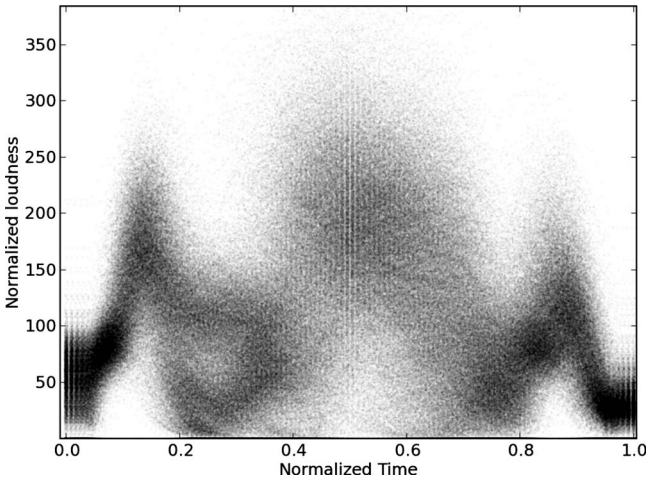


FIG. 10. Normalized loudness vs normalized time for the entire corpus. Plotted as Fig. 11, except that all dots are the same size.

1. Flat-topped profiles

In many utterances, the f_0 peak position is not completely clear, usually because the f_0 curve has a flat top with no obvious peak. These are commonly known as plateaus (House *et al.*, 1999; Ogden *et al.*, 2000; Wichmann *et al.*, 1997). Figure 12 shows one such example. (Note that this utterance was chosen to display a plateau; for f_0 data more representative of the corpus as a whole, see Fig. 14.)

To count flat-topped utterances, I looked for utterances where many points would likely to be indistinguishably high. I defined top points to be voiced points within 3 Hz of the

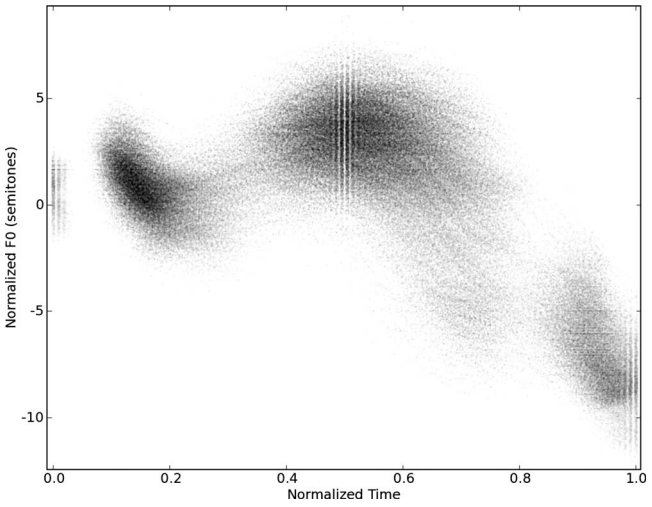


FIG. 11. Fundamental frequency of the entire corpus. This plots normalized f_0 against normalized time for the region of interest. The horizontal axis goes from the beginning of the syllable before the variable syllable to the end of the syllable after. The vertical axis is f_0 deviation from 170 Hz, in semitones. This is a smoothed scatter-plot of f_0 measurements. (The vertical stripes, e.g., near x=0.5, are the result of the 10 ms interval between f_0 measurements.)

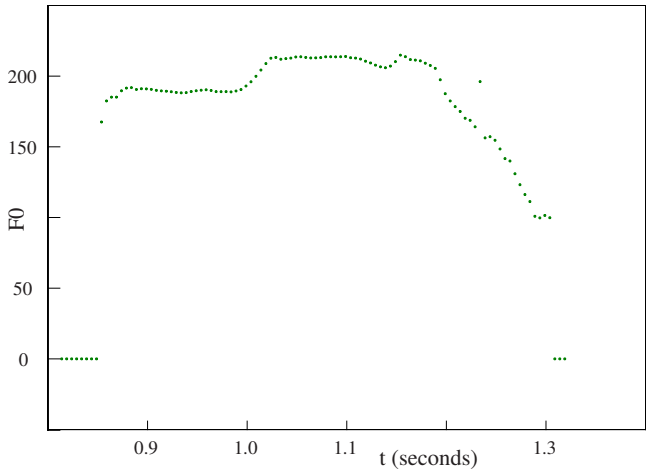


FIG. 12. (Color online) Sample flat-topped f_0 profile.

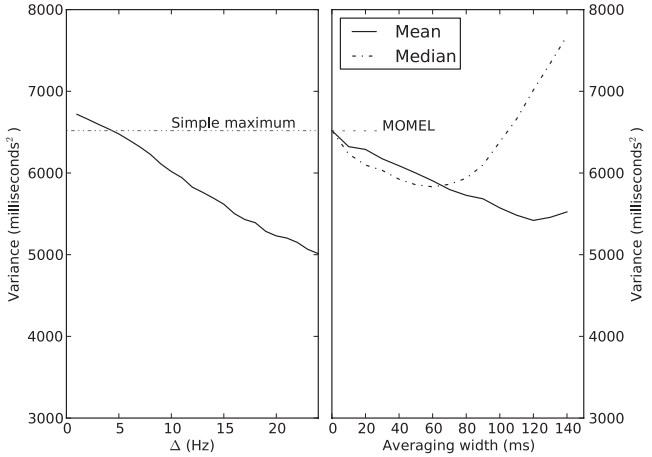


FIG. 13. Left: variance of the bracketed maximum estimate of the prosodic peak position, plotted against Δ. The solid line shows the variance of the entire corpus. The simple maximum is shown by the dashed-dotted line, for comparison. Right: variance of algorithms that smooth the f_0 data and then take the maximum. The variance is plotted against the width of the smoothing window. The best MOMEL result is the gray dot. (MOMEL does not have a comparable smoothing parameter, so the horizontal axis is irrelevant.)

second-highest f_0 value. (The point with highest f_0 was ignored on the grounds that it was commonly at the edge of a voiced region and the waveform was usually not stationary there.) The f_0 differences among these top points are small enough so that the listener is unlikely to be able to reliably pick one as higher than another (see Sec. II E). Considering the speaker, small muscle tremors or other linguistically unimportant changes in production could push f_0 up or down by 3 Hz, so any top point could plausibly have been an intended intonation maximum (this follows House *et al.*, 1999; Knight and Nolan, 2006).

In each utterance in the corpus, a computer program found the largest interval starting and ending on top points that contains at least 50% of top points. The length of this interval provides an estimate of the uncertainty in the time of the maximum: It measures the length of the flat top of the utterance. Using this criterion, 9% of the utterances have intervals of plausible maxima longer than 100 ms and 36% have intervals longer than 60 ms. This criterion, while somewhat arbitrary, shows that a substantial number of utterances have tops that are flat enough so that segmental f_0 perturbations could lead to large changes in the timing of the maximum.

D. Results: Comparison of algorithms

To compare different approaches, we computed the variance of the peak position over the corpus. Since the underlying accent is expected to be in a stable location, the algorithm that provides the most stable measurement should be the most accurate (see Appendix A for a justification of this assumption). This idea is a common way of choosing a measurement technique; see Kochanski and Orphanidou, 2008 in phonetics and Sachs *et al.*, 1995; Krolik, 1996; and Abel, 1990 in other fields. However, this paper is not a full-blown competitive evaluation, primarily because the testing corpus is produced by a single speaker.

In our corpus, there is a built-in dependence of the frame sentence on the initial or final segments of the variable word. This could possibly cause the prosodic peak position to depend on the segmental structure if the type or alignment of the accent depended on the choice of frame (see Fig. 15). However, this problem is avoided by splitting the corpus into four sub-corpora, one for each possible frame, and then computing the variances separately within each sub-corpus.

Figure 13 shows this variance for different algorithms (see Appendix C for comparisons of the various sections of the corpus).

As the measurement offset is increased, the variance of the bracketed maximum progressively decreases. Thus, the best measurement of the underlying accent position is made with Δ as large as possible. For all $\Delta > 3$ Hz in this data set, the variance is reduced by a statistically significant amount. The significance level improves from $P < 0.05$ at $\Delta = 3$ Hz to $P < 10^{-6}$ at $\Delta = 10$ Hz. For $\Delta = 10$ Hz, the variance is reduced by 9%, and for $\Delta = 20$ Hz, the variance is reduced by 21%.

Note that this reduction is a fraction of the total variance, which includes both the errors induced by segmental effects and also the intrinsic variability in the peak alignment (see Appendix A for discussion). Thus, the fractional reduction quoted here gives a lower bound to the reduction in the segmentally induced measurement error.

The largest practical Δ is set because the algorithm can only measure peak positions if the peaks are larger than Δ. For this data set the fraction of unmeasurable syllables is small (less than 1%) as long as $\Delta \leq 20$ Hz. This is consistent with Sec. II E because the psychophysical and linguistic evidence only establish minimum peak sizes for intelligibility; nothing prohibits a speaker from using an f_0 peak larger than the minimum. However, this 20 Hz upper limit for Δ should be treated as an estimate, as substantial inter-subject and inter-style differences in f_0 range are not unknown.

When this new procedure is used, the largest possible value of Δ should be used, so long as linguistically important f_0 peaks are not lost. Users should be aware of the tradeoff between large Δ, which provides the most reliable timing

measurements but cannot measure smaller peaks, vs small Δ, which provides less advantage over a simple maximum but can operate on small peaks.

The right side of Fig. 13 shows corresponding results for two algorithms that smooth the f_0 contour and then take the simple maximum. The performance of these algorithms is plotted as a function of the width of the smoothing window. Using an arithmetic smooth with a 110-ms-wide window (i.e., taking the mean of all points that are voiced and within 55 ms of the point under consideration) reduces the variance by 13%.

However, the variance worsens for large smoothing windows. Presumably, the smoothing window is becoming wide enough to significantly distort the shape of $p(t)$. This optimum point may be the width where the increasing systematic distortion becomes more important than further reductions in segmental effects. Since this optimum point is likely to depend on the speech rate, mean spacing between accents, and other characteristics of the corpus, in practice, it would be safer to use a smaller smoothing window. With a 90 ms window, an arithmetic smooth improves the variance by 12%. This improvement is very similar to the bracketed maximum algorithm's performance when operated with a relatively safe value of $\Delta = 10$ Hz.

Results for the median smoothing algorithm are similar, but not as good (for the particular speaker who produced our corpus). The optimal window width is then 60 ms, which yields an improvement of 10% in the variance. The median smooth behaves very badly for large window widths; for windows 100 ms or wider, it is worse than the simple maximum. This is doubtless related to the flat tops that are typically produced on intonation peaks when $f(t)$ is subjected to a median smooth. After a median smoothing operation, there is often no unique maximum, with several points near the top of the peak having mathematically identical values. The final step of applying the simple maximum to the smoothed curve then can behave badly.

Both of these smoothing-based algorithms can improve the accuracy of the timing estimation, but the bracketed maximum can out-perform them under some conditions (i.e., for this corpus when $\Delta > 14$ Hz).

MOMEL (Hirst and Espesser, 1993; Di Cristo and Hirst, 1986) does not yield an improvement to the peak location accuracy. We ran MOMEL in 180 different ways; the best performance in terms of timing variance is plotted in Fig. 13 (right), which is nearly the default (Hirst and Espesser, 1993) parameters. All the runs used f_0 data sampled at 10 ms intervals. The best-performing six sets of parameters are all close to the recommended defaults for the program, with $1.032 \leq \text{max err} \leq 1.044$, and other parameters falling between 80% and 110% of their default values of win1=30, win2=20, min d=5, and min r=0.05; none of the top six runs used the "–non-elim-glitch" flag. All of the top 45 runs (including the run plotted) mask MOMEL's output with a voicing indicator, so that it does not interpolate f_0 into unvoiced regions. The alternative (treating MOMEL's result as valid in unvoiced regions) leads to a peak position variance, which is 20% larger than applying either MOMEL or the simple maximum over voiced regions.

This result is perhaps not surprising, because MOMEL is here used differently from its intended application. It was intended to represent entire utterances, and intended to provide a smooth f_0 curve that is intonationally equivalent to the input data. Here, it is being used on a small fragment of an utterance, extracted from the middle, and we are hoping for it to smooth away segmental effects.

IV. CONCLUSION

Segmental effects are surprisingly important to experiments that measure the timing of f_0 peaks, if a superposition model of intonation is adopted. It has been shown that

- peak positions can be strongly influenced by the segmental structure near the peak, and
- very small changes in the underlying intonation can lead to large jumps in the measured peak position.

The model shows that there can be unexpectedly large correlations of peak positions with the segmental content of the utterance. These problems arise because in a superposition model, the measurable quantity (f_0) is different from the underlying prosodic contour. Even with fairly large intonational swings (such as three semitones), segmental effects should not be ignored, and they are increasingly important when the pitch range becomes smaller. This effect can lead to the "anchoring" of f_0 peaks to segments that boost f_0. Conversely, peaks in f_0 will be repelled from segments with especially low f_0. Under plausible assumptions, this can generate a bimodal distribution of f_0 peak positions that is unrelated to any underlying phonological distinctions.

This work shows that any f_0 peak much less than 10 Hz tall will have its height and alignment strongly affected by segmental effects. Even if there are lexically identical comparison data, the peaks in f_0 will be systematically biased toward segments with a higher intrinsic f_0. (Also, a review of literature suggests that such a small peak is unlikely to reliably communicate anything to the listener.) Consequently, f_0 peaks with small excursions may best be ignored, or used only in situations where segmental effects are well understood. Overall, peak timing measurements need to be conducted with careful consideration of segmental effects.

These observations imply that peak timing measurements should not be simply interpreted as the alignment of the peak of an underlying prosodic contour in superposition models of intonation. Overall, segmentally influenced differences between underlying and measured peak positions are large enough so that to properly interpret many experiments in terms of intonational phonology, a numerical model of segmental effects will be necessary.

Of course, superposition models are not mandatory. However, if one wishes to maintain the distinction between the surface f_0 and an underlying phonologically determined intonation contour, some sort of model is necessary to connect from one to the other. It seems likely that similar effects will be seen for a broader class of models, including all those referenced herein. An alternative is that phonological analysis would be conducted on the observable f_0 directly. This

would have the implication that f_0 should be identified directly with phonological intent, thus erasing any sort of competence/performance distinction.

This paper also presents a technique for measuring the timing of peaks, the "bracketed maximum." In simulations, it yields a substantially better estimate of the underlying prosodic peak position than simply taking the time of the maximum f_0. The bracketed maximum is tested on speech data, on a corpus where the accent positions are believed to be known a priori. It can give significant reductions in the variance of the estimated peak position, and can behave as well or better than existing techniques that involve smoothing the f_0 contour.

ACKNOWLEDGMENTS

I thank John Coleman, Elinor Payne, Burton Rosner, Chilin Shih, and Christina Orphanidou for their valuable comments. John Coleman also kindly supplied the data set used here. I thank Daniel Hirst for release and explanations for the MOMEL code. I thank the editorial office for protecting the world from undated URLs. I gratefully acknowledge financial support from the United Kingdom's Economic and Social Research Council under Project No. RES-000-23-0149.

APPENDIX A: INDEPENDENCE AND MINIMUM VARIANCE

For a superposition model of intonation, it is possible to show that the best algorithm is the one that gives the least variance. Thus, one can compare two ways of estimating the underlying peak position by simply comparing the variance of the estimates.

Formally, we will compute the variance of the estimated peak positions from each algorithm. We consider the output of algorithm X, t_X, to be the sum of two random variables: the time at which $p(t)$ has a maximum (P) plus an error caused by segmental effects (S_X). If S_X is independent of P, the variances will add nicely so that

$$\mathrm{var}(t_X) = \mathrm{var}(P_X) + \mathrm{var}(S_X). \tag{A1}$$

Since we compare all the algorithms on the same corpus, var(S) is the same for all algorithms. Thus the algorithm with the smallest var(t_X) will have the smallest var(S_X), and therefore it will have the smallest mean-squared error in estimating the timing of the prosodic peak.

For S to be independent of P, it must meet three conditions.

(1) The prosodic peak position must have no intrinsic dependence on the segmental structure. This is true for all superposition models; it would not be true for a model where $p(t)$ explicitly depends on the segments.
(2) The corpus must have no built-in correlations between the segmental structure and the underlying prosodic contour (i.e., with the expected type of accent). This should be approximately true for the corpus used here, especially after we split the corpus into sub-corpora that have uniform frames. See Secs. III B and III D.

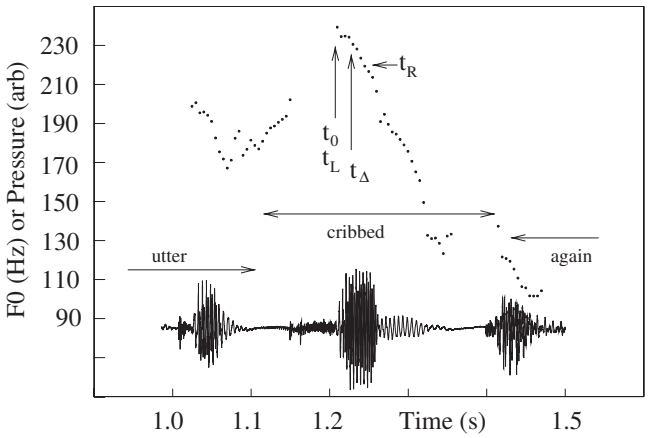


FIG. 14. Sample audio data, plotted as Fig. ??. These data were chosen to show the operation of the bracketed maximum algorithm, when the f_0 maximum is at the edge of a voiced region.

(3) The algorithm must have no dependence on the segmental structure. Any algorithm Q has that desirable property if (1) it is computed just from f_0, and (2) if the algorithm does not depend on the choice of origin for the time axis, i.e., $Q\{f(t+\eta)\}=Q\{f(t)\}+\eta$ for any η. Thus, if one were to delay the speech by 1 s, the result of the algorithm f_0 should also be delayed by 1 s. All the algorithms we consider have both properties.

Thus, the algorithm that has the smallest variance of its estimates will have the smallest errors between its estimates and the underlying position of the peaks of the prosodic contour. This will be true even if there is some intrinsic variability in the underlying prosody.

APPENDIX B: EDGE CASES

The discussion of the operation of the bracketed maximum algorithm in Sec. II D assumed voicing everywhere. This section discusses cases where unvoiced regions are important. Ultimately, though, its value will be tested experimentally in Sec. III.

1. Prosodic maximum within, but near the edge of voiced region

Figure 14 shows sample data annotated to show the operation of the bracketed maximum algorithm. In this situation, one of the bracketing measurements may end up at the voiced/unvoiced transition. Under some common conditions, such as a large positive segmental shift at the edge of a voiced region (e.g., van Santen and Hirschberg, 1994, Fig. 4), the bracketed maximum should provide a substantially better estimate of the position of the peak of $p(t)$ than would the simple maximum. However, the average accuracy relative to the simple maximum is not trivial to predict.

2. Prosodic maximum in unvoiced region

In this case, neither the simple maximum nor the algorithm presented here can accurately mark the prosodic peak. The bracketed maximum will typically be less accurate than the simple maximum, as it is biased away from unvoiced

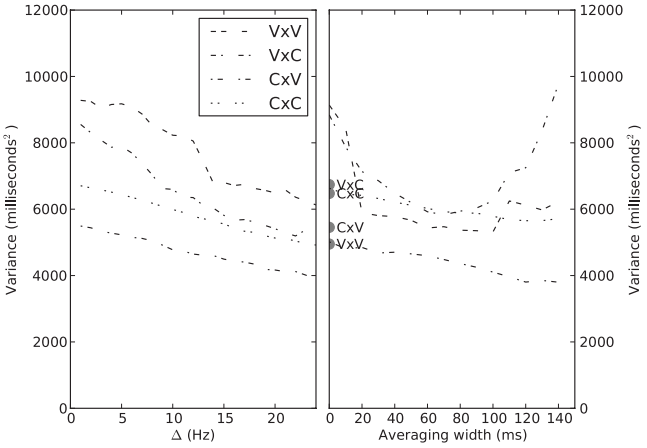


FIG. 15. Left: variance of the bracketed maximum estimate of the prosodic peak position, plotted against Δ. The lines show the variance of each section; the dotted-dashed pattern indicates the form of the word (Sec. III B). Right: variance of algorithms that median smooth the f_0 data and then take the simple maximum. The variance is plotted against the width of the smoothing window. (The appropriately weighted average of these curves corresponds to the dotted-dashed line in Fig. 13, right.) Results for MOMEL are shown by the large gray dots, and are labeled by the form of the word.

regions. However, from the point of view of human-to-human communication, this ought not to be an important case. If the precise timing of f_0 peaks is indeed an important part of the language, it seems unlikely that the language's phonological rules would evolve in such a way as to put the peaks where they cannot be heard.

APPENDIX C: SECTIONS OF THE CORPUS

The corpus (Sec. III B) has a mixture of four different frames, and Fig. 13 shows the average over the entire corpus. However, the different frames could have different behaviors. We check this by re-plotting Fig. 13 with the data from each frame separately (see Figs. 15 and 16).

Figures 15 and 16 show that the various components of the corpus have qualitatively similar behaviors, especially for the bracketed maximum algorithm. Variances are generally lower for words beginning with a consonant than for those that begin with vowels. (These words are also more common, so they dominate the overall average.) Possibly, f_0 shifts caused by consonants constrain the peak position, thus reducing the variance. Figure C shows somewhat different behavior for *Cx* words vs *Vx* words: The variance of the latter rises when a large smoothing window is used.

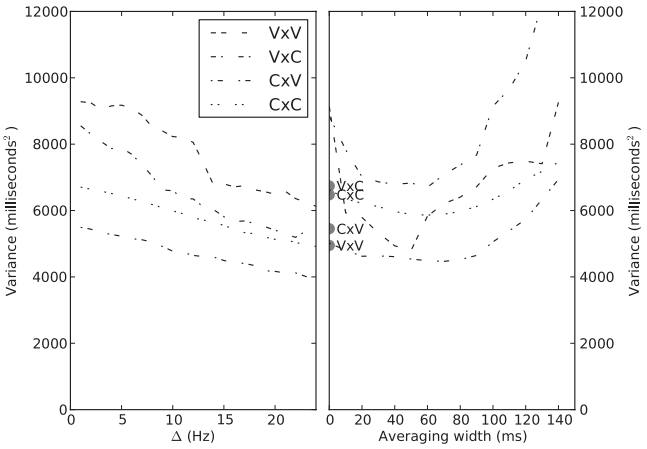


FIG. 16. Left: variance of the bracketed maximum, plotted as per Fig. 15. Right: variance of algorithms that mean smooth the f_0 data and then take the simple maximum.

[1]This term assumes that there is an underlying prosody that you are attempting to produce (e.g., this is your linguistic competence), but that the actual performance is disturbed by other effects. "Prosodic peak" refers to the peak of the underlying prosody. Equation (1) provides a mathematical model of this idea.

[2]For instance, one cannot describe animals with a superposition model of heads onto bodies. This is because some combinations, such as putting a horse's head on a hummingbird's body, do not yield biologically valid animals.

[3]The model in van Santen and Hirschberg, 1994 and van Santen *et al.*, 1998 is not strictly a superposition model because the segmental effects are not independent of the prosodic contour. The prosodic contour in their model explicitly depends on the segmental structure.

[4]This independence assumption implies that segment-related frequency shifts are the same in high f_0 regions as in regions with low f_0. For the purposes of this paper, it does not much matter whether this is true on a linear or a logarithmic (or other) frequency scale.

[5]Alternatively, if segmental frequency shifts are a core part of the language, then some of the available intonation information is being used to help listeners discriminate vowels. This alternative implies that, to be unambiguous to a listener, prosodically meaningful pitch motions would be at least as large as segmentally related f_0 shifts, which would strengthen the arguments in Sec. II E.

[6]However, one can imagine that the example might correspond to a nonsense phrase like "mamimamimami…."

[7]The possibility of seeing spurious f_0 maxima due to segmental effects was mentioned in van Santen *et al.*, 1998.

[8]For example, Gussenhoven (1999) argued (based on Pierrehumbert and Steele, 1989) that "…, if subjects were to produce a bimodal distribution of peak times in their imitations, then the difference must be categorical." Their deduction can now be seen to be incorrect, given the counterexample presented here that is bimodal but not discrete.

[9]The misinterpretation need not be on the part of a linguist. To the extent that peak positions are perceptually important, the listener might also perceive a sharp, phonological distinction that was not necessarily intended by the speaker. Possibly, this effect could lead to listeners later intentionally producing a bimodal distribution of peak positions, which might eventually lead to a phonological distinction becoming part of the language. Such an effect is presumably most likely on common phrases, where the same segmental structure might be used with a variety of intonation patterns [Elinor Payne, private communication (9/2008)]

[10]By "simple maximum," I mean the time of the peak f_0. Mathematically, this is often written as $\arg\max\{f(t)\}$.

[11]Hawkins and White (1988) argued that the builders of Stonehenge used this idea to mark the day of most northerly sunrise. So, this technique may have been discovered before c. 2800 BCE.

[12]Two measurements are approximately independent if knowledge of one thing is only very slightly useful in predicting the other.

[13]If declination and/or down-step were well enough understood to be predictable without reference to the speech data, it might make sense to use a larger Δ on the forward side. However, such an algorithm might misbehave badly at places where the right side of a peak is higher than the left.

[14]Note that these experiments involve comparisons of sentences with identical sequences of phones, so the listener does not need to compensate for segmental effects. The listener's task in these experiments is thus easier than in natural speech, and these discrimination thresholds should thus provide a lower bound on how much f_0 shift is necessary to indicate an accent under more realistic conditions.

[15]Note that high f_0 peaks are neither the only nor the best indicator of prominence. See Kochanski *et al.*, 2005 for discussion of this point.

Abel, J. S. (**1990**). "Optimal sensor placement for passive source localization," in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP-90, Albuquerque, NM, 3–6 April, Vol. **5**, pp. 2927–2930.

Arvaniti, A., Ladd, D. R., and Mennen, I. (**1998**). "Stability of tonal align-

ment: The case of Greek prenuclear accents," J. Phonetics **26**, 3–25.
Baken, R. J. (**1987**). *Clinical Measurement of Speech and Voice* (Taylor & Francis, London).
Braun, B., Kochanski, G., Grabe, E., and Rosner, B. S. (**2006**). "Evidence for attractors in English intonation," J. Acoust. Soc. Am. **119**, 4006–4015.
Chen, A., Gussenhoven, C., and Rietveld, T. (**2004**). "Language-specificity in the perception of paralinguistic intonational meaning," Lang Speech **47**, 311–349.
Chuang, C.-K., and Wang, W. S.-Y. (**1978**). "Psychophysical pitch biases related to vowel quality, intensity difference, and sequential order," J. Acoust. Soc. Am. **64**, 1004–1014.
Crandall, J. B. (**1925**). "The sounds of speech," Bell Syst. Tech. J. **4**, 586–626.
Di Cristo, A., and Hirst, D. J. (**1986**). "Modeling French micromelody: Analysis and synthesis," Phonetica **43**, 11–30.
Dilley, L. C., Ladd, D. R., and Schepman, A. (**2005**). "Alignment of L and H in bitonal pitch accents: Testing two hypotheses," J. Phonetics **33**, 115–119.
Fujisaki, H. (**1983**). "Dynamic characteristics of voice fundamental frequency in speech and singing," in *The Production of Speech*, edited by P. F. MacNeilage (Springer, New York), pp. 39–55.
Greenberg, S., Chang, S., and Hitchcock, L. (**2001**). "The relation between stress accent and vocalic identity in spontaneous American English discourse," in *Prosody-2001*, presented at the Prosody in Speech Recognition and Understanding Conference, Red Bank, NJ, 22–24 October, Paper No. 9.
Gussenhoven, C. (**1999**). "Discreteness and gradience in intonational contrasts," Lang Speech **42**, 283–305.
Hawkins, G. S., and White, J. B. (**1988**). *Stonehenge Decoded* (Hippocrene Books, New York).
Hirst, D., and Espesser, R. (**1993**). "Automatic modelling of fundamental frequency using a quadratic spline function," *Travaux de l'Institute de Phonétique d'Aix* (Universite de Provence, Aix de Provence, France), Vol. **15**, pp. 75–85.
Holm, B., and Bailly, G. (**2002**). "Learning the hidden structure of intonation: Implementing various functions of prosody," in Proceedings of the Speech Prosody 2002.
House, A. S., and Fairbanks, G. (**1953**). "The influence of consonant environment upon the secondary acoustic characteristics of vowels," J. Acoust. Soc. Am. **25**, 105–113.
House, D. (**2003**). "Perceiving question intonation: The role of pre-focal pause and delayed focal peak," in Proceedings of the 15th International Congress of the Phonetic Sciences (ICPhS), Barcelona, Spain, pp. 755–758.
House, J., Dankovicova, J., and Huckvale, M. (**1999**). "An integrated prosodic approach to speech synthesis," in *Proceedings of the XIV ICPhS (International Congress of the Phonetic Sciences)*, San Francisco, CA (University of California, Berkeley, CA), Vol. **3**, pp. 2343–2346.
Knight, R.-A. (**2002**). "The influence of pitch span on intonational plateaux," in *Proceedings of the Speech Prosody Conference* (Laboratoire Parole et Langage, Université de Provence, Aix en Provence, France), pp. 439–442.
Knight, R.-A., and Nolan, F. (**2006**). "The effect of pitch span on intonational plateaux," J. Int. Phonetic Assoc. **36**, 21–38.
Kochanski, G., Grabe, E., Coleman, J., and Rosner, B. (**2005**). "Loudness predicts prominence: Fundamental frequency lends little," J. Acoust. Soc. Am. **118**, 1038–1054.
Kochanski, G., and Orphanidou, C. (**2008**). "What marks the beat of speech?," J. Acoust. Soc. Am. **123**, 2780–2791.
Krolik, J. L. (**1996**). "The performance of matched-field beamformers with Mediterranean vertical array data," IEEE Trans. Signal Process. **44**, 2605–2611.
Ladd, D. R. (**1996**). *Intonational Phonology* (Cambridge University Press, Cambridge).
Ladd, D. R., Faulkner, D., Faulkner, H., and Schepman, A. (**1999**). "Constant 'segmental anchoring' of f_0 movements under changes in speech rate," J. Acoust. Soc. Am. **106**, 1543–1554.
Ladd, D. R., Mennen, I., and Schepman, A. (**2000**). "Phonological conditioning of peak alignment in rising pitch accents in Dutch," J. Acoust. Soc. Am. **107**, 2685–2696.
Ladefoged, P. (**1964**). "Some possibilities in speech synthesis," Lang Speech **7**, 205–214.
Mack, M. A., and Gold, B. (**1984**). "The discrimination of pitch in pulse trains and speech," Technical Report No. 680, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA.
Morlec, Y., Bailly, B., and Aubergé, V. (**1996**). "Generating intonation by superposing gestures," in *Proceedings of the Fourth International Conference on Spoken Language Processing, ICSLP 96* (International Speech Communications Association, Philadelphia, PA).
Ogden, R., Hawkins, S., House, J., Huckvale, M., Local, J., Carter, P., Dancovicova, J., and Heid, S. (**2000**). "Prosynth: An integrated prosodic approach to device independent, natural-sounding speech synthesis," Comput. Speech Lang. **14**, 177–210.
Peng, S.-H. (**2000**). "Lexical versus 'phonological' representations of Mandarin sandhi tones," in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, edited by M. B. Broe and J. B. Pierrehumbert (Cambridge University Press, Cambridge), pp. 152–167.
Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**, 175–184.
Pierrehumbert, J. B., and Steele, S. A. (**1989**). "Categories of tonal alignment in English," Phonetica **46**, 181–196.
Ross, K. N., and Ostendorf, M. (**1999**). "A dynamical system model for generating fundamental frequency for speech synthesis," IEEE Trans. Speech Audio Process. **7**, 295–309.
Sachs, T. S., Meyer, C. H., Irarrazabal, P., Hu, B. S., Nishimura, D. G., and Macovski, A. (**1995**). "The diminishing variance algorithm for real-time reduction of motion artifacts in MRI," Magn. Reson. Med. **34**, 412–422.
See EPAPS supplementary material at http://dx.doi.org/10.1121/1.3268511 Document No. E-JASMAN-127-018001 for Python source code for the bracketed maximum algorithm in Sec. II D.
Silverman, K. (**1987**). "The structure and processing of fundamental frequency contours," Ph.D. thesis, Cambridge University, Cambridge.
Silverman, K., and Pierrehumbert, J. (**1990**). "The timing of prenuclear high accents in English," in *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, edited by J. Kingston and M. Beckman (Cambridge University Press, Cambridge), pp. 72–106.
Slater, A., and Coleman, J. (**1996**). "Non-segmental analysis and synthesis based on a speech database," in Proceedings of the ICSLP 96, Fourth International Conference on Spoken Language Processing, edited by H. T. Bunnell and W. Idsardi, Vol. **4**, pp. 2379–2382.
Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT, Cambridge, MA).
Taylor, H. C. (**1933**). "The fundamental pitch of English vowels," J. Exp. Psychol. **16**, 565–582.
Taylor, P. (**1993**). "Automatic recognition of intonation from f_0 contours using the rise/fall/connection model," in *Proceedings of the Third European Conference on Speech Communication and Technology (EUROSPEECH '93)* (International Speech Communications Organization, Berlin), Vol. **2**.
van Santen, J., and Hirschberg, J. (**1994**). "Segmental effects on timing and height of pitch contours," in Proceedings of the ICSLP-94, Third International Conference on Spoken Language Processing, Yokohama, Japan, pp. 719–722.
van Santen, J. P. H., and Möbius, B. (**1999**). "A quantitative model of f_0 generation and alignment," in *Intonation Analysis, Modelling and Technology*, edited by A. Botinis (Kluwer, Dordrecht), pp. 269–288.
van Santen, J. P. H., Möbius, B., Venditti, J. J., and Shih, C. (**1998**). "Description of the Bell Labs intonation system," in *Proceedings of the Third International Workshop on Speech Synthesis* (International Speech Communications Organization, Jenolan Caves House, Blue Mountains, NSW, Australia), pp. 293–398.
Whalen, D. H., and Levitt, A. G. (**1995**). "The universality of intrinsic f_0 of vowels," J. Phonetics **17**, 193–203.
Wichmann, A., House, J., and Rietveld, T. (**1997**). "Peak displacement and topic structure," *Intonation: Theory, Models, and Applications* (International Speech Communications Organization, Athens, Greece).
Xu, C. X., and Xu, Y. (**2003**). "f_0 perturbations by consonants and their implications on tone recognition," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '03* (IEEE, Los Alamitos, CA), Vol. **1**, pp. 456–459.
Xu, Y., and Sun, X. J. (**2000**). "How fast can we really change pitch? Maximum speed of pitch change revisited," in Proceedings of the International Conference on Spoken Language Processing 2000.

Acoustic censusing using automatic vocalization classification and identity recognition

Kuntoro Adi
Santa Dharma University, Mrican, Yogyakarta 55002, Indonesia

Michael T. Johnson[a)]
Speech and Signal Processing Laboratory, Marquette University, 1515 West Wisconsin Avenue, Milwaukee, Wisconsin 53201-1881

Tomasz S. Osiejuk
Department of Behavioural Ecology, Adam Mickiewicz University, Umultowska 89, 61–614 Poznań, Poland

(Received 29 June 2009; revised 16 October 2009; accepted 16 November 2009)

This paper presents an advanced method to acoustically assess animal abundance. The framework combines supervised classification (song-type and individual identity recognition), unsupervised classification (individual identity clustering), and the mark-recapture model of abundance estimation. The underlying algorithm is based on clustering using hidden Markov models (HMMs) and Gaussian mixture models (GMMs) similar to methods used in the speech recognition community for tasks such as speaker identification and clustering. Initial experiments using a Norwegian ortolan bunting (*Emberiza hortulana*) data set show the feasibility and effectiveness of the approach. Individually distinct acoustic features have been observed in a wide range of animal species, and this combined with the widespread success of speaker identification and verification methods for human speech suggests that robust automatic identification of individuals from their vocalizations is attainable. Only a few studies, however, have yet attempted to use individual acoustic distinctiveness to directly assess population density and structure. The approach introduced here offers a direct mechanism for using individual vocal variability to create simpler and more accurate population assessment tools in vocally active species.
 [DOI: 10.1121/1.3273887]

PACS number(s): 43.60.Bf, 43.80.Ka [WWA] Pages: 874–883

I. INTRODUCTION

Individually distinct acoustic features have been observed in a wide range of vocally active animal species, for example, cetaceans (Janik, 2000), bats (Masters *et al.*, 1995), and primates (Butynski *et al.*, 1992). There is strong evidence to suggest that individual identification from vocalizations is possible in many species, just as it is in humans, and that many of the state-of-the-art techniques for robust human speaker identification and clustering (Reynolds and Rose, 1995; Tranter and Reynolds, 2006) can be applied equally well to animal vocalizations.

Within birds, the presence of vocal individuality has been shown in the European bitterns (*Botaurus stellaris*) and Black-throated divers (*Gavia arctica*) (Gilbert *et al.*, 1994), American woodcock (*Scolopax minor*) (Beightol and Samuel, 1973), Australian kingfishers (*Dacelo novaeguineae*) (Saunders and Wooler, 1978), and Tawny owls (*Strix aluco*) (Galeotti and Pavan, 1991). Birds use vocal differences to identify other members of their species nearby and to identify individual birds in their immediate vicinity. They have been shown to use vocalizations in recognizing their mates, their parents, and in differentiating between neighbors and strangers (Holschuh, 2004). While a wide variety of approaches has been used to count and monitor bird populations within a species (Peake and McGregor, 2001), most of those approaches do not use individual vocal variability or require the identification of individual birds.

For rare or elusive species that are hard to monitor or to mark visually, the possibility of recognizing individuals by their vocalizations may provide a useful census tool, e.g., Saunders and Wooler, 1978 and Gilbert *et al.*, 1994, but only a few researchers have used individual variation to assess population structure, abundance and density, seasonal distribution and trends, or impact of human-made noise on animals (Mellinger and Barlow, 2003). Peake and McGregor (2001) employed a statistical Pearson-correlation approach to identify corncrake (*Crex crex*) vocal individuality and to estimate numbers of individuals in species. Holschuh (2004) used discriminant function analysis (DFA) to explore vocal individuality of the saw-whet owl (*Aegolius acadicus*) to monitor its habitat quality. Terry and McGregor (2002) suggested neural network models to monitor and census male corncrake species, using a backpropagation and probabilistic network to re-identify the members of the known population (monitoring task) and a Kohonen network to count a population of unknown size. In cetaceans, there are several examples of the use of vocalizations in assessment, including sperm whales (*Physeter macrocephalus*), humpback whales (*Megaptera novaeangliae*), Cuvier's beaked whales (*Ziphius cavirostris*), and harbor porpoises (*Phocoena phocoena*)

[a)]Author to whom correspondence should be addressed. Electronic mail: mike.johnson@marquette.edu

(Barlow and Taylor, 1998; Mellinger and Barlow, 2003; Marques *et al.*, 2009), where combined visual and acoustic methods have significantly improved the population estimate.

Numerous quantitative approaches for analyzing vocal individuality exist. Otter (1996) was able to differentiate individual birds through a series of nested Analysis of Variance (ANOVA) Holschuh (2004) did the same using DFA. In marine mammals, Buck and Tyack (1993) utilized Dynamic Time Warping (DTW) to classify 15 bottlenose dolphin (*Tursiops truncatus*) signature whistles into five groups. The research presented here adapts a well established automatic human speech recognition framework to the task of acoustic censusing, i.e., estimating the abundance of animals in a specified survey area. Previous and current studies show the feasibility of a hidden Markov model (HMM) method to automatically classify ortolan bunting song-types, to identify individual birds (Trawicki *et al.*, 2005; Adi, 2008; Adi *et al.*, 2008), to identify individual African elephants (*Africana Loxodonta*) (Clemins *et al.*, 2005), and to cluster beluga whale (*Delphinapterus leucas*) repertoires (Adi, 2008). The proposed framework is based on HMMs and Gaussian mixture models (GMMs), both commonly used in the speech processing community to perform speech recognition and speaker identification and verification. This approach has advantages in its flexibility and robustness to duration and temporal alignment differences between training and testing examples. An integration of several techniques for supervised and unsupervised classification is proposed and combined with the mark-recapture approach to estimate animal abundance.

Following this introduction, Sec. II gives an overview of the study population, introduces two protocols for estimating the number of animals in a population, and discusses the methods behind the key individual subtasks. Section III presents the experimental results, and Sec. IV finishes with overall conclusions.

II. METHODS

A. Study population: Ortolan bunting (*Emberiza hortulana*)

The ortolan bunting is a migratory passerine bird distributed from Western Europe to Mongolia (Cramp and Perrins, 1994). They winter in Africa. The species inhabits open agricultural areas, raised peat bogs, clear-cut forest on poor sand, and cleared farmland and forest burn (Dale and Hagen, 1997).

Ortolan buntings are classified as an endangered species (Steinberg, 1983; Dale, 2001a) and have shown a major population decline both in individual numbers and in their distribution. In Finland, Vepsalainen *et al.* (2005) studied their population density changes and environment associations in years 1984–2002. They observed a population crash between 1990 and 1993, resulting in a 54% reduction in population density, with a total density reduction of 72% between years of 1984–2002. The Norwegian ortolan bunting, meanwhile, currently numbers approximately 150 singing males and has shown decline over the past 50 years as well. In years 1996–2000 the decline rate was 8% per year (Dale, 2001a). The decline is most likely related to female-biased dispersal away from the population which results in many unpaired males and low population productivity (Dale, 2001b; Steifetten and Dale, 2006). Recently, it was revealed that ortolan bunting males are able to discriminate vocally between neighbors and strangers based on single song derived from the repertoire of a particular male (Skierczyński *et al.*, 2007).

The ortolan bunting vocalizations being examined for this study were collected from County Hedmark, Norway in May of 2001 and 2002 (Osiejuk *et al.*, 2003). The male vocalizations were recorded on 11 out of 25 sites within an area of about 500 km^2. The total number of males in the covered area of the years 2001 and 2002 was about 150. Individual identity was determined for a high percentage of the vocalizations, based on visual observation of individuals using wing markings. For the purposes of this study the 2001 data are used as training data ("marking" data in the mark-recapture protocol) for building song-type and individual identity classification models and for determining repertoire statistics, while the 2002 data are used as test data ("recapture" data) for classification and overall censusing evaluation. It is also used to demonstrate the method of overall abundance estimation using an acoustic mark-recapture model, although because the data were not collected originally for that purpose and does not meet the time and locality guidelines for a mark-recapture study, the mark-recapture model estimates given here are only illustrative.

The ortolan bunting has a relatively simple song and small repertoire size of typically 2–3 song-types for each individual. Individual repertoires are relatively stable but do change some over time. However, in the Norwegian population males use on average 4.2 song-types and have a repertoire size of between 1 and 24 (Łosak, 2007). Song frequencies are in a range between 1.9 and 6.7 kHz. As described by Osiejuk *et al.* (2003), these ortolan vocalizations were recorded between 04:00 and 11:00 am by using an HHB PDR 1000 professional Digital Audio Tape (DAT) recorder. All recordings were transferred to a PC using 48 kHz/16 bit sampling.

The ortolan bunting has a relatively simple repertoire, with songs described by syllable, song-type, and song-variant. A syllable, the minimal unit of song production, is described using letter notation, as illustrated in Fig. 1. Syllables are grouped together into patterns, with each general pattern called a song-type and each unique song called a song-variant. For example, *b* and *c* are examples of syllables, *cb* is an example of a song-type pattern consisting of one or more instances of a *c* syllable followed by one or more instances of a *b* syllable, and *cccccb* is an example of a specific song-variant within the *cb* song-type, consisting of exactly five repetitions of a *c* syllable followed by exactly one repetition of a *b* syllable. Although syllables of the same type differ in length and frequency between individuals, they have the same basic spectrogram shape. Figure 2 illustrates an example of this song-variant for two different individuals in

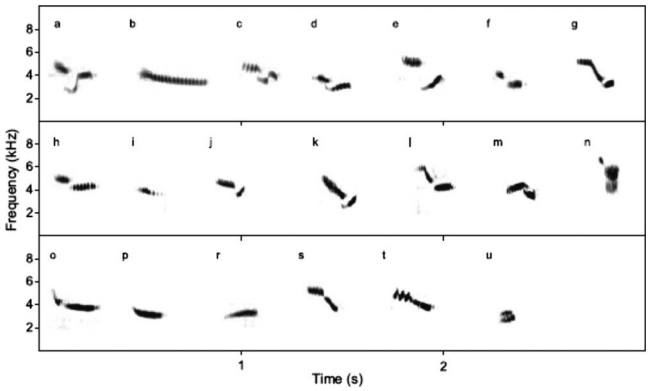


FIG. 1. Ortolan bunting syllables (after Osiejuk *et al.*, 2003).

this data set. In total, the data sets used here include 63 different song-types and 234 different song-variants composed of 20 different syllables.

B. Overview of population assessment protocols

This research proposes two separate population assessment protocols. Both protocols assume single species data, which may require the preprocessing step of species classification (in most cases a somewhat simpler task than song-type and identity classification). In the simpler single-pass protocol illustrated in Fig. 3, a single acoustic data set is used for analysis, with training data limited to enough song-type labeled repertoire examples that classification models can be built, typically five to ten examples of each song-type. In this case song-type classification and individual clustering methods can be used to estimate the total number of individuals within each song-type, and repertoire statistics can then be used to give an overall population estimate within that data set, with confidence intervals.

In the more complex mark-recapture protocol illustrated in Fig. 4, two acoustic data sets are collected in accordance with a closed population mark-recapture survey design. The first set is labeled with song-type and individual identity and used to build song-type and identity classification models.

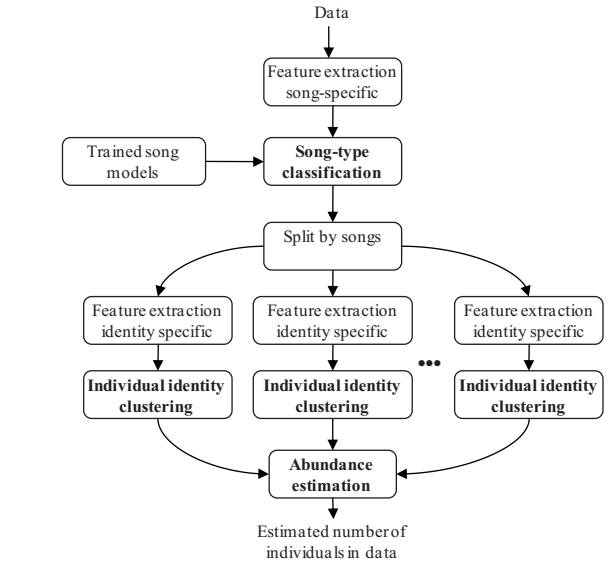


FIG. 3. Single pass protocol, resulting in an estimate of the number of individuals present in the acoustic data set under analysis. Required subtasks include song-type classification (Sec. II C 1), individual clustering (Sec. II C 2), and abundance estimation weighted by repertoire statistics (Sec. II C 3).

The second data set then uses those models to automatically split vocalizations into song-types using song-type classification, cluster individuals within each song type using identity classification, and then match up the individuals in each set to obtain the number of "marked" and "recaptured" animals for total population estimation.

There are four distinct tasks needed to implement these protocols: song-type classification, identity clustering, identity cluster matching, and abundance estimation. Song-type classification trains repertoire models using labeled data and then classifies unknown data. Identity clustering groups data, within one specific song-type, to find the number of clusters representing unique individuals. Identity matching then matches those identity clusters in the unlabeled data to a specific known individual in the labeled training set, using a speaker verification model. The final task is the population

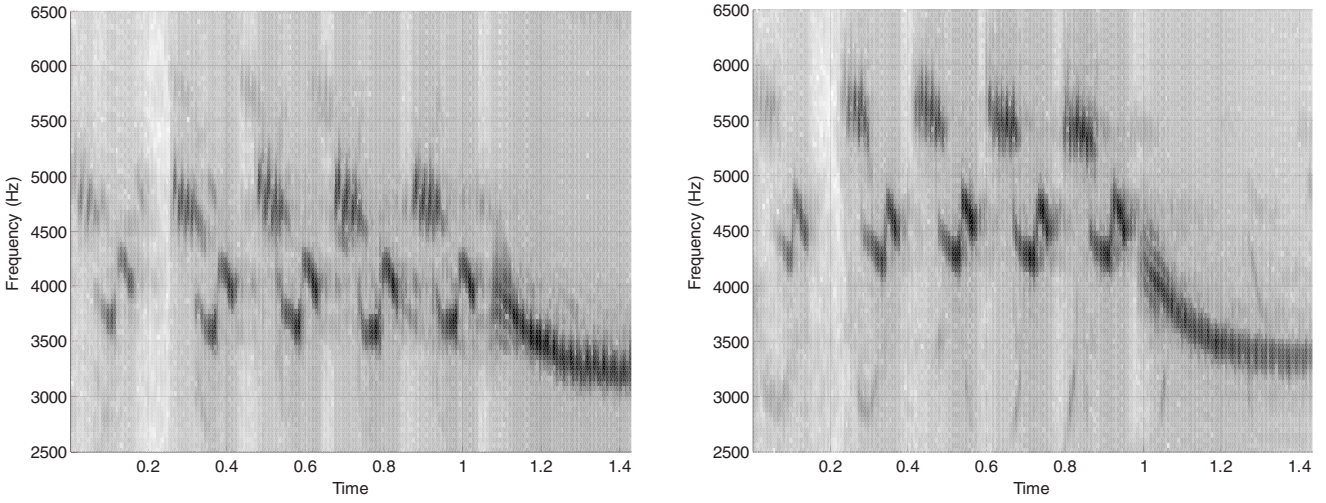


FIG. 2. Examples of song-type *cb*, song-variant *ccccb*, from two different individuals. Note the similarity in basic pattern but difference in timing and mean frequency level.

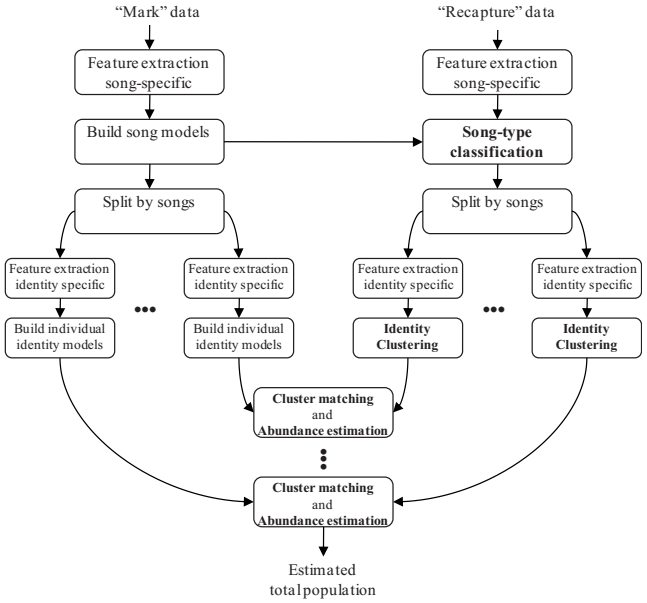


FIG. 4. Mark-recapture protocol, resulting in a net population estimate. The data set on the left represents the "mark" data set used for building both song-type models and individual identity models. The data set on the right represents the "recapture" data set, which is separated into songs using song-type classification (Sec. II C 1), separated by individual using individual clustering (Sec. II C 2), matched across data sets to identify "recaptures" using cluster matching (Sec. II C 3), after which a mark-recapture abundance estimation model (Sec. II C 4) is used to estimate total population.

assessment itself, which is accomplished in the single-pass protocol using a weighted averaging of sub-population estimates within each song-type and is accomplished in the mark-recapture protocol using established maximum-likelihood methods for abundance estimation, based on the estimate of individuals in the two data sets plus how many were present in both.

Section II C discusses in more detail the separate tasks involved in the above scenarios, including song-type classification, individual identity clustering, individual cluster matching, and population size estimation.

C. Individual sub-task methodology

1. Song-type classification using HMMs

Song-type classification is implemented using HMMs (Clemins and Johnson, 2003; Trawicki *et al.*, 2005). Classification features are based on Greenwood function cepstral coefficients (GFCCs) (Clemins, 2005; Clemins and Johnson, 2006), normalized to minimize individual vocal variability. Specifically, a 39-element feature vector is calculated consisting of the 12 GFCCs plus energy, appended with first and second derivatives. The waveforms are first Hamming windowed with frame-size of 3 ms and overlap of 1.5 ms. The Greenwood frequency warping constants are calculated using 26 filter banks spaced across the ortolan bunting hearing range of $f_{\min}$ 400 Hz to $f_{\max}$ 7400 Hz (Edwards, 1943) for each frame.

Song-type classification is implemented using syllables as the base unit. Each syllable is represented by a 15-state HMM to capture the temporal pattern of that syllable. The

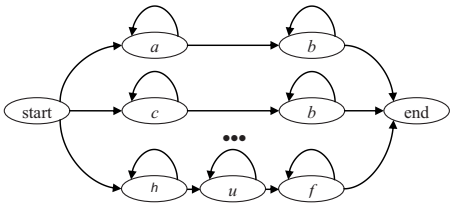


FIG. 5. Song-type classification language model constraints.

corresponding HMMs are connected together for training and recognition using a song-type language model which constrains the output to a valid song-type. Figures 5 and 6 illustrate the language model and waveform-to-HMM matching process.

The HMMs are trained using the Baum–Welch algorithm (Baum *et al.*, 1970), a maximum likelihood estimation (MLE) method based on expectation maximization. Classification on new data is accomplished using the Viterbi algorithm (Forney, 1973) to identify the most likely syllable sequence given the waveform. All HMM tasks for these experiments were implemented using the Cambridge University HMM Toolkit (HTK) version 3.2 (Cambridge University Engineering Department, 2002).

2. Individual identity clustering using GMMs and deltaBIC analysis

The clustering of vocalizations according to individual identity focuses on accurately estimating the number of unique individuals, each of which is represented by a cluster. Because the number of individuals is completely unknown and some individuals may vocalize very few times, this task is much more difficult than that of song-type clustering. The approach used here is based on a method called deltaBIC (BIC denotes Bayesian information criterion) analysis, commonly used in human speech recognition systems for speaker diarization, the task of associating dialog segments in a conversation with specific individuals (Trantor and Reynolds, 2006). DeltaBIC analysis (Ajmera and Wooters, 2003) uses GMMs rather than HMMs as a model for each individual.

The deltaBIC method is based on differential values of the BIC as the number of clusters is increased. The BIC value itself is a similarity measure between two probability

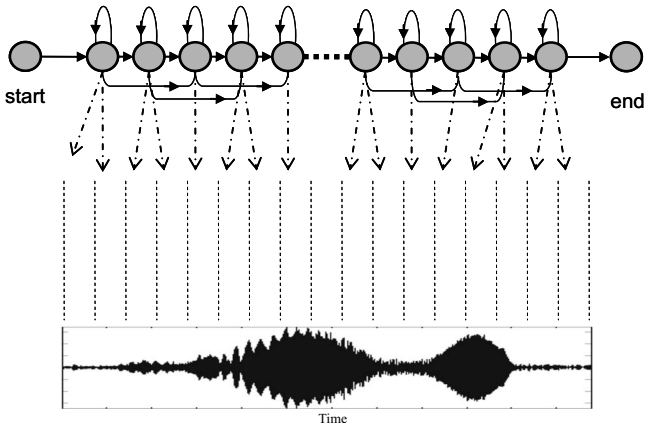


FIG. 6. Illustration of HMM to waveform matching for classification.

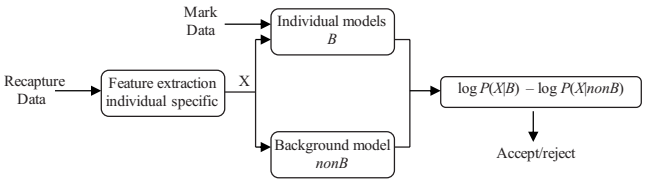


FIG. 7. Individual verification process for individual-to-cluster matching.

density functions, here GMMs. The process starts with over-clustering of the data sets and iteratively merges clusters and retrains a new cluster until no pair of clusters is left with a positive deltaBIC distance measure. The deltaBIC measure is given by

$$\text{deltaBIC} = \sum_{X \in D} \log p(X|\theta) - \sum_{X \in D_1} \log p(X|\theta_1) - \sum_{X \in D_2} \log p(X|\theta_2), \quad (1)$$

where x represents the feature vectors for each frame and D_1, θ_1 and D_2, θ_2 represent the two clusters and cluster parameters being considered for merging. When the clustering process is complete, the remaining number of clusters is used as an estimate of the number of individuals.

Features for individual bird clustering consist of a 39-element GFCC feature vector similar to that used for song-type classification, except implemented without mean and variance normalization in order to preserve individually specific vocal characteristics. The individual bird models are 15 mixture GMMs, implemented in HTK using a single-state HMM with a GMM observation model. The modeling and clustering process is always done on data that consist of a single song-type, i.e., individual identity clustering is always performed after song-type classification or clustering, so that differences in feature characteristics can be reliably associated with individual differences rather than vocalization differences.

3. Identity cluster matching using speaker verification models

To implement an acoustic mark-recapture protocol, it is necessary to match the identity clusters from the recapture data set to known individuals in the marking data set in order to find the overlap, or "recaptures," between the sets. The process is similar to the HMM-based classification as done for song-types, but is implemented using un-normalized GFCC features as with the individual clustering, and also adds a verification step to allow for classifying a vocalization or group of vocalizations as unknown. This is accomplished using a basic likelihood-ratio speaker verification approach, as used in the field of human speaker recognition. The process is implemented separately for each song-type, and then individual results are globally combined. Figure 7 gives an overview of this task.

To implement this, 15-state HMMs (one for each specific song-type) are created for each known individual in the training set. In addition, a verification model called a universal background model is created for each song-type across all individuals in the data set. A likelihood-ratio test is then implemented using all vocalizations in a cluster to discriminate whether the vocalizations in that cluster come from a specific known individual or represent a new unseen individual. The threshold of this accept/reject decision can be varied to control the degree of confidence required to verify that the cluster vocalizations match a known individual.

4. Abundance estimation

a. Single-pass protocol: Data set population estimation using song distribution statistics. The single-pass protocol is able to estimate the number of individuals in the data set under analysis based on a known repertoire distribution for each song-type. Within each song-type, an overall local population estimate is obtained from the estimated number of individuals within that song-type, found through identity clustering, combined with knowledge of how many individuals within the population make that particular song. The final estimated population is then the estimated number of birds within a song type group divided by the percentage of birds that typically use that song type.

These population estimates are then combined using an average or a weighted average according to overall song-type occurrence, leading to a final data set estimate as well as a variance that represents margin of estimation error due to dissimilarity in individual estimates.

If repertoire information regarding the percentage of individuals who make specific song-types is unknown, it is not possible to directly combine the population estimates for each song-type into an overall local population estimate. However, in this case upper and lower bounds on the population can still be established, with a lower bound equal to the maximum number of individuals in any one category (an implicit assumption that all individuals make that song-type) and an upper bound equal to the sum of individuals in each category (an implicit assumption that no individuals make more than one type of song).

b. Abundance estimation using a mark-recapture model. The mark-recapture protocol addresses the bird abundance estimation problem using the MLE framework of a standard mark-recapture model. A two sample mark-recapture involves one session of catching and marking, and another session of recapturing. In the context of this study, catching refers to recording bird vocalizations in an initial session, marking refers to labeling the vocalizations with specific identities, and recapturing refers to recording a second set of vocalizations and acoustically comparing identities.

The process of labeling and recapture or re-labeling involves the tasks of supervised recognition, unsupervised clustering, and identity cluster matching discussed in sections II C I, II C 2, and II C3. The previous steps, therefore, provide the number of individuals (u_1) in one data set, the estimated number of individuals (u_2) in the second data set, and the estimated number of individuals present in both data sets (m_2).

Given the variables u_1, m_2, and u_2, the likelihood of population estimate is computed using the Jolly–Seber (Seber, 1982) equation

TABLE I. Song-type recognition confusion matrix. Rows represent correct song-type categories, and columns represent the classifications made by the algorithm so that entries along the diagonal represent correct classifications (accuracy of 89.6%).

	ab	*c*	*cb*	*cd*	*eb*	*ef*	*gb*	*guf*	*h*	*hb*	*hd*	*huf*	*jd*	*kb*
ab	1561	3	24	1	0	0	1	0	0	5	0	0	0	9
c	0	53	1	11	0	0	1	0	2	0	0	7	0	0
cb	1	2	706	11	0	0	85	1	0	2	0	2	1	1
cd	0	0	9	434	0	0	0	0	0	0	0	0	5	0
eb	1	1	4	0	384	11	1	0	0	0	0	0	0	0
ef	0	0	0	0	0	57	0	0	0	0	0	0	0	1
gb	3	1	19	1	0	0	320	5	0	3	0	0	0	37
guf	0	0	2	0	0	0	1	130	0	0	0	6	0	0
h	0	32	44	0	0	0	3	0	138	27	0	0	0	17
hb	0	0	0	0	0	0	0	0	0	32	0	0	0	1
hd	0	0	0	0	0	0	0	0	0	5	8	3	0	0
huf	0	2	7	41	0	0	0	1	0	2	22	285	0	1
jd	0	0	1	2	0	0	0	2	0	0	0	0	47	0
kb	0	1	0	0	0	0	0	0	0	0	0	0	0	87

$$L(N,p) = \prod_{s=1}^{2} \binom{U_s}{u_s} p^{u_s}(1-p)^{U_s-u_s} \binom{M_s}{m_s} p^{m_s}(1-p)^{M_s-m_s} \tag{2}$$

as a function of the unknown variables N (population) and p (the probability of recapture). The N and p where the likelihood function achieves its maximum value is the maximum likelihood estimate of the population N.

III. EXPERIMENTAL RESULTS

Before system integration, the individual task components are evaluated separately, including song-type classification, individual identity clustering, and identity cluster matching. Following this each of the two protocols is evaluated separately.

A. Evaluation of separate sub-tasks

1. Song-type classification

Speaker independent song-type classification experiments are performed across 14 of the most common song-types. Each song-type contains multiple song-variants. As described previously, the experiment uses the 2001 data set for training and the 2002 data set for evaluation (experiments with those designations reversed yielded similar results).

Results are summarized in the classification confusion matrix in Table I. Individual song-type classification accuracy ranges from a low of 50% (song-type *hd*, which had only 16 examples, 5 of which were confused for type *hb*) to a high of 98.8% (song-type *kb*), with an average overall accuracy of 89.6%. Nearly all errors were made between two-syllable song-types where one syllable was the same and the other was closely related (often syllables *b* and *d*). Full confusion matrices and more detailed analysis of these results are available in Adi (2008). It is worth noting that perfect song-type separation is not necessary to do the larger task of counting individuals. Error in song-type classification tends to lead to an upward bias in the ensuing step of individual identity recognition and clustering. This can be compensated for by downward calibration of the final population estimation methods, although no such adjustment has been made in these experiments.

2. Individual identity clustering

The number of individuals in the test set was estimated using identity clustering via the deltaBIC analysis discussed in Sec. III A 1. Results of the deltaBIC clustering are shown in Fig. 8 (only a subset of the calls are shown for readability), with final estimated number of clusters, true number of individuals, and error shown in Table II. Error ranges from 2.9% to 50% and is notably lower for the more common song-types.

3. Identity cluster matching

An example of identity cluster matching is given in Table III for the *cb* song-type. Identity clusters from the target data are shown in each row, with known individuals from the training data in each column. The number in each cell represents the "acceptance value" from the likelihood-ratio

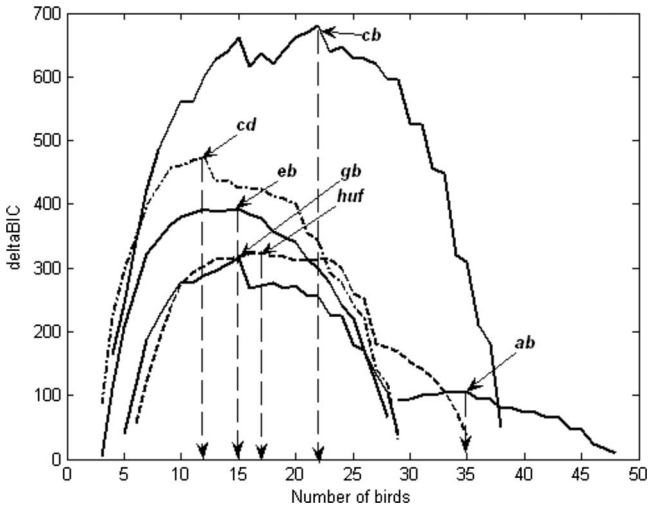


FIG. 8. DeltaBIC analysis curves for selected song-types. The peak value in each curve is used as the estimated number of individuals for that song-type.

TABLE II. Comparison of deltaBIC cluster estimates to known populations.

Song-type	Estimated number of individuals for song	Known number of individuals for song	Percent error
ab	35	34	2.9
cb	22	20	10.0
cd	12	9	33.3
eb	15	12	25.0
ef	3	2	50.0
gb	15	13	15.4
guf	6	4	50.0
huf	17	20	15.0
kb	9	7	28.6

test, with a positive value indicating that the cluster matched more closely to a specific individual than the background model. Here we have arbitrarily chosen a threshold of 1.0 as a cutoff point, which results in a match to two different individuals in the original set, noted b3 and b11. The actual number of individuals present in both the training and test data sets is 3 for a matching error of 33%.

B. Population estimation evaluation

1. Single-pass protocol results

In the single-pass protocol, the goal is to estimate the number of individuals within each song-type, and then use those estimates to arrive at an overall estimate of the population present within the data set. To do this, the calls are first separated by song-type and clustered for individual identity using the deltaBIC approach, and then the number of clusters is used to arrive at a local population estimate.

The song type-specific estimates for the test data were shown previously in Table II. These can be used to project overall local population estimates by dividing the song-type estimate by the percentage of individuals in the population that make each song. These separate estimates can then be combined through an average or weighted average (to emphasize the more frequent song-types which give more accurate estimates), as shown in Table IV.

TABLE IV. Single pass protocol estimates of number of individuals in data set.

Song	Proportion of individuals for each song (%)	Estimated number of individuals for each song	Estimated number of total individuals (true value=81)
ab	51.8	35	67.6
cb	26.8	22	82.1
cd	21.4	12	56.1
eb	12.5	15	120.0
ef	5.3	3	56.6
gb	26.8	15	56.0
guf	8.9	6	67.4
huf	10.7	17	158.9
kb	5.3	9	169.8
Average ± stdev (error)			92.7 ± 45 (14.4%)
Weighted average			79.0 ± 31 (2.4%)

The results indicate that the projected population estimates tend to be biased on the high side, likely due to addi-

TABLE III. Song-type *cb* identity matching results (overlap=2 individuals, *b*3 and *b*11).

		Known individuals from training set												
		b1	b2	b3	b4	b5	b6	b7	b8	b9	b10	b11	b12	b13
Test set cluster	c1	−6.1911	−4.8384	0.2598	−4.8071	−4.8664	−1.7287	−1.4306	−2.2258	−7.4660	−6.6986	−2.0243	−3.5749	−4.4040
	c2	−3.1360	−2.2224	**1.0759**	−3.5349	−3.7813	−2.4249	−1.1016	−2.9164	−5.1645	−3.5034	0.0294	−4.0007	−4.9570
	c3	−1.2931	−4.0089	−0.9428	−7.4533	0.5253	−5.3972	−5.5260	−2.3992	−7.3673	−6.3985	0.7298	−2.5238	−7.1605
	c4	−1.7191	−4.5681	0.1981	−7.0703	0.3497	−5.6525	−5.4115	−3.0178	−6.9223	−5.9974	**1.1603**	−3.5598	−7.2065
	c5	−3.7740	−3.0045	−0.9646	−3.2053	−2.8519	0.1312	−1.1781	−1.4876	−4.4815	−4.4405	−2.4751	−3.5888	−1.6194
	c6	−4.1879	−3.8325	−2.2324	−2.5864	−1.4738	−2.3203	−1.7527	−1.4110	−4.8890	−5.3197	−1.9983	−2.9655	−2.9405
	c7	−4.0524	−3.7983	−0.1179	−1.9972	−3.6972	−3.7014	0.0847	−1.5080	−5.6273	−6.7853	−1.0934	−3.2686	−4.7191
	c8	−1.6285	−1.0683	−2.5166	−3.7007	−1.7817	−2.0835	−2.3154	−1.9211	−4.3459	−4.5966	−0.9438	−3.1081	−2.8231
	c9	0.4203	−3.9631	−3.2250	−6.2160	−0.4092	−4.9076	−3.0624	−3.7383	−5.5757	−4.7141	0.8304	−4.6439	−6.6938
	c10	−6.1358	−4.7771	0.5048	−2.9390	−5.8793	−0.0643	−0.9827	−2.1916	−7.4385	−6.1976	−2.9593	−5.2991	−4.8353
	c11	−6.0416	−5.6546	−4.8061	−1.5755	−1.3735	−3.8146	−1.4443	−1.3929	−6.0052	−7.2179	−4.3271	−1.7333	−4.4404
	c12	−3.1280	−3.3624	−0.9467	−3.5546	−3.8155	−2.6841	−1.1312	−3.4759	−1.5937	−4.5810	−1.0143	−4.8189	−4.4083
	c13	−5.5911	−4.8649	0.2842	−6.4990	−4.7271	−4.2375	−2.6389	−1.0592	−7.7188	−6.7524	−0.0592	−3.8018	−6.1107
	c14	−6.4177	−4.8428	0.1138	−5.4141	−4.8679	−3.0886	−2.9369	−2.7645	−6.2715	−7.2595	−0.6921	−4.7532	−5.7999
	c15	−5.5897	−6.7359	0.1789	−3.8895	−4.4871	−1.3617	−2.4674	−4.0671	−6.9669	−1.7588	−3.1082	−6.2582	−5.5221
	c16	−3.1855	−5.3260	0.6257	−4.6915	−3.3572	−2.0442	−1.1992	−3.6259	−5.3410	−1.8490	−2.0575	−5.4783	−5.9606
	c17	−5.0415	−6.5233	−2.2900	−1.9306	−4.1976	−4.9813	−1.2144	−3.6470	−4.0607	−7.0770	−2.6779	−6.1275	−5.9435
	c18	−5.8263	−6.6272	−4.4404	−2.7344	−0.0812	−2.7804	−2.1755	−0.2400	−7.9769	−7.7126	−4.6130	0.1282	−4.1389
	c19	−1.8148	−1.8384	−1.4369	−3.8289	−1.8039	−3.3647	−1.9384	−3.0285	−4.5124	−3.7390	**1.1541**	−3.6040	−4.5305
	c20	−0.9310	−3.3875	−4.1278	−3.6522	0.4725	−4.5987	−1.8132	−1.5103	−4.6590	−5.4152	−1.8644	−3.8411	−4.7288
	c21	−3.6516	−4.8461	−0.7648	−2.9481	−3.1399	−3.7093	−0.2985	−1.6854	−4.7744	−6.5691	−1.6969	−3.7253	−4.4330
	c22	−4.6355	−4.2748	0.4865	−2.9744	−3.8810	−3.0470	−0.1036	−1.6721	−6.2508	−7.0564	−1.7838	−3.6584	−5.1683

tional non-individual variability present in the separate song-type clusters due to the song-type classification error (which as seen previously is typically in the 5%–15% range). There is also significant variance in the projected estimates, which occurs due to the original error in the per-song estimates combined with errors in the song distribution statistics. The song distribution statistics can have significant impact, especially for infrequently occurring song-types whose population estimates are amplified by a large factor when projecting to an overall estimate. For the data given here, there were, in fact, substantial differences between the training set call distribution used for statistics and the test set used for population estimates. For example, song-type *huf* had a distribution of just over 10% in the first data set, but that increased to nearly 25% in the second set. This resulted in a larger number of individuals in the test set for that song-type, but the results were projected using the original statistics, so that even though the per-call population estimate error was only 15%, the projected population error was over 90%. However, since changes in repertoire statistics by necessity go both ways, with some increasing and others decreasing, this source of error tends to balance itself out.

Weighting the individual population estimates toward the more frequent song-types not only significantly decreases the error, but also tends to lower the variance. It is also possible to compensate for the upward bias with a global adjustment factor (which would likely be species dependent but not data set dependent), but that has not been done here.

2. Mark-recapture protocol illustrative results

In the mark-recapture protocol, illustrated in Fig. 4, song-type classification and speaker identification are performed on the recapture data set to separate the calls by type and remove vocalizations from known individuals in the training set. Following this, identity clustering is performed on the remaining vocalizations and used to determine the total number of new individuals, and the mark-recapture method is used to estimate overall population abundance. Each song may be used to determine an abundance estimate, and the results combined to get a single final estimate.

As discussed previously, the two ortolan data sets used here were not collected in a specific mark-recapture protocol, but instead are representative data sets over two successive breeding seasons across multiple locations, too separated in time and location to be used for a mark-recapture abundance survey. Thus the results given here for this scenario should be considered illustrative of the method rather than true estimates, and for this reason cannot be compared to overall abundance numbers obtained through other survey methods on this population.

The individual per-call population estimates and accuracies, and the overlap determination using the identity matching technique, were presented previously in Tables II and III. From these tables, we can see that for the *cb* song-type, the number of individuals in the training set was 13 and the number of identity clusters over the test set was 22, with two individuals matched between the data sets. Given these numbers, the likelihood of the population estimate is computed using the mark-recapture formula in Eq. (2). The likelihood

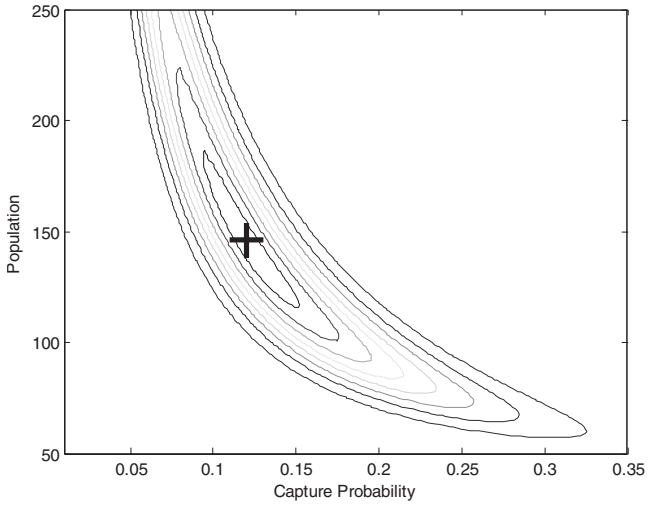


FIG. 9. Likelihood contour plot for song *cb* (ML estimate=146, $p=0.12$).

function, shown in Fig. 9, reaches a maximum value when $p=0.12$ and $N=146$, yielding a total abundance estimate of 146 individuals that use the *cb* song-type. A likelihood confidence interval (95% probability level) is then constructed using the variance of Chapman's modified estimator (Seber, 1982), with a resulting estimate of $\hat{N}=146\pm44$.

Given a properly implemented mark-recapture survey, this approach can be used to give an overall abundance estimate for individuals that make a specific song-type. Repertoire information can then be used to combine these estimates into an overall abundance estimate for the population, as described previously and illustrated in Sec. III B 3.

3. Discussion

The results presented here are intended to illustrate the feasibility of applying advanced methods for vocalization and individual classification to the task of population estimation. It is important to note that the two specific protocols addressed in this work are by no means the only possible approaches to incorporating models of individual vocal variability within an acoustic population assessment design. The concept of "acoustic mark-recapture" can, in fact, be extended to essentially any type of mark-recapture survey, and there are a wide variety of ways in which call-type or identity classification and/or clustering methods can be incorporated into these designs.

For example, within the single-pass protocol of Fig. 3, it was assumed that initial training data were available for the purposes of building the necessary song-type classification models. However, it is possible to implement this step of the process even if substantial training data are not available, using song-type clustering rather than classification as a front-end step. [This has, in fact, been implemented, for details see (Adi, 2008)]. Additionally, within the mark-recapture protocol presented in Fig. 4, it was assumed that individual identity was expertly labeled in the "marking" data set in order to build individual identity models, which negates some of the value of the automated methods, since half of the data still require manual identity analysis. By using identity clustering within both data sets, however, it is

possible to extend the mark-recapture protocol and make it more fully automated. This requires a different mechanism for matching the individuals across the two data sets, which is a focus of continuing work. A third example lies with the way in which repertoire data were used for the purpose of combining abundance estimates within individual song-types into an overall population estimate, using repertoire statistics. While the existence and knowledge of a stable repertoire allow for higher estimation accuracy as well as providing a much more accurate idea of confidence intervals for the estimate, it is not a necessary part of the protocol. In the case of most terrestrial or marine species, the vast majority of which have less sophisticated vocalization repertoires than those of song-birds, it would be sufficient and much simpler to identify one or two basic vocalization types which exist across all or nearly all of the population and to use the resulting population estimates directly.

There are number of key factors that must be considered in generalizing this kind of approach to other species, populations, and habitat regions. Because the technique relies on separation of vocalizations into call-type categories before identity analysis, it is important to have some prior knowledge of the underlying vocalization repertoire of a species as well as the repertoire's consistency across subpopulations and geographic regions. The larger and more complex the repertoire, the more difficult it is to separate call-type differences from individual differences. Species where individuals have relatively large repertoires, with a corresponding fewer number of vocalizations of each type, will thus be more difficult to accurately census. Clearly, in order for acoustic censusing techniques to be feasible, it is necessary that a significant majority of individuals within the species under study vocalize relatively often and loudly enough to allow design of a complete acoustic survey. An equally important but less problematic requirement is that individuals within the species possess sufficient individual vocal variability that they can be separated and identified, which although generally present in a very wide range of species, is certainly not guaranteed, and has not yet been carefully studied outside of terrestrial and marine mammals and songbirds (for example, in insects). This individual distinctiveness must be measurable and methods must be robust to the presence of environmental and microphone channel noise, an area which is also the focus of much research for human speaker identification and verification tasks.

There are a number of parameter settings within the classification and particularly clustering methods that may have impact on the accuracy of results. The most substantial of these currently is the threshold for matching the individuals within the marking data set to those within the recapture data set, because the number of overlapping individuals is the single biggest factor in resulting population estimate.

Overall, though, the promising initial results presented here suggest that there is great potential for the inclusion of individual vocal variability analysis into population assessment designs. This work has significant potential for further extension as well, including, in particular, application to distance sampling by using a point or line transect protocol for recording the vocalizations combined with source localization methods for determining distances to the survey line, which could yield accurate estimates of population density. In general, the underlying approach introduced here creates a means for integrating acoustics and individual vocal variability into many different tools for population assessment.

IV. CONCLUSIONS

The results from this work strongly indicate that the individual distinctiveness of vocalizations can be used to accurately estimate abundance within a data set as well as to match individuals across data sets, and furthermore that this methodology can be incorporated into a larger mark-recapture survey design for overall population assessment. We have illustrated this idea here for the ortolan bunting.

It is likely that nearly all vocally active animals have individually distinct vocalization characteristics, as has already been observed across many different species. The framework for abundance estimation presented in this paper is thus applicable to any vocally active species with a distinct vocal repertoire. This approach addresses the problem of population assessment in a new way, employing algorithms for automatic human speech and speaker recognition to estimate animal abundance. The method has advantages over physical marking techniques, as it is less invasive and is more cost and labor effective. It also has the potential to be more autonomous than current approaches to acoustic or visual surveying, since once basic repertoire models have been built local population can be estimated from any continuous recording. In general, if obtained under the larger umbrella of a well-designed mark-recapture or distance sampling survey protocol, this mechanism may allow for a substantially more accurate understanding of overall population structure and abundance on a larger scale.

ACKNOWLEDGMENT

The authors would like to thank the National Science Foundation (Grant No. IIS-0326395 "The Dr. Dolittle Project") for supporting this work.

Adi, K. (**2008**). "Hidden Markov model based animal acoustic censusing: Learning from speech technology," Ph.D. thesis, Marquette University, Milwaukee, WI.

Adi, K., Sonstrom, K., Scheifele, P., and Johnson, M. T. (**2008**). "Unsupervised validity measures for vocalization clustering," in ICASSP, Las Vegas, NV.

Ajmera, J., and Wooters, C. (**2003**). "A robust speaker clustering algorithm," in Automatic Speech Recognition and Understanding Workshop, St. Thomas, VI.

Barlow, J., and Taylor, B. L. (**1998**). "Preliminary abundance of sperm whales in the northeastern temperate pacific estimate from a combined visual and acoustic survey," International Whaling Commission Working Paper No. SC/50/CAWS20.

Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (**1970**). "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," Ann. Math. Stat. **41**, 164–171.

Beightol, D. R., and Samuel, D. E. (**1973**). "Sonographic analysis of the American woodcock's peent call," J. Wildl. Manage. **37**, 470–475.

Buck, J. R., and Tyack, P. L. (**1993**). "A quantitative measure of similarity for *tursiops truncatus* signature whistles," J. Acoust. Soc. Am. **94**, 2497–2506.

Butynski, T. M., Chapman, C. A., Chapman, L. J., and Weary, D. M. (**1992**). "Use of male blue monkey 'pyow' calls for long-term individual identification," Am. J. Primatol **28**, 183–189.

Cambridge University Engineering Department (**2002**). Hidden Markov Model Toolkit (HTK) Version 3.2.1 User's Guide, Cambridge, MA.
Clemins, P. J. (**2005**). "Automatic speaker identification and classification of animal vocalizations," Ph.D. thesis, Marquette University, Milwaukee, WI.
Clemins, P., and Johnson, M. T. (**2006**). "Generalized perceptual linear prediction (gPLP) features for animal vocalization analysis," J. Acoust. Soc. Am. **120**, 527–534.
Clemins, P. J., and Johnson, M. T. (**2003**). "Application of speech recognition to African elephant (*Loxodonta africana*) vocalizations," in 2003 International Conference on Acoustics, Speech, and Signal Processing, Hong Kong.
Clemins, P. J., Johnson, M. T., Leong, K. M., and Savage, A. (**2005**). "Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations," J. Acoust. Soc. Am. **117**, 956–963.
Cramp, S., and Perrins, C. M. (**1994**). *The Birds of the Western Palearctic* (Oxford University Press, Oxford).
Dale, S. (**2001a**). "Causes of population decline in ortolan bunting in Norway," in Proceedings in the Third International Ortolan Symposium, pp. 33–41.
Dale, S. (**2001b**). "Female-biased dispersal, low female recruitment, unpaired males, and the extinction of small and isolated bird populations," Oikos **92**, 344–356.
Dale, S., and Hagen, O. (**1997**). "Population size, distribution, and habitat choice of the ortolan bunting *Emberiza hortulana* in Norway," Fauna Norv. Ser. C, Cinclus **20**, 93–103.
Edwards, E. P. (**1943**). "Hearing ranges of four species of birds," Auk **60**, 239–241.
Forney, G. D. (**1973**). "The Viterbi algorithm," Proc. IEEE **61**, 268–278.
Galeotti, P., and Pavan, G. (**1991**). "Individual recognition of male tawny owls (*Strix aluco*) using spectrograms of their territorial calls," Ethol. Ecol. Evol. **3**, 113–126.
Gilbert, G., McGregor, P. K., and Tyler, G. (**1994**). "Vocal individuality as a census tool: Practical considerations illustrated by a study of two rare species," J. Field Ornithol. **65**, 335–348.
Holschuh, C. (**2004**). "Monitoring habitat quality and condition of Queen Charlotte saw-whet owls (*Aegolius Acadicus Brooksi*) using vocal individuality," University of Northern British Columbia.
Janik, V. M. (**2000**). "Whistle matching in wild bottlenose dolphins (Tursiops truncatus)," Science **289**, 1355–1357.
Łosak, K. (**2007**). "A comparative analysis of song variation in ortolan bunting (Emberiza hortulana) from populations of different status and quality," Ph.D. thesis, Adam Mickiewicz University, Poznan, Poland.
Marques, T., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales," J. Acoust. Soc. Am. **125**, 1982–1994.
Masters, W. M., Raver, K. A., and Kazial, K. A. (**1995**). "Sonar signals of big brown bats, *Eptesicus fuscus*, contain information about individual identity, age, and family affiliation," Anim. Behav. **50**, 1243–1260.
Mellinger, D. K., and Barlow, J. (**2003**). "Future directions for acoustic marine mammal surveys: Stock assessment and habitat use," NOAA OAR Special Report No. NOAA/PMEL 2557, NOAA, La Jolla, CA.
Osiejuk, T. S., Ratynska, K., Cygan, J. P., and Dale, S. (**2003**). "Song structure and repertoire variation in ortolan bunting (Emberiza hortulana L.) from isolated Norwegian population," Ann. Zool. Fenn. **40**, 3–19.
Otter, K. (**1996**). "Individual variation in the advertisement call of male saw-whet owls," J. Field Ornithol. **67**, 398–405.
Peake, T. M., and McGregor, P. K. (**2001**). "Corncrake Crex crex census estimates: A conservation application of vocal individuality," Animal Biodiversity and Conservation **24**, 81–90.
Reynolds, D. A., and Rose, R. C. (**1995**). "Robust text-independent speaker identification using Gaussian mixture speaker models," IEEE Trans. Speech Audio Process. **3**, 72–83.
Saunders, D. A., and Wooler, R. D. (**1978**). "Consistent individuality of voice in birds as management tool," Emu **88**, 25–32.
Seber, G. A. F. (**1982**). *The Estimation of Animal Abundance and Related Parameters* (MacMillan, New York).
Skierczynski, M., Czarnecka, K. M., and Osiejuk, T. S. (**2007**). "Neighbor-stranger song discrimination in territorial ortolan bunting *Emberiza hortulana* males," J. Avian Biol. **38**, 415–420.
Steifetten, O., and Dale, S. (**2006**). "Viability of endangered populations of ortolan buntings: The effect of a skewed sex ratio," Biol. Conserv. **132**, 88–97.
Steinberg, B. D. (**1983**). *Microwave Imaging with Large Antenna Arrays* (Wiley, New York).
Terry, A. M., and McGregor, P. K. (**2002**). "Census and monitoring based on individually identifiable vocalizations: The role of neural networks," Animal Conservation **5**, 103–111.
Tranter, S. E., and Reynolds, D. A. (**2006**). "An overview of automatic speaker diarization systems," IEEE Trans. Audio, Speech, Lang. Process. **14**, 1557–1565.
Trawicki, M. B., Johnson, M. T., and Osiejuk, T. S. (**2005**). "Automatic song-type classification and speaker identification of Norwegian ortolan bunting (Emberiza hortulana)," in IEEE International Conference on Machine Learning in Signal Processing (MLSP).
Vepsalainen, V., Pakkala, T., Piha, M., and Tiainen, J. (**2005**). "Population crash of the ortolan bunting Emberiza hortulana in agricultural landscapes of Southern Finland," Ann. Zool. Fenn. **42**, 91–107.

Orthogonal-like fractional-octave-band filters

Jérôme Antoni[a)]
Laboratory Roberval CNRS UMR 6253, University of Technology of Compiègne, 60200 Compiègne, France

(Received 4 May 2009; revised 14 November 2009; accepted 17 November 2009)

This paper addresses the design of digital fractional-octave-band filters with energy conservation and perfect reconstruction—i.e., whose outputs to each fractional-octave-band correctly sum up to the original signal and whose partial energies at each output correctly sum up to the overall signal energy—a combination of properties that cannot be met by any current design despite its considerable importance in many applications. A solution is devised based on the introduction of complex basis functions that span the outputs of the fractional-octave bands and whose real and imaginary parts form two individually—but not mutually—orthogonal bases. This imposes a "partition-of-unity" condition on the design of the filter frequency gains such that they exactly sum up to one over the frequency axis. The practical implementation of the proposed solution uses the discrete Fourier transform, and a fast algorithm is implemented using the fast Fourier transform. The proposed filters are well suited to any application involving the post-processing of finite-energy signals. They closely match the international standard templates, except for a small departure at the bandedge frequencies which can be made arbitrarily small.
 [DOI: 10.1121/1.3273888]

PACS number(s): 43.60.Hj, 43.58.Fm, 43.58.Kr [EJS] Pages: 884–895

I. INTRODUCTION

Fractional-octave-band filters are of wide-spread use in acoustics and vibrations since they closely approximate the human's ear logarithmic sensitivity to sounds of different frequency contents. A typical example is 1/3-octaves which are about the level of discrimination of the ear for narrow-band noises. For discriminating between tonal sounds, much finer fractions of octaves are usually required such as 1/12 or 1/24. Since the advent of digital signal processors, the implementation of fractional-octave-band filters of arbitrary bandwidth designator—say, $1/n$—has become an easy task.[1–7] Specifications relating to their design have been compiled in the ANSI S1.11 and IEC 61260 international standards[8,9] which ensure a number of performance requirements relating to the (digital) filter shapes (typically Chebyshev) and their responses to test signals. Surprisingly, none of these specifications addresses the capability of fractional-octave-band filters to conserve energy and to be invertible. Yet, lacking one of these two requirements may lead to undesirable effects in practical applications. On the one hand, the non-conservation of energy means that the sum of the partial energies in all $1/n$-octave bands does not generally equal the overall signal energy: the filter-bank "looses" or "produces" energy, a serious flaw in all applications where the filters are specifically aimed at analyzing how a signal energy decomposes in the frequency domain. On the other hand, the non-invertibility of the filter-bank means that the sum of the filtered signals at the output of each $1/n$-octave band does not reproduce the original input signal, which is unsatisfactory in all applications concerned with the analysis of the signal waveforms in the time domain. To the best of the author's knowledge, the issue of designing (exact) fractional-octave-band filters that conserve energy and are perfectly invertible at the same time has never been addressed before, although fundamental in all applications involving accurate measurements of physical quantities. Most filter designs that comply with the ANSI S1.11 and IEC 61260 standards actually happen to fall short in this prospect, often with large discrepancies as will be shown later in the paper, thus leaving room for significant improvement.

The design of fractional-octave-band filters with energy conservation and perfect reconstruction seems to be a theoretical challenge as evidenced by a rapid literature survey on related topics. As a matter of fact, the particular case of octave-band analysis ($n=1$) falls within the realm of multiresolution analysis (MRA) in which the classical discrete wavelet transform (DWT) offers an infinity of solutions.[10] In the DWT, orthogonality is the key property that guarantees energy conservation and perfect reconstruction. This is a very encouraging result, which—although limited to a particular value of the bandwidth designator, that surely will not be fine enough in many acoustical applications—motivates the research of similar solutions in the general case. Unfortunately, the generalization of MRA to finer-than-octave frequency-resolutions is restricted to rational ratios $1+1/p$, $p \in \mathbb{N}$,[11] which do not match the fractional-octave-ratios $2^{1/n}$ in the general case $n>1$. A number of orthogonal "rational approximations" to fractional-octave-band filters have yet been proposed based on this idea.[12–15] For instance, 1/2-octave-band filters have been approximated by the geometric ratio 3/2 ($2^{1/2}=1.4142 \simeq 1.5$) and 1/3-octave-band filters by 5/4 ($2^{1/3}=1.2599 \simeq 1.25$). More accurate approximations are still possible by cascading several multiresolution analyses with different frequency ratios. Such solutions are usually at the expense of serious departures from the standard filter specifications, as evidenced by the shape of the corresponding frequency gains. Indeed, orthogonality is such as strin-

a)Electronic mail: antoni@utc.fr

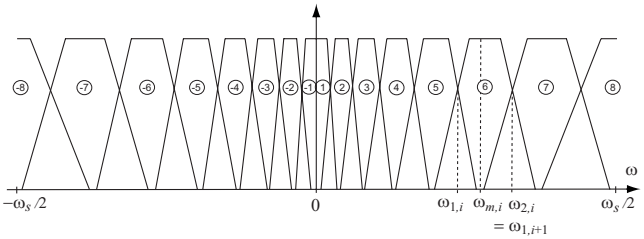


FIG. 1. Representation of a fractional-octave filter-bank.

gent property that it allows little flexibility in designing the shape of the filters. Thus, in applications where meeting the current standards is really of concern, sacrificing the constraint of orthogonality seems to prevail. While this will no longer guarantee the conservation of energy, perfect reconstruction can still be preserved. Design examples of some non-orthogonal yet perfectly invertible filters that are exact or approximate fractional-octave-band filters are found in Refs. 16–19.

Motivated by the solutions that the DWT provides in the octave case, the object of this paper is to propose a general solution in the fractional-octave case. The idea proceeds from the introduction of complex basis functions that span the output subspace of the fractional-octave-filter-bank and for which orthogonality can be replaced by a less restrictive property—which will be referred to as "orthogonal-like"—that still ensures energy conservation and perfect reconstruction. Inspired by Auscher's[20] seminal work, a trial solution is then proposed to build these complex basis functions from a subband decomposition in the discrete Fourier transform (DFT) domain with overlapping frequency gains of arbitrary bandwidth designator that is carefully designed to match the standard templates as close as possible. The proposed solution is fully original in many respects. First because it addresses fractional-octave-band filters for the first time. Second because it introduces a new sufficient condition for energy conservation and perfect reconstruction. Third because a fast algorithm is provided for the analysis and synthesis of signals through the proposed fractional-octave-filter-bank, with similar complexity as the fast Fourier transform (FFT). This makes the algorithm well suited to any application involving the post-processing of recorded data (i.e., finite-energy signals), but not in its present form to the real-time analysis of a continuous stream of data.

II. ORTHOGONAL-LIKE FRACTIONAL-OCTAVE-BAND FILTERS

A. Generalities

As depicted in Fig. 1, a bank of $1/n$-octave-band filters is made of a cascade of constant percentage bandwidth (CPB) bandpass filters that divide octaves into n subbands. Denoting by $\omega_{1,i}$ and $\omega_{2,i}$ the lower and upper bandedge frequencies in radians/s and by $\omega_{m,i}=\sqrt{\omega_{1,i}\omega_{2,i}}$ the exact mid-band frequency of the ith band, the following relation fully characterizes the CPB frequency partition achieved by $1/n$-octave-filter-banks:

$$\omega_{\cdot,i+1} = 2^{1/n}\omega_{\cdot,i}, \tag{1}$$

where $\cdot$ stands indifferently for index 1, 2, or m. In practice, the partition of the frequency axis toward zero will not continue below some user-defined lower frequency—typically the lower bound of the audible range—and it therefore turns out convenient to introduce a "complementary" low-pass filter that closes the low-frequency range, as shown in Fig. 1. With these specifications, it is possible to design $I-1$ digital bandpass filters and one low-pass filter that completely covers the positive frequency axis up to the folding frequency $\omega_s/2$ (i.e., half Shannon's sampling frequency). As mentioned in the Introduction, the ANSI S1.11 and IEC 61260 standards provide many guidelines toward that goal. To keep notations simple, the low-pass filter will be numbered as the first band $(i=1)$ of the filter-bank and then the fractional-octave-bands will be numbered from $i=2$ to $i=I$ with crescent frequencies. The same numbering will be mirrored to the negative frequency axis, with negative indices. The frequency-dependent gain assigned to band i will be denoted as $H_i(\omega)$. The usual Hermitian-symmetry property $H_{-i}(\omega)=H_i(-\omega)^*$ (where * stands for the complex conjugate symbol) will be assumed between paired negative and positive frequency bands so as to guarantee a real impulse response.

In this paper, the input signal to a (digital) fractional-octave-filter-bank, say, $x(t_n)$, $n=-N/2+1,\ldots,N/2$ (N even), will be assumed real-valued, of finite energy $\mathcal{E}_x=\Sigma_n|x(t_n)|^2$, and sampled at time $t_n=2\pi n/\omega_s$. The output of the ith band to $x(t_n)$ will be denoted as $x_i(t_n)$. It is important to realize from the onset that $x_i(t_n)$, $i>0$ (respectively, $i<0$), will have its whole spectrum supported on the positive (respectively, negative) frequency axis only, thus being a complex-valued signal in general. However, due to the realness of the input signal and the Hermitian-symmetry property of the frequency gain $H_i(\omega)$, the outputs from positive and negative frequency bands of indices i and $-i$ will be related as

$$x_{-i}(t_n) = x_i^*(t_n) \tag{2}$$

so that their summation in paired bands will systematically return the real signal:

$$x_i(t_n) + x_{-i}(t_n) = x_i(t_n) + x_i^*(t_n) = 2\Re\{x_i(t_n)\}. \tag{3}$$

B. Filter requirements

With the notations just introduced, one is now in a position to make more precise the requirements of energy conservation and perfect reconstruction.

1. Perfect reconstruction (filter invertibility)

Perfect reconstruction requires that the sum of all outputs to the fractional-octave-filter-bank equals the original signal, namely,

$$x(t_n) = \sum_{\substack{i=-I \\ i\neq 0}}^{I} x_i(t_n) = 2\sum_{i=1}^{I} \Re\{x_i(t_n)\}, \tag{4}$$

where the Hermitian property of real signals—Eq. (2)—was used to establish the equality between the last two terms. Note that in Eq. (4) the dc component is counted half in band

$i=1$ and half in band $i=-1$. The same happens to the folding-frequency component at $\omega=\omega_s/2$ in bands $i=I$ and $i=-I$.

2. Conservation of energy (Plancherel theorem)

The energy carried by a signal is defined as

$$\mathcal{E}_x = \sum_{n=-N/2+1}^{N/2} |x(t_n)|^2. \tag{5}$$

Although not strictly related to the notion of energy in physics, the signal energy $\mathcal{E}_x$ can in general be converted to the latter (or at least to its time-average) if multiplied by a factor in appropriate units of measure. In most applications $1/n$-octave-band filters are used to decompose the signal energy into contiguous CPB bands. This amounts to computing the *partial energies* $\mathcal{E}_{x_i}$ through each band of the filter-bank such that

$$\mathcal{E}_x = \sum_{\substack{i=-I \\ i\neq 0}}^{I} \mathcal{E}_{x_i} = 2\sum_{i=1}^{I} \mathcal{E}_{x_i}, \tag{6}$$

where again the Hermitian property of real signals—Eq. (2)—was used to establish the second equality.

Designing a filter that *simultaneously* fulfills Eqs. (4) and (6) is a difficult task. Section II C proposes a design structure which, by construction, will be perfectly invertible. Conditions will then be found so that it also conserves energy.

C. Design principles

1. Fractional-octave-band basis functions

The proposed approach proceeds from introducing the set of complex basis functions $\{\psi_{ij}(t_n)\}_{j\in\mathbb{Z}}$, $i=\pm 1,\ldots,\pm I$, that spans the vector subspace $\mathcal{S}_i$ at the ith output of the fractional-filter-bank. That is, for any input $x(t_n)$, there exists a set of complex coefficients $\{X_{ij}\}_{j\in\mathbb{Z}}$, $i=\pm 1,\ldots,\pm I$, such that

$$x_i(t_n) = \sum_j X_{ij}\psi_{ij}(t_n). \tag{7}$$

(As before, index i refers to frequency localization, whereas index j will be shown later to refer to time-localization.)

Since $x_i(t_n)$ on the left-hand-side of Eq. (7) is a band-pass signal confined in the fractional-octave band i, the same must hold true for $\psi_{ij}(t_n)$ on the right-hand-side. Specifically, upon taking the Fourier transforms of both sides of Eq. (7),

$$\mathcal{F}\{x_i(t_n)\} = H_i(\omega)X(\omega) = \sum_j X_{ij}\Psi_{ij}(\omega), \tag{8}$$

with $X(\omega)=\mathcal{F}\{x(t_n)\}$ and $\Psi_{ij}(\omega)=\mathcal{F}\{\psi_{ij}(t_n)\}$ (where $\mathcal{F}$ stands for the Fourier transform symbol). This shows that $\Psi_{ij}(\omega)$ has the same frequency support as $H_i(\omega)$, the frequency gain of the ith band. Thus $\psi_{ij}(t_n)$ is an oscillating function of time, hereby justifying its later quotation as a "wavelet."

At this stage, it is readily seen that decomposition (7) is perfectly invertible if

$$x(t_n) = \sum_{\substack{i=-I \\ i\neq 0}}^{I} x_i(t_n) = \sum_{\substack{i=-I \\ i\neq 0}}^{I} X_{ij}\psi_{ij}(t_n) = 2\sum_{i=1}^{I}\sum_j \Re\{X_{ij}\psi_{ij}(t_n)\}. \tag{9}$$

As for energy conservation, let us seek the energy of the ith band in the form $\mathcal{E}_{x_i}=\Sigma_j|X_{ij}|^2$ in anticipation of a Plancherel identity. Then, energy is conserved provided that

$$\mathcal{E}_x = \sum_{\substack{i=-I \\ i\neq 0}}^{I} \mathcal{E}_{x_i} = \sum_{\substack{i=-I \\ i\neq 0}}^{I}\sum_j |X_{ij}|^2, \tag{10}$$

where $\mathcal{E}_x$ is as defined in Eq. (5). Physically, this means that the total energy $\mathcal{E}_x$ of the signal decomposes as a summation of many "energy packets" $|X_{ij}|^2$ well localized in frequency (index i) and, as will be shown shortly, in time (index j).

It now remains to investigate which properties the basis functions $\psi_{ij}(t_n)$ should satisfy to meet conditions (9) and (10).

2. Orthogonal-like condition

It would be an easier matter to verify that having orthogonal complex basis functions $\psi_{ij}(t_n)$ would make conditions (9) and (10) always satisfied. For instance, the complex exponentials[21] of the Fourier series (FS) provide a well-known example of a set of orthogonal complex basis functions with the requested properties. The limitation of the FS, however, would shortly arise due to the division of the frequency axis with a constant infinitesimal frequency resolution that is fundamentally different from the logarithmic frequency resolution of fractional-octave-band filters.

As discussed in the Introduction, orthogonality is also the mechanism used by the DWT—an octave-band analysis—and its various related transforms to meet conditions (9) and (10), yet its generalization to arbitrary $1/n$-octave-filter-bank ($n>1$) is theoretically not feasible. Moreover, the DWT involves *real* basis functions and its extension to orthogonal *complex* basis functions—a property which will be shown later to be of great importance in the present work—is definitively difficult.[22,23]

It turns out that orthogonality is actually not a necessary condition for energy conservation and perfect reconstruction. One important result of this paper is that, in the case of complex basis functions, it suffices that the following condition holds:

$$2\Re\{\langle\psi_{ij},\psi_{pq}\rangle\} = \sum_{n=-N/2+1}^{N/2} 2\Re\{\psi_{ij}(t_n)\psi^*_{pq}(t_n)\} = \delta_{ip}\delta_{jq}, \tag{11}$$

where δ denotes Kronecker's symbol and $\langle\cdot,\cdot\rangle$ the inner product. It is proved in Appendix A that the above condition implies Eqs. (9) and (10) and that, as a consequence, coefficients X_{ij} are simply computed by projecting the input signal onto the basis functions,

$$X_{ij} = \langle x,\psi_{ij}\rangle = X^*_{-i,j}. \tag{12}$$

It is remarkable that, although not strictly orthogonal due to the presence of the real-part operator on the inner product, basis functions that satisfy the *sufficient condition* (11) still enjoy most of the nice properties as if they were actually orthogonal. However, being subjected to less stringent constraints than strict orthogonality, they will considerably simplify the design of fractional-octave-band filters that can meet the requested properties.

3. Partition of the frequency axis

Let us remind the reader that, at this stage, the issue of designing fraction-octave-band filters with energy conservation and perfect reconstruction has been confined to finding complex basis functions $\psi_{ij}(t_n)$ that satisfy the orthogonal-like condition (11) subject to predefined frequency supports as specified by Eq. (8). Elaborating on Auscher's[20] seminal work, a closed-form solution for $\psi_{ij}(t_n)$ is presented here based on its DFT,

$$\Psi_{ij}(\omega_k) = \mathrm{DFT}\{\psi_{ij}(t_n)\} = \sum_{n=-N/2+1}^{N/2} \psi_{ij}(t_n)\exp\left(-J2\pi\frac{nk}{N}\right), \tag{13}$$

where $J^2=-1$ and $\omega_k=k\omega_s/N$ (please note that the J symbol, without a dot, is used for the imaginary number to avoid any confusion with the summation index j).

The idea proceeds from factoring $\Psi_{ij}(\omega_k)$ into the product of a real window-function $0\leq G_i(\omega_k)\leq 1$ that shares the same frequency support as the requested filter frequency-gain $H_i(\omega_k)$ and an orthogonal complex exponential. Specifically, let us introduce

$$k_i = \mathrm{round}\left[N\frac{\omega_{1,i}}{\omega_s}\right], \tag{14}$$

the DFT line that corresponds to the lower bandedge frequency $\omega_{1,i}$ of $H_i(k)$—with round$[\cdot]$ standing for the nearest integer—and

$$N_i = k_{i+1} - k_i, \tag{15}$$

the number of DFT bins in the frequency band $[\omega_{2,i};\omega_{1,i}]$. Then a tentative solution inspired from the theory of lapped transforms is

$$\Psi_{ij}(\omega_k) = (-J)^i\sqrt{\Delta_i}\cdot G_i(\omega_k)\exp\left(J\frac{\pi(k-k_i)j}{N_i}\right) \tag{16}$$

for $j=-N_i+1,\ldots,N_i$ with

$$\Delta_i = \frac{N}{2N_i}. \tag{17}$$

The rationale beyond expression (16) is that the window-function $G_i(\omega_k)$ has most of its energy concentrated within the DFT interval $[k_i,k_i+N_i]$ so that it effectively selects the ith fractional-octave-band as required by Eq. (8). The normalization by $\sqrt{\Delta_i}$ is to force $\psi_{ij}(t_n)$ to have a constant norm, $\|\psi_{ij}\|^2=1/2$, as required by condition (11). Finally the complex exponential term is the mechanism to build a collection of time shifted functions $\psi_{ij}(t_n)$ when incrementing index j. From the theory of lapped transforms, this is a complete representation for all functions on the interval $0\leq\omega\leq\omega_s/2$. With the proposed structure, a simple requirement for the set of functions in Eq. (16) to be orthogonal-like in the sense of Eq. (11) is to impose that

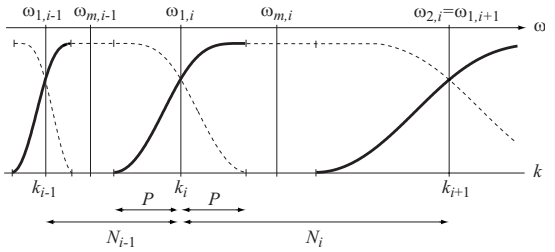


FIG. 2. Construction of the window-functions $G_i(\omega)$ in the DFT domain so as to partition the unity.

(1) the window-functions $G_i(\omega_k)$, $i=1,\ldots,I$, do not overlap one another by more than 50%,
(2) they satisfy the "partition-of-unity" property,

$$\sum_{i=1}^{I}|G_i(\omega_k)|^2 = 1 - \frac{1}{2}\delta_{k,0} - \frac{1}{2}\delta_{k,N/2}, \tag{18}$$

where the one-half fractions at the dc and folding frequencies guarantee that the corresponding components are counted half in bands $i=\pm 1$ $i=\pm I$ in Eq. (4), and thus only once in the paired bands $(+1,-1)$ and $(+I,-I)$.

The proof of this result is reported in Appendix B.

It is noteworthy that there exists an infinity of window-functions $G_i(\omega_k)$ that will meet the above specifications, thus making the design of fractional-octave-band filters with energy conservation and perfect reconstruction much more flexible than expected. One such solution is to build the crescent part of $G_i(\omega_k)$ around the lower bandedge DFT line k_i as

$$G_i(\omega_{k_i}+p) = \sin\left(\frac{\pi}{2}\phi_l(p)\right) \tag{19}$$

for $p=-P,\ldots,P$ with $P<\mathrm{round}[N(\omega_{m,i}-\omega_{1,i})/\omega_s]$, and where $\phi_l(p)$, $l\geq 0$, is obtained from the recursion:

$$\begin{cases}\phi_0(p) = \frac{1}{2}(p/P+1),\\ \phi_l(p) = \sin\left(\frac{\pi}{2}\phi_{l-1}(p)\right), \quad l\geq 1.\end{cases} \tag{20}$$

Then, by building the decreasing part of $G_{i-1}(\omega_k)$ as $G_{i-1}(\omega_{k_i}+p)=\cos(\pi/2\phi_l(p))$, one automatically fulfills Eq. (18) since $\sin^2+\cos^2=1$. Tuning the value of l will directly impact on the transition bandwidths of the fractional-octave-band filter, while the value of P will determine the fraction of overlap (from 0% to a maximum of 50%) between adjacent frequency gains. The principle of the construction is depicted in Fig. 2. Figure 3 shows different design solutions obtained from the proposed algorithm in the case $n=3$.

4. Corresponding complex wavelets

The inverse DFT (IDFT) of $\Psi_{ij}(\omega_k)$ in Eq. (16) is readily found as

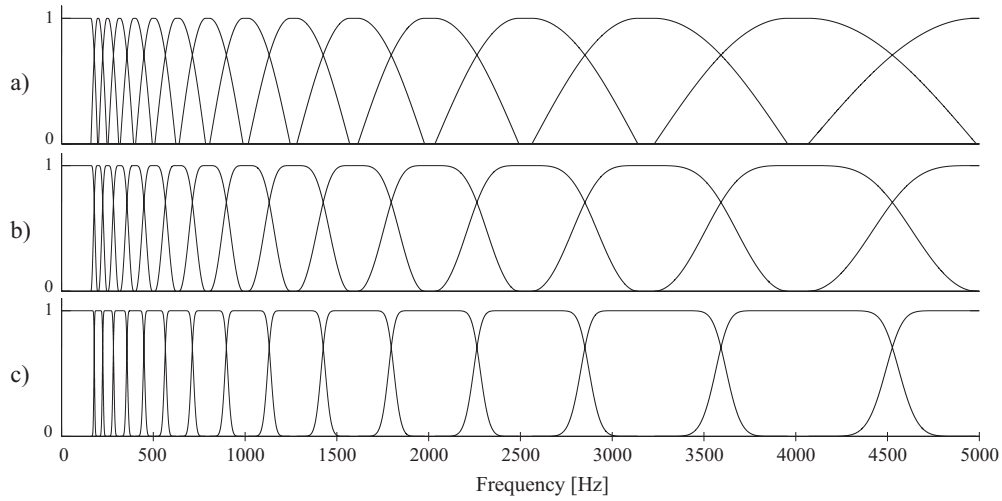


FIG. 3. Examples of 1/3-octave-band filter designs—window-functions $G_i(\omega_k)$—with 50% overlap and (a) $l=0$, (b) $l=1$, and (c) $l=3$. (Note that vertical scales are linear.).

$$\psi_{ij}(t_n) = (-J)^i\sqrt{\Delta_i}\cdot g_i\left(t_n + \frac{j2\pi\Delta_i}{\omega_s}\right)\exp\left(-J\frac{\pi k_i j}{N_i}\right) = \psi^*_{-i,j}(t_n), \tag{21}$$

where $g_i(t_n)=\mathrm{IDFT}\{G_i(\omega_k)\}$ and Δ_i, as defined in Eq. (17), now takes an explicit interpretation as a time-delay inversely proportional to the ith bandwidth. Several remarks arrive at this junction.

a. Time-localization. First, it is seen that for a given fractional-octave-band i, the corresponding basis functions $\psi_{ij}(t_n)$ are shifted versions—by multiples of Δ_i—of the same wavelet. This clearly confers to index j an information on the time-localization of $\psi_{ij}(t_n)$. As well-known from the Fourier transform duality, the smoother the design of the window-function $G_i(\omega_k)$ in the frequency domain, the faster the decay of $g_i(t_n+j2\pi\Delta_i/\omega_s)$ in the time domain, thus potentially making the basis functions $\psi_{ij}(t_n)$ well-localized "wave-packets" in both frequency and time.

By way of an example, Fig. 4 displays the discrete grid formed by the time-frequency positions $j2\pi\Delta_i/\omega_s$ of the wavelets $\psi_{ij}(t_n)$ in the case of an $N=256$ long signal analyzed in third-octave bands $i=1,\ldots,10$ with $k_1=20$. It is emphasized here that if the filters are carefully designed with $k_{I+1}=N/2$, the grid will have in general a total of $2N$ points (including negative frequency bands not shown in Fig. 4) for an N-long signal, a different complex coefficient X_{ij} being assigned to each of them. This double redundancy is the price to pay for the transform to be orthogonal-like instead of exactly orthogonal, see Appendix A.

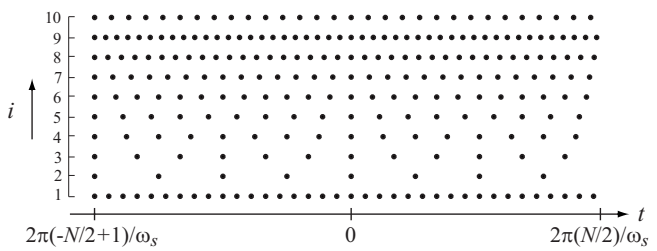


FIG. 4. Grid formed by the time-frequency positions $j2\pi\Delta_i/\omega_s$ in the 1/3-octave-filter-bank analysis of an $N=256$ long signal (indices i of positive bands only are represented).

b. Shape-invariance of $|\psi_{ij}(t_n)|$. Second, it should be noted that the shape of $\psi_{ij}(t_n)$ is not constant with respect to the "scale" index i owing to the presence of the complex exponential. This precludes it from being a wavelet in the strict sense of MRA, where all basis functions are obtained from translated and dilated versions of the same mother wavelet. Nevertheless, the modulus $|\psi_{ij}(t_n)|=\sqrt{\Delta_i}|g_i(t_n+j\Delta_i/\omega_s)|$ is nearly shape-invariant through scales since, with the conventions of this paper, $G_i(\omega_k)\simeq G_2(2^{-(i-2)}\omega_k)$, $i\geq 2$, and therefore $|g_i(\omega_k)|\simeq 2^{(i-2)}|g_2(2^{(i-2)}\omega_k)|$ [where an approximation may result from roundoff errors in Eq. (14)].

c. Interpretation as an envelope. Third, the physical interpretation of $|\psi_{ij}(t_n)|$ is that of the *envelope* of the wave-packet $\psi_{ij}(t_n)$, i.e., that function which smoothly relates all the maxima of the oscillations in $\psi_{ij}(t_n)$. It is noteworthy that for fractional-octave-band filters, the ratio of the filter bandwidth to the exact midband frequency is always less than 1, thus implying that the envelope is a well-defined quantity that can always be factored out of the oscillations it encloses.[24] Moreover, since the basis functions $\psi_{ij}(t_n)$ are complex—actually analytic signals—$|\psi_{ij}(t_n)|$ provides a *unique* definition to the notion of an envelope, a peculiarity which does not happen in the real case.

d. Known particular cases. As a final remark, it is proved in Appendix A that the real and imaginary parts of $\psi_{ij}(t_n)$ form two orthogonal basis, albeit not mutually orthogonal. In the octave-band case ($n=1$), setting $G_i(\omega_k)=1$ when $k_i\leq k\leq k_i+N_i$ and zero otherwise (i.e., the ideal band-pass filter) then makes $\Re\{\psi_{ij}(t_n)\}$ identical to the Littlewood–Paley wavelet.[25] Designing $G_i(\omega_k)$ as a polynomial with a crescent transition bandwidth of 1/3 and a decreasing transition bandwidth of 2/3 makes $\Re\{\psi_{ij}(t_n)\}$ identical to Meyer wavelet.[26] Figure 5 displays three examples of complex basis functions $\psi_{ij}(t_n)$ for $l=0$, 1, 3, and 50% overlap. It is seen from these pictures that the real part of $\psi_{ij}(t_n)$ tends to the Littlewood–Paley wavelet as $l\to\infty$.

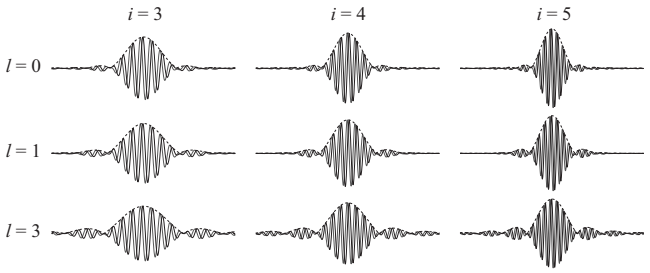


FIG. 5. Examples of 1/3-octave-band "wavelets" for three different filter designs $l=0,1,3$ and for three consecutive bands $i=3,4,5$. The real and imaginary parts, $\Re\{\psi_{ij}(t_n)\}$ and $\Im\{\psi_{ij}(t_n)\}$, are drawn in black and gray, respectively, and the modulus $|\psi_{ij}(t_n)|$ is superposed in dotted line.

D. Compliance with ANSI S1.11 and IEC 61260 standards

1. Corresponding filter frequency gains $H_i(\omega)$

To summarize hitherto the findings of the paper, the proposed fractional-octave-band filter makes use of the *synthesis equation* (7) to compute the filtered output $x_i(t_n)$ to the ith fractional-octave-band, wherein the coefficients X_{ij} are computed from the *analysis equation* (12). A sufficient condition for the fractional-octave-filter-bank to conserve energy and be perfectly invertible is to design the complex basis functions entering into Eq. (7) according to the *design equation* (16) with window-functions satisfying the *partition-of-unity* property (18). Deeper insights into the connection between the complex basis functions $\psi_{ij}(t_n)$ and the filter frequency gains $H_i(\omega)$ may now be gained by noting that Eq. (7) implies that

$$x_i(t_n) = \sum_{j=-N_i+1}^{N_i} X_{ij}\psi_{ij}(t_n) = \mathrm{IDFT}\{X(\omega_k)H_i(\omega_k)\}. \tag{22}$$

Making use of the analysis and design equations to expand X_{ij} and $\psi_{ij}(t_n)$ in the above, the following equality is then arrived at:

$$H_i(\omega_k) = |G_i(\omega_k)|^2, \tag{23}$$

which relates the frequency gain of the fractional-octave-band filter to the squared-magnitude of the window-function of the complex basis functions. In particular, merging Eqs. (18) and (23) results in

$$\sum_{i=1}^{I} H_i(\omega_k) = 1 - \frac{1}{2}\delta_{k,0} - \frac{1}{2}\delta_{k,N/2}. \tag{24}$$

At this stage, it is interesting to note that commercial solutions that comply with ANSI S1.11 and IEC 61260 standards do not satisfy the partition-of-unity property (24), the reason why they cannot conserve energy and be perfectly invertible. By way of an example, Fig. 6(a) displays the frequency gains of a freely available code for 1/3-octave-filter-bank analysis (order-3 Chebyshev filters).[27] Significant departure from unity is observed when summing up all frequency gains $H_i(\omega)$. Figure 6(b) displays the 1/3-octave spectrum of a unit-energy pink noise signal using the same filters (the signal was sampled at $f_s=10$ kHz and further high-pass filtered above $f_c=88$ Hz with a sharp type-I Chebyshev filter in order to maintain a fair comparison with the traditional filter-banks which did not have a complementary low-pass filter; the 1/3-octave-band analysis was then carried out from band $i=2$ at 100 Hz—the first that did not contained f_c—to band $i=21$ at 8000 Hz—the first that contained $f_s/2$). When summing up the partial energies in all 1/3-octave-bands, one recovered an overall energy of 0.9771 instead of 1, that is, a mismatch error of 2.3%. When summing up the signal outputs of all 1/3-octave-bands, one recovered a signal differing by 176% from the input signal according to the normalized mean-square-error measure $\Sigma_n|x(t_n)-\Sigma_i x_i(t_n)|^2/\mathcal{E}_x$. All these defaults were naturally fixed when using the design strategies proposed in the present study.

2. Comparison with standard templates

In turn, a legitimate question is whether the proposed orthogonal-like fractional-octave-band filters are compliant with the ANSI S1.11 and IEC 61260 standards. Figure 7 shows the template for the 1/3-octave-band class-2 filter—which is actually the same for both standards ANSI S1.11 and IEC 61260 and applies as well to other bandwidth des-

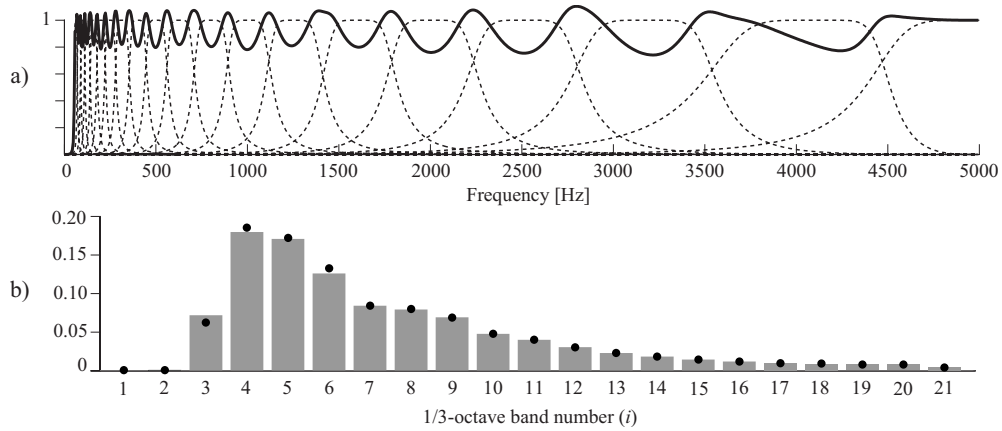


FIG. 6. (a) Demonstration that traditional 1/3-octave-filter-banks do not satisfy the partition-of-unity property [$H_i(\omega_k)$ displayed in gray dotted lines and $\Sigma_{i=1}^{I}H_i(\omega_k)$ in continuous black line, both on a linear vertical scale]. (b) Comparison of 1/3-octave-band spectra of a high-pass filtered pink-noise returned by the filter-bank in (a) (gray bars) and by the proposed algorithm with 50% overlap and $l=2$ (black bullets).

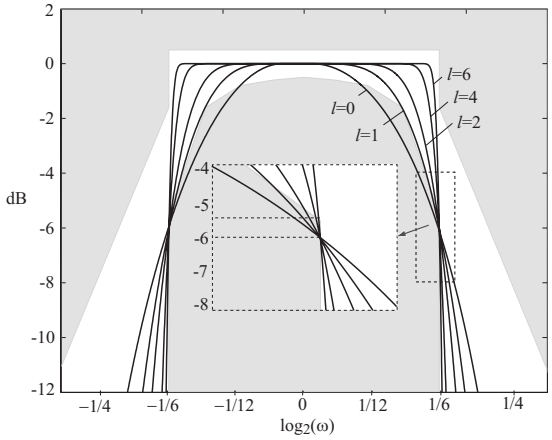


FIG. 7. Frequency gains of the proposed fractional-octave-band filter with 50% overlap and $l=0,1,2,4,6$ against the standard template of ANSI S1.11 and IEC 61260 standards.

ignators after proper scaling of the abscissa axis—together with some frequency gains $20\log_{10}|G_i(\omega)|^2$ designed according to the algorithm provided in Eq. (20). It is seen that for $l>1$, the frequency gains match the template except near the bandedge frequencies $\omega_{1,i}$ and $\omega_{2,i}$ where they are slightly narrower than allowed. This is necessarily so since the partition-of-unity property implies that $|G_{i-1}(\omega_{2,i})|^2=|G_i(\omega_{1,i})|^2=1/2$ and therefore $20\log_{10}|G_i(\omega)|^2=-6$ dB at the bandedge frequencies where the standards require an attenuation of no more than −4.5, −5.0, and −5.5 dB for class 0, 1, and 2 filters, respectively. Increasing the value of l will make this discrepancy arbitrarily small, as evidenced by Fig. 7. This places a tradeoff with time-localization, however, because for large values of l, the filters will tend to ideal bandpass filters, that is, slowly evanescent sinc functions in time. As a matter of fact, $l=2$ or $l=3$ will probably be a reasonable choice in any application where a small departure from the standard templates is acceptable—it is reminded here that this departure is negligible as compared to that induced by orthogonal filters based on rational approximation of fractional-octave-bands.[13–15] Whether fractional-octave-band filters exist—outside the class of solutions proposed in the present paper—that conserve energy, are invertible, and comply exactly with the ANSI S1.11 and IEC 61260 standards remains an open question for future research.

E. Fast algorithm

It is a remarkable fact that the analysis and synthesis of signals with the proposed fractional-octave-band filters can be implemented in a fast and numerically efficient way by making a systematic use of the FFT algorithm. This is a prime reason that makes the results of the present work of practical relevance.

1. Direct transform (analysis step)

Invoking Parseval identity, the analysis equation (12) may be expressed as

$$X_{ij}=J^i\frac{\sqrt{\Delta_i}}{N}\sum_{k=k_i^l}^{k_i^u}X(\omega_k)G_i(\omega_k)\exp\left(-J\frac{\pi(k-k_i)j}{N_i}\right), \tag{25}$$

where k_i^l and k_i^u denote the lower and upper DFT lines, respectively, of the interval over which the frequency gain $G_i(\omega_k)$ is non-zero, i.e., such that $k_i^l\le k_i$ and $k_{i+1}\le k_i^u$. Upon a change in variable and because $k_i^u-k_i^l<2N_i$ when the weighting functions do not overlap by more than 50%,

$$X_{ij}=J^i\frac{\sqrt{\Delta_i}}{N}\exp\left(J\frac{\pi(k_i-k_i^l)j}{N_i}\right)\times\sum_{k=0}^{2N_i-1}G_i(\omega_{k_i^l+k})X(\omega_{k_i^l+k})\exp\left(-J\frac{2\pi kj}{2N_i}\right). \tag{26}$$

The computational demand is in the evaluation of the above summation for all values of j. This can be implemented as two $2N_i$-long DFTs,

$$\mathrm{DFT}_{2N_i}\{G_i(\omega_{k_i^l+k})X(\omega_{k_i^l+k})\}_j \tag{27}$$

for positive indices $j=0,\ldots,N_i$, and

$$\mathrm{DFT}_{2N_i}\{G_i(\omega_{k_i^l+k})X(\omega_{k_i^l+k})^*\}^*_{|j|} \tag{28}$$

for negative indices $j=-N_i+1,\ldots,-1$. Whenever possible, these DFTs will be computed by the FFT algorithm, the efficiency of which will depend on the number of terms in the prime factorization of N_i.

2. Inverse transform (synthesis step)

In the DFT domain, the synthesis equation (7) becomes $X_i(\omega_k)=\Sigma_{j=-N_i+1}^{N_i}X_{ij}\Psi_{ij}(\omega_k)$ or, upon inserting Eq. (16) and changes in variable $k=k'+k_i$ and $j=q-N_i+1$,

$$X_i(\omega_{k'+k_i})=(-J)^i\frac{\sqrt{\Delta_i}}{N}G_i(\omega_{k'+k_i})\exp\left(-J\frac{\pi k'(N_i-1)}{N_i}\right)\times\sum_{q=0}^{2N_i-1}X_{i,q-N_i+1}\exp\left(J\frac{2\pi k'q}{2N_i}\right), \tag{29}$$

where the above summation is directly evaluated from the DFT

$$\mathrm{DFT}_{2N_i}\{X^*_{i,q-N_i+1}\}^*_{k'} \tag{30}$$

at $k'=k_i^l-k_i,\ldots,k_i^u-k_i$ only since $G_i(\omega_{k'+k_i})$ is zero elsewhere. As before, the computation of the DFT (at $k'=0,\ldots,k_i^u-k_i$, wherein $k_i^u-k_i<2N_i$) will be performed by the FFT algorithm whenever possible. Owing to the periodicity of the DFT, a suitable rearrangement of that result will then return the evaluation of the DFT at $k'=k_i^l-k_i,\ldots,-1$ at no additional cost. Finally, the time-domain signal is obtained as $x_i(t_n)=\mathrm{IDFT}\{X_i(\omega_k)\}$ which involves an N-long inverse DFT, with N the length of the input signal $x(t_n)$. Since this is a user-defined parameter, the natural recommendation is to select N a power of 2 (i.e., by zero-padding) so that once again, the FFT algorithm can be used in its most favorable configuration.

F. The concept of instantaneous power

There are numerous applications in acoustics and vibrations where one is not only interested in the global energy $\mathcal{E}_x$ conveyed by a signal but also in how this energy flows as function of time, that is, the *instantaneous power* $\Pi_x(t_n)$. For instance, suppose that $x(t_n)$ is the acoustic pressure of a pure progressing plane-wave; then $\Pi_x(t_n)/(\rho_0 c_0)$, with $\rho_0 c_0$ the specific impedance of the medium, reflects the instantaneous value of energy per unit time and of surface at the measurement point (the so-called "instantaneous intensity"). Physically, the instantaneous power $\Pi_x(t_n)$ must be such that its integration over time returns the overall energy of the signal, that is, for an N-long signal,

$$\sum_{n=-N/2+1}^{N/2} \Pi_x(t_n) = \mathcal{E}_x. \tag{31}$$

Moreover, since energy is typically carried by oscillating (propagating) waves, its flow should display slower variations than the oscillations of the carrier waves. For instance, in the special case where the carrier is a pure tone of amplitude A, one expects the instantaneous power to be a constant function of time with magnitude $A^2/2$. Similarly, in the case the carrier is a narrow-band noise with rms value σ, one expects the instantaneous power to "envelop" the oscillations of the signal, i.e., to vary no faster than the curve that connects the successive maxima (respectively, minima) of the signal, with an average magnitude σ.

The proposed orthogonal-like fractional-octave-band filters are perfectly suited to this task. First, because of their built-in ability to conserve energy and thus instantaneous power. Second, because they decompose the signal into a linear combination of complex wave-packets $\psi_{ij}(t_n)$, each carrying an "atomic" quantity of energy $|X_{ij}|^2$ that smoothly flows up and down in time like function $|\psi_{ij}(t_n)|^2$. Specifically, it is proved in Appendix C that

$$\Pi_x(t_n) = 4\sum_{i=1}^{I} \sum_{j=-N_i+1}^{N_i} |X_{ij}|^2 |\psi_{ij}(t_n)|^2, \tag{32}$$

which shows that instantaneous power is obtained from a linear combination of many elementary energy flows $|X_{ij}|^2|\psi_{ij}(t_n)|^2$, each well localized in time and in frequency. In addition, the expression of $\Pi_x(t_n)$ as given in Eq. (32) implies that

$$\Pi_x(t_n) = 2\sum_{i=1}^{I} \Pi_{x_i}(t_n), \tag{33}$$

where $\Pi_{x_i}(t_n) = 2\Sigma_j |X_{ij}|^2 |\psi_{ij}(t_n)|^2$ is interpreted as the *instantaneous partial power* at time t_n in the fractional-octave-band i. As a final remark, it should be highlighted that $\Pi_{x_i}(t_n)$—and therefore $\Pi_x(t_n)$—can be computed with the fast algorithm described in Sec. II E 2, where X_{ij} and $\psi_{ij}(t_n)$ are simply substituted for $|X_{ij}|^2$ and $|\psi_{ij}(t_n)|^2$.

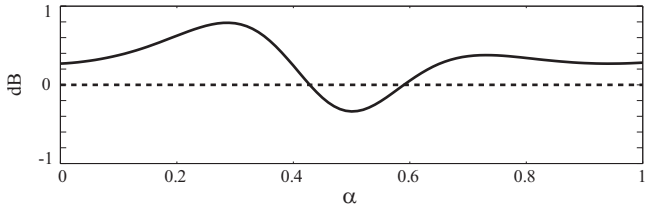


FIG. 8. Deviations from the actual energy of a pure tone analyzed in a third-octave when the tone frequency varies linearly from the lower ($\alpha=0$) to the upper ($\alpha=1$) bandedge frequency (continuous line: a classical filter-bank compliant with ANSI S1.11 and IEC 61260 standards; dotted line: the proposed orthogonal-like filter-bank).

III. EXAMPLES

A. Analysis of a pure tone

As a first pedagogical example, let us investigate how the proposed orthogonal-like fractional-octave-band filters decompose a pure tone, i.e., a signal of the form

$$x(t_n) = A\sin(\omega_0 t_n + \phi), \quad n = -N/2+1, \ldots, N/2, \tag{34}$$

where $\omega_0 = 2\pi k_0/N$ with $0 \leq k_0 \leq N/2$ an integer, and A and ϕ arbitrary amplitude and phase components, respectively. From Eq. (25) it is readily found that, for $i>0$,

$$X_{ij} = \frac{A}{2J}\exp(J\phi)J^i\sqrt{\Delta_i}G_i(\omega_{k_0})\exp\left(-J\frac{\pi(k_0-k_i)j}{N_i}\right), \tag{35}$$

with Δ_i as given in Eq. (17). Therefore, the partial energy in band i and at time index j is simply

$$|X_{ij}|^2 = \frac{A^2}{4}\Delta_i G_i(\omega_{k_0})^2, \tag{36}$$

which turns out independent of the time position j, as expected for a tone. The resulting total energy is, according to Eq. (10), the sum of all partial energies in each band,

$$\mathcal{E}_x = 2\sum_{i=1}^{I}\sum_{j=-N_i+1}^{N_i}|X_{ij}|^2 = \frac{NA^2}{2}\sum_{i=1}^{I}G_i(\omega_{k_0})^2 = \frac{NA^2}{2}, \tag{37}$$

where the partition-of-unity property (18) has been used to arrive at the last equality. Note that this is the correct energy for an N-long tone, i.e., as would have been returned directly by the definition in Eq. (5). It is emphasized here that this result holds true whatever the (integer) value of k_0, and, in particular, even when the latter falls near or exactly on a bandedge frequency, thus intersecting with two window-functions $G_i(\omega)$. This is in contrast with the result returns by classical fractional-octave-filter-banks which, according to ANSI S1.11 and IEC 61260 standards, can allow up to 2 dB difference with the actual energy of a pure tone depending on where its frequency falls between two consecutive fractional octaves. This is illustrated in Fig. 8 for a traditional fractional-octave-band filter (order-3 Chebyshev[27]), where the tone frequency is moved linearly from the lower to the upper bandedge frequency as $\omega_0 = (1-\alpha)*\omega_{1,i} + \alpha*\omega_{2,i}$, and the difference with the actual signal energy is computed as a function of $0 \leq \alpha \leq 1$.

Let us now investigate the instantaneous power of the pure tone. From definition (32) applied to Eq. (36),

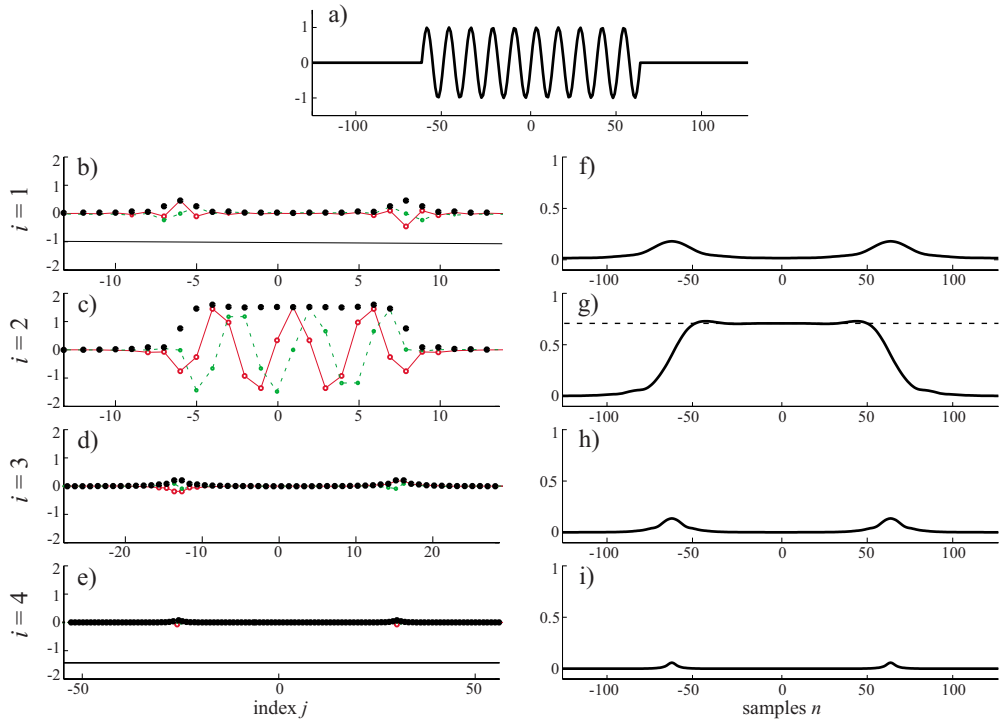


FIG. 9. (Color online) (a) Short-lived tone. [(b)–(e)] Complex coefficients X_{ij} in octave-bands $i=1,2,3,4$ (black bullets: magnitudes; bullets connected by a thin line: real parts; bullets connected by a dotted line: imaginary parts). [(f)–(i)] Square-root of the partial instantaneous powers $2\Pi_{x_i}(t_n)$ in the same octave-bands.

$$\Pi_x(t_n) = \frac{A^2}{2}\sum_{i=1}^{I} \Delta_i |G_i(\omega_{k_0})|^2 \sum_j |\psi_{ij}(t_n)|^2. \tag{38}$$

Now, according to Eq. (21),

$$\begin{aligned}\sum_j |\psi_{ij}(t_n)|^2 &= \Delta_i \sum_j |g_i(t_n + j2\pi\Delta_i/\omega_s)|^2 \\ &= \Delta_i \sum_j \left| \frac{1}{N}\sum_k G_i(\omega_k)\exp\left(J\frac{2\pi k(n+j\Delta_i)}{N}\right)\right|^2 \\ &= \frac{1}{N},\end{aligned} \tag{39}$$

where again the partition-of-unity property (18) has been used to arrive at the last equality after expanding the square. Therefore,

$$\Pi_x(t_n) = \frac{A^2}{2} = \frac{\mathcal{E}_x}{N}. \tag{40}$$

as expected for an N-long pure tone.

B. Analysis of a short-lived tone

Let us now investigate the decomposition of a short-lived tone,

$$x(t_n) = \begin{cases} \sin(\omega_0 t_n), & n \in \{-64, \ldots, 64\} \\ 0, & n \in \{-127, \ldots, -65\} \cup \{65, \ldots, 128\}, \end{cases} \tag{41}$$

as depicted in Fig. 9(a). The analysis is now carried out in octave-bands ($n=1$) with the first bandedge frequency set at $k_1=14$, with 50% overlap and $l=1$. The real and imaginary parts and the magnitude of the coefficients X_{ij} resulting from this decomposition are displayed in Figs. 9(b)–9(e). It is seen that the coefficients have their highest magnitudes in frequency band $i=2$ which contains the tone frequency and at the time indices j where the tone is non-zero. Significant magnitudes are also seen at the onset and release of the tone in all frequency bands owing to the wide-band frequency content of such discontinuities.

This example nicely illustrates how the representation of coefficients X_{ij} can provide time-frequency information. As a matter of fact, this is identical to the classical "scalogram" of wavelet analysis. However, by comparison, any scalogram computed with a real wavelet will return coefficients oscillating with respect to the time index j, actually just as the real and imaginary parts of X_{ij} do in Figs. 9(b)–9(e). The fact that the proposed fractional-octave-band wavelets return complex coefficients whose magnitude $|X_{ij}|$ smoothly envelops the signal—without any oscillation—is quite a unique feature which makes much easier the interpretation of the scalogram.

Another benefit of the proposed approach is linked to the availability of the partial instantaneous powers $\Pi_{x_i}(t_n)$, see Figs. 9(f)–9(i). Even though they contain the same time-frequency information as the $|X_{ij}|$, the partial instantaneous powers enjoy a physical meaning the latter do not have. For instance, the square-root of the instantaneous power $2\Pi_{x_2}(t_n)$—which corresponds to the rms value in the paired bands $i=(-2;+2)$ containing the tone frequency—correctly returns the value $1/\sqrt{(2)}$, as demonstrated in Fig. 9(g), which may be subsequently converted to physical units. In addition,

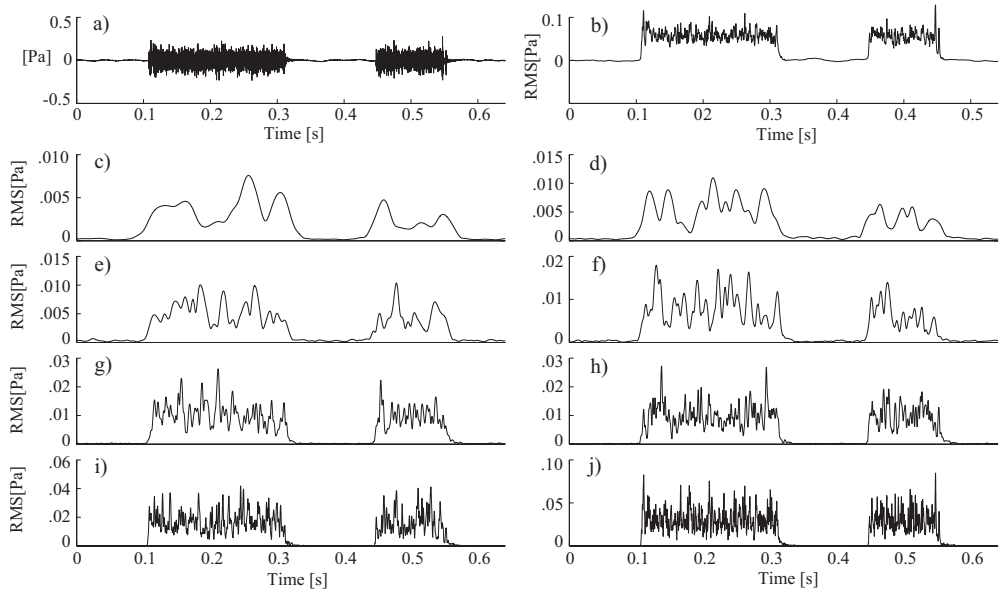


FIG. 10. (a) Pressure signal measured 70 cm from a loudspeaker. (b) Square-root of the overall instantaneous power $\Pi_x(t_n)$. [(c)–(j)] Square-root of the partial instantaneous powers $2\Pi_{x_i}(t_n)$ in 1/3-octave-bands 200, 315, 500, 800, 1250, 2000, 3150, and 5000 Hz.

the partial instantaneous powers all share the same sampling period $1/\omega_s$ as $x(t_n)$, which makes easier their representation as time signals.

C. Analysis of a non-stationary noise

Figure 10 shows the results of the instantaneous power analysis of a pressure signal measured 70 cm from a small loudspeaker driven with alternatively switching on-and-off white noise. The signal ($N=10\,000$) in Fig. 10(a) was sampled at 16 384 Hz and analyzed with a 1/3-octave-filter-bank with 50% overlap and $l=1$. Figures 10(c)–10(j) display the square-roots of the partial instantaneous powers $\Pi_{x_i}(t_n)$ in 1/3-octave-bands 200, 315, 500, 800, 1250, 2000, 3150, and 5000 Hz. Figure 10(b) returns the square-root of the overall instantaneous power $\Pi_x(t_n)$ as obtained from the summation over all partial instantaneous powers in Eq. (33). It is emphasized that qualitatively similar results could surely be obtained from other techniques—e.g., such as from the Hilbert transform—nevertheless those would not conserve energy in general, neither in the sense of Eq. (31) nor of Eq. (33). Moreover, as already discussed in several places of the paper, it is a unique feature of *complex* basis functions to return an envelope, as opposed to other real (orthogonal) basis functions that typically show excessively oscillating results with period equal to twice the midband frequency. In that respect, the proposed definition of instantaneous power returns the smoothest possible result subject to the imposed physical constraints.

IV. CONCLUSION

This paper addressed the issue of designing fractional-octave-band filters which conserve energy and have perfect reconstruction. The proposed solution nearly matches the ANSI S1.11 and IEC 61260 standards at the exception of a small departure at the bandedge frequencies, which can be made arbitrarily small. Clearly, the extension to other auditory filter-banks such as the Mel or the Bark scale is a straightforward matter following the principles of this paper. One question that remains for further research is whether fractional-octave-band filters exist that conserve energy, are invertible, and match exactly the templates of international standards. Another question is whether such filters can be designed with finite impulse responses so that they can be run in real-time to analyze continuous flows of data.

APPENDIX A: PROOF OF THE ORTHOGONAL-LIKE CONDITION (11)

It is first proved that $2\Re\{\langle\psi_{ij},\psi_{pq}\rangle\}=\delta_{ip}\delta_{jq}$ *if and only if* twice the real and imaginary parts of $\psi_{ij}(t_n)$ each form an orthogonal basis.

1. Sufficient condition

Let us expand $\psi_{ij}(t_n)$ as

$$\psi_{ij}(t_n)=\tfrac{1}{2}(\psi^R_{|i|,j}(t_n)+J\,\mathrm{sign}(i)\psi^I_{|i|,j}(t_n)), \tag{A1}$$

$i=\pm1,\ldots,\pm I$, where $\psi^R_{|i|,j}(t_n)$ and $\psi^I_{|i|,j}(t_n)$ are real functions [note that the absolute value on index i forces the Hermitian symmetry $\psi_{ij}(t_n)=\psi^*_{-i,j}(t_n)$]. If $\{\psi^R_{ij}(t_n)\}$ and $\{\psi^I_{ij}(t_n)\}$, $i>0$, each form an orthogonal basis such that $\langle\psi^R_{ij},\psi^R_{pq}\rangle=\langle\psi^I_{ij},\psi^I_{pq}\rangle=\delta_{ip}\delta_{jq}$, then $2\Re\{\langle\psi_{ij},\psi_{pq}\rangle\}=\delta_{ip}\delta_{jq}$ immediately follows.

2. Necessary condition

Conversely, by decomposing $2\psi_{ij}(t_n)$ into real and imaginary parts $\psi^R_{|i|,j}(t_n)$ and $\psi^I_{|i|,j}(t_n)$, it holds in general that $2\Re\{\langle\psi_{ij},\psi_{ij}\rangle\}=\frac{1}{2}\langle\psi^R_{ij},\psi^R_{pq}\rangle+\frac{1}{2}\langle\psi^I_{ij},\psi^I_{pq}\rangle$. Now, because $\psi_{ij}(t_n)$ must vanish for negative frequencies, $\psi^R_{ij}(t_n)$ and $\psi^I_{ij}(t_n)$ are necessarily related by a Hilbert transform, $\psi^I_{ij}(t_n)=\mathcal{H}\{\psi^R_{ij}(t_n)\}$ or, equivalently in the frequency domain, $\Psi^I_{ij}(\omega)=-J\,\mathrm{sign}(\omega)\Psi^R_{ij}(\omega)$. Since the Hilbert transform is a unitary operator, $\langle\psi^R_{ij},\psi^R_{pq}\rangle=\langle\psi^I_{ij},\psi^I_{pq}\rangle$. Therefore, $2\Re\{\langle\psi_{ij},\psi_{pq}\rangle\}$

$=\langle\psi^R_{ij},\psi^R_{pq}\rangle=\langle\psi^I_{ij},\psi^I_{pq}\rangle$. This proves that for having $2\Re\{\langle\psi_{ij},\psi_{pq}\rangle\}=\delta_{ip}\delta_{jq}$, $\{\psi^R_{|i|,j}(t_n)\}$ and $\{\psi^I_{|i|,j}(t_n)\}$ are necessarily two orthogonal bases (note, however, that these two bases are not mutually orthogonal in general, i.e., $\langle\psi^R_{ij},\psi^I_{pq}\rangle \neq \delta_{ip}\delta_{jq}$).[28]

It now remains to prove that the orthogonal-like property (11) is also a sufficient condition to meet the filter requirements (9) and (10). Because $\{\psi^R_{ij}(t_n)\}$ and $\{\psi^I_{ij}(t_n)\}$ are two orthogonal bases, any signal $x(t_n)$ will expand as

$$x(t_n)=\sum_{i=1}^{I}\sum_{j}X^R_{ij}\psi^R_{ij}(t_n)=\sum_{i=1}^{I}\sum_{j}X^I_{ij}\psi^I_{ij}(t_n), \tag{A2}$$

where $X^R=\langle x,\psi^R_{ij}\rangle$ and $X^I=\langle x,\psi^I_{ij}\rangle$. By taking the average of the above two expansions,

$$\begin{aligned}x(t_n)&=\sum_{i=1}^{I}\sum_{j}\frac{1}{2}X^R_{ij}\psi^R_{ij}(t_n)+\frac{1}{2}X^I_{ij}\psi^I_{ij}(t_n)\\&=\frac{1}{2}\sum_{i=1}^{I}\sum_{j}\Re\{(X^R_{ij}-JX^I_{ij})(\psi^R_{ij}(t_n)+J\psi^I_{ij}(t_n))\}\\&=2\sum_{i=1}^{I}\sum_{j}\Re\{X_{ij}\psi_{ij}(t_n)\}=\sum_{\substack{i=-I\\ i\neq 0}}^{I}\underbrace{\sum_{j}X_{ij}\psi_{ij}(t_n)}_{x_i(t_n)}\end{aligned} \tag{A3}$$

with

$$X_{ij}=\tfrac{1}{2}(X^R_{|i|,j}-J\ \mathrm{sign}(i)X^I_{|i|,j})=\langle x,\psi_{ij}\rangle, \tag{A4}$$

and $\psi_{ij}(t_n)$, as defined in Eq. (A1). Hence, by construction, perfect reconstruction is satisfied.

The conservation of energy is now proved. On the one hand, from the first line of Eq. (A3),

$$\begin{aligned}\mathcal{E}_x&=\frac{1}{4}\sum_{i,j,p,q}(X^R_{ij}X^R_{pq}\langle\psi^R_{ij},\psi^R_{pq}\rangle+X^I_{ij}X^I_{pq}\langle\psi^I_{ij},\psi^I_{pq}\rangle\\&\quad+X^R_{ij}X^I_{pq}\langle\psi^R_{ij},\psi^I_{pq}\rangle+X^I_{ij}X^R_{pq}\langle\psi^I_{ij},\psi^R_{pq}\rangle)\\&=\frac{1}{4}\sum_{i,j}(|X^R_{ij}|^2+|X^I_{ij}|^2)+\frac{1}{2}\sum_{i,j,p,q}X^I_{ij}X^R_{pq}\langle\psi^I_{ij},\psi^R_{pq}\rangle.\end{aligned} \tag{A5}$$

On the other hand, from Eq. (A2) and because $\{\psi^R_{i,j}(t_n)\}$ and $\{\psi^I_{i,j}(t_n)\}$ each form an orthogonal basis, $\mathcal{E}_x=\Sigma_{ij}|X^R_{ij}|^2=\Sigma_{ij}|X^I_{ij}|^2$. By taking the average, $\mathcal{E}_x=\frac{1}{2}\Sigma_{ij}|X^R_{ij}|^2+\frac{1}{2}\Sigma_{ij}|X^I_{ij}|^2$, and upon equating with expression (A5), $\frac{1}{2}\Sigma_{ijpq}X^I_{ij}X^R_{pq}\langle\psi^I_{ij},\psi^R_{pq}\rangle=\frac{1}{4}\Sigma_{ij}(|X^R_{ij}|^2+|X^I_{ij}|^2)$. Therefore,

$$\mathcal{E}_x=2\sum_{i=1}^{I}\sum_{j}|X_{ij}|^2=\sum_{\substack{i=-I\\ i\neq 0}}^{I}\sum_{j}|X_{ij}|^2=\sum_{\substack{i=-I\\ i\neq 0}}^{I}\mathcal{E}_{x_i}, \tag{A6}$$

with $\mathcal{E}_{x_i}=\Sigma_j|X_{ij}|^2$.

APPENDIX B: PROOF OF THE PARTITION-OF-UNITY PROPERTY (18)

From Parseval identity, the orthogonal-like property (11) reads $2/N\Sigma_{k=0}^{N/2}\Re\{\Psi_{ij}(\omega_k)\Psi_{pq}(\omega_k)^*\}=\delta_{ip}\delta_{jq}$ in the frequency domain. Substituting for $\Psi_{ij}(\omega_k)$ as given in Eq. (16), this becomes

$$\frac{1}{\sqrt{N_iN_p}}\sum_{k=0}^{N/2}G_i(\omega_k)G_p(\omega_k)\Re\left\{(-J)^{i-p}\exp\left(J\frac{\pi(k-k_i)j}{N_i}\right)\times\exp\left(-J\frac{\pi(k-k_p)q}{N_p}\right)\right\}=\delta_{ip}\delta_{jq}. \tag{B1}$$

1. Auto-orthogonality

Let us first prove Eq. (B1) when the same frequency band $i=p$ is considered, $1<i<I$, yet with arbitrary time-lags j and q. Taking advantage of the fact that $G_i(\omega_k)$ is zero for $k<k_i-N_i$ and $k>k_i+2N_i$, the left-hand-side of Eq. (B1) is split into three summations over intervals $k_i-N_i<k\leq k_i$, $k_i\leq k<k_i+N_i$, and $k_i+N_i<k\leq k_i+2N_i$, that is, after a suitable change in variable,

$$\begin{aligned}&\frac{1}{N_i}\sum_{k=0}^{N/2}G_i(\omega_k)^2\cos\left(\frac{\pi(k-k_i)(j-q)}{N_i}\right)\\&=\frac{1}{N_i}\left\{\sum_{k'=0}^{N_i-1}[G_i(\omega_{k_i-k'})^2+G_i(\omega_{k_i+k'})^2+G_i(\omega_{k_i+2N_i-k'})^2]\times\cos\left(\frac{\pi k'(j-q)}{N_i}\right)-G_i(\omega_{k_i})^2+(-1)^{j-q}G_i(\omega_{k_i+N_i})^2\right\},\end{aligned} \tag{B2}$$

where care has been taken to remove $G_i(\omega_{k_i})^2$ that is counted twice in the summation and to add $(-1)^{j-q}G_i(\omega_{k_i+N_i})^2$ which is not counted in the summation. Taking into account the symmetries $G_i(\omega_{k_i-k'})^2=G_{i-1}(\omega_{k_i+k'})^2$, $G_i(\omega_{k_i+2N_i-k'})^2=G_{i+1}(\omega_{k_i+k'})^2$, $G_i(\omega_{k_i})=\frac{1}{2}$, and $G_i(\omega_{k_i+N_i})=\frac{1}{2}$ implied by the partition-of-unity property—see Fig. 2—the former equation simplifies in

$$\frac{1}{N_i}\sum_{k'=0}^{N_i-1}\left\{\cos\left(\frac{\pi k'(j-q)}{N_i}\right)-\frac{1}{2}+\frac{(-1)^{j-q}}{2}\right\}=\delta_{jq}, \tag{B3}$$

which proves the auto-orthogonality. The case with $i=1$ (respectively, $i=I$) returns the same expression as in Eq. (B2) without the first term (respectively, third) in the summation and with $G_i(\omega_{k_i})^2$ [respectively, $G_i(\omega_{k_i+N_i})^2$] substituted for $\frac{1}{2}$, thus finally reproducing the result of Eq. (B3).

2. Cross-orthogonality

The case where two different and non-adjacent frequency bands are considered—i.e., $|i-p|>1$—trivially returns a zero inner product since $G_i(\omega_k)$ and $G_p(\omega_k)$ then have disjoint support sets—remember that the proposed fractional-octave-bands are constrained to overlap by no more than 50%. When two adjacent frequency bands are considered—e.g., $p=i-1$ without loss of generality—then the summation in the left-hand-side of Eq. (B1) becomes, after a similar change in variable as before,

$$\sum_{k'=-N_i}^{N_i} G_i(\omega_{k_i+k'})G_{i-1}(\omega_{k_i+k'}) \times \Re\left\{-J\exp\left(J\frac{\pi k' j}{N_i}\right)\right.$$
$$\left.\times\exp\left(-J\frac{\pi(k'+N_{i-1})q}{N_{i-1}}\right)\right\}$$
$$= \sum_{k'=-N_i}^{N_i} G_i(\omega_{k_i+k'})G_{i-1}(\omega_{k_i+k'})$$
$$\times(-1)^q \sin\left(\pi k'\left(\frac{j}{N_i}-\frac{q}{N_{i-1}}\right)\right). \quad \text{(B4)}$$

The partition-of-unity property implies that $G_i(\omega_{k_i+k'})\times G_{i-1}(\omega_{k_i+k'})$ is an even function of k'; on the other hand, $\sin(\pi k' x)$ is odd, and thus the above summation is zero.

APPENDIX C: PROOF OF EQ. (32)

Physically, the instantaneous power is embodied by all components in $|x(t_n)|^2$ which do not average to zero in $\Sigma_n|x(t_n)|^2=\Sigma_n\Pi_x(t_n)=\mathcal{E}_x$, as requested by Eq. (31). In order to express the rate of energy as a function of time, a putative solution is to insert the "interpolating functions" $|\psi_{ij}(t_n)|^2$ into Eq. (10), i.e., $\Pi_x(t_n)=4\Sigma_{i>0}\Sigma_j|X_{ij}|^2|\psi_{ij}(t_n)|^2$, which is motivated by the property of these functions of being smooth envelopes, as discussed in Sec. II C 4 c. Now, because the complex basis functions have squared norm $\|\psi_{ij}(t_n)\|^2=1/2$ from Eq. (11) and because $2\Sigma_{i>0}\Sigma_j|X_{ij}|^2=\mathcal{E}_x$ according to Eq. (10), the time-average of $4\Sigma_{i>0}\Sigma_j|X_{ij}|^2|\psi_{ij}(t_n)|^2$ correctly reproduces Eq. (31).

[1]R. Randall and R. Upton, *Digital Filters and FFT Technique in Real-Time Analysis* (Bruel & Kjaer, Denmark, 1981), pp. 45–67.
[2]S. Davis, "Octave and fractional octave band digital filtering based on the proposed ANSI standard," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Tokyo, Japan (1986), pp. 945–948.
[3]J. C. Brow, "Calculation of a constant q spectral transform," J. Acoust. Soc. Am. **88**, 425–434 (1990).
[4]J. W. Waite, "A multirate bank of digital bandpass filters for acoustic applications—Design of HP's 3569A real-time frequency analyzer," Hewlett-Packard J. **44**, 73–81 (1993).
[5]A. Lozano and A. Carlosena, "DSP-based implementation of an ANSI S1.11 acoustic analyzer," IEEE Trans. Instrum. Meas. **52**, 1213–1219 (2003).
[6]K. Yu-Ting, L. Tay-Jyi, L. Yueh-Tai, C. Wei-Han, L. Chih-Wei, and Y. Shuenn-Tsong, "Design of ANSI S1.11 filter bank for digital hearing aids," in Proceedings of the 14th IEEE International Conference on Electronics, Circuits and Systems, Marrakech, Morocco (2007), pp. 242–245.
[7]K. Yu-Ting, L. Tay-Jyi, L. Yueh-Tai, L. Chou-Kun, and L. Chih-Wei, "Low-power ANSI S1.11 filter bank for digital hearing aids," in Proceedings of the International Conference on Signals and Electronic Systems, Krakow, Poland (2008), pp. 347–350.
[8]ANSI, *Specifications for Octave-Band and Fractional-Octave-Band Analog and Digital Filters* (Acoustical Society of America, New York, 1993), reaffirmed 13.
[9]IEC, *Electroacoustics—Octave-Band and Fractional-Octave-Band Filters* (International Electrotechnical Commission, Geneva, 1995).
[10]S. Mallat, *A Wavelet Tour of Signal Processing* (Academic, New York, 1999).
[11]P. Auscher, *Wavelets and Their Applications* (Jones and Barlett, Boston, MA, 1992).
[12]T. Blu, "Iterated filter banks with rational rate changes connection with discrete wavelet transforms," IEEE Trans. Signal Process. **41**, 3232–3244 (1993).
[13]T. Blu, "An iterated rational filter bank for audio coding," in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France (1996), pp. 81–84.
[14]T. Blu, "A new design algorithm for two-band orthonormal rational filter-banks and orthonormal rational wavelets," IEEE Trans. Signal Process. **46**, 1494–1504 (1998).
[15]G. F. Choueiter and J. R. Glass, "An implementation of rational wavelets and filter design for phonetic classification," IEEE Trans. Audio, Speech, Lang. Process. **15**, 939–948 (2007).
[16]C. Heegard and T. Shamoon, "High-fidelity audio compression: Fractional-band wavelets," in IEEE International Conference on Acoustics, Speech, and Signal Processing, Tokyo, Japan (1992), pp. 201–203.
[17]J. Kovacevic and M. Vetterli, "Perfect reconstruction filter banks with rational sampling factors," IEEE Trans. Signal Process. **41**, 2047–2066 (1993).
[18]S. L. Stanford and S. N. Levine, "Critically sampled third octave filter banks," in International Computer Music Conference, Hong-Kong (1996), pp. 201–203.
[19]M. Schonle and N. Fliege, "1/m-octave-band filter banks and wavelets," in IEEE International Symposium on Circuits and Systems, Hong-Kong (1997), pp. 2413–2416.
[20]P. Auscher, *Wavelets: A Tutorial in Theory and Application* (Academic, New York, 1993), pp. 439–452.
[21]$\psi_{ij}(t_n)=\psi_i(t_n)=\exp(-J2\pi in/N)$, $J^2=-1$, which in this particular case happens to not depend on index j.
[22]N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," Appl. Comput. Harmon. Anal. **10**, 234–253 (2001).
[23]I. W. Selesnick, "Hilbert transform pairs of wavelet bases," IEEE Signal Process. Lett. **8**, 170–173 (2001).
[24]E. Bedrosian, "A product theorem for Hilbert transform," Proc. IEEE **51**, 868–869 (1963).
[25]I. Daubechies, *Ten Lectures on Wavelets* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992).
[26]Y. Meyer, *Wavelets and Operators* (Cambridge University Press, Cambridge, 1992).
[27]C. Couvreur, "Implementation of a one-third-octave filter bank in Matlab," http://citeseer.ist.psu.edu/24150.html (Last viewed January 5, 2009).
[28]Since $\{\psi_{ij}(t_n)\}$ is built from the union of two orthogonal bases, $\{\psi^R_{ij}(t_n)\}$ and $\{\psi^I_{ij}(t_n)\}$, it is actually an overcomplete basis that gathers twice as many basis functions as necessary, i.e., a "frame" with bound 2.

Automatic parameter optimization in epsilon-filter for acoustical signal processing utilizing correlation coefficient

Tomomi Abe[a] and Shuji Hashimoto
Department of Applied Physics, Waseda University, 55N-4F-10A, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan

Mitsuharu Matsumoto
The Education and Research Center for Frontier Science, the University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan

(Received 3 August 2009; revised 4 December 2009; accepted 4 December 2009)

ε-filter can reduce most kinds of noise from a single-channel noisy signal while preserving signals that vary drastically such as speech signals. It can reduce not only stationary noise but also nonstationary noise. However, it has some parameters whose values are set empirically. So far, there have been few studies to evaluate the appropriateness of the parameter settings for ε-filter. This paper employs the correlation coefficient of the filter output and the difference between the filter input and output as the evaluation function of the parameter setting. This paper also describes the algorithm to set the optimal parameter value of ε-filter automatically. To evaluate the adequateness of the obtained parameter, the mean absolute error is calculated. The experimental results show that the adequate parameter in ε-filter can be obtained automatically by using the proposed method. [DOI: 10.1121/1.3278601]

PACS number(s): 43.60.Hj, 43.60.Wy, 43.60.Ac [EJS] Pages: 896–901

I. INTRODUCTION

In speech recognition and individual identification, noise reduction plays an important role. There have been many studies about multichannel signal processing for noise reduction in acoustical signal processing such as microphone array,[1–3] independent component analysis,[4,5] and sparseness approaches.[6–8] However, when we consider downsizing the system, single-channel approaches have several advantages compared to multichannel approaches. An approach utilizing a single-channel signal can also be simpler than one employing a multichannel system.

Spectral subtraction (SS) is a well-known approach for reducing the noise signal of the monaural sound.[9–11] It can reduce the noise effectively despite the simplicity of the procedure. However, it can handle only the stationary noise. It also needs noise estimation in advance. Although noise reduction utilizing Kalman filter has also been reported,[12,13] the calculation cost is large. Some authors have reported a model based approach for noise reduction.[14] In this approach, we can extract the objective sound by constructing the sound model in advance. However, it is not applicable to signals with unknown noise as well as SS. There are some approaches utilizing comb filter.[15] In this approach, we first estimate the pitch of the speech signal, and then reduce the noise signal utilizing comb filter. However, the estimation error results in degradation of the speech quality of the objective sound.

As a promising approach, we look to a nonlinear filter labeled ε-filter.[16,17] ε-filter is a nonlinear filter, which can reduce the noise signal while preserving the speech signal. ε-filter is simple and has some desirable features for noise reduction. It does not need a model of the signal nor of the noise in advance. It is easy to design and the calculation cost is small. It can reduce not only stationary noise but also nonstationary noise. However, ε-filter has some parameters whose values are set empirically. Moreover, we only have a single-channel noisy signal, and therefore it is difficult to evaluate whether the parameter is optimal or not. This is because we cannot know the difference between the original signal and the filter output from the observed signal. So far, there have been few studies on the appropriateness of the parameter setting of ε-filter.

As a simple criterion, we employ the correlation coefficient of the filter output and the difference between the input signal and the filter output, assuming that the signal and noise are noncorrelated. We introduce a method to determine the parameter utilizing the correlation coefficient. When we utilize the proposed method, we can set the parameters adequately without information about the noise and the signal in advance and can alter the parameters of ε-filter to the optimal one automatically.

In Sec. II, we explain ε-filter to clarify the problem. In Sec. III, we describe the algorithm of the method to optimize the parameter automatically. In Sec. IV, we show the experimental results. Experimental results show that the proposed method can estimate the adequate parameter of the ε-filter automatically. Discussions and conclusion are given in Sec. V.

[a]Author to whom correspondence should be addressed. Electronic mail: tomomi@shalab.phys.waseda.ac.jp

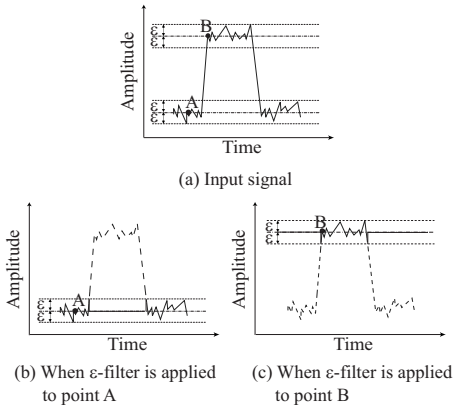


FIG. 1. Basic concept of ε-filter.

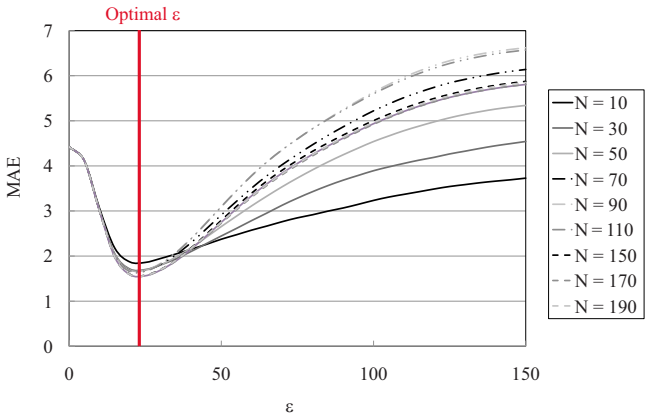


FIG. 2. (Color online) MAE with changing ε and window size N.

II. ε-FILTER

To clarify the problems of an ε-filter, we first explain the ε-filter algorithm. Let us define x_i as the input signal at discrete time i. Let us also define y_i as the output signal of the ε-filter at time i as follows:

$$y_i = x_i + \sum_{k=-N}^{N} a_k F(x_{i+k} - x_i), \tag{1}$$

where a_k represents the filter coefficient. a_k is usually constrained as follows:

$$\sum_{k=-N}^{N} a_k = 1. \tag{2}$$

The window size of the ε-filter is $2N+1$. $F(x)$ is the nonlinear function described as follows:

$$|F(x)| \le \varepsilon : -\infty \le x \le \infty, \tag{3}$$

where ε is a constant. This method can reduce noise while preserving the speech signal that varies drastically. For example, we can set the nonlinear function $F(x)$ as follows:

$$F(x) = \begin{cases} x & (-\varepsilon \le x \le \varepsilon) \\ 0 & (\text{else}). \end{cases} \tag{4}$$

Figure 1 shows the basic concept of an ε-filter when Eq. (4) is utilized as $F(x)$. Figure 1(a) shows the waveform of the input signal. Executing the ε-filter at point A in Fig. 1(a), we first replace all the points where the difference from A is larger than ε. We then summate the signals in the same window. Figure 1(b) shows the basic concept of this procedure. In Fig. 1(b), the broken line represents the points whose difference from point A is larger than ε. In Fig. 1(b), the solid line represents the values replaced through this procedure. As a result, if the difference between point A and the other point is large, the points are ignored. On the other hand, if the difference between point A and the other point is small, the points are smoothed. Due to this procedure, the ε-filter reduces noise while preserving the precipitous attack and decay of the speech signal. In the same way, when we execute the ε-filter at point B, the points are ignored if the difference from point B is larger than ε. The points are smoothed if the difference from point B is smaller than ε. Consequently, we can reduce noise nearby the processed point while preserving the speech signal.

In ε-filter, ε is a crucial parameter to reduce the noise appropriately. If ε is set to an excessively large value, the ε-filter becomes similar to linear filter and smooths not only the noise but also the signal. On the other hand, if ε is set to an excessively small value, it does nothing to reduce the noise anymore. Due to these reasons, the ε value should be set adequately. Note that the window size is another parameter in ε-filter. However, it does not affect the performance of ε-filter compared to the value of ε. We conducted some pre-experiments to show that. Figure 2 shows the experimental result when we changed the window size and the value of ε. As shown in Fig. 2, the optimal ε becomes similar values even when the window size changed. Moreover, even when the window size changes, if ε is set at optimal value, mean absolute error (MAE) does not change so much, as shown in Fig. 2. We, therefore, focus on the parameter setting of ε value in this paper.

III. AUTOMATIC PARAMETER OPTIMIZATION UTILIZING CORRELATION COEFFICIENT

As described in Sec. II, when the ε-filter is employed, we need to set adequate ε values to reduce the noise effectively. However, we cannot estimate the optimal parameter because the noise and signal are not known throughout all the procedures.

To solve the problem, we pay attention to the correlation of the speech signal and the noise signal. We make the following assumption concerning the sound source and noise.

Assumption 1: The speech signal is noncorrelated with the noise signal.

Let us define s_i and n_i as the objective signal and the noise, respectively. Let $R(s_i, n_i)$ be the correlation coefficient of s_i and n_i described as follows:

$$R(s_i, n_i) = \frac{\sum_{i=1}^{L} (s_i - \overline{s_i})(n_i - \overline{n_i})}{\sqrt{\sum_{i=1}^{L} (s_i - \overline{s_i})^2} \sqrt{\sum_{i=1}^{L} (n_i - \overline{n_i})^2}}, \tag{5}$$

where L is the data length. $\overline{s_i}$ and $\overline{n_i}$ represent the statistical means of s_i and n_i, respectively. We estimate $\overline{s_i}$ and $\overline{n_i}$ as time-average and describe them as follows:

$$\overline{s_i} = \frac{1}{L}\sum_{j=1}^{L} s_j, \tag{6}$$

$$\overline{n_i} = \frac{1}{L}\sum_{j=1}^{L} u_j. \tag{7}$$

When L is sufficiently large, Assumption 1 can be expected to satisfy

$$R(s_i,n_i) = 0. \tag{8}$$

As described above, s_i and n_i are unknown throughout the filtering procedures. Instead of s_i and n_i, we consider the correlation coefficient of the filter output and the difference between the input signal and the filter output. Let us consider x_i and y_i as the input signal and the output signal of ε-filter, respectively. x_i can be described as follows:

$$x_i = s_i + n_i. \tag{9}$$

If the ε-filter can wholly reduce the noise, while it preserves the signal completely, the filter output y_i equals the signal s_i. The noise n_i can also be described as follows:

$$n_i = x_i - s_i = x_i - y_i. \tag{10}$$

Although the actual ε-filter does not reduce the whole noise but reduces the signal, if ε value is set optimally, it is expected that the absolute value of the correlation coefficient of y_i and $x_i - y_i$, $|R(y_i, x_i - y_i)|$, has a smaller value than $|R(y_i, x_i - y_i)|$ in other ε. Hence, the optimal parameter $\varepsilon_{\mathrm{opt}}$ can be obtained as

$$\varepsilon_{\mathrm{opt}} = \arg\min_{\varepsilon} |R(y_i, x_i - y_i)|, \tag{11}$$

where

$$R(y_i, x_i - y_i) = \frac{\sum_{i=1}^{L} (y_i - \overline{y_i})(x_i - y_i - \overline{x_i - y_i})}{\sqrt{\sum_{i=1}^{L} (y_i - \overline{y_i})^2}\sqrt{\sum_{i=1}^{L} (x_i - y_i - \overline{x_i - y_i})^2}}, \tag{12}$$

where $\overline{x_i}$ and $\overline{x_i - y_i}$ represent the means of x_i and $x_i - y_i$, respectively. $\overline{x_i}$ and $x_i - y_i$ are described as follows:

$$\overline{x_i} = \frac{1}{L}\sum_{j=1}^{L} x_j, \tag{13}$$

$$\overline{x_i - y_i} = \frac{1}{L}\sum_{j=1}^{L} (x_j - y_j). \tag{14}$$

y_i and x_i are determined by the value of ε. To obtain the optimal ε, we regard the correlation coefficient $R(y_i, x_i - y_i)$ as the function of ε and define $R(\varepsilon)$ as follows:

$$R(\varepsilon) \equiv R(y_i(\varepsilon), x_i(\varepsilon) - y_i(\varepsilon)) = R(y_i, x_i - y_i). \tag{15}$$

Based on the above prospects, the updating of the ε in each iteration is described in the gradient descent method as follows:

$$\Delta\varepsilon = -\eta \frac{\partial |R(\varepsilon)|}{\partial \varepsilon}, \tag{16}$$

where η represents the learning parameter. In practice, the update process is discrete and we cannot obtain the gradient with a $R(\varepsilon)$ value because the shape of the function is unknown. Hence, we update $\varepsilon(T)$ to obtain $\varepsilon_{\mathrm{opt}}$ as follows;

$$\varepsilon(T+1) = \varepsilon(T) + \Delta\varepsilon, \tag{17}$$

where

$$\Delta\varepsilon = -\eta(|R(\varepsilon(T)+1)| - |R(\varepsilon(T))|). \tag{18}$$

We employed a unity as incrementation parameter because $\varepsilon(T)$ is integer. T is the sample number and is used for iteration. In other words, we check two adjacent $R(\varepsilon)$ and update the ε depending on the difference of $R(\varepsilon(T))$ and $R(\varepsilon(T)+1)$. As the initial parameter $\varepsilon(0)$, we use $\sigma(x_i)$, the standard deviation of x_i, as follows:

$$\varepsilon(0) = \sigma(x_i). \tag{19}$$

$\sigma(x_i)$ is described as follows:

$$\sigma(x_i) = \sqrt{\frac{1}{L}\sum_{i=1}^{L} (x_i - \overline{x_i})^2}. \tag{20}$$

This is because $\sigma(x_i)$ represents the fluctuation of n_i when x_i is equal to n_i. It can be considered that most noise will be reduced by the ε-filter under the above situation. In practice, x_i includes s_i and therefore it does not correspond to the correct fluctuation of n_i. However, we consider that the $\sigma(x_i)$ is useful as a first order approximation of ε as we only have the input signal x_i and filter output y_i. The iteration process stops when the following condition is satisfied:

$$||R(\varepsilon(T))| - |R(\varepsilon(T)+1)|| < d, \tag{21}$$

where d is the threshold and a sufficiently small number. We can obtain $\varepsilon_{\mathrm{opt}}$ as the smaller value of $|R(\varepsilon(T))|$ and $|R(\varepsilon(T)+1)|$ in the end.

IV. EXPERIMENT

A. Experimental condition

To clarify the adequateness of the proposed method, we conducted the experiments utilizing monaural sounds with the speech signal and the noise signal. In the experiments, we updated the ε value by using the proposed method and checked whether the proposed method worked well. We also calculated the MAE between the original signal s_i and the filter output y_i to evaluate the proposed method. MAE is defined as follows:

$$\mathrm{MAE} = \frac{1}{L}\sum_{i=1}^{L} |s_i - y_i|. \tag{22}$$

As the sound source, we used "Japanese Newspaper Article Sentences" edited by the Acoustical Society of Japan. We used white noise with uniform distribution as the noise. The signal and the noise were mixed in the computer. The sampling frequency and quantization bit rate were set at 44.1 kHz and 8 bits, respectively. We set the window size of

TABLE I. SNR of each input signal.

	SNR (dB)
Female A	11.0
Female B	11.0
Male A	8.6
Male B	6.6

ε-filter at 21. This paper focuses on a method for adequate setting of ε value which determines the nonlinearity of the filter. Therefore we used the uniform a_k to perform a simple smoothing to avoid the other effects of the weight of a_k.

B. Relation between *T* and ε

We prepared four noisy signals to confirm the robustness of the proposed method with regard to differences among speakers and gender. We used two samples of female speech and two of male speech as the objective sounds. They are labeled "Female A," "Female B," "Male A," and "Male B," respectively. Signal to noise ratios (SNRs) of the signals are shown in Table I. As the $\varepsilon(0)$ value, we used σ value and the values from 10 to 50 with ten intervals to show the robustness of the proposed method concerning the change in $\varepsilon(0)$ and to evaluate the adequateness of σ value as the initial $\varepsilon(0)$. ε value was updated from the $\varepsilon(0)$ and ε_{opt} was obtained as the convergent value by the proposed method. Because of the quantization bit rate, the value of ε is an integer.

Figure 3(a) shows the MAE when we used Female A and applied the ε-filter to it while changing the value of ε. Figure 3(b) shows the relation between T and ε. Figures 4–6 show the experimental results of the other three input signals. These figures also show that the MAE and the relation between T and ε are the same as Fig. 3. As shown in Fig. 3(b), ε converged to the same ε_{opt} concerning all the $\varepsilon(0)$ when Female A was used. We added the line of $\varepsilon_{\text{opt}}(=16)$ calculated by the proposed method to Fig. 3(a) to compare it

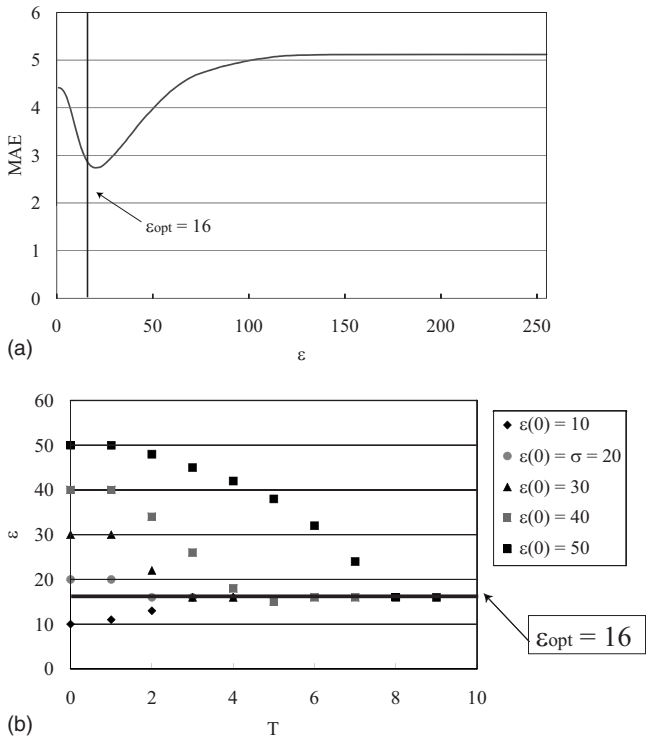


FIG. 3. Female A: (a) MAE and (b) relation between iteration number T and ε.

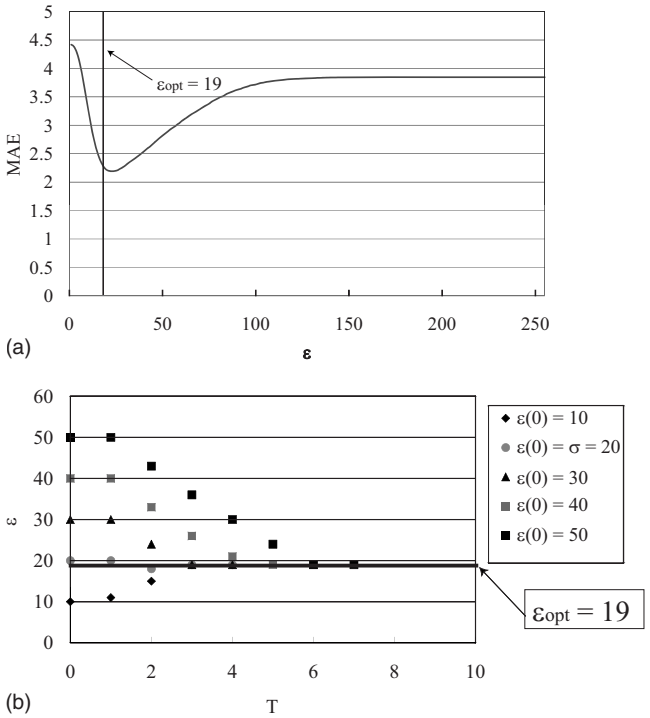


FIG. 4. Female B: (a) MAE and (b) relation between iteration number T and ε.

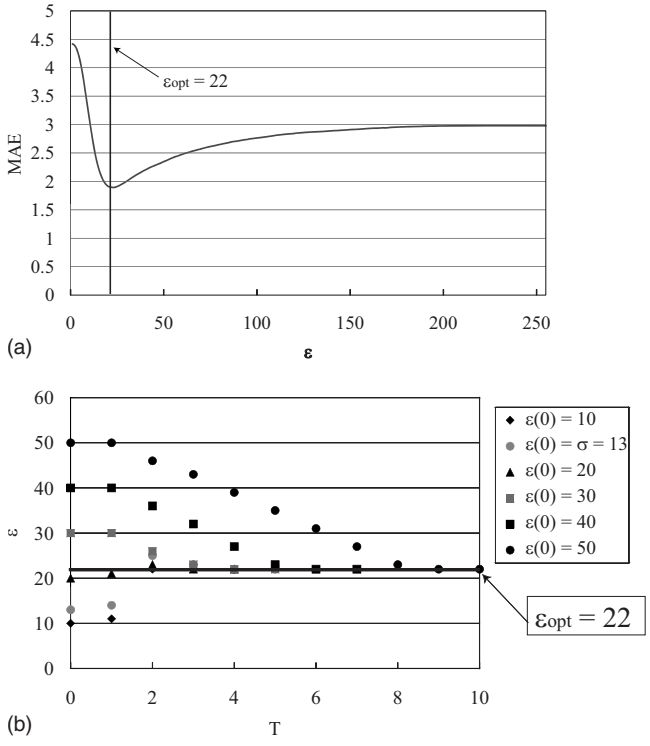


FIG. 5. Male A: (a) MAE and (b) relation between iteration number T and ε.

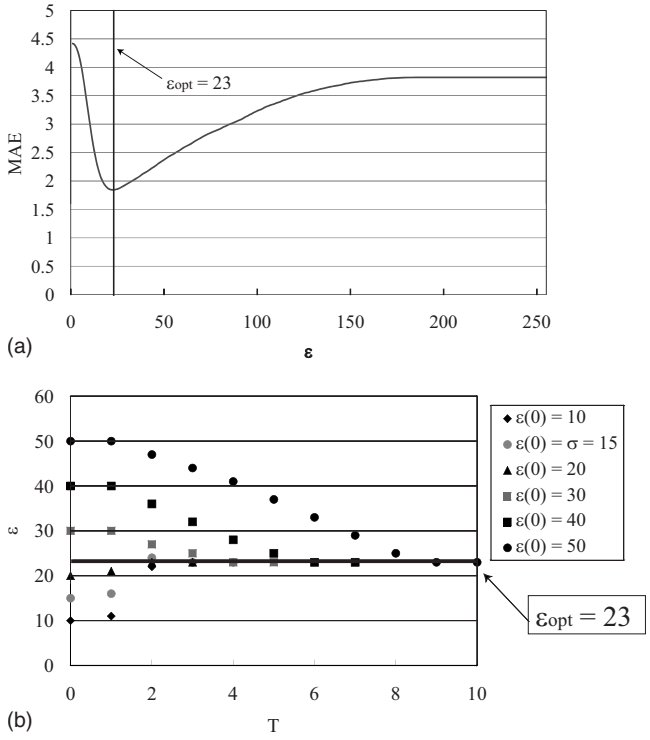


FIG. 6. Male B: (a) MAE and (b) relation between iteration number T and ε.

to the ε that has the minimal MAE. As shown in Fig. 3(a), although ε_{opt} is not equal to the ε that has the minimal MAE, it is sufficiently close to the optimal value. Similar results were also obtained when we used other signals, as shown in Figs. 4(a), 5(a), and 6(a).

It should be noted that the σ value works well as the initial ε. As shown in Figs. 3–6, ε converged to a constant value regardless of the difference in speakers and gender. The difference between σ and ε_{opt} is within 3.5% of the quantization bits in all the experiments. The adequate ε_{opt} was obtained concerning all the cases. These results show positive effects for the adequateness of σ as $\varepsilon(0)$. All the ε_{opt} were sufficiently close to the optimal value of ε with regard to MAE. The proposed method could determine the adequate ε even when the SNR of the input signal was relatively small such as Male A and Male B. As shown in Figs. 3–6, the adequate ε was obtained regardless of the speakers' gender.

C. Experimental results of subjective evaluation

We conducted an experiment using subjective evaluation. We employed a male signal and a female signal used in Sec. IV B as input signals. In each signal, the examinees listened to the five signals: an input signal, a filter output of linear filter, and filter outputs of three types of ε-filters. As three ε values, ε_{opt}, ε_{small}, and ε_{large} are employed. ε_{small} is smaller than ε_{opt} and was set to 10 in the experiment. ε_{large} is larger than ε_{opt} and was set to 50 in the experiment. The examinees rated each signal on a scale of 1–5. Note that score 1 is the worst rating while score 5 is the best auditory impression. Ten examinees participated in this experiment. Table II shows the experimental results of subjective evaluation. As shown in Table II, the signal processed by the proposed method shows better results than any other signals.

TABLE II. Experimental results using subjective evaluation.

	Input	ε_{small}	ε_{large}	Linear	ε_{opt}
Female B	3.3	3.2	2.3	3.1	3.4
Male B	3.1	3.1	3.0	2.5	3.4

V. DISCUSSIONS AND CONCLUSION

In this paper, we employed the correlation coefficient of the filter output and the difference between the input and the filter output as the evaluation function of the parameter setting of ε-filter. We also proposed an algorithm to determine the parameter of ε-filter automatically. The experimental results show that we can automatically determine the adequate parameters of ε-filter by utilizing our method.

As the proposed method only assumes the decorrelation of the signal and noise, it is expected that the application range of the proposed method is large. Although we only have the single-channel noisy signal, our method enables us to obtain an adequate ε value automatically. The proposed method does not require an estimation of the noise in advance. The features will help us to use ε-filter in a practical situation.

To handle nonstationary noise, we need to change ε adaptively depending on the noise. Hence, we aim to improve our method to solve this problem. For future studies, we would like to evaluate robustness when changing the window size of the ε-filter. We also would like to determine all parameters in ε-filter, that is, not only the ε value but also the window size adequately based on automatic control. An improved method should also be developed to handle not only small amplitude noise but also large amplitude noise.

ACKNOWLEDGMENTS

This research was supported by Special Coordination Funds for Promoting Science and Technology, by research grant from Support Center for Advanced Telecommunications Technology Research (SCAT), by research grant from Tateisi Science and Technology Foundation, by research grant from Foundation for the Fusion of Science and Technology, and by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B), Contract No. 20700168, 2008. This research was also supported by the CREST project "Foundation of technology supporting the creation of digital media contents" of JST, by the Grant-in-Aid for the WABOT-HOUSE Project by Gifu Prefecture, and the Global-COE Program, "Global Robot Academia," Waseda University.

[1]K. Sasaki and K. Hirata, "3D-localization of a stationary random acoustic source in near-field by using 3 point-detectors," Transactions of the Society of Instrument and Control Engineers **34**, 1329–1337 (1998).
[2]Y. Yamasaki and T. Itow, "Measurement of spatial information in sound fields by the closely located four point microphone method," J. Acoust. Soc. Jpn. **10**, 101–110 (1990).
[3]M. Matsumoto and S. Hashimoto, "A miniaturized adaptive microphone array under directional constraint utilizing aggregated microphones," J. Acoust. Soc. Am. **119**, 352–359 (2006).
[4]A. J. Bell and T. J. Sejowski, "An information maximization approach to blind separation and blind deconvolution," Neural Comput. **7**, 1129–1159 (1995).

[5]H. Saruwatari, S. Kurita, and K. Takeda, "Blind source separation combining frequency-domain ICA and beamforming," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Process 2001, Utah, Canada (2001), pp. 146–157.
[6]T. Ihara, M. Handa, T. Nagai, and A. Kurematsu, "Multi-channel speech separation and localization by frequency assignment," IEICE Trans. Fundamentals **J86-A**, 998–1009 (2003).
[7]S. Rickard and O. Yilmaz, "On the approximate w-disjoint orthogonality of speech," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Process 2002, Orlando, FL (2002), pp. 529–532.
[8]M. Aoki, Y. Yamaguchi, K. Furuya, and A. Kataoka, "Modifying SAFIA: Separation of the target source close to the microphones and noise sources far from the microphones," IEICE Trans. Fundamentals **J88-A**, 468–479 (2005).
[9]S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust., Speech, Signal Process. **ASSP-27**, 113–120 (1979).
[10]J. S. Lim, *Speech Enhancement* (Prentice-Hall, Englewood Cliffs, NJ, 1983).
[11]J. S. Lim, "Evaluation of a correlation subtraction method for enhancing speech degraded by additive white noise," IEEE Trans. Acoust., Speech, Signal Process. **ASSP-26**, 471–472 (1978).
[12]R. E. Kalman, "A new approach to linear filtering and prediction problems," Trans. of the American Society of Mechanical Engineers **82**, 35–45 (1960).
[13]M. Fujimoto and Y. Ariki, "Speech recognition under noisy environments using speech signal estimation method based on Kalman filter," IEEE Trans. Comput. **J85-D-II**, 1–11 (2002).
[14]P. Daniel, W. Ellis, and R. Weiss, "Model-based monaural source separation using a vector-quantized phase-vocoder representation," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Process 2006 (2006), pp. V-957–V-960.
[15]J. S. Lim, A. V. Oppenheim, and L. D. Braida, "Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition," IEEE Trans. Acoust., Speech, Signal Process. **ASSP-26**, 419–423 (1978).
[16]H. Harashima, K. Odajima, Y. Shishikui, and H. Miyakawa, "ε-separating nonlinear digital filter and its applications," IEICE Trans. Fundamentals **J65-A**, 297–303 (1982).
[17]K. Arakawa, K. Matsuura, H. Watabe, and Y. Arakawa, "A method of noise reduction for speech signals using component separating ε-filters," IEICE Trans. Fundamentals **J85-A**, 1059–1069 (2002).

Monaural sound-source-direction estimation using the acoustic transfer function of a parabolic reflection board

Ryoichi Takashima[a)]
Graduate School of Engineering, Kobe University, Kobe, 657-8501, Japan

Tetsuya Takiguchi and Yasuo Ariki
Organization of Advanced Science and Technology, Kobe University, Kobe, 657-8501, Japan

(Received 6 June 2009; revised 4 December 2009; accepted 4 December 2009)

This paper presents a sound-source-direction estimation method using only a single microphone with a parabolic reflection board. A simple signal-power-based method using a parabolic antenna has been proposed in the radar field. But the signal-power-based method is not effective for finding the direction of a talking person due to the varying power of the uttered speech signals. In this paper, the sound-source-direction estimation method focuses on the acoustic transfer function instead of the signal power. The use of the parabolic reflection board leads to a difference in the acoustic transfer functions of the target direction and the non-target directions, where the parabolic reflector and its associated microphone rotate together and observe the speech at each angle. The acoustic transfer function is estimated from the observed speech using the statistics of clean speech signals. Its effectiveness has been confirmed by monaural sound-source-direction estimation experiments in a room environment. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3278603]

PACS number(s): 43.60.Jn [WMC] Pages: 902–908

I. INTRODUCTION

Many systems using microphone arrays have been tried in order to localize sound sources. Conventional techniques, such as MUSIC (MUltiple SIgnal Classification), CSP (Cross-power Spectrum Phase), and so on (e.g., Refs. 1 and 2), use simultaneous phase information from microphone arrays to estimate the direction of the arriving signal. Also, sound source localization techniques focusing on the auditory system have been described in Refs. 3 and 4.

Single-microphone source separation is one of the most challenging scenarios in the field of signal processing, and some techniques have been described (e.g. Refs. 5–8). In our previous work,[9] we discussed a sound source localization method using only a single microphone. In that report, the acoustic transfer function was estimated from observed (reverberant) speech using the statistics of clean speech signals without using texts of the user's utterance, where a Gaussian mixture model (GMM) was used to model the features of the clean speech. This estimation is performed in the cepstral domain employing a maximum-likelihood-based approach. This is possible because the cepstral parameters are an effective representation for retaining useful clean speech information. The experiment results of our talker-localization showed its effectiveness. However, the previous method required the measurement of speech for each room environment in advance. Therefore, this paper presents a new method that uses parabolic reflection that is able to estimate the sound source direction without any need for such prior measurements.

In this paper, we introduce the concept of an active microphone that achieves a good combination of active-operation and signal processing. The active microphone has a parabolic reflection board, which is extremely simple in construction. The reflector and its associated microphone rotate together, perform signal processing, and seek to locate the direction of the sound source.

A simple signal-power-based method using a parabolic antenna has been proposed in the radar field. But the signal-power-based method is not effective for finding the direction of a person talking in a room environment. One of the reasons is that the power of the speech signal varies for all directions of the parabolic antenna since a person does not utter the same power (word) for all directions of the parabolic antenna. Therefore, in this paper, our new sound-source-direction estimation method focuses on the acoustic transfer function instead of the signal power. The use of the parabolic reflection board results in a difference in the acoustic transfer functions of the target direction and the non-target directions, where the active microphone with the parabolic reflection board rotates and observes the speech at each angle. The sound source direction is detected by comparing the acoustic transfer functions observed at each angle, which are estimated from the observed speech using the statistics of clean speech signals. Its effectiveness is confirmed by sound-source-direction estimation experiments in a room environment.

II. ACTIVE MICROPHONE

A. Parabolic reflection board

In this paper, an active microphone with a parabolic reflection board is introduced for estimation of sound source direction, where the reflection board has the shape of a para-

a)Author to whom correspondence should be addressed. Electronic mail: takashima@me.cs.scitec.kobe-u.ac.jp

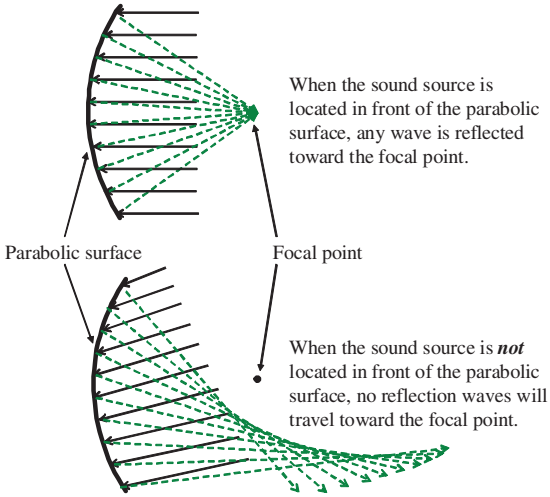


FIG. 1. Concept of parabolic reflection.

bolic surface. The parabolic reflector has been used for estimation of the direction of arrival in the radar field.[10] As shown in Fig. 1, under the assumptions associated with plane waves, any line (wave) parallel to the axis of the parabolic surface is reflected toward the focal point. On the other hand, if the sound source is not located at 90 deg (in front of the parabolic surface), no reflection wave will travel toward the focal point. Therefore, the use of the parabolic reflection board will be able to give us the difference in the acoustic transfer function between the target direction and the nontarget directions.

B. Signal observed using parabolic reflection

Next, we consider the signal observed using parabolic reflection.[11] In Ref. 11, a simple signal-power-based method using a parabolic reflection board has been described. Its effectiveness has been confirmed on white noise signals, but the signal-power-based method was not effective for finding the direction of a talking person due to the varying power of the uttered speech signals.

As shown in Fig. 2, when the sound source is located directly in front of the parabolic surface and there is no background noise, the observed signal at the focal point at time t can be expressed by the addition of the waves arriving at the focal point directly (direct wave) and those arriving at the focal point after being reflected by the parabolic surface (reflection waves):

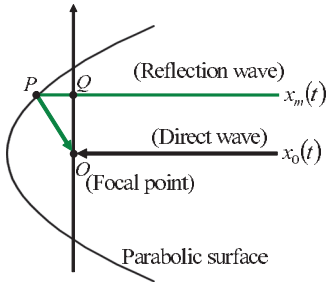


FIG. 2. Observed signal at the focal point, where the input signal comes from directly in front of the parabolic surface.

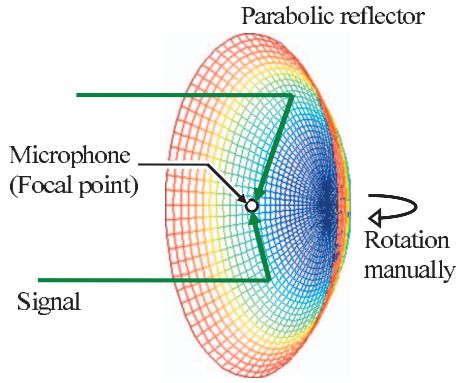


FIG. 3. Active microphone with parabolic reflection.

$$o(t) = x_p(t) + \sum_{m=1}^{M} x_m(t), \tag{1}$$

where $o(t)$, x_p, and x_m $(m=1,\ldots,M)$ are observed sound, direct sound, and reflection sound, respectively. Based on the properties of a parabola, the time difference to the focal point between the direct and reflection waves is constant without depending on m. Therefore, Eq. (1) can be written as

$$o(t) = s(t) * h_p(t) + \sum_{m=1}^{M} s(t-\tau) * h_m(t), \tag{2}$$

where $s(t)$ and τ are clean speech and the time difference, respectively. h_p is the acoustic transfer function of a direct wave and h_m is that of a reflection wave. By applying the short-term Fourier transform, the observed spectrum at frame n is given by

$$O(\omega;n) \approx S(\omega;n) \cdot \left(H_p(\omega;n) + e^{-j2\pi\omega\tau} \cdot \sum_{m=1}^{M} H_m(\omega;n) \right)$$
$$= S(\omega;n) \cdot (H_p(\omega;n) + H_r(\omega;n)). \tag{3}$$

Here H_p is the acoustic transfer function of the direct sound that is not influenced by parabolic reflection. H_r is the acoustic transfer function resulting from parabolic reflection.

On the other hand, when the sound source is not located in front of the parabolic surface, parabolic reflection does not influence the acoustic transfer function since no reflection waves will travel toward the focal point:

$$O(\omega;n) \approx S(\omega;n) \cdot H_p(\omega;n). \tag{4}$$

C. Estimation of sound source direction

As shown in Fig. 3, a new active microphone with a parabolic reflection board was constructed with the microphone located at the focal point. In order to obtain the signal observed at each angle, the angle of the microphone was changed manually in research carried out for this paper. Then, from Eqs. (3) and (4), the spectrum of the signal observed at a microphone angle θ can be expressed as

$$O_\theta(\omega;n) \approx S_\theta(\omega;n) \cdot H_\theta(\omega;n)$$

$$H_\theta(\omega;n) = \begin{cases} H_p(\omega;n) + H_r(\omega;n) & (\theta = \hat{\theta}) \\ H_p(\omega;n) & (\theta \neq \hat{\theta}), \end{cases} \quad (5)$$

where S_θ and H_θ are spectra of clean speech and acoustic transfer function at the angle θ, and $\hat{\theta}$ is the sound source direction. Assuming that H_p is nearly constant for each angle, when the active microphone does not face the sound source, the value of H_θ will be almost the same for every non-target direction. On the other hand, the only condition under which H_θ will have a different value from that obtained at the other angles is when the active microphone faces the sound source.

Therefore, the acoustic transfer function is estimated at each discrete direction θ, and the sound source direction can be estimated by selecting the direction whose acoustic transfer function is the farthest from the acoustic transfer functions of other directions:

$$\hat{\theta} = \arg\max_{\theta} \sum_{\theta'} (\bar{H}_\theta - \bar{H}_{\theta'})^2, \quad (6)$$

where θ' is all directions of the microphone except θ, and $\bar{H}$ is the expectation of H in regard to the time frame. Actually, in this research, the cepstrum of acoustic transfer function is used to calculate this equation. In Sec. III, we will describe how to estimate H_θ from observed speech signals.

III. ESTIMATION OF THE ACOUSTIC TRANSFER FUNCTION

In our previous work,[9] we proposed a method to estimate the acoustic transfer function from the reverberant speech (any utterance) using the clean speech acoustic model, where a GMM is used to model the feature of the clean speech. The clean speech GMM enables us to estimate the acoustic transfer function from the observed speech without needing to have texts of the user's utterance (text-independent estimation). However, because an active microphone with parabolic reflection board was not used, the previous method required the measurement of speech for each room environment in advance in order to be able to determine the direction of a talking person. In this paper, we can estimate the sound source direction without any need for prior measurements by information fusion of an active microphone and an estimation of an acoustic transfer function.

A. Cepstrum representation of reverberant speech

The observed signal (reverberant speech), $o(t)$, in a room environment is generally considered as the convolution of clean speech and acoustic transfer function. The spectral analysis of the acoustic modeling is generally carried out using short-term windowing. Therefore, the observed spectrum is approximately represented by $O(\omega;n) \approx S(\omega;n) \cdot H(\omega;n)$, where the length of the acoustic transfer function may be greater than that of the window. Here $O(\omega;n)$, $S(\omega;n)$, and $H(\omega;n)$ are the short-term linear spectra in the analysis window n.

Cepstral parameters are an effective representation for retaining useful speech information in speech recognition. Therefore, we use the cepstrum for acoustic modeling necessary to estimate the acoustic transfer function. The cepstrum of the observed signal is given by the inverse Fourier transform of the log spectrum:

$$O_{\text{cep}}(d;n) \approx S_{\text{cep}}(d;n) + H_{\text{cep}}(d;n), \quad (7)$$

where d is the cepstral index. O_{cep}, S_{cep}, and H_{cep} are cepstra for the observed signal, clean speech signal, and acoustic transfer function, respectively. As shown in Eq. (7), if O and S are observed, H can be obtained by

$$H_{\text{cep}}(d;n) \approx O_{\text{cep}}(d;n) - S_{\text{cep}}(d;n). \quad (8)$$

However, S cannot be observed actually. Therefore H is estimated by maximizing the likelihood (ML) of observed speech using clean speech GMM.

B. Maximum-likelihood-based parameter estimation

The sequence of the acoustic transfer function in Eq. (8) is estimated in a ML manner[12] by using the expectation maximization (EM) algorithm, which maximizes the likelihood of the observed speech:

$$\hat{H} = \arg\max_{H} \Pr(O|H,\lambda_S). \quad (9)$$

Here, λ denotes the set of GMM parameters of the clean speech, while the suffix S represents the clean speech in the cepstral domain. The GMM of clean speech consists of a mixture of Gaussian distributions.

$$\lambda_S = \{w_k, N(\mu_k^{(S)}, \sigma_k^{(S)^2})\}, \quad \sum_k w_k = 1, \quad (10)$$

where w_k, μ_k, and σ_k^2 are the weight coefficient, mean vector, and variance vector (diagonal covariance matrix) of the kth mixture component, respectively. These parameters are estimated using the EM algorithm using a clean speech database.

The estimation of the acoustic transfer function in each frame is performed in a maximum likelihood fashion by using the EM algorithm. The EM algorithm is a two-step iterative procedure. In the first step, called the expectation step, the following auxiliary function Q is computed:

$$\begin{aligned} Q(\hat{H}|H) &= E[\log \Pr(O,c|\hat{H},\lambda_S)|H,\lambda_S] \\ &= \sum_c \frac{\Pr(O,c|H,\lambda_S)}{\Pr(O|H,\lambda_S)} \cdot \log \Pr(O,c|\hat{H},\lambda_S). \end{aligned} \quad (11)$$

Here c represents the unobserved mixture component labels corresponding to the observation sequence O.

The joint probability of observing sequences O and c can be calculated as

$$\Pr(O,c|\hat{H},\lambda_S) = \prod_{n^{(v)}} w_{c_{n^{(v)}}} \Pr(O_{n^{(v)}}|\hat{H},\lambda_S), \quad (12)$$

where w is the mixture weight and $O_{n^{(v)}}$ is the cepstrum at the nth frame for the vth training data (observation data). Since we consider the acoustic transfer function as additive noise in the cepstral domain, the mean to mixture k in the

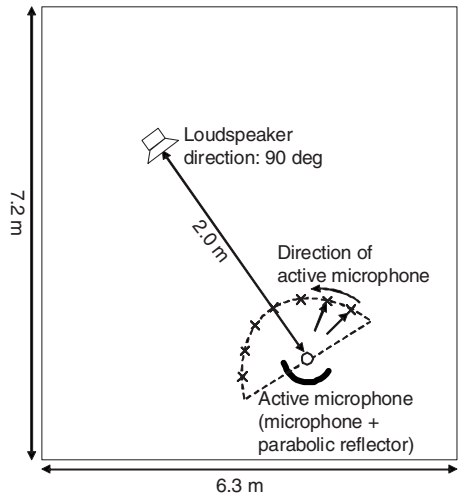


FIG. 4. Experimental conditions.

model λ_O is derived by adding the acoustic transfer function. Therefore, Eq. (12) can be written as

$$\Pr(O,c|\hat{H},\lambda_S)=\prod_{n^{(v)}} w_{c_{n^{(v)}}}\cdot N(O_{n^{(v)}};\mu^{(S)}_{k_{n^{(v)}}}+\hat{H}_{n^{(v)}},\Sigma^{(S)}_{k_{n^{(v)}}}), \tag{13}$$

where $N(O;\mu,\Sigma)$ denotes the multivariate Gaussian distribution. It is straightforward to derive that

$$\begin{aligned}Q(\hat{H}|H)=&\sum_k\sum_{n^{(v)}}\Pr(O_{n^{(v)}},c_{n^{(v)}}=k|\lambda_S)\log w_k\\&+\sum_k\sum_{n^{(v)}}\Pr(O_{n^{(v)}},c_{n^{(v)}}=k|\lambda_S)\log N(O_{n^{(v)}};\mu^{(S)}_k\\&+\hat{H}_{n^{(v)}},\Sigma^{(S)}_k).\end{aligned} \tag{14}$$

Here $\mu^{(S)}_k$ and $\Sigma^{(S)}_k$ are the kth mean vector and the (diagonal) covariance matrix in the clean speech GMM, respectively. It is possible to train those parameters by using a clean speech database. Next, we focus only on the term involving H.

$$\begin{aligned}Q(\hat{H}|H)=&-\sum_k\sum_n\gamma_k(n)\sum_{d=1}^{D}\left\{\frac{1}{2}\log(2\pi)^D\sigma^{(S)^2}_{k,d}\right.\\&\left.+\frac{(O(d;n)-\mu^{(S)}_{k,d}-\hat{H}(d;n))^2}{2\sigma^{(S)^2}_{k,d}}\right\},\end{aligned} \tag{15}$$

$$\gamma_k(n)=\Pr(O(n),k|\lambda_S). \tag{16}$$

Here $O(n)$ is the cepstrum at the nth frame for observed speech data. D is the dimension of the $O(n)$, and $\mu^{(S)}_{k,d}$ and $\sigma^{(S)^2}_{k,d}$ are the dth mean value and the dth diagonal variance value of the kth component in the clean speech GMM, respectively.

The maximization step (M-step) in the EM algorithm becomes max $Q(\hat{H}|H)$. The re-estimation formula can, therefore, be derived, knowing that $\partial Q(\hat{H}|H)/\partial\hat{H}=0$ as

$$\hat{H}(d;n)=\frac{\sum_k\gamma_k(n)\dfrac{O(d;n)-\mu^{(S)}_{k,d}}{\sigma^{(S)^2}_{k,d}}}{\sum_k\dfrac{\gamma_k(n)}{\sigma^{(S)^2}_{k,d}}}. \tag{17}$$

Therefore, the frame sequence of the acoustic transfer function $\hat{H}_\theta(d;n)$ is estimated from the signal $O_\theta(d;n)$ observed at direction θ in the cepstral domain using Eq. (17).

We obtain $\hat{H}_\theta(d;n)$ at a discrete direction. Next, the dth dimension of the mean vector $\bar{H}_\theta(d)$ is obtained by averaging $\hat{H}_\theta(d;n)$ per frame n.

$$\bar{H}_\theta(d)=\sum_n\hat{H}_\theta(d;n). \tag{18}$$

In a similar way, we obtain the mean vector $\bar{H}_\theta(d)$ at all discrete directions, and the sound source direction is estimated using Eq. (6) using the cepstral vector $\bar{H}_\theta$. In this paper, the angle of the parabolic reflection microphone was changed manually from 30 to 150 deg in increments of 20 deg.

IV. EXPERIMENT

A. Experiment conditions

The direction estimation experiment was carried out in a real room environment. The parabolic reflection microphone shown in Fig. 3 was used for the experiments. The diameter was 24 cm, and the distance to the focal point was 9 cm. The microphone located at the focal point was an omnidirectional type (SONY ECM-77B). The target sound source was located at 90 deg and 2 m from the microphone. The angle of the parabolic reflection microphone was changed manually from 30 to 150 deg in increments of 20 deg. Then the acoustic transfer function of the target signal at each angle was estimated for the following speech lengths: 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 s. The size of the recording room was about 6.3 m × 7.2 m (width × depth). Figure 4 shows the environment of the experiment.

The speech signal was sampled at 12 kHz and windowed with a 32-ms Hamming window every 8 ms. The clean speech GMM was trained by using 50 sentences (spoken by a female) in the ASJ Japanese speech database. The trained GMM has 64 Gaussian mixture components. Then second-order mel-frequency cepstral coefficients were used as feature vectors. The test data were spoken by the same female who recorded the training data. The text utterances, however, were different.

B. Experiment results

Figure 5 shows the direction accuracy performance using the acoustic transfer function estimated at various speech lengths, and the performance is compared to the power-based technique. The top figure shows the accuracy for the same text utterance at each angle of the active microphone, and the bottom figure shows the accuracy for a different text utterance at each angle of the active microphone. The test data for the same text utterance consisted of 100 segments. The test

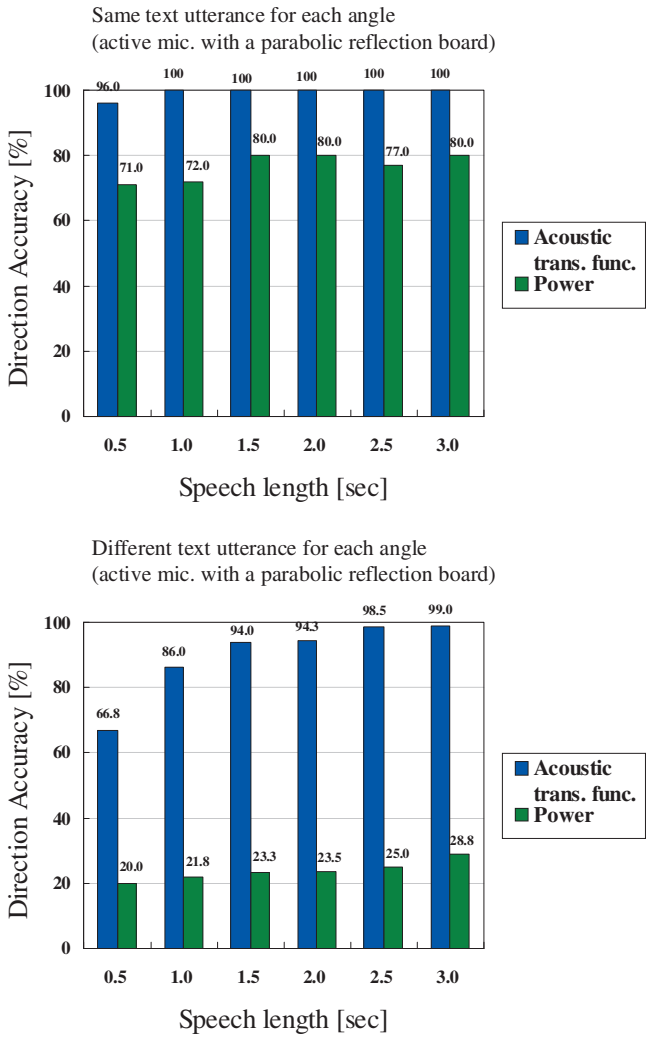


FIG. 5. Performance of an active microphone with a parabolic reflection board.

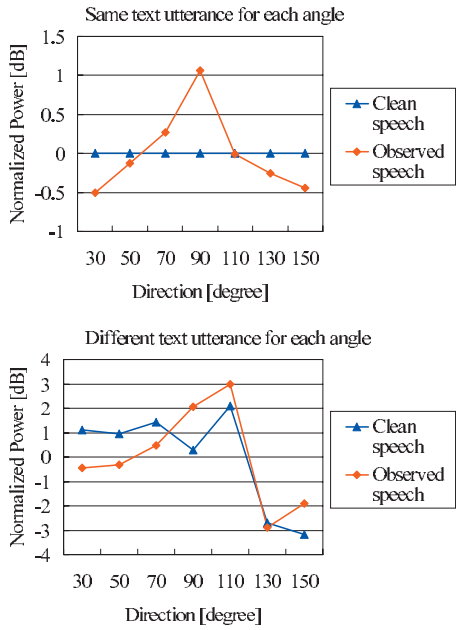


FIG. 6. Power of a clean speech segment that did not estimate the direction of arrival correctly and the speech segment observed by the active microphone with a parabolic reflection board for each angle. The power was normalized so that the mean value of all directions was 0 dB.

data for the different utterance consisted of 600, 300, 200, 150, 120, and 100 segments, where one segment has time lengths of 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 s, respectively. The test for the same text utterance was conducted 100 times, and that for the different text utterance was conducted 600 times by changing the combination of the text utterances for each direction. As shown in the top figure, the performance for both the techniques based on the power and the acoustic transfer function is high. However, the possibility of there being an identical text utterance at each angle of the active microphone will be very small in a real environment.

In the bottom portion of Fig. 5, we can see that the performance of the power-based technique degrades drastically when the utterance text differs at each angle of the active microphone because the power of the speech signal varies for all directions of the active microphone. Figure 6 shows the power of a clean speech segment that did not estimate the direction of arrival correctly and the observed speech segment for each angle. The power was normalized so that the mean values of all directions were 0 dB. As shown in this figure, the power of the observed speech was most enhanced by the parabolic reflection board at 90 deg. But when the utterance text differs at each angle of the active microphone, the power of observed speech at 90 deg did not have the maximum power since the power of input speech for another direction had a higher power than that for 90 deg.

On the other hand, the performance of the new method based on the acoustic transfer function is high, even for different text utterances. This is because the new method uses the information of the acoustic transfer function, which depends on the direction of the active microphone only and does not depend on the utterance text. Also, we can see that the shorter the speech length for each angle is, the more the direction accuracy decreases. One reason for this is that the statistics for the observed speech are not readily available if there are not enough samples used to estimate the acoustic transfer function.

Figure 7 shows the performance of a shotgun microphone (SONY ECM-674) without a parabolic reflection board. The power-based method can provide good performance for the same text utterance at each angle of the shotgun microphone due to the directivity of the shotgun microphone, but the performance degrades when the utterance text differs at each angle of the shotgun microphone. On the other hand, the performance of the new method based on the acoustic transfer function is even lower. The directivity of the shotgun microphone changes drastically as the sound-source direction changes from the front direction to the side directions of the shotgun microphone, and, as a result, the acoustic transfer function that is farthest from all the other acoustic transfer functions comes to be that at 30 or 150 deg in Eq. (6). The mean values of all acoustic transfer functions for the shotgun microphone and a parabolic reflection board are plotted in Figs. 8 and 9, respectively, where the acoustic transfer function is computed by Eq. (8) using true clean speech signal $S_{\mathrm{cep}}(d;n)$ and the total number of frames is 36 600. Then the mean values are computed. As shown in

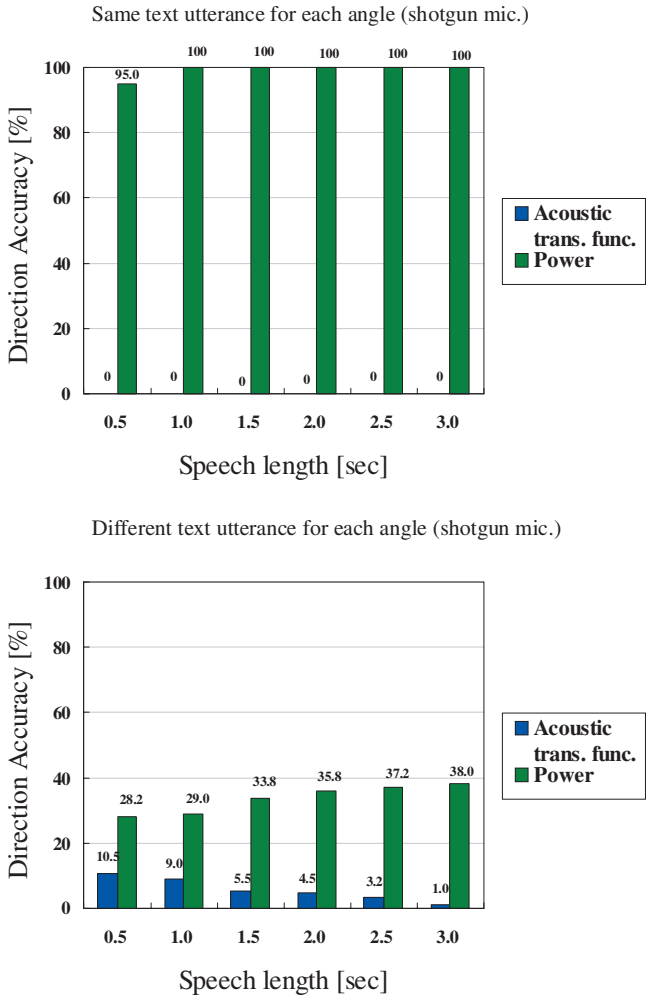


FIG. 7. Performance of a shotgun microphone without a parabolic reflection board.

Fig. 8, we can see that the acoustic transfer function that is farthest from all the other acoustic transfer functions is that at 30 or 150 deg. As shown in Fig. 9, on the other hand, the acoustic transfer function that is farthest from all the other acoustic transfer functions is that at 90 deg to the target direction.

Figures 10 and 11 show the plot of acoustic transfer function for 300 segments of observed speech for the case of the active microphone. In Fig. 10, the acoustic transfer function H_{sub} was computed by Eq. (8) using true clean speech signal $S_{\mathrm{cep}}(d;n)$. On the other hand, in Fig. 11, the acoustic transfer function H_{est} was estimated by Eq. (17) using only the statistics of clean speech GMM. As shown in Fig. 10, when the active microphone does not face the sound source, H_{sub} is distributed in almost the same place, and H_{sub} of the sound source direction is distributed away from the H_{sub} of other directions. In Fig. 11, though the distribution of the estimated H_{est} may have some slight variations, it can be said that the distribution of H_{est} is similar to that of H_{sub}.

Figure 12 shows the difference in the direction accuracy between the use of H_{sub} (the true clean speech data) and H_{est} (the statistics of clean speech model: GMM). As shown in this figure, when the utterances for each angle consist of the same text, the direction accuracy was 100%. However, when

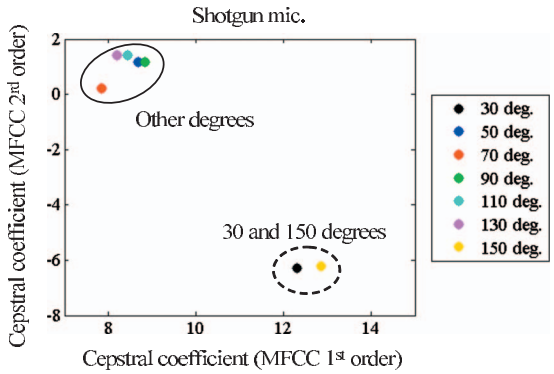


FIG. 8. Mean values of the acoustic transfer functions for the shotgun microphone.

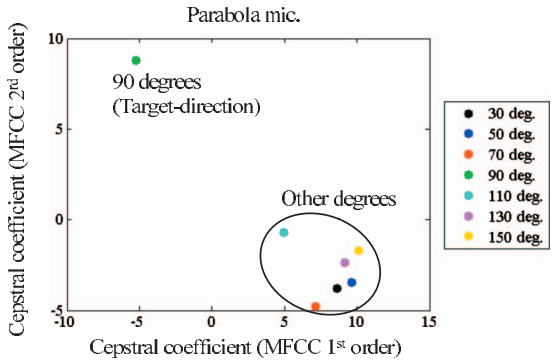


FIG. 9. Mean values of the acoustic transfer functions for the microphone with a parabolic reflection board.

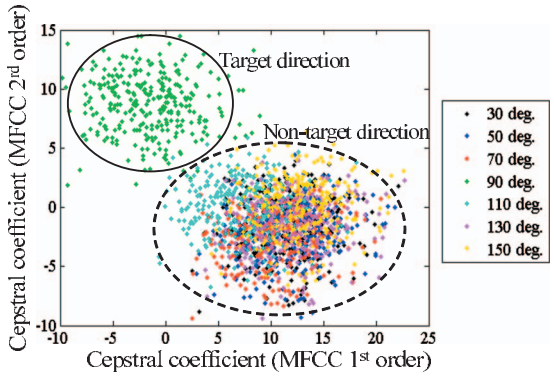


FIG. 10. Acoustic transfer function computed by using true clean speech data at each angle in the cepstral domain.

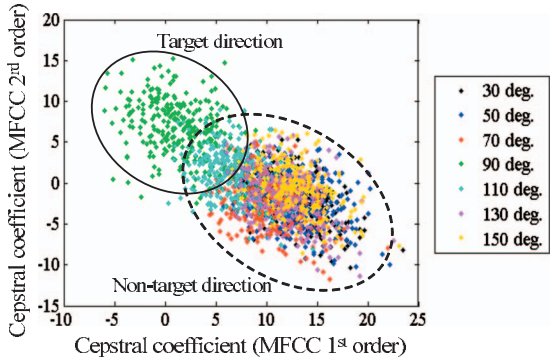


FIG. 11. Acoustic transfer function estimated by the proposed method using only the statistics of clean speech GMM at each angle in the cepstral domain.

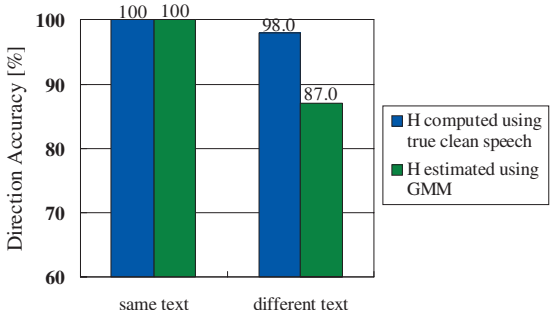


FIG. 12. Comparison of true clean speech data and clean speech model.

the texts of utterances for each angle are different, the direction accuracy obtained using H_{est} decreased. This is because the value of H_{est} was influenced to some extent by the phoneme sequence of clean speech.

V. CONCLUSIONS

This paper has introduced the concept of an active microphone that achieves a good combination of active-operation and signal processing, and described a sound-source-direction estimation method using a single microphone with a parabolic reflection board. The experiment results in a room environment confirmed that the acoustic transfer function influenced by parabolic reflection can clarify the difference between the target direction and the non-target direction. In future work, more research will be needed in regard to different utterances and direction estimation in short intervals. Also, we intend to investigate the performance of the proposed system in noisy environments, such as with multiple sound sources and when the orientation of the speaker's head changes, and to test the performance of the system in a speaker-independent speech model.

[1]D. Johnson and D. Dudgeon, *Array Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1996).
[2]M. Omologo and P. Svaizer, "Acoustic event localization in noisy and reverberant environment using csp analysis," in Proceedings of the ICASSP (1996), pp. 921–924.
[3]O. Ichikawa, T. Takiguchi, and M. Nishimura, "Sound source localization using a pinna-based profile fitting method," in Proceedings of the International Workshop on Acoustic Echo and Noise Control (2003) pp. 263–266.
[4]N. Ono, Y. Zaitsu, T. Nomiyama, A. Kimachi, and S. Ando, "Biomimicry sound source localization with fishbone," IEEJ Trans. Sens. Micromachines **121-E**, 313–319 (2001).
[5]T. Kristjansson, H. Attias, and J. Hershey, "Single microphone source separation using high resolution signal reconstruction," in Proceedings of the ICASSP (2004), pp. 817–820.
[6]B. Raj, M. V. S. Shashanka, and P. Smaragdis, "Latent dirichlet decomposition for single channel speaker separation," in Proceedings of the ICASSP (2006), pp. 821–824.
[7]G.-J. Jang, T.-W. Lee, and Y.-H. Oh, "A subspace approach to single channel signal separation using maximum likelihood weighting filters," in Proceedings of the ICASSP (2003), pp. 45–48.
[8]T. Nakatani and B.-H. Juang, "Speech dereverberation based on probabilistic models of source and room acoustics," in Proceedings of the ICASSP (2006), pp. I-821–I-824.
[9]T. Takiguchi, Y. Sumida, R. Takashima, and Y. Ariki, "Single-channel talker localization based on discrimination of acoustic transfer functions," EURASIP Journal on Advances in Signal Processing **2009**, 918404 (2009).
[10]B. Saka and A. Kaderli, "Direction of arrival estimation and adaptive nulling in array-fed reflectors," in Electrotechnical Conference (1998), pp. 274–277.
[11]T. Takiguchi, R. Takashima, and Y. Ariki, "Active microphone with parabolic reflection board for estimation of sound source direction," in Proceedings of the Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA2008) (2008), pp. 65–68.
[12]B.-H. Juang, "Maximum likelihood estimation for mixture multivariate stochastic observations of Markov chains," AT&T Tech. J. **64**, 1235–1250 (1985).

Deconvolution of sparse underwater acoustic multipath channel with a large time-delay spread

Wen-Jun Zeng[a)]
Key Laboratory of Underwater Acoustic Communication and Marine Information Technology of the Ministry of Education, Xiamen University, Xiamen 361005, China

Xue Jiang
Sunplus mMobile Inc., Beijing Branch, Beijing 100085, China

Xi-Lin Li
Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, Maryland 21250

Xian-Da Zhang
Department of Automation, Tsinghua University, Beijing 100084, China

(Received 17 April 2009; revised 8 November 2009; accepted 4 December 2009)

The deconvolution of multipath underwater acoustic channel with a large time-delay spread is investigated. The channel deconvolution involves estimating the multipath time-delays and attenuation factors from a noisy received signal consisting of multiple overlapped signals. Similar to conventional deconvolution methods, the proposed method estimates channel impulse response based on least-squares criterion. However, the proposed method harnesses the sparse structure of an underwater acoustic channel, and ℓ_1-norm of the channel impulse response is adopted as the cost function to be minimized. In addition, the available *a priori* knowledge of support constraint and attenuation factor constraint are imposed and channel deconvolution problem is converted to a convex optimization problem. Instead of employing the existing standard algorithms, which require huge storage space and high computational complexity, a simple iterative algorithm for solving the optimization problem with fast convergence rate and low complexity is developed. The computational complexity of the proposed algorithm is $\mathcal{O}(N \log_2(N))$ per iteration with N being the length of the received signal. Simulation results confirm that the proposed method provides better performance in terms of temporal resolution and robustness to noise compared with other extant multipath channel deconvolution techniques.
 [DOI: 10.1121/1.3278604]

PACS number(s): 43.60.Pt, 43.60.Jn, 43.60.Lq, 43.60.Rw [WMC] Pages: 909–919

I. INTRODUCTION

Multipath propagation phenomenon often arises in fields such as sonar, radar, wireless communications, and geophysics. In an underwater acoustic channel,[1] due to multiple reflections or scattering from the boundaries, the received signal can be viewed as the superposition of a number of delayed and amplitude-weighted replicas of the original transmitted signal. In many applications such as oceanic acoustic tomography,[2] underwater acoustic communications,[3] underwater acoustic source localization,[4,5] and underwater target classification,[6,7] it is of interest to estimate the multipath channel parameters (including multipath time-delays and attenuation factors) and remove the multipath effects. The task of channel deconvolution is to estimate the channel impulse response using the received signal and the known transmitted waveform.

The classical method for time-delay estimation is matched filtering, which is equal to the cross correlation of the received and transmitted signals. It has been shown that matched filtering is equivalent to the maximum likelihood (ML) estimator if there exists only one single path.[8] However, the resolution of matched filtering is limited by the reciprocal of the transmitted signal bandwidth and it is unable to distinguish the multipath components that are spaced closer than the resolution limit. To improve the temporal resolution, various techniques have been proposed. A class of approaches first transforms the time-delay estimation problem into a spectral estimation problem via Fourier transform;[9–12] then high-resolution spectral estimation algorithms, e.g., MUSIC,[13] are applied to estimate the time-delay parameters. These methods require the number of paths to be known and they are unable to attain the estimates of attenuation factors. Several optimization strategies have been proposed to minimize a complicated nonlinear least-squares (LS) cost function for joint time-delay and attenuation estimation.[14–16] However, since the nonlinear least-squares cost function is not convex and highly oscillatory, its global minimum is difficult to seek and a nearby local minimum is often obtained.

[a)]Author to whom correspondence should be addressed. Also at Department of Automation, Tsinghua University, Beijing 100084, China. Electronic mail: cengwj06@mails.tsinghua.edu.cn

Deconvolution is a typical inverse problem whose most notable feature is ill-conditioning.[17] The elementary deconvolution method is inverse filtering. The inverse filtering is simple but is sensitive to noise and difficult to be applied to practical problems.[8] To overcome the ill-conditioned problem of deconvolution, a constraint least-squares method[18] exploits *a priori* support knowledge, which is obtained from the matched filter to regularize and improve the performance. The set-theoretic deconvolution method using projection onto convex sets (POCS) (Ref. 19) uses not only the support constraint information but also variance of residual and amplitude constraints. The POCS method finds a feasible solution that satisfies these convex constraints as the estimate of the channel response. However, the performances of both constraint least-squares and POCS depend on the size of the available support constraint set. When the support constraint set is not small enough, the performances of these two methods are unsatisfactory.

By considering the sparseness of the underwater acoustic channel, the sparse representation based approaches are proposed to estimate the channel response. The matching pursuit (MP) algorithm[20] is first applied to the sparse channel estimation problem.[21] The order-recursive LS-MP algorithm was proposed to estimate the time-varying sparse underwater acoustic channel.[22] MP is a well-known algorithm for sparse representation, but it uses greedy search and the global optimum is not guaranteed. In channel estimation, MP iteratively subtracts channel path contributions, which may result in errors if the channel paths are closely spaced.

The accurate metric of sparsity is counting the nonzero numbers of a vector, which also equals the ℓ_0-norm[23] of the vector. However, ℓ_0-norm is not convex and its minimization is non-deterministic polynomial-time (NP)-hard. To relax the NP-hard problem, the basis pursuit (BP) (Ref. 23) is proposed and it adopts ℓ_1-norm[23] instead of ℓ_0-norm. Since the ℓ_1-norm of a vector is convex, the ℓ_1-norm minimization is a convex optimization and can be efficiently solved. The BP algorithm is designed for overcomplete, i.e., *underdetermined* signal representation, whereas the channel estimation problem is based on *overdetermined* or *well-determined* model. The reconstructed error is adopted as the cost function, and the ℓ_1-norm is added to this cost function as a regularization term.[24,25] The channel deconvolution criterion[24] can be converted into a quadratic program (QP) and a standard program for solving QP problem is used. However, this deconvolution criterion does not take into account the support information and the attenuation factor constraint. Moreover, the standard algorithm for the QP problem requires huge store space and has a high computational complexity.

In this paper, a new deconvolution method for estimating the response of sparse underwater acoustic multipath channel is developed. Then the multipath time-delays and amplitudes can be obtained from the estimated channel response. The proposed method casts the channel estimation to a convex optimization problem in which ℓ_1-norm of the channel impulse response is used as the cost function and the residual fitting error constraint, support, and attenuation constraints are taken into account.

The remainder of this paper is organized as follows. In Sec. II, the deconvolution problem and channel model are introduced. In Sec. III, the limits of the widely used least-squares deconvolution method are pointed out. We present the proposed channel deconvolution method in detail in Sec. IV. In Sec. V, extensive numerical simulations are conducted to compare the performances of our new method with other extant techniques. Finally, this paper is concluded in Sec. VI.

II. DECONVOLUTION PROBLEM AND CHANNEL MODEL

A. Channel deconvolution problem

The received signal $x(t)$ of an underwater acoustic channel in response to the predetermined probing signal $s(t)$ can be written as

$$x(t) = h(t) \otimes s(t) + v(t) = \int h(\tau)s(t-\tau)\mathrm{d}\tau + v(t), \tag{1}$$

where $\otimes$ denotes convolution, $h(t)$ is the impulse response of the underwater acoustic channel, and $v(t)$ is the additive white Gaussian noise (AWGN) with variance σ_v^2. In this paper, we develop an efficient algorithm for identifying the channel response $h(t)$ from the known source $s(t)$ and the received signal $x(t)$. The channel identification problem is also referred to as deconvolution problem. By exploiting the least-squares criterion, the deconvolution problem can be described as a functional extremum problem

$$\min_{h(t)} \int \left\| x(t) - \int h(\tau)s(t-\tau)\mathrm{d}\tau \right\|^2 \mathrm{d}t. \tag{2}$$

It is known that deconvolution belongs to the inverse problem; therefore it exhibits the ill-conditioned property.

B. Underwater acoustic channel model

The impulse response of an underwater acoustic multipath channel is modeled as

$$h(t) = \sum_{k=1}^{M} a_k \delta(t-\tau_k), \tag{3}$$

where $\delta(\cdot)$ denotes the delta function, M is the number of different paths, and a_k and τ_k $(k=1,\ldots,M)$ denote the attenuation factors (amplitudes) and time-delays of path k, respectively. Then the observation $x(t)$ can be written as the sum of M scaled and delayed replicas of the source $s(t)$ as follows:

$$x(t) = \sum_{k=1}^{M} a_k s(t-\tau_k) + v(t). \tag{4}$$

Under the multipath channel model assumption, the channel deconvolution problem is reduced to estimating the multipath parameters $\{a_k, \tau_k\}_{k=1}^{M}$.

Different from the wireless radio channel, there are two distinctive properties of a typical underwater acoustic channel.

(1) The acoustic channel has a large time-delay spread (i.e., a long impulse response) due to the low acoustic velocity. Therefore the taps of the impulse response of such a channel will be large, which makes identifying the channel of this type a challenging task.
(2) The impulse response of the underwater acoustic channel has a small nonzero support since the number of the propagation acoustic rays is limited. In other words, a typical underwater acoustic channel is sparse.

After sampling with sampling period T_s, we obtain the discrete-time version of Eq. (1) as follows:

$$\boldsymbol{x} = \boldsymbol{S}\boldsymbol{h} + \boldsymbol{v}, \tag{5}$$

where the vectors $\boldsymbol{x}=[x(0),\ldots,x(N-1)]^T$, $\boldsymbol{h}=[h(0),\ldots,h(N-1)]^T$, and $\boldsymbol{v}=[v(0),\ldots,v(N-1)]^T$ are formed from the samples of $x(t)$, $h(t)$, and $v(t)$, respectively. Note that all of the signals are sufficiently zero-padded to length N; therefore the linear convolution equals the circular convolution and the $N\times N$ circulant matrix is

$$\boldsymbol{S} = \begin{bmatrix} s(0) & s(N-1) & \cdots & s(1) \\ s(1) & s(0) & \ddots & \vdots \\ \vdots & \vdots & \ddots & s(N-1) \\ s(N-1) & s(N-2) & \cdots & s(0) \end{bmatrix}. \tag{6}$$

For an M-path channel, $\boldsymbol{h}$ has only M nonzero elements whose values are the M corresponding attenuation factors if the time-delays are integer multiples of the sampling period. It is clear that $\boldsymbol{h}$ is sparse due to $M \ll N$. Note that if the time-delays are integer multiples of the sampling period T_s, the location of nonzero elements of $\boldsymbol{h}$ indicates these time-delays. However, $\boldsymbol{h}$ will be smeared if there exist fractional time-delays. The fractional time-delays can be estimated by interpolation of the discrete estimate $\hat{\boldsymbol{h}}$, which will be discussed in detail in Sec. IV D. The channel deconvolution goal is to obtain an estimate $\hat{\boldsymbol{h}}$ from which the multipath parameters can be obtained.

III. LEAST-SQUARES DECONVOLUTION AND ITS LIMITS

The least-squares estimate of Eq. (5) without any constraints on the impulse response is

$$\hat{\boldsymbol{h}}_{\mathrm{LS}} = \arg\min_{\boldsymbol{h}} \|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{h}\|^2. \tag{7}$$

The least-squares solution can be expressed as

$$\hat{\boldsymbol{h}}_{\mathrm{LS}} = \boldsymbol{S}^{\dagger}\boldsymbol{x} = \boldsymbol{S}^{-1}\boldsymbol{x}, \tag{8}$$

where the superscript † denotes the Moore–Penrose pseudoinverse. Since $\boldsymbol{S}$ is a nonsingular squared matrix, we have $\boldsymbol{S}^{\dagger}=\boldsymbol{S}^{-1}$. We will point out that $\hat{\boldsymbol{h}}_{\mathrm{LS}}$ is equivalent to the inverse filtering. The eigenvalue decomposition of the circulant matrix $\boldsymbol{S}$ is given by[26]

$$\boldsymbol{S} = \boldsymbol{F}^H \boldsymbol{\Lambda}_s \boldsymbol{F}, \tag{9}$$

where the superscript ^{H} denotes the conjugate transpose and $\boldsymbol{F}$ is the $N\times N$ Fourier matrix whose (k,l)th element is

$$\boldsymbol{F}_{kl} = \frac{1}{\sqrt{N}} e^{-j(2\pi/N)kl}. \tag{10}$$

Premultiplying matrix $\boldsymbol{F}$ is equivalent to performing discrete Fourier transform (DFT) and premultiplying $\boldsymbol{F}^H$ means carrying out inverse discrete Fourier transform (IDFT). $\boldsymbol{\Lambda}_s$ is a diagonal matrix whose diagonal elements are the DFT of $\boldsymbol{s}=[s(0),\ldots,s(N-1)]^T$; i.e.,

$$\boldsymbol{\Lambda}_s = \mathrm{diag}\{\mathrm{DFT}[\boldsymbol{s}]\}. \tag{11}$$

The eigenvalue decomposition of $\boldsymbol{S}^{-1}$ is

$$\boldsymbol{S}^{-1} = \boldsymbol{F}^H \boldsymbol{\Lambda}_s^{-1} \boldsymbol{F}. \tag{12}$$

Substituting Eq. (12) to Eq. (8) leads to

$$\hat{\boldsymbol{h}}_{\mathrm{LS}} = \boldsymbol{F}^H \boldsymbol{\Lambda}_s^{-1} \boldsymbol{F}\boldsymbol{x}. \tag{13}$$

Since $\boldsymbol{F}$ is unitary, we have $\boldsymbol{F}\boldsymbol{F}^H=\boldsymbol{I}$ with $\boldsymbol{I}$ denoting the identity matrix. Hence Eq. (13) becomes

$$\boldsymbol{F}\hat{\boldsymbol{h}}_{\mathrm{LS}} = \boldsymbol{\Lambda}_s^{-1}\boldsymbol{F}\boldsymbol{x}. \tag{14}$$

By exploiting $\boldsymbol{F}\boldsymbol{x}=\mathrm{DFT}[\boldsymbol{x}]$ and $\boldsymbol{F}\hat{\boldsymbol{h}}_{\mathrm{LS}}=\mathrm{DFT}[\hat{\boldsymbol{h}}_{\mathrm{LS}}]$, we obtain

$$\hat{\boldsymbol{h}}_{\mathrm{LS}} = \mathrm{IDFT}[\mathrm{DFT}[\boldsymbol{x}] \oslash \mathrm{DFT}[\boldsymbol{s}]], \tag{15}$$

where $\oslash$ denotes element-by-element division of two vectors. Equation (15) means that $\hat{\boldsymbol{h}}_{\mathrm{LS}}$ is equivalent to the result of inverse filtering. The inverse filtering does not exploit any *a priori* information or any constraints to regularize; therefore it suffers from the ill-conditioned problem and is very sensitive to noise.

Manickam and Vaccaro[18] proposed to use *a priori* support constraint to regularize and improve the performance. The range of values where the delays could possibly occur can be estimated from the matched-filter output. The support constraint is given by

$$C_s = \{\boldsymbol{h}:\boldsymbol{h}(i) = 0, i \notin E\}, \tag{16}$$

with E being the indicator set of indices for $\boldsymbol{h}$. This indicator set is expected to contain the indices corresponding to the nonzero elements of $\boldsymbol{h}$. Let the size of set E be N_E and it is clear that $N_E<N$. Then the data model can be rewritten as

$$\boldsymbol{x} = \boldsymbol{S}_E\boldsymbol{h}_E + \boldsymbol{v}, \tag{17}$$

where $\boldsymbol{S}_E$ contains only those columns of $\boldsymbol{S}$ corresponding to the indicator set E and $\boldsymbol{h}_E$ is a vector of N_E elements. Note that the $N\times N_E$ matrix $\boldsymbol{S}_E$ is a "tall" matrix; hence the data model of Eq. (17) is overdetermined. The least-squares solution of Eq. (17) is

$$\hat{\boldsymbol{h}}_E = \boldsymbol{S}_E^{\dagger}\boldsymbol{x} = (\boldsymbol{S}_E^T\boldsymbol{S}_E)^{-1}\boldsymbol{S}_E^T\boldsymbol{x}. \tag{18}$$

Equation (18) is the result of constraint least-squares deconvolution, and its performance is better than the inverse filtering if $N_E<N$. However, the success of this constraint least-squares deconvolution depends on the size of the indicator set. Only when the indicator set is small enough, the performance of this method is satisfactory.

IV. THE PROPOSED METHOD

In this section, we elaborate the new algorithm for fast estimation of sparse channel via convex optimization (FESCCO).

A. Constraints and cost function

The proposed FESCCO method harnesses both the sparse structure of the multipath channel and the available *a priori* information, which can be used as the constraints. The residual fitting error constraint is important since it reflects the fitting quality of the estimated result. We expect to find such an $\boldsymbol{h}$ satisfying the residual fitting error constraint[19,27]

$$C_v = \{\boldsymbol{h}: \|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{h}\|^2 \le \epsilon\}, \tag{19}$$

where the error bound $\epsilon = (\gamma\sqrt{2N} + N)\sigma_v^2$ is related to the noise variance σ_v^2 with the constant γ reflecting the confidence interval with which the ideal solution is in set C_v. Denote this confidence interval as P_d. Note that P_d is a probability and $0 \le P_d \le 1$. Since the noise is white and Gaussian, the relationship between the confidence P_d and γ can be expressed by

$$\mathrm{erf}\left(\frac{\gamma}{\sqrt{2}}\right) = P_d,$$

where $\mathrm{erf}(u) = 2/\sqrt{\pi}\int_0^u e^{-z^2}\mathrm{d}z$ is the error function. Then γ can be calculated by

$$\gamma = \sqrt{2}\ \mathrm{erf}^{-1}(P_d),$$

with $\mathrm{erf}^{-1}(\cdot)$ denoting the inverse error function. Clearly, C_v is a convex set.[27] Moreover, we should consider the constraint of the estimated attenuation factors since the true attenuation factors are not larger than 1. Hence we define the attenuation factors' constraint set

$$C_a = \{\boldsymbol{h}: a_l \le \boldsymbol{h}(i) \le a_u,\ \ i = 1, \ldots, N\}, \tag{20}$$

with a_l and a_u representing the lower bound and the upper bound of the attenuation factors. If there does not exist phase reversal, the attenuation factor bounds take values as $a_l = 0$ and $a_u = 1$. If phase reversal is possible, e.g., the reflection coming from the surface of an ocean, the bounds should be modified as $a_l = -1$ and $a_u = 1$.

The support constraint C_s, the residual fitting error constraint C_v, and the attenuation factors' constraint C_a are taken into consideration in our FESCCO method. A set-theoretic deconvolution method using POCS was proposed.[19] The POCS method finds a feasible solution, which is in the intersection of the three convex sets: $C_s \cap C_v \cap C_a$. The POCS method is regularized but it does not consider the sparsity of $\boldsymbol{h}$. Therefore the channel estimate obtained by POCS is thorny and pseudopeaks representing non-true paths will appear when the signal-to-noise ratio (SNR) is relatively low.

In this paper, considering the sparsity of $\boldsymbol{h}$, we aim to find a sparse solution of $\boldsymbol{h}$, which satisfies three convex constraints mentioned above. The measurement of sparsity is the number of the nonzero elements of a vector, which is also equal to the ℓ_0-norm of the vector. Therefore the sparsest solution of $\boldsymbol{h}$ is given by

$$\min \|\boldsymbol{h}\|_0,$$

$$\text{s.t. } \boldsymbol{h} \in C_s \cap C_v \cap C_a, \tag{21}$$

where $\|\cdot\|_0$ denotes the ℓ_0-norm of a vector. The combinatory optimization problem of Eq. (21) is non-convex and is NP-hard, which means that no known algorithm for finding the sparsest solution is significantly more efficient than exhausting all subsets of the entries for $\boldsymbol{h}$. The MP algorithm[20] is a greed search strategy to find a sparse solution. It is simple but it cannot guarantee to attain the global optimal solution. In multipath channel estimation, MP iteratively subtracts channel path contributions and may lead to errors in the presence of very close paths.

We use the ℓ_1-norm instead of the ℓ_0-norm, and the channel estimation problem can be reformulated as

$$\min \|\boldsymbol{h}\|_1,$$

$$\text{s.t. } \|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{h}\|^2 \le \epsilon,$$

$$\boldsymbol{h}(i) = 0, \quad i \notin E,$$

$$a_l \le \boldsymbol{h}(i) \le a_u, \quad i = 1, \ldots, N, \tag{22}$$

where $\|\boldsymbol{h}\|_1 = \Sigma_{i=1}^N |h(i)|$ is the ℓ_1-norm of $\boldsymbol{h}$ with $|\cdot|$ denoting the absolute value of a real number. It is clear that Eq. (22) is a convex optimization problem and the global convergence can be guaranteed. Moreover, it can be solved with a polynomial time complexity.

It should be pointed out that our model shown in Eq. (22) is different from the ℓ_1-norm regularization deconvolution method,[24] which is proposed by Fuchs and is shown as

$$\min \|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{h}\|^2 + \eta\|\boldsymbol{h}\|_1, \tag{23}$$

where η is the regularization factor. The criterion of Eq. (23) is known in statistical literature as the Lasso.[28] The regularization factor η is equal to the inverse of Lagrange multiplier corresponding to the constraint $\|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{h}\|^2 \le \epsilon$. For each η, there exists a corresponding ϵ. However, ϵ is easily determined from the noise variance but η is difficult to determine. In addition, the Lasso model does not take into account the support knowledge and the attenuation factor constraint, which is different from our model. Fuchs[24] proposed that the optimization problem of Eq. (23) was first transformed into a QP and a standard program for solving QP problem was used. However, the standard algorithm for the QP problem requires one to store the huge $N \times N$ matrix $\boldsymbol{S}$ and has a high computational complexity. In Sec. IV B, we will develop a new simple algorithm for solving Eq. (22), which has much lower complexity than the deconvolution approach by virtue of QP.

B. New iterative algorithm for channel deconvolution

There exist several techniques,[29] e.g., the standard primal-dual interior point method,[29,30] to solve the convex optimization problem shown in Eq. (22). However, directly applying the existing standard algorithms to solve this convex optimization problem leads to the following two drawbacks.

(1) It requires one to construct and store the $N\times N$ matrix $\boldsymbol{S}$ for the standard optimization algorithms. It will cost tremendous storage space when the length of the received signal N is large.
(2) As a universal program for convex optimization, the standard optimization algorithm does not take into account the special circulant structure of matrix $\boldsymbol{S}$. However, it is expected that exploiting such a structure will reduce the computational complexity.

Herein we propose a new simple iterative algorithm to solve Eq. (22) with a fast convergence rate. Moreover, the computational complexity of the new algorithm is reduced by harnessing the circulant structure of matrix $\boldsymbol{S}$. We can use the output of the matched filter as the initial value of our algorithm, which is denoted as $\hat{\boldsymbol{h}}_0$.

The gradient of the cost function can be calculated as

$$\nabla_{\boldsymbol{h}}\|\boldsymbol{h}\|_1 = \mathrm{sgn}(\boldsymbol{h}), \tag{24}$$

where $\mathrm{sgn}(\cdot)$ is the sign function. The channel estimate result of the kth iteration is represented by $\hat{\boldsymbol{h}}_k$ $(k=1,2,\ldots)$. In the kth iteration, we first update $\hat{\boldsymbol{h}}_{k-1}$ along the steep descent direction

$$\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_{k-1} - \mu_k\,\mathrm{sgn}(\hat{\boldsymbol{h}}_{k-1}), \tag{25}$$

where μ_k is the step size in the kth iteration. The optimal step size $\mu_k^{[\mathrm{opt}]}$ in the kth iteration is obtained via the following line search:

$$\mu_k^{[\mathrm{opt}]} = \underset{\mu}{\mathrm{argmin}}\|\hat{\boldsymbol{h}}_{k-1} - \mu\,\mathrm{sgn}(\hat{\boldsymbol{h}}_{k-1})\|_1.$$

This line search problem can be easily solved by the existing line search methods, such as Golden search.[29] To circumvent this line search problem, one can use a small positive constant μ as a fixed step size as long as μ is small enough to avoid divergence. However, a fixed step size usually results in a trade-off between the residual error and the convergence rate. To attain both fast convergence rate and small residual error, the variable step size strategy is adopted. In each iteration, the value of the cost function is recorded. When the algorithm detects that the cost function oscillates, the step size is halved and the algorithm begins a new linear search starting from the best solution ever found.

After being updated by Eq. (25), $\hat{\boldsymbol{h}}_k$ may be not in the feasible region, i.e., the intersection of three constraint sets. Therefore we need to project $\hat{\boldsymbol{h}}_k$ back onto the feasible set $C_s\cap C_v\cap C_a$. The projection operators onto the convex sets C_s, C_a, and C_v are represented by $\boldsymbol{P}_s$, $\boldsymbol{P}_a$, and $\boldsymbol{P}_v$, respectively. It is straightforward to give the expressions of projection operators

$$\boldsymbol{P}_s\hat{\boldsymbol{h}}_k(i) = \begin{cases} \hat{\boldsymbol{h}}_k(i) & \text{for } i\in E \\ 0 & \text{for } i\notin E \end{cases} \tag{26}$$

and

$$\boldsymbol{P}_a\hat{\boldsymbol{h}}_k(i) = \begin{cases} \hat{\boldsymbol{h}}_k(i) & \text{if } a_l\le\hat{\boldsymbol{h}}_k(i)\le a_u \\ a_l & \text{if } \hat{\boldsymbol{h}}_k(i)<a_l \\ a_u & \text{if } \hat{\boldsymbol{h}}_k(i)>a_u. \end{cases} \tag{27}$$

By denoting the projection onto C_v of $\hat{\boldsymbol{h}}_k$ as $\boldsymbol{y}=\boldsymbol{P}_v\hat{\boldsymbol{h}}_k$, then we have

$$\min\|\boldsymbol{y}-\hat{\boldsymbol{h}}_k\|^2,$$

$$\text{s.t. } \|\boldsymbol{x}-\boldsymbol{S}\boldsymbol{y}\|^2=\epsilon. \tag{28}$$

The Lagrange function of Eq. (28) is

$$l(\boldsymbol{y}) = \|\boldsymbol{y}-\hat{\boldsymbol{h}}_k\|^2 + \lambda(\|\boldsymbol{x}-\boldsymbol{S}\boldsymbol{y}\|^2-\epsilon), \tag{29}$$

where λ is the Lagrange multiplier. Solving $\partial l(\boldsymbol{y})/\partial\boldsymbol{y}=0$ leads to

$$\boldsymbol{y} = \boldsymbol{P}_v\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_k + \left(\boldsymbol{S}^T\boldsymbol{S}+\frac{1}{\lambda}\boldsymbol{I}\right)^{-1}\boldsymbol{S}^T(\boldsymbol{x}-\boldsymbol{S}\hat{\boldsymbol{h}}_k). \tag{30}$$

Substituting Eq. (30) into the constraint $\|\boldsymbol{x}-\boldsymbol{S}\boldsymbol{y}\|^2=\epsilon$, we obtain

$$\left\|\left(\boldsymbol{I}-\boldsymbol{S}\left(\boldsymbol{S}^T\boldsymbol{S}+\frac{1}{\lambda}\boldsymbol{I}\right)^{-1}\boldsymbol{S}^T\right)(\boldsymbol{x}-\boldsymbol{S}\hat{\boldsymbol{h}}_k)\right\|^2 = \epsilon. \tag{31}$$

The Lagrange multiplier λ can be obtained according to Eq. (31). However, computing the left side of Eq. (31) directly will result in a high computational complexity. The computational amount can be reduced by exploiting the circulant structure of $\boldsymbol{S}$. Define the residual error vector in the kth iteration as

$$\boldsymbol{r}_k = \boldsymbol{x}-\boldsymbol{S}\hat{\boldsymbol{h}}_k. \tag{32}$$

In this paper, the notation "$\tilde{}$" is used to represent the DFT of a vector. Therefore we have

$$\tilde{\boldsymbol{r}}_k = \mathrm{DFT}[\boldsymbol{x}-\boldsymbol{S}\hat{\boldsymbol{h}}_k]. \tag{33}$$

It can be proven that Eq. (31) can also be expressed as

$$\sum_{i=0}^{N-1}\frac{|\tilde{\boldsymbol{r}}_k(i)|^2}{(\lambda|\tilde{\boldsymbol{s}}(i)|^2+1)^2} = \epsilon \tag{34}$$

in the frequency domain.[30] Since the DFT and IDFT can be efficiently implemented by fast Fourier transform (FFT), λ can be easily solved with a low computational complexity based on Eq. (34). Once λ is obtained, the projection onto C_v can be computed according to Eq. (30). Herein we propose a low complexity method to compute $\boldsymbol{P}_v\hat{\boldsymbol{h}}_k$. First we calculate $\boldsymbol{r}_k$ as follows:

$$\boldsymbol{r}_k = \boldsymbol{x}-\boldsymbol{S}\hat{\boldsymbol{h}}_k = \boldsymbol{x}-\boldsymbol{F}^H\boldsymbol{\Lambda}_s\boldsymbol{F}\hat{\boldsymbol{h}}_k = \boldsymbol{x}-\mathrm{IDFT}[\tilde{\hat{\boldsymbol{h}}}_k\odot\tilde{\boldsymbol{s}}], \tag{35}$$

where $\odot$ denotes element-by-element multiplication of two vectors. Then $\boldsymbol{P}_v\hat{\boldsymbol{h}}_k$ can be computed as

$$\boldsymbol{P}_v\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_k + \mathrm{IDFT}\left[\tilde{\boldsymbol{s}}^*\odot\tilde{\boldsymbol{r}}_k\oslash\left(|\tilde{\boldsymbol{s}}|^2+\frac{1}{\lambda}1\right)\right], \tag{36}$$

where $1=[1,\ldots,1]^T$ is an $N\times 1$ vector with all elements being one. The proof procedure of Eq. (36) is given in the

Appendix. Equation (36) means that only FFT operations are required and it does not involve any matrix multiplication or matrix inverse operations.

It is clear that projecting $\hat{\boldsymbol{h}}_k$ back onto the feasible set $C_s \cap C_v \cap C_a$ can be exactly achieved by *repeatedly* applying the three projection operators: $\hat{\boldsymbol{h}}_k \leftarrow \boldsymbol{P}_v \boldsymbol{P}_a \boldsymbol{P}_s \hat{\boldsymbol{h}}_k$. Generally speaking, about ten iterations (like in the POCS method) are needed to guarantee $\hat{\boldsymbol{h}}_k$ in the feasible region. To decrease the computational amount, one can apply the three projection operators *once* as an approximation.

C. Summarization of the proposed algorithm

The detailed steps for implementing the FESCCO algorithm are summarized as follows.

The following is FESCCO.

- Input: the transmitted signal $\boldsymbol{s}$, the received signal $\boldsymbol{x}$, the estimate of noise variance σ_v^2, and the confidence factor γ.
- Output: multipath channel response estimate $\hat{\boldsymbol{h}}$.
- Algorithm:

(a) Initialization:

(1) Take the output of the matched filter as the initial value $\hat{\boldsymbol{h}}_0$.

(2) Determine the support set E from the output of the matched filter.

(b) For $k=1,2,\ldots$:

(1) Update $\hat{\boldsymbol{h}}_{k-1}$ along the steep descent direction using a variable step size strategy

$$\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_{k-1} - \mu_k \operatorname{sgn}(\hat{\boldsymbol{h}}_{k-1}).$$

(2) Project $\hat{\boldsymbol{h}}_k$ back onto the feasible set $C_s \cap C_v \cap C_a$, which can be approximately achieved by performing the three projection operators once:

$$\hat{\boldsymbol{h}}_k \leftarrow \boldsymbol{P}_v \boldsymbol{P}_a \boldsymbol{P}_s \hat{\boldsymbol{h}}_k,$$

where the three projection operators $\boldsymbol{P}_s$, $\boldsymbol{P}_a$, and $\boldsymbol{P}_v$ are given by Eqs. (26), (27), and (36), respectively.

If one of the termination conditions $\|\hat{\boldsymbol{h}}_k - \hat{\boldsymbol{h}}_{k-1}\| / \|\hat{\boldsymbol{h}}_k\| \le \delta$ or $k \ge k_m$ is satisfied, the iterative process can be stopped. Here k_m is the maximum iteration number and δ is a small admissible tolerance.

The FESCCO algorithm needs to compute the output of matched filter as an initial value. Note that matched filtering is equivalent to convolution operation and it can be implemented by FFT with a low complexity. The leading computational amount of the FESCCO method is computing the projection onto C_v, which can be efficiently performed by FFT. Therefore the computational complexity per iteration of the proposed method is $\mathcal{O}(N \log_2(N))$.

Remarks. We find that the FESCCO method has a fast convergence rate through the numerical simulations. It always converges in ten (at most several tens) iterations.

D. Interpolation for fractional time-delay estimation

If the time-delays are integer multiples of the sampling period T_s, the location of nonzero elements of $\boldsymbol{h}$ indicates these time-delays. However, $\boldsymbol{h}$ will be smeared if there exist fractional time-delays. The fractional time-delays can be estimated by interpolation of the discrete estimate $\hat{\boldsymbol{h}}$. According to the sampling theorem,[31] $h(t)$ can be reconstructed from its discrete sampling

$$h(t) = \sum_{n=-\infty}^{\infty} h(nT_s)\phi\left(\frac{t}{T_s} - n\right), \tag{37}$$

where $\phi(t) = \sin(\pi t)/\pi t$ is the sinc function. It has been proven that Eq. (37) can be simplified to[18]

$$\hat{h}(t) = \sum_{n=0}^{N-1} \hat{\boldsymbol{h}}(n+1)\phi_N(t-n), \tag{38}$$

where

$$\phi_N(t) = \begin{cases} \dfrac{\sin(\pi t)}{N \sin(\pi t/N)} & \text{for odd } N \\ \dfrac{\sin(\pi t)}{N \sin(\pi t/N)} \cos\left(\dfrac{\pi t}{N}\right) & \text{for even } N \end{cases} \tag{39}$$

is the digital-sinc function.

V. SIMULATION RESULTS

In this section, we conduct several numerical simulations to demonstrate the performance of the proposed channel deconvolution algorithm. In all of the simulations, we adopt a confidence interval of $P_d=0.95$ and the resulting value of γ is 1.96.

A. Simulation 1

In the first simulation, we use a typical example adopted in many references.[18,19] The transmitted signal $s(n)$ is a linear frequency modulated (LFM) signal with a discrete time-bandwidth product equal to 450, which can be written as

$$s(n) = \cos(2\pi(\alpha n^2 + \beta n)), \quad n = 0, 1, \ldots, N_s - 1. \tag{40}$$

The length of $s(n)$ is $N_s=450$. The parameters are set as $\alpha = 1/4\pi(N_s-1)$ and $\beta=0.1$ such that the normalized bandwidth of $s(n)$ is equal to 1.

Three-component multipath channel with time-delays $\{\tau_1, \tau_2, \tau_3\}$ and attenuation factors $\{a_1, a_2, a_3\}$ is used as the channel model. The attenuation parameters are fixed as $a_1 = 1$, $a_2=0.9$, and $a_3=0.5$ in this simulation. We simulate four channels with different time-delay parameters, named channel I, channel II, channel III, and channel IV. The time-delays of the four channels are listed in Table I. The minimum multipath time-delay difference, which is denoted as $\Delta\tau = \tau_2 - \tau_1$, is varied to test the time-delay resolution of different algorithms. We adopt $\Delta\tau = 2T_s$, $4T_s$, $5T_s$, and $10T_s$ for the four channels. The SNR is defined as

TABLE I. Time-delays (in samples) of the four channels.

	τ_1	τ_2	τ_3
Channel I	10	12	50
Channel II	10	14	50
Channel III	10	15	50
Channel IV	10	20	50

$$\mathrm{SNR} = 10 \log_{10}\left(\frac{P_s}{\sigma^2}\right), \tag{41}$$

where P_s is the variance of the transmitted signal $s(t)$. We set SNR=15 dB for all of the four channels.

We compare the estimation results of the FESCCO method with those obtained by matched filtering, matching pursuit algorithm,[21] MUSIC based time-delay estimation method,[9] constraint least-squares,[18] and POCS method.[19] The support indicator set E can be determined from the output of the matched filter. We set $E=\{0:20\}\cup\{40:60\}$ for channels I, II, and III, and $E=\{5:25\}\cup\{40:60\}$ for channel IV. Figures 1–4 illustrate the estimates of channels I, II, III, and IV, respectively. The resolution of the matched filter is the worst: It cannot distinguish the two closely spaced paths when $\Delta\tau = 2T_s$, $4T_s$, and $5T_s$. The matching pursuit algorithm delivers better performance than the matched filter in terms of resolution. But it still cannot identify the two close paths correctly when $\Delta\tau=2T_s$ and $4T_s$: Some wrong peaks, which indicate incorrect multipath components, appear. The matching pursuit algorithm is greedy. Hence there is no guarantee for matching pursuit algorithm finding the global optimal solution, which leads to that the performance breaks down when the multipath components are closely spaced. It can be seen that the performance of the MUSIC based time-delay estimation method also degrades when $\Delta\tau$ is small. Moreover, the MUSIC based method cannot provide the estimates of the attenuation factors. The POCS, constraint least-squares, and the FESCCO method give the correct estimation results for all of the four channels. However, the estimates of the POCS and constraint least-squares are spinous. A prickly estimate result means the sensitivity to noise. One needs to determine a threshold to distinguish the pseudopeaks from the peaks that indicate the true multipath components. Nevertheless, it is difficult to determine such a threshold, especially in the case of low SNR. Since the FESCCO method takes the sparsity of the channel into account, it gives smooth

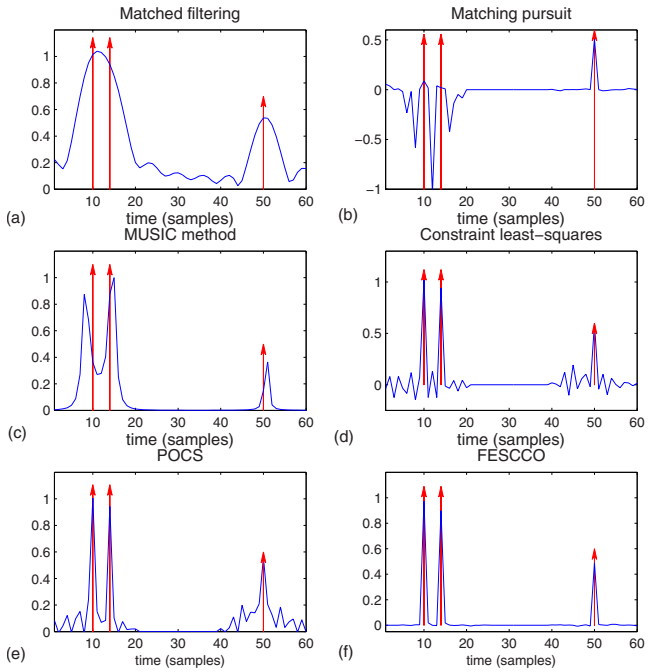


FIG. 2. (Color online) Estimation results of channel II.

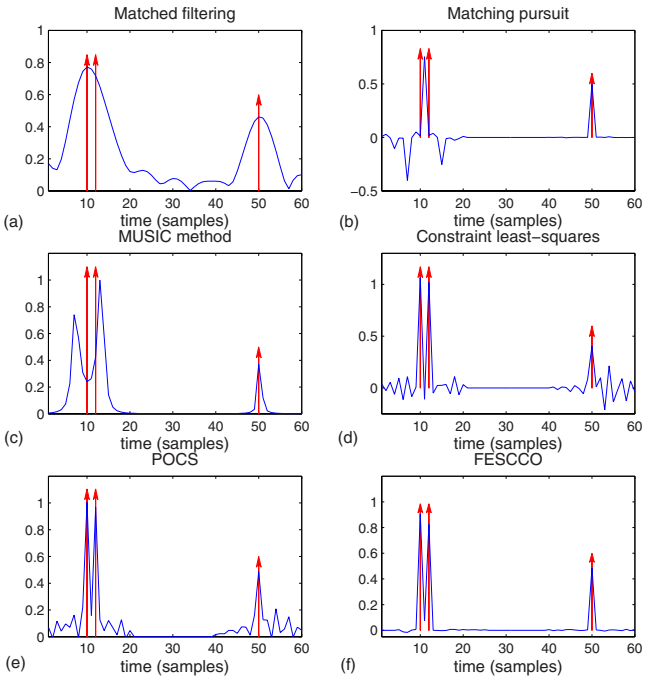


FIG. 1. (Color online) Estimation results of channel I.

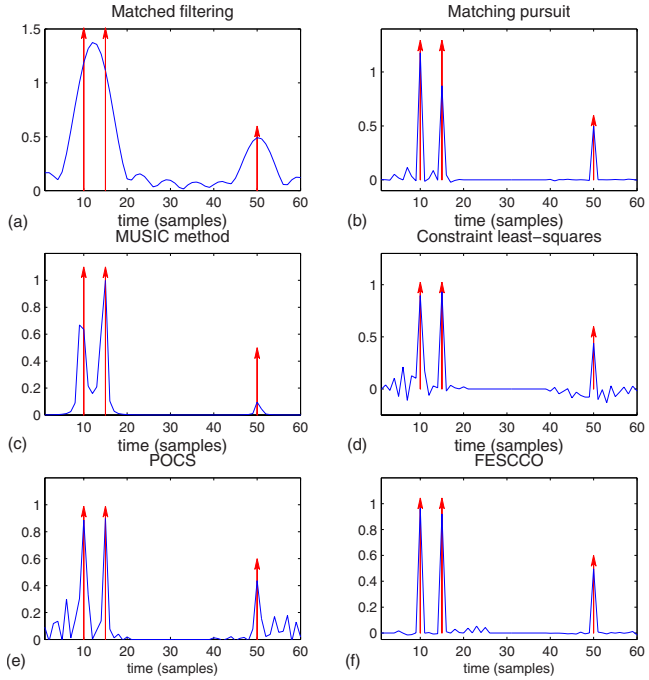


FIG. 3. (Color online) Estimation results of channel III.

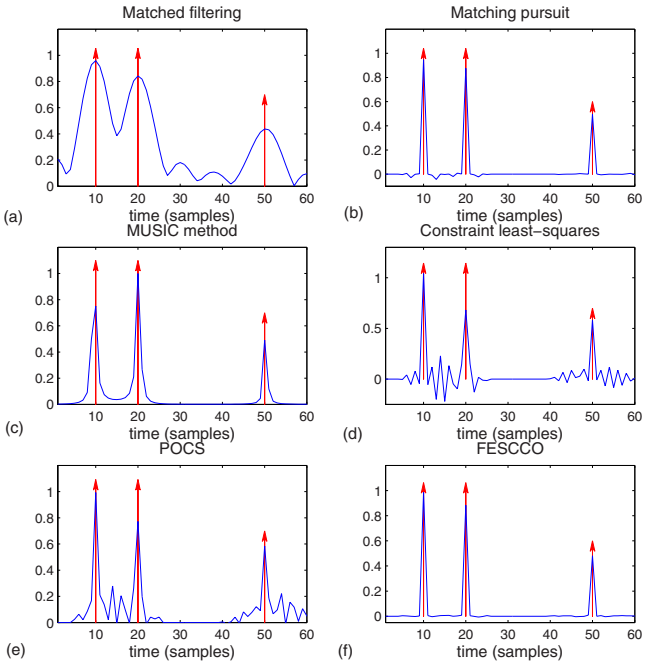


FIG. 4. (Color online) Estimation results of channel IV.

estimate results and the incorrect peaks will not appear. Therefore the FESCCO method is more robust to noise.

B. Simulation 2

In the second simulation example, we use the same LFM signal shown in Eq. (40) but the multipath time-delays are non-integer multiples of T_s. The fractional time-delays of three paths are $\tau_1=10.5T_s$, $\tau_2=12.4T_s$, and $\tau_3=50.6T_s$ with the corresponding attenuation factors being $a_1=1$, $a_2=0.9$, and $a_3=0.5$. The SNR is 15 dB. First we use matching pursuit, constraint least-squares, POCS, and the FESCCO method to obtain the discrete estimate $\hat{\boldsymbol{h}}$. Then Eq. (38) is used to get the interpolated estimate $\hat{h}(t)$. Figure 5 shows the interpolated estimation results of the four methods. The matching pursuit algorithm gives the incorrect result again. The three dominant peaks of the interpolated curves obtained by constraint least-squares, POCS, and the FESCCO method can give good estimations of the fractional time-delays. Compared to POCS and constraint least-squares, the main advantage of the FESCCO method is that it gives more smooth result and can void pseudopeaks.

C. Simulation 3

The third simulation is designed to evaluate the robustness to noise of the FESCCO algorithm. The transmitted signal is the same as the one used in simulation 1. The time-delays of three paths are fixed as $\tau_1=10T_s$, $\tau_2=12T_s$ and $\tau_3=50T_s$ with the corresponding attenuation factors being $a_1=1$, $a_2=0.9$, and $a_3=0.5$. We perform M_c Monte Carlo trials to evaluate the statistical performances of the proposed FESCCO algorithm, the constraint least-squares, and the POCS method. We use two performance indices. One index is the root mean squared error (RMSE) of the estimated channel response

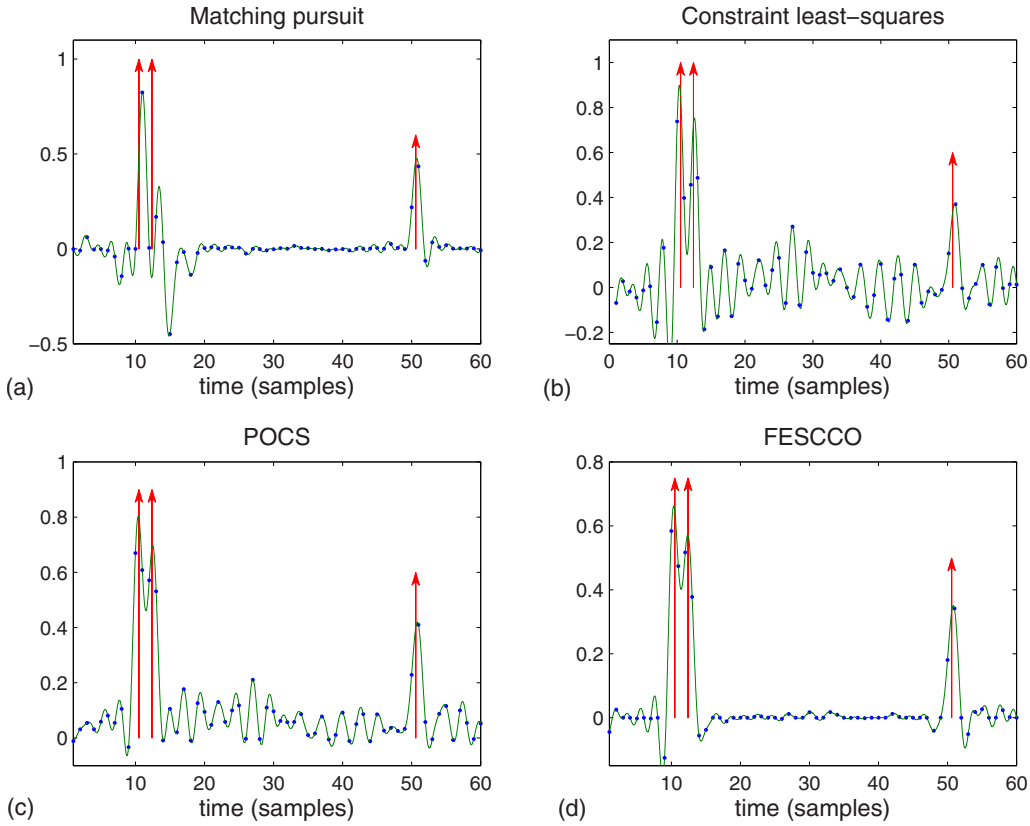


FIG. 5. (Color online) Interpolation results. The vertical lines with arrows show the true time-delays. The dots are the estimates of $\hat{\boldsymbol{h}}$ and the solid lines plot the interpolated $\hat{h}(t)$.

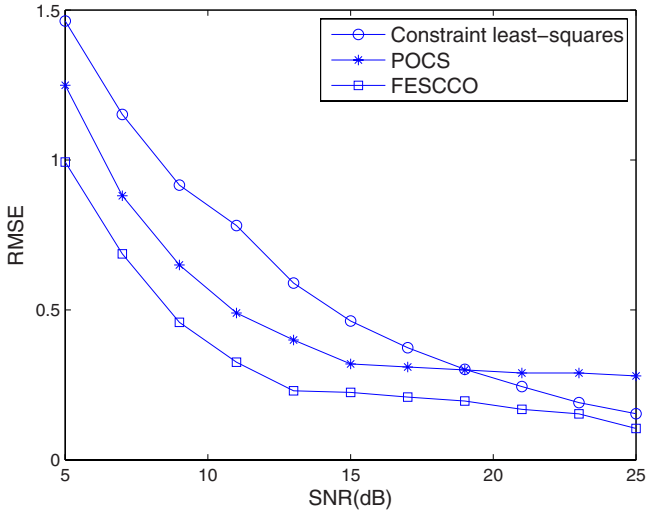


FIG. 6. (Color online) RMSE of the channel estimates versus SNR.

$$\text{RMSE} = \frac{1}{M_c}\sum_{m=1}^{M_c} \|\hat{\boldsymbol{h}}^{[m]} - \boldsymbol{h}\|,$$

where $\hat{\boldsymbol{h}}^{[m]}$ is the estimated result of the mth trial. The other performance index is the mean absolute error (MAE), which is defined as

$$\text{MAE} = \frac{1}{M_c}\sum_{m=1}^{M_c} \|\hat{\boldsymbol{h}}^{[m]} - \boldsymbol{h}\|_1 = \frac{1}{M_c}\sum_{m=1}^{M_c}\sum_{i=1}^{N} |\hat{\boldsymbol{h}}^{[m]}(i) - \boldsymbol{h}(i)|.$$

Since $\|\hat{\boldsymbol{h}}^{[m]} - \boldsymbol{h}\|_1$ is the ℓ_1-norm of the error vector $\hat{\boldsymbol{h}}^{[m]} - \boldsymbol{h}$, the MAE is an appropriate performance index when the channel is sparse. The SNR is varied from 5 to 25 dB. At each SNR, $M_c = 200$ Monte Carlo trials are performed.

Figure 6 illustrates the RMSE while Fig. 7 shows the MAE of the estimated results of the three methods, respectively. As one can see, the channel estimations are much more accurate by using the FESCCO algorithm rather than the constraint least-squares and the POCS method at various SNRs, which substantiates that the proposed FESCCO algorithm is more robust to noise.

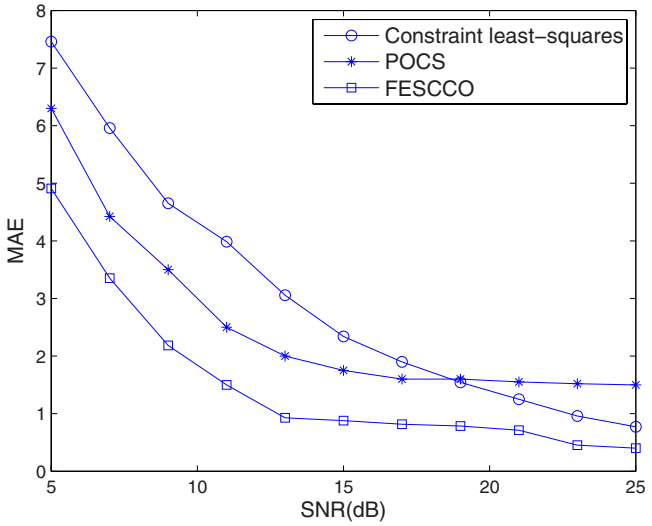


FIG. 7. (Color online) MAE of the channel estimates versus SNR.

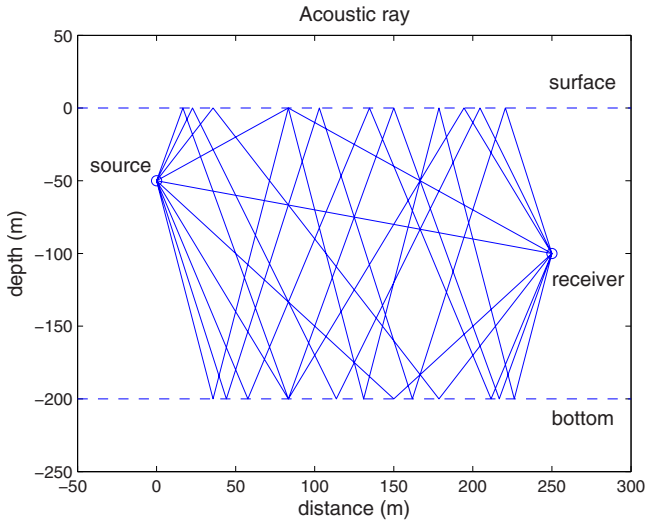


FIG. 8. (Color online) Acoustic ray propagation model for a shallow water environment.

D. Simulation 4

The fourth numerical example simulates a shallow water environment, where the continuous waveform (CW) pulse is used as the probing signal. The expression of CW pulse with frequency f_0 and pulse duration T_m can be written as

$$s(t) = \begin{cases} \cos(2\pi f_0 t) & \text{if } 0 \le t \le T_m \\ 0 & \text{otherwise.} \end{cases} \tag{42}$$

We set $f_0 = 100$ Hz, $T_m = 0.1$ s, and the sampling rate $f_s = 8f_0 = 800$ Hz.

The acoustic source is positioned at 50 m depth, and the receiver is located at 100 m depth with 250 m range of the source, as illustrated in Fig. 8. The depth of the shallow ocean is 200 m and the bottom is assumed to be perfectly rigid. The acoustic velocity is a constant: $c = 1500$ m/s. Based on the acoustic ray propagation model,[32] the multiple rays between the source and the receiver are plotted, as shown in Fig. 8.

The first ten dominant rays are taken into account and the attenuation factor of the direct ray is normalized to 1. The length of the received signal is $N = 1024$; i.e., the corresponding observed time is $T_s N = 1.28$ s. The waveforms of the transmitted and received signals are shown in Fig. 9 at SNR = 15 dB. The received signal is severely corrupted due to superposition of multiple echoes. From the output of the matched filter, the support indicator set E is adopted as $E = \{60:650\}$ (corresponding to the time duration of [0.075, 0.8125] s).

Figure 10 displays the channel estimation results of matched filtering, matching pursuit, POCS, constraint least-squares, and the FESCCO method. Clearly, only the FESCCO method gives a reasonable and fine resolution estimate result. The performances of the POCS and constraint least-squares methods are not satisfactory. Owing to the large time-delay spread of the channel, the support of the output of matched filter spreads a wide time range. Therefore the matched filter cannot provide a small enough support indicator set. The successes of POCS and constraint least-squares

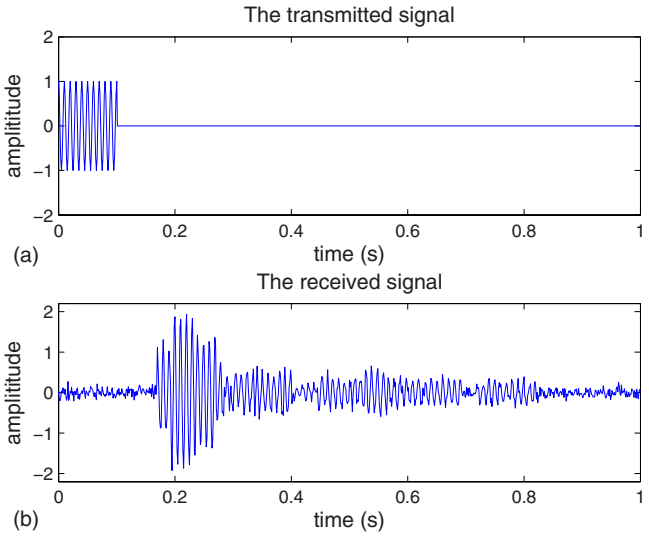


FIG. 9. (Color online) Waveforms of the transmitted and received signals.

are dependent on the size of the support indicator set. When the size of the support indicator set is not small enough, both POCS and constraint least-squares do not work well. Since the FESCCO method harnesses the sparsity of the channel, it still delivers good performance when the support indicator set cannot be effectively determined.

VI. CONCLUSION

A method for deconvolution of underwater acoustic channel response was developed. Not only the sparse structure of underwater acoustic channel but also the available *a priori* information are harnessed. A low computational complexity algorithm is designed to solve the resultant convex optimization problem. One interesting advantage of the FESCCO method is that it does not require the number of paths to be known. Besides, the FESCCO method obtains a smooth estimate result and the pseudopeaks, which indicate the incorrect paths will not appear. When the channel has severe multipath components and a long impulse response, the POCS and constraint least-squares method fail since the matched filter cannot offer sufficient support information, but the FESCCO method still works well. Therefore the FESCCO method is especially suitable to estimate the underwater acoustic channel with a large time-delay spread in a severe multipath environment. Moreover, simulation results demonstrate that the FESCCO method possesses fine temporal resolution and is robust to noise.

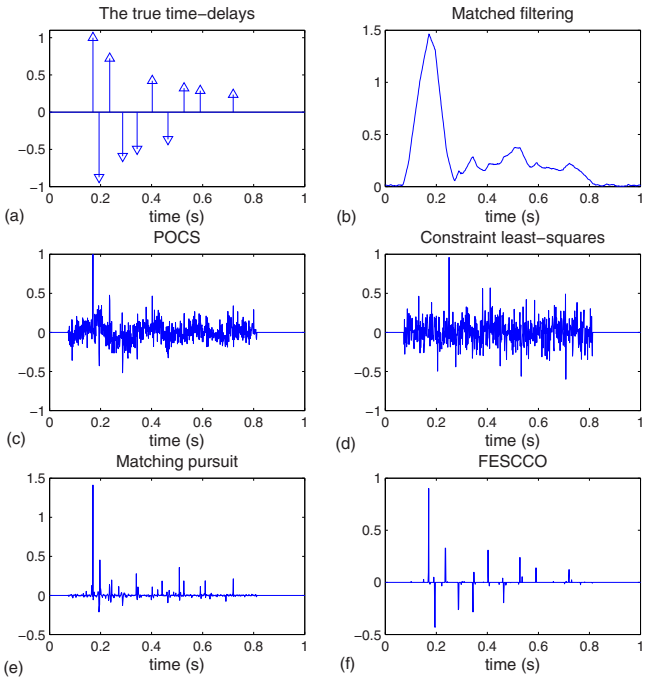


FIG. 10. (Color online) Estimation results of the simulated shallow water acoustic channel.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant No. 60975041 and 60672046, and in part by Basic Research Foundation of Tsinghua National Laboratory for Information Science and Technology (TNList). The authors would like to thank Professor En Cheng of Xiamen University, and Professor Sheng-Chun Piao and Dr. Yi-Wang Huang of Harbin Engineering University for their helpful discussion and advice.

APPENDIX: DERIVATIONS

In this appendix, we will proof the fast algorithm for computing the projection onto C_v, which is given by Eq. (36). According to Eqs. (30) and (32), we obtain

$$\boldsymbol{P}_v\hat{\boldsymbol{h}}_k=\hat{\boldsymbol{h}}_k+\left(\boldsymbol{S}^T\boldsymbol{S}+\frac{1}{\lambda}\boldsymbol{I}\right)^{-1}\boldsymbol{S}^T\boldsymbol{r}_k. \tag{A1}$$

By exploiting $\boldsymbol{S}=\boldsymbol{F}^H\boldsymbol{\Lambda}_s\boldsymbol{F}$ and $\boldsymbol{S}^T=\boldsymbol{F}^H\boldsymbol{\Lambda}_s^*\boldsymbol{F}$, we get

$$\left(\boldsymbol{S}^T\boldsymbol{S}+\frac{1}{\lambda}\boldsymbol{I}\right)^{-1}=\boldsymbol{F}^H\left(|\boldsymbol{\Lambda}_s|^2+\frac{1}{\lambda}\boldsymbol{I}\right)^{-1}\boldsymbol{F} \tag{A2}$$

and

$$\boldsymbol{S}^T\boldsymbol{r}_k=\boldsymbol{F}^H\boldsymbol{\Lambda}_s^*\boldsymbol{F}\boldsymbol{r}_k=\boldsymbol{F}^H\boldsymbol{\Lambda}_s^*\tilde{\boldsymbol{r}}_k=\boldsymbol{F}^H\tilde{\boldsymbol{s}}^*\odot\tilde{\boldsymbol{r}}_k. \tag{A3}$$

Using $\boldsymbol{F}\boldsymbol{F}^H=\boldsymbol{I}$, we obtain

$$\begin{aligned}\left(\boldsymbol{S}^T\boldsymbol{S}+\frac{1}{\lambda}\boldsymbol{I}\right)^{-1}\boldsymbol{S}^T\boldsymbol{r}_k&=\boldsymbol{F}^H\left(|\boldsymbol{\Lambda}_s|^2+\frac{1}{\lambda}\boldsymbol{I}\right)^{-1}\tilde{\boldsymbol{s}}^*\odot\tilde{\boldsymbol{r}}_k\\&=\boldsymbol{F}^H\left(\tilde{\boldsymbol{s}}^*\odot\tilde{\boldsymbol{r}}_k\oslash\left(|\tilde{\boldsymbol{s}}|^2+\frac{1}{\lambda}1\right)\right)\\&=\mathrm{IDFT}\left[\tilde{\boldsymbol{s}}^*\odot\tilde{\boldsymbol{r}}_k\oslash\left(|\tilde{\boldsymbol{s}}|^2+\frac{1}{\lambda}1\right)\right],\end{aligned} \tag{A4}$$

which means that

$$\boldsymbol{P}_v\hat{\boldsymbol{h}}_k=\hat{\boldsymbol{h}}_k+\mathrm{IDFT}\left[\tilde{\boldsymbol{s}}^*\odot\tilde{\boldsymbol{r}}_k\oslash\left(|\tilde{\boldsymbol{s}}|^2+\frac{1}{\lambda}1\right)\right].$$

[1]W. Munk and C. Wunsch, "Ocean acoustic tomography: A scheme for large scale monitoring," Deep-Sea Res. **26**, 123–161 (1979).
[2]O. Rabaste and T. Chonavel, "Estimation of multipath channels with long impulse response at low SNR via an MCMC method," IEEE Trans. Signal

Process. **55**, 1312–1325 (2007).
[3]M. Stojanovic, J. Catipovic, and J. G. Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," J. Acoust. Soc. Am. **94**, 1621–1631 (1993).
[4]Y. X. Yuan, "Passive localization of an underwater acoustic source using directional sensors," Ph.D. thesis, University of Saskatchewan, Saskatoon, Canada (1994).
[5]Y. X. Yuan and J. E. Salt, "Range and depth estimation using a vertical array in a correlated multipath environment," IEEE J. Ocean. Eng. **18**, 500–507 (1993).
[6]H. S. Anderson and M. R. Gupta, "Joint deconvolution and classification with applications to passive acoustic underwater multipath," J. Acoust. Soc. Am. **124**, 2973–2983 (2008).
[7]M. J. Roan, M. R. Gramann, J. G. Erling, and L. H. Sibul, "Blind deconvolution applied to acoustical systems identification with supporting experimental results," J. Acoust. Soc. Am. **114**, 1988–1996 (2003).
[8]J. E. Ehrenberg, T. E. Ewart, and R. D. Morris, "Signal processing techniques for resolving individual pulses in a multipath signal," J. Acoust. Soc. Am. **63**, 1861–1865 (1978).
[9]M. A. Hasan, M. R. Azimi-Sadjadi, and G. J. Dobeck, "Separation of multiple time delays using new spectral estimation schemes," IEEE Trans. Signal Process. **46**, 1580–1590 (1998).
[10]M. Pallas and G. Jourdain, "Active high resolution time delay estimation for large BT signals," IEEE Trans. Signal Process. **39**, 781–788 (1991).
[11]K. Pahlavan, X. Li, and J.-P. Makela, "Indoor geolocation science and technology," IEEE Commun. Mag. **40**, 112–118 (2002).
[12]F.-X. Ge, D. Shen, Y. Peng, and V. Li, "Super-resolution time delay estimation in multipath environments," IEEE Trans. Circuits Syst. I: Regular Papers **54**, 1977–1986 (2007).
[13]R. O. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas Propag. **34**, 276–280 (1986).
[14]R. J. Vaccaro, C. S. Ramalingam, D. W. Tufts, and R. L. Field, "Least-squares time-delay estimation for transient signals in a multipath environment," J. Acoust. Soc. Am. **92**, 210–218 (1992).
[15]T. G. Manickam, R. J. Vaccaro, and D. W. Tufts, "A least-squares algorithm for time-delay estimation," IEEE Trans. Signal Process., **42**, 3229–3233 (1994).
[16]J. Li and R. Wu, "An efficient algorithm for time delay estimation," IEEE Trans. Signal Process. **46**, 2231–2235 (1998).
[17]A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM, Philadelphia, PA, 2005).
[18]T. G. Manickam and R. J. Vaccaro, "A non-iterative deconvolution method for estimating multipath channel responses," Proceedings of the ICASSP (1993), pp. 333–336.
[19]Z. Kostic, M. I. Sezan, and E. L. Titlebaum, "Estimation of the parameters of a multipath channel using set-theoretic deconvolution," IEEE Trans. Commun. **40**, 1006–1011 (1992).
[20]S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," IEEE Trans. Signal Process. **41**, 3397–3415 (1993).
[21]S. Cotter and B. Rao, "Sparse channel estimation via matching pursuit with application to equalization," IEEE Trans. Commun. **50**, 374–377 (2002).
[22]W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," IEEE J. Ocean. Eng. **32**, 927–939 (2007).
[23]S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM Rev. **43**, 129–159 (2001).
[24]J.-J. Fuchs, "Multipath time-delay detection and estimation," IEEE Trans. Signal Process. **47**, 237–243 (1999).
[25]J.-J. Fuchs, "Minimal L_1-norm reconstruction function for oversampled signals: Applications to time-delay estimation," IEEE Trans. Inf. Theory **46**, 1666–1673 (2000).
[26]M. Robert, "Gray, Toeplitz and circulant matrices: A review," Foundations and Trends in Communications and Information Theory **2**, 155–239 (2006).
[27]H. J. Trussell and R. Civanlar, "The feasible solution in signal restoration," IEEE Trans. Acoust., Speech, Signal Process. **32**, 201–212 (1984).
[28]R. Tibshirani, "Regression shrinkage and selection via the LASSO," J. R. Stat. Soc. Ser. B (Methodol.) **58**, 267–288 (1996).
[29]S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2006).
[30]S. J. Wright, *Primal-Dual Interior Point Methods* (SIAM, Philadelphia, PA, 2005).
[31]A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing* (Prentice-Hall, New York, 1999).
[32]L. M. Brehovsikh and Y. P. Lysanov, *Fundamentals of Ocean Acoustics* (Springer-Verlag, Berlin, 2003).

A model-based approach to dispersion and parameter estimation for ultrasonic guided waves

James S. Hall and Jennifer E. Michaels[a)]
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0250

(Received 2 May 2009; revised 20 August 2009; accepted 20 November 2009)

A model-based algorithm is presented that adaptively estimates *in situ* ultrasonic guided wave system parameters. Dispersion curves, propagation loss, transducer distances, transmitted signal, and mode weighting coefficients are estimated using minimal *a priori* information and assumptions. The five-part algorithm is scalable to accommodate two or more receivers and one or more propagating modes, provided that mode separation can be achieved prior to use of the algorithm. Algorithmic performance is demonstrated on signals obtained both from theoretical dispersion curves and finite element modeling. Quantitative performance curves are presented that are based on algorithmic performance from multiple simulated test cases with varying amounts of additive noise. Results show excellent agreement between estimated and actual parameters, as well as between modeled and received signals. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3273894]

PACS number(s): 43.60.Uv, 43.35.Cg, 43.60.Pt, 43.60.Jn [PEB] Pages: 920–930

I. INTRODUCTION

Most ultrasonic systems employing guided waves for flaw detection require knowledge of the operating environment, which may include dispersion curves, transducer locations, and the transmitted signal. Errors in this *a priori* information typically degrade performance and limit the ability of the system to detect and characterize defects.[1] For example, nominal dispersion curves are typically computed by numerically solving the Rayleigh–Lamb equations. Errors in bulk-wave speeds and the plate thickness are compounded with numerical errors in the final dispersion computation. Variations in temperature also change wave speeds and may also perturb the spatial relationship of transducers, introducing additional errors. By characterizing the entire propagating environment at the time of test and foregoing potentially erroneous *a priori* estimates, performance of ultrasonic guided wave systems can be improved.[2–4]

The motivation for this research is *in situ* characterization of the source, receivers, and propagating environment as applied to a spatially distributed array of guided wave sensors such as those that might be appropriate for nondestructive evaluation and structural health monitoring applications.[4–6] A model-based approach is proposed for which a propagation model is assumed and model parameters (distances, dispersion curves, etc.) are obtained, allowing the model to describe the received signals as closely as possible. Such an approach takes advantage of the inherent constraints of the system model in an effort to simplify the parameter search.

Various aspects of this problem have been considered previously for both bulk waves and guided waves. In 1977, Sachse and Pao[7] proposed a phase spectral analysis method using received signals from two different distances from the transmitter to extract phase velocity of longitudinal waves. The change in phase of the received signals was analyzed, assuming a known transmitted signal and propagation distance. Their work was extended by Peters and Petit[8] to avoid the use of a reference medium, and by Hutchins *et al.*[9] to apply the technique to Lamb waves with an excitation spectrum that does not extend to 0 Hz. Hutchins *et al.*[9] used a model-based approach to separate the problems of determining plate thickness and dispersion relations by performing a nonlinear least-squares fit to an assumed propagation model. Prosser and Gorman[10] used the phase spectral analysis method with Lamb waves to underscore the limitations of classical plate theory for composite materials by showing experimental data that agreed much more closely with Mindlin plate theory.[11] Schumacher *et al.*[12] further demonstrated the capabilities of the phase spectral method to estimate phase velocities up to 4 MHz-mm for laser-generated Lamb waves.

Time-frequency representations (TFRs) have also been applied to estimate frequency-dependent group velocities for Lamb waves.[13–15] One benefit that is unique to TFR analysis is the ability to extract both group velocity and attenuation information from a single multi-mode signal. However, without determination of an unknown integration constant, group velocity cannot be translated to phase velocity or a wavenumber versus frequency dispersion curve.

As yet another example of dispersion estimation, Alleyne and Cawley[16] successfully demonstrated a two-dimensional (2D) Fourier transform method that is capable of accurately estimating dispersion relations from a set of multi-mode signals. Costley and Berthelot[17] and Eisenhardt *et al.*[18] further validated the method using laser-generated Lamb waves to estimate dispersion curves up to several megahertz. One of the few limitations associated with the two-dimensional Fourier transform method is the requirement for linear spatial sampling at sufficiently close intervals to avoid spatial aliasing.

[a)]Author to whom correspondence should be addressed. Electronic mail: jennifer.michaels@ece.gatech.edu

In all of the above approaches, the propagation distances are assumed known. In contrast, He and Zheng[19] found that both the phase velocity and the sample thickness can be determined for normal incident longitudinal waves from the phase spectra of four recorded pulses. Their approach took advantage of the different pulse echo and through transmission propagation paths traveled by each recorded signal.

Holland *et al.*[20] demonstrated that the propagation distance can be estimated from a single received guided wave signal via analysis of its TFR. Their method required a search for the distance that provided the best match between theoretical dispersion curves and features of the TFR. The matching procedure was based on a visual comparison and was not automated.

The algorithm presented in this paper builds upon the work of Sachse and Pao[7] in that the difference in phase between received signals is used to obtain the dispersion relationships. By placing bounds on the problem through an assumed model, however, the model-based approach allows a significant amount of information to be extracted from a set of received signals. Specifically, the algorithm adaptively estimates dispersion curves, propagation loss, relative transducer distances, transmitted signal function, and mode weighting coefficients in a multi-mode propagation environment using a minimal number of ultrasonic transducers and as little *a priori* information as possible.

This paper is organized as follows. Section II presents the problem statement, which includes a description of the propagation model and a discussion of assumptions and limitations. Section III is a detailed description of the algorithm itself. Performance results are presented and discussed in Sec. IV, and concluding remarks are made in Sec. V.

II. PROBLEM STATEMENT

It is assumed that waveform data are obtained by generating an ultrasonic guided wave with one transducer and by recording the response with multiple receivers. The received signals are windowed in time to isolate the direct arrivals and to remove any reflections. Assuming a homogeneous plate with N identical receivers and a common transmitted signal, a frequency domain model of the recorded signal at the ith receiver, $R_{DAi}(\omega)$, is

$$R_{DAi}(\omega) = T(\omega)H_{DAi}(\omega) + \nu = T(\omega)\sum_{m=1}^{M} \alpha_m(\omega)e^{\gamma_m(\omega)d_i} + \nu. \quad (1)$$

Here M is the number of modes and $\gamma_m(\omega)=p(\omega)+jk_m(\omega)$. The function $T(\omega)$ incorporates the excitation signal, the transfer function relating the excitation signal to the guided wave, and the combined transfer function of the receiving transducer and its instrumentation. Since it is assumed that the transmitter is isotropic, the receivers are all identical, and the propagation medium is homogeneous, $T(\omega)$ is independent of both mode and receiver. The transfer function for the ith receiver $H_{DAi}(\omega)$ is composed of a mode and frequency-dependent weighting term $\alpha_m(\omega)$, an exponential propagation loss term due to both geometric spreading and attenuation $p(\omega)$, and a mode and frequency-dependent dispersion term $k_m(\omega)$. The additive noise term ν is a complex Gaussian random variable $\nu \sim N(0,\sigma_\nu^2)$ that represents the incoherent or electronic noise associated with the receiving transducer. Table I summarizes key assumptions and defines model parameters.

TABLE I. Model assumptions and description of model parameters.

Assumptions	Identical receivers
	Homogeneous and isotropic propagation medium
	Isotropic transmission
	Far-field propagation model [see Eq. (1)]
Known	$R_{DAi}(\omega)$: direct arrival received at ith transducer
	M: total number of propagating modes
	N: total number of receiving transducers
Noise	ν: Gaussian random variable (frequency independent variance)
Solve for	$T(\omega)$: combined isotropic transducer function
	$H_{DAi}(\omega)$: transfer function for ith receiver
	d_i[a]: distance from transmitter to ith receiver
	$\alpha_m(\omega)$: mth mode weighting coefficient $\Sigma_{m=1}^{M}\lvert\alpha_m(\omega)\rvert^2=1$
	$p(\omega)$[b,a]: mode independent propagation loss
	$k_m(\omega)$[b,a]: mth mode dispersion curve

[a]Can only be determined to within a common scaling factor.
[b]Always negative because of sign convention and causality.

A waterfall plot of a typical set of received signals is shown in Fig. 1. These signals were calculated using nominal S_0 and A_0 dispersion curves for a 3 mm thick aluminum plate at 25 °C. The transmitted signal is a three-cycle, Hanning-windowed sinusoid with a center frequency of 200 kHz. The weighting coefficient for the S_0 mode is 0.5, and that for the A_0 mode is 0.866. Receivers are located at 400, 500, 617, and 803 mm from the transmitter, and Gaussian noise was added to each signal.

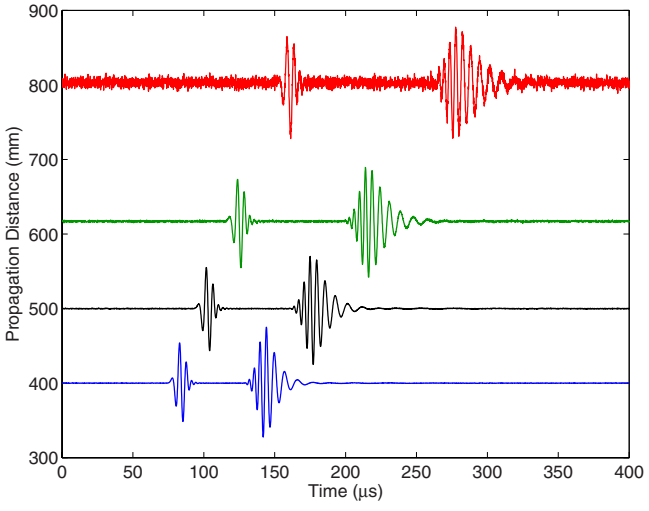


FIG. 1. (Color online) Simulated received signals corresponding to a single transmitter and multiple receivers at different distances. The test case shown consists of four receivers located at 400, 500, 617, and 803 mm, and each signal is normalized by its peak amplitude for display purposes. Note that the signal to noise ratio decreases as the propagation distance increases.

Although *a priori* information is kept to a minimum, it cannot be avoided entirely. Note that the exponential term $\gamma_m(\omega)d_i$ in Eq. (1) cannot be separated mathematically from $[C\gamma_m(\omega)][d_i/C]$, where C is an unknown constant. Thus, these model parameters can only be determined within a common scaling factor, which is resolved after-the-fact by using *a priori* information about d_i. Similarly, to remove the ambiguity in the product of $T(\omega)$ and $H_{DAi}(\omega)$, it is assumed that the squares of the $\alpha_m(\omega)$ sum to one at each frequency.

The signal-to-noise ratio (SNR) for all frequencies being considered is assumed to be sufficiently high so that signal power exceeds noise power. This requirement is a consequence of using both power ratios and phase-response information. Therefore, transmitted signals with power spectral densities that contain nulls extending below the receiver noise floor must be handled as two separate signals.

At this time, it is assumed that mode separation has been achieved and that single-mode sub-signals are available at each individual receiver. The model for each sub-signal (mode) is

$$R^m_{DAi}(\omega) = T(\omega)\alpha_m(\omega)e^{\gamma_m(\omega)d_i} + \nu \tag{2}$$

for $i=1,\ldots,N$ and $m=1,\ldots,M$. Although mode separation is not a trivial requirement, several methods do exist to separate the sub-signals. The simplest technique is to position transducers sufficiently far apart from one another to achieve mode separation in the time domain. For systems that cannot support such distances, another method is to use sensor pairs located at the same position but on opposite surfaces, which provides the ability to separate overlapping symmetric and antisymmetric modes. This requirement is not unreasonable when considering only the S_0 and A_0 modes, but may be impractical if additional modes are also present.

In summary, the specific problem considered here is to estimate transducer distances d_i and frequency-dependent functions $T(\omega)$, $\alpha_m(\omega)$, $p(\omega)$, and $k_m(\omega)$ given measured direct arrivals $R^m_{DAi}(\omega)$ for multiple modes $(m=1,\ldots,M)$ and at multiple receivers $(i=1,\ldots,N)$.

III. ALGORITHM

The assumption of mode separation allows many of the model variables to be isolated and estimated independently. As such, the algorithm is split into five distinct stages: (1) mode weighting coefficient estimation, (2) relative distance and group velocity estimation, (3) common distance estimation, (4) constant offset estimation, and (5) magnitude response and propagation loss estimation. Each stage is described in detail in Secs. III A–III E, and data used for discussion purposes are that of Fig. 1.

Although the receiving transducers, and therefore their respective noise levels, are expected to be identical, signal power at each receiver is different because of propagation loss. Therefore, as illustrated in Fig. 1, SNRs are not identical between receivers. Additionally, when considering frequency-dependent parameters, it is important to note that the SNR varies over the bandwidth of the signal. Thus, frequency-dependent estimates tend to have higher errors at frequencies for which the signal has low signal power. For this reason, at various stages of the algorithm, weighted means are frequently used across receivers and/or frequencies to provide preference to higher SNR data. The weights are always a function of signal power.

Each calculation in the algorithm affects system noise. For many of the calculations, the noise remains a Gaussian white noise distribution that corresponds to the frequency- and receiver-dependent SNR. To avoid the confusion associated with introducing a new random variable at each step, the random variable ν is used to represent any noise that is consistent across the frequency spectrum, as in Eq. (1), and μ corresponds to a SNR dependent Gaussian distribution, such as that observed in phase-response measurements. It is important to note that even though the same noise variable appears in multiple equations, the variance changes at each stage and is not necessarily identical to the initial variance.

A. Mode weighting coefficient

Mode weighting coefficients reflect the frequency-dependent power distribution between propagating modes. Using Eq. (2), the magnitudes of the received mth-mode sub-signals are

$$\begin{bmatrix} |R^m_{DA1}(\omega)| \\ \vdots \\ |R^m_{DAN}(\omega)| \end{bmatrix} = \begin{bmatrix} |T(\omega)\alpha_m(\omega)e^{p(\omega)d_1+jk_m(w)d_1} + \nu| \\ \vdots \\ |T(\omega)\alpha_m(\omega)e^{p(\omega)d_N+jk_m(w)d_N} + \nu| \end{bmatrix}$$
$$= |T(\omega)||\alpha_m(\omega)| \begin{bmatrix} e^{p(\omega)d_1} \\ \vdots \\ e^{p(\omega)d_N} \end{bmatrix} + \begin{bmatrix} \nu \\ \vdots \\ \nu \end{bmatrix}. \tag{3}$$

The algebraic simplification is possible by noting that the complex random variable of Eq. (2), $\nu \sim N(0,\sigma^2_\nu)$, can be replaced by a real random variable that represents the noise of the magnitude response. At frequencies for which $|\text{noise}| < |\text{signal}|$, the magnitude of each $|R^m_{DAi}(\omega)|$ is guaranteed positive and ν of Eq. (3) is another Gaussian random variable $\nu \sim N(0,\sigma^2_\nu/2)$. If the SNR is not sufficiently high, ν is no longer a Gaussian random variable and further analysis becomes much more complex. Assuming that the noise term of Eq. (3) is Gaussian, each of the received signal magnitudes can be described by a Gaussian random variable with a nonzero mean:

$$\begin{bmatrix} |R^m_{DA1}(\omega)| \\ \vdots \\ |R^m_{DAN}(\omega)| \end{bmatrix} = \begin{bmatrix} \tilde{R}^m_1(\omega) \\ \vdots \\ \tilde{R}^m_N(\omega) \end{bmatrix}, \tag{4}$$

where $\tilde{R}^m_i(\omega) \sim N(|T(\omega)||\alpha_m(\omega)|e^{p(\omega)d_i}, \sigma^2_\nu/2)$.

Taking advantage of the fact that the propagation loss $p(\omega)$ and transmitted signal term $T(\omega)$ are common to all modes, weighting coefficient ratios between modes m_i and m_j are obtained for each receiver via a simple division operation:

$$\vec{A}_{m_i,m_j}(\omega) = \begin{bmatrix} |R^{m_i}_{DA1}(\omega)|/|R^{m_j}_{DA1}(\omega)| \\ \vdots \\ |R^{m_i}_{DAN}(\omega)|/|R^{m_j}_{DAN}(\omega)| \end{bmatrix} \approx \begin{bmatrix} |\alpha_{m_i}(\omega)|/|\alpha_{m_j}(\omega)| \\ \vdots \\ |\alpha_{m_i}(\omega)|/|\alpha_{m_j}(\omega)| \end{bmatrix}. \quad (5)$$

Since the propagating environment is homogeneous and isotropic and the excitation is assumed to be isotropic, weighting coefficient ratios should be identical at each receiver. Therefore, each element of the column vector in Eq. (5) can be combined to obtain a composite coefficient ratio:

$$\rho_{m_i,m_j}(\omega) = E_{\mathrm{SNR}}\{\vec{A}_{m_i,m_j}(\omega)\}, \quad (6)$$

where $E_{\mathrm{SNR}}\{\cdot\}$ is the expected value operator, incorporating total power from each receiver as a weighting basis.

Hinkley[21] showed that the ratio of uncorrelated, nonzero mean Gaussian random variables has a closed-form probability density function. This density function, however, can be heavy-tailed, causing a bias in the expected value. Intuitively, the heavy-tailed behavior is accentuated with small denominator values. Therefore, as confirmed by Hinkley, the ratio of Eq. (5) is more numerically stable if the mode with greater power is used as the denominator. Two methods of compensation for the heavy-tailed behavior were considered when combining ratio estimates across receiver pairs in Eq. (6): (1) use of the median value and (2) a weighted mean. While both methods showed improvement in SNR environments that produced a biased estimate, the weighted mean produced more consistent results.

With a single, data-driven, zero-bias approximation of each mode weighting coefficient ratio from Eq. (6), a least-squares approach at each frequency can be used to produce the final frequency-dependent estimate of $\alpha_m(\omega)$. It is assumed that the sum of the squares of the mode weighting coefficients is 1. Therefore, the governing equations are

$$\begin{aligned} |\alpha_1(\omega)|^2 + \cdots + |\alpha_M(\omega)|^2 &= 1, \\ |\alpha_2(\omega)|^2 \rho^2_{1,2}(\omega) &= |\alpha_1(\omega)|^2, \\ |\alpha_3(\omega)|^2 \rho^2_{2,3}(\omega) &= |\alpha_2(\omega)|^2, \\ &\vdots \\ |\alpha_M(\omega)|^2 \rho^2_{M-1,M}(\omega) &= |\alpha_{M-1}(\omega)|^2. \end{aligned} \quad (7)$$

These equations can be rewritten in matrix form as $A\vec{x}=\vec{b}$, where

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & -\rho^2_{1,2}(\omega) & 0 & \\ 0 & 1 & -\rho^2_{2,3}(\omega) & \\ \vdots & & & \ddots \end{bmatrix},$$

$$\vec{x} = \begin{bmatrix} |\alpha_1(\omega)|^2 \\ \vdots \\ |\alpha_M(\omega)|^2 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (8)$$

Since the columns of A are linearly independent, a unique, least-squares solution to Eq. (8) can be obtained for each frequency with sufficient signal power. Note that potential bias in the $\rho_{m_i,m_j}(\omega)$ terms may cause the least-squares estimate of $\vec{x}$ to also reflect this bias. The use of a conditioning pre-transformation or a total least-squares solution may yield improved results.

B. Relative distance/group velocity

Group velocity and receiver distance estimation is based on measured differences in the phase-response of each received signal. The phase-response of a single-mode subsignal is

$$\measuredangle R^m_{DAi}(\omega) = (\measuredangle T(\omega) + k_m(\omega)d_i + \mu) \bmod 2\pi \quad (9)$$

for $i=1,\ldots,N$ and $m=1,\ldots,M$.

The challenges associated with the modulo nature of the phase information can be addressed through the use of an unwrapped phase-response. An assumption must be made that the maximum angular change between two adjacent frequencies of the frequency response of each received signal is less than π. Ignoring the impact of noise, the angular change between two adjacent frequencies of a Fast Fourier Transform (FFT) is obtained from Eq. (9) as

$$\begin{aligned} \left(\frac{\partial}{\partial\omega}\measuredangle R^m_{DAi}(\omega)\right)\Delta\omega &= \left(\frac{\partial}{\partial\omega}\measuredangle T(\omega) + d_i\frac{\partial}{\partial\omega}k_m(\omega)\right)\Delta\omega \\ &= \left(\frac{\partial}{\partial\omega}\measuredangle T(\omega) + \frac{d_i}{v_{\mathrm{gr}}(\omega)}\right)\Delta\omega, \end{aligned} \quad (10)$$

where $\Delta\omega$ is the spacing between FFT frequency bins and $v_{\mathrm{gr}}(\omega)$ is the group velocity $[\partial\omega/\partial k_m(\omega)]$. The impact of noise on the phase-response is related to the SNR for that specific frequency. The maximum angular deviation in phase for a specific frequency due to noise with amplitude a_μ is

$$\Delta\theta_{im}(\omega) = \sin^{-1}\left(\frac{a_\mu}{|R^m_{DAi}(\omega)|}\right) = \sin^{-1}\left(\frac{1}{\mathrm{SNR}_{im}(\omega)}\right). \quad (11)$$

For the angular change between frequencies to be less than π for any ω, the angular change between two adjacent frequencies due to the signal, Eq. (10), and twice the maximum angular deviation that is possible from either bin due to noise, Eq. (11), must sum to less than π:

$$2\sin^{-1}\left(\frac{1}{\mathrm{SNR}_{\min}}\right) + \left|\frac{\partial}{\partial\omega}\measuredangle T(\omega) + \frac{d_{\max}}{v_{\min}}\right|\Delta\omega < \pi. \quad (12)$$

Here, the minimum SNR of any frequency $\mathrm{SNR}_{\min}$, the maximum distance between transmitter and any receiver $d_{\max}$, and the minimum group velocity for any propagating mode at any excited frequency $v_{\min}$ are used to preserve the inequality and extend the bound to any receiver, propagating mode, and excited frequency. Note that for many cases, including the case described in this paper, the derivative of the

transmitted phase-response in Eq. (12) can be safely ignored because of the magnitude difference between the two terms in the absolute value function. Substituting $\Delta\omega = 2\pi F_s/n$, where F_s is the sampling frequency and n is the number of samples, Eq. (12) can be rearranged as

$$n > \frac{2\pi F_s d_{\max}}{v_{\min}\left[\pi - 2\sin^{-1}\left(\frac{1}{\mathrm{SNR}_{\min}}\right)\right]}. \tag{13}$$

The number of samples can be increased to n or greater by padding the direct arrival signals with zeros and, thus, does not translate to any operational system requirements.

By unwrapping the phase-response data and placing it into a single mode-specific vector, the structure of the equations can be better understood:

$$\begin{bmatrix} \measuredangle R^m_{DA1}(\omega) \\ \vdots \\ \measuredangle R^m_{DAN}(\omega) \end{bmatrix} = \begin{bmatrix} \measuredangle T(\omega) + k_m(\omega)d_1 + \mu - 2\pi b_1 \\ \vdots \\ \measuredangle T(\omega) + k_m(\omega)d_N + \mu - 2\pi b_N \end{bmatrix}$$

$$\measuredangle \vec{R}^m(\omega) = \measuredangle T(\omega)\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + k_m(\omega)\begin{bmatrix} d_1 \\ \vdots \\ d_N \end{bmatrix} - 2\pi\begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix} + \begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix}. \tag{14}$$

Since the phase-responses have been unwrapped, the b_i represent unknown integers that are consistent across the entire frequency spectrum. Although each b_i is unknown, it represents the appropriate 2π offset to account for the inherent modulo operation. By subtracting the vector elements in Eq. (14) from one another using a projection matrix P such as

$$P = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}, \tag{15}$$

the transmitted signal term $T(\omega)$ can be eliminated:

$$\vec{B}^m(\omega) = P\measuredangle \vec{R}^m(\omega) = \vec{D}k_m(\omega) - 2\pi\vec{\delta} + \vec{\mu}. \tag{16}$$

Here $\vec{D}$ represents the relative-distance vector, defined as $\vec{D} = P\vec{d}$. The addition of noise terms, which are both frequency and receiver dependent, produces a Gaussian noise vector $\vec{\mu}$, in which each element is expected to have increased variance compared to Eq. (14). Finally, $\vec{\delta}$ is an all-integer vector resulting from the multiplication of P with the vector of b_i integers.

Let $\hat{D}$ be the unit-norm direction of the relative distance vector $\vec{D}$ as follows:

$$\hat{D} = \frac{\vec{D}}{\|\vec{D}\|}. \tag{17}$$

Then, $\hat{D}\hat{D}^T\vec{a}$ describes a projection of a vector $\vec{a}$ onto $\hat{D}$, and $(I - \hat{D}\hat{D}^T)\vec{a}$ describes the part of vector $\vec{a}$ without any component in the $\hat{D}$ direction. To rephrase, $(I - \hat{D}\hat{D}^T)\vec{a}$ represents a projection of vector $\vec{a}$ onto the null-space of $\hat{D}$. Observe that the $\vec{D}k_m(\omega)$ term in Eq. (16) is the only source of nonzero-mean frequency dependence. Therefore, the variance over frequency is minimized when $\vec{B}^m(\omega)$ is projected onto the null-space of $\vec{D}$. This allows a search for $\hat{D}$ to be performed by minimizing

$$\begin{aligned} e(\hat{D}) &= \operatorname*{var}_{\omega}[(I - \hat{D}\hat{D}^T)\vec{B}^m(\omega)] \\ &= \operatorname*{var}_{\omega}[(\vec{D} - \hat{D}\hat{D}^T\vec{D})k_m(\omega) + (I - \hat{D}\hat{D}^T)\vec{\mu}]. \end{aligned} \tag{18}$$

Note that $e(\hat{D})$ is minimized when $\hat{D}$ is aligned with the relative-distance vector $\vec{D}$, as described in Eq. (17), since $\vec{D} - \hat{D}\hat{D}^T\vec{D} = \vec{0}$. Although an N-dimensional search can be computationally intensive, the error-metric is fairly well behaved, and the search can be performed using a gradient method over the restricted search space of N-dimensional unit vectors.

One method to improve estimation accuracy is to use frequency-dependent weighting in the variance computation of Eq. (18). By providing preference for higher SNR data, the search metric is more robust to noise at the frequencies with low signal power. Since the relative distance vector is identical for all modes, further improvement can also be obtained by incorporating information from all modes into a single vector,

$$\vec{B}(\omega) = \sum_{m=1}^{M} \vec{B}^m(\omega). \tag{19}$$

The minimization in Eq. (18) is then performed using the above $\vec{B}(\omega)$ vectors.

Once the relative distance vector is estimated, the mode-specific dispersion curve $k_m(\omega)$ can be estimated within an unknown constant offset C_{k_m}:

$$\begin{aligned} \hat{k}_m(\omega) &= \hat{D}^T\vec{B}^m(\omega) = (\hat{D}^T\vec{D})k_m(\omega) - 2\pi(\hat{D}^T\vec{\delta}) + \hat{D}^T\vec{\mu} \\ &\approx \|\vec{D}\|k_m(\omega) - C_{k_m} + \mu. \end{aligned} \tag{20}$$

Taking the derivative with respect to ω of the dispersion curve estimate of Eq. (20) yields a zero-bias, data-driven approximation of the inverse scaled group velocity, $v_{\mathrm{gr}}^{-1}(\omega)$. Figure 2 reflects the group velocity estimate. Note that although the inverse group velocity approximation is zero-bias, the group velocity estimate is biased away from zero with increased noise variance at the low and high frequencies indicative of the frequency-dependent SNR. Additionally, the noise level observed in $v_{\mathrm{gr}}(\omega)$ is higher than that of the $k_m(\omega)$ estimate because of the differentiation with respect to ω.

For reference to actual group velocity curves, the $\|\vec{D}\|$ scaling factor shown in Eq. (20) must be known. An estimate of $\|\vec{D}\|$ is obtained by incorporating information from the measured distance vector $\hat{d}_m$. A measured relative distance vector $\hat{D}_m$ is found using the projection matrix P of Eq. (15), i.e., $\hat{D}_m = P\hat{d}_m$. An estimate of $\|\vec{D}\|$ is then the absolute value of the inner product between the measured relative distance vector $\hat{D}_m$ and estimated relative distance vector $\hat{D}$:

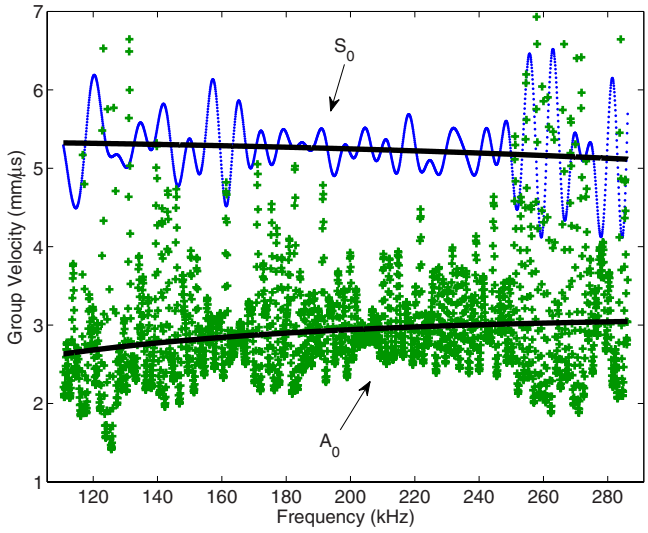


FIG. 2. (Color online) Group velocity estimate as a function of frequency. Effects of frequency-dependent SNR are clearly present below 150 and above 250 kHz, particularly for the A_0 estimate.

$$|\hat{D}_m^T \hat{D}| = \frac{|\hat{D}_m^T \vec{D}|}{\|\vec{D}\|} \approx \|\vec{D}\|. \tag{21}$$

C. Common distance

The next step in the algorithm is a search to characterize the distance vector by leveraging the information contained in the relative distance vector $\vec{D}$. Since it can be shown that the null-space of P in Eq. (15) is spanned by a single, constant-valued vector, the scaled distance vector $\hat{d}$ can be described as a linear combination of the null-space of P and the Moore–Penrose pseudo-inverse[22] of P (denoted by †) applied to the relative-distance vector:

$$\hat{d}(\hat{\gamma}) = P^{\dagger}\hat{D} + \hat{\gamma}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \approx \frac{\vec{d}}{\|\vec{D}\|}. \tag{22}$$

Here, $\hat{\gamma}$ is referred to as the scaled common distance variable since it represents a distance that is common to all elements of the scaled distance vector.

The scaled common distance is found by subtracting the estimated phase changes due to dispersion from the phase-response of the received signals:

$$\begin{aligned} \measuredangle \vec{R}^m(\omega) - \hat{k}_m(\omega)\hat{d}(\hat{\gamma}) &= \left(\measuredangle T(\omega)\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \vec{d}k_m(\omega) - 2\pi\vec{b} \right) - (\|\vec{D}\|k_m(\omega) - C_{k_m})\hat{d}(\hat{\gamma}) + \vec{\mu} \\ &= \measuredangle T(\omega)\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + (\vec{d} - \|\vec{D}\|\hat{d}(\hat{\gamma}))k_m(\omega) + \vec{d}(\hat{\gamma})C_{k_m} - 2\pi\vec{b} + \vec{\mu}. \end{aligned} \tag{23}$$

Let $\vec{t}^m(\hat{\gamma},\omega)$ be the vector defined as the derivative of Eq. (23) with respect to frequency. Assuming a frequency spacing of $\Delta\omega$, the derivative is approximated as

$$\begin{aligned} \vec{t}^m(\hat{\gamma},\omega) &= \frac{\Delta}{\Delta\omega}\measuredangle \vec{R}^m(\omega) - \frac{\Delta}{\Delta\omega}\hat{k}_m(\omega)\hat{d}(\hat{\gamma}) \\ &= \frac{\Delta}{\Delta\omega}\measuredangle T(\omega)\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + (\vec{d} - \|\vec{D}\|\hat{d}(\hat{\gamma}))\frac{\Delta}{\Delta\omega}k_m(\omega) + \vec{\mu}. \end{aligned} \tag{24}$$

Equation (24) reveals that if $\vec{d} - \|\vec{D}\|\hat{d}(\hat{\gamma}) = 0$, then each $\vec{t}^m(\hat{\gamma},\omega)$ vector represents a set of N independent estimates of the same value $\Delta\measuredangle T(\omega)/\Delta\omega$. These mode-specific estimates can be combined into a single $(NM)\times W$ matrix $T(\hat{\gamma})$:

$$T(\hat{\gamma}) = \begin{bmatrix} \vec{t}^1(\hat{\gamma},\omega_1) & \cdots & \vec{t}^1(\hat{\gamma},\omega_W) \\ \vdots & & \vdots \\ \vec{t}^M(\hat{\gamma},\omega_1) & \cdots & \vec{t}^M(\hat{\gamma},\omega_W) \end{bmatrix}, \tag{25}$$

where W is the number of frequencies. The scaled common distance variable $\hat{\gamma}$ is found by searching for the value that creates the most agreement between the rows of $T(\hat{\gamma})$, which can be done quantitatively by several different techniques. An intuitive approach to measuring the agreement between rows is to measure the variance along each column of $T(\hat{\gamma})$ and perform a SNR-based weighted average across the resulting values. More accurate results, however, were consistently obtained by first taking the SNR-weighted average across all frequencies, and then analyzing the variance across receivers and modes. The resulting error metric is

$$e(\hat{\gamma}) = \mathrm{var}(E_{\mathrm{SNR}}\{T(\hat{\gamma})\}). \tag{26}$$

The improvement associated with this technique is likely a result of the noise averaging that occurs prior to the variance calculation ($NM \ll W$).

This error-metric, like the error-metric in Eq. (18) for the relative-distance vector, is well-behaved and takes into account data from all receivers and all modes. Unlike the

error-metric in Eq. (18), however, the search-space is one-dimensional and can be performed using any linear search algorithm.

Once the scaled common distance value $\hat{\gamma}$ is found, the scaled distance vector $\hat{d}$ is obtained from Eq. (22), and the derivative of the phase-response of the transmitted signal is found using a SNR-weighted sum of the rows of $T(\hat{\gamma})$ in Eq. (25). Integrating with respect to frequency, the transmitted signal phase-response is then

$$\measuredangle \hat{T}(\omega) = \measuredangle T(\omega) + C_T + \mu, \tag{27}$$

where C_T is an unknown constant of integration.

D. Constant offsets

Although both the dispersion curves and transmitted signal phase-response have now been estimated, these estimates have unknown additive constants (i.e., C_{k_m} and C_T) that must be resolved. A frequency-dependent phase term is calculated as follows:

$$\vec{\theta}^m(\omega) = \measuredangle \vec{R}^m(\omega) - \measuredangle \hat{T}(\omega) - \hat{d}(\hat{\gamma})\hat{k}_m(\omega). \tag{28}$$

By substituting Eqs. (9), (20), (22), and (27) into Eq. (28), this phase term can be expressed as

$$\vec{\theta}^m(\omega) = C_{k_m}\hat{d}(\hat{\gamma}) - C_T\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix} - 2\pi\begin{bmatrix}b_1\\ \vdots \\ b_N\end{bmatrix} + \begin{bmatrix}\mu\\ \vdots \\ \mu\end{bmatrix}, \tag{29}$$

which should not change with frequency. A SNR-based weighted average can then be taken across the frequency spectrum to arrive at a single phase vector estimate:

$$\vec{\theta}^m = E_{\mathrm{SNR}}\{\vec{\theta}^m(\omega)\}. \tag{30}$$

Similar to the approach taken previously, the C_T term of Eq. (29) is removed by subtracting one row from another using the projection matrix P from Eq. (15),

$$P\vec{\theta}^m = C_{k_m}P\hat{d}(\hat{\gamma}) - 2\pi\vec{\delta} + P\vec{\mu}, \tag{31}$$

where $\vec{\delta}$ represents an all-integer vector. Ignoring the noise term, these equations can be restructured into matrix form:

$$P\vec{\theta}^m = [P\hat{d}(\hat{\gamma}) \mid -2\pi I_{N-1}]\begin{bmatrix}C_{k_m}\\ \vec{\delta}\end{bmatrix} = X\vec{f}, \tag{32}$$

where I_{N-1} is the $(N-1)\times(N-1)$ identity matrix. It can be verified that since P has a single-vector null-space, the matrix X also has a single-vector null-space. Therefore, let $\vec{f}(\lambda)$ be defined as the following vector:

$$\vec{f}(\lambda) = \begin{bmatrix}C_{k_m}\\ \vec{\delta}\end{bmatrix} = X^{\dagger}P\vec{\theta}^m + \lambda\ \ \mathrm{null}(X). \tag{33}$$

This equation shows that valid values of C_{k_m} are directly related to $\vec{\delta}$. Specifically, since each element of $\vec{\delta}$ must be an integer, any value of λ that produces all-integer elements of $\vec{\delta}$ also produces a value of C_{k_m} that satisfies the assumed model. Therefore, a search for C_{k_m} can be performed by minimizing the following error metric:

$$e(\lambda) = \|\vec{f}^*(\lambda) - \mathrm{round}(\vec{f}^*(\lambda))\|, \tag{34}$$

where $\vec{f}^*(\lambda)$ corresponds to the $\vec{f}(\lambda)$ vector of Eq. (33) without the first element.

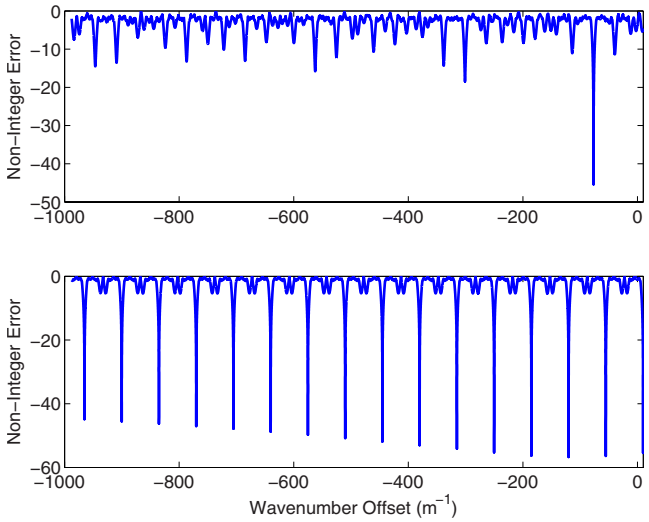


FIG. 3. (Color online) Illustration of the presence of a unique minimum (top) and multiple minima (bottom) in the error metric for estimating the dispersion curve offset C_{k_m}. Multiple minima are a result of spatial subsampling and can be avoided by carefully selecting receiver locations relative to the transmitter.

Unfortunately, one of the implications of Eq. (33) is that the resulting C_{k_m} is not unique. Two examples are provided in Fig. 3 that demonstrate the existence or lack of multiple minima from the error-metric of Eq. (34). The receiver distances are critical to the existence (or lack) of multiple minima. The ambiguity is a direct result of spatial subsampling and is a phenomenon observed in earlier work by Ting and Sachse[23] in the context of bulk-wave dispersion estimation. In addition, irregularly spaced receivers (especially with irrational spacings) can maximize the distance between minima, allowing reasonable bounds to be used to select the appropriate value for C_{k_m}.

It is possible to establish a lower-bound on the ambiguity of C_{k_m}. To do so, notice that the impact on the phase term of Eq. (29) due to an error in C_{k_m} of Δ is simply $\vec{d}\Delta/\|\vec{D}\|$. Therefore, ambiguous solutions result when

$$\vec{d}\Delta/\|\vec{D}\| = 2\pi\vec{\delta}, \tag{35}$$

where $\vec{\delta}$ is an all-integer vector.

Let $\Delta_{\min}$ be the smallest value of Δ that satisfies Eq. (35). Considering the special case when C_Δ is some scalar value and $C_\Delta\vec{d}$ is an all-integer vector,

$$\Delta_{\min} = \frac{2\pi C_\Delta\|\vec{D}\|}{\mathrm{gcd}(C_\Delta\vec{d})}. \tag{36}$$

Here $\mathrm{gcd}(\cdot)$ represents the greatest common denominator for all vector elements within the parentheses.

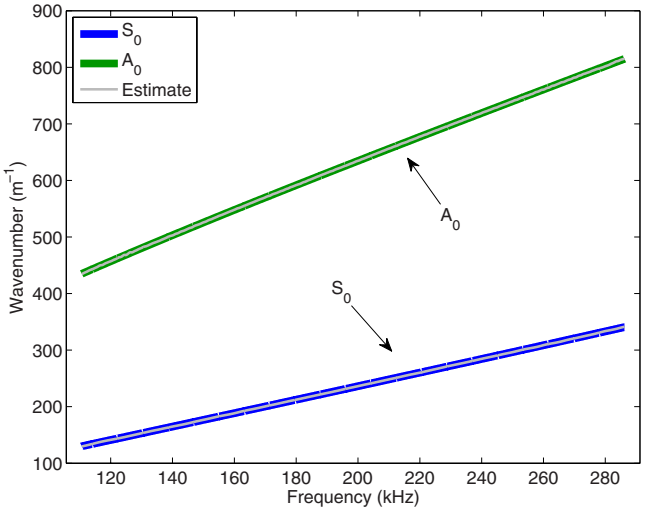


FIG. 4. (Color online) Comparison of estimated and actual dispersion curves.

Since there are some uncertainties in receiver distances, both due to measurement uncertainty and the finite size of the transducers, a path accuracy is defined as

$$\Delta p = 10^{\lceil \log_{10}(2d_t) \rceil}, \tag{37}$$

where d_t is the uncertainty of the effective transducer location and $\lceil \cdot \rceil$ represents the ceiling function. Let $\hat{d}_m$ be the measured (nominal) distance vector and $\|\hat{D}_m\|$ be an approximation of the norm of the relative distance vector (obtained using Eq. (21) or more directly as $\|P\hat{d}_m\|$). Given the uncertainty in path measurement indicated by Δp, it is possible that $\vec{d}/\Delta p$ is very close to an all-integer vector. Therefore, Eq. (36) can be modified to bound $\Delta_{\min}$:

$$\Delta_{\min} \geq \frac{2\pi\|\hat{D}_m\|}{\Delta p \ \gcd\left[\text{round}\left(\frac{1}{\Delta p}\hat{d}_m\right)\right]}, \tag{38}$$

As $\Delta p \to 0$, $\Delta_{\min}$ reaches a stable value.

In Fig. 3, the top graph corresponds to the test case of Fig. 1, where $\Delta_{\min} \to 37\,600$ as $\Delta p \to 0$. The bottom graph, however, reflects alternative receiver distances of 400, 500, 600, and 750 mm, where $\Delta_{\min} \to 65$ as $\Delta p \to 0$.

With an estimate of the constant offset C_{k_m}, the scaled dispersion curves $\|\vec{D}\|k_m(\omega)$ of Eq. (20) can be obtained. Figure 4 compares estimated $k_m(\omega)$ dispersion curves with the actual curves used to generate the signals. Estimates for both modes exhibit excellent agreement with the reference values.

The constant C_T of Eq. (27) can be found, modulo 2π, by substituting the C_{k_m} term into Eq. (29), performing a modulo 2π operation on every element to remove the $2\pi b_i$ terms, and then performing a SNR-based weighted average across all modes and frequencies.

E. Magnitude response and propagation loss

The transmitted signal magnitude and propagation loss are estimated using the magnitude response of the received signals. Subtracting the log of the estimated mode weighting coefficients from the log of the received magnitude response yields

$$\log|R^m_{DAi}(\omega)| - \log|\alpha_m(\omega)| = \log|T(\omega)| + d_i p(\omega) + \nu. \tag{39}$$

All received signals from the same mode can be expressed in vector form:

$$\vec{X}^m(\omega) = \begin{bmatrix} \log|R^m_{DA1}(\omega)| \\ \vdots \\ \log|R^m_{DAN}(\omega)| \end{bmatrix} - \log|\alpha_m(\omega)| \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{40}$$

Substituting Eq. (39) into Eq. (40):

$$X^m(\omega) = \begin{bmatrix} 1 & d_1/\|\vec{D}\| \\ \vdots & \vdots \\ 1 & d_N/\|\vec{D}\| \end{bmatrix} \begin{bmatrix} \log|T(\omega)| \\ \|\vec{D}\|p(\omega) \end{bmatrix} + \vec{\nu}. \tag{41}$$

Note that a scaling constant has been introduced so that the estimated distance vector can be used in the rightmost column of the matrix in the right-hand-side of Eq. (41). Also, it is worth noting that the two columns of the right-hand-side matrix are linearly independent, but are not orthogonal. This can be seen in Eq. (22) by the fact that one of the components of the estimated distance vector is the all-ones vector multiplied by the common distance variable. Therefore, the numerical stability of this system of equations is largely dependent on the magnitude relationship between the common distance variable and relative distance vector.

All received modes can be further combined into a composite system of equations:

$$\begin{bmatrix} \vec{X}^1(\omega) \\ \vdots \\ \vec{X}^M(\omega) \end{bmatrix} = \begin{bmatrix} 1 & \vec{d}/\|\vec{D}\| \\ \vdots & \vdots \\ 1 & \vec{d}/\|\vec{D}\| \end{bmatrix} \begin{bmatrix} \log|T(\omega)| \\ \|\vec{D}\|p(\omega) \end{bmatrix} + \vec{\nu} \tag{42}$$

The two columns of the matrix on the right-hand-side of Eq. (42) are again linearly independent. Therefore, the Moore–Penrose pseudo-inverse can be used to obtain the least-squares approximation of both the transmitted-signal amplitude response and the scaled propagation loss,

$$\begin{bmatrix} \log|\hat{T}(\omega)| \\ \|\vec{D}\|\hat{p}(\omega) \end{bmatrix} = Y^{\dagger}\hat{X}(\omega) \quad \text{where} \quad Y = \begin{bmatrix} 1 & \hat{d}(\hat{\gamma}) \\ \vdots & \vdots \\ 1 & \hat{d}(\hat{\gamma}) \end{bmatrix} \tag{43}$$

and $\hat{X}(\omega)$ is defined as in Eq. (40), using the estimated mode weighting coefficients, $\hat{\alpha}_m(\omega)$.

The top plot of Fig. 5 shows the composite (magnitude and phase) estimation performance for the transmitted signal. Recall that this transmitted signal is not necessarily the same as the excitation signal because it incorporates the transmitter and receiver transducer transfer functions and the guided wave excitation. The bottom plot of Fig. 5 demonstrates typical model-fit accuracy for this input SNR level. An estimate of the received signal at 803 mm is generated by forward propagating the estimated transmitted signal using the previously estimated parameters and is compared with the actual received signal at 803 mm.

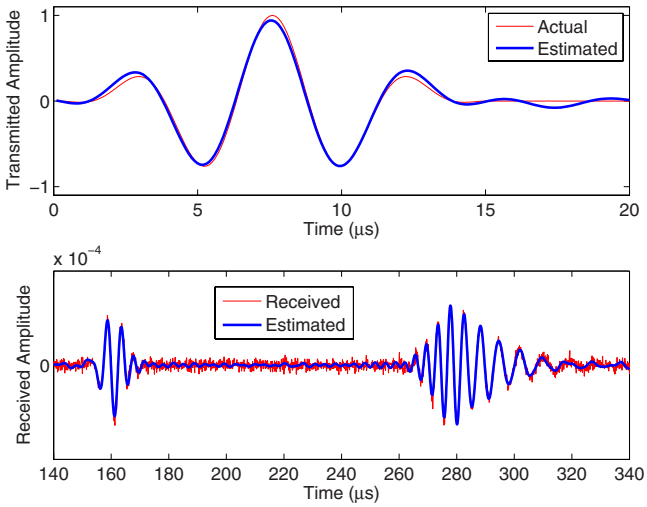


FIG. 5. (Color online) Comparison of estimated and actual transmitted signals (top) and estimated and actual received signals (bottom) for the test case. Signals in the bottom plot correspond to the receiver located at 803 mm.

IV. RESULTS

Multiple simulations were used to characterize the expected performance of the algorithm whereby signals were computed using numerical estimates of theoretical dispersion curves. Performance results are presented for noisy, synthetic test data that match the assumed model, and thus represent best-case performance for the specific SNR level in use. Although the actual parameters associated with the propagating environment are not available for many experimental situations, such simulations allow the comparison of estimated parameters to "true" parameters to obtain performance metrics. Signals obtained from finite element modeling are also considered to verify algorithmic performance on data that do not exactly match the model.

To quantitatively analyze performance, compare estimation performance across parameters, and provide some degree of consistency between model-fit and parameter estimation performance, a generalized error-metric, referred to as the parameter-error-ratio (PER), is utilized; it is analogous to SNR and is the reciprocal of the normalized least-squares error:

$$\mathrm{PER} = \frac{\sum_{k=0}^{k-1} |w_k||p_k|^2}{\sum_{k=0}^{k-1} |w_k||p_k - \hat{p}_k|^2}. \tag{44}$$

Here p_k is the parameter being estimated (i.e., distance vector, mode weighting coefficients, etc.), w_k is a weighting coefficient, $\hat{p}_k$ is the estimated parameter, and k is an index used to span the length of the comparison. For all frequency-dependent parameters, w_k is based on SNR; however, distance vector estimates use uniform weights for the PER calculation. Model fit SNR is computed in a similar fashion to Eq. (44), where k now spans all samples over all N time-domain signals:

$$\mathrm{SNR} = \frac{\sum_{n=1}^{N} \sum_{t} |r_n[t]|^2}{\sum_{n=1}^{N} \sum_{t} |r_n[t] - \hat{r}_n[t]|^2}. \tag{45}$$

Input SNR is calculated as in Eq. (45) with the exception that the denominator (Gaussian noise term) is scaled to account for the nominal bandwidth of interest, which is a fraction of the total sampling bandwidth.

Algorithm performance is evaluated based on 2100 separate simulations. The transmitted signal and dispersion curves are those used for Fig. 1 and are held constant across all iterations to provide some consistency and simplify coding. The simulations are based on 100 variations of mode weighting coefficients, propagation loss, and receiver locations. Mode weighting coefficients are selected at random with the constraint that the sum of the squares equals 1 and magnitudes are constant across all frequencies. Values for propagation loss are selected from a uniform distribution between -7 and -12 and are also constant across the frequency spectrum. Five receiver distances are chosen randomly with the constraint that the worst-case minimum wavenumber-frequency offset from Eq. (38), $\Delta_{\min}$, is sufficiently high to avoid ambiguity. One set of synthetic received signals is generated for each of the 100 variations. The case of four receivers is considered by dropping one of the five signals, where the dropped signal is selected such that the $\Delta_{\min}$ for the remaining four receivers is sufficiently high to avoid ambiguity. Three receivers are considered by dropping an additional signal in the same manner. Gaussian noise is added to each set of synthetic received signals to achieve seven desired input SNR ratios, calculated over a nominal bandwidth of 200 kHz. Note that the noise level and SNR for each specific receiver varies as a function of the number of receivers because the composite signal power is changed.

For each simulation, the noisy signals are presented to the estimation algorithm and the estimated parameters are used to model the original received signals. The estimated parameters and modeled received signals are then compared to the original parameters and noise-free synthetic received signals to determine algorithmic performance.

Figure 6 presents average parameter estimation performance of the proposed algorithm. These results provide an indication of how accurately model parameters are estimated given an input SNR value. It can be seen that for a given input SNR, the inverse group velocity values are estimated with the lowest PER of all parameters, while the distances and dispersion offset values are obtained with the highest PER values.

Similarly, Fig. 7 presents averaged model-fit performance of the proposed algorithm. One can see that for a given input SNR, additional receivers tend to improve model-fit performance, which is to be expected since additional signals provide a means to further average out noisy measurements. Note that some of the larger errors (i.e., low model-fit SNR) are attributed to very asymmetric mode weighting coefficients, which can result in degraded dispersion curves and transmitted signal phase estimation.

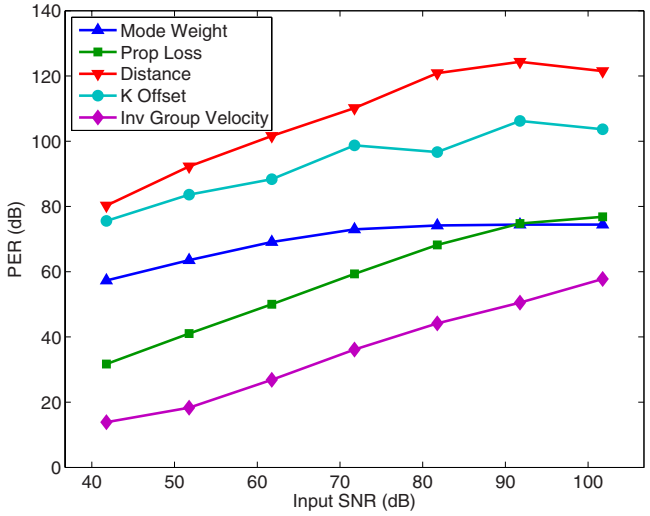


FIG. 6. (Color online) Average parameter estimation performance from 2100 simulations. Each simulation is based on randomly selected distances, mode weighting coefficients, and propagation loss.

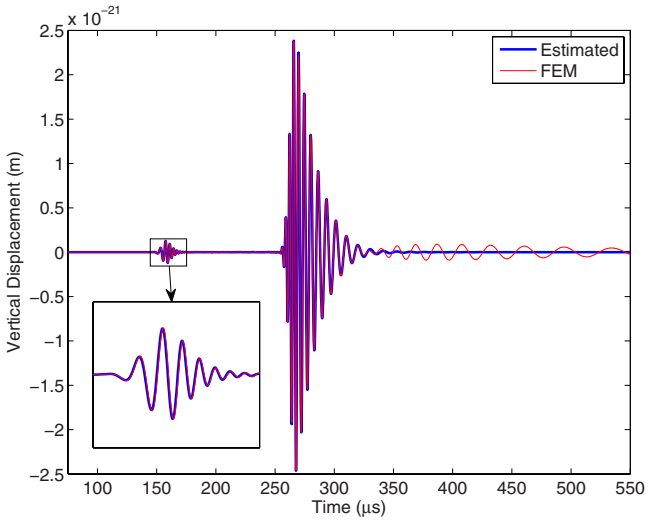


FIG. 8. (Color online) Comparison of the FEM signal to the signal estimated from the model for a receiver distance of 803 mm. The lower frequency oscillations after ~320 μs are present in the FEM signal but not the model, and reside outside of the analyzed signal bandwidth.

Figure 8 illustrates model-fit performance with input data from a 2D finite element model (FEM). The FEM simulations were conducted using ABAQUS/EXPLICIT, which uses explicit time integration. Receiver distances are identical to the simulation case used throughout this paper. The FEM data represent transmission and reception that are free of electronic noise, although model-mismatch and numerical noise are present. The 2D simulation was performed on a 3 mm thick aluminum plate using a time resolution of 0.05 μs and a grid spacing of 0.5 mm. The excitation was a vertical force whose time history was a three-cycle, Hanning-windowed sinusoid with a center frequency of 250 kHz. All received signals were out-of-plane displacements obtained from the side of the plate opposite the transmitter. Since the 2D simulations did not reflect geometric spreading loss or material absorption, the FEM signals were artificially attenuated according to the assumed propagation model. There is clear agreement between the FEM received signals and those calculated from the model, achieving a model-fit SNR of 18 dB. Note that beginning at approximately 320 μs, some oscillations in the FEM data are not matched in the forward propagated model. These ~36 kHz oscillations, which are probably numerical artifacts, are not in the analyzed spectrum (~100–400 kHz) and are therefore not incorporated into the model. Figure 9 compares dispersion curves obtained from the algorithm to nominal dispersion curves calculated for the same material properties and thickness as used for the finite element model. The agreement is excellent for both modes despite the much lower amplitude of the S_0 mode as compared to A_0.

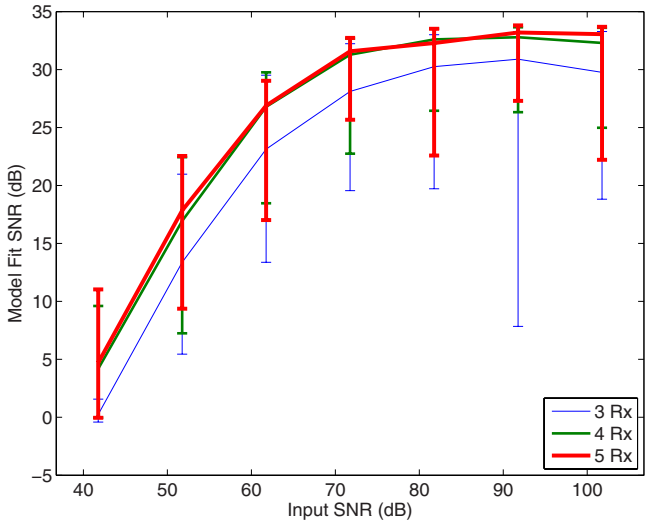


FIG. 7. (Color online) Average model-fit performance from 2100 simulations (100 per data point). Each simulation is based on randomly selected distances, mode weighting coefficients, and propagation loss. Error bars correspond to 1 standard deviation, and the asymmetry about the mean is caused by the linear-log mapping.

V. SUMMARY AND CONCLUSIONS

This paper presents an algorithm that leverages an assumed propagation model to adaptively estimate dispersion curves, propagation loss, transducer distances, and mode

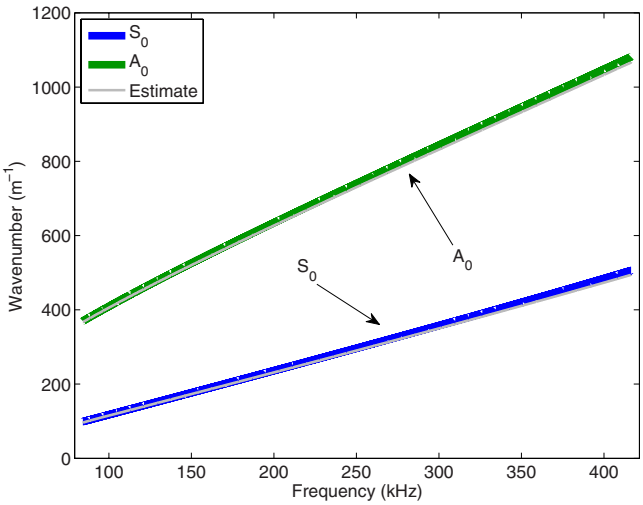


FIG. 9. (Color online) Comparison of dispersion curves estimated from FEM data to those calculated using the nominal thickness and material properties.

weighting coefficients with minimal *a priori* information. The five-part algorithm is described in detail and performance results are provided with both numerically simulated and FEM data. The primary contributions of this paper are the mathematical development of the algorithm, the ability to simultaneously estimate multiple guided wave parameters, and the applicability of the algorithm to *in situ* measurements using a small number of spatially distributed sensors. As is shown in the paper, estimated parameters are available in a form where they can be directly used to compute received signals. It is expected that parameters adaptively estimated at the time of test can be used in nondestructive evaluation and structural health monitoring algorithms to reduce errors caused by *a priori* assumptions such as the use of nominal dispersion curves and inaccurate transducer locations.

The limitations of the algorithm are also discussed. First and foremost, the algorithm is model-based in nature, and therefore all model assumptions must hold in order to achieve satisfactory results. Additionally, distances and dispersion curves can only be established to within a common scaling factor, which must be combined with measured distances to resolve the uncertainty. Further, transducer spacing plays a direct role in the ambiguity (or lack thereof) of dispersion curves, and judicious selection of transducer distances is necessary. Finally, the requirement for mode separation is a limitation that has not yet been addressed and may limit algorithm applicability, especially when higher order guided wave modes are present.

Although results for two modes are shown here, the algorithm is also suitable for single-mode signals. This case is of significant interest because of the emphasis on achieving mode purity for both nondestructive evaluation and structural health monitoring applications. Additionally, a single mode approach would be appropriate to apply if the model does not hold, such as may be the case if propagation loss values or combined isotropic transducer functions are not common for all measured modes.

Although the algorithm was developed for isotropic media and transducers, it could be applied to anisotropic media under some circumstances. For example, if multiple receivers were placed along directions of propagation with the same dispersion characteristics, then the propagation model would apply and the algorithm could be used.

Future work should be directed at applying the algorithm to experimental data, further improving algorithmic computational performance, and adapting the algorithm to eliminate the mode-separation requirement.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of NASA through the Graduate Student Research Program, under Grant No. NNX08AY93H, and the Air Force Office of Scientific Research, under Grant No. FA9550-08-1-0241. The authors also thank Mr. Joannes Pezant for his assistance with the derivation of Eq. (38).

[1] P. Wilcox, M. Lowe, and P. Cawley, "The effect of dispersion on long-range inspection using ultrasonic guided waves," NDT Int. **34**, 1–9 (2001).
[2] P. D. Wilcox, "A rapid signal processing technique to remove the effect of dispersion from guided wave signals," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **50**, 419–427 (2003).
[3] P. S. Tua, S. T. Quek, and Q. Wang, "Detection of cracks in plates using piezo-actuated Lamb waves," Smart Mater. Struct. **13**, 643–660 (2004).
[4] J. E. Michaels, "Detection, localization and characterization of damage in plates with an in situ array of spatially distributed ultrasonic sensors," Smart Mater. Struct. **17**, 035035 (2008).
[5] J.-B. Ihn and F.-K. Chang, "Detection and monitoring of hidden crack growth using a built-in piezoelectric sensor/actuator network," Smart Mater. Struct. **13**, 609–620 (2004).
[6] A. J. Croxford, P. D. Wilcox, B. W. Drinkwater, and G. Konstantinidis, "Strategies for guided-wave structural health monitoring," Proc. R. Soc. A **463**, 2961–2981 (2007).
[7] W. Sachse and Y.-H. Pao, "On the determination of phase and group velocities of dispersive waves in solids," J. Appl. Phys. **49**, 4320–4327 (1978).
[8] F. Peters and L. Petit, "A broad band spectroscopy method for ultrasound wave velocity and attenuation measurement in dispersive media," Ultrasonics **41**, 357–363 (2003).
[9] D. A. Hutchins, K. Lundgren, and S. B. Palmer, "A laser study of transient Lamb waves in thin materials," J. Acoust. Soc. Am. **85**, 1441–1448 (1989).
[10] W. H. Prosser and M. R. Gorman, "Plate mode velocities in graphite/epoxy plates," J. Acoust. Soc. Am. **96**, 902–907 (1994).
[11] R. D. Mindlin, "Influence of rotatory inertia and shear on flexural motions of isotropic, elastic plates," J. Appl. Mech. **18**, 31–38 (1951).
[12] N. A. Schumacher, C. P. Burger, and P. H. Gien, "A laser-based investigation of higher-order modes in transient Lamb waves," J. Acoust. Soc. Am. **93**, 2981–2984 (1993).
[13] H. Kwun and K. A. Bartels, "Experimental observation of elastic-wave dispersion in bounded solids of various configurations," J. Acoust. Soc. Am. **99**, 962–968 (1996).
[14] W. H. Prosser, M. D. Seale, and B. T. Smith, "Time-frequency analysis of Lamb modes," J. Acoust. Soc. Am. **105**, 2669–2676 (1999).
[15] M. Neithammer, L. J. Jacobs, J. Qu, and J. Jarzynski, "Time-frequency representation of Lamb waves," J. Acoust. Soc. Am. **109**, 97–102 (2001).
[16] D. Alleyne and P. Cawley, "A two-dimensional Fourier transform method for the measurement of propagating multimode signals," J. Acoust. Soc. Am. **89**, 1159–1168 (1991).
[17] R. D. Costley, Jr. and Y. H. Berthelot, "Dispersion curve analysis of laser-generated Lamb waves," Ultrasonics **32**, 249–253 (1994).
[18] C. Eisenhardt, L. J. Jacobs, and J. Qu, "Application of laser ultrasonics to develop dispersion curves for elastic plates," J. Appl. Mech. **66**, 1043–1045 (1999).
[19] P. He and J. Zheng, "Acoustic dispersion and attenuation measurement using both transmitted and reflected pulses," Ultrasonics **39**, 27–32 (2001).
[20] S. Holland, T. Kosel, R. Weaver, and W. Sachse, "Determination of plate source, detector separation from one signal," Ultrasonics **38**, 620–623 (2000).
[21] D. V. Hinkley, "On the ratio of two correlated normal random variables," Biometrika **56**, 635–639 (1969).
[22] R. Penrose, "A generalized inverse for matrices," Proc. Cambridge Philos. Soc. **51**, 406–413 (1955).
[23] C. S. Ting and W. Sachse, "Measurement of ultrasonic dispersion by phase comparison of continuous harmonic waves," J. Acoust. Soc. Am. **64**, 852–857 (1978).

Influence of inhibitory synaptic kinetics on the interaural time difference sensitivity in a linear model of binaural coincidence detection

Christian Leibold
Division of Neurobiology, University of Munich, and Bernstein Center for Computational Neuroscience Munich, Großhaderner Strasse 2, 82152 Planegg-Martinsried, Germany

(Received 27 May 2009; revised 4 December 2009; accepted 9 December 2009)

Temporal correlations between the sound waves arriving at the two ears are used to extract the azimuthal position of sound sources. Nerve cells in the mammalian medial superior olive (MSO) that extract these binaural correlations are sensitive to interaural time differences (ITDs) in the range of about 10 μs. These neurons receive inputs from the two ears via four pathways, two excitatory and two inhibitory ones. In this paper, a simple linear model is fitted to the frequency dependence of ITD sensitivity of MSO neurons, which is quantified by the two parameters, characteristic phase and characteristic delay. The fit parameters are the relative delays and the relative strengths of the two inhibitory pathways and thus specify the underlying ITD-detecting circuit assuming a non-Jeffress-like situation, i.e., no excitatory delay lines but phase-locked inhibition. The fitting procedure finds the parameters of these inhibitory pathways such that they account for a desired frequency dependence of ITD sensitivity. It is found that positive characteristic delays require a finite amount of ipsilateral inhibition that arrives at roughly the same time as ipsilateral excitation as well as contralateral inhibition that lags contralateral excitation so much that it effectively leads excitation of the next cycle. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3282997]

PACS number(s): 43.64.Bt, 43.64.Qh, 43.66.Qp, 43.66.Pn [JCM] Pages: 931–942

I. INTRODUCTION

Humans localize the azimuthal position of a low-frequency sound based on the difference of arrival times of the sound wave at the two ears (Sandel *et al.*, 1955; Blauert, 1997). These interaural time differences (ITDs) are up to several hundred μs. For animals with smaller heads such as cats and gerbils, this physiological ITD range is correspondingly smaller. The acuity with which humans and low-frequency hearing animals can resolve ITDs is on the order of only 10 μs (Mills, 1958; Heffner and Heffner, 1988). Although there is general agreement that the basic mechanism that underlies ITD acuity is coincidence detection of the binaural inputs, the biophysics that allows this most precise temporal processing is still a matter of debate (McAlpine and Grothe, 2003; Joris and Yin, 2006). The traditional model by Jeffress (Jeffress, 1948) proposes ITD sensitivity to be generated by combining coincidence detector neurons with axonal delay lines that compensate for the disparities in acoustic traveling times. There is strong experimental evidence in favor of Jeffress-type architectures in chicken (Young and Rubel, 1983) and barn owls (Carr and Konishi, 1990) and a lot of labor went into the search for evidence in mammals (Smith *et al.*, 1993; Beckius *et al.*, 2000). Several years ago data from guinea pig (McAlpine *et al.*, 1996, 2001) and gerbil (Brand *et al.*, 2002) raised doubts about whether the Jeffress model is valid for mammals (see Sec. IV). Furthermore, Brand *et al.* (2002) showed some evidence that, instead of delay lines, the observed time shifts might be generated by *phase-locked* inhibitory inputs. In this paper, a theoretical analysis is presented to further test the potential of phase-locked inhibition to explain the observed ITD sensitivity.

Low-frequency hearing is characterized by neuronal responses that are phase-locked to the sound pressure signal of the acoustic stimulus. The time structure of the stimulus is preserved directly within the timing of action potentials (APs) (Galambos and Davis, 1943; Rose *et al.*, 1967; Joris *et al.*, 1994). In the first binaural stage, which in mammals is called the medial superior olive (MSO), the sound-evoked neuronal activities from the two ears are combined (Goldberg and Brown, 1969; Yin and Chan, 1990). The MSO cells are sensitive to the ITD of a stimulus in that their discharge rate is strongly modulated by the ITD. Most MSO cells respond strongest if the sound arrives first at the contralateral ear (Brand *et al.*, 2002; Pecka *et al.*, 2008). To understand this sensitivity on a mechanistic basis, the present study focuses on modeling the temporal interference of the different phase-locked synaptic inputs impinging on the cell in response to an acoustic pure tone stimulus.

Neurons in the MSO receive phase-locked inputs via four principal pathways. Excitation is conveyed via the spherical bushy cells of the ventral cochlear nuclei in both brain hemispheres. Ipsilateral inhibition originates from the lateral nucleus of the trapezoid body (LNTB), and contralateral inhibition is mediated by the medial nucleus of the trapezoid body (MNTB) (Kuwabara and Zook, 1992; Cant and Hyson, 1992; Grothe and Sanes, 1994). The decay time constants of the excitatory synaptic currents are in the range of only few hundred μs (Smith, 1995; Scott *et al.*, 2005) and therefore would suffice to explain the observed localization acuity (Brand *et al.*, 2002; Gerstner *et al.*, 1996; Kempter *et al.*, 1998; Joris *et al.*, 1998; Svirskis *et al.*, 2003; Leibold and van Hemmen, 2005; Zhou *et al.*, 2005). The inhibitory syn-

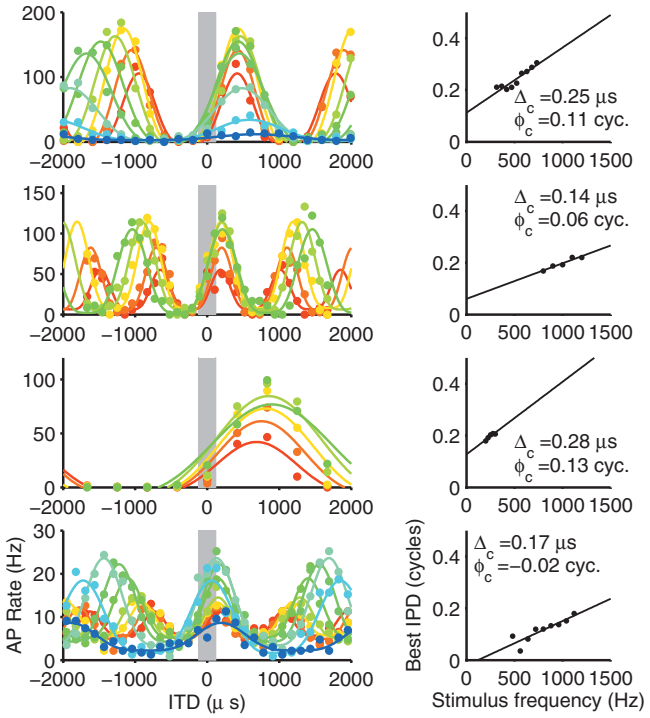


FIG. 1. Frequency dependence of ITD sensitivity in four MSO neurons (rows) of the Mongolian gerbil. (Left column) Tone delay functions: Extracellularly recorded AP discharge rate as a function of the ITD of a pure tone stimulus. Lines are derived from $\cos^2$ fits to the data (dots). Different colors indicate different stimulus frequencies (blue: low; red: high). (Right column) Phase-frequency curves: Best IPD as a function of stimulus frequency derived from the tone delay functions on the left (dots). Lines indicate the best linear fit to the data with slope Δ_c and offset ϕ_c. Data replotted from Pecka *et al.* (2008) by courtesy of the authors.

aptic currents decay with time constants of around 1–2 ms (Smith, 1995; Magnusson *et al.*, 2005). It is unclear whether these synaptic kinetics is sufficient to allow the inhibitory currents to impose phase shifts on the excitatory inputs and interfere with coincidence detection or not.

In this paper, it is shown that, despite the slow synaptic kinetics, phase locking of inhibition heavily influences the phase of the oscillating total input current and thus indeed contributes significantly to the temporal processing underlying coincidence detection. The model therefore provides an explanation of ITD sensitivity of MSO cells *in vivo* (Brand *et al.*, 2002) without the need to either employ delay lines for the excitatory pathways as they are needed in the classical Jeffress model (Jeffress, 1948) or to assume additional cellular mechanisms such as morphological asymmetries (Zhou *et al.*, 2005). This, of course, does not mean that the mechanisms beyond the present model would not also be able to contribute to ITD tuning in the MSO.

A. Phase-frequency curves

The best ITD τ_{best} at which a MSO neuron exhibits the highest firing rate typically changes with stimulus frequency f (Fig. 1, left). This frequency dependence is traditionally quantified via phase-frequency curves (Yin and Kuwada, 1983) in which the best interaural phase difference (best IPD), defined as $\phi_{best}=2\pi f \tau_{best}$, is plotted versus stimulus frequency f (Fig. 1, right). The phase-frequency curve is classically described by two parameters, the characteristic phase ϕ_c and the characteristic delay Δ_c derived from the best linear fit

$$\phi(f) = \phi_c + \Delta_c 2\pi f \quad (1)$$

to the measured phase-frequency curve. The quantities ϕ_c and Δ_c are thus cell-specific parameters that generalize the concept of the best IPD (and ITD) to multiple stimulus frequencies.

Phase-frequency curves of MSO neurons are rather linear (Pecka *et al.*, 2008; Pecka, private communication), at least as compared to the midbrain where non-linear phase-frequency curves have also been observed (Yin and Kuwada, 1983). Characteristic phases of MSO cells typically have absolute values smaller than 0.25 cycles. Most of the cells have small characteristic phase $\phi_c \lessapprox 0.1$ cycle with a bias toward positive values (Yin and Chan, 1990; Pecka *et al.*, 2008). Best IPDs ϕ_{best} are typically positive (contralaterally leading) and take values smaller than 0.25 (Yin and Chan, 1990; Brand *et al.*, 2002; Pecka *et al.*, 2008). As a consequence of ϕ_c and ϕ_{best}, characteristic delays of MSO cells are biased toward positive values, although a considerable fraction of neurons exhibits negative Δ_c (those with large positive ϕ_c and small ϕ_{best}).

The main constraint on the model discussed in this paper is how well it fits a given phase-frequency curve in a 2/3-octave range of stimulus frequencies centered at best frequency. The fits are furthermore constrained such that they ensure both a sufficient amount of ITD information in the amplitude of the total input current and a minimal inhibitory synaptic conductance.

II. MODEL

A. Preliminary remarks on linearity

The present paper refrains from constructing a detailed biophysical model of the MSO neuron. It is mainly concerned with the synaptic (circuit) parameters and how much of the *in vivo* data on ITD sensitivity they are able to explain. The fundamental assumption is that ITD sensitivity of a neuron is reflected by the ITD dependence of the amplitude A of the cycle-by-cycle fluctuations in the total synaptic input, i.e., the firing probability positively correlates with this input amplitude A. Two possible physiological mechanisms of MSO neurons may account for this assumption, low-threshold activated potassium channels (KLT) (Svirskis *et al.*, 2002) and HCN channels (h-currents) (Koch *et al.*, 2004): If the synaptic input is small, a slow rise of the membrane voltage evokes the opening of the KLT channels, which inhibits spiking owing to the outflux of potassium. For large synaptic inputs, the membrane voltage may rise faster than the opening of the KLT channels and finally cross the sodium activation threshold to evoke an action potential. The h-currents, which are too slow to follow the ac component in the input, may counterbalance the dc component of the inhibitory synaptic current. Therefore, the slow components of the inputs possibly do not influence firing probability.

In addition to the cellular mechanisms, the present modeling approach also neglects all kinds of non-linearities up-

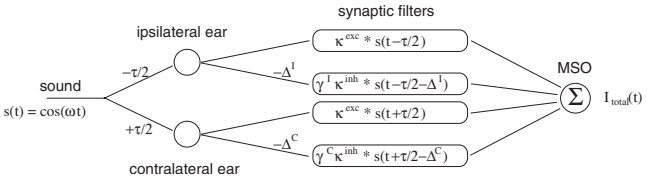


FIG. 2. Linear superposition model. A pure tone $s(t)$ arrives at the two ears with an ITD τ. Both ears project to the coincidence detector neurons in the MSO via two pathways, an excitatory one and an inhibitory one. The temporal delay of inhibition with respect to excitation is denoted by the Δs, the relative strengths of the inhibitory inputs are given by parameters γ. In addition, distinct synaptic filtering induces a pathway-specific convolution with a kernel κ. The total current $I_{\text{total}}(t)$ at a MSO cell is modeled as a linear superposition of the currents from the four pathways.

stream of the MSO, such as compression and half-wave rectification in the periphery, temporal jitter in the auditory nerve, the biophysics of the bushy cells in the antero-ventral cochlear nucleus, as well as the MNTB and LNTB neurons. However, the pathways from the ear to the MSO are subject to several instances of low-pass filtering which, as in every periodically driven dynamical system, amplify the fundamental frequency as compared to its harmonics. Thus, to understand the summation of different inputs in terms of their major components, one has to primarily study the summation of sinusoidal oscillations. All the results in this paper are hence to be seen as first-order approximations in this sense.

B. Linear superposition of sine waves

The simplest model to understand the summation of the four MSO input pathways is a linear superposition $I_{\text{total}}(t)$ of the respective synaptic currents. The currents are considered to be evoked by a pure tone $s(t)=\cos(\omega t)$ with frequency $f=\omega/(2\pi)$. The contributions from the both ears split up into ipsilateral (I) and contralateral (C) ones,

$$I_{\text{total}}(t) = I^I(t) + I^C(t),$$

in which the monaural parts $I^{I/C}$ are again sums of two pathways, excitation and inhibition (Fig. 2 and Appendix A). As shown in Appendix A, the summation of the four input pathways can be written as

$$\begin{aligned} I_{\text{total}}(t) = {} & \cos[\omega(t+\tau/2)] + \cos[\omega(t-\tau/2)] \\ & - \Gamma^C(\omega)\cos[\omega(t+\tau/2) - \delta^C(\omega)] \\ & - \Gamma^I(\omega)\cos[\omega(t-\tau/2) - \delta^I(\omega)]. \end{aligned} \tag{2}$$

The first two cosines reflect excitatory synaptic currents, and the second two cosines model phase-locked inhibition. The ipsi- and contralateral inhibitory synaptic strengths relative to excitation are denoted with $\Gamma^{I/C}(\omega)$ and the respective relative phase delays are called $\delta^{I/C}(\omega)$. Both the relative strengths of inhibition $\Gamma^{I/C}(\omega)$ and the phase shifts $\delta^{I/C}(\omega)$ between inhibition and excitation are functions of the stimulus frequency, see Eqs. (A5) and (A6). These functions are derived from the synaptic kinetics $\kappa(t)$ which is modeled according to

$$\kappa(t) \propto \begin{cases} (e^{-t/\tau_1} - e^{-t/\tau_2}), & t \geq 0 \\ 0, & t < 0. \end{cases}$$

For details see Eq. (A3) in Appendix A. The relative strengths $\Gamma^{I/C}(\omega)$ incorporate low-pass filtering of oscillation amplitudes. Their frequency independent scaling factors are denoted by $\gamma^{I/C}$, such that

$$\Gamma^{I/C}(\omega) = \gamma^{I/C} F(\omega),$$

as derived in Eq. (A5). The inhibitory phase shifts $\delta^{I/C}(\omega)$ are induced via the different synaptic kinetics and the additional inhibitory time delays $\Delta^{I/C}$ (negative delays are interpreted as advance). To reduce the number of free parameters, synaptic kinetics are assumed to be identical for both ipsi- and contralateral inputs; however, they are different for excitation and inhibition. The ITD is denoted by τ, where $\tau>0$ points to a contralaterally leading sound and $\tau<0$ to an ipsilaterally leading sound. In Appendix A, it is shown that the stimulus-induced oscillation $I_{\text{total}}(t)$ has an amplitude

$$A(\tau) = \sqrt{(A^I)^2 + (A^C)^2 + 2A^I A^C \cos(\psi^C - \psi^I + \omega\tau)}, \tag{3}$$

with

$$A^{I/C} = 1 + (\Gamma^{I/C})^2 - 2\Gamma^{I/C}\cos(\delta^{I/C}) \tag{4}$$

and

$$\psi^{I/C} = -\delta^{I/C}/2 + \arg[\exp(i\delta^{I/C}/2) - \Gamma^{I/C}\exp(-i\delta^{I/C}/2)].$$

Assuming the firing rate to be proportional to the amplitude $A(\tau)$, according to Eq. (3), the best IPD equals

$$\phi_{\text{best}} = \omega\tau_{\text{best}} = \psi^I - \psi^C. \tag{5}$$

By analogy, the minimum of A occurs at

$$\phi_{\min} = \phi_{\text{best}} \pm \pi.$$

An example of the total synaptic current $I_{\text{total}}(t)$ of the linear superposition model is illustrated in Fig. 3. Here the best IPD increases with frequency; i.e., the characteristic delay Δ_c is positive. The four parameters $\Delta^{I/C}, \gamma^{I/C}$ have been determined by the fitting procedure from Appendix B to obtain a characteristic phase of $\phi_c=0.145$ cycles and a characteristic delay of $\Delta_c=100$ μs. Since the model was fitted for a best frequency of $f_b=800$ Hz, the best phase of the example neuron is $\Phi_{\text{best}}=\phi(f_b)=\phi_c+f_b\Delta_c=0.125$ cycles.

C. Overview on the fitting procedure

Here it is briefly sketched how the four fit parameters $\Delta^{I/C}, \gamma^{I/C}$ are obtained. A more detailed description can be found in Appendix B.

A naive fitting approach in the full four-dimensional parameter space did not work out, since the root mean squared error E_ϕ of the linear fit $\phi(f)$ to the phase-frequency curve turned out to have many local minima. Moreover, the best fit in terms of E_ϕ is maybe also not a very physiological objective, since there is no genuine reason for perfect linearity of the phase-frequency curve. Fitting was therefore done in a two-step process.

In a first step, the two delay parameters $\Delta^{I/C}$ were obtained from a conventional fit (minimizing E_ϕ) of $\phi(f)$ to the

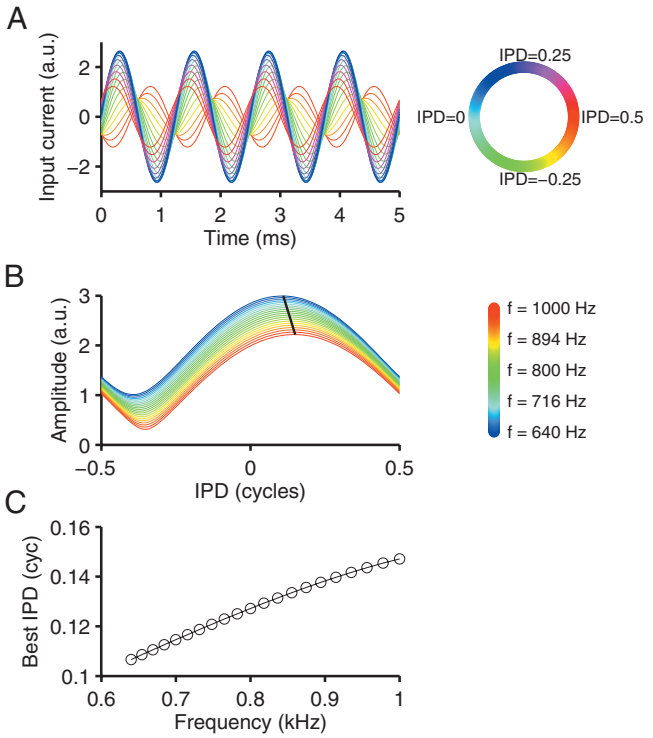


FIG. 3. IPD sensitivity of a linear superposition model. (A) Total input current I_{total} as a function of time for a stimulus frequency of $f=800$ Hz and different IPDs (colors see right). The parameters $\gamma^C=1.20$, $\gamma^I=0.59$, $\Delta^C=0.58$ ms, $\Delta^I=0.04$ ms were obtained by the optimization process given in Appendix B for desired characteristic phase $\phi_c=0.045$ cycles and characteristic delay $\Delta_c=0.1$ ms (and hence a best phase of 1/8 cycles). Synaptic kinetics had time constants $\tau_1^{\text{exc}}=0.1$ ms, $\tau_2^{\text{exc}}=0.2$ ms, $\tau_1^{\text{inh}}=0.25$ ms, and $\tau_2^{\text{inh}}=1.5$ ms. The phase inversion of the time course as the stimulus IPD passes through 0.5 cycles is due to the fact that the oscillation of the total input is periodic with frequency $f/2$, whereas the amplitude A is periodic with frequency f. This is explained for a simple example in Appendix A. (B) Amplitude of the input current oscillations from (A) as a function of the IPD for different input frequencies f (colors see right). Black line marks the best IPDs at which the amplitude is maximal. (C) Phase-frequency curve: Best IPD as a function of stimulus frequency.

phase-frequency curve for a fixed pair of strength parameters $\gamma^{I/C}$. The fit was further constrained by a minimal coding range A_0; i.e., the maximal difference in amplitudes $A(\tau_{\text{best}})-A(\tau_{\text{min}})$ obtained via Eq. (3) had to be at least as large as some value A_0. This constraint was applied to rule out solutions for which the input amplitude A bears only little information in the ITD. In this paper, calculations were being done for a minimal range of $A_0=1$.

In a second step, the summed inhibitory strength $E_\gamma=\gamma^I+\gamma^C$ was minimized under the constraint that the root mean squared fitting error E_ϕ is smaller than some bound E_0. This minimization step can be motivated by energy efficiency. The constraint ensures that the phase-frequency curve is sufficiently well approximated by a linear function as it is usually observed in MSO cells. As a bound on linearity, I used 5% of a typical best IPD, i.e., $E_0=0.05\times0.1$ cycles $=0.005$ cycles.

III. OPTIMAL CIRCUITS

A. Four pathways

As a result of the two-step optimization procedure introduced in Appendix B, one is able to generate models for a

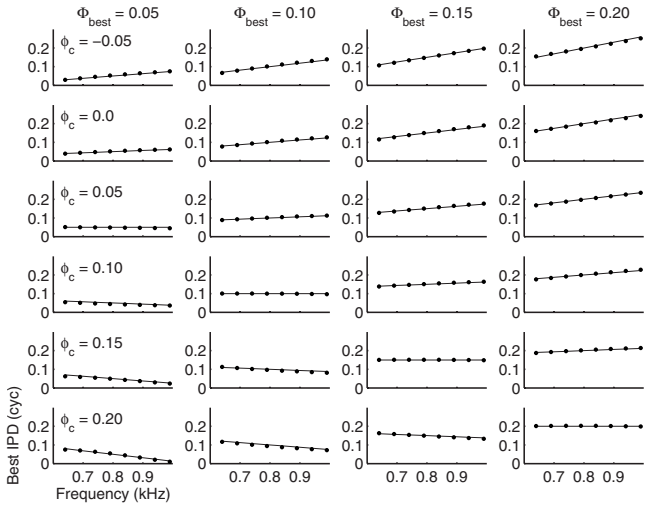


FIG. 4. Phase-frequency curves for given characteristic phase ϕ_c (rows) and best phases Φ_{best} (columns). The dots are derived from the linear superposition model using the set of fit parameters obtained by the optimization process described in Appendix B. The lines indicate the best linear fit. The best frequency in all of these examples was 800 Hz.

desired pair of characteristic phase and delay (Fig. 4). As pointed out in appendix B the fit works for characteristic phases $\lessapprox 0.25$. The four fit parameters $\Delta^{I/C}$, $\gamma^{I/C}$ that are best to achieve a given phase-frequency curve define the respective optimal circuits. Examples for such optimal circuits are given in Fig. 5 for four different best phases (0.05, 0.1, 0.15, and 0.2 cycles) and a best frequency of 800 Hz. The optimization results allow the following five conclusions.

(1) For all considered phase-frequency curves, the optimization process converges within the given acuity E_0 (see Appendix B). This includes positive and negative values for characteristic phases and delays.
(2) For low ϕ_c the best fits are in the contralaterally dominated regime $\gamma^C>\gamma^I$ [Fig. 5(B)]; i.e., contralateral inhibition is stronger than ipsilateral inhibition. For higher ϕ_c, the best fits are in the ipsilaterally dominated regime $\gamma^C<\gamma^I$. The point of transition between the two regimes increases with the best IPD.
(3) The total inhibitory strength E_γ increases with the best phase Φ_{best} [solid line in Fig. 5(B)], i.e., the further the peak is required to shift toward positive IPDs, the more inhibition is necessary.
(4) The optimal delay Δ^I of the ipsilateral inhibition is relatively small (≈ 0) for all considered conditions [Fig. 5(C), dashed lines].
(5) The optimal contralateral delay Δ^C of the contralateral inhibition is mostly relatively large (around 0.5 ms) and rather independent of ϕ_c and Δ_c [Fig. 5(C), dots]. Note that for some values of ϕ_c the optimal contralateral inhibition is zero. At those values the delay parameter Δ^C is omitted.

As shown in Fig. 6, the overall situation does not change much if faster inhibitory kinetics of $\tau_2^{\text{inh}}=0.5$ ms are considered. The most obvious difference as compared to the situation with $\tau_2^{\text{inh}}=1.5$ ms is that the required total inhibitory amplitude is somewhat smaller, reflecting the "low-pass" fil-

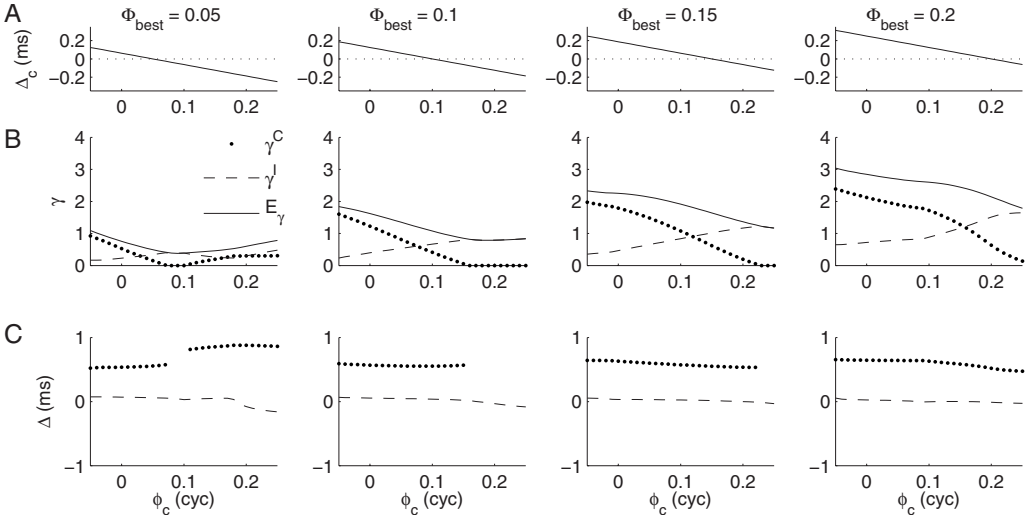


FIG. 5. Optimal circuits. (A) Characteristic delay obtained from $\Delta_c=(\Phi_{\text{best}}-\phi_c)/(2\pi 800\ \text{Hz})$. Dotted line depicts $\Delta_c=0$. (B) Inhibitory strength $\gamma^{I/C}$ and total strength E_γ. (C) Inhibitory delays Δ^I (dashed lines) and Δ^C (dots).

tering through the inhibitory synaptic kinetics. In addition, for small Φ_{best} and negative characteristic delays, there appears a new regime with contralaterally leading inhibition $\Delta^C<0$.

1. On inhibitory delays

From the above optimal circuits, one can derive mechanistic ideas of how the synaptic parameters give rise to phase-frequency curves. Here, let us concentrate on the situation that is most frequently considered (McAlpine and Grothe, 2003; Joris and Yin, 2006; Jeffress, 1948; McAlpine *et al.*, 2001; Harper and McAlpine, 2004) for which best phases are between 0.1 and 0.15 cycles and characteristic phases are small ($\lessapprox$0.05 cycles) such that the characteristic delay has to be positive $\Delta_c>0$. In this case the optimal circuits arise from the contralaterally dominated regime $\gamma^C>\gamma^I$ (see Figs. 5 and 6). Let us first consider the contralateral pathways. A positive time delay $\Delta^C\approx+0.5$ ms of the contralateral inhibition translates to a phase delay $\Delta^C\omega$, which is positively correlated with frequency (Fig. 7, left). Adding the relative phase delay induced by synaptic filtering further shifts the peaks toward positive phases such that contralateral inhibition effectively precedes excitation of the next cycle. Again the phase shift is positively correlated with frequency.

For the ipsilateral pathways, the best fits predict a small positive delay $\Delta^I\approx0.05$ cycles. The additional phase delay from synaptic filtering shifts the ipsilateral inhibition such that the *total* ipsilateral current peaks at an earlier phase than the excitation (Fig. 7, right). The amplitude and phase of the ipsilateral current are almost independent of frequency.

The best IPD follows from the sum of the ipsi- and the contralateral currents. This sum current has maximal amplitude if the ipsilateral current is delayed with respect to the contralateral one by a positive IPD ϕ_{best}. Due to the frequency-dependent phase shift of the contralateral currents, the best IPD ϕ_{best} increases with stimulus frequency ω; i.e., the characteristic delay Δ_c is positive. The basis for this positive characteristic delay is a strongly delayed contralateral

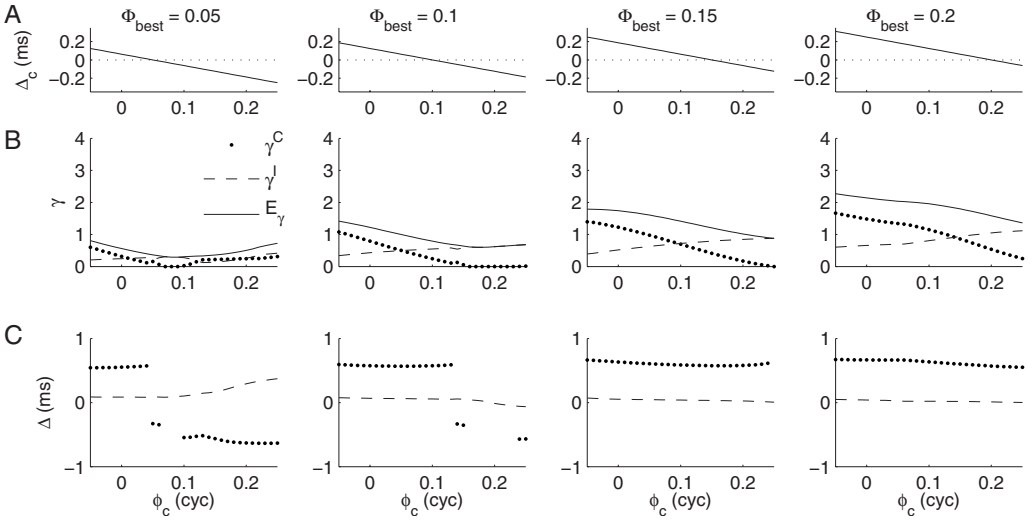


FIG. 6. Optimal circuits. Same as Fig. 5 for $\tau_2^{\text{inh}}=0.5$ ms.

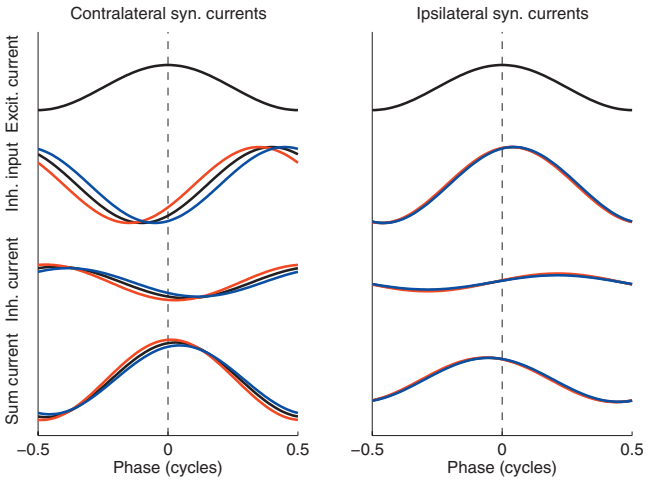


FIG. 7. How inhibitory time delays shape synaptic currents (left: contralateral; right: ipsilateral). Excitatory currents: Phase zero is defined as the peak. Inhibitory inputs: A temporal delay Δ induces frequency-dependent phase delays (red: 700 Hz; black: 800 Hz; blue 900 Hz). Inhibitory currents: synaptic filtering induces a frequency-dependent amplitude and phase change. The relative amplitudes are chosen $\gamma^C=1$ and $\gamma^I=0.5$ according to Fig. 5 for $\Phi_{best}=0.15$ cycle and $\phi_c\approx 0.05$. Sum current: the sum of the excit. and negative inh. currents (first and third rows).

inhibition that effectively leads the next cycle of excitatory currents.

2. Influence of best frequency

All previous results in this paper were derived for an intermediate best frequency of 800 Hz. To rule out that the four-pathway model breaks down for particularly low or high best frequencies, optimal circuits have also been derived for best frequencies between 200 and 1200 Hz (Fig. 8). The fitting process described in Appendix B identifies optimal parameter regimes of the linear four-pathway model for all tested best frequencies. The best frequency dependence of the optimal parameter is gradual with few exceptions, which argues for the robustness of the optimization procedure. As expected, the (total) relative inhibitory strength increases with best frequency (owing to the low-pass filtering by the

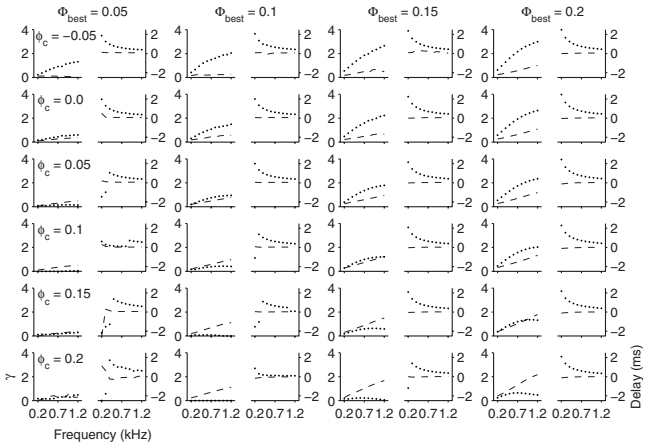


FIG. 8. Influence of best frequency. Relative inhibitory strengths γ (graphs with y axis on the left) and delays (y axis on the right) for several values of ϕ_c (rows) and Φ_{best} (principal columns). Contralateral parameters are depicted as dots, and ipsilateral parameters by dashed lines.

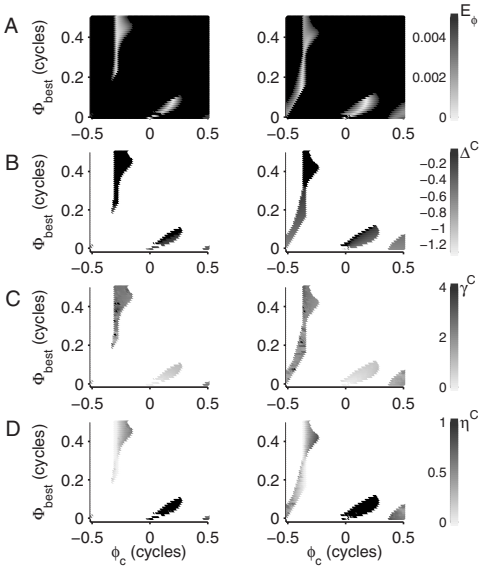


FIG. 9. Optimal circuits for three input pathways. Left column: $\tau_2^{inh}=1.5$ ms; right column: $\tau_2^{inh}=0.5$ ms. (A) Fitting error E_ϕ as a function of Φ_{best} and ϕ_c. Regions for which $E_\phi>0.005$ cycle are marked in black. (B) Optimal inhibitory delays Δ^C (in ms). (C) Optimal inhibitory strengths γ^C. (D) Optimal contralateral strength η^C of excitation. In (B)–(D), regions for which $E_\phi>0.005$ cycle are marked in white.

inhibitory kinetics), whereas the relative inhibitory delays generally decrease with frequency, at least their moduli. For the lowest tested best frequency (200 Hz), the optimal inhibitory delays may become as large as 2.5 ms. A delay of that magnitude is quite long and it is unclear whether differences in axonal geometry and myelination may account for it. An alternative to effectively generate this delay would be to slow down all synaptic kinetics for low frequencies. Tonotopic gradients of synaptic time constants, however, have not yet been reported in the MSO. It thus might well be that for low best frequencies ($\lessapprox$400 Hz), the ITD tuning is shaped by additional mechanisms other than phase-locked inhibition, for example asymmetries in the excitatory pathways.

B. Restriction to three pathways

To study whether optimal circuits can also be fitted with only one inhibitory pathway, next optimal circuits are derived without ipsilateral inhibition (via the LNTB). This pathway is generally assumed to be less prominent in the MSO than contralateral inhibition (via the MNTB). Technically this restriction was achieved by setting $\gamma^I=0$. In addition, the fitting algorithm is also allowed to change the amplitude η^C of contralateral excitation within the bounds $0\leqslant\eta^C\leqslant 1$. The latter degree of freedom is of interest because it allows the possibility of EI cells, i.e., neurons that receive excitation from ipsilateral and only inhibition from contralateral. Such neurons are found in the low-frequency region of the lateral superior olive (LSO) (Glendenning *et al.*, 1991) and exhibit trough-type phase-frequency curves (Tollin and Yin, 2005), meaning they have a characteristic phases around 0.5 cycles.

In contrast to four pathways, the three-pathway model is simple enough for the optimization to be done at once with the remaining three parameters η^C, γ^C, and Δ^C. Figure 9

shows that for both $\tau_2^{inh}=1.5$ ms and $\tau_2^{inh}=0.5$ ms the range of best IPDs and characteristic phases for which a sufficiently small E_ϕ can be achieved is rather limited as compared to the situation with four pathways (Fig. 5).

At characteristic phases smaller than 0.25 cycles (the so-called "peaker-type" neurons) fits only work for small best phases (<0.1 cycles) and negative characteristic delays Δ_c. As a consequence, also $\Delta^C<0$, i.e., contralateral inhibition is advanced with respect to contralateral excitation. The latter regime is in line with traditional models of inhibitory-based ITD sensitivity (Brand *et al.*, 2002; Svirskis *et al.*, 2003; Leibold and van Hemmen, 2005) that have not been checked for their phase-frequency curves; however, the present three-pathway model has clear limits to the maximal best phase (for peakers) as compared to the model with four input pathways (Fig. 5) if one enforces a linear phase-frequency curve. Figure 9 also shows that ipsilateral inhibition is necessary for peakers to achieve positive characteristic delays, although there the circuit is in a contralaterally dominated regime $\gamma^C > \gamma^I$.

In addition to the limited peaker regime, the three-pathway fits also reveals a so-called "trougher-type" regime with the modulus of the characteristic phase larger than 0.25 cycles (Marquardt and McAlpine, 2006). There, the contralateral excitation η^C is low, which is in line with the situation of EI cells of the low-frequency LSO. The fit reveals an optimal contralateral inhibitory delay $\Delta^C<0$; i.e., for these troughers the contralateral inhibition should lead.

IV. DISCUSSION

The present paper discusses a linear model for the binaural coincidence detector neurons in the MSO for the situation of ongoing (non-transient) pure tone input. The model takes into account the modulation of interaural phase sensitivity by stimulus frequency, which is expressed by the phase-frequency relation. The quality of fit to the phase-frequency relation is used as a criterion to identify optimal circuits, i.e., to specify the delay and strength parameters of the inhibitory pathways from the two ears. All results were derived under the primary assumption of symmetric bilateral excitation. The main finding is that ipsilateral inhibition is necessary to allow for positive characteristic delays, i.e., an increase in the best IPD with stimulus frequency. Moreover, positive characteristic delays additionally require contralateral inhibition to lag contralateral excitation. If the contralateral inhibition leads excitation, the model predicts negative characteristic delays, i.e., a decrease in the best IPD with stimulus frequency.

Previous models that employ phase-locked inhibition to explain the shift of best ITD assume the contralateral inhibition to lead the contralateral excitation (Brand *et al.*, 2002; Svirskis *et al.*, 2003; Leibold and van Hemmen, 2005). This assumption is consistent with findings (Siveke *et al.*, 2006) showing that backward masking of the contralateral response using short binaural chirp stimuli with ipsilaterally leading ITD. In the present modeling study, the optimal contralateral inhibitory delay of about 0.5 ms (together with the synaptic time constant of 1.5 ms) produces an inhibitory lag that is long enough to act as leading inhibition for the next oscillation cycle. In such a regime, the backward masking (Siveke *et al.*, 2006) may be explainable by a ringing of the cochlear band-pass filter, in which the response to the first excitatory cycle is too small to evoke a response.

For the sake of simplicity, the model is formulated in terms of synaptic currents rather than conductances. From the reversal potentials for excitation (generally assumed to be ≈ 0 mV) (Svirskis *et al.*, 2003; Zhou *et al.*, 2005) and inhibition (≈ -90 mV) (Magnusson *et al.*, 2005), and the resting potential (≈ -55 mV) (Scott *et al.*, 2005), the ratio between excitatory and inhibitory driving forces can be estimated as ≈ 1.5. For a best phase between 0.1 and 0.15 cycles, Fig. 5 predicts a maximal (unilateral) inhibitory current amplitude of about twice the unilateral excitatory current amplitude. This corresponds to a maximal ratio between the (unilateral) inhibitory and excitatory synaptic conductances amplitudes of about $2\times 1.5=3$; i.e., inhibitory pathways should be stronger than excitatory ones. In most cases, however, the optimal (unilateral) inhibitory current amplitudes are below 1.5 times the excitatory currents (Fig. 5), which predicts a conductance ratio between inhibitory and excitatory pathways of $\lessapprox 2$.

The conclusions drawn from the studied model are derived under several assumptions that constrain its validity. The assumptions can be subdivided into those that are imposed by considering the experimental model of the gerbil and those that arise from the applied mathematical framework.

The restriction that arises from constraining the model by gerbil data is that the shift of best IPDs toward positive values must be mediated by inhibition, since it was shown that blocking inhibition (Brand *et al.*, 2002) as well as reducing inhibitory phase locking (Pecka *et al.*, 2008) gives rise to tuning curves which have peaks closer to the midline ($\tau=0$). This finding restricts the model to two symmetric excitatory pathways including symmetric delays across hemispheres between the site of cochlear excitation and the MSO. As a result, the classical Jeffress model (Jeffress, 1948) has been ruled out for the gerbil (at least for best frequencies above 500 Hz) and hence is not part of the present theory. The question of bilateral excitatory delay differences is, however, still unresolved since the geometric distance to the contralateral ear is necessarily longer than that to the ipsilateral ear. Any delay of contralateral excitation would add to the characteristic delay and evidently also to the best ITD. However, the geometric distance alone cannot explain the whole phenomenology, in particular, characteristic phases distinctly different from zero and negative characteristic delays (Pecka *et al.*, 2008). Moreover, the geometry is roughly the same for neurons independent of their best frequency, whereas the best ITDs are negatively correlated to best frequencies (Brand *et al.*, 2002; McAlpine *et al.*, 2001). So even for positive characteristic delays, there must be an additional mechanism that tunes the axonal conduction delays according to their best frequency (Gerstner *et al.*, 1996; Kempter *et al.*, 2001; Leibold and van Hemmen, 2005).

Instead of axonal delays, an effective asymmetry of the excitatory pathways has been shown to be possibly induced

by an asymmetry of the origin of the axon (Zhou *et al.*, 2005). However, such a systematic morphological asymmetry could not be confirmed in a detailed analysis in gerbils (Rautenberg *et al.*, 2009). It remains open to explore, how much the observed morphological variability accounts for the variability in response properties.

Furthermore, there are restrictions to the present results that arise from the employed modeling framework. First, the excitatory synaptic kinetics are assumed to be the same for both the ipsilateral and the contralateral pathway. Allowing for a bilateral disparity of the EPSC kinetics in combination with excitatory delays may also contribute to the best ITD (Jercog *et al.*, 2008). Another possibility that was suggested to induce a bilateral asymmetry of the excitatory phases is via the contributions of mismatched frequencies of the ipsi- and contralateral input fibers such that, on average, ipsi- and contralateral inputs are preprocessed by different cochlear filters (Joris *et al.*, 2006).

Second, phase-frequency curves may not resemble a linear relationship as well as those obtained by the presented optimization procedure. This means the error threshold of $E_\phi > 0.005$ cycles is somewhat arbitrary and real neurons may also reside in a "less optimal" parameter regime.

Third, the model only deals with ongoing stimuli. Real acoustic environments, however, generally comprise many onset-type stimuli. The circuit parameters derived in this study may thus be considered as steady-state values of an adaptive process. The translation from the onset to the ongoing regime is, however, not very well investigated physiologically. Yet, even stimuli with distinct onsets generate oscillatory behavior (ringing) in the cochlear band pass filters.

Fourth, the biophysical properties of the MSO neurons are not accounted for in the model. This is particularly a problem for the "long" inhibitory time constant $\tau_2^{\mathrm{inh}} = 1.5$ ms, which produces a considerable hyperpolarizing dc component. In order to compensate for this hyperpolarization, the neuron requires strong additional inward currents, such as *h*-currents (HCN channels) (Koch *et al.*, 2004). Another current that is prevalent in MSO neurons is the low-threshold potassium current that has a relatively fast activation kinetics and thus influences the transformation of synaptic currents into membrane voltage, particularly for very low frequencies (Svirskis *et al.*, 2002). *In vivo*, MSO neurons are even leakier than under *in vitro* conditions owing to the large synaptic, HCN, and low-threshold potassium conductances that are opened by the synaptic bombardment and the presumed concurrent cellular counteractions. From a biophysical perspective, generation of APs is difficult in such a high conductance state. The spatial segregation between soma and axonal site of AP generation may be essential for AP firing in binaural coincidence detector neurons (Kuba *et al.*, 2006; Scott *et al.*, 2007) *in vivo*.

Fifth, the large glycinergic dc component accounts for a large flux of chloride ions into the cell. The present model thus has to assume that the cell's chloride extrusion copes with this steady-state current. For the rat LSO, the high-frequency counterpart of the MSO with one excitatory and one inhibitory input pathway (via the MNTB), it is known that chloride extrusion becomes quite fast during maturation (Ehrlich *et al.*, 1999). Moreover, since the inhibitory/excitatory conductance ratios of ($\gamma \lessapprox 2$) estimated by the present model seem to be in an experimentally realistic range (Felmy, private communication), the cell obviously has to be able to maintain chloride homeostasis under high-frequency inputs independent of whether the currents provide timing information or not.

To conclude, the present optimization approach allows one to select circuit parameters such that they account for a prescribed frequency dependence of ITD sensitivity. Such an algorithm is necessary, since the "four-dimensional" parameter space is already too complicated to find the right solutions by trial and error. The algorithm thus can be a helpful tool to tune the input parameters for more elaborate biophysical models of binaural coincidence detection in the MSO.

ACKNOWLEDGMENTS

The author is grateful to Michael Pecka and Ida Siveke for many fruitful discussions about characteristic delays and phases and to Marion Sobotka and Lutz Wiegrebe for comments on the manuscript. This work was supported by the Bundesministerium für Bildung und Forschung (BMBF) under Grant No. 01GQ0440 (Bernstein Center for Computational Neuroscience, Munich).

APPENDIX A: MODEL DETAILS

1. Superposition of sine waves

MSO neurons collect synaptic inputs from four pathways, two contralateral (C) ones and two ipsilateral (I) ones (Fig. 2). The inputs from the two ears are shifted with respect to each other by an ITD τ and consist of an excitatory (exc) and an inhibitory (inh) pathway each. To analytically approximate the response of a MSO neuron, a simple linear model is employed that is based on a sum of phase-locked sinusoidal synaptic currents (Batra *et al.*, 1997; Leibold *et al.*, 2002). As stimuli I assume pure tones with frequency $f = \omega/(2\pi)$. The ipsi- and contralateral currents, I^I and I^C, respectively, are thus written as

$$I^{I/C}(t) = I^{I/C}_{\mathrm{exc}}(\omega)\cos[\omega(t \mp \tau/2) + \varphi^{I/C}_{\mathrm{exc}}] - I^{I/C}_{\mathrm{inh}}(\omega)\cos[\omega(t \mp \tau/2) + \varphi^{I/C}_{\mathrm{inh}}(\omega)].$$

Both pathways involve a frequency-dependent amplitude attenuation $I^{I/C}_x(\omega)$ and a frequency-dependent phase shift $\varphi^{I/C}_x(\omega)$, $(x \in \{\mathrm{exc}, \mathrm{inh}\})$.

In what follows, it is assumed that the excitatory pathways are bilaterally symmetric such that $I^{I/C}_{\mathrm{exc}}(\omega) = I_{\mathrm{exc}}(\omega)$ and $\varphi^{I/C}_{\mathrm{exc}}(\omega) = \varphi_{\mathrm{exc}}(\omega)$. One therefore can remove four degrees of freedom from the model, by considering the input traces relative to the excitatory phase and amplitude.

$$\tilde{I}_{\mathrm{total}}(t) = I_{\mathrm{total}}[t - \varphi_{\mathrm{exc}}(\omega)/\omega][I_{\mathrm{exc}}(\omega)]^{-1}.$$

In the main text and in what follows, it is not distinguished between the relative input $\tilde{I}_{\mathrm{total}}$ and the total input I_{total} since they differ by only trivial factors. Moreover, to simplify notation, all the tildes are omitted, though the total current is

always meant to be the relative one. Introducing relative inhibitory amplitudes $\Gamma^{I/C}(\omega)=I_{\mathrm{inh}}^{I/C}(\omega)/I_{\mathrm{exc}}(\omega)$ and phase delays $\delta^{I/C}(\omega)=\varphi_{\mathrm{exc}}(\omega)-\varphi_{\mathrm{inh}}^{I/C}(\omega)$, the input is then computed as

$$I_{\mathrm{total}}(t)=\cos[\omega(t+\tau/2)]-\Gamma^{C}(\omega)\cos[\omega(t+\tau/2)-\delta^{C}(\omega)]+\cos[\omega(t-\tau/2)]-\Gamma^{I}(\omega)\cos[\omega(t-\tau/2)-\delta^{I}(\omega)].$$

This formula can be simplified using the identity for the sum of two cosine waves

$$a_1\cos(x+\xi)+a_2\cos(x-\xi)=\sqrt{a_1^2+a_2^2+2a_1a_2\cos(2\xi)}\cos(x+\arg z), \tag{A1}$$

in which

$$z=a_1e^{i\xi}+a_2e^{-i\xi}.$$

Combining excitation and inhibition from one ear, Eq. (A1) yields

$$\cos[\omega(t\pm\tau/2)]-\Gamma\cos[\omega(t\pm\tau/2)-\delta]=\sqrt{1+\Gamma^2-2\Gamma\cos\delta}\cos[\omega(t\pm\tau/2)-\delta/2+\arg(e^{i\delta/2}-\Gamma e^{-i\delta/2})]. \tag{A2}$$

Applying Eq. (A1) a second time to the sum of ipsilateral and contralateral inputs, i.e., Eq. (A2), results in Eq. (3) given in Sec. II.

a. Linear superposition model for two pathways

To illustrate the present approach for the simplest coincidence detection circuit, the situation is discussed for a hypothetic neuron that receives two symmetric excitatory pathways. Given an ITD τ, the ipsilateral pathway provides a signal $\cos[\omega(t-\tau/2)]$ and the contralateral pathway gives rise to $\cos[\omega(t+\tau/2)]$. Summing up both contributions yields the total input I_{total} a function of time t as

$$I_{\mathrm{total}}(t)=\cos[\omega(t-\tau/2)]+\cos[\omega(t+\tau/2)]=\sqrt{2+2\cos(\omega\tau)}\cos(\omega t+\alpha).$$

The first factor on the right-hand side determines the (positive) amplitude of the combined signal, which depends on the ITD τ and has a period of $T=1/f$. The second factor describes the oscillation in time, with the phase

$$\alpha=\arg(z)\quad\text{and}\quad z=\cos(\omega\tau/2).$$

Thus α vanishes for $|\tau|<T/2$ ($\cos\omega\tau/2>0$) and equals π for $T/2<|\tau|<T$ ($\cos\omega\tau/2<0$).

The best ITD is obtained where the amplitude factor $\sqrt{2+2\cos(\omega\tau)}$ is maximal, i.e., for $\tau=0$. Because $\cos(x)=2\cos^2(x/2)-1$, the amplitude factor can also be rewritten as

$$\sqrt{2+2\cos(\omega\tau)}=2|\cos(\omega\tau/2)|,$$

such that one obtains the somewhat simpler equivalence

$$\cos[\omega(t-\tau/2)]+\cos[\omega(t+\tau/2)]=2\cos(\omega\tau/2)\cos(\omega t).$$

Note that though this function is obviously periodic in τ with period $2T$, its envelope (=amplitude or response rate) is periodic with period T.

2. Synaptic kinetics

The synaptic currents are modeled by the kernel

$$\kappa(t)=\begin{cases}I_0\dfrac{N(\tau_1,\tau_2)}{\tau_1-\tau_2}(e^{-t/\tau_1}-e^{-t/\tau_2}), & t\geq 0\\ 0, & t<0,\end{cases} \tag{A3}$$

which is parametrized by two time constants $\tau_1\neq\tau_2$ that determine the rise and decay times. The normalization factor

$$N(\tau_1,\tau_2)=\frac{\tau_2^{\tau_2/(\tau_2-\tau_1)}}{\tau_1^{\tau_1/(\tau_2-\tau_1)}}$$

is chosen such that $\bar{I}=\max_t\kappa(t)$ denotes the current amplitude. Note that for this normalization it does not matter whether $\tau_1>\tau_2$ or vice versa.

The effective synaptic currents are modeled as a convolution of the sinusoidal inputs with the synaptic kernels κ. The further analysis is therefore done in the frequency (Fourier) domain, in which the synaptic kernel reads

$$\hat{\kappa}(\omega)=\bar{I}\frac{N(\tau_1,\tau_2)}{(1+i\omega\tau_1)(1+i\omega\tau_2)}. \tag{A4}$$

To study the interaction of inhibition and excitation, one has to introduce two time constants for excitation $\tau_{1/2}^{\mathrm{exc}}$ and two time constants for inhibition $\tau_{1/2}^{\mathrm{inh}}$ modeling rise and decay of the respective synaptic currents. Here, the time constants are chosen $\tau_1^E=0.1$ ms, $\tau_2^E=0.2$ ms (Scott *et al.*, 2005), $\tau_1^I=0.25$ ms, $\tau_2^I=1.5$ ms (Magnusson *et al.*, 2005), unless otherwise specified.

The relative strength $\Gamma(\omega)$ between inhibitory and excitatory-mediated input currents is thus given by the frequency-dependent ratio of synaptic currents.

$$\Gamma(\omega)=\frac{\bar{I}_{\mathrm{inh}}}{\bar{I}_{\mathrm{exc}}}\left|\frac{N(\tau_1^{\mathrm{inh}},\tau_2^{\mathrm{inh}})(1+i\omega\tau_1^{\mathrm{exc}})(1+i\omega\tau_2^{\mathrm{exc}})}{N(\tau_1^{\mathrm{exc}},\tau_2^{\mathrm{exc}})(1+i\omega\tau_1^{\mathrm{inh}})(1+i\omega\tau_2^{\mathrm{inh}})}\right|, \tag{A5}$$

which reflects low-pass filtering by the synaptic kinetics. The ratio of peak currents is abbreviated by

$$\gamma=\frac{\bar{I}_{\mathrm{inh}}}{\bar{I}_{\mathrm{exc}}}$$

and serves as a circuit (fit) parameter.

Besides amplitude information, the complex nature of $\hat{\kappa}(\omega)$ also allows to obtain the relative phase shift $\delta(\omega)$ between inhibition and excitation. The latter includes two factors: (1) the synaptic phase shift δ^{syn} given by the complex argument of $\hat{\kappa}^I(\omega)/\hat{\kappa}^E(\omega)$ and (2) an additional temporal delay Δ due to differences in conduction times. Together the phase shift is

$$\delta(\omega)=\omega\Delta-\underbrace{\arg\left(\frac{(1+i\omega\tau_1^{\mathrm{inh}})(1+i\omega\tau_2^{\mathrm{inh}})}{(1+i\omega\tau_1^{\mathrm{exc}})(1+i\omega\tau_2^{\mathrm{exc}})}\right)}_{\delta^{\mathrm{syn}}}. \tag{A6}$$

The minus sign in Eq. (A6) reflects the convention that de-

layed inhibition yields positive phase, whereas advanced inhibition yields negative phase.

APPENDIX B: PARAMETER IDENTIFICATION

To be able to construct a neuron model for a given *pair* of parameters characteristic phase ϕ_c and delay Δ_c, one requires a method to appropriately choose the *four* circuit parameters $\Delta^{I/C}$, $\gamma^{I/C}$. Since this problem is underdetermined, one needs to employ additional constraints. Here it is shown how these constraints can be implemented in a hierarchy of optimization steps.

1. Fitting the phase-frequency-frequency curve

The primary objective (error) function E_ϕ is the root mean square difference between the phase-frequency curve $\phi_{\text{best}}(\omega)$ from the superposition model and the linear approximation $\phi_c+\Delta_c\omega$,

$$E_\phi = \left[|\mathcal{B}_c|^{-1} \int_{\mathcal{B}_c} d\omega (\phi_c + \Delta_c\omega - \phi_{\max}(\omega))^2 \right]^{1/2}. \quad \text{(B1)}$$

The error E_ϕ is an average across the frequency range $\mathcal{B}_c=[\omega_-,\omega_+]$ to which the neuron is responsive. For the Mongolian gerbil, the critical band has been estimated about half an octave (for best frequencies around $f=1$ kHz) (Schmiedt, 1989; Kittel *et al.*, 2002). The present model uses a slightly simpler estimate for the frequency range with limiting frequencies $\omega_+=1.25\omega_0$ and $\omega_-=\omega_0/1.25$, in which ω_0 is the best frequency of the neuron. Minimizing the error E_ϕ thus ensures a faithful approximation of the phase-frequency curve. A good fit to $\phi(\omega)$ is, however, useless if the amplitude difference $\Delta A=A(\phi_{\text{best}}/\omega)-A(\phi_{\min}/\omega)$ of the subthreshold oscillations is too small. The IPD of a stimulus is encoded in the amplitude difference, which, according to Eq. (3), is computed as

$$\Delta A = (\sqrt{(A^I+A^C)})^2 - (\sqrt{(A^I-A^C)})^2 = A^I + A^C - |A^I - A^C|$$
$$= 2\min(A^I, A^C).$$

To achieve a detectably large amplitude difference ΔA, the minimization of E_ϕ is additionally constrained by the equality

$$C(\gamma^{I/C}, \Delta^{I/C}) = \int_{\mathcal{B}_c} d\omega \mathcal{H}[A_0 - \Delta A(\omega)] = 0, \quad \text{(B2)}$$

in which $\mathcal{H}(x)=[\text{sign}(x)+1]/2$ is the Heaviside step function and A_0 is the minimum detectable amplitude difference. If the amplitude difference ΔA is larger than or equal to A_0, the cost function $C=0$, if ΔA is smaller than A_0 then $C>0$. The minimal difference A_0 is a parameter of the optimization problem.

To minimize E_ϕ in the Δ^I-Δ^C space the MATLAB function fminsearch was employed. The best fit was independent of the initial values for $\Delta^{I/C}$ as long as they were in the interval between -1 and $+0.1$ ms. For larger initial values the algorithm converged to suboptimal local minima. The constraint from Eq. (B2) was implemented by adding a quadratic penalty of $50C^2$ to the objective function E_ϕ to obtain smooth behavior at the constraint boundaries.

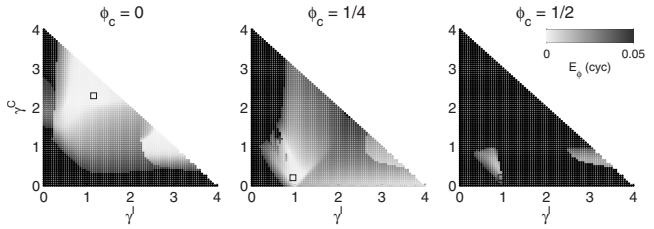


FIG. 10. Fitting error for fixed $\gamma^{I/C}$. The fitting error E_ϕ is plotted as a function of the inhibitory strengths $\gamma^{I/C}$ for three characteristic phases ϕ_c (see titles). Here, the best phase was $\Phi_{\text{best}}=0.125$ cycle, the best frequency was $\omega_0=2\pi 800$ Hz, and the total inhibitory strength $\gamma^I+\gamma^C$ is limited to values smaller than 4. The black squares depict the position of minimum error.

Figure 10 shows the fitting error of the constrained optimization problem, where only the time delays $\Delta^{I/C}$ were taken as fit parameters and the relative strengths $\gamma^{I/C}$ were held fixed during optimization. As an example, the best phase Φ_{best} was taken 1/8 of a cycle, a typical value reported in the literature (Brand *et al.*, 2002; Pecka *et al.*, 2008; McAlpine *et al.*, 2001). For characteristic phase $\phi_c=0$ (and thus $\Delta_c=0.156$ ms), there are large regions in the (γ^I,γ^C) space in which the fitting error is similarly small. There is both a contralaterally dominated $(\gamma^C>\gamma^I)$ and an ipsilaterally dominated $(\gamma^C<\gamma^I)$ regime in which the given phase-frequency curve can be realized. For $\phi_c=0.25$ $(\Delta_c=-0.156$ ms), the region with low fitting error is smaller and mainly ipsilaterally dominated $(\gamma^I>\gamma^C)$. For even larger characteristic phases $\phi_c=0.5$ $(\Delta_c=-0.469$ ms), the fits never achieve an error E_ϕ below 0.01 cycles, which already corresponds to an average error of about 10% given the typical best phases (0.1 cycles) from the literature. Thus for such high ϕ_c, the model with four input paths and symmetric excitation only gives sub-optimal fits.

2. Minimizing total inhibitory input

Figure 10 shows that the error function E_ϕ in general has multiple local minima. Therefore, to reduce the number of optimal solutions for the fitting parameters $\Delta^{I/C}$, $\gamma^{I/C}$, another error function is introduced, which is the total inhibitory input

$$E_\gamma = \gamma^I + \gamma^C.$$

This cost function is motivated by two considerations: (i) synaptic transmission generates metabolic costs. Among all solutions it is thus reasonable to choose the one with the smallest energy use. (ii) Optimization results for a fixed E_γ show that a reasonably small fitting error $E_\phi<0.005$ cycles is reached for a finite value of E_γ (Fig. 11). The bottom panel of Fig. 11 reveals that for $\phi_c=0$, the minimal fitting error concurs with a contralaterally dominated regime, whereas for $\phi_c=1/4$ it concurs with an ipsilaterally dominated regime. For intermediate $\phi_c=1/8$ good fits can be achieved with both ipsi- and contralaterally dominating inhibitions.

3. Synthesis

The four fit parameters are obtained by the following procedure: For each pair of values E_γ and γ^C, one identifies

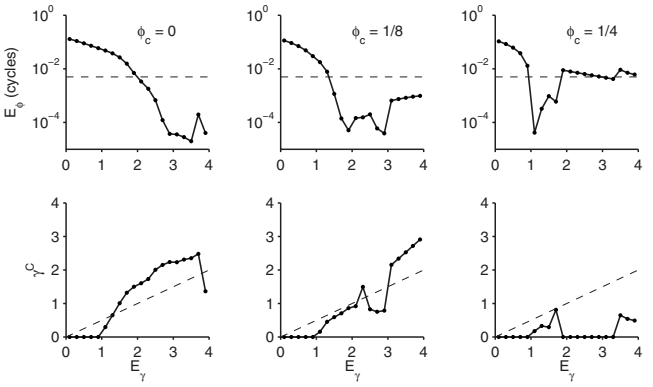


FIG. 11. Minimum total inhibitory input. Top: Fitting error E_ϕ as a function of the fixed total inhibitory input E_γ for three characteristic phases ϕ_c and a best phase $\Phi_{best}=0.125$ cycle. Dashed line marks an error threshold of 0.005 cycles. Bottom: Corresponding contralateral inhibitory strength γ^C. The ipsilateral inhibitory strength can be obtained from $\gamma^I=E_\gamma-\gamma^C$. Dashed line indicates balance of ipsi- and contralateral inhibitions.

the delays $\Delta^{I/C}$ by the constrained minimization of E_ϕ [Eqs. (B1) and (B2)], i.e., fitting the phase-frequency curve. Next, the minimum of the phase error E_ϕ is determined by varying γ^C while leaving E_γ constant. Here the possible values for γ^C are subject to the constraints that both $\gamma^{I/C}$ are non-negative (otherwise the pathways would be excitatory). Finally, E_γ (and hence γ^I) is chosen as the minimal E_γ for which the smallest phase error E_ϕ is below a given accuracy threshold E_0 (dashed line in Fig. 11 top panel). The whole fitting procedure thus has two parameters, the minimal signal A_0 (from minimizing E_ϕ) and the error threshold E_0. In this paper, all numerical results are derived with $A_0=1$ and $E_0=0.005$ cycles, the latter corresponding to about 5% of the typical best IPDs (0.1 cycles).

Batra, R., Kuwada, S., and Fitzpatrick, D. C. (**1997**). "Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex I. Heterogeneity of responses," J. Neurophysiol. **78**, 1222–1236.

Beckius, G. E., Batra, R., and Oliver, D. L. (**2000**). "Axon from anteroventral cochlear nucleus that terminates in medial superior olive of the cat: Observations related to delay lines," J. Neurosci. **19**, 3146–3161.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge, MA).

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (**2002**). "Precise inhibition is essential for microsecond interaural time difference coding," Nature (London) **417**, 543–547.

Carr, C. E., and Konishi, M. A. (**1990**). "Circuit for detection of interaural time differences in the brainstem of the barn owl," J. Neurosci. **10**, 3227–3246.

Cant, N. B., and Hyson, R. L. (**1992**). "Projections of the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil," Hear. Res. **58**, 26–34.

Ehrlich, I., Lohrke, S., and Friauf, E. (**1999**). "Shift from depolarizing to hyperpolarizing glycine action in rat auditory neurones is due to age-dependent Cl- regulation," J Physiol. **520**, 121–137.

Galambos, R., and Davis, H. (**1943**). "The responses of single auditory nerve fibers to acoustic stimulation," J. Neurophysiol. **6**, 39–57.

Gerstner, W., Kempter, R., van Hemmen, J. L., and Wagner, H. (**1996**). "A neuronal learning rule for sub-millisecond temporal coding," Nature (London) **383**, 76–81.

Glendenning, K. K., Masterton, R. B., Baker, B. N., and Wenthold, R. J. (**1991**). "Acoustic chiasm. III: Nature, distribution, and sources of afferents to the lateral superior olive in the cat," J. Comp. Neurol. **310**, 377–400.

Goldberg, J. M., and Brown, P. B. (**1969**). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization," J. Neurophysiol. **32**, 613–636.

Grothe, B., and Sanes, D. H. (**1994**). "Synaptic inhibition influences the coding properties of medial superior olivary neurons: An in-vitro study," J. Neurosci. **14**, 1701–1709.

Harper, N. S., and McAlpine, D. (**2004**). "Optimal neural population coding of an auditory spatial cue," Nature (London) **430**, 682–686.

Heffner, R. S., and Heffner, H. E. (**1988**). "Sound localization acuity in the cat: Effect of azimuth, signal duration, and test procedure," Hear. Res. **36**, 221–232.

Jeffress, L. A. (**1948**). "A place theory of sound localization," J. Comp. Physiol. Psychol. **41**, 35–39.

Jercog, P., Bezrutczyk, M. G., Kotak, V. C., Sanes, D. H., and Rinzel, J. (**2008**). "Asymmetry in electrotonic length of dendrites can position the ITD response function in gerbil MSO neurons," Society for Neuroscience Abstract No. 664.18.

Joris, P. X., and Yin, T. C. (**2007**). "A matter of time: Internal delays in binaural processing," Trends Neurosci. **30**, 70–78.

Joris, P. X., Smith, P. H., and Yin, T. C. (**1998**). "Coincidence detection in the auditory system: 50 Years after Jeffress," Neuron **21**, 1235–1238.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994**). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency," J. Neurophysiol. **71**, 1022–1036.

Joris, P. X., Van de Sande, B., Louage, D. H., and van der Hejden, M. (**2006**). "Binaural and cochlear disparities," Proc. Natl. Acad. Sci. U.S.A. **103**, 12917–12922.

Kempter, R., Gerstner, W., van Hemmen, J. L., and Wagner, H. (**1998**). "Extracting oscillations. Neuronal coincidence detection with noisy periodic spike input," Neural Comput. **10**, 1987–2017.

Kempter, R., Leibold, C., Wagner, H., and van Hemmen, J. L. (**2001**). "Formation of temporal-feature maps by axonal propagation of synaptic learning," Proc. Natl. Acad. Sci. U.S.A. **98**, 4166–4171.

Kittel, M., Wagner, E., and Klump, G. M. (**2002**). "An estimate of the auditory-filter bandwidth in the Mongolian gerbil," Hear. Res. **164**, 69–76.

Koch, U., Braun, M., Kapfer, C., and Grothe, B. (**2004**). "Distribution of HCN1 and HCN2 in rat auditory brainstem nuclei," Eur. J. Neurosci. **20**, 79–91.

Kuba, H., Ishii, T. M., and Ohmori, H. (**2006**). "Axonal site of spike initiation enhances auditory coincidence detection," Nature (London) **444**, 1069–1072.

Kuwabara, N., and Zook, J. M. (**1992**). "Projections to the medial superior olive from the medial and lateral nuclei of the trapezoid body in rodents and bats," J. Comp. Neurol. **324**, 522–538.

Leibold, C., and van Hemmen, J. L. (**2005**). "Spiking neurons learning phase delays: how mammals may develop auditory time-difference sensitivity," Phys. Rev. Lett. **94**, 168102.

Leibold, C., Kempter, R., and van Hemmen, J. L. (**2002**). "How spiking neurons give rise to a temporal-feature map: from synaptic plasticity to axonal selection," Phys. Rev. E **65**, 051915.

Magnusson, A. K., Kapfer, C., Grothe, B., and Koch, U. (**2005**). "Maturation of glycinergic inhibition in the gerbil medial superior olive after hearing onset," J. Physiol. **568**, 497–512.

Marquardt, T., and McAlpine, D. (**2006**). "A π-Limit for Coding ITDs: Implications for binaural models," in *Hearing From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer, Heidelberg), pp. 407–416.

McAlpine, D., and Grothe, B. (**2003**). "Sound localization and delay lines—do mammals fit the model?" Trends Neurosci. **26**, 347–350.

McAlpine, D., Jiang, D., and Palmer, A. R. (**1996**). "Interaural delay sensitivity and the classification of low best-frequency binaural responses in the inferior colliculus of the guinea pig," Hear. Res. **97**, 136–152.

McAlpine, D., Jiang, D., and Palmer, A. R. (**2001**). "A neural code for low-frequency sound localization in mammals," Nat. Neurosci. **4**, 396–401.

Mills, A. W. (**1958**). "On the minimum audible angle," J. Acoust. Soc. Am. **30**, 237–246.

Pecka, M., Brand, A., Behrend, O., and Grothe, B. (**2008**). "Interaural time difference processing in the mammalian medial superior olive: the role of glycinergic inhibition," J. Neurosci. **28**, 6914–6925.

Rautenberg, P. L., Grothe, B., and Felmy, F. (**2009**). "Quantification of the three-dimensional morphology of coincidence detector neurons in the medial superior olive of gerbils during late postnatal development," J. Comp. Neurol. **517**, 385–396.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 769–793.
Sandel, T. T., Teas, D. C., Feddersen, W. E., and Jeffress, L. A. (**1955**). "Localization of sound from single and paired sources," J. Acoust. Soc. Am. **27**, 842–852.
Schmiedt, R. A. (**1989**). "Spontaneous rates, thresholds and tuning of auditory-nerve fibers in the gerbil: Comparisons to cat data," Hear. Res. **42**, 23–35.
Siveke, I., Pecka, M., Seidl, A. H., Baudoux, S., and Grothe, B. (**2006**). "Binaural response properties of low-frequency neurons in the gerbil dorsal nucleus of the lateral lemniscus," J. Neurophysiol. **96**, 1425–1440.
Scott, L. L., Hage, T. A., and Golding, N. L. (**2007**). "Weak action potential backpropagation is associated with high-frequency axonal firing capability in principal neurons of the gerbil medial superior olive," J Physiol. **583**, 647–661.
Scott, L. L., Mathews, P. J., and Golding, N. L. (**2005**). "Posthearing developmental refinement of temporal processing in principal neurons of the medial superior olive," J. Neurosci. **25**, 7887–7895.
Smith, P. H. (**1995**). "Structural and functional differences distinguish principal from non-principal cells in the guinea pig MSO slice," J. Neurophysiol. **73**, 1653–1667.
Smith, P. H., Joris, P. X., and Yin, T. C. (**1993**). "Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive," J. Comp. Neurol. **331**, 245–260.
Svirskis, G., Dodla, R., and Rinzel, J. (**2003**). "Subthreshold outward currents enhance temporal integration in auditory neurons," Biol. Cybern. **89**, 333–340.
Svirskis, G., Kotak, V., Sanes, D. H., and Rinzel, J. (**2002**). "Enhancement of signal-to-noise ratio and phase locking for small inputs by a low-threshold outward current in auditory neurons," J. Neurosci. **22**, 11019–11025.
Tollin, D., and Yin, T. C. (**2005**). "Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive," J. Neurosci. **25**, 10648–10657.
Yin, T. C., and Chan, J. C. (**1990**). "Interaural time sensitivity in medial superior olive of cat," J. Neurophysiol. **64**, 465–488.
Yin, T. C., and Kuwada, S. (**1983**). "Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency," J. Neurophysiol. **50**, 1020–1042.
Young, S. R., and Rubel, W. E. (**1983**). "Frequency-specific projections of individual neurons in chick brainstem auditory nuclei," J. Neurosci. **3**, 1373–1378.
Zhou, Y., Carney, L. H., and Colburn, H. S. (**2005**). "A model for interaural time difference sensitivity in the medial superior olive: Interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology," J. Neurosci. **25**, 3046–3058.

A computer model of auditory efferent suppression: Implications for the recognition of speech in noise

Guy J. Brown[a)]
Department of Computer Science, University of Sheffield, Sheffield S1 4DP, United Kingdom

Robert T. Ferry and Ray Meddis
Department of Psychology, University of Essex, Colchester CO4 3SQ, United Kingdom

(Received 11 June 2009; revised 4 November 2009; accepted 20 November 2009)

The neural mechanisms underlying the ability of human listeners to recognize speech in the presence of background noise are still imperfectly understood. However, there is mounting evidence that the medial olivocochlear system plays an important role, via efferents that exert a suppressive effect on the response of the basilar membrane. The current paper presents a computer modeling study that investigates the possible role of this activity on speech intelligibility in noise. A model of auditory efferent processing [Ferry, R. T., and Meddis, R. (2007). J. Acoust. Soc. Am. **122**, 3519–3526] is used to provide acoustic features for a statistical automatic speech recognition system, thus allowing the effects of efferent activity on speech intelligibility to be quantified. Performance of the "basic" model (without efferent activity) on a connected digit recognition task is good when the speech is uncorrupted by noise but falls when noise is present. However, recognition performance is much improved when efferent activity is applied. Furthermore, optimal performance is obtained when the amount of efferent activity is proportional to the noise level. The results obtained are consistent with the suggestion that efferent suppression causes a "release from adaptation" in the auditory-nerve response to noisy speech, which enhances its intelligibility.
 [DOI: 10.1121/1.3273893]

PACS number(s): 43.64.Bt, 43.71.Rt [WPS] Pages: 943–954

I. INTRODUCTION

The detection of communication sounds against a background of environmental noise is a fundamental problem that affects many animal species. Among humans, this problem is particularly acute for listeners with impaired hearing, who frequently complain of difficulties in hearing speech in noisy places such as offices, shops, bars, and restaurants. An understanding of the mechanisms that underlie the ability of normal human listeners to recognize speech in the presence of background noise is therefore of considerable theoretical interest, and has an important practical application in the development of aids for the hearing impaired.

Our understanding of hearing is based mainly on our knowledge of the afferent system, where speech sounds are processed and passed through the auditory nervous system in the direction of the cerebral cortex. However, there have been numerous recent suggestions that the efferent system may make an important contribution (see Guinan, 1996, 2006 for reviews). The efferent system consists of nerve fibers whose direction of information flow appears to be away from the cortex. The most peripheral part of the auditory efferent system consists of fibers in the auditory nerve that project from the brainstem to the cochlea itself. It is now generally agreed that one component, the medial olivocochlear (MOC) system, indirectly exerts a suppressive influence on the response of the basilar membrane (BM) to sounds (Dallos, 1992). This has the effect of shifting the auditory-nerve (AN) rate/level function toward higher sound levels, and may be a way of extending the dynamic range of the auditory system. This effect has been recorded in small mammals at the level of the basilar membrane (Dolan *et al.*, 1997; Russell and Murugasu, 1997; Cooper and Guinan, 2006), the auditory nerve (Wiederhold and Kiang, 1970; Guinan and Stankovic, 1996), and the compound action potential (Winslow and Sachs, 1988; Dolan and Nuttall, 1988).

In addition, it has been suggested that the efferent system confers robustness to noise. For example, Dolan and Nuttall (1988) suggested that the activity of the efferent system may increase the detectability of tones in noise. They demonstrated that the compound action potential (CAP) response to a tone in noise was enhanced when the crossed olivocochlear bundle (OCB) was electrically stimulated. The mechanism responsible for this effect is complex, but is likely to involve adaptation. In prolonged background noise, the auditory nerve response becomes adapted and less responsive to new sounds. The efferent system has the potential to reduce the response to the continuous noise, reduce adaptation, and, therefore, enhance the response to a new sound presented in that background. Liberman and Guinan (1998) suggested that when the noise is continuous but the signal is transient, the MOC reflex acts to minimize the response to long-lasting stimuli while maximizing the response to novel stimuli. Other effects related to level-dependent compression may also be involved (Russell and Murugasu, 1997).

[a)]Author to whom correspondence should be addressed. Electronic mail: g.brown@dcs.shef.ac.uk

Relatively few studies give direct support to the idea that the efferent system contributes to the intelligibility of speech in noise. May and McQuone (1995) and Hienz *et al.* (1998) found that severing the olivocochlear bundle in cats reduced performance in tasks involving formant discrimination or intensity discrimination of tones in noise. Dewson (1968) showed that MOC lesions impair the ability of monkeys to discriminate vowel sounds when presented in noise, but have no effect on discrimination in silence. Giraud *et al.* (1997), using human subjects, found that contralateral noise improved speech-in-noise intelligibility in normal ears. They suggested that the crossed olivocochlear efferents were responsible. Kumar and Vanaja (2004) also found that contralateral acoustic stimulation improved speech intelligibility in noise when noise was presented to the contralateral ear, for ipsilateral signal-to-noise ratios (SNRs) of +10 and +15 dB. They showed that the same contralateral noise could suppress ipsilateral otoacoustic emissions, suggesting a role for efferent fibers. Kim *et al.* (2006) investigated the relationship between MOC efferent processing and speech intelligibility in noise for normal hearing listeners of different ages, using distortion product otoacoustic emissions (DPOAEs) as an index of efferent activity. Their findings suggest that the decline in ability to understand speech in noise with increasing age is associated with a corresponding decline in the function of the MOC efferent system.

It should be noted that not all of the evidence supports a role for the efferent system in improving the intelligibility of noisy speech. Wagner *et al.* (2008) found no correlation between efferent activity and speech intelligibility in noise, as judged by a speech reception threshold (SRT) test and measurements of contralateral suppression of DPOAEs. Other studies suggest a relatively minor role for the OCB in hearing; for example, Scharf *et al.* (1997) studied patients with sectioned crossed olivocochlear bundles and found that they had no audiological impairment. However, one of the few changes noted in these patients concerned the absence of an attention effect seen in normal listeners. In this effect, normal listeners had raised thresholds for stimulus tones at frequencies that were unexpected (i.e., had a low probability of occurrence). In their patient group, by contrast, Scharf *et al.* (1997) found no increase in threshold for unexpected stimuli. It is possible that efferent fibers were suppressing the BM response in regions most sensitive to the low probability frequencies in normal listeners. If this is the case, the patient group would be less likely to show this "attention" effect.

One way of critically assessing the claims made for the role of the efferent system in improving speech intelligibility in noise is to build and evaluate a computer model. Ghitza and co-workers (Ghitza *et al.*, 2007; Ghitza, 2007; Messing *et al.*, 2009) proposed a computer model of auditory efferent processing, and evaluated it on a speech recognition task. They described a closed-loop model of the auditory periphery in which the mechanical filtering properties of the cochlea are regulated by feedback based on short-term measurements of the dynamic range of simulated AN fibers. Ghitza *et al.* (Ghitza *et al.*, 2007; Ghitza, 2007) modeled consonant-confusions made by listeners in noise, for a diphone discrimination task that used synthetic speech stimuli with restricted phonemic variation. This was achieved by coupling the auditory model to a simple speech recognizer, which performed template-matching based on a minimum mean-squares error distance metric.

The current study also investigates the possible effect of efferent activity on speech intelligibility by using a computer model of auditory efferent processing as the "front-end" acoustic processor for an automatic speech recognition (ASR) system. Our approach differs from that of Ghitza *et al.* (2007) in several important respects. First, we use the auditory model of Ferry and Meddis (2007) which is open-loop (i.e., the amount of efferent suppression is fixed directly by the experimenter). This allows a systematic study of the effect of different amounts of suppression in order to identify the optimum level of efferent activity as a function of the level of the background noise. A similar analysis is not possible with Ghitza *et al.*'s (2007) model because of its closed-loop design [it should be noted, however, that the model of Ferry and Meddis (2007) is greatly simplified because the efferent system operates as a closed-loop system in practice].

Second, the current study uses a conventional statistical ASR system that is trained on a large corpus of naturalistic speech (spoken digits from the TIDigits speech corpus; Pearce and Hirsch, 2000). This contrasts with Ghitza *et al.*'s (2007) study, which used a simple template-matching recognizer and synthetic speech. Ghitza *et al.*'s (2007) choice of speech material and recognizer was made in order to ensure that errors due to the recognizer were minimized, so that the consonant-confusions that occurred were mainly due to the auditory model. A limitation of our approach is that we are unable to discriminate errors due to the "back end" recognizer from those that originate in the front end auditory model. However, our approach also has advantages; the speech material used is naturalistic and therefore more representative of the phonemic variation that is typically encountered in speech. Also, the statistical speech recognizer that we use is typical of modern ASR systems; the results of the current study therefore indicate whether, in principle, an auditory model that incorporates efferent processing could serve as a noise-robust front-end for a practical ASR system.

Finally, we note that Ghitza *et al.* (2007) used a simplistic model of neuromechanical transduction by inner hair cells and made no reference to the role of adaptation in explaining their findings. The model of Ferry and Meddis (2007) used here incorporates a detailed model of adaptation (and recovery from adaptation) at the auditory nerve synapse, and this will be shown to be an important factor in explaining the effect of efferent suppression.

It is important to stress that the proposed model is intended purely to illustrate the principle benefits of efferent stimulation when recognizing speech in noisy backgrounds. It is not proposed as a working model of the auditory efferent system for general use. Such a model would need to operate on a closed-loop basis such as that of Ghitza *et al.* (Ghitza *et al.*, 2007; Ghitza, 2007) and take into account the considerable body of knowledge recently accumulated concerning the time constants of efferent activation and the different amounts of inhibition observed across frequency. This will be the focus of future work. However, the results of Liber-

man (1988, Fig. 12C) suggested that noise, when it is present, is the dominant influence on the amount of efferent activity compared to the influence of accompanying pure tones. It follows that the model used here might well be a useful representation of what happens in the presence of continuous background noise of unchanging level.

The remainder of the article is structured as follows. In Sec. II, the computer model is described and it is informally demonstrated that efferent suppression of the BM response leads to an improved representation of speech in noise. An analysis is then presented which concludes that the beneficial effects of efferent suppression can largely be explained in terms of release from adaptation. After a description of the ASR system and speech corpus in Sec. III, experiments are presented in Secs. IV–VI that quantify the speech intelligibility gain associated with efferent processing, and the extent to which this depends on the noise level and the speech level.

II. THE COMPUTER MODEL

A. Model description

The computer model is shown schematically in Fig. 1, and consists of two main stages (delineated in the figure by gray boxes). The first stage is a model of peripheral auditory processing, which takes a digitally sampled mixture of speech and noise as its input and produces a simulation of action potential generation in the AN. The second stage is an ASR system that uses statistical word models to decode the AN firing pattern into its corresponding word sequence.

The computer model of the auditory periphery consists of a cascade of modules representing the resonances of the outer/middle ear, the response of the basilar membrane, coupling by inner hair cell stereocilia, the inner hair cell receptor potential, calcium dynamics, and transmitter release and adaptation at the inner hair cell auditory-nerve synapse. The final stage of the model produces a probabilistic representation of firing rate in the AN. Detailed discussions regarding the implementation and evaluation of each of these stages can be found in Meddis *et al.*, 2001; Lopez-Poveda and Meddis, 2001; Sumner *et al.*, 2002; Sumner *et al.*, 2003a, 2003b; Meddis, 2006.

The model of the basilar membrane used here is a modification of the dual resonance nonlinear (DRNL) filterbank proposed by Ferry and Meddis (2007) (see also Meddis *et al.*, 2001). The DRNL receives its input (stapes velocity) from a model of the outer/middle ear, and produces an output (basilar membrane velocity) that drives a simulation of inner hair cell function. A single DRNL filter is shown in Fig. 2. The output of the DRNL is the sum of two signal pathways, which represent linear and nonlinear components of the basilar membrane response. Each pathway consists of a sequence of bandpass (gammatone) and lowpass (Butterworth) filters.

The nonlinear path also contains a compressive nonlinearity, implemented by a "broken-stick" function that compresses the input signal (i.e., stapes velocity) when it exceeds a threshold level.[1] The nonlinear path also begins with an attenuation stage, introduced by Ferry and Meddis (2007) to model the effect of efferent suppression from the MOC. The amount of attenuation is determined by the parameter ATT

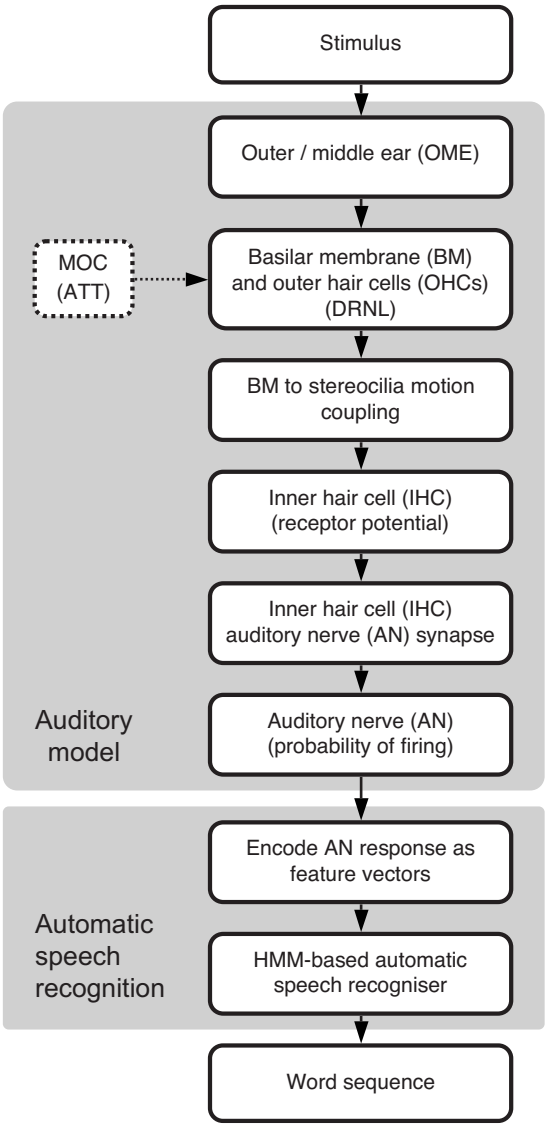


FIG. 1. Schematic of the computer model used for speech recognition experiments. The first major component is a model of the auditory periphery, which includes a stage representing efferent suppression from the MOC. The input to the auditory model is a digitally sampled stimulus (a mixture of speech and noise) and the output is the AN representation of the stimulus, in terms of firing probability. The second major component is an ASR system based on statistical word models. Data reduction is performed so that the AN response is encoded by a time-series of small feature vectors. These features are then decoded by an ASR system based on HMMs, producing a word sequence that is scored against a reference transcription.

(in decibels). This model has been shown to be in good agreement with physiological measurements of the basilar membrane, auditory-nerve and CAP responses when the value of ATT is chosen to be proportional to the amount of MOC activity (Ferry and Meddis, 2007).

In the following simulations, the parameters of the model differ from those used by Ferry and Meddis (2007). Whereas their study modeled physiological data from the guinea pig, our study addresses the representation of speech in human hearing. The outer/middle ear stage of the model was configured using data from Huber *et al.* (2001),[2] whereas the DRNL filterbank parameters were taken from Lopez-Poveda and Meddis (2001). The parameters for subsequent stages of the model were those given by Meddis

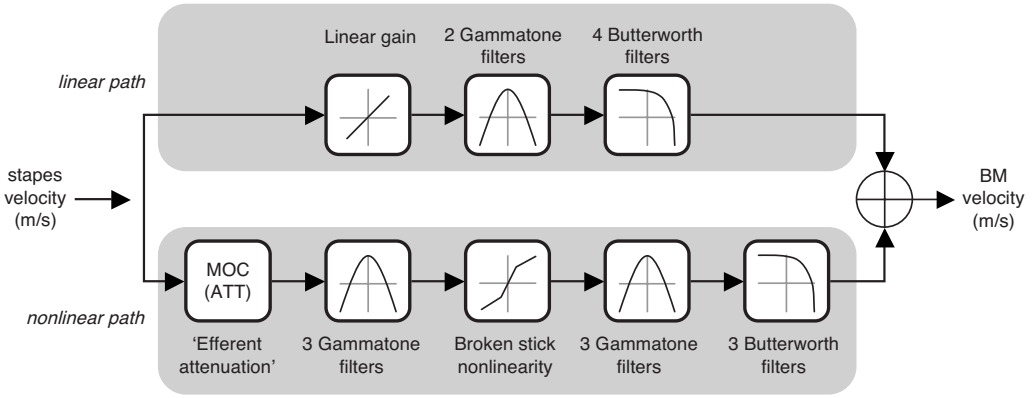


FIG. 2. Schematic diagram of the DRNL filterbank, modified to include an "efferent attenuation" stage (adapted from Ferry and Meddis, 2007). The DRNL consists of parallel linear and nonlinear signal paths, and the suppressive role of the MOC is modeled by inserting an attenuator at the start of the nonlinear path. The degree of efferent activity is determined by tuning the parameter ATT; larger values of ATT correspond to greater suppression by the MOC. The DRNL receives its input (stapes velocity) from a model of the outer/middle ear. Output from the DRNL (basilar membrane velocity) is subsequently processed by a model of inner hair cell function to give a representation of auditory-nerve activity.

(2006). Guinan and Stankovic (1996, Fig. 1) showed six different types of rate/level functions with and without electrical stimulation, all of which have been simulated in an earlier publication (Ferry and Meddis, 2007). One of these was chosen for the present study on the basis that it gave a good representation of the speech in quiet but a poor representation of speech in noise. The model fiber had a low spontaneous rate (LSR), a threshold of 20 dB and a narrow dynamic range saturating at 50 dB sound pressure level (SPL), and simulated in all respects the fiber in their Fig. 1D. These characteristics were obtained by setting the calcium clearance time constant $\tau_{Ca}=0.75\times10^{-4}$ s, as given in Table III of Meddis, 2006 (see also Sumner *et al.*, 2002, Table II). Clearly, many different kinds of fibers or mixtures of them could have been used. However, our main purpose was to illustrate how the efferent system could benefit the recognition of speech in noise even when the dynamic range is restricted.

To provide an auditory time-frequency representation of the noisy speech stimuli, 30 frequency channels were used with best frequencies (BFs) distributed between 100 and 4500 Hz on a logarithmic scale. A detailed list of model parameters can be found in Appendix A of Ferry (2008). The model was implemented in the MATLAB programming language: the source code is available from the authors on request.

In the second stage of the computer model, shown in Fig. 1, the simulated AN firing patterns provide the input to an automatic speech recognizer. The recognizer is a conventional statistical speech recognition system in which whole words are modeled by hidden Markov models (see, for example, Gales and Young, 2008). To provide a suitable input to the recognizer, the AN firing patterns are encoded as a sequence of feature vectors, each of which describes the spectral shape of the AN response at a certain point in time. Details of the encoding strategy and recognizer architecture are given in Sec. III B.

B. Analysis

Figure 3 shows the output of the peripheral auditory model in the form of an "auditory spectrogram," obtained by integrating overlapping 25 ms Hann-windowed segments of the simulated AN firing probability within each frequency channel at intervals of 10 ms. The grayscale value corresponds to firing rate (darker tones indicate higher firing rate). Panel (a) of the figure shows the auditory spectrogram for the utterance "two eight four one" spoken by a male talker and presented at a level of 60 dB SPL. Acoustic-phonetic features that are known to be important for speech intelligibility

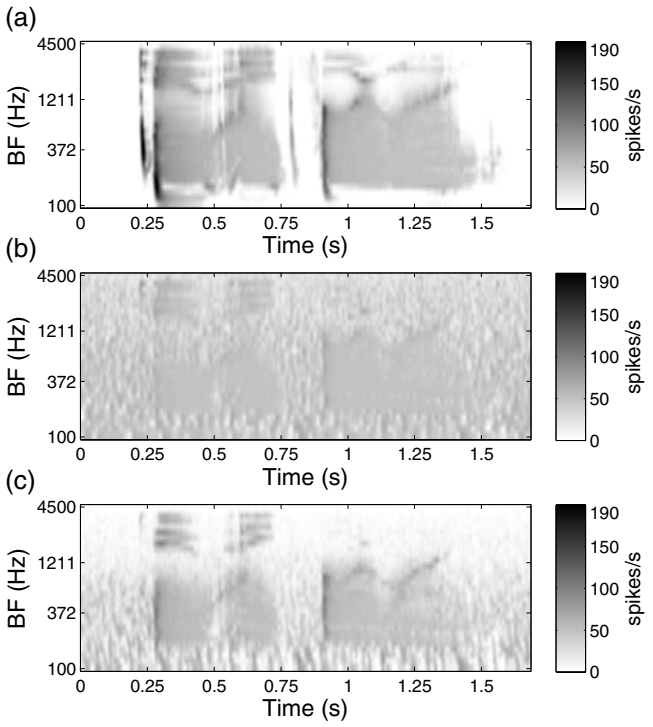


FIG. 3. Simulated auditory-nerve firing rate representations ("auditory spectrograms") for the utterance "two eight four one" spoken by a male talker. Darker regions represent higher firing rate, and the level of the speech is 60 dB SPL in all panels. (a) Clean speech with no efferent activity. (b) Speech with pink noise added at a level of 50 dB SPL (giving a signal-to-noise ratio of 10 dB), with no efferent activity. (c) As in panel (b), but with an efferent activity of 15 dB applied to the model. Efferent suppression reduces the masking effect of the noise. For clarity, the first 1 s of the auditory-nerve response has been omitted from the display; the speech was preceded by 1 s of silence in panel (a) and 1 s of noise in panels (b) and (c).

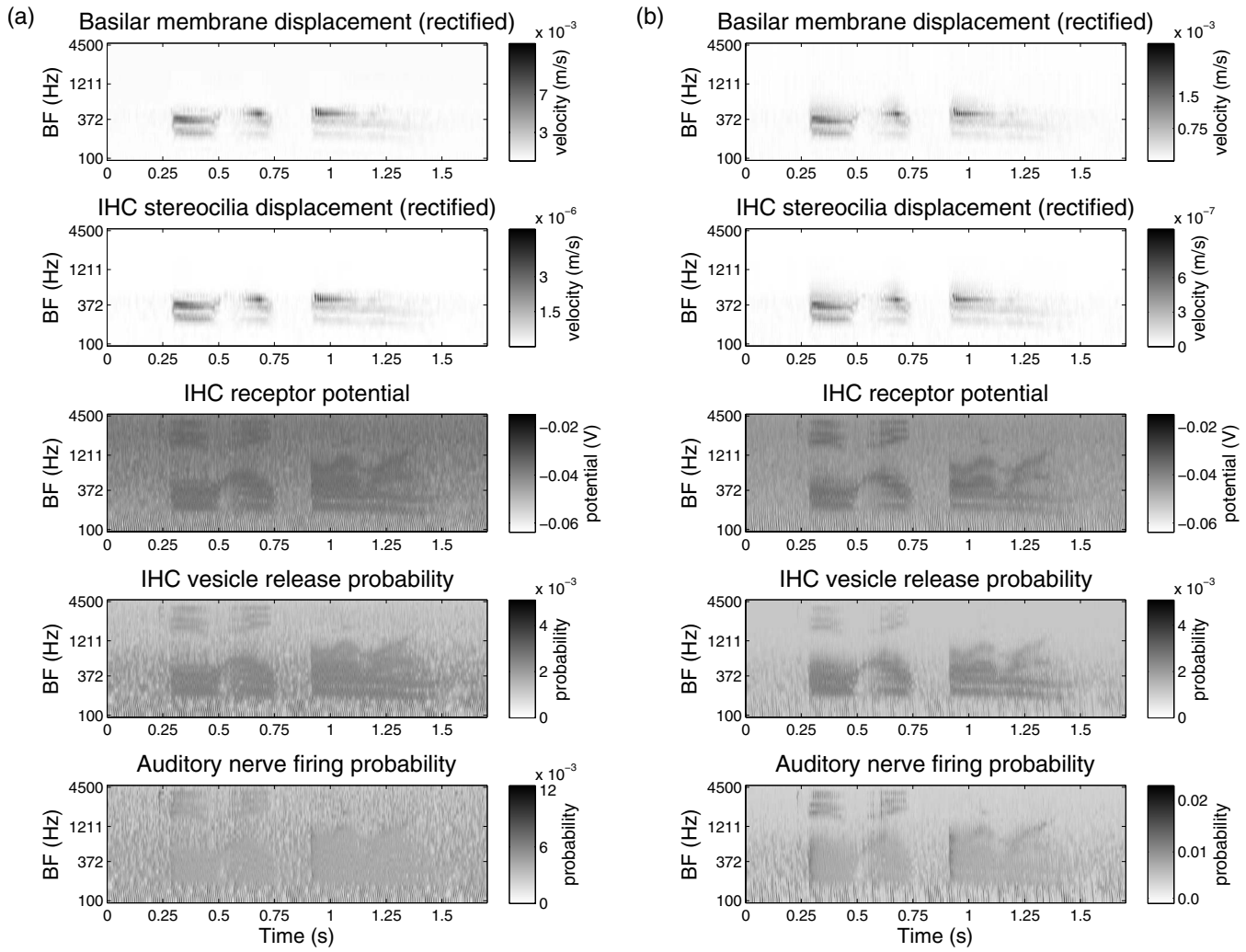


FIG. 4. Output from major stages of the auditory model, for the utterance "two eight four one" presented in pink noise. The sound levels of the speech and noise were 60 and 50 dB SPL, respectively, giving a signal-to-noise ratio of 10 dB. The response to 1 s of noise preceding the speech has been omitted from each figure. (a) No efferent attenuation (ATT=0 dB). (b) Simulated MOC stimulation giving an efferent attenuation of 15 dB (ATT=15 dB). From top to bottom, the plots in (a) and (b) show basilar membrane displacement, cilia displacement, IHC receptor potential, IHC vesicle release probability, and auditory-nerve firing probability. To improve the quality of the grayscale display, the basilar membrane displacement and IHC stereocilia displacement are full-wave rectified, and different scales are used in the left and right panels. Note that the auditory-nerve firing probability is higher in panel (b), because efferent attenuation reduces the adaptation caused by the preceding noise.

(e.g., formant transitions, release bursts, and frication) are well represented. However, panel (b) shows that much of this structure is lost when pink noise with a level of 50 dB SPL is added to the utterance (corresponding to a SNR of 10 dB). Weak time-frequency structure is masked by the noise, and high-intensity parts of the signal now drive the simulated auditory-nerve fibers close to their saturated firing rate. Panel (c) shows that the masking effect of the noise can be partially reversed by efferent suppression of the basilar membrane response. Here, efferent activity was simulated by setting ATT=15 dB, which reduces the gain in the nonlinear path of the DRNL by 15 dB (this amount of attenuation was found to be optimal for a speech level of 60 dB SPL and noise level of 50 dB SPL, as described in Sec. V). Note that the speech was preceded by 1 s of silence in panel (a), and 1 s of noise in panels (b) and (c); for clarity, the first 1 s of the auditory-nerve response has been omitted from the display.

Further insight into the effect of efferent activity can be gained by considering the representation of noisy speech in each stage of the auditory periphery model. Figure 4 shows the output from each stage of processing for the same mixture of speech and noise used in Fig. 3, for cases in which (a) no efferent attenuation is applied and (b) the efferent attenuation is 15 dB. In each case, the speech was preceded by 1 s of noise (which is not shown in the display). At stages of the model up to the inner hair cell (IHC) receptor potential, the effect of efferent activity resembles scaling by a constant factor (although constant scaling is not specifically expected due to the effect of BM compression). However, efferent activity has a more complex effect at the stage of the IHC vesicle release probability and beyond. Vesicle release due to the noise is suppressed, reducing adaptation and allowing the speech regions to elicit a larger vesicle release. This is reflected in the AN response, which shows a greater probability of firing and an increased dynamic range. The "unmasking" of noisy speech by efferent suppression can therefore be explained in terms of release from adaptation.

An alternative way of understanding this effect is to con-

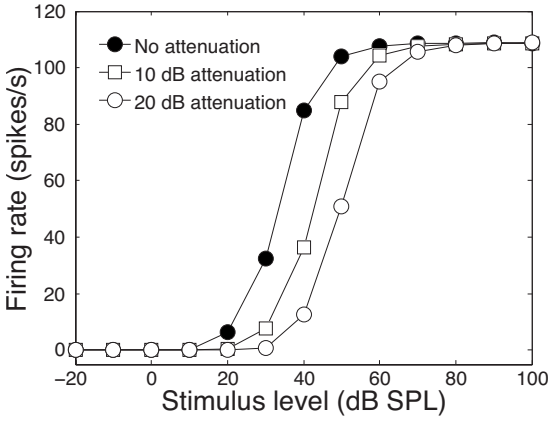


FIG. 5. Rate-level functions for a simulated auditory-nerve fiber with a best frequency of 1 kHz and a low spontaneous rate. The stimulus was a pure tone presented at BF with a duration of 100 ms and onset/offset ramps of 5 ms. Firing rate was averaged over the last 50 ms of the stimulus. Conditions are shown in which no efferent attenuation was applied (ATT =0 dB), and in which MOC stimulation was modeled by applying efferent attenuations of 10 and 20 dB (ATT=10, 20 dB). The rate-intensity curve shifts to the right as the amount of MOC stimulation increases.

sider the change in the rate-level function of a simulated auditory-nerve fiber when efferent activity is applied. Figure 5 shows a rate-level function generated by presenting brief (100 ms) pure tones to a simulated auditory-nerve fiber with a BF of 1 kHz. The frequency of the tone was set to the BF of the simulated fiber, and the firing rate was measured over the last 50 ms of each stimulus. The rate-level function has a typical sigmoidal shape, which progressively shifts to the right when increasing amounts of efferent attenuation are applied (conditions for ATT=10 dB and ATT=20 dB are shown in the figure). When speech is presented in a less intense noise background (i.e., at a positive SNR), the effect of such a shift in the rate-level function will be to reduce the AN response to the noise, since the lower-level noise will be relegated to the toe of the curve. Likewise, a shift in the rate-level function moves high-energy regions of the speech and noise mixture from the shoulder of the rate-level curve back to its linear portion. This reduces saturation and restores the dynamic range of the fiber. Note that the amount of unmasking produced by such a mechanism will depend on the level of the speech; this point is addressed later (Sec. VI).

III. AUTOMATIC SPEECH RECOGNIZER

The above discussion suggests that speech intelligibility in background noise should be improved by efferent suppression and raises the question of how the amount of unmasking is related to the speech level and noise level. The remainder of the paper investigates these issues by using the auditory model as the front-end processor for an ASR system. The speech and noise corpus and recognizer architecture are now described.

A. Corpus

Speech material for the following experiments was drawn from the Aurora 2.0 corpus (Pearce and Hirsch, 2000), which consists of sequences of between one and seven digits ("oh," "zero," and "one" to "nine") spoken by male and female talkers. Three sets of utterances were used. The recognizer was trained on the "clean" training set, which consists of 8440 utterances. For testing the recognizer, 1001 utterances from the "clean1" section of Aurora test set A were used. In addition, a small development set of 200 utterances drawn from the "clean2" section of test set A was used to tune the auditory model and ASR system. The training, testing, and development sets were completely independent, and each contained an approximately equal number of recordings from male and female talkers.

The Aurora speech material was modified in two respects to suit the experiments described here. First, the neuromechanical transduction stage of the auditory model involves numerical integration that must be performed at a high sample rate. Accordingly, all utterances were upsampled to a rate of 44.1 kHz (from the Aurora sample rate of 20 kHz) using the MATLAB resample function. Second, all utterances were scaled to the same root-mean-square level (60 dB SPL) in order to minimize changes in the spectral representation obtained from the (nonlinear) auditory model due to variations in sound level.

Noisy speech was generated by adding pink noise to the test utterances at a range of SNRs between 200 dB (clean) and 0 dB. Broadband noise was employed because it is known to be a particularly effective speech masker (e.g., Miller, 1947). The pink noise was band-passed between 100 Hz and 10 kHz in order to ensure that noise energy above the Nyquist frequency of the Aurora speech signals did not influence the SNR. Prior to adding the noise, 1 s of silence was appended to the start of each utterance; this allowed the auditory model to adapt before the onset of the speech. The corresponding second of simulated auditory-nerve response was removed before speech recognition.

B. Automatic speech recognizer

Speech recognition was performed by a conventional continuous-density hidden Markov model (HMM) system (e.g., see Gales and Young, 2008). Such systems require the acoustic input to be encoded as a sequence of feature vectors, each of which (a "frame") encodes the shape of the speech spectrum over a brief time window. The goal of the recognizer is to find the most likely word sequence that corresponds to an observed sequence of feature vectors.

The recognizer represents speech units (e.g., words) by trained HMMs that model the speech as a sequence of stationary states. Each state is characterized by a multivariate Gaussian mixture distribution over the observed acoustic feature vectors. During training, the Baum–Welch algorithm is used to learn the parameters of the HMMs from a large corpus of annotated speech. During testing, the VITERBI algorithm is applied to find the most likely sequence of HMM states (and hence words) given an observed sequence of feature vectors and the trained HMMs. For an accessible review of the Baum–Welch and Viterbi algorithms, see Rabiner, 1989. Here, a modified version of the Aurora baseline recognizer was used (Pearce and Hirsch, 2000), in which observations were modeled by Gaussian mixtures with diagonal covariance. Gaussian mixtures with seven components were

used, as these were found to give good performance on the development set.

It was necessary to perform some data reduction in the output of the auditory model in order to obtain feature vectors that were suited to the HMM speech recognizer. As previously shown in Fig. 3, the auditory-nerve firing probability emanating from each channel of the model was integrated over a 25 ms Hann window at intervals of 10 ms (i.e., successive windows overlapped by 60%), giving a temporal resolution that is typical for ASR systems. However, the resulting spectral features are not well modeled by a small number of Gaussian mixture components with diagonal covariance, because features from adjacent frequency channels are highly correlated. Accordingly, further data reduction was performed by applying a discrete cosine transform (DCT) to each frame, giving feature vectors that contain approximately independent components (Oppenheim *et al.*, 1999). The first 14 DCT coefficients were retained. To improve performance, time derivatives of the static DCT coefficients were also included; specifically, first-order and second-order regression coefficients (referred to as "deltas" and "accelerations") were appended to each vector, to give a total of 42 features per frame. A similar approach has been used in numerous other studies that employ auditory models as acoustic front-end processors for ASR systems (e.g., Jankowski *et al.*, 1995; Holmberg *et al.*, 2007).

HMMs with 16 emitting states were trained for each word in the Aurora corpus (i.e., the digits zero, oh, and one to nine). Models were also trained for silence (three states) and short pauses (one state). To reduce the number of insertion errors, a simple grammar was used to constrain all hypotheses so that they started and ended with the silence model. The hidden Markov model toolkit (HTK) was used to train the models and perform decoding (Young *et al.*, 2009). The ASR system was always trained on clean speech (i.e., without added noise) and no efferent attenuation was applied during training (ATT=0 dB).[3]

Word sequences produced by the recognizer were scored using the HTK HResults tool, which compares the transcript produced by the recognizer with a hand-labeled reference transcription. Recognition accuracy is computed as $(H-I)/N\times 100\%$, where H is the number of correct words ("hits"), I is the number of incorrectly inserted words ("insertions"), and N is the total number of words in the reference transcription.

IV. EXPERIMENT I: EFFECT OF EFFERENT ACTIVITY

It is well known from experiments with human listeners that speech intelligibility declines in the presence of broadband noise (e.g., Miller, 1947). The same is true of ASR systems, and hence the performance of an ASR system in noise can be taken as indicative of human speech intelligibility in noise. The comparison is only a qualitative one, however, because the error rate of an ASR system is typically an order of magnitude greater than that of a human listener under the same conditions (Lippmann, 1997). Additionally, humans and ASR systems differ in the rate at which their performance falls in the presence of increasing amounts of noise; ASR systems degrade much earlier, and much quicker, as the SNR worsens.

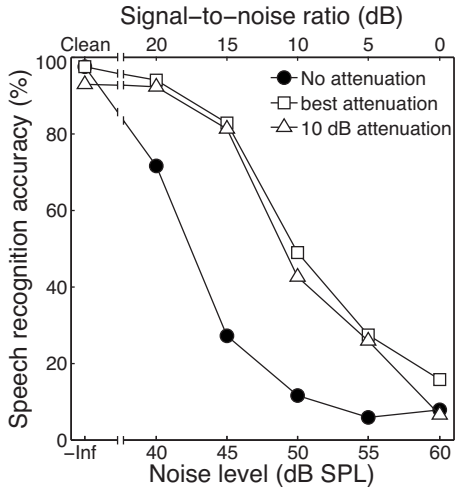


FIG. 6. Speech recognition accuracy (in percent) as a function of noise level, for conditions in which no efferent attenuation is applied, when 10 dB of efferent attenuation is applied, and when the "optimal" efferent attenuation for each noise level is used. Each point represents an average over the test set of 1001 utterances. Speech stimuli were presented at a sound level of 60 dB SPL during training and testing. Efferent activity gives a substantial performance gain in some noisy conditions. Note, however, that efferent activity degrades performance in the clean condition, suggesting that less activity is required in clean conditions and more in noisy conditions; this can be achieved by applying an optimal attenuation that is proportional to the noise level.

Noise is detrimental to ASR performance because it introduces a discrepancy between the training and testing conditions (i.e., there is a mismatch between the statistical models that are derived from clean speech during training, and the noisy speech features that are encountered during testing). A first question is whether efferent activity is able to compensate for this mismatch, by providing a representation of noisy speech that is closer to the ideal clean-speech models.

Figure 6 shows speech recognition accuracy for a range of noise levels, obtained from the auditory model and ASR system as described above. When efferent activity is disabled by setting ATT=0 dB (recall Fig. 2), speech recognition accuracy is high (97.5%) for clean speech but declines sharply with increasing noise level. For noise levels between 40 and 55 dB SPL, a substantial improvement in recognition accuracy is obtained by introducing an efferent activity of 10 dB (i.e., ATT=10 dB). This result confirms that for a speech level of 60 dB SPL, efferent suppression serves to reduce the effects of the noise, yielding acoustic features that more closely resemble those of clean speech. Experiment III (Sec. VI) investigates whether this conclusion holds across a range of different speech levels.

A notable feature of Fig. 6 is that recognition accuracy of clean speech declines (to 92.4%) when an efferent attenuation of 10 dB is introduced. This suggests that efferent activity is undesirable when noise is absent, because it warps the auditory representation of the speech away from the clean-speech models. More generally, this raises the issue of whether there is an "optimal" amount of efferent activity that

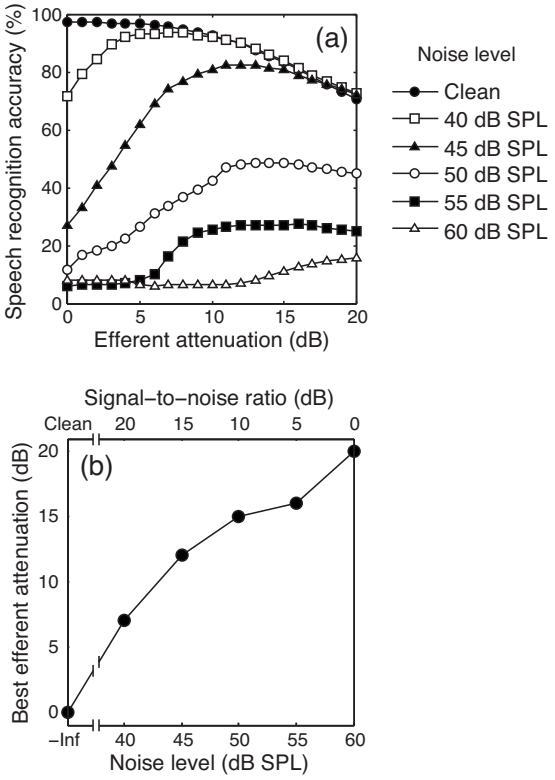


FIG. 7. (a) The effect of efferent activity on speech recognition accuracy. Each curve corresponds to a different noise level, with the speech level fixed at 60 dB SPL. (b) Best efferent attenuation as a function of noise level. The best efferent attenuation corresponds to the peak of each curve in panel (a); i.e., it is the efferent attenuation that maximizes speech recognition accuracy for the given noise level. Speech stimuli were presented at a sound level of 60 dB SPL during training and testing of the speech recognizer.

maximizes speech recognition accuracy, and whether this depends on the noise level. Similarly, an efferent attenuation of 10 dB gives little performance gain when the noise level is 60 dB SPL, and it is possible that recognition accuracy could be improved in this condition by increasing the amount of efferent activity. These issues are addressed in the following experiment.

V. EXPERIMENT II: EFFECT OF NOISE LEVEL

In this experiment, the relationship between noise level, level of efferent activity, and speech recognition accuracy was systematically investigated. Speech recognition accuracy was obtained for different configurations of the auditory model in which the efferent attenuation was set to a value between 0 and 20 dB in steps of 1 dB. The upper bound on the efferent activity was set in accordance with the physiological study of Liberman and Guinan (1998) in the cat, which found that the maximum suppression obtained with sound-evoked activity was approximately 20 dB. For each configuration of the model, speech recognition accuracy was evaluated in a range of noise conditions (clean speech, and speech with pink noise added at levels between 40 and 60 dB SPL).

Figure 7(a) shows the results from this experiment. For clean speech, recognition accuracy is highest when there is no efferent attenuation (ATT=0 dB). In the remaining conditions, the curve relating efferent activity to speech recognition accuracy generally shows a broad peak that indicates the optimal efferent attenuation for that noise level. Figure 7(b) plots these optimal efferent attenuation values against noise level. The results suggest that the best efferent attenuation is proportional to the noise level (i.e., if the level of the speech is held constant and noise is added, higher noise levels require more efferent activity).

Speech recognition accuracy is plotted in Fig. 6 (open squares) when the optimal efferent attenuation is used for each noise level. The resulting performance curve is indicative of the speech recognition accuracy achievable by a system that adjusts the efferent activity to its optimum value based on an assessment of the noise level. However, the optimal efferent attenuation is likely to depend on speech level (and overall sound level) in addition to the noise level. These factors are considered in the next experiment.

VI. EXPERIMENT III: EFFECT OF SPEECH LEVEL

In the two previous experiments, the speech was presented at a sound level of 60 dB SPL during training and testing of the ASR system. A further question is whether the benefits of efferent activity observed at a speech level of 60 dB SPL are also apparent at other speech levels. To address this, the ASR system was trained on clean speech and tested on mixtures of speech and pink noise for which the level of the speech was varied between 40 and 80 dB SPL in steps of 10 dB.

It should be noted that the aim of this experiment was to investigate the likely benefit of efferent activity at different sound levels, rather than to determine the robustness of the ASR system to discrepancies between the level of the speech in the training set and test set. Because the auditory model is nonlinear, it provides the recognizer with acoustic features that are level-dependent; hence, recognizer performance declines if there is a difference in speech level between the training and test sets. To avoid this confound, the sound level of the speech was always held the same during training and testing of the recognizer.

The results from this experiment are shown in Fig. 8. Speech recognition accuracy was determined using the optimum value of the efferent attenuation for each noise level, which was obtained using the procedure described in Sec. V. Substantial benefits of efferent activity are obtained at speech levels between 40 and 70 dB SPL. However, speech recognition accuracy is poor when the speech is presented at a level of 80 dB SPL, even with efferent activity. This can be explained by the rate-level function shown in Fig. 5. At a sound level of 80 dB SPL, the simulated auditory-nerve fibers are driven close to their saturated firing rate, even when a substantial efferent attenuation is applied. As a result, the efferent suppression does not result in a reduced firing rate during the noise and no reduction in adaptation is achieved. Conversely, speech recognition accuracy improves at lower sound levels, because the simulated auditory-nerve fibers respond in the linear portion of their rate-level functions. It should be noted, however, that the model does not currently include a simulation of the acoustic reflex, which would be active at sound levels above 75 dB SPL (Liberman and

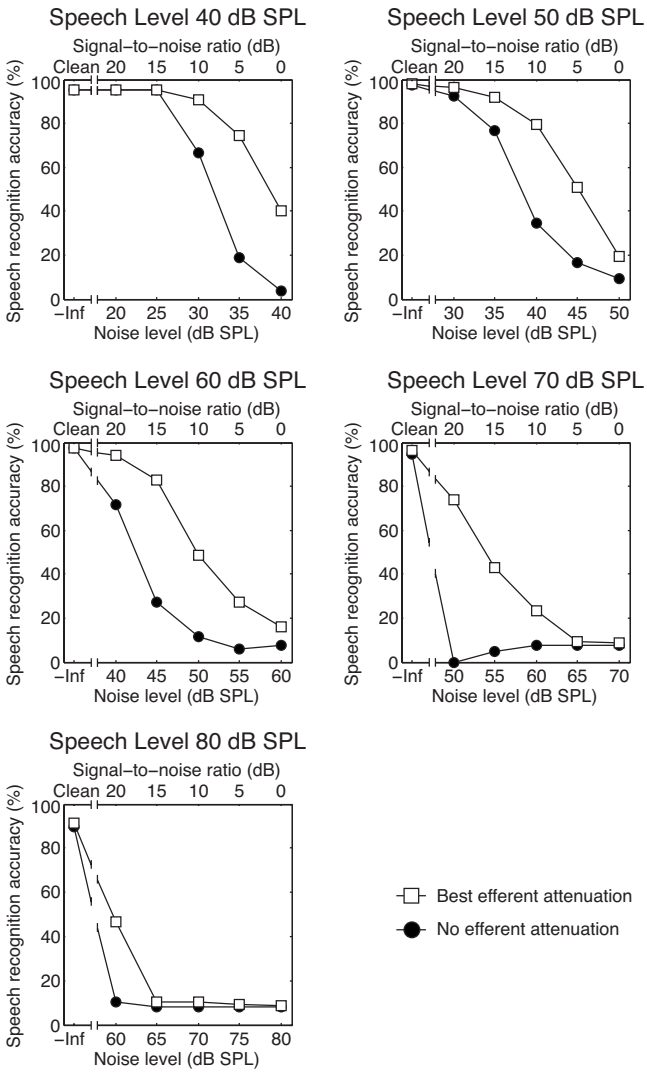


FIG. 8. Effect of speech level and noise level on word recognition accuracy. The plots show performance on the 1001-word test set when the recognizer was trained and tested on speech presented at sound levels between 40 and 80 dB SPL. The figures show performance without efferent activity and when the best efferent attenuation is used. The best efferent attenuation was determined separately for each experimental condition (i.e., for each combination of speech level and noise level).

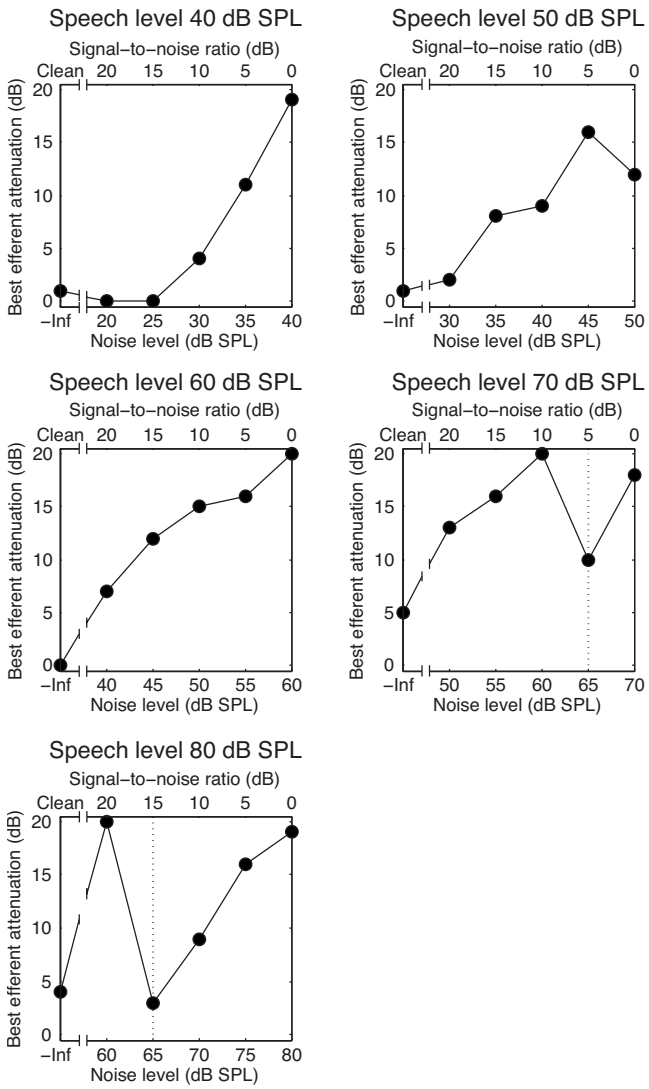


FIG. 9. Effect of speech level and noise level on best efferent attenuation. The plots show the best efferent attenuation for speech levels between 40 and 80 dB SPL. The best efferent attenuation was determined separately for each experimental condition (i.e., for each combination of speech level and noise level). At speech levels of 60 dB SPL and below, the best efferent attenuation generally increases with increasing noise level. For speech levels above 60 dB SPL, best efferent attenuation generally increases with increasing noise level until the performance of the recognizer falls to chance level. The point at which speech recognition degrades to chance performance is marked by a vertical dotted line in the panels for speech levels of 70 and 80 dB SPL.

Guinan, 1998). If present in the model, the acoustic reflex would reduce the effective level of the 80 dB SPL stimulus, leading to reduced adaptation and improved speech recognition accuracy.

Figure 9 illustrates the effect of speech level and noise level on best efferent attenuation. For speech levels of 40, 50, and 60 dB SPL, there is generally a monotonic increase in best efferent attenuation as the noise level is increased. Additionally, the best efferent attenuation rises more steeply with increasing noise level as the speech level is increased, indicating that the best efferent attenuation is also influenced by overall sound level. For speech levels of 70 and 80 dB SPL, the best efferent attenuation increases with increasing noise level only up to the point where the speech recognition performance degrades to chance level (approximately 9%; see Fig. 8). Beyond this point (indicated by a vertical dotted line in Fig. 9), efferent attenuation is unable to counteract the saturation caused by the high overall sound level, and many different values of ATT will give the same (poor) speech recognition performance.

VII. DISCUSSION

The aim of this modeling study was to investigate the potential role of the auditory efferent system in improving the intelligibility of speech that is masked by broadband noise. By using the auditory model as the front-end for an ASR system, it has been shown that speech recognition accuracy is improved by attenuating the response of the simulated basilar membrane when noise is present. Efferent activity has the effect of shifting the rate-level curve of the model auditory-nerve fibers to the right, which improves the dy-

namic range of the firing rate response and provides a release from adaptation. It was found that the optimum efferent attenuation was proportional to the noise level in most experimental conditions. This finding is compatible with a model in which the efferent system adjusts itself to the background noise level in order to prevent excessive adaptation and therefore optimizes speech recognition. Like the previous studies of Ghitza and colleagues (Ghitza *et al.*, 2007; Ghitza, 2007), our results therefore support the notion that the auditory efferent system contributes to the robustness of speech perception in adverse acoustic conditions.

The model suppresses the auditory response to broadband noise, providing a better spectro-temporal representation of speech components. Figure 6 shows that the effect of efferent activity in the model declines at high noise levels, and is marginal when the SNR is 0 dB. The ability of human listeners to recognize speech in noise is better than that of our ASR system. However, human speech intelligibility is poor below a SNR of 0 dB when the speech is masked by broadband noise (word intelligibility less than 50%; see Miller, 1947). We hypothesize that negative SNRs are particularly difficult for human listeners because the efferent system is less effective in unmasking the speech under such conditions. Although efferent activity will still reduce adaptation caused by noise when the speech level lies below the noise level, the model simulation suggests that the benefit will be small.

A limitation of the current study is that it only considers a pink noise masking sound. Human listeners are able to exploit temporal fluctuations in the envelope of a masker in order to "listen in the dips" or "glimpse" the speech (Miller and Licklider, 1950; Cooke, 2006). The effect of efferent suppression on "glimpsing" mechanisms remains an interesting issue for further study. Without efferent suppression, the availability of glimpses may be reduced because the peripheral auditory system is too adapted to respond during dips in the temporal envelope of the background noise. Listeners with impaired hearing have a reduced ability to listen in the dips (e.g., Hopkins *et al.*, 2008). Hence, a deficiency in efferent suppression (caused by outer hair cell damage or by a deficiency in the efferent system itself) is one factor that could contribute to the difficulty that hearing-impaired listeners experience when listening to speech in fluctuating background noise.

In the current model, the same level of efferent activity is applied at all frequencies. In fact, physiological data from the cat suggest that efferent suppression is greatest at frequencies above 2 kHz (Guinan and Gifford, 1988; Liberman and Guinan, 1998). However, the physiological data also show that efferent suppression is effective over a wide frequency range; for low spontaneous rate fibers (as used in the computer model), Fig. 4 of Guinan and Gifford (1988) indicates that the maximum efferent attenuation is at least 12 dB over the range of best frequencies used in the computer model (i.e., between 100 and 4500 Hz). The current model is therefore a reasonable approximation, although there is scope for more detailed modeling of the frequency-dependent effects of efferent activity.

Conventional front-end processors for ASR provide acoustic features that encode the spectral shape of speech but are largely independent of sound level. Typical approaches are cepstral mean subtraction (Liu *et al.*, 1993) and RASTA filtering (Hermansky and Morgan, 1994), both of which aim to remove slowly varying channel characteristics. Without the efferent attenuation component, the DRNL filterbank exhibits level-dependent behavior that complicates its use as a front-end for ASR; in particular, the best frequency associated with a filterbank channel is subjected to change at high sound levels, and the filter bandwidths broaden (Meddis *et al.*, 2001). Efferent activity moderates this behavior to some degree and extends the effective dynamic range of the auditory model (as would a simulation of the acoustic reflex, which is currently absent from the model). However, the results of Experiment III (Fig. 8) show that the performance of the model is still level-dependent, due to the saturating rate-level function of the simulated auditory-nerve fibers. For this reason, the speech recognition accuracy obtained with the auditory model is below that typically obtained with state-of-the-art signal processing front-ends (e.g., Cui and Alwan, 2005).

A related issue is the use of low spontaneous rate (LSR) auditory-nerve fibers in the computer model. In the simulations reported here, the calcium clearance time constant τ_{Ca} (Meddis, 2006; Sumner *et al.*, 2002) was configured to give fibers with a low spontaneous rate. The choice of a LSR fiber type to illustrate the consequences of efferent suppression was determined by the availability of relevant physiological data showing how efferent activity affects a fiber's response. Guinan and Stankovic (1996) showed the effect of efferent suppression on six different fibers, and all of these showed low spontaneous rates. All six fibers were simulated in the modeling study of Ferry and Meddis (2007). By choosing one of these fibers, it was possible to assume some physiological realism. Unfortunately, the most common type of fiber in small mammals shows high spontaneous rates. To compensate for this discrepancy, it was decided to use the fiber with a low threshold (20 dB SPL) and the narrowest dynamic range (20 dB) found in Fig. 1D of Guinan and Stankovich's (1996) report. In these two respects, at least, the fiber used in this study was similar to typical HSR fibers. In contrast, most LSR fibers have high thresholds and wide dynamic ranges.

The relative contribution of low and high spontaneous fiber types to the representation of speech sounds is of considerable interest but beyond the scope of this study (see, however, Winslow *et al.*, 1987; Sachs *et al.*, 2006). Our primary purpose here was to demonstrate improvements in speech-in-noise performance when efferent suppression is added to the model. This demonstration is particularly important with respect to fibers with narrow dynamic ranges. At first sight, auditory-nerve fibers that saturate at levels above 40 dB SPL would appear to be ill-suited to represent speech presented at levels of 60 dB SPL and above, especially when presented against a background of noise at similar levels. The modeling study described above has shown that this apparently unpromising approach can nevertheless give useful representations of speech in quiet up to 80 dB SPL (Fig.

8). Moreover, the poor performance of the system in background noise is considerably ameliorated when efferent suppression is applied at strengths related to the intensity of the noise.

VIII. CONCLUSIONS

It has been shown that efferent suppression improves the intelligibility of speech masked by broadband noise in a model that combines auditory efferent processing with an ASR system. Optimum speech intelligibility is achieved in the model using a level of efferent attenuation that is proportional to the noise level, other than when the noise is so intense that the recognizer degrades to chance performance. Unmasking due to efferent suppression occurs across a wide range of sound levels. However, the amount of unmasking depends both on the speech level and the noise level.

ACKNOWLEDGMENTS

This work was supported by grants from the Royal National Institute for Deaf People (RNID) to Guy Brown and Ray Meddis, and by a grant to Ray Meddis from the UK Engineering and Physical Sciences Research Council (Grant No. EP/E064590/1). The authors would like to thank Oded Ghitza and two anonymous reviewers for their helpful comments.

[1]The broken-stick nonlinearity is defined by $y(t) = \text{sign}[x(t)]\min[a|x(t)|, b|x(t)|^c]$, where a, b, and c are parameters and $x(t)$ and $y(t)$ represent the input and output signals, respectively. For details of the parameter values see Lopez-Poveda and Meddis (2001).

[2]The filtering of the human outer/middle ear is modeled by passing the acoustic signal through three first-order linear bandpass Butterworth filters arranged in series. The first filter has lower and upper cutoff frequencies of 1900 and 4200 Hz. The second and third filters have lower/upper cutoff frequencies of 4500/6300 and 8000/12 000 Hz, respectively.

[3]In practice, the efferent system is likely to be activated by speech alone, and hence it would be appropriate to apply some degree of efferent attenuation during training of the recognizer with clean speech. In this study, efferent activity was suppressed during training of the recognizer in order to simplify our study of the benefit of efferent activity during speech recognition.

Cooke, M. (**2006**). "A glimpsing model of speech perception in noise," J. Acoust. Soc. Am. **119**, 1562–1573.
Cooper, N. P., and Guinan, J. J. (**2006**). "Medial olivocochlear efferent effects on basilar membrane response to sound," in *Auditory Mechanisms: Processes and Models*, edited by A. L. Nuttall, T. Ren, P. G. Gillespie, K. Grosch, and E. de Boer (World Scientific, Singapore), pp. 86–92.
Cui, X., and Alwan, A. (**2005**). "Noise robust speech recognition using feature compensation based on polynomial regression of utterance SNR," IEEE Trans. Speech Audio Process. **13**, 1161–1172.
Dallos, P. (**1992**). "The active cochlea," J. Neurosci. **12**, 4575–4585.
Dewson, J. H. (**1968**). "Efferent olivocochlear bundle: Some relationships to stimulus discrimination in noise," J. Neurophysiol. **31**, 122–130.
Dolan, D., and Nuttall, A. (**1988**). "Masked cochlear whole-nerve response intensity functions altered by electrical-stimulation of the crossed olivocochlear bundle," J. Acoust. Soc. Am. **83**, 1081–1086.
Dolan, D. F., Guo, M. H., and Nuttall, A. L. (**1997**). "Frequency-dependent enhancement of basilar membrane velocity during olivocochlear bundle stimulation," J. Acoust. Soc. Am. 3587–3596.
Ferry, R. T. (**2008**). "Auditory processing and the medial olivocochlear efferent system," Ph.D. thesis, University of Essex, Colchester, UK.
Ferry, R. T., and Meddis, R. (**2007**). "A computer model of medial efferent suppression in the mammalian auditory system," J. Acoust. Soc. Am. **122**, 3519–3526.
Gales, M., and Young, S. (**2008**). "The application of hidden Markov models in speech recognition," Foundations and Trends in Signal Processing **1**, 195–304.
Ghitza, O. (**2007**). "Using auditory feedback and rhythmicity for diphone discrimination of degraded speech," in Proceedings of the International Conference on Phonetic Sciences (ICPhS), Saarbrucken, Germany, pp. 163–168.
Ghitza, O., Messing, D., Delhorne, L., Braida, L., Bruckert, E., and Shondhi, M. (**2007**). "Towards predicting consonant confusions of degraded speech," in *Hearing—From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer, Berlin), pp. 541–550.
Giraud, A., Garnier, S., Micheyl, C., Lina, G., Chays, A., and CheryCroze, S. (**1997**). "Auditory efferents involved in speech-in-noise intelligibility," NeuroReport **8**, 1779–1783.
Guinan, J. J. (**1996**). "Physiology of olivocochlear efferents," in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer-Verlag, Berlin), pp. 435–502.
Guinan, J. J. (**2006**). "Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans," Ear Hear. **27**, 589–607.
Guinan, J. J., and Gifford, M. L. (**1988**). "Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory-nerve fibers. III. Tuning curves and thresholds at CF," Hear. Res. **37**, 29–45.
Guinan, J. J., and Stankovic, K. M. (**1996**). "Medial efferent inhibition produces the largest equivalent attenuations at moderate to high sound levels in cat auditory-nerve fibers," J. Acoust. Soc. Am. **100**, 1680–1690.
Hermansky, H., and Morgan, N. (**1994**). "Rasta processing of speech," IEEE Trans. Speech Audio Process. **2**, 578–589.
Hienz, R., Stiles, P., and May, B. (**1998**). "Effects of bilateral olivocochlear lesions on vowel formant discrimination in cats," Hear. Res. **116**, 10–20.
Holmberg, M., Gelbart, D., and Hemmert, W. (**2007**). "Speech encoding in a model of peripheral auditory processing: Quantitative assessment by means of automatic speech recognition," Speech Commun. **49**, 917–932.
Hopkins, K., Moore, B. C. J., and Stone, M. A. (**2008**). "Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech," J. Acoust. Soc. Am. **123**, 1140–1153.
Huber, A., Linder, T., Ferrazzini, M., Schmid, S., Dillier, N., Stoeckli, S., and Fisch, U. (**2001**). "Intraoperative assessment of stapes movement," Ann. Otol. Rhinol. Laryngol. **110**, 31–35.
Jankowski, C. R. J., Vo, H.-D., and Lippmann, R. P. (**1995**). "A comparison of signal processing front ends for automatic word recognition," IEEE Trans. Speech Audio Process. **3**, 286–293.
Kim, S., Frisina, R. D., and Frisina, D. R. (**2006**). "Effects of age on speech understanding in normal hearing listeners: Relationship between the auditory efferent system and speech intelligibility in noise," Speech Commun. **48**, 855–862.
Kumar, U., and Vanaja, C. (**2004**). "Functioning of olivocochlear bundle and speech perception in noise," Ear Hear. **25**, 142–146.
Liberman, M. C. (**1988**). "Response properties of cochlear efferent neurons: Monaural vs. binaural stimulation and the effects of noise," J. Neurophysiol. **60**, 1779–1798.
Liberman, M. C., and Guinan, J. J. (**1998**). "Feedback control of the auditory periphery: Anti-masking effects of middle ear muscles vs. olivocochlear efferents," J. Commun. Disord. **31**, 471–482.
Lippmann, R. P. (**1997**). "Speech recognition by machines and humans," Speech Commun. **22**, 1–16.
Liu, R., Stern, R., Huang, X., and Acero, A. (**1993**). "Efficient cepstral normalization for robust speech recognition," in Proceedings of ARPA Speech and Natural Language Workshop, (Princeton, NJ), pp. 69–74.
Lopez-Poveda, E. A., and Meddis, R. (**2001**). "A human nonlinear cochlear filterbank," J. Acoust. Soc. Am. **110**, 3107–3118.
May, B. J., and McQuone, S. J. (**1995**). "Effects of bilateral olivocochlear lesions on pure-tone intensity discrimination in cats," Aud. Neurosci. **1**, 385–400.
Meddis, R. (**2006**). "Auditory-nerve first-spike latency and auditory absolute threshold: A computer model," J. Acoust. Soc. Am. **119**, 406–417.
Meddis, R., O'Mard, L., and Lopez-Poveda, E. (**2001**). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**, 2852–2861.
Messing, D. P., Delhorne, L., Bruckert, E., Braida, L. D., and Ghitza, O. (**2009**). "A non-linear efferent-inspired model of the auditory system; matching human confusions in stationary noise," Speech Commun. **51**, 668–683.
Miller, G. A. (**1947**). "The masking of speech," Psychol. Bull. **44**, 105–129.

Miller, G. A., and Licklider, J. (**1950**). "The intelligibility of interrupted speech," J. Acoust. Soc. Am. **22**, 167–173.
Oppenheim, A. V., Schafer, R. W., and Buck, J. R. (**1999**). *Discrete-Time Signal Processing* (Pearson Education, Upper Saddle River, NJ).
Pearce, D., and Hirsch, H.-G. (**2000**). "The AURORA experimental framework for the performance evaluation of speech recognition systems under noisy conditions," in Proceedings of the International Conference on Spoken Language Processing, Vol. **IV**, pp. 29–32.
Rabiner, L. R. (**1989**). "A tutorial on hidden Markov models and selected applications," Proc. IEEE **77**, 257–286.
Russell, I. J., and Murugasu, E. (**1997**). "Medial efferent inhibition suppresses basilar membrane responses to near characteristic frequency tones of moderate to high intensities," J. Acoust. Soc. Am. **102**, 1734–1738.
Sachs, M. B., May, B. J., Prell, G. S. L., and Hienz, R. D. (**2006**). "Adequacy of auditory-nerve rate representations of vowels: Comparison with behavioural measures in cat," in *Listening to Speech: An Auditory Perspective*, edited by S. Greenberg and W. A. Ainsworth (Lawrence Erlbaum Associates, Hillsdale, NJ), pp. 115–127.
Scharf, B., Magnan, J., and Chays, A. (**1997**). "On the role of the olivocochlear bundle in hearing: 16 case studies," Hear. Res. **103**, 101–122.
Sumner, C., Lopez-Poveda, E., O'Mard, L., and Meddis, R. (**2002**). "A revised model of the inner-hair cell and auditory-nerve complex," J. Acoust. Soc. Am. **111**, 2178–2189.
Sumner, C., Lopez-Poveda, E., O'Mard, L., and Meddis, R. (**2003a**). "Adaptation in a revised inner-hair cell model," J. Acoust. Soc. Am. **113**, 893–901.
Sumner, C., O'Mard, L., Lopez-Poveda, E., and Meddis, R. (**2003b**). "A nonlinear filterbank model of the guinea-pig cochlear nerve: Rate responses," J. Acoust. Soc. Am. **113**, 3264–3274.
Wagner, W., Frey, K., Heppelmann, G., Plontke, S. K., and Zenner, H.-P. (**2008**). "Speech-in-noise intelligibility does not correlate with efferent olivocochlear reflex in humans with normal hearing," Acta Oto-Laryngol. **128**, 53–60.
Wiederhold, M. L., and Kiang, N. Y. S. (**1970**). "Effects of electric stimulation of the crossed olivocochlear bundle on single auditory-nerve-fibers in the cat," J. Acoust. Soc. Am. **48**, 950–965.
Winslow, R. L., Barta, P. E., and Sachs, M. B. (**1987**). "Rate coding in the auditory nerve," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Lawrence Erlbaum Associates, Hillsdale, NJ), pp. 212–224.
Winslow, R. L., and Sachs, M. B. (**1988**). "Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle," Hear. Res. **35**, 165–190.
Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., and Woodland, P. (**2009**). *The Hidden Markov Model Toolkit (HTK)*, Cambridge University Engineering Department, http://htk.eng.cam.ac.uk/ (Last viewed 10/06/09).

Properties of a nonlinear version of the stimulus-frequency otoacoustic emission

Kyle P. Walsh,[a] Edward G. Pasanen, and Dennis McFadden
Department of Psychology and Center for Perceptual Systems, 1 University Station A8000, University of Texas, Austin, Texas 78712-0187

(Received 1 June 2009; revised 4 December 2009; accepted 7 December 2009)

A procedure for extracting the nonlinear component of the stimulus-frequency otoacoustic emission (SFOAE) is described. This nSFOAE measures the amount by which the cochlear response deviates from linear additivity when the input stimulus is doubled in amplitude. When a 4.0-kHz tone was presented alone, the magnitude of the nSFOAE response remained essentially constant throughout the 400-ms duration of the tone; response magnitude did increase monotonically with increasing tone level. When a wideband noise was presented alone, nSFOAE magnitude increased over the initial 100- to 200-ms portion of the 400-ms duration of the noise. When the tone and the wideband noise were presented simultaneously, nSFOAE magnitude decreased momentarily, then increased substantially for about the first 100 ms and then remained strong for the remainder of the presentation. Manipulations of the noise bandwidth revealed that the low-frequency components were primarily responsible for this rising, dynamic response; no rising segment was seen with bandpass or highpass noise. The rising, dynamic nSFOAE response is likely attributable to activation of the medial olivocochlear efferent system. This perstimulatory emission appears to have the potential to provide information about the earliest stages of auditory processing for stimuli commonly used in psychoacoustical tasks.
 [DOI: 10.1121/1.3279832]

PACS number(s): 43.64.Jb, 43.64.Kc, 43.64.Bt, 43.64.Nf [BLM] Pages: 955–969

I. INTRODUCTION

One type of otoacoustic emission (OAE) is the stimulus-frequency OAE (SFOAE). Kemp was the first to measure SFOAEs (Kemp and Chum, 1980; Kemp, 1980) and since then various procedures have been used to extract versions of the SFOAE. Fundamental characteristics of SFOAEs are that they are measured during the presentation of a (much-stronger) acoustic stimulus, typically a tone, and the frequency components present in the SFOAE correspond to frequency components in the acoustic stimulus. Here we describe a procedure that extracts the nonlinear component of the SFOAE response and that appears to have promise for studying cochlear processes during the presentation of waveforms commonly used in behavioral studies of hearing.

The auditory system consists of a series of early mechanical stages of processing of the incoming sound at the auditory periphery followed by numerous neural stages of processing in the brain, all culminating in auditory perception and the initiation of responses to the acoustical world. Each successive stage of processing has the potential to perpetuate, elaborate, offset, and/or augment aspects of the earlier stages of processing. An ultimate goal of auditory science is to understand how and where in this series of processing stages various details of auditory perception arise and are refined. For example, one might ask where and how the critical band first arises, and where and how it is refined to have the final characteristics it has in detailed behavioral tests.

To date, our greatest knowledge about auditory performance and perception has come from psychophysical studies of humans, and our greatest knowledge about the processing characteristics of the auditory periphery and brain has come from physiological studies on non-human species. This partitioning of knowledge by species is likely to persist for some time because of the inherent difficulties associated with obtaining appropriate measurements of the successive neural stages in humans. However, the physiological measure described here appears able to provide auditory science with knowledge about some early stages of processing in humans during the presentation of a wide array of stimulus waveforms, and offers a noninvasive technique for animal research as well.

The procedures described here are closely related to those used by Keefe (e.g., Keefe, 1998; Schairer *et al.*, 2003; Schairer and Keefe, 2005; Keefe *et al.*, 2009) and Guinan (e.g., Guinan, 2006; Guinan *et al.*, 2003; Backus and Guinan, 2006) to measure SFOAEs, but the stimuli, procedures, and analyses do differ in several significant details. More importantly, several of our key outcomes differ from those reported previously. Procedures used by previous investigators have measured a combination of the linear and nonlinear components of the SFOAE, whereas our procedure extracts only the nonlinear component. Ultimately, this difference may explain some of the differences in outcomes that have been observed with the various procedures. We use the term nSFOAE when referring to our measure, both to distinguish it from other SFOAE measures and to emphasize that it contains only the nonlinear component of the physiological response.

Our long-term goal is to compare the measurements obtained using certain forms of OAE with behavioral perfor-

[a]Author to whom correspondence should be addressed. Electronic mail: kylewalsh@mail.utexas.edu

mance measured in various psychophysical tasks when using the same ears, the same acoustic stimuli, and the same basic procedures in both domains. Specifically, the purpose of this research is to determine how much the processing characteristics of the active human cochlea contribute to performance in certain psychoacoustical tasks, and how much of the individual differences that exist psychoacoustically can be explained by individual differences in cochlear response. However, prior to reporting on specific psychoacoustical tasks, such as overshoot, forward masking, critical bands, etc., some of the basic properties of the nSFOAE need to be documented, and that is the purpose of this report. Additional nSFOAE measurements made as part of this study are reported in the second paper in this series.[1]

II. METHODS

A. Subjects

One female aged 21, two males aged 19, and one male aged 26 participated in this study. All had audiometrically normal hearing sensitivity ($\leq$15 dB Hearing Level) in both ears for the standard audiometric frequencies between 250 and 8000 Hz and normal middle-ear function as measured by a clinical audiometric screening device (Auto Tymp 38, GSI/VIASYS, Madison, WI). No subject had an SOAE stronger than −9.0 dB sound-pressure level (SPL) any closer than 640 Hz to the 4.0-kHz tone used for most measurements here. The male subjects all had extensive experience with auditory psychophysics, and all four subjects had their OAEs measured at least once prior to this study. Informed consent was obtained from all subjects prior to the study, and all were paid for their participation except for author KW.

B. Procedures

The OAE measure used here is related to the SFOAE that was first measured by Kemp and Chum (1980) and later studied by Kemp (1980), Zwicker and Schloth (1984), Dallmayr (1987), and Lonsbury-Martin *et al.* (1990), among others (see Probst *et al.*, 1991, for an early review). In recent times, SFOAEs have been used productively by a number of prominent investigators studying an assortment of basic phenomena in various species. Numerous procedures for extracting the SFOAE have been reported, but the common factor is that the measurements are made *during* the presentation of an acoustic stimulus, typically a stimulus of long duration. As noted, this perstimulatory response contains both linear and nonlinear components, and most procedures measure a combination of the two, whereas the procedure to be described here extracts only the nonlinear component, hence the name nSFOAE.

The primary goal of the work reported here was to explore how the nSFOAE measure typically responds to manipulations of a number of basic stimulus parameters. Complete sets of data were collected from all subjects for the majority of conditions, but for some conditions, data were collected from only one or two subjects, and those conditions are identified where appropriate. The data are described largely in terms of general patterns of the nSFOAE response, and representative results are shown in the figures. Comparisons are made between subjects' data, but the possible implications of individual differences on psychophysical performance are not explored here.

For the OAE measurements, an individual subject was seated in a comfortable reclining chair in a double-walled, sound-attenuated room. Prior to any data collection, the subject relaxed in isolation in this room for 15 min. This initialization period has been shown to stabilize OAE measurements (Whitehead, 1991; McFadden and Pasanen, 1994). The OAEs were measured in the right ear in all subjects.

The equipment used for measuring nSFOAEs included an Etymotic ER-10A microphone system and two Etymotic ER-2 earphones (Etymotic, Elk Grove Village, IL). The earphones were attached to small plastic sound-delivery tubes that terminated at the sound-delivery ports located at the outboard end of the microphone capsule. Metal tubes passed through the probe tip of the microphone system for the delivery of sound into the external ear canal. The acoustic stimuli were produced, and the nSFOAEs recorded, using a Macintosh G4 computer running custom-written LabVIEW® software.

The microphone output was amplified by 20 dB by the Etymotic preamplifier, then passed to a custom-built amplifier/filter that highpass filtered the waveform at about 400 Hz to eliminate low-frequency noise. The waveform also was lowpass filtered at about 15 kHz prior to being delivered to an analog-to-digital converter. All acoustic stimuli were generated digitally in the computer, delivered to a digital-to-analog converter, passed through a custom-built earphone amplifier, and then delivered to the earphones. The sampling rate for both input and output was 50 kHz with 16-bit resolution. Digitizing was accomplished using a National Instruments board (PCI-MIO-16XE-10) installed in the Macintosh G4 computer.

Our procedure for extracting the nSFOAE borrows elements from several previous reports (e.g., Keefe, 1998; Schairer *et al.*, 2003; Guinan *et al.*, 2003). For data collection, the stimulus waveforms were organized in sets of three successive presentations (a triplet), and a block consisted of at least 50 triplets. The first stimulus presentation of each triplet involved activating only one of the two Etymotic ER-2 earphones, the second presentation involved activating only the other earphone, and the third presentation involved the simultaneous activation of both earphones playing the same digitized waveform (a version of the "double-evoked" procedure described by Keefe, 1998). For all triplets in a block, the stimulus waveforms were identical. The voltage delivered to the individual earphones was always exactly the same for all triplets for a particular condition; so in a strictly linear system, the instantaneous pressure fluctuations for the third presentation should have been the exact acoustic sum of the instantaneous fluctuations in the first two presentations.

The sound in the ear canal was recorded during all three presentations for each triplet. Those sounds consisted of the (quite strong) acoustic stimulus plus whatever (weak) sound was being produced inside the cochlea in response to the acoustic stimulus. The sounds recorded during each of the first two presentations of a triplet were summed, and that sum was subtracted from the sound recorded during the third

presentation of that triplet. (This subtraction was performed prior to the presentation of the next triplet in order to assess that sample's freedom from artifact and noise; see below.) The acoustic stimuli always were the same digitized waveforms, so ideally the result of this subtraction was the elimination of both the physical stimulus and any linear components of the cochlear response. Assuming that the recording system itself is linear, any residual left in the difference waveform can be attributed to nonlinearity in the cochlear response (see Keefe, 1998)—here called the nSFOAE. The nSFOAE response might best be thought of as the difference between the actual measurement in the ear canal and the expectation about that measurement based on the process of additivity in a strictly linear system.

Those difference waveforms determined to be artifact- and noise-free (see below) were averaged, and data collection continued until at least 50 such waveforms were collected. The resulting averaged waveform was saved and later subjected to an analysis (see below) that yielded our estimate of the nSFOAE to the acoustic stimuli used for that block of triplets. The assumption that the difference waveform originates primarily from the cochlea is strengthened by the fact that the earphones and microphone used here were highly linear in their responses when our system was tested with a (strictly linear) 0.5-cc syringe instead of an ear canal. In the syringe, the nonlinear response was about −15 to −20 dB SPL in the 400-Hz band used for analysis (see below), which was well below the values typically measured in an ear canal, and was not distinguishable from the noise floor of the measurement system with stimulus amplitudes set to zero.

The sound-pressure levels produced by the ER-2 earphones when driven by the waveforms produced by the computer were calibrated using a 0.5-in. pressure microphone (B&K 4134) and sound-level meter (B&K 2215) mounted to a Zwislocki coupler (DB4005). The AC output of the sound-level meter was delivered to a spectrum analyzer (Hewlett-Packard model 35665). In all cases the ER-2 earphones were connected to the ER-10A OAE microphone in the same configuration used for the OAE measurements. An ER10-14 foam ear tip was placed on the OAE microphone, and the ear tip was completely inserted into the ear-canal extension of the coupler. Swept-sine measurements of the frequency response of the ER-2 itself in this configuration revealed a uniform response within ±1 dB between 0.5 and 6.5 kHz. The net frequency response of the entire sound-delivery system was evaluated using computer-generated white noise. Irregularities in the frequency response were corrected by adjusting the magnitude spectrum in the noise-generation program. When the calibration procedure was checked by placing the microphone assembly in the coupler as described above, the stimulus levels were verified to be within 1 dB of the intended values.[2]

Immediately prior to each data-collection block, with the microphone assembly fitted in the ear canal, the level of a 500-Hz calibration tone was adjusted to produce 65 dB SPL, and the resulting calibration factor was used to set presentation levels for the tone and noise waveforms to be used during that block of triplets. Next, a series of about 16 triplets was presented over the course of about 50 s, the rms value of each nSFOAE was calculated and saved, and that distribution of values was used later when deciding about which individual responses to include in the accumulating average nSFOAE and which to exclude as being possibly contaminated by physiological noise. Then a recording was obtained during a 20-s period of silence to provide an estimate of the ambient noise level in the ear canal in the absence of our acoustic stimulus. Actual data collection followed these preliminaries.

During the data-collection process, the difference waveforms obtained from individual triplets were evaluated in multiple ways prior to being accepted for use in the accumulating average. For the first 20 of the 50 individual nSFOAE responses making up an averaged nSFOAE, the difference waveform from each triplet had to satisfy the noise and artifact criteria in two ways. First, the difference waveform was compared to the distribution of 16 responses described in the preceding paragraph and was included in the accumulating nSFOAE average only if its rms level was less than 0.25 standard deviation (SD) above the median of that distribution. Second, the difference waveform from each triplet was subtracted point-for-point from the accumulating nSFOAE average, the rms of that difference was calculated, and only if the magnitude of that rms value was less than 6 dB above the ambient noise level in the quiet (as measured earlier) was that individual nSFOAE response added to the accumulating average. For the final 30 of the individual nSFOAE responses, the difference waveform from a triplet was added to the accumulating average only if it satisfied the second of these criteria. Typically about 60–65 triplets needed to be presented to acquire 50 usable difference waveforms from our highly experienced subjects.

Three basic stimulus configurations were used for testing: (1) a tone presented alone, (2) a sample of noise presented alone, and (3) a tone presented along with a sample of noise. In the majority of conditions, the tone was 4.0 kHz and 500 ms in duration, but for some conditions, the duration was reduced to 10 ms. The tone always was gated using a 5-ms cosine-squared rise and decay, and the noise was gated using a 2-ms rise and decay. For the majority of the measurements reported, the onset of the noise lagged the onset of the tone by 100 ms. Typically, the noise was band-limited between 0.1 and 6.0 kHz and had an overall level of about 63 dB SPL (a spectrum level of about 25 dB), but for some conditions, only lowpass, bandpass, or highpass versions of the noise were used. Lowpass and highpass noise bands had filter cutoffs 200 Hz below or above the tone frequency, respectively, and the bandpass noise was the complement of this arrangement—a 400-Hz noise band centered on the tone frequency. The noise was digitally synthesized using an inverse fast-Fourier-transform (FFT) procedure, which provided infinitely steep spectral cutoffs prior to transduction by the ER-2 earphones. The duration of the noise was varied across conditions, but typically was 400 ms.

The tone and noise always were the same synthesized waveforms for all triplets of a condition, and their starting phases always were the same across presentations, conditions, and subjects. Stimulus presentations within and between triplets were separated by silent intervals of approxi-

mately 500-ms duration except that after every sixth presentation (two triplets), the silent interval was about 2000 ms to allow for the necessary calculations regarding inclusion. To minimize any periodicities from the stimulus train, the actual separations between successive presentations varied randomly between 490 and 510 ms in 1-ms steps.

The averaged nSFOAE waveforms were analyzed offline by taking successive 20-ms time segments and measuring their rms levels at the output of a 400-Hz-wide elliptic bandpass digital filter (sixth order) centered on the frequency of the 4.0-kHz tone. This filter bandwidth corresponds approximately to one critical band at 4.0 kHz. The time window was rectangular and typically it was moved in 1-ms steps beginning at stimulus onset. This succession of rms values was converted to decibels sound-pressure level and used as the nSFOAE response over time. When the duration of the tone was 10 ms, a 10-ms window was used to analyze the nSFOAE response in order to improve temporal resolution. (The 20-ms window typically used here was a compromise between the amount of smoothing and the precision of temporal resolution.) For purposes of presentation, only every fifth value from the moving analysis window was plotted in the figures.

Various aspects of the procedure contributed to stable results. The subjects reclined in a comfortable armchair, and the head was held in position using a pillow. The subjects all were highly experienced and had learned to remain quite still during testing. The levels of the tone and/or noise were measured in the ear canal and adjusted as necessary prior to every block of triplets. The difference waveform from each triplet was compared with the accumulated average of difference waveforms for that block before being added to that average (see above), meaning that changes in the fit of the probe tip in the ear canal easily could be detected, at which point the block was restarted.

The Institutional Review Board at The University of Texas at Austin approved this research protocol.

III. RESULTS

A. Tone alone

Tones of 4.0 kHz and 400 ms in duration were presented at multiple sound-pressure levels to each of the subjects. Representative results are shown in Fig. 1. As can be seen, the strength of the nSFOAE remained essentially constant throughout the duration of the tone, and the nSFOAE increased monotonically in magnitude as the level of the tone was increased; that is, the magnitude of the nonlinear component of the SFOAE increased. Similar results were obtained for the other subjects, although two subjects had low-frequency fluctuations in the nSFOAE response magnitude when a tone of about 70 or 75 dB was presented, likely reflecting activation of the middle-ear reflex (MER).

B. Noise alone

Typically, the same sample of synthesized wideband noise (0.1–6.0 kHz) was used across subjects and for all conditions. When this sample of noise was presented alone, the nSFOAE response did increase monotonically with the level of the noise, but the response was more irregular than the nSFOAE to a tone presented alone (see Fig. 2). Much of this irregularity apparently is attributable to the nSFOAE "following" the fluctuations in amplitude that are inherent in the envelope of the noise waveform. This interpretation is supported by the similarity in the shapes of the nSFOAEs obtained for different subjects and at different noise levels (also, see Fig. 3 below).

As evidence that the wide fluctuations in the nSFOAE

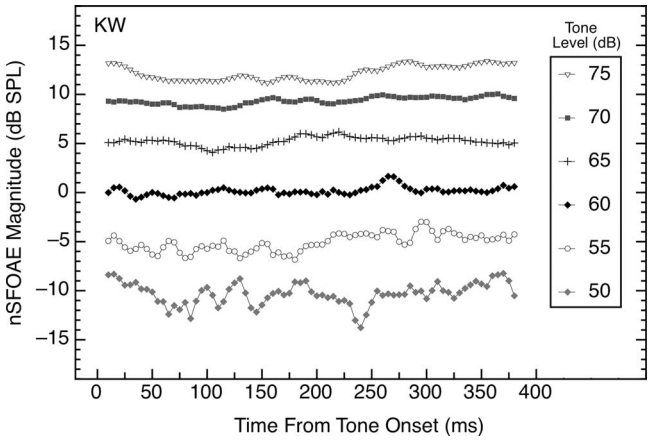


FIG. 1. Magnitudes and time courses of the nSFOAE responses to tones (4.0 kHz, 400 ms) of differing sound-pressure levels. Data are from subject KW, but they are representative of data from other subjects. The overall nSFOAE response was obtained by averaging the difference waveforms from 50 triplets; then the first 20-ms segment of that averaged waveform was bandpass filtered at 4.0 kHz (filter width of 400 Hz) and the rms level calculated. The filter window was advanced in 1-ms steps, and at each step, the rms level was calculated again. For this, and for all other figures, only every fifth value from the moving window is plotted. The first data point of each function was deleted because it reflected the 5-ms rise time of the stimulus. Additional data points at the end of each function also were deleted once there were too few values remaining to fill the 20-ms analysis window. Note that here, and in all subsequent figures, the abscissa value of each data point marks the beginning of a 20-ms (or 10-ms) window of analysis. Thus, the data point at 20 ms represents the strength of the nSFOAE response from 20 to 40 ms. For this and all subsequent figures, the noise floor of the measurement system was about −15 dB SPL.

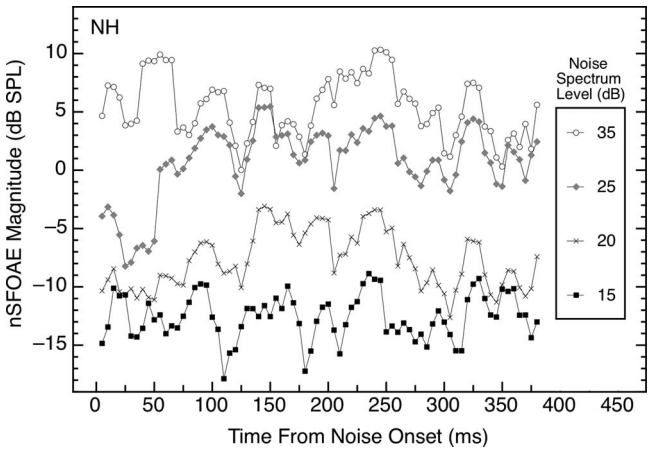


FIG. 2. Magnitudes and time courses of the nSFOAE responses to the same sample of wideband noise (0.1–6.0 kHz, 400 ms) presented at differing spectrum levels. Data are from subject NH, but they are representative of data from other subjects. The average nSFOAE waveform to wideband noise-alone was processed in the same way as for tone-alone (see caption for Fig. 1).

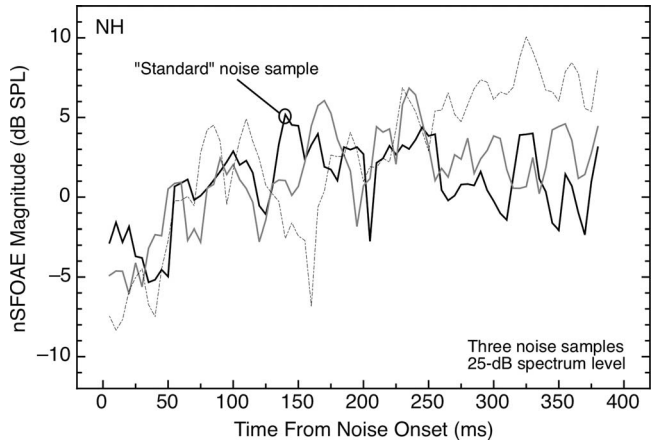


FIG. 3. Magnitudes and time courses of the nSFOAE responses to three different samples of synthesized noise (each 0.1–6.0 kHz, 25 dB in spectrum level, and 400 ms in duration). The basic forms of the results are highly similar, but the patterns of fluctuation differed across the samples of noise. The "standard" sample was the one used for Fig. 2 and elsewhere. Results are shown for subject NH; one additional subject was tested and showed similar results.

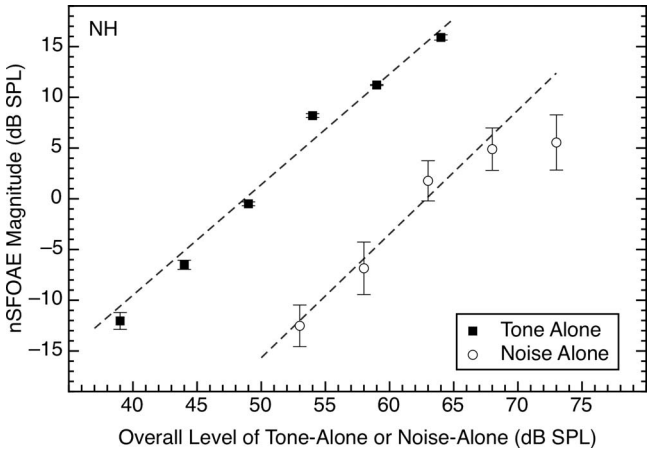


FIG. 4. Effect of stimulus level on the nSFOAE response when the 4.0-kHz tone or the wideband noise was presented alone. The abscissa values for noise-alone denote overall level. The estimates of nSFOAE were obtained from the output of a 400-Hz-wide filter centered at 4.0 kHz. The values plotted are the nSFOAE responses averaged over a 200-ms time period beginning 150 ms after the onset of the 400-ms tone or noise. Error bars show one standard deviation. The moment-to-moment variability in the tone-alone responses (see Fig. 1) was so small that the error bars are barely discriminable from the symbols themselves except at the lowest tone level, while the moment-to-moment variability in the response to the noise clearly was affected by the variability in amplitude in the wideband noise sample at all levels (see Fig. 2). For this subject, the slopes of the best-fitting straight lines were 1.09 and 1.22 for the tone-alone and noise-alone data, respectively.

response to noise-alone in Fig. 2 are attributable to the envelope characteristics of the specific noise sample used to collect the data, nSFOAEs were collected for the noise-alone condition using two additional samples of wideband noise. The results are shown in Fig. 3, where it can be seen that the three different noise waveforms produced responses having unique irregularities but generally similar overall patterns.

Part of the overall pattern of response to noise-alone in Fig. 3 is that the nSFOAE can exhibit a slowly rising, dynamic segment over at least the first 100–200 ms of presentation. This effect occurs only for a relatively narrow range of noise spectrum levels. Careful examination of Fig. 2 reveals a rising, dynamic segment to the response for spectrum levels of 20 and 25 dB, but not for noises stronger or weaker than that, and this was true for other ears as well. For the weakest noise level used (15 dB spectrum level), a horizontal line appeared to fit the nSFOAE responses over the full 400-ms duration of the noise-alone presentation for all four subjects, just like the tone-alone response.

Figure 4 summarizes the effect of stimulus level on the magnitude of the nSFOAE response for both the tone-alone and the noise-alone conditions (Figs. 1 and 2). Similar data were obtained for all four subjects, and regression lines were fitted to the individual data. Across subjects, the average slopes were 1.01 (SD=0.24) and 1.13 (SD=0.18) for tone-alone and noise-alone conditions, respectively, and for each of the subjects, the pairs of slopes were quite similar. For comparison, Shairer *et al.* (2006, p. 910) reported increases of about 0.75 dB per decibel of tone level using their SFOAE procedure with a 4-kHz tone.

To obtain the values used for the fits plotted in Fig. 4, the magnitudes of the nSFOAE responses were averaged across successive 20-ms analysis windows over the range of 150–350 ms of the 400-ms common duration of the sounds. This range was selected in order to emphasize the steady-state segments of the responses. Note in Fig. 4 that when the overall noise level was about 73 dB (35 dB spectrum level), the magnitude of the nSFOAE response did not increase the way it had at weaker noise levels; accordingly, that value was not included in the fitting of the straight line to the noise-alone data.

C. Tone plus noise

When the tone and the wideband noise were presented together, the result was distinctly different from the response to tone-alone even though the noise was weak compared to the tone. After a decline in nSFOAE magnitude that began immediately upon noise onset and lasted about 25 ms, the nSFOAE to the tone-plus-noise increased substantially in magnitude and then asymptoted at a level about 4–15 dB above where it began, depending upon the individual subject. Typical results for two subjects are shown in Fig. 5 for the condition in which the wideband noise (0.1–6.0 kHz) had a spectrum level of 25 dB and the 4.0-kHz tone was 60 dB; the tone was 500 ms in duration and the 400-ms noise was gated on 100 ms after the onset of the tone. For the subject in the top panel of Fig. 5, the immediate decline in the nSFOAE response following noise onset was sharper than for the subject in the bottom panel. (Evidence presented below suggests that this immediate decline upon noise onset is attributable to lateral, or two-tone, suppression.) The difference between the nSFOAE response to tone-alone and the response at asymptote to tone-plus-noise was about 7 dB for subject JZ (top panel) and about 8 dB for subject NH (bottom panel). Note that in this condition, the level of the tone in the 400-Hz analysis band centered at 4.0 kHz averaged about 9 dB higher than the overall level of the noise in that band, suggesting that in the tone-plus-noise condition, the nSFOAE

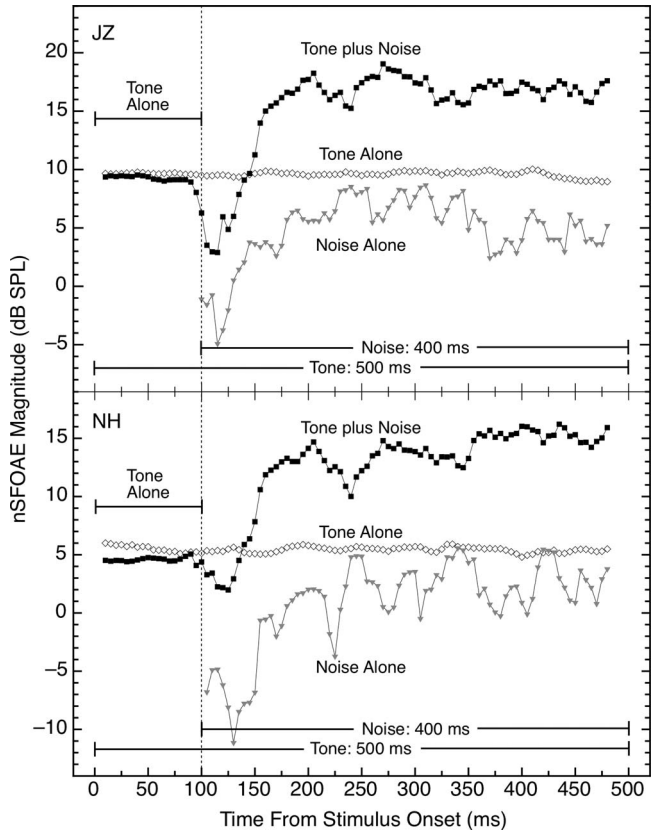


FIG. 5. Magnitude and time course of the nSFOAE response to tone-plus-noise compared to the response for tone-alone and noise-alone. For both subjects, all data shown were collected in the same test session. The magnitude of the immediate decline in nSFOAE response following noise onset was different for the two subjects. The level of the 4.0-kHz tone was 60 dB and the spectrum level of the wideband noise was 25 dB; the tone was 500 ms in duration and the noise was 400 ms in duration. Note the hesitation of about 25 ms prior to the beginning of the dynamic rise in the nSFOAE response for tone-plus-noise during which the nSFOAE magnitude decreased. The analysis procedure was the same as for Fig. 1. Note that the specific values on the ordinates are different for the two panels but the scales are the same.

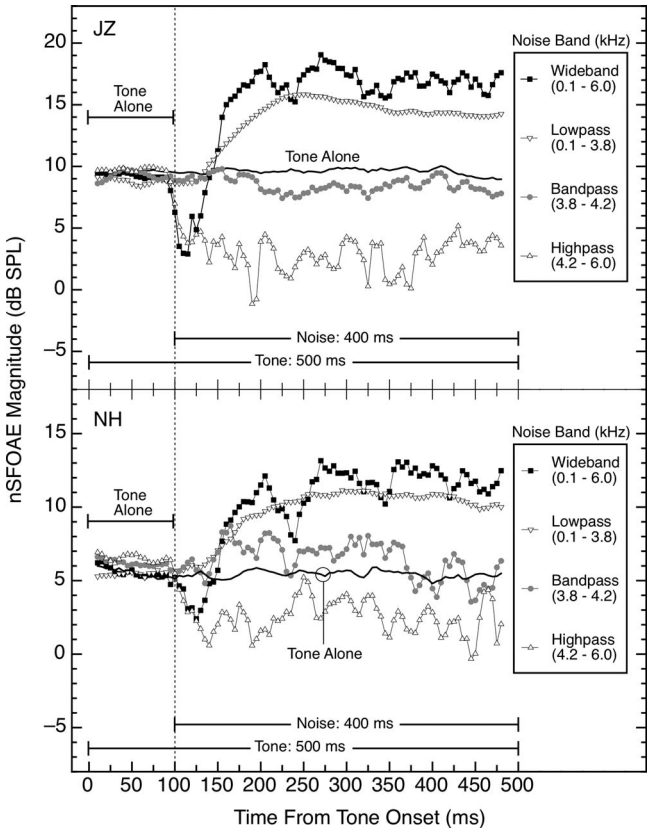


FIG. 6. Magnitudes and time courses of the nSFOAE responses in the tone-plus-noise condition for noises of differing spectral configurations. For both subjects, all data shown were collected in the same test session, and for subject JZ, the tone-alone and tone-plus-wideband noise data also are shown in Fig. 5. For these conditions, the 4.0-kHz tone was 60 dB and the spectrum levels of the noise bands were fixed at 25 dB. Note that the ordinate scale is expanded in the bottom panel.

response consisted primarily of the cochlear response to the tone. Also, the nSFOAE response to tone-plus-noise was not a simple combination of the responses to tone-alone and noise-alone (below we suggest that activation of the efferent system in the tone-plus-noise condition may explain this departure from simple additivity).

For each subject, the gradually rising nSFOAE response to tone-plus-noise was fitted with an exponential function of the form $y=a(1-e^{bx})+c$, beginning 25 ms after noise onset. When the exponential function was fitted to the averaged nSFOAE waveforms obtained using our standard 20-ms time window, the fits generally were very good, and the resulting time constants were 30.3, 23.5, 38.5, and 23.0 ms for subjects KW, JZ, SC, and NH, respectively. However, these values are strongly dependent upon the length of the time window used for the analysis. When the exponential function was fitted to the rising, dynamic response obtained from the same averaged waveforms using a 40-ms analysis window, the estimated time constants were in the range of about 70–90 ms.

D. Spectral characteristics of the noise

Additional measurements revealed that the magnitude and time course of the dynamic response illustrated in Fig. 5 are highly dependent upon the spectral characteristics of the noise band used. Specifically, when the noise was low-passed below the frequency of the 4.0-kHz tone (i.e., from 0.1 to 3.8 kHz), the nSFOAE response was similar to the tone-plus-noise response in Fig. 5, but was generally somewhat reduced in maximum magnitude. However, when the noise was either high-passed above the frequency of the tone (i.e., from 4.2 to 6.0 kHz) or band-passed around the frequency of the tone (i.e., from 3.8 to 4.2 kHz), the nSFOAE response did not show the characteristic dynamic rise seen in the wideband condition. That is, the mechanism operating to produce the rising, dynamic response at each specific frequency region apparently cannot be activated by acoustic energy just anywhere in the spectrum; rather, there needs to be acoustic energy in the frequency region below (apical to) the specific region of interest. These spectral effects are illustrated in Fig. 6 for two subjects. For all of these noise bandwidths, the spectrum level of the noise was held constant at about 25 dB; the corresponding overall levels for the wideband, lowpass, bandpass, and highpass noises were approximately 63, 61, 51, and 58 dB SPL, respectively. When

overall level was equated across the different noise bands (at 63 dB SPL, same as the wideband noise), the same pattern of results was observed.

Not only did the highpass noise not produce a rising, dynamic response, but at noise onset, the nSFOAE for tone-plus-highpass-noise typically fell rapidly to a value lying somewhere below the responses for both tone-alone and tone-plus-bandpass-noise and remained there for the duration of the stimuli. Similar effects were observed in other subjects with this combination of levels for tone and highpass noise. When the level of the tone was increased by 6 dB and the spectrum level of the highpass noise was decreased by about 5 dB, the rapid fall in nSFOAE magnitude was not observed; rather, the nSFOAE was approximately equal to that seen for tone-alone or tone-plus-bandpass-noise. These facts suggest that the weakening of the nSFOAE response with highpass noise depends upon the relative levels of the tone and noise. We believe that this immediate decline represents a form of lateral (or two-tone) suppression (Kemp and Chum, 1980; Shannon, 1976) operating from the highpass noise to the tone; additional evidence is presented in Sec. III E.

Note that the failure of our bandpass noise to produce a rising, dynamic response surely was dependent, in part, upon the relatively narrow bandwidth of that noise. Bandpass noises of greater bandwidth than used here would activate the low-frequency region that clearly has considerable power to initiate a rising, dynamic response (see Fig. 6). In accord with this reasoning, Lilaonitkul and Guinan (2009a) reported that noise centered on their tone *was* highly effective at triggering their SFOAE measure, unlike the result here, but their noise had a bandwidth of one-half octave. When they used noises of half-octave bandwidth centered either below or above the tone, both were capable of initiating an SFOAE response but the lower noise band was the more effective. That asymmetry is similar to the present results except that our highpass noise produced either a suppressive effect or no change in the nSFOAE response from its strength for tone-alone, depending upon the individual subject and the levels of the tone and the highpass noise.

Supplementary bandwidth-manipulation measurements were made for subject NH using a 2.5-kHz tone and lowpass, bandpass, and highpass noise bands correspondingly shifted down in frequency. The bandpass noise was 400-Hz wide, centered on the 2.5-kHz tone, and the lowpass and highpass noise bands had frequency cutoffs 200 Hz below and above the frequency of the 2.5-kHz tone, respectively. These nSFOAE responses were analyzed using a 250-Hz-wide elliptic bandpass filter centered on the frequency of the 2.5-kHz tone, but otherwise were analyzed the same as were the responses to the 4.0-kHz tone. The pattern of results was basically the same as in Fig. 6, suggesting that the spectral asymmetry is an effect of relative frequency, not absolute frequency.

The lowpass noise used to collect the data shown in Fig. 6 had a bandwidth of 0.1–3.8 kHz. In supplementary measurements with subject KW, the upper shoulder frequency of the lowpass noise was decreased gradually, thereby increasing the frequency separation between the 4.0-kHz tone and the shoulder frequency. The cutoff frequencies for the noise were 3.8, 3.5, 3.0, 2.5, and 2.0 kHz, and the attenuation beyond the cutoff frequency was quite steep (because the noise was digitally synthesized using an inverse FFT). The spectrum level of the noise was held constant across these bandwidths. Once the bandwidth of the lowpass noise was 0.1–2.0 kHz, the nSFOAE response ceased to exhibit a rising segment following noise onset and instead looked like the no-rise responses measured with the bandpass and highpass noise bands in Fig. 6.

E. Contralateral noise

The results in Fig. 6 revealed that the nSFOAE response to tone-plus-highpass-noise could be substantially weaker than the response to tone-alone. One interpretation of this outcome was that the highpass noise failed to activate the mechanism responsible for the rising, dynamic segment of the nSFOAE response, but it did produce lateral (two-tone) suppression (Kemp and Chum, 1980; Shannon, 1976) on the 4.0-kHz region of the basilar membrane. Because lateral suppression is a monaural mechanism, a test of this interpretation would be to move the noise stimulus to the opposite ear.

In accord with this reasoning, additional nSFOAE data were collected from three subjects using a modified procedure for stimulus presentation. Within each block, half of the triplets were the same as used for most of the conditions described so far in this paper; the tone and noise both were presented simultaneously to the right ear and the nSFOAE was extracted from that ear. The remaining triplets involved presenting only the tone to the right (ipsilateral) ear, in which the nSFOAE was measured, while simultaneously presenting the noise to the left (contralateral) ear. The timing and levels of the tone and noise matched those described so far in this paper.[3]

For maximal control of the situational and subject variables, the ipsilateral- and contralateral-noise conditions were alternated within each block of triplets. One pair of successive triplets involved ipsilateral noise, the next pair involved contralateral noise, and so on, for the remainder of the block. Within each block of triplets, the noise was either wideband, lowpass, bandpass, or highpass. The nSFOAE responses were stored and analyzed separately for the two modes of noise presentation, and data were collected from at least 30 triplets for both the ipsilateral- and the contralateral-noise conditions in each block. Representative data are shown in Fig. 7.

The data in the top panel of Fig. 7 were obtained from those triplets in which the noise was ipsilateral to the tone, and thus represent a replication of the data for subject JZ in the top panel of Fig. 6 collected with the same sample of noise. The two sets of results are similar in that both the wideband and the lowpass noises produced rising, dynamic responses and the bandpass noise did not. Also evident in both data sets is the immediate and sustained diminution in nSFOAE magnitude produced by the highpass noise that we believe to be attributable to lateral suppression.

The data in the bottom panel of Fig. 7 were collected using noises contralateral to the nSFOAE measurement ear, which received just the tone. As can be seen, the wideband

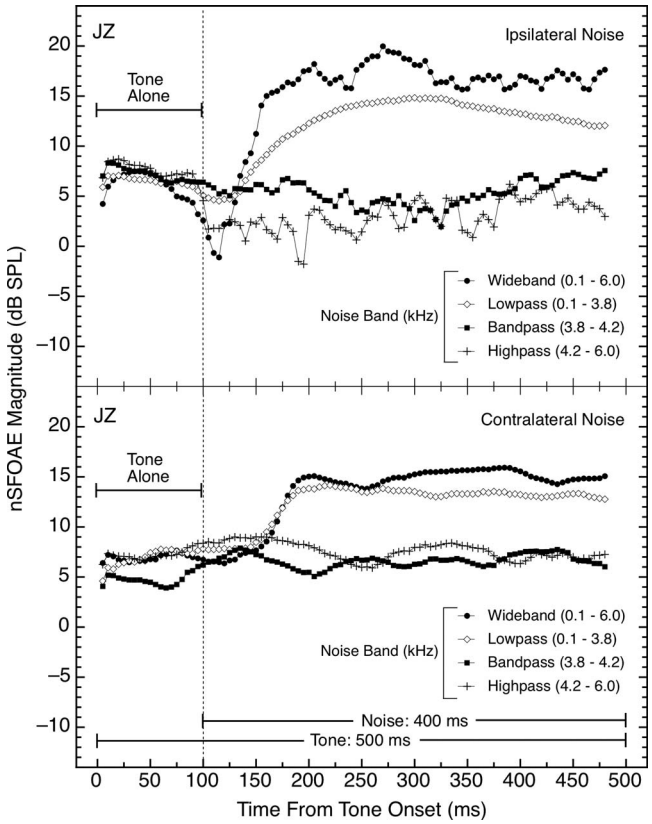


FIG. 7. Magnitudes and time courses of the nSFOAE responses in the tone-plus-noise condition for different noise configurations. The data in the top panel were collected with the noise in the same ear as the tone (ipsilateral), and the data in the bottom panel were collected with the noise in the ear contralateral to the tone and the measured nSFOAE. These two types of triplets were interleaved within each test block. In the contralateral conditions, there was no immediate decline in nSFOAE magnitude with the onset of the wideband noise nor suppression of the nSFOAE with the onset of the highpass noise. For these conditions, the 4.0-kHz tone was 60 dB and the spectrum levels of all the noise bands were fixed at 25 dB. Here only about 30 triplets were averaged for both the ipsilateral- and contralateral-noise conditions rather than the typical 50.

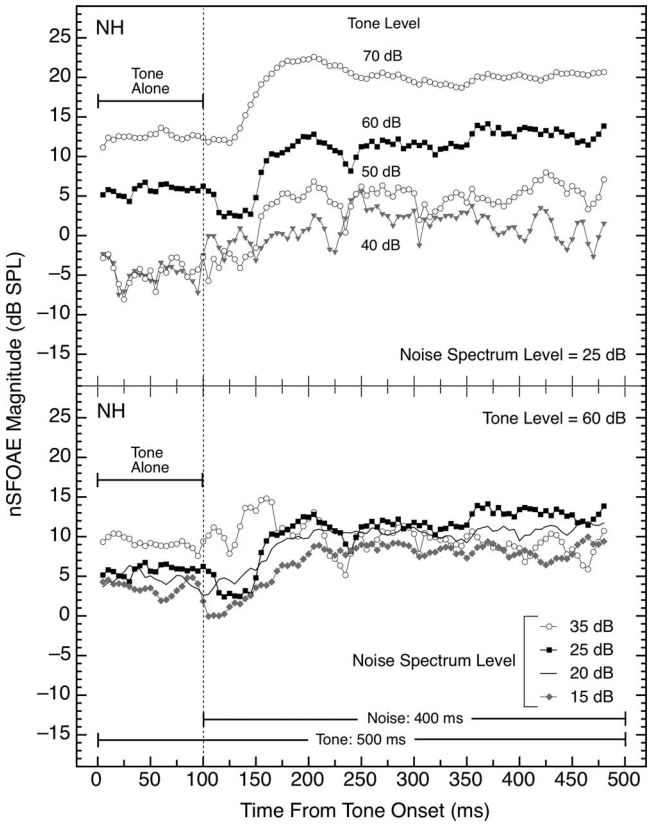


FIG. 8. Magnitudes and time courses of the nSFOAE responses when either the level of the tone was varied and the level of the simultaneous wideband noise was held constant (top panel) or the reverse (bottom panel). The tone was 4.0 kHz and 500 ms in duration; in the top panel, its level was varied as indicated and in the bottom panel, its level was 60 dB SPL. The noise was 0.1–6.0 kHz in width and 400 ms in duration; in the top panel, its spectrum level was 25 dB and in the bottom panel, its level was varied as indicated. The level of the tone had more effect on the magnitude of the nSFOAE response than did the level of the noise. The solid black symbols in the two panels are the same data. Similar data were obtained from additional subjects.

and lowpass noises again produced rising, dynamic responses for the tone (in the ipsilateral, tone-alone ear), and the bandpass and lowpass noises both led to no change in the nSFOAE response from the initial tone-alone interval. This suggests that the sustained, weakened response to ipsilateral highpass noise (top panel of Fig. 7) may have been attributable to lateral suppression (Kemp and Chum, 1980; Shannon, 1976); additional data collection on this point is in progress. Contralateral-noise presentations also were used with subjects NH and KW, and the outcomes were the same as shown in the bottom panel of Fig. 7. Guinan (2006) also demonstrated the essential equivalence of ipsilateral and contralateral noises on the dynamic behavior of SFOAEs. Note in Fig. 7 that the hesitation between noise onset and the beginning of the rising, dynamic response was slightly longer for the contralateral noises than for the ipsilateral noises, but this was not true for all subjects.

The realization that lateral suppression operates from a higher-frequency region onto the 4.0-kHz region of our nSFOAE measurement explains the momentary declines in nSFOAE magnitude often seen immediately after the onset of a wideband noise (see Figs. 5 and 6 and below). These declines can be attributed to the fast action of lateral suppression that is then gradually offset by an opposing mechanism triggered by the wideband noise.

F. Graded nSFOAE responses

When only the tone or only the wideband noise was presented, the nSFOAE responses were graded in magnitude according to the level of the tone or noise (Fig. 4). When tone-plus-noise was presented, the magnitude of the asymptotic nSFOAE response could be substantially larger than to either tone-alone or noise-alone, and the response also differed dynamically from the responses to tone-alone and noise-alone (Fig. 5). The contributions of stimulus level to nSFOAE magnitude were further explored in an attempt to gain further insight into the underlying mechanisms.

For one set of measurements, the level of the wideband noise was held constant (at 25 dB spectrum level) while the level of the tone was manipulated across blocks of triplets. The result was that the asymptotic magnitudes of the rising nSFOAE response also were graded according to the level of the tone (see top panel of Fig. 8). It is important to realize that, in these tone-plus-noise conditions, there were nSFOAE

responses being produced to both the tone and the noise, and the output of the 400-Hz analysis filter contained both nSFOAE components. As the tone level was decreased, the relative contribution of the noise to the overall nSFOAE began to predominate, and accordingly the responses at 40 and 50 dB became more variable and also more similar in magnitude both to each other and to the noise-alone responses seen in Figs. 2 and 3.

For another set of measurements, the level of the tone was held constant (at 60 dB SPL) while the level of the wideband noise was manipulated across blocks of triplets. The result was that the maximum magnitude of the nSFOAE response changed relatively little over a 20-dB change in the spectrum level of the wideband noise (see bottom panel of Fig. 8). Note that for the highest noise level tested, the rising, dynamic response disappeared in this subject; that is, the response to the tone-plus-noise was nearly constant throughout the 400-ms duration of tone-plus-noise. This absence of a rising, dynamic response with the 35-dB noise level was seen for other subjects as well, even when the level of the tone was raised to 66 dB in the 35-dB noise condition. [Note that the nSFOAE response to noise-alone (Fig. 2) also was essentially constant once the spectrum level of the noise was raised to 35 dB.] Some subjects did show a rising, dynamic response with the 15-dB noise level (like NH in Fig. 8), but other subjects showed an essentially flat nSFOAE, resembling the response to tone-alone.

For both of the manipulations in Fig. 8, it is important to remember that the nSFOAE response is the magnitude of the waveform required to make the sound in the ear canal during the two-earphone presentation equal to the linear sum of the sounds from the two single-earphone presentations. The top panel of Fig. 8 shows that this "correction waveform" (the nSFOAE) must increase in magnitude as the level of the tone is increased, and the bottom panel of Fig. 8 shows that the correction waveform is not much affected by the level of the noise. A way to think about these effects in terms of the input/output functions of the cochlea is described below and in the second paper in this series.[1]

G. Recovery following noise termination

Because the nSFOAE to tone-plus-noise is strong (Fig. 5) and the nSFOAE to tone-alone is weak (Fig. 1), it is logical to expect that if the tone were to outlast the duration of the noise, then the nSFOAE to the tone should begin to weaken, perhaps rapidly, and move toward the response magnitude seen with tone-alone. A test of that expectation proved it wrong. For all subjects, after the 400-ms noise was terminated, the nSFOAE response to a 60-dB tone remained strong, and approximately the same magnitude as during the noise, with the return to tone-alone values requiring hundreds of milliseconds after noise termination. To further explore this persistence of the nSFOAE response following termination of the noise, we varied the duration of the noise from 25 to 200 ms while retaining the duration of the tone at 500 ms. The results are shown in Fig. 9.

As Fig. 9 reveals, when the wideband noise lasted 100 or 200 ms, the elevated nSFOAE response to the tone per-

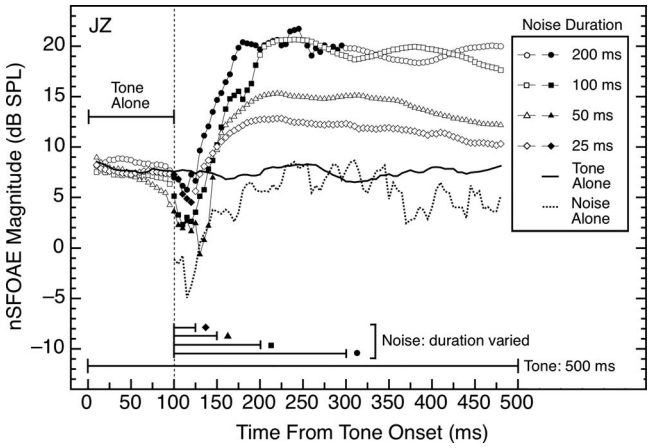


FIG. 9. Magnitudes and time courses of the nSFOAE responses to a 4.0-kHz, 60-dB tone of 500-ms duration when the duration of the wideband noise (0.1–6.0 kHz, 25 dB spectrum level) was varied from 25 to 200 ms. Noise onset was 100 ms after onset of the tone. Solid symbols indicate the presence of the noise for each condition. Note the hesitation of about 25 ms between noise onset and the beginning of the rise of the dynamic response. The noise-alone data from Fig. 5 are shown at the bottom of the figure for comparison. Except for the noise-alone data, all conditions were measured in the same test session. Similar data were obtained from other subjects.

sisted unabated through the remainder of the 400-ms observation period, similar to the persistence seen when the duration of the noise was 400 ms and the tone outlasted it (not shown). When the noise duration was shortened to 50 or 25 ms, the nSFOAE response to the tone did decline somewhat, but it still was elevated slightly at the end of the 500-ms tone when compared with the initial tone-alone period. For some other subjects, the 25-ms noise produced a response to the tone that did decline to the value seen in the initial tone-alone period by the end of the 500-ms observation period. Thus, there are conditions for which the nSFOAE response to the tone does decline gradually following the termination of the noise. However, the protracted persistence in the other conditions surely is more notable than the fact that in extreme conditions, the response does recover in accord with intuition. Apparently the wideband noise (or at least its low-frequency region) is necessary to activate the mechanism responsible for the gradual rise in the nSFOAE response to the tone-plus-noise (tone-alone is inadequate to activate that mechanism; see Fig. 1), but once the mechanism has been activated, the nSFOAE response can persist for hundreds of milliseconds. When straight lines were fitted to the segments of the responses between 225 and 500 ms in Fig. 9, the slopes were about 1 dB per 100 ms for both the 25- and 50-ms noises.

In passing we note that the responses during the tone-alone periods that preceded noise onset were very similar in magnitude for all noise durations in Fig. 9 (and in other subjects). This suggests that the 500-ms silent intervals between successive stimulus presentations in each triplet were adequate to allow recovery of whatever physiological mechanism is responsible for the rising, dynamic response to the tone-plus-noise.

Notice in Fig. 9 that a 25-ms burst of noise was adequate to initiate an increase in magnitude of the nSFOAE response even though that noise duration did not exceed the approxi-

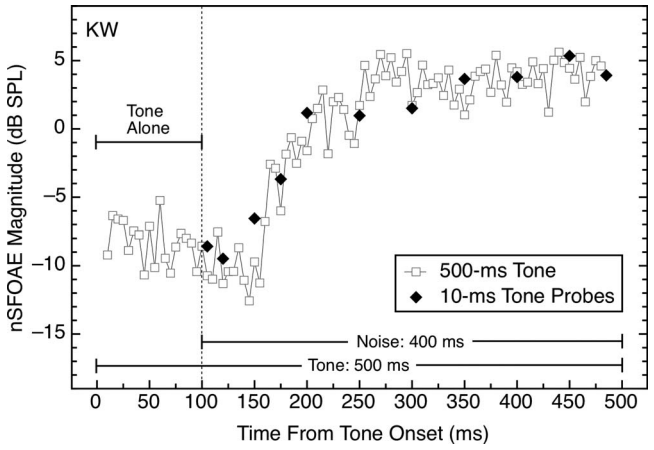


FIG. 10. Magnitudes and time courses of the nSFOAE responses to long and short tones (at 60 dB SPL) presented with a wideband noise (0.1–6.0 kHz; 25 dB spectrum level). The 10-ms tone bursts produced essentially identical nSFOAEs as the 500-ms tone at corresponding points in time after noise onset, demonstrating that short probe tones are capable of revealing the state of the dynamic process activated in the tone-plus-noise conditions. The duration of the analysis window was 10 ms for both conditions.

mate duration of the hesitation present in all the responses shown in Fig. 9. That is, once triggered, the rising nSFOAE response takes time to develop.

H. Short vs long tone durations

All of the data shown so far were collected using tones having durations of 400 ms or longer. Data also were collected using 10-ms tones presented at different delays relative to the onset of the noise. One basic question was whether probe tones of this sort would be adequate to reveal the functional relationships shown in Figs. 5 and 6. The basic answer was yes; whatever mechanisms are activated by the presentation of the noise, they produced nSFOAE responses to short tones that were highly similar to those for long tones, both in magnitude and time course. This is further evidence that the relatively intense, long-duration tone used in all preceding demonstrations contributed little or nothing to the *initiation* of the dynamic response to that tone.

A comparison of short- and long-duration tones is provided in Fig. 10. The close agreement between the responses is representative of the data for other subjects. Although not shown, short tones also behaved similarly to a long tone in the time period following the *offset* of the wideband noise (see Fig. 9 above). Thus, short tones apparently can be used safely as probes of the state of the dynamic process illustrated in Figs. 5, 6, 8, and 9. For the purposes of Fig. 10, the duration of the rectangular analysis window was shortened from 20 to 10 ms in order to match the duration of the 10-ms tone bursts. This brief window explains the higher variability in the long-tone data than was present in Fig. 5, 6, or 8.

IV. DISCUSSION

The purpose of this report was to describe the results obtained when a new, nonlinear procedure for measuring SFOAEs was used with stimuli of the sort commonly employed in psychoacoustical studies of such phenomena as simultaneous and forward masking. In a companion paper,[1] nSFOAEs and behavioral performance were measured using the same acoustic waveforms, the same subjects, and the same ears.

To summarize the results of this report, when the stimulus was a 4.0-kHz test tone presented alone, the nSFOAE response increased monotonically in magnitude as tone level was increased, but for all levels, the nSFOAE magnitude was essentially constant throughout the 400-ms duration of the tone (see Fig. 1). When the stimulus was a wideband noise alone, the magnitude of the response again varied directly with level, but the form of the response differed depending upon noise level. At moderate noise levels for most subjects, there was a gradual increase in the nSFOAE response followed by a segment with essentially constant magnitude (see Figs. 2 and 3). When the tone and the wideband noise were presented simultaneously, the result seemingly was synergistic in that the maximum magnitude of the nSFOAE response was much greater than would be expected from simple additivity of the responses to tone-alone and noise-alone (see Fig. 5). This dynamic response showed an initial hesitation of about 25 ms followed by a rapid rise in the response that required about 100 ms to complete; then the nSFOAE response was approximately constant until the end of the 400-ms noise presentation (see Figs. 5, 6, 8, and 10). During the hesitation period, the nSFOAE often showed an immediate, sharp decline in magnitude that was not present when the noise was only in the contralateral ear or when an ipsilateral noise was lower in frequency than the tone, suggesting that the sharp decline is attributable to lateral suppression operating from a higher frequency region on to the tone. The contralateral noise did produce a rising, dynamic response to the ipsilateral tone like that seen with an ipsilateral noise.

Subsequent manipulations revealed that it was primarily the low-frequency components of the noise that were responsible for the rising, dynamic segment of the nSFOAE response; highpass and bandpass versions of the noise were incapable of producing the dynamic response, but rather left the magnitude of the nSFOAE close to or weaker than the level of response to the tone alone (see Fig. 6). Changing the tone to 2.5 kHz, and changing the cutoff frequencies for the various noise bands accordingly, demonstrated that these bandwidth-manipulation outcomes were a matter of relative not absolute frequency. When the tone was continued past the termination of the noise, the nSFOAE response remained elevated in magnitude for hundreds of milliseconds (see Fig. 9). Finally, when the long-duration tone was replaced by a 10-ms tone burst during or after the noise presentation, the nSFOAE responses were essentially identical to those observed with the long-duration tone (see Fig. 10).

A. Comparison of SFOAE methods

The fact that the dynamic nSFOAE response to tone-plus-noise shows a 25-ms hesitation at noise onset and is well characterized by a short time constant suggests that this measure is strongly related to the similar OAE measures long studied by Guinan and his colleagues (e.g., Guinan, 2006; Guinan *et al.*, 2003; Backus and Guinan, 2006; Lilaonitkul

and Guinan, 2009a, 2009b). Thus, the rising, dynamic segment of the nSFOAE response likely is attributable to the medial olivocochlear (MOC) reflex that Guinan so elegantly demonstrated to underlie the dynamic response he saw in a version of the SFOAE extracted using a heterodyne procedure. Although Guinan typically presented the noise to the contralateral ear, typically used a diotic 1.0-kHz tone, and did not employ a triplet procedure, among other differences from our procedure, nevertheless he did extract a perstimulatory SFOAE from the sound in the ear canal, so the existence of similarities across the two data sets is encouraging. The similarities can be summarized as follows:

- Both procedures yield a constant response to a tone presented alone (which Guinan called the SFOAE), and a rising response to tone-plus-noise (which Guinan called the ΔSFOAE; see Guinan *et al.*, 2003; Backus and Guinan, 2006). [The numerous differences in procedure have led us not to adopt the Guinan terminology for fear of possibly confusing (clearly related) phenomena.]
- After the onset of an ipsilateral noise, there is a hesitation of about 25 ms before the beginning of the rising, dynamic response (see Figs. 5, 6, and 9; Backus and Guinan, 2006).
- Although Guinan and colleagues (e.g., Guinan, 2006; Backus and Guinan, 2006) were able to extract multiple time constants from their measures, presumably attributable to multiple underlying mechanisms, the ears with the strongest dynamic responses did reveal an initial, fast time constant much like those shown by our subjects. The fast time constants reported by Guinan and colleagues were about 60–80 ms; ours were about 23–39 ms when the exponential function was fitted to the response obtained with our typical 20-ms analysis window and about 70–90 ms when fitted to the response obtained with a 40-ms window. By comparison, Backus and Guinan (2006) noted that a time constant of about 70 ms emerges from a wide array of different measures of efferent activity and thus may represent a "fundamental time constant" of the MOC system.
- When a sample of typical nSFOAE responses was analyzed with a heterodyne procedure similar to Guinan's, the form of the data, including the hesitation, the rising, dynamic response, the magnitude of the rise, and the fitted time constants were remarkably similar to the values obtained with our standard analysis procedure.
- The nSFOAE response sometimes shows large variation across test sessions (described below), and Lilaonitkul and Guinan (2009a) reported the same.

There also are some differences in the outcomes obtained with Guinan's and our procedures:

- According to Guinan (2006, p. 601), when the bandwidth of the noise was manipulated, significant MOC effects could be observed (at least "in some cases") with noise bands located either well below or well above their 1.0-kHz tone. Lilaonitkul and Guinan (2009a) also reported that lowpass noise was more effective than highpass noise at producing a dynamic change in their SFOAE measure, but the most effective noises were those centered on the test tone. By comparison, both our bandpass and highpass noises were ineffective at triggering a rising, dynamic response. Some of this difference surely is attributable to Lilaonitkul and Guinan's use of half-octave bandwidths compared to the narrower bandpass noise used here, but the failure of our highpass noise to initiate a rising, dynamic response is unlikely to be simply a bandwidth effect.
- During the hesitation period prior to the beginning of the rising, dynamic response, the nSFOAE sometimes shows an immediate, sharp decline following noise onset (see Figs. 5, 6, and 9). Multiple lines of evidence suggest that this decline is attributable to lateral (two-tone) suppression acting on the tone from components of the noise higher in frequency than the tone (see Figs. 6 and 7). Guinan and colleagues took steps to avoid complications from suppression in their measurements (e.g., Lilaonitkul and Guinan, 2009a, 2009b). Commonly they employ a tonal suppressor on the low-frequency side of their test tone, and their measurement window is located after the suppression has declined away. In other studies, for their ipsilateral conditions, their MOC-eliciting noises exclude that band of frequencies surrounding their tone that are capable of producing suppression on their tone. For a 1.0-kHz target tone, Backus and Guinan (2006) estimated that the spectral notch had to be about 2.1 octaves.
- Although Guinan and colleagues reported similar values of hesitation for ipsilateral- and contralateral-noise bands (Backus and Guinan, 2006), the hesitation for the nSFOAE response was longer for contralateral- than ipsilateral-noise bands for some of our subjects (see Fig. 7).
- Following noise termination, the Guinan response decays relatively rapidly (Backus and Guinan, 2006), unlike the marked persistence seen for the nSFOAE response (see Fig. 8). Note that in primary auditory nerve fibers, the onset of efferent activity also is considerably faster than is the offset of efferent activity (Wiederhold and Kiang, 1970).
- The dynamic response observed by Backus and Guinan (2006, Fig. 5) increased in magnitude with increases in noise level, whereas for the nSFOAE, the level of the noise appears to matter less than the level of the tone (compare top and bottom panels of Fig. 8).
- We have yet to see any subjects having the medium time constants (about 330 ms) reported by Backus and Guinan (2006).

Our procedure for obtaining nSFOAEs has several potential advantages compared to Guinan's measure: when the noise is ipsilateral, it does not need to be notched around the frequency of the tone; short-duration tones can be used as easily as long-duration tones; and the sound of interest can be more complex than a tone. Also, the nSFOAE procedure allows an investigator easily to work in whatever frequency region provides the best signal-to-noise ratio. A potential weakness of the nSFOAE measure is that it is dependent upon the behavior of the SFOAE being nonlinear. Our procedure eliminates the stimulus waveform plus all components of the cochlear response that are linear and leaves only those components not obeying strict additivity. Thus, our

procedure reveals properties of the *behavior* of the underlying SFOAE (e.g., in response to changes in tone and/or noise level), as opposed to the actual *magnitude* of the underlying SFOAE, which is revealed only indirectly. These characteristics of the nSFOAE response may be an advantage in situations where the linear response obscures a phenomenon of interest, but they may be a disadvantage in situations where the actual magnitude of the underlying SFOAE is relevant.

The procedure used here also has marked similarities and differences with the procedure used by Keefe *et al.* (2009), but a detailed comparison is deferred until the second paper in this series.[1]

B. Effects of noise on OAEs

Other investigators also have studied the effect of noise bands of various sorts on an OAE produced by a tone. Unlike the present study, Maison *et al.* (2001) used a transient-evoked OAE (TEOAE), their noise bands were in the ear contralateral to the tone, the effect observed was a diminution in the magnitude of a TEOAE, and those diminutions typically were only 1 dB or less in magnitude. Also unlike the present study, the noise bands producing the greatest attenuation of the TEOAE were those centered on the frequency of the tone (either 1.0 or 2.0 kHz), with little evidence of an asymmetry in the effectiveness of noise bands below or above the tone frequency. Apparently, TEOAEs are not affected by noise in the same way as the perstimulatory nSFOAE described here. The time course of this contralateral attenuation was about 60 ms (Maison *et al.*, 2001).

Both Kim *et al.* (2001) and Bassim *et al.* (2003) measured decreases of about 1–2 dB in a distortion-product OAE when noise was presented to the contralateral ear (this effect is much weaker in humans than in other mammalian species; see discussion by Guinan, 2006). The time constants fitted to those declines showed large individual differences, but generally were about 70 ms, which is similar to the time constants reported by Backus and Guinan (2006) and to ours when a 40-ms analysis window was used (see above). Other work involving the effects of noise on OAEs is discussed in the companion paper.[1]

C. Middle-ear reflex

The parameters for the tones and noise used here were chosen in large part because of our eventual interest in certain psychoacoustical phenomena such as auditory masking.[1] As a result, the noise levels used may have been sufficient to activate the MER in some subjects. In the most commonly used condition, the noise levels were 63 and 69 dB SPL for the single-earphone and two-earphone presentations of each triplet, respectively, the larger of which is close to the nominal MER threshold for broadband noise (Wilson and Margolis, 1999). In adult ears, Shairer *et al.* (2007) found that the range of broadband noise levels capable of eliciting a statistically significant change in acoustic admittance in the ear canal was 64–80 dB Hearing Level. These facts make it logically possible that some or all of the effects reported here were attributable to the MER rather than to cochlear mechanisms and the MOC system.

We are not unequivocally able to rule out a potential influence of the MER on our nSFOAE measures, but several facts suggest that those influences were minimal. (1) The MER primarily acts to attenuate the transmission of low frequencies through the middle ear (Dallos, 1973; Goodman and Keefe, 2006; Shairer *et al.*, 2007). In the present study, the frequency of the probe tone and the frequency at which the nSFOAE was measured was 4.0 kHz. Shairer *et al.* (2007) found that above 2.0 kHz, shifts in acoustic admittance were not significantly different from zero for *any* level of noise activator. Also, shifts in acoustic reflectance from the tympanic membrane were positive—more energy reflected back into the ear canal—only for frequencies below 1.26 kHz. For frequencies between 1.26 and 5.0 kHz, the shift in reflectance was negative, implying that the transmission of these frequencies *improved* following a broadband noise activator. (2) Higher frequencies are more effective elicitors of the MER compared to lower frequencies (Dallos, 1973). When the bandwidth of the noise was manipulated here, the high-frequency components of the noise were found *not* to be effective elicitors of the dynamic tone-plus-noise response, while the low-frequency components were effective elicitors. This clearly is a reversal of the expectations if the rising, dynamic segment of our nSFOAE response were attributable primarily to the MER. (3) The onset latency of the MER is commonly taken to be about 100 ms for a strong stimulus, with shorter latencies at even higher levels (Dallos, 1973; Church and Cudahy, 1984). Goodman and Keefe (2006) measured SFOAEs to a high-frequency probe tone while simultaneously measuring the middle-ear reflex using a low-frequency probe tone. They classified a shift in response to their low-frequency probe as having an MER origin (i.e., noncochlear) only if the onset latency was $\geq$70 ms. By comparison, the rising, dynamic segment of our nSFOAE response begins well before 70–100 ms, and the time constants we obtained were similar to those reported by Backus and Guinan (2006) and Kim *et al.* (2001), both of whom took rigorous steps to rule out contributions from the MER in their investigations of the temporal characteristics of the MOC response. (4) The change in acoustic impedance attributable to the MER increases as the duration of the eliciting stimulus increases, at least up to approximately half a second (Church and Cudahy, 1984). By comparison, our nSFOAE response to tone-plus-noise does asymptote after about 100 ms even when the durations of the tone and noise were as long as half a second. (5) Goodman and Keefe (2006) showed that when the MER was observable at a low frequency, there sometimes was a coincident rapid shift in the magnitude of their high-frequency SFOAE that followed approximately the same time course as the MER. In contrast, once our nSFOAE response had asymptoted, no additional rapid changes in response magnitude ever were observed. (6) In subjects NH and KW, a 66-dB tone was varied in 5-Hz steps around 4.0 kHz in the presence of a wideband noise of 25 dB spectrum level. When the phase of the nSFOAE response could be estimated reliably using a heterodyne procedure similar to Guinan's, the phase did shift systematically with frequency. Guinan used outcomes of this sort to rule out strong contributions from the middle-ear reflex on the dy-

namic component of his SFOAE response (see Guinan *et al.*, 2003). (7) Using a procedure they regarded to be superior to Guinan's measure of phase shift with frequency, Keefe *et al.* (2009) reported no evidence of an MER with noises having overall levels of about 69 and 81 dB compared to the 63 and 69 dB typically used here. (8) In one subject, clear evidence of activation of the MER was observed using noise levels comparable to those reported here; specifically, her nSFOAE to tone-plus-noise oscillated unlike anything seen in the responses of the subjects reported here. This individual difference led us to exclude this subject from further measurements.

Taken together, these facts suggest that the MER was not the primary source of the rising, dynamic response reported here for the tone-plus-noise conditions. However, the work of Goodman and Keefe (2006) suggests caution with interpretation until additional tests are developed to rule out MER effects completely.

D. MOC system and cochlear input/output functions

Following the lead of von Klitzing and Kohlrausch (1994) and Strickland (2004), these results can be explained by assuming that the MOC system operates to modulate the gain of the cochlear amplifiers and hence the form of the input/output function of the cochlea (also see Oxenham and Bacon, 2004). The presumption is that, in its initial, resting state, the gain of the cochlear-amplifier system is high and the relevant input/output function is highly compressive for the middle range of sound-pressure levels. When only a single tone is turned on, there is no activation of the MOC system, and thus no change in gain or in the amount of compression; for tone-alone, the initial input/output function is relevant for all three presentations of each triplet. When the tone is accompanied by a noise, such as the wideband and lowpass noises used here, some mechanism is activated (presumably the MOC efferent system; see Guinan *et al.*, 2003; Guinan, 2006) that, after a short delay, leads to a decrease in the gain of the cochlear amplifiers and an accompanying shift toward input/output functions that are less compressive than the initial function. The two single-earphone presentations of each triplet for tone-plus-noise, being acoustically equivalent, are processed by the same input/output function equally, but the final, two-earphone presentation of each triplet is processed by an input/output function that has slightly lower gain and is slightly less compressive. The result is that, after a short hesitation, the nSFOAE response for tone-plus-noise shows the characteristic dynamic rise seen in Fig. 5 (and in the data of Backus and Guinan, 2006), and that response remains high until the MOC system gradually relaxes (following stimulus offset) and allows the cochlear gain to return to its initial state. This mechanism is discussed in more detail in the companion paper.[1]

Our bandwidth-manipulation results revealed an important point that deserves mention in the context of this MOC explanation of our nSFOAE outcomes. For both 4.0- and 2.5-kHz tones, and for both ipsilateral and contralateral noises, the bandpass noise and the highpass noise were not capable of activating the rising, dynamic response, but the lowpass noise was. This suggests that the rising, dynamic segment of the nSFOAE response is initiated and controlled by a mechanism that operates relatively locally along the basilar membrane. Namely, each local frequency region apparently is controlled by a frequency region lying just below (apical to) it. The width of that lower-frequency region has yet to be defined, but it appears to be relatively wide, and the width is likely to depend upon the absolute frequency. A wideband sound apparently is capable of activating all those local mechanisms, but each mechanism only controls activity in a frequency region lying slightly above (basal to) it. Even a relatively wide noise band apparently is incapable of initiating a rising, dynamic response in a frequency region below that band. These facts clearly carry strong implications about the wiring of the underlying mechanism, be it the MOC system or some other mechanism. Also note that, when a wideband noise is presented with a tone, there seem to be two opposing forces acting simultaneously on the frequency region containing the tone. Frequency components lying above the tone exert a fast-acting suppressive effect on the frequency region of the tone, while frequency components lying below the tone act more slowly to initiate a rising, dynamic response in the frequency region of the tone (see Fig. 6). The interaction of these opposing forces is not simple, however. If a simple algebraic summation were at work, then one would expect the nSFOAE response with the lowpass noise to be stronger than that with the wideband noise, and that did not happen (see Fig. 6). An understanding of this interaction awaits further work.

E. Weaknesses of the nSFOAE measure

Although the stimulus procedure and extraction method described here has produced generally stable and consistent nSFOAE responses, occasional irregularities have been observed. For example, on one occasion, at the end of a long test session, subject SC suddenly showed no rising, dynamic response to the tone-plus-noise stimulus even though she routinely had exhibited that response previously, including earlier that same session. On this occasion, various checks on the apparatus and the procedures revealed nothing awry, and in subsequent sessions, SC's dynamic response again was strong and consistent. In another case, subject JZ was clearly drowsy one day, and his rising, dynamic response was noticeably weaker than typical for him. After a break that included walking and hydration, his dynamic response returned to its typical strength (data shown below). In yet another case, subject KW exhibited no rising, dynamic response to tone-plus-noise when tested soon after participating in a fatiguing sports activity, yet his response returned to its typical value in subsequent sessions. These examples suggest that fatigue of physiological, cognitive, or cochlear origin can alter the mechanisms responsible for the typical cochlear response to tone-plus-noise. These examples also point out a strength of the repeated-measures strategy used in this study; without knowledge from prior and subsequent test sessions, incorrect conclusions surely would have been drawn from these anomalous episodes.

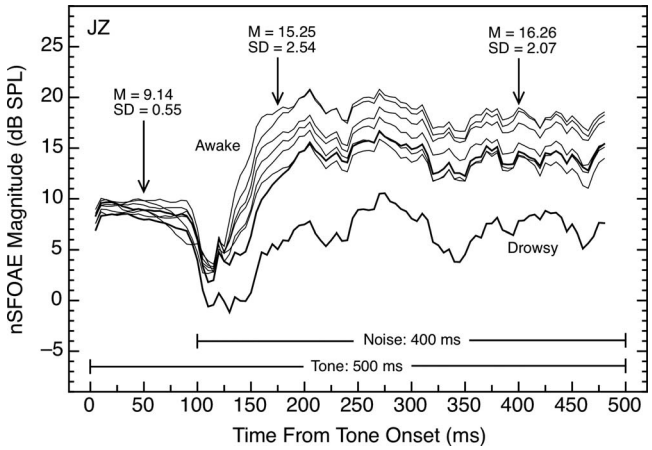


FIG. 11. Magnitudes and time courses of nSFOAE responses to tone-plus-noise for one subject in four different test sessions. The magnitude of the nSFOAE response was quite similar for the initial 100 ms of tone-alone but showed greater variability during the dynamic segment of the response initiated by the onset of the noise. The response obtained while the subject was drowsy was not included in the summary statistics shown; the typically shaped, heavy trace above was obtained after steps were taken to reverse his drowsiness. The average time constant for the six gray traces at the top of the figure was 22.3 ms. Similar patterns of response were seen in other subjects.

The magnitude of the nSFOAE measured to the same stimulus also can vary considerably, both within and across test sessions. Figure 11 contains seven nSFOAE responses to tone-plus-noise collected from the same ear over four test sessions. As can be seen, the rising, dynamic segments initiated by the onset of the noise differed in maximum magnitude (although not much in shape). For this subject, the standard deviation of nSFOAE magnitude across tests was similar at the 75- and 300-ms points following noise onset, suggesting that response magnitude simply differed by a constant across tests. When time constants were estimated for the six gray traces at the top of Fig. 11, the range was 17.2–25.4 ms, and the mean value was 22.3 ms.

Also evident in Fig. 11 is that repeated measurements of the response to the initial 100 ms of tone-alone can be quite similar in strength across tests and sessions, unlike the variation seen for the dynamic segment of the nSFOAE response. Concretely, the standard deviation across the seven tests was about 0.6 dB for the response to tone-alone. Occasional exceptions to this regularity for tone-alone also have been observed (e.g., bottom panel of Fig. 8), but when those nSFOAE responses were normalized to the typical magnitude seen for the tone-alone stimulus, the shape of the response was similar to that seen in more typical sessions. Although there surely were small differences in the placement of the probe tip across these test runs showing variability, the level of the stimuli should have been affected relatively little because stimulus level was adjusted in the ear canal before every test run.

Also shown in Fig. 11, but not included in the summary statistics provided, is the episode described above in which subject JZ was tested while drowsy and then again after a brief, brisk walk and hydration. Note that subject JZ's nSFOAE response to tone-alone was affected much less than his response to tone-plus-noise during his drowsy test session. These examples bolster the implication that the mechanisms responsible for the rising, dynamic segment of the nSFOAE response appear to be affected by everyday fluctuations in alertness or physical fatigue. Guinan and his colleagues also commented about large differences across individuals and test sessions with their SFOAE measure (e.g., Lilaonitkul and Guinan, 2009a).

As mentioned above, one weakness of our nSFOAE measure is that it exists only when a nonlinear component is present in the SFOAE. Also, while our nSFOAE measure does reveal the existence of a nonlinear component in the SFOAE (a failure of simple additivity), it does not reveal the origin of that nonlinear component. For example, the same magnitude of nSFOAE response could arise from a compressive process, as described above, or from an expansive one.

F. Final comment

It is likely that some auditory psychophysical tasks are heavily affected by the initial responses in the cochlea and other psychophysical tasks less so. Also, some characteristics of particular psychophysical tasks might be determined in the cochlea and other characteristics might be determined at higher neural levels. Having some simple, noninvasive measures that are capable of helping investigators determine when the initial cochlear responses are crucial and when they are not obviously would be extremely valuable for both basic and clinical auditory research. At this point it is unclear whether the perstimulatory nSFOAE measure described here will prove to have such value. In this report we have described the basic characteristics of the nSFOAE response as some fundamental parameters of the acoustic stimulus were manipulated. In a companion paper,[1] we report that the nSFOAE response does behave in accord with psychophysical performance in certain auditory masking conditions, but it also behaves contrary to psychophysical performance in some other regards. The interim judgment, then, appears to be that the glass is half full. Whether that is a good thing or an irrelevant thing, only more research can tell.

ACKNOWLEDGMENTS

This work was supported by a research grant awarded to D.M. by the National Institute on Deafness and other Communication Disorders (NIDCD 00153). K.P.W. conducted this and additional research on this topic while working on a Master's degree at The University of Texas (Walsh, 2009). Early stages of the work were reported (Walsh *et al.*, 2008, 2009). The work profited greatly from discussions with Dr. C. A. Champlin, Dr. E. A. Strickland, Dr. M. Wojtczak, and Dr. N. F. Viemeister. Comments from two anonymous reviewers are gratefully acknowledged.

[1]A companion paper that serves as the second paper in this series has been submitted to Hearing Research and is under revision ("Overshoot measured physiologically and psychophysically in the same human ears"). The interested reader is referred to this paper for detailed comparisons between our nSFOAE measure and performance on some common psychoacoustical tasks.

[2]Initially the sound-pressure levels of the stimuli in the ear canal were estimated and adjusted using just the ER-10A microphone to monitor the

outputs of the ER-2 earphones, and an entire set of data was collected using that procedure. Later use of the Zwislocki coupler revealed that the noise spectrum was not as flat as desired and the tones were stronger than believed. The uncorrected noise spectrum "drooped" from the target value between about 1.5 and 6.5 kHz, with the maximum deviation being about −4 dB in the range from about 3.5 to 4.5 kHz. All of the data were re-collected with those errors in level corrected, and all basic outcomes were confirmed. This reveals that these findings are not critically dependent upon the levels of the sounds; also, every outcome shown is a verification of an earlier, but unshown, measurement in that same subject.

[3]Noise was presented contralaterally by passing the synthesized noise waveform to a second digital-to-analog converter board (PCI-4451, National Instruments, Austin, TX) installed in the G4 computer. The analog output was amplified and delivered to a third ER-2 insert earphone that was fitted into the subject's left (contralateral) ear canal using an ER3-14A foam eartip. The gain and frequency response of this system was measured with the coupler and sound-level meter described above. Prior to each test run, the level of the noise was set using the calibration factor measured in the left ear. The wideband, lowpass, bandpass, and highpass noises differed in overall level by less than 1 dB for the ipsilateral and contralateral presentations.

Backus, B. C., and Guinan, J. J., Jr. (**2006**). "Time-course of the human medial olivocochlear reflex," J. Acoust. Soc. Am. **119**, 2889–2904.

Bassim, M. K., Miller, R. L., Buss, E., and Smith, D. W. (**2003**). "Rapid adaptation of the 2f1-f2 DPOAE in humans: Binaural and contralateral stimulation effects," Hear. Res. **182**, 140–152.

Church, G. T., and Cudahy, E. A. (**1984**). "The time course of the acoustic reflex," Ear Hear. **5**, 235–242.

Dallmayr, C. (**1987**). "Stationary and dynamical properties of simultaneous evoked otoacoustic emissions (SEOAE)," Acustica **63**, 243–255.

Dallos, P. (**1973**). *The Auditory Periphery: Biophysics and Physiology* (Academic, New York).

Goodman, S. S., and Keefe, D. H. (**2006**). "Simultaneous measurement of noise-activated middle-ear muscle reflex and stimulus frequency otoacoustic emissions," J. Assoc. Res. Otolaryngol. **7**, 125–139.

Guinan, J. J., Jr. (**2006**). "Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans," Ear Hear. **27**, 589–607.

Guinan, J. J., Jr., Backus, B. C., Lilaonitkul, W., and Aharonson, V. (**2003**). "Medial olivocochlear efferent reflex in humans: Otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs," J. Assoc. Res. Otolaryngol. **4**, 521–540.

Keefe, D. H. (**1998**). "Double-evoked otoacoustic emissions. I. Measurement theory and nonlinear coherence," J. Acoust. Soc. Am. **103**, 3489–3498.

Keefe, D. H., Schairer, K. S., Ellison, J. C., Fitzpatrick, D. F., and Jesteadt, W. (**2009**). "Use of stimulus-frequency otoacoustic emissions to investigate efferent and cochlear contributions to temporal overshoot," J. Acoust. Soc. Am. **125**, 1595–1604.

Kemp, D. T. (**1980**). "Towards a model for the origin of cochlear echoes," Hear. Res. **2**, 533–548.

Kemp, D. T., and Chum, R. A. (**1980**). "Observations on the generator mechanism of stimulus frequency acoustic emissions—Two tone suppression," in *Psychophysical, Physiological, and Behavioral Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft University, Delft, The Netherlands), pp. 34–42.

Kim, D. O., Dorn, P. A., Neely, S. T., and Gorga, M. P. (**2001**). "Adaptation of distortion product otoacoustic emission in humans," J. Assoc. Res. Otolaryngol. **2**, 31–40.

Lilaonitkul, W., and Guinan, J. J., Jr. (**2009a**). "Reflex control of the human inner ear: A half-octave offset in medial efferent feedback that is consistent with an efferent role in the control of masking," J. Neurophysiol. **101**, 1394–1406.

Lilaonitkul, W., and Guinan, J. J., Jr. (**2009b**). "Human medial olivocochlear reflex: Effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths," J. Assoc. Res. Otolaryngol. **10**, 459–470.

Lonsbury-Martin, B. L., Harris, F. P., Stagner, B. B., Hawkins, M. D., and Martin, G. K. (**1990**). "Distortion-product emissions in humans: II. Relations to acoustic immittance and stimulus-frequency and spontaneous otoacoustic emissions in normally hearing subjects," Ann. Otol. Rhinol. Laryngol. Suppl. **147**, 15–29.

Maison, S., Durrant, J., Gallineau, C., Micheyl, C., and Collet, L. (**2001**). "Delay and temporal integration in medial olivocochlear bundle activation in humans," Ear Hear. **22**, 65–74.

McFadden, D., and Pasanen, E. G. (**1994**). "Otoacoustic emissions and quinine sulfate," J. Acoust. Soc. Am. **95**, 3460–3474.

Oxenham, A. J., and Bacon, S. P. (**2004**). "Psychophysical manifestations of compression: Normal-hearing listeners," in *Compression: From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper (Springer, New York), pp. 62–106.

Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (**1991**). "A review of otoacoustic emissions," J. Acoust. Soc. Am. **89**, 2027–2067.

Schairer, K. S., Fitzpatrick, D., and Keefe, D. H. (**2003**). "Input-output functions for stimulus-frequency otoacoustic emissions in normal-hearing adult ears," J. Acoust. Soc. Am. **114**, 944–966.

Schairer, K. S., and Keefe, D. H. (**2005**). "Simultaneous recording of stimulus-frequency and distortion-product otoacoustic emission input-output functions in adult ears," J. Acoust. Soc. Am. **117**, 818–832.

Shairer, K. S., Ellison, J. C., Fitzpatrick, D., and Keefe, D. H. (**2006**). "Use of stimulus-frequency otoacoustic emission latency and level to investigate cochlear mechanics in human ears," J. Acoust. Soc. Am. **120**, 901–914.

Shairer, K. S., Ellison, J. C., Fitzpatrick, D., and Keefe, D. H. (**2007**). "Wideband ipsilateral measurements of middle-ear muscle reflex thresholds in children and adults," J. Acoust. Soc. Am. **121**, 3607–3616.

Shannon, R. V. (**1976**). "Two-tone unmasking and suppression in a forward-masking situation," J. Acoust. Soc. Am. **59**, 1460–1470.

Strickland, E. A. (**2004**). "The temporal effect with notched-noise maskers: Analysis in terms of input-output functions," J. Acoust. Soc. Am. **115**, 2234–2245.

von Klitzing, R., and Kohlrausch, A. (**1994**). "Effect of masker level on overshoot in running- and frozen-noise maskers," J. Acoust. Soc. Am. **95**, 2192–2201.

Walsh, K. P. (**2009**). "Psychophysical and physiological measures of dynamic cochlear processing," Master's thesis, The University of Texas at Austin, Austin, TX.

Walsh, K. P., Pasanen, E. G., and McFadden, D. (**2008**). "Overshoot measured psychophysically and physiologically in the same ears," Assoc. Res. Otolaryngol. Abstr. **31**, 927.

Walsh, K. P., Pasanen, E. G., and McFadden, D. (**2009**). "Evidence for dynamic cochlear processing in otoacoustic emissions and behavior (A)," J. Acoust. Soc. Am. **125**, 2720.

Whitehead, M. L. (**1991**). "Slow variations of the amplitude and frequency of spontaneous otoacoustic emissions," Hear. Res. **53**, 269–280.

Wiederhold, M. L., and Kiang, N. Y. S. (**1970**). "Effects of electric stimulation of the crossed olivocochlear bundle on single auditory-nerve fibers in the cat," J. Acoust. Soc. Am. **48**, 950–965.

Wilson, R. H., and Margolis, R. H. (**1999**). "Acoustic-reflex measurements," in *Contemporary Perspectives in Hearing Assessment*, edited by F. E. Musiek and W. F. Rintelmann (Allyn and Bacon, Boston, MA), pp. 131–166.

Zwicker, E., and Schloth, E. (**1984**). "Interrelation of different oto-acoustic emissions," J. Acoust. Soc. Am. **75**, 1148–1154.

Sensitization to masked tones following notched-noise correlates with estimates of cochlear function using distortion product otoacoustic emissions

Xiang Zhou
Department of Biomedical Engineering, The City College of the City University of New York, Steinman Hall, Room T-401, Convent Avenue, 140th Street, New York, New York 10031

Simon Henin, Suzanne E. Thompson, and Glenis R. Long
Speech and Hearing Sciences, Graduate Center, City University of New York, 365 Fifth Avenue, New York, New York 10016

Lucas C. Parra[a)]
Department of Biomedical Engineering, The City College of the City University of New York, Steinman Hall, Room T-401, Convent Avenue, 140th Street, New York, New York 10031

(Received 17 September 2009; revised 23 November 2009; accepted 23 November 2009)

Neuronal gain adaptation has been proposed as the underlying mechanism leading to the perception of phantom sounds such as Zwicker tones and tinnitus. In this gain-adaptation theory, cochlear compression plays a significant role with weaker compression leading to stronger phantom percepts. The specific aim of this study was to find a link between the strength of neuronal gain adaptation and cochlear compression. Compression was assessed using distortion product otoacoustic emissions (DPOAEs). Gain adaptation is hypothesized to manifest itself in the sensitization observed for the detection of masked tones when preceded by notched noise. Perceptual thresholds for pure tones in notched noise were measured at multiple frequencies following various priming signals. The observed sensitization was larger than expected from the combined effect of the various maskers. However, there was no link between sensitization and compression. Instead, across subjects, stronger sensitization correlated with stronger DPOAEs evoked by low-level primaries. In addition, growth of DPOAEs correlated reliably with perceptual thresholds across frequencies within subjects. Together, the data suggest that short-term dynamic adaptation leading to perceptual sensitization is the result of an active process mediated by the outer hair cells, which are thought to modulate the gain of the cochlear amplifier via efferent feedback.
 [DOI: 10.1121/1.3277156]

PACS number(s): 43.64.Jb, 43.66.Dc [BLM] Pages: 970–976

I. BACKGROUND

The subjective sensation of sound in the absence of a real stimulus is referred to as a phantom percept. Tinnitus and the Zwicker tone are phantom percepts induced in some subjects at frequencies of reduced auditory stimulation. We argued previously that these phantom percepts result from central neuronal gain adaptation, which increases sensitivity to a point where background neuronal activity is perceived as a phantom sound (Parra and Pearlmutter, 2007). The corresponding computational model suggested a link between the strength of the perceived phantom sounds and the compression factor of cochlear dynamics. Indeed, this previous study found empirically that Tinnitus subjects, who as a group have been shown to have reduced compression (Janssen *et al.*, 1998), are significantly more likely to perceive the Zwicker tone.

The basilar membrane responds to sound in a nonlinear fashion, providing an intensity-dependent gain to incoming sounds (see Cooper *et al.*, 2008 and Neely and Kim, 2008 for a review). This adaptive amplification is thought to be mediated by the outer hair cells (OHCs), which are in the position to modify basilar membrane mechanics on a cycle-by-cycle basis and is often referred to as the cochlear amplifier (see Cooper *et al.*, 2008). At lower signal levels amplification is stronger than at high signal levels at which point the basilar membrane is believed to become purely passive. This results in a compressive non-linearity, which provides a reduction in the output range of the incoming sound and an increase in dynamic range required for proper transmission to the auditory nerve. Adaptive gains have also been documented for central auditory processing stages. For instance, Dean *et al.* (2005, 2008) showed that rate-response curves of neurons in the inferior colliculus adjust in thresholds (sensitivity) and slope (gain) to the auditory stimulus intensity on a time scale of 100 ms.

The specific aim of this study was to find a link between the strength of this central neuronal gain adaptation and the instantaneous cochlear compression. Cochlear compression was assessed with high frequency resolution using a newly

[a)]Author to whom correspondence should be addressed. Electronic mail: parra@ccny.cuny.edu

developed method for obtaining distortion product otoacoustic emissions (DPOAEs), which extracts the generator component of the DPOAE (Long *et al.*, 2008).

To measure gain adaptation we took an indirect approach. Initial preliminary experiments (not shown here) indicated that following a notched noise, the sensitivity to faint sounds in the notch region should be increased. In the extreme, we hypothesized that this increased sensitivity results in the perception of a tone even in the absence of a sound, a phantom percept known as the Zwicker tone. However, we did not find an increase in sensitivity for pure tones in quiet, but for tones that were simultaneously masked by notched noise—a phenomenon that had been observed previously (Carlyon, 1989; Strickland, 2004). The surprising observation is that the addition of a long-duration masker, or precursor, can improve detection thresholds of a stimulus as the stimulus is delayed from the onset of the masker. This phenomenon is referred to as overshoot (e.g., Zwicker, 1965). Viemeister and Bacon (1982) and Thibodeau (1991) reported a similar effect with the use of an adaptor stimulus on the enhancement of a masker in a forward masking paradigm. When a preceding adaptor stimulus lacked the spectral content of the masker, an increase in masking was observed. In this scenario the enhancement effect was attributed to an increase in gain at the masker frequencies resulting from an adaptation to suppression of the components contained in the adaptor stimulus. Here we hypothesize that this sensitization is the result of neuronal gain adaptation, which increases sensitivity in the missing frequency band of the notch. We therefore use this sensitization as a measure of neuronal gain adaptation. In order to measure this sensitization psychophysically, a paradigm is established here that measures perceptual thresholds of a simultaneously masked pure tone in the presence of a long-duration precursor stimulus.

Two experiments were conducted: Experiment I aimed to establish a link between compression and sensitization, and experiment II aimed to confirm that this sensitization is not just a result of a linear process. The results of these experiments suggest that the observed sensitization is the result of an active process mediated by outer hair cell function, but that cochlear compression by itself does not necessarily affect its strength. The experiments also revealed a reliable correlation of compression with perceptual thresholds across frequency.

II. METHODS

A. Subjects and procedures

Eighteen subjects were recruited for this study (11 for experiment I and 7 for experiment II, see below). Initial audiograms were performed to exclude cases of moderate and severe hearing loss. All but two subjects had hearing thresholds of less than 20 dB hearing level (HL) at all audiometric frequencies. Two subjects each had mild hearing losses of less than 30 dB HL at one single frequency. Subjects were between 20 and 45 years of age and were recruited from the main campus of the City University of New York. All subjects were paid $10 an hour for participating in the experiment. An Institutional Review Board consent form was signed before the experiment. The experiment consisted of a psychoacoustic task as well as DPOAE measurement. The total experiment time per subject was approximately 4 h.

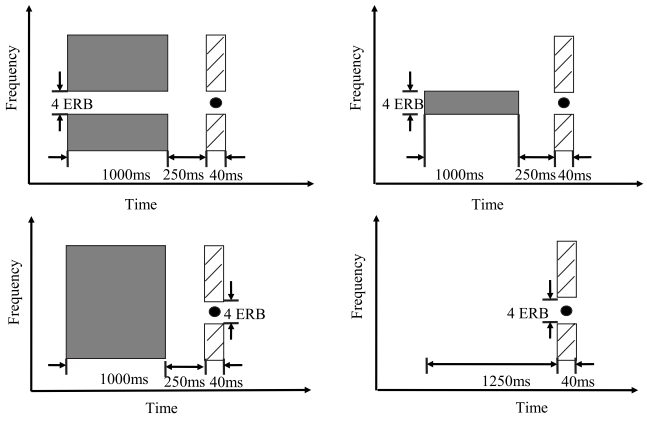


FIG. 1. Schematic of spectral content in psychoacoustic experiment. Thresholds were obtained for a brief probe tone (indicated by the black dot) simultaneously masked by notched noise (light shaded area). The four panels show the different precursors (dark shaded area) clockwise from the top left: notched-noise, bandpass-noise, no precursor/quiet, and white noise.

B. Psychoacoustics: Primed notched-noise masking

Masked-thresholds were measured using a three-interval three-alternative forced-choice (3I3AFC) paradigm with an adaptive threshold-tracking procedure. During the experiment subjects were seated in a Industrial Acoustics Company (IAC) sound-treated booth. For a given trial, the listener pressed one of three keys to indicate the interval in which the probe tone was perceived. A visual aid marked each interval and feedback was provided after each response to indicate if the response was correct or incorrect.

The stimulus contained the following three components, which are represented schematically in Fig. 1: (1) a 1000 ms precursor period which can have one of four precursors: notched noise, bandpass noise, white noise, or no precursor ("quiet"). The notched- and bandpass-noise precursors had a 4-ERB bandwidth and the noise was fixed at 50 dB sound pressure level (SPL) overall [$\mathrm{ERB}(f)=0.108f+24.7$ at center frequency f]; (2) a variable-level probe tone (initial level of 50 dB SPL), and (3) a fixed-level (50 dB SPL overall) simultaneous notched-noise masker (4-ERB notch width). In all conditions the total durations (including onset/offset ramps) of the precursor, masker, and signal were 1000, 40, and 40 ms, respectively. The probe tone and simultaneous notched-noise masker had 5 ms Hanning window onset and offset ramps. The precursor had 50 ms Hanning window ramps. There was a 250 ms time delay between the offset of the precursor and onset of the simultaneous notched-noise masker. All stimuli were generated digitally and played via an M-audio USB sound-card with 24-bit resolution at a sampling rate of 44.1 kHz. These stimuli were routed through a headphone buffer (TDT HB7) before being presented to the listeners via Sony headphones MDR-7506. All signals were filtered to equalize the spectrum of the specific pair of headphones. Equalization filters were obtained by recording a

white noise signal emitted by the headphones with a calibrated microphone (Brüel & Kjær, Nærum Denmark, model 2218) inside a KEMAR head and torso simulator. Filter coefficients were computed from this using linear prediction coefficients of order 20. During each experiment, masked thresholds were measured at signal frequencies ranging from 1 to 4 kHz in steps of 250 Hz. Each trial consisted of three observation intervals. The primed precursor and simultaneous masker were presented in all three intervals, and the probe tone was randomly presented in one of three intervals. The threshold was measured using a modified version of the threshold-tracking procedure known as Parameter Estimation by Sequential Testing (PEST) (Taylor and Creelman, 1967), which estimated the threshold level at the 70% correct point on the psychometric function. In this procedure, the initial tone level was set to 50 dB SPL and decremented with a step size of 8 dB. After the first reversal the step size was reduced to 4 dB and after an additional reversal to 2 dB. The threshold estimate was taken as the mean of the last four reversals with a 2 dB step size. Data collection did not begin until a listener had several practice trials with the experimental paradigm. Presentation of a single frequency condition was randomized across subjects. Each threshold reported here represents the mean over two repetitions of this procedure. In experiment I the range of these two measures (max-min) pooled across frequencies and subjects was 2.7 dB SPL.

In the first experiment (experiment I), 11 subjects were tested in the bandpass- and notched-noise precursor conditions. Two threshold estimates were obtained for each condition and averaged to obtain the final threshold. One subject was excluded from the analysis because a second estimate could not be obtained. An additional seven subjects participated in a second experiment (experiment II) and were tested using all four precursor conditions (bandpass, notched, white, and quiet).

C. Distortion product otoacoustic emissions

DPOAE input-output functions were obtained from all subjects and were used to estimate basilar membrane response. There are several ways to estimate basilar membrane input/output (I/O) functions such as measuring a growth of masking (GOM) function or measuring temporal masking curves (TMCs) (e.g., Rosengard *et al.*, 2005). In this paper, DPOAEs provide an objective measure of cochlear compression. A technical challenge for DPOAEs is to ensure that one measures the response from just one frequency region in the cochlea. DPOAEs are generated in the cochlea in the region where two nearby primary tone stimuli maximally overlap (Shera, 2004). Once the DPOAE is generated, the signal travels both basally toward the oval window and also apically to its own characteristic place on the basilar membrane, where it generates an OAE similar to that generated by the external stimulus. The resulting components have the same frequency but originate from two different regions of the cochlea. To evaluate nonlinear growth, one must extract the component from the generator (maximum overlap) region alone (Mauermann and Kollmeier, 2004).

DPOAEs were obtained from one ear of each subject, who was seated in a recliner in a double-walled IAC sound-treated booth. Custom MAC software (OSX) was used to generate the primaries and to record the ear canal signals. Sweeps were presented via etymotic ER-10A earphones connected to the computer via a MOTU828 firewire interface (24 bit, 44.1 kHz). Ear canal signals were recorded with an Etymotic ER-10A microphone/preamplifier system and amplified by a Stanford SR560 low-noise amplifier connected to and controlled by the same computer. The stimuli used for DPOAE measurement were continuously sweeping primaries with a fixed primary ratio (f_2/f_1) of 1.22, as described by Long *et al.* (2008). Primary frequencies f_1 and f_2 ($f_1<f_2$) were logarithmically swept from an f_2 frequency of 1000–4000 Hz at a rate of 2 s/octave. Primary tone presentation levels were set based on the scissors level paradigm (Kummer *et al.*, 1998) according to the equation $L_1=\max(0.4L_2+39\ \text{dB SPL}, L_2)$. DPOAE levels were measured as a function of input signal level ($L_2=25-75$ dB SPL, 5 dB step).

Several sweeps were obtained for each primary level and averaged to increase the signal-to-noise ratio between the measured DPOAEs and the background noise. The number of sweeps obtained for each level depended on the primary level, with the lowest presentation levels requiring more sweeps ($L_2=25$, $N=60$) than the highest presentation levels ($L_2=75$, $N=12$).

Spectrograms of the individual sweeps were visually inspected, and noisy sweeps were eliminated before averaging at each level. A least-squares fit (LSF) procedure was used to extract the level of the DPOAE generator component for each averaged sound file using overlapping analysis windows of 1/2 s and a step size of 1/80 s (see Long *et al.*, 2008 for a review of the LSF procedure), resulting in an estimate of the magnitude and phase of the generator component of the DPOAEs.

III. RESULTS

The goal of the first experiment was to establish a link between cochlear compression and threshold sensitization. DPOAEs were obtained at various primary levels (25–75 dB SPL in steps of 5 dB) and perceptual thresholds for pure tones masked by notched-noise preceded by one of two precursors (bandpass or notched noise, see Fig. 1). Figure 2 shows the resulting data for one of the ten subjects that participated in this experiment. The top panel of Fig. 2 shows the DPOAE level in the range from 1 to 4 kHz. The bottom panel shows the perceptual thresholds obtained at 12 frequencies within that same range.

One measure of compression can be obtained by measuring the slope of DPOAE level growth as a function of the L_2 level (input/output slope). Low slope values correspond to a more compressive growth function and hence stronger cochlear amplification. In this instance, a compression factor was determined for each frequency as the difference in DPOAE levels between the highest and lowest input L_2 levels (75 and 25 dB SPL)—essentially the spread of the DPOAE curves—divided by 50 dB to obtain an input-output slope. Generally, for various subjects, these compression fac-

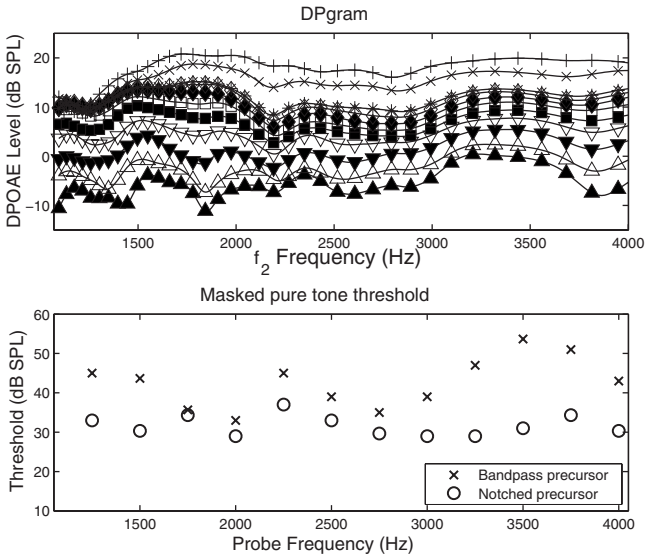


FIG. 2. Example of measured DPOAE and perceptual thresholds for one subject. Top: Each curve shows DPOAEs for one of 11 primary levels (25–75 dB SPL). A larger spread between levels indicates weaker compression. Bottom: Perceptual thresholds for tones under simultaneous notched-noise masking measured with two different precursor conditions (notched and bandpass noises). Distance between the data from two conditions indicates perceptual sensitization (L_2 primary levels: ▲: 25, △: 30, ▼: 35, ▽: 40, ■: 45, □: 50, ◆: 55, ◇: 60, *: 65, x: 70, and +: 75).

tors were within 0.25 and 0.5, indicating normal hearing (see Fig. 4). Compression factors above 0.5 coincided with mild hearing loss in two subjects (e.g., subjects 1 and 9). These values are in agreement with previous literature (Williams and Bacon, 2005). In addition to slope, we evaluated the mean DPOAE level as the average across all input L_2 levels (mean over all curves in Fig. 2, top).

Estimation of sensitization is based on the difference between perceptual thresholds with the notched- and bandpass-noise precursors. Preliminary unpublished experiments with no precursor established that thresholds are reduced by 5 ± 2 dB ($p < 0.0001$, $N=4$) when the masked tone is preceded by a 3 s notched noise. For the present experiment, with a 1 s precursor duration, thresholds are reduced by 5.5 ± 1.1 dB ($p < 0.0001$, $N=10$) when compared to the bandpass-noise condition. (The relationship between the no-precursor and bandpass-noise precursor conditions will be analyzed below.) These results are consistent with previously reported results using similar stimuli and precursor durations (Carlyon, 1989; Strickland, 2004).

Various psychoacoustic and DPOAE data were compared across frequency. Figure 3 shows these comparisons for subject 2. The heading of each panel gives the correlation coefficient (and the corresponding p-value) across frequency for each pair of measures. We found that elevated thresholds in the notched-noise condition were significantly correlated with reduced compression for eight of ten subjects (average correlation coefficient $c=0.67$ for subjects with $p < 0.05$; Fig. 4). As a group, this correlation is highly significant, as shown in Fig. 5, top panel. Additionally, as a group, a significant correlation between compression and masked thresholds following the bandpass-noise precursor was found ($p < 0.01$; Fig. 5, middle panel). However, the correlation between sensitization and compression was not significant. In-

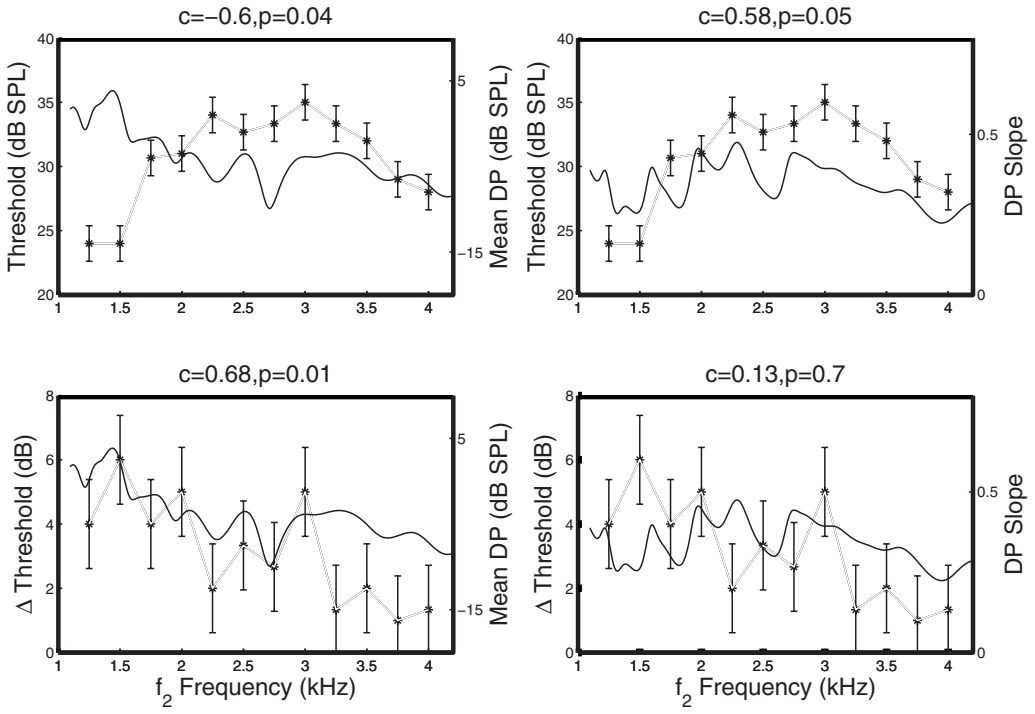


FIG. 3. Within-subject comparison of DPOAEs and psychoacoustic threshold measures. Top left panel plots correlation between masked threshold and mean DPOAE level (averaged across all L_2 levels) as a function of frequency. The top right panel shows the correlation between the threshold and compression (DPOAE I/O slope) as a function of frequency. The lower left panel shows the correlation between △Threshold (the difference between notched- and bandpass-noise thresholds) and the mean DPOAE level. The lower right panel shows the correlation between △Threshold and DPOAE compression (symbols represent psychoacoustic results, solid curve represents DPOAE measures, and error bars indicate range of two repeated measures pooled over frequencies).

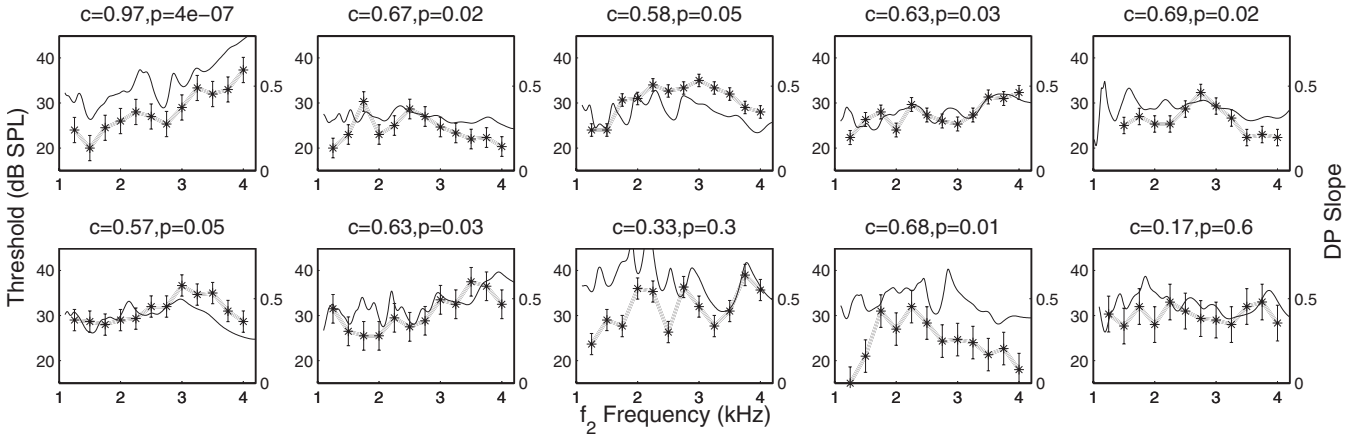


FIG. 4. Within-subject correlation of compression and masked thresholds. Panels show masked thresholds for each subject with notched precursor (symbols) and estimates of DP compression (solid curve). Error bars indicate range of two repeated measures pooled over frequencies. Significant correlations across frequency ($p<0.05$) were observed for eight of the ten subjects. For these subjects, $c=0.67\pm0.04$.

stead, stronger sensitization correlated with stronger DPOAEs for the lowest primary level measured ($c=0.24$, $p=0.008$, $N=10$, Fig. 5, bottom panel).

The sensitization of masked thresholds 250 ms after a notched-noise precursor is interpreted here as the result of neuronal gain adaptation (see Sec. IV for more details).

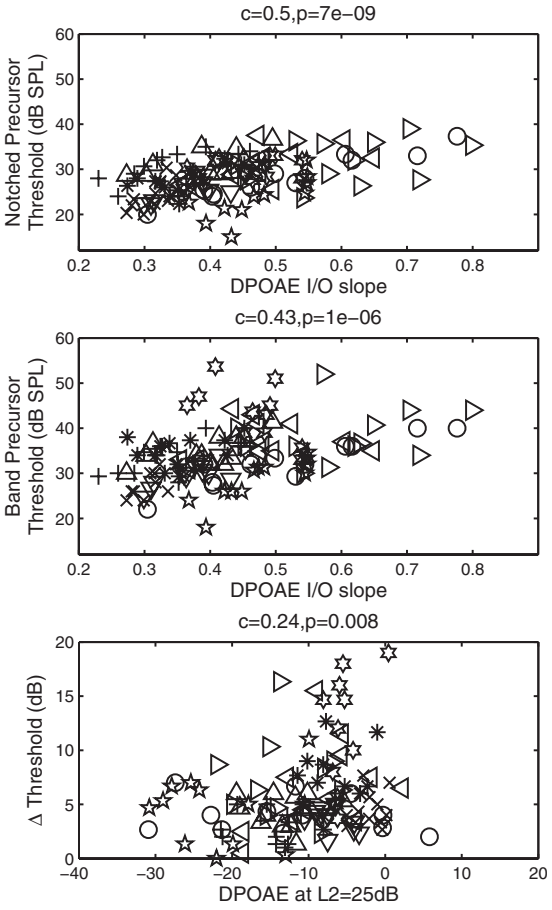


FIG. 5. Comparison of DPOAEs and masked thresholds for all subjects and all frequencies. The top panel shows the correlation between masked thresholds with notched-noise and estimates of compression for all ten subjects. The middle panel shows that thresholds with the bandpass-noise precursor correlate with estimates of compression. The bottom panel shows the correlation for DPOAEs obtained at the lowest primary level and sensitization (bandpass precursor-notched precursor).

To further quantify this sensitization and assess the differential effects of the on- and off-frequency bands in the precursor, an additional two precursor conditions were tested, namely, a quiet and a white noise condition. The top panel of Fig. 6 shows the results obtained in all four precursor conditions. Indeed these data confirm the preliminary observation that the notched-noise precursor significantly improves detection thresholds as compared to the quiet condition [two-way analysis of variance (ANOVA) with frequency and condition as factor, $p=0.004$, $N=7$, df=1, and $F=21.0$]. On average the improvement was 3.2 dB.

Can this improvement be explained as the effect of the off-frequency bands alone or does the contrast in the notched-noise precursor matter? The white noise condition has the same off-frequency power as the notched-noise condition, but also contains power in the on-frequency band. Therefore, its effect was compared to the combined effect of the notched- and bandpass-noise precursors by examining whether the threshold under the notched- and bandpass-noise conditions, I_N+I_B, was equivalent to the thresholds obtained in the white-noise and quiet conditions, I_W+I_Q. In other

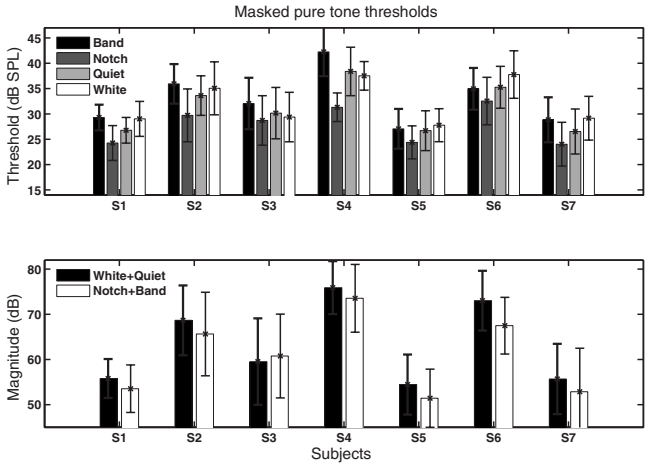


FIG. 6. Masked pure tone thresholds in each precursor condition for seven subjects. The top panel shows individual subjects thresholds in each condition averaged across frequencies. The bottom panel shows the combined thresholds I_N+I_B and I_W+I_Q for each subject (error bars represent standard deviation across frequency).

words, whether the relation $I_N+I_B=I_W+I_Q$ holds (I_Q is included to factor in the effect to the simultaneous notched-noise masker). The improvement due to the notch is significantly larger than this combined effect by 2.5 dB (two-way ANOVA with combined thresholds and frequency as factors, $p=0.015$, $N=7$, df=1, and $F=11.4$). Note that this comparison implies that additivity does not hold under these conditions and that the combined effects of these various precursors point to an underlying non-linear mechanism.

IV. DISCUSSION

The premise of this work was that the perception of a tone is affected by central gain-adaptation mechanisms and that this adaptation would be affected by the amount of cochlear compression: A more compressive cochlea would reduce the intensity difference between a loud and a quiet stimulus, and thus, changes in neuronal gain following changes in signal intensity should be less pronounced. Indeed, neuronal gain adaptation has been demonstrated in the inferior colliculus (Dean *et al.*, 2005) and auditory nerve (Wen *et al.*, 2009).

Here, gain adaptation was assessed by the sensitization observed following a notched-noise precursor, and cochlear compression was assessed using DPOAEs emitted at various primary input levels. Contrary to our expectation, no significant correlation was found between the measure of DPOAE compression and sensitization as measured by the difference between primed narrow-band and primed notched-noise thresholds. However, this sensitization did correlate with the DPOAEs measured with the lowest level primary (see Fig. 5, bottom panel). This suggests that this sensitization is affected or depends on the amplification mechanism of outer hair cells.

The sensitization phenomenon, which can be considered a form of release from masking, is unique in that sensitivity is increased when adding a sound to a probe signal—rather than decreased as is typically the case in masking.

The interpretation of this effect as the result of a neurally mediated adaptive gain is supported by the non-linearity of the combined effects of the notched- and bandpass-noise precursor conditions, as well as the long time scale over which this effect occurs. Note that sensitization is observed here as late as 250 ms after the precursor signal. This is much longer than the instantaneous effect of cochlear compression, longer than the rate adaptation for single auditory nerve fibers, and still longer than forward masking (Meddis and O'Mard, 2005). Thus, we speculate that sensitization is determined more centrally but mediated to some extent by adjusting OHC amplification gains via an efferent feedback loop.

An alternative hypothesis is that the decrease in threshold using a notched-noise precursor may result from a perceptual grouping of the precursor and masker. In this scenario the notched-noise precursor and masker are grouped together in the auditory system, resulting in the increased detectability of the target stimulus (Bregman, 1990). To test this hypothesis, a notched-noise precursor with a different spectrum from that of the masker could be used in order to control such effects and will be investigated in a future experiment.

Overall, the results show that sensitization occurs after presentation of a notched-noise masker and the strength of this sensitization correlated with DPOAEs elicited by the lowest primary level. Additionally, evidence of cochlear compression based on DPOAE I/O functions correlated strongly with both notched- and bandpass-noise psychoacoustic thresholds. In eight of ten subjects these measures correlated significantly across frequency (significance could be established despite a relatively small sample of frequencies).

V. CONCLUSIONS

In normal hearing subjects, notched noise increases sensitivity to tones embedded in noise. This increased sensitization was found to be a non-linear effect, consistent with our hypothesis of a neurally mediated gain-adaptation mechanism. However, perceptual grouping cannot be ruled as a potential mechanism. The correlation between this sensitization and DPOAEs measured with the lowest primary level is interpreted as a link between outer hair cell function and gain adaptation. Cochlear compression correlated strongly with various perceptual thresholds. This establishes DPOAE compression as a potential candidate for the objective evaluation of hearing.

ACKNOWLEDGMENTS

The authors would like to thank Zhenya Yubiler and Naomi Eichorn for their contribution to data collection efforts. This work is supported by the Tinnitus Research Initiative (TRI) and the Graduate Center of the City University of New York.

Bregman, A. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Carlyon, R. (**1989**). "Changes in the masked thresholds of brief tones produced by prior bursts of noise," Hear. Res. **41**, 223–235.

Cooper, N., Pickles, J., and Manley, G. (**2008**). "Traveling waves, second filters, and physiological vulnerability: A short history of the discovery of active processes in hearing," *Springer Handbook of Auditory Research* (Springer, New York).

Dean, I., Harper, N. S., and McAlpine, D. (**2005**). "Neural population coding of sound level adapts to stimulus statistics," Nat. Neurosci. **8**, 1684–1689.

Dean, I., Robinson, B., Harper, N. S., and McAlpine, D. (**2008**). "Rapid neural adaptation to sound level statistics," J. Neurosci. **28**, 6430–6438.

Janssen, T., Kummer, P., and Arnold, W. (**1998**). "Growth behavior of the 2 f1-f2 distortion product otoacoustic emission in tinnitus," J. Acoust. Soc. Am. **103**, 3418–3430.

Kummer, P., Janssen, T., and Arnold, W. (**1998**). "The level and growth behavior of the 2 f1 f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss," J. Acoust. Soc. Am. **103**, 3431–3444.

Long, G. R., Talmadge, C. L., and Lee, J. (**2008**). "Measuring distortion product otoacoustic emissions using continuously sweeping primaries," J. Acoust. Soc. Am. **124**, 1613–1626.

Mauermann, M., and Kollmeier, B. (**2004**). "Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source," J. Acoust. Soc. Am. **116**, 2199–2212.

Meddis, R., and O'Mard, L. (**2005**). "A computer model of the auditory-nerve response to forward-masking stimuli," J. Acoust. Soc. Am. **117**, 3787.

Neely, S., and Kim, D. (**2008**). "Cochlear models incorporating active processes," *Springer Handbook of Auditory Research* (Springer, New York).
Parra, L., and Pearlmutter, B. (**2007**). "Illusory percepts from auditory adaptation," J. Acoust. Soc. Am. **121**(3), 1632–1641.
Rosengard, P., Oxenham, A., and Braida, L. (**2005**). "Comparing different estimates of cochlear compression in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **117**(5), 3028–3041.
Shera, C. A. (**2004**). "Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions," Ear Hear. **25**, 86–97.
Strickland, E. A. (**2004**). "The temporal effect with notched-noise maskers: Analysis in terms of input-output functions," J. Acoust. Soc. Am. **115**, 2234–2245.
Taylor, M. M., and Creelman, C. D. (**1967**). "Pest: Efficient estimates on probability functions," J. Acoust. Soc. Am. **41**, 782–787.
Thibodeau, L. M. (**1991**). "Performance of hearing-impaired persons on auditory enhancement tasks," J. Acoust. Soc. Am. **89**, 2843–2850.
Viemeister, N., and Bacon, S. P. (**1982**). "Forward masking by enhanced components in harmonic complexes," J. Acoust. Soc. Am. **71**(6), 1502–1507.
Wen, B., Wang, G. I., Dean, I., and Delgutte, B. (**2009**). "Dynamic range adaptation to sound level statistics in the auditory nerve," J. Neurosci. **29**, 13797–13808.
Williams, E. J., and Bacon, S. P. (**2005**). "Compression estimates using behavioral and otoacoustic emission measures," Hear. Res. **201**, 44–54.
Zwicker, E. (**1965**). "Temporal effects in simultaneous masking by white-noise bursts," J. Acoust. Soc. Am. **37**, 653–663.

Effects of simulated spectral holes on speech intelligibility and spatial release from masking under binaural and monaural listening

Soha N. Garadat,[a] Ruth Y. Litovsky,[b] and Gongqiang Yu[c]
Waisman Center, University of Wisconsin, 1500 Highland Avenue, Madison, Wisconsin 53705

Fan-Gang Zeng
University of California-Irvine, 364 Med Surge II, Irvine, California 92697

(Received 8 December 2008; revised 20 September 2009; accepted 22 November 2009)

The possibility that "dead regions" or "spectral holes" can account for some differences in performance between bilateral cochlear implant (CI) users and normal-hearing listeners was explored. Using a 20-band noise-excited vocoder to simulate CI processing, this study examined effects of spectral holes on speech reception thresholds (SRTs) and spatial release from masking (SRM) in difficult listening conditions. Prior to processing, stimuli were convolved through head-related transfer-functions to provide listeners with free-field directional cues. Processed stimuli were presented over headphones under binaural or monaural (right ear) conditions. Using Greenwood's [(1990). J. Acoust. Soc. Am. **87**, 2592–2605] frequency-position function and assuming a cochlear length of 35 mm, spectral holes were created for variable sizes (6 and 10 mm) and locations (base, middle, and apex). Results show that middle-frequency spectral holes were the most disruptive to SRTs, whereas high-frequency spectral holes were the most disruptive to SRM. Spectral holes generally reduced binaural advantages in difficult listening conditions. These results suggest the importance of measuring dead regions in CI users. It is possible that customized programming for bilateral CI processors based on knowledge about dead regions can enhance performance in adverse listening situations. © *2010 Acoustical Society of America.*
[DOI: 10.1121/1.3273897]

PACS number(s): 43.66.Pn, 43.66.Ts, 43.71.Ky, 43.66.Qp [ADP] Pages: 977–989

I. INTRODUCTION

Bilateral cochlear implantation has been provided to an increasing number of deaf patients with the goal that bilateral stimulation would enable optimal performance in adverse listening situations. This approach is mainly driven by the substantial benefit that binaural hearing is known to provide to normal-hearing listeners (NHLs) as compared with monaural hearing, in particular, in complex listening situations (e.g., Bronkhorst and Plomp, 1988; Blauert, 1997; Hawley *et al.*, 1999, 2004). Specifically, when listening under binaural conditions, NHLs experience significant improvement in speech recognition when an interfering sound is spatially separated from target speech. This improvement in performance is known as the advantage of spatial separation or *spatial release from masking* (SRM) (Hawley *et al.*, 1999, 2004; Drullman and Bronkhorst, 2000; Freyman *et al.*, 2001; Litovsky, 2005).

In recent years, some benefits of bilateral cochlear implants (CIs) have been demonstrated when performance is compared to that with a single CI. These benefits include enhancement of speech recognition in a multi-source environment, whether target and interfering sources are co-located or spatially separated (Muller *et al.*, 2002; Schon *et al.*, 2002; Tyler *et al.*, 2002; Vermeire *et al.*, 2003; Schleich *et al.*, 2004; Litovsky *et al.*, 2006a, 2009). In addition, improved sound localization abilities have been reported in both adults (e.g., Tyler *et al.*, 2002; van Hoesel and Tyler, 2003; Litovsky *et al.*, 2009) and children (Litovsky *et al.*, 2006b; Litovsky *et al.*, 2006c). Nonetheless, the measured benefits from bilateral implantation vary amongst individuals (e.g., Muller *et al.*, 2002; Wilson *et al.*, 2003; Litovsky *et al.*, 2004, 2006a, 2009; van Hoesel, 2004), with a large range of outcomes and effect sizes. Furthermore, few patients reach the level of performance seen in NHLs. These findings suggest that there exist limitations in the ability of bilateral CIs to provide substantial benefits to all users, but the sources of these limitations remain to be identified. In the context of the current work, two of these limitations are addressed: reduced spectral resolution and loss of spectral information, with the latter being the primary focus.

Reduced spectral resolution or inability to utilize the full spectral information provided by the actual number of electrodes that are stimulated along the implant array has been noted in a number of studies (Friesen *et al.*, 2001; Fu *et al.*, 2004). The extent to which this problem can be ameliorated by providing two implants vs a single implant is not well understood and remains an open question. Within each electrode array, the utility of spectral information is hampered by channel interaction (Throckmorton and Collins, 2002; Fu

[a]Present address: Kresge Hearing Research Institute, University of Michigan-Ann Arbor, MI.
[b]Author to whom correspondence should be addressed. Electronic mail: litovsky@waisman.wisc.edu
[c]Present address: University of Connecticut, Farmington, CT.

and Nogaki, 2005) as well as alteration in the spectral-tonotopic mapping (Fu and Shannon, 1999) which are mediated monaurally at the peripheral level. In the context of bilateral implantation, one can also argue that across-ear differences in monaural processes might partially reverse some of these effects in a summation-like manner due to the use of two independent CIs. However, it is likely that the limitations in utilizing the available spectral information would persist to some extent, even when a patient has a CI in each ear. This is because the extent to which bilateral implants result in an overall increase in the number of functionally independent channels is not clear.

Another source of limitation is related to the local loss of hair cells or auditory neurons. This is known as "dead regions" (Nadol, 1997; Kawano *et al.*, 1998; Moore *et al.*, 2000) and is highly relevant to cochlear implantation. Modern CI speech processors use multiple-electrode arrays specifically designed to take advantage of the natural tonotopic organization of the cochlea and its innervation patterns. The presence of local anomaly along the tonotopic axis is likely to disrupt these innervation patterns (Liberman and Kiang, 1978; Hartmann *et al.*, 1984; Shepherd and Javel, 1997); hence, CI listeners are forced to integrate information from frequency bands that are non-contiguous, i.e., "disjoint bands." The perceptual consequences of this effect can be examined by creating "holes" in the spectrum (Shannon *et al.*, 2002).

An intriguing question that has been the focus of studies done with non-CI users is how do listeners integrate information from disjoint bands, and what is the contribution of each band to overall speech intelligibility. A hallmark of this work has been the development of the articulation index (AI), which represents an attempt to predict speech intelligibility from individual spectral components of the speech segment (Fletcher and Galt, 1950). In general, the AI is based on the assumption that each frequency band contributes to the total AI independent of all other frequency bands and that the contribution of these individual bands is additive. However, subsequent studies demonstrated a synergistic effect when listeners integrated information from widely disjoint bands (Kryter, 1962; Grant and Braida, 1991; Warren *et al.*, 1995). The results of these studies suggest that the AI theory does not take into account the fact that listeners combine information from disjoint bands, thus failing to predict intelligibility of pass-band speech.

A similar line of investigation is needed to understand how CI listeners integrate auditory information from disjoint bands, given that they have been shown to weigh frequency information in a different manner than NHLs (Mehr *et al.*, 2001). While CI users can tolerate a relatively large loss of spectral information (e.g., Shannon *et al.*, 2002; Başkent and Shannon, 2006), studies to date were conducted in quiet, which may not predict CI users' performance in realistic, complex auditory environments. In addition, given the growing number of individuals who undergo bilateral implantation, it is important to further understand the extent to which benefits that are provided through bilateral CIs, such as better speech understanding when target and interfering sounds are spatially separated, are affected by loss of spectral information.

The current study was aimed at investigating the effect of spectral holes on speech intelligibility and on benefits of spatial separation of target and interferes. The paradigm used here builds on the concept of the "cocktail party" environment, in which multiple sources are presented from a number of predetermined locations. First, by including conditions in quiet and with interferers we can compare effects of holes with background noise and in previously-studied quiet conditions. Second, by simulating spatial locations using virtual-space stimuli over headphones, we can directly examine performance under monaural vs binaural conditions and assess binaural benefit. The approach used here is similar to that used previously in absence of spectral holes or degraded speech (Hawley *et al.*, 2004; Culling *et al.*, 2004). By using a noise-excited CI vocoder in which spectrum is degraded already, and then imposing spectral holes, we tested the hypothesis that spectral holes would be detrimental to sentence recognition due to the fragile nature of spectrally sparse information in the vocoder. Further, we tested the hypothesis that the extent of degradation in speech understanding would be dependent on the location of the spectral hole along the simulated cochlear array. Thus, the approach used here would enable identification of regions along the cochlea that, if not intact, place the listener in a position of being particularly susceptible to disruption. Finally, it was hypothesized that a loss of spectral information caused by spectral holes would interact significantly with listening mode (binaural vs monaural) such that binaural hearing may serve to reduce the deleterious effects of spectral hole due to binaural redundancy and the availability of spatial cues.

While the effect of spectral holes should ultimately be tested directly in CI users, currently the identification of dead regions in these listeners remains challenging (Bierer, 2007). As a first step, vocoder simulation of CI listening was used to identify some of the important variables that could potentially impact performance under idealized conditions. Another clinical issue to note is that, because in clinical situations tonotopic information matching across the ears is difficult to achieve and to verify, the precise matching in the present study is idealized relative to true CI bilateral frequency mapping. In addition, in the binaural conditions, spectral holes were deliberately created at matched locations in the two ears, which again, may be unrealistic, as spectral holes present in CI patients due to the existence of inactive or atrophied auditory nerve fibers are unlikely to occur at identical places in the right and left cochleae. This general approach was chosen because of the growing concern regarding the need to maximize salience of binaural cues in CI patients by matching the stimulated cochleotopic regions in the right and left ears (Long, 2000; Long *et al.*, 2003; Wilson *et al.*, 2003). Matching information across the two ears has been a hallmark of studies on binaural sensitivity, in which electrically pulsed signals are delivered to a single pair of binaurally matched electrodes (van Hoesel and Clark, 1997;

Litovsky *et al.*, 2010; Long *et al.*, 2003; van Hoesel, 2004; van Hoesel *et al.*, 2009). However, similar work with speech stimuli remains to be done to understand the potential effectiveness of matching in the clinical realm.

II. METHODS

A. Listeners

Twenty adults with normal hearing (4 males and 16 females; 19–30 yrs old) participated in this study. Participants were native speakers of American English who were recruited from the student population at the University of Wisconsin-Madison. Normal-hearing sensitivity was verified by pure tone, air conduction thresholds of 15 dB hearing level (HL) or better at octave frequencies ranging from 250 and 8000 Hz (ANSI, 2004). In addition, no asymmetry in hearing thresholds exceeded 10 dB HL at any of the frequencies tested. All participants signed approved Human Subjects (University of Wisconsin Health Sciences IRB) consent forms and were paid for their participation.

B. Setup

Testing was conducted in a standard double-walled Industrial Acoustics Company (IAC) sound booth. Participants sat at a small desk facing a computer monitor with a mouse and keyboard for response input. Target and interfering stimuli were first convolved through head-related transfer functions (HRTFs)[1] (described in Sec. II E), digitally mixed and then processed through CI simulation filters to create vocoded speech (see Sec. II F). A Tucker-Davis Technologies (TDT) system 3 RP2.1 real-time processor was used to attenuate the stimuli before sending them to the TDT headphone buffer (HB7). Stimuli were then delivered to listeners through headphones (Sennheiser HD 580). Stimuli were calibrated using a Larsen-Davis "system 824" digital precision sound level meter and a 6-cc coupler (AEC101 IEC 318 artificial ear coupler manufactured by Larson Davis, Depew, New York). Listening conditions were either binaural or monaural (right ear only).

C. Stimuli and procedure

Open set sentences from the MIT recordings of the Harvard IEEE corpus (Rothauser *et al.*, 1969) were used. The complete corpus consists of 72 phonetically balanced lists with ten sentences each, and each sentence has five key words that require identification. These sentences are grammatically and semantically appropriate; however, they have relatively low predictability. Examples of these sentences include "The wide road shimmered in the hot sun," "The small pop gnawed a hole in the sock," or "The glow deepened in the eyes of the sweet girl." This corpus was chosen based on the large number of sentences available. Because they have previously been used in a number of studies with NHLs (e.g., Hawley *et al.*, 1999, 2004) and with CI users (e.g., Stickney *et al.*, 2004; Loizou *et al.*, 2009) under difficult listening conditions, they are valid for the type of tasks used in the present study. As was done in previous studies, short sentences from the corpus were used for the target stimuli, while 15 of the longest sentences were reserved for the interferer sentences and used throughout the experiment in random order. In addition, on each trial interferer sentences had earlier onset (~1 s) than target sentences. Target sentences were recorded with two different male voices; each male voice contributed to half of the lists. The testing order of the target lists was fixed across all participants; however, the order of the sentences within each list was randomized. Interferer stimuli consisted of two-female talkers, each uttering different sentences from the interferer corpus, in random order from trial to trial.

Listeners were instructed to listen to the sentences spoken by a male voice and to ignore those sentences spoken by the females; they were encouraged to guess when not sure and to type everything that they judged to be spoken by the male voice. Spelling and typos were checked online by an examiner seated in the observation room. Other than on the first sentence in each list, once the listeners completed typing the content of the target sentence they pressed the return key and the correct target sentence was displayed on the computer monitor. The first sentence was treated in a special manner, as it was repeated at incrementally higher levels until three key words were identified; only after this first correct response was the sentence revealed to the listener. All other remaining sentences were displayed following the listener's response regardless of number of correct words, as they were only used once. The key words in each sentence were identified with uppercase letters in the transcript. These responses were self-marked by listeners and were verified online by an examiner. Listeners were instructed to compare the two transcripts and to count the number of correct key words, then to enter that number and press the return key to progress to the next trial. Each speech recognition threshold (SRT) run was logged in a data file that could be used for verifying data and scoring reliability by the examiner if needed. This method has been used in previous studies (Hawley *et al.*, 2004; Culling *et al.*, 2004).

D. Estimation of speech reception threshold

The intensity level of the interferers was fixed at 50 dB sound pressure level (SPL) throughout the experiment. SRTs denoting target intelligibility were measured using a similar approach to that described by Plomp (1986) and used by us previously (Hawley *et al.*, 2004; Culling *et al.*, 2004; Loizou *et al.*, 2009). For each SRT measurement, one list of ten sentences was used. At the beginning of each list, the target sentence was presented at a level that yielded poor performance, i.e., a disadvantageous signal-to-noise ratio (SNR). If the target was judged to be inaudible by the listener, the instruction was to press the return key and the same target sentence was repeated at a SNR that was more favorable by 4 dB; this was repeated until 3 or more key words (out of 5) were correctly identified. Following, the remaining nine sentences were presented at varying SNRs using a one-down/one-up adaptive tracking algorithm targeting 50% correct (Levitt, 1971). The rule for varying the levels was as follows: level decreased by 2 dB if three or more key words identified correctly, else the level was increased by 2 dB. A single SRT

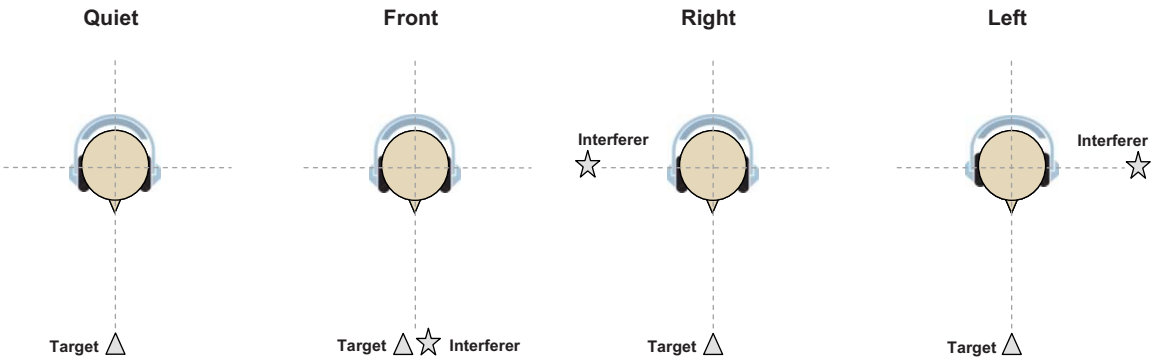


FIG. 1. (Color online) Schematic representation of the virtual spatial conditions used in the study. Quiet, (target at 0° and no interferers), front (target and interferers at 0°), right (target at 0° and interferers at 90°), and left (target at 0° and interferers at −90°). These setups were used in the binaural and monaural conditions.

value was determined by averaging the level presented on the last eight trials (trials 4–11). Given that each list allowed only ten trials to be measured, the level at which the 11th trial would have been presented was estimated based on the result of the tenth trial (Hawley *et al.*, 1999; 2004; Culling *et al.*, 2004).

E. Virtual spatial configurations

The unprocessed stimuli were convolved with non-individualized HRTFs (Gardner and Martin, 1994) to provide listeners with spatial cues regarding both target and interferers. Stimuli for each intended virtual spatial configuration were convolved through HRTFs for the right and left ears. Target and interfering stimuli were then digitally mixed and subsequently passed through the CI simulation filters described below. Measurements were obtained under headphones for each listener using the following virtual spatial configurations: (1) Quiet: target at 0° azimuth and no interferers, (2) front: target and interferers both at 0° azimuth, (3) right: target at 0° azimuth and interferers at 90° azimuth, and (4) left: target at 0° azimuth and interferers at −90° azimuth (see Fig. 1). SRM, which provides a measure of the improvement in SRT that occurs when the target and interfering stimuli are spatially separated, was calculated from SRTs in conditions (2)–(4) such that SRMright equals $[SRT_{front} - SRT_{right}]$ and SRMleft is equal to $[SRT_{front} - SRT_{left}]$.

F. CI signal processing

Stimuli were processed through MATLAB software simulations using the signal processing strategies described by Shannon *et al.* (1995). Speech signals covering the frequency range of 300–10 300 Hz were divided into 20 contiguous frequency bands using a sixth order elliptical infinite impulse response (IIR) filter. The frequency cutoffs in these bands were calculated using the Greenwood map (Greenwood, 1990). The envelope from each band was extracted by full-wave rectification, followed by low-pass filtering using second order Bessel IIR filters at a 50 Hz cutoff frequency. The envelope from each band was then used to modulate a broadband white noise carrier, which was then subject to the same filter as in the analysis filters to remove spectral splatter. Finally, to create vocoded speech, the modulated outputs from each band were summed. For each given trial, target and interferers were processed in the same manner.

In the spectral hole conditions, holes were created by simply setting the output of the corresponding analysis bands to zero. These omitted bands were manipulated using the following frequency ranges which correspond to three tonotopic locations: 150–1139 Hz (basal), 749–3616 Hz (middle), and 2487–10 800 Hz (apical). Using the above frequency ranges, two hole sizes (6 and 10 mm) were created for each tonotopic location by varying the range of the dropped central frequencies (see Table I for specifics). The locations of spectral holes were determined using Greenwood's frequency-position function which assumes a 35 mm cochlear length (Greenwood, 1990). The overall level of stimuli with holes was normalized to the level of the condition without a hole; thus the hole size had no effect on level.

G. Experimental conditions and design

Participants were divided into two equal-*N* size groups and randomly assigned to one of two listening mode groups (binaural or monaural). Each listener completed the study in five, 2 hour sessions. SRTs were measured for seven conditions (baseline+2 hole sizes × 3 locations). For each of these conditions, testing was conducted at four spatial configurations, as described in Sec. II F. Thus, each listener contributed data from 28 SRTs. Data collection was blocked by processing condition, thus seven testing blocks; these blocks were presented in random order for each subject, and within each block spatial conditions were randomized.

Prior to each testing block, listeners received training using the same processing condition as in that block. Four SRTs (two in quiet and two with speech interferers in the front) were collected; these data were discarded from the

TABLE I. Summary of the frequency range *dropped* to create the different experimental spectral holes. Central frequencies of the corresponding output bands were set to zero to eliminate stimulation in those regions.

Spectral hole size (mm)	*Apex* (Hz)	*Middle* (Hz)	*Base* (Hz)
6	150–600	1139–2487	4350–10 800
10	150–1139	749–3616	2487–10 800

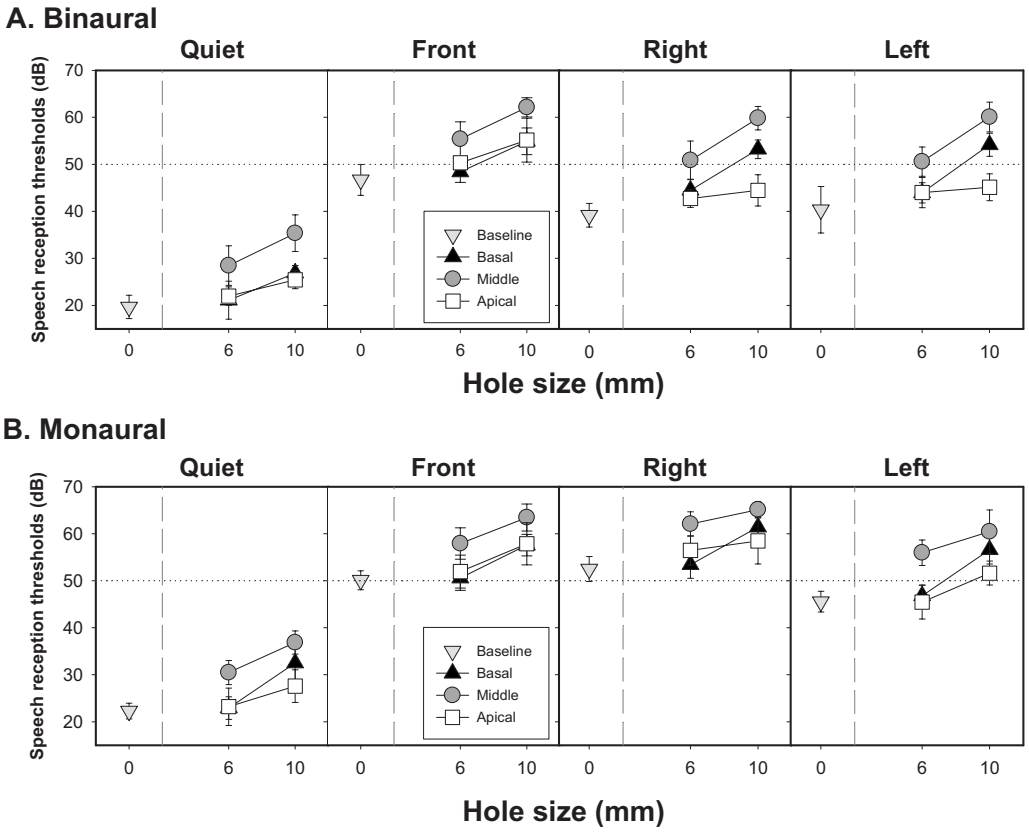


FIG. 2. Average SRTs (±SD) in decibels are plotted as a function of spectral hole size for the baseline and the three different hole placement conditions (basal, middle, and apical). Upper panels represent binaural data obtained for the four spatial configurations and bottom panels represent monaural data. The horizontal dotted line represents 0 dB SNR.

final analysis. This training procedure has been shown to help stabilize subjects' performance when listening to vocoded speech (Garadat *et al.*, 2009). After completing all testing conditions, each subject was tested again on the first testing block to control for possible learning occurring in the first session. This approach increased the total number of SRT conditions tested to 32 per subject, but did not change the number of SRTs per subject used in the data analysis.

III. RESULTS

Results were analyzed separately for the two groups of listeners as well as for across-group effects of listening modes. Within each group, SRT and SRM values were subjected to two-way repeated measure analyses of variance (ANOVAs) with *spectral hole placement* (baseline, basal, middle, and apical) and *spatial configuration* (quiet, front, right, and left) as the within-subject variables. The analyses were conducted separately for the 6 and 10 mm hole data, and within each analysis the same baseline data (where no holes were present) were used. An α criterion of 0.05 was used to determine statistical significance in the omnibus F tests. Across-group comparisons were conducted as mixed-nested ANOVAs (detailed below). *Post hoc* *t*-test α values were corrected for multiple comparisons using the Holm–Bonferroni procedure.

A. Effect of spectral holes on SRTs and SRM in binaural listening

Figure 2 (top panel) shows SRT values obtained in the 6 and 10 mm binaural conditions. A significant main effect of hole placement was found for both 6 mm [$F(3,9)=13.6$, $p<0.001$] and 10 mm [$F(3,9)=51.35$, $p<0.0001$] hole conditions. For both hole sizes, SRTs increased significantly (i.e., worse performance) relative to the baseline for all hole locations ($p<0.0001$). Middle holes, regardless of hole size, resulted in higher SRTs than basal ($p<0.0001$) and apical ($p<0.0001$) holes. While 6 mm apical and basal holes produced comparable SRTs, in the 10 mm condition SRTs were higher with basal than with apical holes ($p<0.0001$).

A significant main effect of spatial configuration was found for both the 6 mm [$F(3,9)=240.7$, $p<0.0001$] and 10 mm [$F(3,9)=378.4$, $p<0.0001$] conditions. Relative to the quiet condition, SRTs increased in the presence of speech interferers, as expected, regardless of the location of the interferers ($p<0.0001$). In addition, front SRTs were higher than both left and right ($p<0.0001$) SRTs, with no differences between left and right. These results were similar for both the 6 and 10 mm conditions. However, in the 10 mm condition, a significant interaction was found for hole placement and spatial configuration [$F(9,81)=7.1$, $p<0.0001$]. SRTs were comparable for the basal and apical holes in the quiet condition and for the front-interferer condition, i.e., when there was no masking or when the interferers and tar-

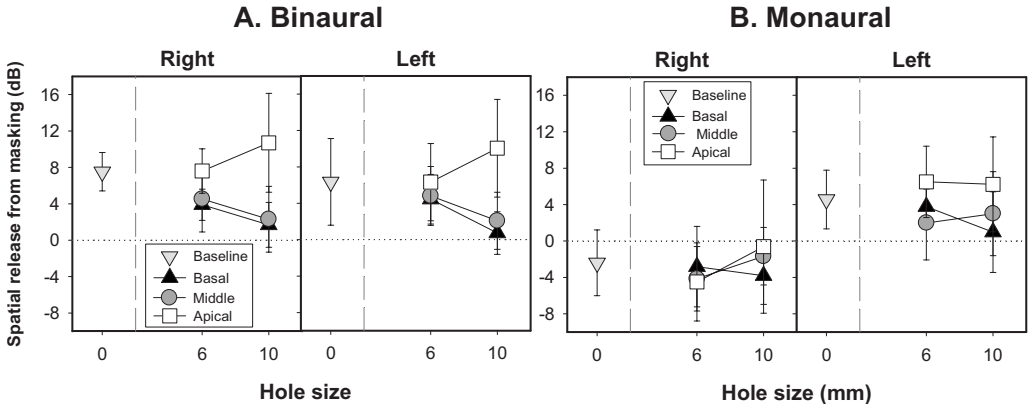


FIG. 3. Average SRM (±SD) in decibels are plotted as a function of spectral hole size for the baseline and the three different hole placement conditions (basal, middle, and apical). Panel A represents binaural data and panel B represents monaural data. The horizontal dotted line represents 0 dB SRM.

get arrived from the same location. However, SRTs were higher for basal than apical holes in the right and left interferer conditions ($p<0.0001$), that is, when the interferers and target were spatially separated. This interaction suggests that the lost high-frequency information due to the presence of 10 mm holes in the base might be necessary in spatial segregation in order to facilitate listening in complex listening environments.

SRM values calculated from the SRTs under binaural conditions are summarized in Fig. 3 (panel A). A main effect of hole placement was not found for 6 mm conditions, suggesting that no differences in amount of SRM are introduced by the presence of 6 mm spectral holes. However, a significant main effect of 10 mm hole placement on SRM was found [$F(3,9)=4.9$, $p<0.05$]. Relative to the no-hole condition, SRM was comparable to the apical condition; however, both conditions produced SRM that was larger than SRM obtained in the basal ($p<0.0001$) and middle ($p<0.005$) hole conditions. No differences in SRM were found between basal and middle holes. There was no main effect of spatial configuration, suggesting that in the binaural conditions, the amount of SRM was comparable for the left and right spatial conditions.

B. Effect of spectral holes on SRTs and SRM in monaural listening

Figure 2 (bottom panels) shows SRT values obtained in the 6 and 10 mm monaural conditions. A significant effect of hole placement on SRTs was found for the 6 mm [$F(3,9)=19.55$, $p<0.0005$] and 10 mm [$F(3,9)=42.1$, $p<0.0005$] conditions. Unlike the binaural condition, 6 mm holes in the apical and basal regions had no effect on SRTs as determined by the results of the *post-hoc* analysis. However, Fig. 2 shows a modest trend for higher SRTs obtained in these two spectral hole conditions; this is discussed further in Sec. IV C. In addition, similar to the binaural conditions, 6 mm middle holes also resulted in higher SRTs compared with no-hole or with holes in the base or apex ($p<0.0001$). In contrast, in the 10 mm case, SRTs were elevated for all conditions with holes relative to the baseline condition ($p<0.0001$). In addition, in the 10 mm case, middle holes resulted in higher SRTs than basal and apical holes ($p<0.0001$), and basal holes led to higher SRTs than apical holes ($p<0.005$).

A main effect of spatial configuration was found for both 6 mm [$F(3,9)=381.43$, $p<0.00001$] and 10 mm [$F(3,9)=311.7$, $p<0.0005$] conditions. Similar to the binaural stimulation mode, masking was evident by the fact that, relative to the quiet condition, SRTs increased in the presence of speech interferers regardless of the location of the interferers ($p<0.0001$). Unlike the binaural listening mode, SRTs were higher when speech interferers were on the right compared with front in the 6 mm condition ($p<0.005$); however, SRTs were comparable for front and right conditions when 10 mm holes occurred. Additionally left SRTs were lower than right ($p<0.0001$) and front SRTs ($p<0.005$) for both 6 and 10 mm holes.

From these monaural SRT values, SRM was computed as the difference between SRTs in the left or right condition and SRTs in the front (see Fig. 3, panel B). A main effect of hole placement was not found for either the 6 or 10 mm conditions, suggesting that SRM was similar in the monaural listening mode regardless of the hole size. These results are partially inconsistent with the binaural data which showed significantly reduced SRM in the presence of 10 mm spectral holes and will be revisited in Sec. IV B for further discussion. As expected for the monaural conditions, a significant main effect was found for spatial configuration in both 6 mm [$F(3,9)=108.39$, $p<0.0001$] and 10 mm conditions [$F(3,9)=103.06$, $p<0.00001$]. As can be seen in Fig. 3 (panel B), there was substantially greater release from masking when the interferers were to the left than the right. In addition, a significant two-way interaction for hole placement by spatial configuration was found for 6 mm conditions [$F(3,27)=294.67$, $p<0.05$]. Specifically, when interferers were to the right, reduction in SRM compared with the baseline condition occurred only with apical holes, with no differences across other conditions. Conversely, when interferers were to the left, apical spectral holes produced larger SRM than that obtained in the baseline ($p<0.05$), basal ($p<0.005$), or middle ($p<0.0001$) conditions. In addition, SRM was smaller in the middle than the baseline conditions ($p<0.01$), and SRM was comparable for the middle and basal conditions (see Fig. 3, panel B). Taken together, these

TABLE II. Results of main effect of listening mode.

Condition	*df*	*F*	*P*
	(a) SRTs		
Baseline	1,18	44.960	<0.0001
6 mm basal	1,18	23.541	<0.0001
10 mm basal	1,18	42.078	<0.0001
6 mm middle	1,18	24.161	<0.0001
10 mm middle	1,18	5.686	<0.05
6 mm apical	1,18	25.624	<0.0001
10 mm apical	1,18	55.924	<0.0001
	(b) SRM		
Baseline	1,18	18.015	<0.0001
6 mm basal	1,18	12.593	<0.005
10 mm basal	1,18	4.220	Not sig.
6 mm middle	1,18	16.263	<0.0001
10 mm middle	1,18	1.433	Not sig.
6 mm apical	1,18	17.529	<0.005
10 mm apical	1,18	9.565	<0.01

results suggest that when interferers were located to the left, spatial segregation of target and interfering speech was facilitated compared to the right conditions. However, the effect size is dependent on information in particular spectral regions.

C. Effect of listening mode

This section focuses on differences observed in binaural vs monaural listening modes. SRT and SRM values were each subjected to a mixed-nested ANOVA with spatial configuration as a within-subject variable and listening mode as a between-subject variable. Analyses were conducted separately for the baseline and each of the six-hole conditions for the base, middle, and apex. Due to the large number of analyses, results are primarily reported in tables (Tables II–IV). Figure 4 is intended to demonstrate the effect of listening mode by plotting differences in SRT values between monaural and binaural conditions; a value of zero represents no difference and positive values indicate better performance (i.e., lower SRTs) for the binaural group than the monaural group. The ANOVA, which was conducted on the raw SRT values, yielded a significant main effect of listening mode, suggesting that the binaural group performed significantly better than the monaural group (see Table IIa for statistical results). A significant main effect of spatial configuration was also found (see Table IIIa).

TABLE III. Results of main effect of spatial configuration.

Condition	*df*	*F*	*P*
	(a) Overall binaural and monaural SRTs		
Baseline	3,18	189.9	<0.0001
6 mm basal	3,18	218.496	<0.0001
10 mm basal	3,18	218.446	<0.0001
6 mm middle	3,18	159.137	<0.0001
10 mm middle	3,18	157.39	<0.0001
6 mm apical	3,18	178.97	<0.0001
10 mm apical	3,18	189.9	<0.0001
	(b) Overall binaural and monaural SRM		
Baseline	1,18	13.408	<0.005
6 mm basal	1,18	21.326	<0.0001
10 mm basal	1,18	8.970	<0.01
6 mm middle	1,18	17.189	<0.005
10 mm middle	1,18	4.552	<0.05
6 mm apical	1,18	27.552	<0.0001
10 mm apical	1,18	10.052	<0.01

TABLE IV. Results for two-way interaction of spatial configuration × listening mode.

Condition	*df*	*F*	*P*
	(a) SRTs		
Baseline	3,54	21.787	<0.0001
6 mm basal	3,54	10.664	<0.0001
10 mm basal	3,54	6.852	<0.005
6 mm middle	3,54	12.403	<0.0001
10 mm middle	3,54	3.322	<0.05
6 mm apical	3,54	28.974	<0.0001
10 mm apical	3,54	11.429	<0.0001
	(b) SRM		
Baseline	1,18	26.541	<0.0001
6 mm basal	1,18	15.269	<0.005
10 mm basal	1,18	20.271	<0.0001
6 mm middle	1,18	14.362	<0.005
10 mm middle	1,18	5.519	<0.05
6 mm apical	1,18	43.896	<0.0001
10 mm apical	1,18	14.709	<0.005

A significant interaction between spatial configuration and listening mode is noteworthy (see Table IVa for *post-hoc* independent sample *t*-test results). In the baseline condition, binaural SRTs were lower than monaural SRTs for all spatial configurations ($p<0.05$); all values shown in Fig. 4 with hole size=0 reflect significant group differences. However, when spectral holes were introduced, group differences depended on the spatial configuration. The two groups had comparable performance in the quiet and front SRT conditions where the SNR was identical across the two ears. As can be seen in Fig. 4 an exception to this occurred in the 10 mm basal condition in which the binaural group had lower quiet SRTs than the monaural group ($p<0.0001$). When the interferers were located to the right, the binaural group performed substantially better than the monaural group ($p<0.0001$) across the different spectral hole conditions. However, when the interferers were located to the left, the binaural group performed better in the 6 mm basal ($p<0.05$), 6 mm middle ($p<0.005$), and 10 mm apical ($p<0.0001$) hole conditions.

In Fig. 5 group differences in SRM are shown. The data are plotted such that zero values would reflect no group differences and positive values indicate greater amount of SRM in the binaural group. Relative to the monaural group, the binaural group had greater SRM in all but the basal and middle 10 mm hole conditions (statistics listed in Table IIb). As can also be seen in Fig. 5, right SRM was greater than left SRM (see Table IIIb for statistical results). The group difference is accounted for by the fact that the binaural group had greater right SRM than the monaural group, but the two

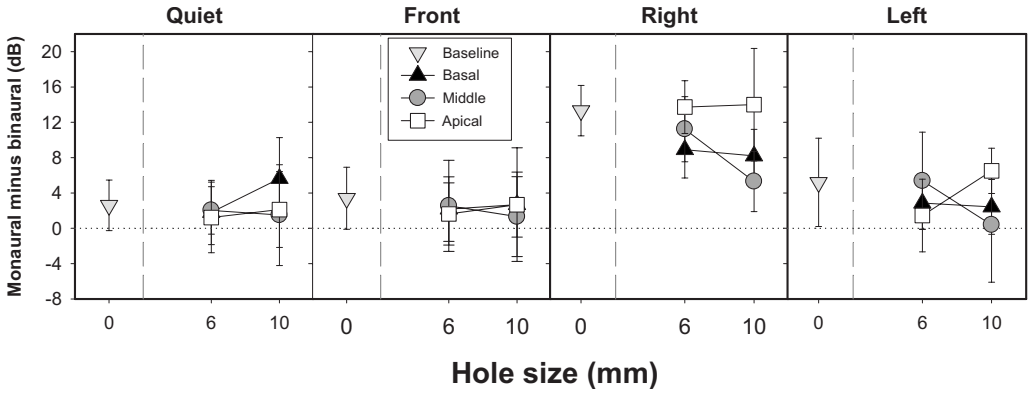


FIG. 4. Average SRTs (±SD) are plotted as the difference in decibels between monaural and binaural SRTs. A positive value indicates improvement in performance due to listening binaurally; consequently, a value near zero indicates minimal differences between monaural and binaural SRTs. Data are plotted as a function of spectral hole size with 0 mm refers to the baseline conditions. These data are plotted for the basal, middle, and apical hole conditions.

groups had comparable left SRM as is evident by values being closer to zero (statistical results are reported in Table IVb). This latter finding is a good example of the robust advantages brought on by the head shadow effect that is naturally available in the monaural conditions, with right ear listening and interferes on the left.

IV. DISCUSSION

In this study, measures of SRT and SRM were used to evaluate the effect of simulated dead regions or spectral holes on sentence recognition. Previous studies addressed this issue in quiet listening conditions, presenting a limited representation of the real world (Kasturi and Loizou, 2002; Shannon *et al.*, 2002). In addition, these issues were addressed under either binaural or monaural listening; hence, role of binaural hearing cannot be determined in these studies. Given that listeners typically operate in less than ideal listening situations, the current study provides an estimate of the effect of spectral holes on performance under conditions that mimic the cocktail party effect (for a review, see Bronkhorst, 2000). By simulating a complex auditory environment using virtual-space stimuli, we were able to look at the interaction between limitations imposed by spectral holes and the contribution of binaural hearing to performance.

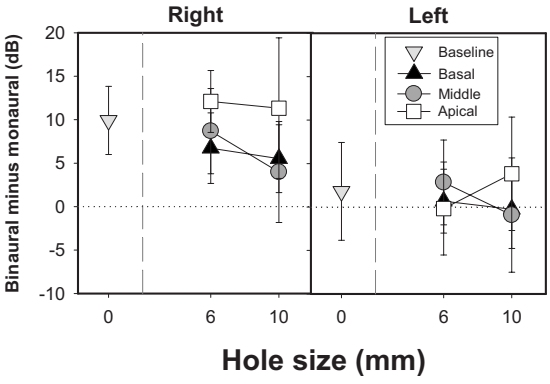


FIG. 5. Average SRM values (±SD) are plotted as the difference in decibels between monaural and binaural SRTs. A positive value indicates greater SRM due to listening binaurally; consequently, a value near zero indicates minimal differences between monaural and binaural SRTs. Data are plotted as a function of spectral hole size with 0 mm refers to the baseline conditions. These data are plotted for the basal, middle, and apical hole conditions.

A. Effect of size and tonotopic location of spectral holes on SRTs

Using a 20-band noise-excited vocoder, current results demonstrated the deleterious effect of spectral holes on SRTs for sentences. As can be seen in Fig. 2, one important finding is the lack of differences in SRTs between 6 mm basal (4350–10 800 Hz) and apical (150–600) holes. However, when the size of the spectral hole was increased to 10 mm, SRTs for the basal hole conditions were higher than SRTs for the apical hole conditions which suggests that basal holes occurred in the frequency range between 2487 and 10 800 Hz are more detrimental to overall performance than apical holes created by omitting the frequency range between 150 and 1139 Hz. These findings suggest the importance of the high-frequency information in the basal region to sentence recognition. Consistent with these data, individuals with high-frequency hearing loss often complain of great difficulty understanding speech particularly in the presence of background noise (Chung and Mack, 1979; Pekkarinen *et al.*, 1990). Current findings are also in agreement with anatomical, electrophysiological, and behavioral data, showing that loss of basal cells in non-human mammals produced a greater deterioration in threshold than loss of apical cells (Schuknecht and Neff, 1952; Stephenson *et al.*, 1984; Smith *et al.*, 1987). One interpretation is that, due to redundancy in mechanisms that encode low-frequency information, a relatively small number of apical hair cells can account for hearing at low-frequencies (Prosen *et al.*, 1990).

More important than basal stimulation, however, seems to be mid-cochlear stimulation. Our results demonstrated a substantial increase in SRTs when spectral holes were created in the middle of the simulated electrode array. This degradation in performance was greater than that produced by basal or apical holes, which suggests the importance of the mid-frequency range for sentence recognition. There is a line of evidence indicating that mid-frequency regions accommodate most of the spectral cues for speech such as the second and the third formants which are important for vowel and consonant identification (Stevens, 1997). In addition, using

filtered speech, frequency bands in the mid-frequency range (1500–3000 Hz) have been shown to produce better speech recognition than those in the lower-frequency range (e.g., French and Steinberg, 1947; Stickney and Assmann, 2001). Direct measures of frequency-importance functions for a variety of speech materials also showed that the frequency band near 2000 Hz contributes the most to speech intelligibility (e.g., Bell *et al.*, 1992; DePaolis *et al.*, 1996). Finally, it appears that while low-frequency information is important for vowel recognition and high-frequency information for consonant recognition, mid-frequency information is important for both vowel and consonant recognitions (Kasturi and Loizou, 2002). Taken together, it seems that the mid-frequency range is the richest information-bearing region of the speech spectrum; hence a significant deterioration in sentence recognition is likely to occur when this frequency range is missing. Overall, current results are consistent with most frequency-importance functions.

However, in contrast to these results, previous work using a band-pass filtering approach in NHLs demonstrated that high level of speech recognition could be achieved even when a large disruption in the mid-frequency range occurs (e.g., Breeuwer and Plomp, 1984; Warren *et al.*, 1995; Lippmann, 1996). Several important issues need to be taken into account when considering this difference. First, in the current study additional reduction in spectral information was introduced by using a vocoder. This finding supports our hypothesis that spectral holes are detrimental to sentence recognition due to the fragile nature of spectrally sparse information in the vocoder and reduced ability to seamlessly combine information from disjointed bands. Future work might delve more deeply into the interactions between number of channels and susceptibility to degradation.

Our results also suggest that the extent of degradation in speech understanding is dependent on the location of the spectral hole along the simulated cochlear array; thus the approach used here would enable identification of regions along the cochlea that, if intact, place the listener in a position of being particularly susceptible to disruption. Nonetheless, one might consider the drawback of using meaningful sentences to measure speech recognition due to increased linguistic redundancy. Meaningful sentences are likely to make available the semantic predictability and the contextual information that listeners could exploit in order to facilitate speech understanding (Miller *et al.*, 1951; Pickett and Pollack, 1963; Rabinowitz *et al.*, 1992). This can be supported by a study by Stickney and Assmann (2001) which replicated a previous investigation showing that speech intelligibility can remain high even when listening through narrow spectral slits (Warren *et al.*, 1995). Using sentences with high and low context predictability, Stickney and Assmann (2001) found that while sentences with high predictability led to high speech recognition rates, performance dropped by 20% when sentences with low predictability were used. It is important to note that the sentence corpus used in this study has relatively low word-context predictability where almost no contextual information is present (Rabinowitz *et al.*, 1992); hence, contextual cues were not a confounding factor in the current study. These results rather support the importance of the middle-frequency information to overall speech recognition.

Findings from the current study are at variance with aspects of previous work using a similar approach (Shannon *et al.*, 2002) in which apical holes were more disruptive to sentence recognition than basal holes. Methodological differences are likely responsible for the variance between studies. Specifically, there are differences between the two studies in the frequency partitioning of bands used to create vocoded speech. Frequency-to-place calculations can be determined by either using the Greenwood equation (this study) or by using one of the 15 frequency allocation tables (Cochlear Corporation 1995) such as Table 9 which uses a linear approach up to 1550 Hz and a logarithmic approach for the remaining bands (Shannon *et al.*, 2002). The range of frequency information that was dropped to create the spectral holes was different between the two studies. Specifically, in the study by Shannon *et al.*, (2002), the frequency range of 350–2031 Hz was eliminated in order to create the largest apical holes. However, this current study eliminated the frequency ranges between 1139–2487 and 749–3616 Hz to create 6 and 10 mm spectral holes, respectively, in the middle region. Therefore, there is an overlap between the two studies in the frequencies used to create the apical and middle spectral holes. Hence, regardless of the discrepancy in the frequency partitioning of the bands, both studies confirm the importance of this frequency range for speech discrimination.

B. Effect of size and tonotopic location of spectral holes on SRM

In adverse listening environments listeners take advantage of the availability of spatial cues to segregate target speech from interfering sounds. SRM is the improvement in speech recognition that occurs when a target source is spatially separated from interfering sounds. In this study we examined the extent to which SRM is affected by speech being spectrally impoverished.

It was hypothesized that despite the presence of spectral holes spatial cues would remain available and provide sufficient cues for SRM. Results from the binaural group (shown in Fig. 3, panel A) indicate that there was no significant impact of the 6 mm holes on SRM. However, as seen in this figure, there seems to be a tendency for a modest decrease in amount of SRM for the 6 mm basal and middle hole conditions which did not reach statistical significance. In contrast, under monaural stimulation, SRM was reduced by the 6 mm spectral holes. However, when 10 mm spectral holes were introduced, SRM was reduced in the binaural conditions, with the extent of hole impact depending on the cochlear place of spectral hole. Specifically, binaural SRM was particularly reduced when 10 mm middle and basal holes were created. In comparison, there was no effect of 10 mm spectral holes on SRM in the monaural conditions. These results leave an important question regarding the binaural/monaural group differences for the 10 mm conditions. However, when considering these data, one should take into account that 6 mm spectral holes resulted in reduced SRM in the monau-

ral conditions. As such it is reasonable to predict that SRM for these monaural conditions would rather decrease with further increase in the spectral hole size. Indeed this decrease in amount of SRM can be noticeable in Fig. 3 for the 10 mm basal conditions (panel B). However, while this decrease in SRM was statistically significant for the binaural condition, it did not reach significance in the monaural conditions. The extent to which inter-subject differences played a role in masking this effect cannot be ruled out. It is noteworthy, however, to mention that these individual differences are larger in extent than the individual differences in NHLs tested under similar conditions using natural speech (Hawley *et al.*, 2004). Part of this variability could be introduced by the uncertainty of listening to spectrally limited stimuli.

Interestingly, the finding of reduced SRM in the presence of spectral holes is consistent with previous reports, demonstrating reduced SRM in hearing impaired individuals (Duquesnoy, 1983; Gelfand *et al.*, 1988). It has been suggested that reduced SRM in hearing impaired individuals is related to their inability to take advantage of interaural level differences due to reduced audibility (Bronkhorst and Plomp, 1989, 1992). Findings from the current study do not support that idea as increasing the level of the target signal did not preserve the spatial cues, rather these results are more consistent with the idea that reduced SRM in hearing impaired individuals is related to reduced or absent spectral cues underlying head shadow effects (Dubno *et al.*, 2002). Head shadow effect is evident starting at approximately 1500 Hz and largest between 2000 and 5000 Hz (Nordlund, 1962; Tonning, 1971; Festen and Plomp, 1986). In this study, to create basal and middle spectral holes, the above frequency range was systematically manipulated depending on the exact location and the extent of the spectral holes. Specifically, basal spectral holes were created in the frequency range of 2487–10 800 Hz, and middle spectral holes were created in the frequency range of 799–3616 Hz. A large reduction in SRM was observed when holes occurred in these conditions. On the other hand, when the spectral holes occurred between 150 and 1139 Hz in the apical conditions, no change in SRM was observed. Therefore, preserving spectral information underlying head shadow cues might aid sound segregation which is essential to overcome the problem of listening in noise.

C. Contribution of binaural hearing

This study was partly motivated by the findings that there are clear and documented limitations to advantages offered by bilateral CIs (e.g., Wilson *et al.*, 2003; Shannon *et al.*, 2004, p. 366; Tyler *et al.*, 2006; Ching *et al.*, 2007). In CI users, speech understanding in the presence of interferers has been shown to improve to some extent when bilateral CIs are provided, although much of this benefit seems to be derived from the head shadow effect (van Hoesel and Tyler, 2003; Litovsky *et al.*, 2006a, 2009), a physical effect that is monaurally mediated. In general, binaural advantages in CI users are minimal or absent relative to those reported in NHLs (e.g., van Hoesel and Clark, 1997; Muller *et al.*, 2002; Schon *et al.*, 2002; Gantz *et al.*, 2002; Litovsky *et al.*, 2006a; Buus *et al.*, 2008; Loizou *et al.*, 2009; Litovsky *et al.*, 2009). These differences indicate the presence of unknown factors that place limitations on CI users' performance. The current study examined the variable of spectral holes.

The binaural group had lower SRTs than the monaural group; however, the extent of the binaural advantage was subject to the presence of spectral holes. In the baseline condition (no holes), the binaural advantage was present in quiet, and with interferers in front, right, and left, consistent with previous reports regarding advantages of binaural hearing in quiet and in adverse listening environment (e.g., MacKeith and Coles, 1971; Bronkhorst and Plomp, 1989; Arsenault and Punch, 1999). With the introduction of spectral holes, the binaural group performed better than the monaural group only in the spatially separated conditions, with the exception of the quiet 10 mm basal condition. These results suggest that the advantages of binaural listening in the quiet and front were reduced by the presence of spectral holes regardless of location and size of hole. It has been shown that listeners exploit other perceptual cues to extract information from multiple speech signals in absence of spatial cues, such as spectral and temporal differences between a target and interfering sources (Tyler *et al.*, 2002; Assmann and Summerfield, 1990; Leek and Summers, 1993; Bird and Darwin, 1998; Bacon *et al.*, 1998; Vliegen and Oxenham, 1999; Summers and Molis, 2004). For example, Hawley *et al.* (2004) reported that effects of binaural listening and F0 differences seem to be interdependent such that advantages of F0 cues were larger in binaural than in monaural and that binaural advantages were greater when F0 cues were present. Hence, these perceptual cues are likely to be limited or less distinct when spectral holes are present, which would explain why binaural advantage was reduced in the quiet and front conditions.

The greatest advantage of binaural listening occurred when interferers were spatially separated to the right. Under monaural stimulation (right ear), listeners had substantially elevated SRTs when speech interferers were ipsilateral to the functional ear; a configuration that created an unfavorable SNR. When interferers were placed on the contralateral side (left), a substantial improvement in monaural SRTs occurred, yet performance was worse than binaural SRTs for some conditions. Specifically, the binaural group had lower left SRTs than the monaural group in the 6 mm hole conditions in the base and middle and 10 mm hole in the apex. These results suggest that binaural hearing might offer additional mechanisms that could improve SRTs in the presence of reduced spectral information. However, these advantages did not occur for 10 mm holes in the base and middle and 6 mm hole in the apex; it is likely that these results are driven by the reduced SRM in the basal and middle conditions. Additionally, given that left SRM was greater for apical hole conditions than for the baseline in the monaural conditions, considerable improvement in speech perception occurred which minimized the differences between binaural and monaural performances for the left spatial configuration. In general, these results indicate that the extent to which binaural benefits occurred for the left spatial conditions was susceptible to the size and location of the spectral holes.

Results further showed a significant main effect of 6 mm basal and apical holes on SRTs for the binaural conditions but not for the monaural conditions. It is important to note that these results do not necessarily indicate that binaural hearing is more susceptible to the presence of spectral holes than monaural hearing. To support this argument, several details need to be taken into account. (1) Although this effect was pronounced in the binaural conditions, right and left SRTs were still significantly lower (better) for binaural than monaural listening mode. (2) Given that effect of spectral holes was determined based on change in SRTs relative to the baseline conditions, these results could be produced by lower baseline SRTs in the binaural conditions and higher baseline SRTs in the monaural conditions. Additionally these results further suggest that the extent of degradation in SRTs that introduced by the presence of spectral holes occurred similarly for binaural and monaural conditions.

Similar comparisons of SRM data showed a robust advantage for binaural hearing in the right spatial configuration as this configuration created unfavorable SNR in the monaural conditions due to the proximity of the interferers to the functional ear. As such, this result points to the susceptibility of unilateral CI users to interfering speech. However, amount of SRM for the left configuration was comparable for the binaural and monaural listening modes; these results were found for the spectral hole conditions as well as the baseline condition. Overall these findings are consistent with previous results that showed similar extent in spatial unmasking for binaural and monaural (shadowed ear) conditions in bilateral CI users (van Hoesel and Tyler, 2003).

V. SUMMARY AND CONCLUSION

Findings from the current study demonstrated that spectral holes are detrimental to speech intelligibility as evidenced by elevated SRTs and reduced SRM; the extent of this deterioration depended on the cochleotopic location and size of the simulated spectral holes as well as the listening environment. A substantial increase in SRTs occurred when spectral holes were created in the middle of the simulated electrode array; this deterioration in performance was greater than that produced by basal and apical holes. Current results further demonstrated that loss of high-frequency information created by the presence of basal holes produced the greatest detriment to SRM. In the presence of limited spectral information, binaural hearing seems to offer greater advantages over monaural stimulation in acoustically complex environments, as demonstrated by lower SRTs. However, these advantages were reduced in the presence of spectral holes. Specifically, compared to the baseline conditions, binaural advantages were mostly observed when target and interfering speech were spatially separated. Regarding SRM, the greatest advantage for binaural over monaural (right ear) hearing was found when interfering speech was placed to the right, near the ear with good SNR. SRM was reduced in the presence of basal and middle holes, but slightly increased with apical holes, suggesting that loss of low-frequency inputs is related to there being greater dependence on spatial cues for unmasking. In general, these findings imply that listening under simulated CI conditions makes for greater susceptibility to holes in the mid-to-high-frequency regions of spectra in speech signals. However, the extent of this degradation in actual CI users may be different due to other factors not controlled here such as the neural-electrode interface, spread of current, and real holes in hearing due to poor neural survival. Finally, results further suggest the importance of spectral cues over temporal cues for speech intelligibility in the presence of limited spectral information.

ACKNOWLEDGMENTS

The authors are grateful to Shelly Godar, Tanya Jensen, and Susan Richmond for assisting with subject recruitment. This work was supported by NIH-NIDCD Grant No. R01DC030083 to R.L.

[1]HRTFs were obtained from the MIT Media Laboratory.

ANSI (**2004**). ANSI S3.6-2004, American National Standards specification for audiometers, American National Standards Institute, New York.

Arsenault, M., and Punch, J. (**1999**). "Nonsense-syllable recognition in noise using monaural and binaural listening strategies," J. Acoust. Soc. Am. **105**, 1821–1830.

Assmann, P., and Summerfield, Q. (**1990**). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," J. Acoust. Soc. Am. **88**, 680–97.

Bacon, S., Opie, J., and Montoya, D. (**1998**). "The effect of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds," J. Speech Lang. Hear. Res. **41**, 549–563.

Başkent, D., and Shannon, R. (**2006**). "Frequency transposition around dead regions simulated with a noiseband vocoder," J. Acoust. Soc. Am. **119**, 1156–1163.

Bell, T. S., Dirks, D. D., and Trine, T. D. (**1992**). "Frequency-importance functions for words in high-and low-context sentences," J. Speech Hear. Res. **35**, 950–959.

Bierer, J. A. (**2007**). "Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration," J. Acoust. Soc. Am. **121**, 1642–1653.

Bird, J., and Darwin, C. J. (**1998**). "Effects of a difference in fundamental frequency in separating two sentences," in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA).

Breeuwer, M., and Plomp, R. (**1984**). "Speechreading supplemented with frequency-selective sound-pressure information," J. Acoust. Soc. Am. **76**, 686–691.

Bronkhorst, A. (**2000**). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," Acust. Acta Acust. **86**, 117–128.

Bronkhorst, A., and Plomp, R. (**1988**). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am. **83**, 1508–1516.

Bronkhorst, A., and Plomp, R. (**1989**). "Binaural speech intelligibility in noise for hearing-impaired listeners," J. Acoust. Soc. Am. **86**, 1374–1383.

Bronkhorst, A., and Plomp, R. (**1992**). "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing," J. Acoust. Soc. Am. **92**, 3132–3139.

BussE., Pillsbury, H., Buchman, C., Pillsbury, C., Clark, M., Haynes, D., Labadie, R., Amberg, S., Roland, P., Kruger, P., Novak, M., Wirth, J., Black, J., Peters, R., Lake, J., Wackym, P., Firszt, J., Wilson, B., Lawson, D., Schatzer, R., D'Haese, PS, and Barco, A. (**2008**). "Multicenter U.S. bilateral MED-EL cochlear implantation study: speech perception over the first year of use," Ear Hear. **29**, 20–32

Ching, T., van Wanrooy, E., and Dillon, H. (**2007**). "Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: A review," Trends Amplif. **11**, 161–192.

Chung, D., and Mack, B. (**1979**). "The effect of masking by noise on word discrimination scores in listeners with normal hearing and with noise-

induced hearing loss," Scand. Audiol. **8**, 139–143.
Cochlear Corporation (**1995**). Technical reference manual, Engle-wood, Co.
Culling, J., Hawley, M., and Litovsky, R. (**2004**). "The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources," J. Acoust. Soc. Am. **116**, 1057–1065.
DePaolis, R. A., Janota, C. P., and Frank, T. (**1996**). "Frequency importance functions for words, sentences, and continuous discourse," J. Speech Hear. Res. **39**, 714–723.
Drullman, R., and Bronkhorst, A. (**2000**). "Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation," J. Acoust. Soc. Am. **107**, 2224–2235.
Dubno, J., Ahlstrom, J., and Horwitz, A. (**2002**). "Spectral contributions to the benefit from spatial separation of speech and noise," J. Speech Lang. Hear. Res. **45**, 1297–1310.
Duquesnoy, A. (**1983**). "Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons," J. Acoust. Soc. Am. **74**, 739–743.
Festen, J., and Plomp, R. (**1986**). "Speech-reception threshold in noise with one and two hearing aids," J. Acoust. Soc. Am. **79**, 465–471.
Fletcher, H., and Galt, R. H. (**1950**). "The perception of speech and its relation to telephony," J. Acoust. Soc. Am. **22**, 89–151.
French, N., and Steinberg, J. (**1947**). "Factors governing the intelligibility of speech sounds," J. Acoust. Soc. Am. **19**(1), 90–119.
Freyman, R., Balakrishnan, U., and Helfer, K. (**2001**). "Spatial release from informational masking in speech recognition," J. Acoust. Soc. Am. **109**, 2112–2122.
Friesen, L., Shannon, R., Başkent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implant," J. Acoust. Soc. Am. **110**, 1150–1163.
Fu, Q., and Nogaki, G. (**2005**). "Noise susceptibility of cochlear implant users, the role of spectral resolution and smearing," J. Assoc. Res. Otolaryngol. **6**, 19–27.
Fu, Q., and Shannon, R. (**1999**). "Recognition of spectrally degraded and frequency shifted vowels in acoustic and electric hearing," J. Acoust. Soc. Am. **105**, 1889–1900.
Fu, Q., Chinchilla, S., and Galvin, J. (**2004**). "The role of spectral and temporal cues in voice gender discrimination by normal hearing listeners and cochlear implant users," J. Assoc. Res. Otolaryngol. **5**, 523–260.
Gantz, B., Tyler, R., Rubinstein, J., Wolaver, A., Lowder, M., Abbas, P., Brown, C., Hughes, M., and Preece, J. (**2002**). "Binaural cochlear implants placed during the same operation," Otol. Neurotol. **23**, 169–180.
Garadat, S., Litovsky, R., Yu, Q., and Zeng, F. G. (**2009**). "Role of binaural hearing in speech intelligibility and spatial release from masking using vocoded speech," J. Acoust. Soc. Am. **126**, 2522–2535
Gardner, W., and Martin, K. (**1994**). HRTF measurements of a KEMAR dummy-head microphone, The MIT media lab machine listening group. http://sound.media.mit.edu/resources/KEMAR.html (Last viewed 2/4/2009).
Gelfand, S., Ross, L., and Miller, S. (**1988**). "Sentence reception in noise from one versus two sources: Effect of aging and hearing loss," J. Acoust. Soc. Am. **83**, 248–256.
Grant, K., and Braida, L. (**1991**). "Evaluating the articulation index for auditory-visual input," J. Acoust. Soc. Am. **89**, 2952–2960.
Greenwood, D. (**1990**). "A cochlear frequency-position function for several species—29 years later," J. Acoust. Soc. Am. **87**, 2592–2605.
Hartmann, R., Top, G., and Klinke, R. (**1984**). "Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea," Hear. Res. **13**, 47–62.
Hawley, M., Litovsky, R., and Colburn, H. (**1999**). "Speech intelligibility and localization in a multi-source environment," J. Acoust. Soc. Am. **105**, 3436–3448.
Hawley, M., Litovsky, R., and Culling, J. (**2004**). "The benefits of binaural hearing in a cocktail party: Effect of location and type of interferer," J. Acoust. Soc. Am. **115**, 833–843.
Kasturi, K., and Loizou, P. (**2002**). "The intelligibility of speech with holes in the spectrum," J. Acoust. Soc. Am. **112**, 1102–1111.
Kawano, A., Seldon, H., Clark, G., Ramsden, R., and Raine, C. (**1998**). "Intracochlear factors contributing to psychophysical percepts following cochlear implantation," Acta Oto-Laryngol. **118**, 313–326.
Kryter, K. (**1962**). "Validation of the articulation index," J. Acoust. Soc. Am. **34**, 1698–1702.
Leek, M., and Summers, V. (**1993**). "The effect of temporal waveform shape on spectral discrimination by normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **94**, 2074–2082.
Levitt, H. (**1971**). "Transformed up-down methods in psychophysics," J. Acoust. Soc. Am. **49**, 467–477.
Liberman, M., and Kiang, N. (**1978**). "Acoustic trauma in cats. Cochlear pathology and auditory-nerve activity," Acta Oto-Laryngol., Suppl. **358**, 1–63.
Lippman, R. (**1996**). "Accurate consonant perception without mid-frequency speech energy," IEEE Trans. Speech Audio Process. **4**, 66–69.
Litovsky, R. (**2005**). "Speech intelligibility and spatial release from masking in young children," J. Acoust. Soc. Am. **117**, 3091–3099.
Litovsky, R. Y., Johnstone, P., and Godar, S. (**2006c**). "Benefits of bilateral cochlear implants and/or hearing aids in children," Int. J. Audiol. **45**, 78–91.
Litovsky, R. Y., Johnstone, P. M., Godar, S., Agrawal, S., Parkinson, A., Peters, R., and Lake, J. (**2006b**). "Bilateral cochlear implants in children: Localization acuity measured with minimum audible angle," Ear Hear. **27**, 43–59.
Litovsky, R. Y., Jones, G. L., Agrawal, S., and van Hoesel, R. (**2010**). "Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans," J. Acoust. Soc. Am., **126**(6) (in press).
Litovsky, R. Y., Parkinson, A., and Arcaroli, J. (**2009**). "Spatial hearing and speech intelligibility in bilateral cochlear implant users," Ear Hear. **30**, 419–431.
Litovsky, R., Parkinson, A., Arcaroli, J., and Sammath, C. (**2006a**). "Clinical study of simultaneous bilateral cochlear implantation in adults: A multicenter study," Ear Hear. **27**, 714–731.
Litovsky, R., Parkinson, A., Arcoroli, J., Peters, R., Lake, J., Johnstone, P., and Yu, G. (**2004**). "Bilateral cochlear implants in adults and children," Arch. Otolaryngol. Head Neck Surg. **130**, 648–55.
Loizou, P., Hu, Y., Litovsky, R. Y., Yu, G., Peters, R., Lake, R., and Roland, P. (**2009**). "Speech recognition by bilateral cochlear implant users in a cocktail party setting," J. Acoust. Soc. Am. **125**, 372–383.
Long, C. (**2000**). "Bilateral cochlear implants: Basic psychophysics," Ph.D. thesis, MIT, Cambridge, MA.
Long, C., Eddington, D., Colburn, S., and Rabinowitz, W. (**2003**). "Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user," J. Acoust. Soc. Am. **114**, 1565–1574.
MacKeith, N., and Coles, R. (**1971**). "Binaural advantages in hearing of speech," J. Laryngol. Otol. **85**, 213–232.
Mehr, M. A., Turner, C. W., and Parkinson, A. (**2001**). "Channel weights for speech recognition in cochlear implant users," J. Acoust. Soc. Am. **109**, 359–366.
Miller, G., Heise, G., and Lichten, W. (**1951**). "The intelligibility of speech as a function of the context of the test material," J. Exp. Psychol. **41**, 329–335.
Moore, B., Huss, M., Vickers, D., Glasberg, B., and Alcantara, J. (**2000**). "A test for the diagnosis of dead regions in the cochlea," Br. J. Audiol. **34**, 205–224.
Muller, J., Schon, F., and Helms, J. (**2002**). "Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system," Ear Hear. **23**, 198–206.
Nadol, J. B., Jr. (**1997**). "Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation," Otolaryngol.-Head Neck Surg. **117**, 220–228.
Nordlund, B. (**1962**). "Physical factors in angular localization," Acta Oto-Laryngol. **54**, 75–93.
Pekkarinen, E., Salmivalli, A., and Suonpaa, J. (**1990**). "Effect of noise on word discrimination by subjects with impaired hearing, compared with those with normal hearing," Scand. Audiol. **19**, 31–36.
Pickett, J., and Pollack, I. (**1963**). "Intelligibility of excerpt from fluent speech: Effect of rate of utterance and duration of Excerpt," Lang Speech **6**, 151–164.
Plomp, R. (**1986**). "A signal-to-noise ratio method for the speech-reception SRT of the hearing impaired," J. Speech Hear. Res. **29**, 146–154.
Prosen, C., Moody, D., Stebbins, W., Smith, D., Sommers, M., Brown, N., Altschuler, R., and Hawkins, J. (**1990**). "Apical hair cells and hearing," Hear. Res. **44**, 179–194.
Rabinowitz, W., Eddington, D., Delhorne, L., and Cuneo, P. (**1992**). "Relations among different measures of speech reception in subjects using a cochlear implant," J. Acoust. Soc. Am. **92**, 1869–1881.
Rothauser, E., Chapman, W., Guttman, N., Nordby, K., Silbigert, H., Urbanek, G., and Weinstock, M. (**1969**). "IEEE recommended practice for speech quality measurements," IEEE Trans. Audio Electroacoust. **17**, 227–246.

Schleich, P., Nopp, P., and D'Haese, P. (**2004**). "Head shadow, squelch, and summation effects in bilateral users of the Med-El COMBI 40/40+ cochlear implant," Ear Hear. **25**, 197–204.
Schon, F., Muller, J., and Helms, J. (**2002**). "Speech reception thresholds obtained in a symmetrical four-loudspeaker arrangement from bilateral users of MED-EL cochlear implant," Otol. Neurotol. **3**, 710–714.
Schuknecht, H., and Neff, W. (**1952**). "Hearing after apical lesions in the cochlea," Acta Oto-Laryngol. **42**, 263–274.
Shannon, R., Fu, Q., Galvin, J., and Friesen, L. (**2004**). in *Cochlear Implants: Auditory Prostheses and Electric Hearing*, edited by F. G. Zeng, A. N. Popper, and R. R. Fay (Springer, New York), pp. 334–376.
Shannon, R., Galvin, J., III, and Baskent, D. (**2002**). "Holes in hearing," J. Assoc. Res. Otolaryngol. **3**, 185–199.
Shannon, R., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.
Shepherd, R., and Javel, E. (**1997**). "Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status," Hear. Res. **108**, 112–144.
Smith, D., Brown, N., Moody, D., Stebbins, W., and Nuttall, A. (**1987**). "Cryoprobe-induced apical lesions in the chinchilla. II. Effects on behavioral auditory thresholds," Hear. Res. **26**, 311–317.
Stephenson, L., Moody, D., Rarey, K., Norat, M., Stebbins, W., and Davis, J. (**1984**). "Behavioral and morphological changes following low-frequency noise exposure in the guinea pig," Assoc. Res. Otolaryngol. St. Petersburg , 88–89.
Stevens, K. (**1997**). "Articulatory-acoustic relationship," in *Handbook of Phonetic Science*, edited by W. J. Hardcastle and J. Laver (Blackwell, Oxford).
Stickney, G., and Assmann, P. (**2001**). "Acoustic and linguistic factors in the perception of band-pass-filtered speech," J. Acoust. Soc. Am. **109**, 1157–1165.
Stickney, G., Zeng, F. G., Litovsky, R., and Assmann, P. (**2004**). "Cochlear implant recognition with speech maskers," J. Acoust. Soc. Am. **116**, 1081–1091.
Summers, V., and Molis, M. (**2004**). "Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level," J. Speech Lang. Hear. Res. **47**, 245–256.
Throckmorton, C., and Collins, L. (**2002**). "The effect of channel interactions on speech recognition in cochlear implant subjects: Predictions from an acoustic model," J. Acoust. Soc. Am. **112**, 285–296.
Tonning, F. (**1971**). "Directional audiometry II. The influence of azimuth on the perception of speech," Acta Oto-Laryngol. **72**, 352–357.
Tyler, R., Gantz, B., Rubinstein, J., Wilson, B., Parkinson, A., Wolaver, A., Precce, J., Witt, S., and Lowder, M. (**2002**). "Three-month results with bilateral cochlear implants," Ear Hear. **23**, 80S–89S.
Tyler, R., Noble, W., Dunn, C., and Witt, S. (**2006**). "Some benefits and limitations of binaural cochlear implants and our ability to measure them," Int. J. Audiol. **45**, S113–S119.
van Hoesel, R. (**2004**). "Exploring the benefits of bilateral cochlear implants," Audiol. Neuro-Otol. **9**, 234–246.
van Hoesel, R. J. M., Jones, G. L., and Litovsky, R. Y. (**2009**). "Interaural time-delay sensitivity in bilateral cochlear implant users: Effects of pulse-rate, modulation-rate, and place of stimulation," J. Assoc. Res. Otolaryngol. **10**, 557–567.
van Hoesel, R., and Clark, G. (**1997**). "Psychophysical studies with two binaural cochlear implant subjects," J. Acoust. Soc. Am. **102**, 495–507.
van Hoesel, R., and Tyler, R. (**2003**). "Speech perception, localization, and lateralization with bilateral cochlear implants," J. Acoust. Soc. Am. **113**, 1617–1630.
Vermeire, K., Brokx, J., van de Heyning, P., Cochet, E., and Carpentier, H. (**2003**). "Bilateral cochlear implantation in children," Int. J. Pediatr. Otorhinolaryngol. **67**, 67–70.
Vliegen, J., and Oxenham, A. (**1999**). "Sequential stream segregation in the absence of spectral cues," J. Acoust. Soc. Am. **105**, 339–346.
Warren, R., Riener, K., Bashford, J., and Brubaker, B. (**1995**). "Spectral redundancy, intelligibility of sentences heard through narrow spectral slits," Percept. Psychophys. **57**, 175–182.
Wilson, B., Lawson, D., Muller, J., Tyler, R., and Kiefer, J. (**2003**). "Cochlear implant: Some likely next steps," Annu. Rev. Biomed. Eng. **5**, 207–249.

Median-plane sound localization as a function of the number of spectral channels using a channel vocoder

Matthew J. Goupell,[a] Piotr Majdak, and Bernhard Laback
Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, A-1040 Vienna, Austria

(Received 25 June 2009; revised 4 December 2009; accepted 9 December 2009)

Using a vocoder, median-plane sound localization performance was measured in eight normal-hearing listeners as a function of the number of spectral channels. The channels were contiguous and logarithmically spaced in the range from 0.3 to 16 kHz. Acutely testing vocoded stimuli showed significantly worse localization compared to noises and 100 pulse/s click trains, both of which were tested after feedback training. However, localization for the vocoded stimuli was better than chance. A second experiment was performed using two different 12-channel spacings for the vocoded stimuli, now including feedback training. One spacing was from experiment 1. The second spacing (called the speech-localization spacing) assigned more channels to the frequency range associated with speech. There was no significant difference in localization between the two spacings. However, even with training, localizing 12-channel vocoded stimuli remained worse than localizing virtual wideband noises by 4.8° in local root-mean-square error and 5.2% in quadrant error rate. Speech understanding for the speech-localization spacing was not significantly different from that for a typical spacing used by cochlear-implant users. These experiments suggest that current cochlear implants have a sufficient number of spectral channels for some vertical-plane sound localization capabilities, albeit worse than normal-hearing listeners, without loss of speech understanding. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3283014]

PACS number(s): 43.66.Qp, 43.66.Ts, 43.66.Ba [JCM] Pages: 990–1001

I. INTRODUCTION

Vertical-plane sound localization (i.e., the perception of elevation and the discrimination of front from back sound sources) depends primarily on directionally-dependent filtering introduced by reflections from the pinnae, head, and torso (e.g., Middlebrooks, 1999a; Algazi *et al.*, 2001). For free-field sources, these vertical-plane cues, together with the binaural cues (interaural time and level differences) for horizontal-plane localization, are typically represented by acoustical transfer functions called head-related transfer functions (HRTFs) (Shaw, 1974; Møller *et al.*, 1995). Due to the size of the pinnae, the relevant peaks and notches resulting from diffraction effects used in vertical-plane localization occur for frequencies between about 4 and 16 kHz (Blauert, 1969; Hebrank and Wright, 1974; Morimoto and Aokata, 1984; Middlebrooks, 1992; Blauert, 1997; Langendijk and Bronkhorst, 2002).

Despite the substantial amount of previous work on understanding the role of HRTFs in vertical-plane sound localization, it is still unclear as to the type, scale, size, and position of spectral features most important to this task. What is clear is that HRTFs are subject dependent (Wightman and Kistler, 1989; Wenzel *et al.*, 1993; Middlebrooks, 1999a). One method to investigate the role of spectral features has been to test localization abilities with spectrally-distorted HRTFs. These spectral distortions have been performed in a fairly uncontrolled way by occluding pinnae with some substance (Gardner and Gardner, 1973; Musicant and Butler, 1984; Oldfield and Parker, 1984; Hofman *et al.*, 1998). In more controlled experiments, signal processing methods have been used to distort spectral localization cues. For example, HRTF spectral distortions have been done by truncating HRTF impulse responses (Zahorik *et al.*, 1995; Senova *et al.*, 2002).

Several studies have shown that relatively broad high-frequency spectral features (on the order of octaves) are the relevant vertical-plane sound localization cues (Asano *et al.*, 1990; Kistler and Wightman, 1992; Kulkarni and Colburn, 1998; Langendijk and Bronkhorst, 2002; Macpherson and Middlebrooks, 2003; Kulkarni and Colburn, 2004; Qian and Eddins, 2008). For example, Kulkarni and Colburn (1998) systematically removed fine-scale HRTF spectral components and showed that substantial spectral smoothing of the HRTF spectrum could be performed before listeners' elevation errors and number of front-back confusions increased. Langendijk and Bronkhorst (2002) performed another type of spectral distortion by flattening 1/2-, 1-, and 2-octave bands in the sound spectrum. They found that there was no effect on sound localization when one 1/2-octave band was flattened in the sound spectrum. However, for the broader bands, a large effect was found. For the flattening of a 2-octave band, sound localization was impossible.

The experiments reported in this paper evaluated the potential of cochlear-implant (CI) users to perform median-plane sound localization via another method of spectral alteration. We tested median-plane sound localization in normal-hearing (NH) listeners using a CI simulation and investigated the number of channels necessary to present adequate spectral localization information. The advantages of using CI simulations with NH listeners were that the popu-

[a]Author to whom correspondence should be addressed. Electronic mail: goupell@waisman.wisc.edu

lation of NH listeners was much more readily available for the long testing times needed for sound localization experiments, and the interindividual variability is typically much less for NH listeners. The latter reason is due to many factors that affect CI users' performance, including the placement of the electrode array and the number of surviving spiral ganglion cells. CI simulations with NH listeners often provide an upper bound on CI user's performance in psychophysical tasks (e.g., Dorman and Loizou, 1997; Fu *et al.*, 1998; Friesen *et al.*, 2001; Carlyon and Deeks, 2002; van Wieringen *et al.*, 2003; Carlyon *et al.*, 2008; Goupell *et al.*, 2008b).

Sound localization in the horizontal plane in CI users has been the topic of many recent papers. In particular, with the ever increasing number of bilaterally-implanted CI users, there has been much work on interaural time and/or level difference sensitivity (Lawson *et al.*, 1998; Long *et al.*, 2003; Laback *et al.*, 2004; Majdak *et al.*, 2006; Laback *et al.*, 2007; van Hoesel, 2007; Grantham *et al.*, 2008; Laback and Majdak, 2008; van Hoesel, 2008) and sound localization in the horizontal plane (van Hoesel and Tyler, 2003; Nopp *et al.*, 2004; Seeber *et al.*, 2004; van Hoesel, 2004; Schoen *et al.*, 2005; Seeber and Fastl, 2008; van Hoesel *et al.*, 2008). However, much less attention has been paid to localization in the vertical planes in CI users. Results from a recent study by Majdak *et al.* (2008) showed that CI listeners using current clinical speech processors with behind-the-ear microphones had a substantial deterioration of vertical-plane localization performance compared to NH listeners.

The speech processors of multi-channel CIs typically separate acoustic signals into several spectral channels. This information is used to stimulate different tonotopic places with electrical pulse trains. The amount of transmitted spectral information is determined by the number of spectral channels used to analyze the signal, the amount of tonotopic overlap that occurs when the electrical pulse trains excite the auditory nerve, and the sensitivity in the excited region. Present day CIs typically have 12–24 electrodes. The bandpass signals typically subdivide the spectrum logarithmically between frequencies on the order of 0.1 and 10 kHz. For example, a MED-EL Combi 40+ implant has 12 electrodes that typically subdivide an incoming acoustic sound spectrum between 0.3 and 8.5 kHz. This means that there are nine electrodes presenting acoustic information below 4 kHz and three electrodes above 4 kHz. Since median-plane sound localization relies on frequencies above 4 kHz, using only three electrodes for this frequency region places CI listeners at a noticeable disadvantage for this task compared to NH listeners. Additionally, a large portion of the relevant sound spectrum (between 8.5 and 16 kHz) is omitted in this example, which may further hinder median-plane localization for CI listeners. Lastly, the typical placement of the microphone behind the ear and above the pinna causes a lack of directional filtering from the pinna, arguably the most important anatomical vertical-plane directional filter. It is obvious that substantial changes to CI speech processing strategies and systems are needed if they are to incorporate vertical-plane localization cues. However, these changes must respect the primary use of a CI, which is to provide speech understanding to profoundly hearing-impaired and deaf individuals. Therefore, new processing strategies should carefully balance the competing needs for speech understanding and sound localization.

Although spectral distortions can cause decreased localization performance, it has been shown that NH listeners can adapt to some localization-cue modifications after long-term, real-world experience (Hofman and Van Opstal, 1998). This is promising for CI users who, if post-lingually deafened, will have to relearn their auditory spatial map. Additionally, as mentioned above, studies show that vertical-plane localization relies upon fairly broadly tuned features in the HRTF spectrum. This is a necessity for CI users who, for this generation of CIs, can be presented only a few channels of spectral localization information.

The goal of this study was to determine the number of spectral channels necessary to perform median-plane sound localization. If this number is sufficiently low, then it may be possible for CI users to perform the task. The experiments tested median-plane localization in NH listeners using a CI simulation. The first experiment consisted of extensive procedural training and localization training to wideband (WB) virtual stimuli. This was followed by the localization of vocoded stimuli, where the number of logarithmically-spaced frequency channels was varied from 3 to 24, which were acutely tested. The second experiment consisted of long-term localization training and testing for the vocoded stimuli, which focused on two 12-channel spacings, one from experiment 1 and one with a custom configuration of channels designed to balance speech and vertical-plane localization cues. An additional experiment investigated a potential deterioration in speech understanding using a clinical spacing and the spacing introduced in experiment 2.

II. EXPERIMENT 1: ACUTE VOCODER MEASUREMENTS

A. Listeners and equipment

Eight listeners participated in the experiment. All eight listeners had audiometrically normal hearing and were between 21 and 46 years old.

The virtual acoustic stimuli were presented via headphones (Sennheiser HD580) in a semi-anechoic room. A digital audio interface (RME ADI-8) was used to present stimuli with a 48-kHz sampling rate and 24-bit resolution.

A visual environment was presented via head-mounted display (HMD; Trivisio 3-Scope). It provided two screens with a field of view of $32° \times 24°$ (horizontal $\times$ vertical dimensions). The HMD did not enclose the entire field of view. The visual environment was presented binocularly, the same picture used for both eyes. Listeners could adjust the interpupillary distance and the eye relief so that they viewed a single focused image. The position and orientation of the listener's head were measured by an electromagnetic tracker (Ascension Flock of Birds) in real-time. One tracking sensor was mounted on the top of the listener's head. The second tracking sensor was mounted on the end of a pointer, which was held by the listener. The tracking device was capable of measuring six degrees of freedom (three translations, three

rotations) at a rate of 100 measurements per second for each sensor. The tracker accuracy was 1.7 mm for translations and 0.5° for rotations.

B. HRTF measurements

The HRTFs were measured for each listener individually. Twenty-two loudspeakers (custom-made boxes with VIFA 10 BGS as drivers) were mounted on a metal arc (1.2-m radius) at fixed elevations from −30° to +80° relative to the listener's eye level. The loudspeakers were driven by amplifiers adapted from Edirol MA-5D active loudspeaker systems. The loudspeakers and the arc were covered with acoustic damping material to reduce the reflections from the construction. The listeners were seated in the center of the arc and had microphones (Sennheiser KE-4-211-2) placed in his/her ear canals, which were connected via pre-amplifiers (RDL FP-MP1) to the digital audio interface. Each HRTF was measured with a 1728.8-ms exponential frequency sweep from 0.05 to 20 kHz. The multiple exponential sweep method (MESM) was used to measure HRTFs in an interleaved and overlapped fashion for one azimuth and all elevations (Majdak *et al.*, 2007). After measuring HRTFs for 0° azimuth, the listener was rotated by 2.5° and the measurement of HRTFs for the next azimuth began. In total, 1550 HRTFs were measured for one listener, where the positions were distributed with a constant spherical angle on the sphere. During the procedure, the head position and orientation were monitored with the head tracker. The entire HRTF measurement procedure lasted approximately 20 min. The HRTFs were calculated from the recordings according to the MESM system identification procedure (Majdak *et al.*, 2007).

The equipment transfer functions were derived from a reference measurement performed by placing the in-ear microphones in the center of the arc and using the system identification procedure as before. The transfer function of the equipment was individually measured for each loudspeaker and removed from the HRTF measurements by filtering each HRTF with the appropriate inverse equipment transfer function.

Directional transfer functions (DTFs) were calculated according to the procedure of Middlebrooks (1999b). The magnitude of the common transfer function (CTF) was calculated by averaging the log-amplitude spectra of all the HRTFs. The phase of the CTF was the minimum phase of the CTF amplitude spectrum. The DTFs were the result of filtering the HRTFs with the inverse complex CTF. Since the headphone transfer function is the same for all positions, removing the CTF removes the headphone transfer function, which is known to be important for proper virtual stimulus externalization (Pralong and Carlile, 1996). No further headphone transfer function compensation was used. Finally, all the DTFs were temporally windowed with a Tukey window to a 5.33-ms duration. A typical set of DTFs for the median plane is shown in Fig. 1(a). More details about the HRTF measurement procedure can be found in Majdak *et al.* (2010).

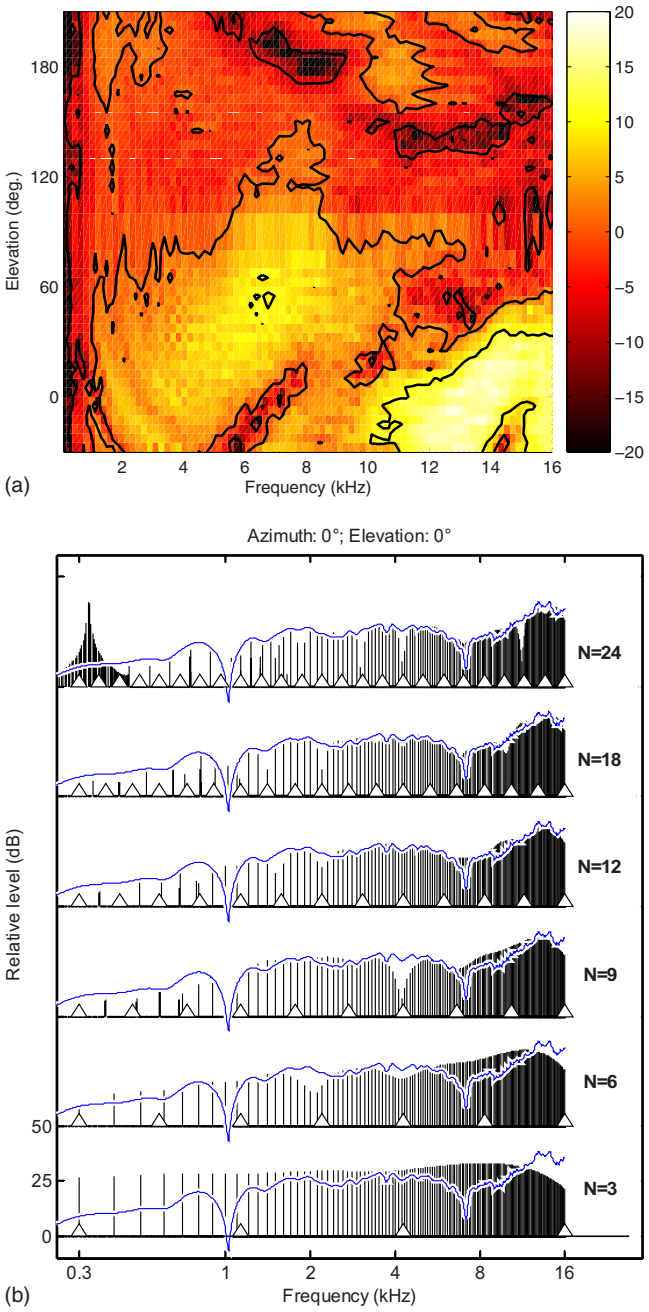


FIG. 1. (Color online) A set of DTFs in the median plane for a typical listener. Panel (a) shows the measured DTF. Panel (b) shows the amplitude spectra of the same DTF for 0° azimuth and 0° elevation (solid thin line) and processed by the GET vocoder (vertical lines) for different numbers of channels (N). Channel corner frequencies are marked by the triangles.

C. Stimuli

Three types of free-field stimuli were used as acoustic targets: WB Gaussian noises, WB click trains, and vocoded pulse trains. They were uniformly distributed along the median plane of a virtual sphere with the listener in the center of this sphere. Positions from −30° to +210° in elevation, relative to the eye level of the listener, were tested. These positions varied within the lateral range of ±10° of the median plane. Therefore, 290 of the 1550 measured DTFs were used.

The level of the stimuli was 50 dB with respect to the hearing threshold. The hearing threshold was estimated in an experimenter-controlled manual up-down procedure using a

target positioned at an azimuth and elevation of 0°. In the experiment, the level for each presentation was randomly roved within the range of ±5 dB to reduce the possibility of localizing targets based on level.

1. Wideband signals

The WB Gaussian white noises (to be called WB noises) and 100 pulse/s WB click trains (to be called WB clicks) had 500-ms duration, which included temporal shaping by a Tukey window with a 10-ms rise-fall time. The stimuli were filtered with the listener-specific DTFs.

2. Vocoded signals

A Gaussian-enveloped tone (GET) vocoder (Lu *et al.*, 2007) was used to simulate CI sound processing. A more conventional vocoder was not used for the following reasons. A sine vocoder has a sparse spectral representation, meaning that there is a single sine tone at the center frequency of the channel and possible sidebands, which may hinder the peak and notch detection necessary for vertical-plane sound localization. A GET vocoder has spectrally full channels with harmonics spaced at the pulse rate of the stimuli. We think that spectrally full channels better reproduce the electrical stimulation of a CI. A noise vocoder generates a nondeterministic signal, where random fluctuations make the characteristic spectral peaks and notches between 4 and 16 kHz difficult to visually identify. On the other hand, a GET vocoder generates a deterministic signal. Thus we needed only one token, and there was no problem in visually identifying the spectral peaks and notches. Lastly, we think that presenting pulsatile stimulation, albeit relatively low-rate acoustic pulses, better represents the electrical pulse trains delivered by a CI.

Detection of GETs has been previously studied in NH listeners (Gabor, 1947; van den Brink and Houtgast, 1990; van Schijndel *et al.*, 1999). Our targets were multi-channel GET trains with incorporated spatial information. The total processing scheme can be seen in Fig. 2. A single WB click was filtered with a DTF corresponding to the particular position of the target. The resulting directional impulse response was filtered into $N=3$, 6, 9, 12, 18, or 24 contiguous channels by a filter bank. The filters were eighth-order Butterworth bandpass filters. The lowest corner frequency was 0.3 kHz. The highest corner frequency was 16 kHz. The other corner frequencies were logarithmically spaced according to the value of N. Each channel, n, had a center frequency, f_n, defined as the geometric mean of the channel's corner frequencies. After filtering the DTF, each channel had energy, E_n.

For a single channel, n, a Gaussian pulse, $A_n(t)$, is given by

$$A_n(t) = \sqrt{\alpha_n f_n} \cdot e^{-\pi(\alpha_n f_n t)^2},$$

where α_n is the shape factor. The value of α_n was chosen so that the equivalent rectangular bandwidth, $B_n = \alpha_n f_n$, equaled the bandwidth of the corresponding bandpass filter for that channel. A GET, $P_n(t)$, is created by modulating a sinusoidal carrier with frequency f_n by the Gaussian pulse:

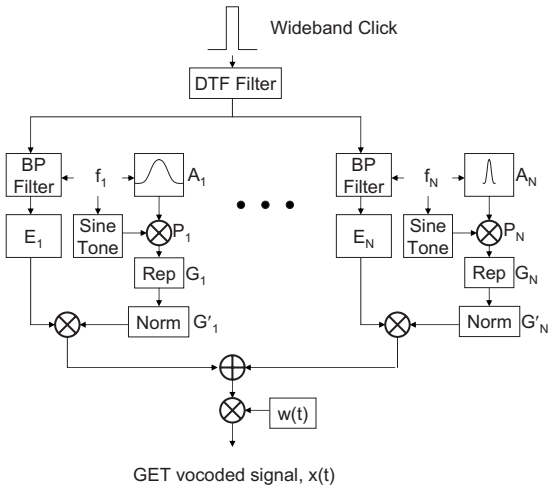


FIG. 2. Processing scheme for the GET vocoder. DTF information is bandpass filtered (BP filter) into N channels, where the energy (E_n) is measured. Gaussian envelopes modulate a sine tone, are replicated, and delayed (Rep) to make a GET train. The GET trains are energy normalized, weighted by the energy E_n from a channel of the DTF, and summed. After temporal windowing, the result is the GET-vocoded signal.

$$P_n(t) = A_n(t) \cdot \sin\left(2\pi f_n t + \frac{\pi}{4}\right),$$

where the phase shift of $\pi/4$ was used to keep $P_n(t)$'s energy independent of f_n (van Schijndel *et al.*, 1999). The single-channel GET train, $G_n(t)$, is the sum of 50 delayed GETs:

$$G_n(t) = \sum_{m=1}^{50} P_n\left(t - mT - \frac{T}{2}\right),$$

where T is the delay between each GET. The delay was 10 ms, which corresponds to a rate of 100 pulses/s. The $T/2$ phase shift was used to have $G_n(t)$ not begin at a maximum.

Note that for low f_n, especially when N is relatively large, $G_n(t)$ would have overlapping GETs. Hence, the modulation depth is not 100%. In such cases, if the Gaussian pulses modulate the carrier before being summed into a GET train, a spurious higher-order modulation of the signal would be introduced from the interfering phases of adjacent pulses. We determined that a modulation depth of 99% occurred if the equivalent rectangular duration of $A_n(t)$ was longer than 3.75 ms. To avoid overlapping pulses and unwanted modulations, if $A_n(t)$ is longer than 3.75 ms, then $G_n(t)$ is the sum of 50 Gaussian pulses, which *then* modulate the carrier:

$$G_n(t) = \sin(2\pi f_n t) \cdot \sum_{m=1}^{50} A_n\left(t - mT - \frac{T}{2}\right).$$

Note that by using this method, in the worst case of 0% modulation depth, the GET vocoder reduces to a sine vocoder.[1]

The amount of energy in $G_n(t)$ depends on f_n. Therefore, $G_n(t)$ was normalized with respect to its total energy, called $G'_n(t)$. Finally, the multi-channel GET train $x(t)$ is the sum over the normalized GET trains $G'_n(t)$ weighted by the energy E_n from the spatial information (i.e., the energy from each channel of the DTF):

$$x(t) = w(t) \cdot \sum_{n=1}^{N} G'_n(t) \cdot E_n,$$

where $w(t)$ is a Tukey window with a 10-ms rise-fall time. Figure 1(b) shows amplitude spectra of the same DTF as in Fig. 1(a) for 0° azimuth and 0° elevation (horizontal line) and processed by the GET vocoder (vertical lines) for different N values.[2]

D. Procedure

The listeners were immersed in a virtual sphere with a 5-m diameter. To facilitate the listeners' orientation, horizontal grid lines were placed every 5° and vertical grid lines every 11.25°. The reference position (azimuth and elevation of 0°), horizontal plane, and medial plane were marked with small spheres. Rotational movements were rendered in the virtual environment but not translational movements. The listeners could see the visualization of the hand pointer and its projection upon the sphere whenever they were in the listeners' field of vision. The projection of the pointer direction on the sphere's surface, calculated from the position and orientation of listeners' head and the pointer, was recorded as the indicated target position.

Prior to the acoustical tests, listeners performed a procedural training. In order to familiarize the listeners with the equipment and virtual environment, the listeners were trained to quickly and accurately respond to visual targets presented on the sphere. After the training, the listeners were able to respond to visual targets with an error smaller than 2°.

Acoustical sound localization training was provided, which was similar to that provided to listeners in Zahorik *et al.* (2006). In the acoustic training, at the beginning of each trial, the listeners aligned their head to the reference position. By pressing a button, the acoustic target was presented. During the presentation, the listeners were instructed not to move. After the acoustic presentation, the listeners were asked to point to the perceived position with the pointer and respond by pressing a button. This response was collected for further analysis. Next, a visualization of the acoustic target appeared on the surface of the sphere. The listeners were instructed to find the visualization, to point and respond to the visualization, and return to the reference position. After pressing the button, the same acoustic target with visualization was presented to the listeners and they had to point and respond to the acoustic target again. In total, for each acoustic target, listeners heard the stimulus twice, responded once to the initial perceived position, and responded twice to the actual position. The training was performed in blocks of 50 targets. Each block lasted for approximately 20–30 min. More details on the procedure and training are given in Majdak *et al.* (2010).

Listeners were first trained to WB noises within ±10° around the median plane (290 possible positions). For six listeners, the training consisted of 500–600 trials. For the other two listeners, the training consisted of 300 trials because they had extensive localization training over the entire sphere from previous studies. After training to WB noises, listeners were trained to WB clicks for 100 trials.

Seven conditions were tested in this experiment, the vocoded stimuli with $N=3$, 6, 9, 12, 18, and 24 channels and the WB clicks. As in the training, stimuli were randomly chosen from ±10° around the median plane. Due to the large number of positions, the same number of responses was not required from each position. Each condition consisted of three blocks of 100 trials and no feedback on the target position was given. The blocks were presented in random order for each listener.

E. Results

The metrics used to evaluate localization ability were the local polar error and the percentage of trials with quadrant errors. Localization errors were calculated by subtracting the target polar angles from the response polar angles. The polar error was the root-mean-square of the localization error. A quadrant error was defined as having an absolute polar error of greater than 90°. In tests with a large number of quadrant errors, the polar error is highly correlated to the quadrant error rate, showing the dominant role of the quadrant error rate in the metric. Therefore, the local polar error was used, which was the polar error after removing all the quadrant errors.[3] Assuming uniformly distributed random responses (i.e., guessing) within the range from −45° to 225°, the local polar error converges at 52° and the percentage of quadrant errors converges at 39%.[4]

Figure 3 shows the individual (top panels) and average (bottom panels) localization results. Data from the WB noise localization training were included with the experimental conditions. The local polar error and quadrant error rate for the WB clicks were similar to those for the WB noises. In general, the local polar error and quadrant error rate increased from the WB conditions to the vocoded conditions. For the local polar error, performance becomes worse for a decreasing number of channels, which appears to plateau around 18 channels. For the quadrant error rate, there seems to be a relatively flat plateau between 9 and 24 channels and increases for fewer channels. All of the conditions showed average local polar errors and average quadrant error rates much better than chance. Only one listener (NH41) showed chance performance for any vocoded condition.

For the WB noises, individual listener quadrant error rates were mostly within the measured range reported by Middlebrooks (1999b) for virtual WB noise stimuli. The local polar error shows some listeners near the lower limit of the Middlebrooks range. However, the average local polar error is near the upper limit of the Middlebrooks range. This discrepancy may be due to the fact that Middlebrooks tested targets distributed in the whole lateral range, not just near the median plane. The average quadrant error rate and standard deviation correspond well to the Middlebrooks range.

A repeated-measures analysis of variance (RM ANOVA) was performed on the factor N, including the WB noises and WB clicks as additional factor levels for the local polar error and quadrant error rate. There was a significant effect of N for both metrics ($p<0.0001$ for both).

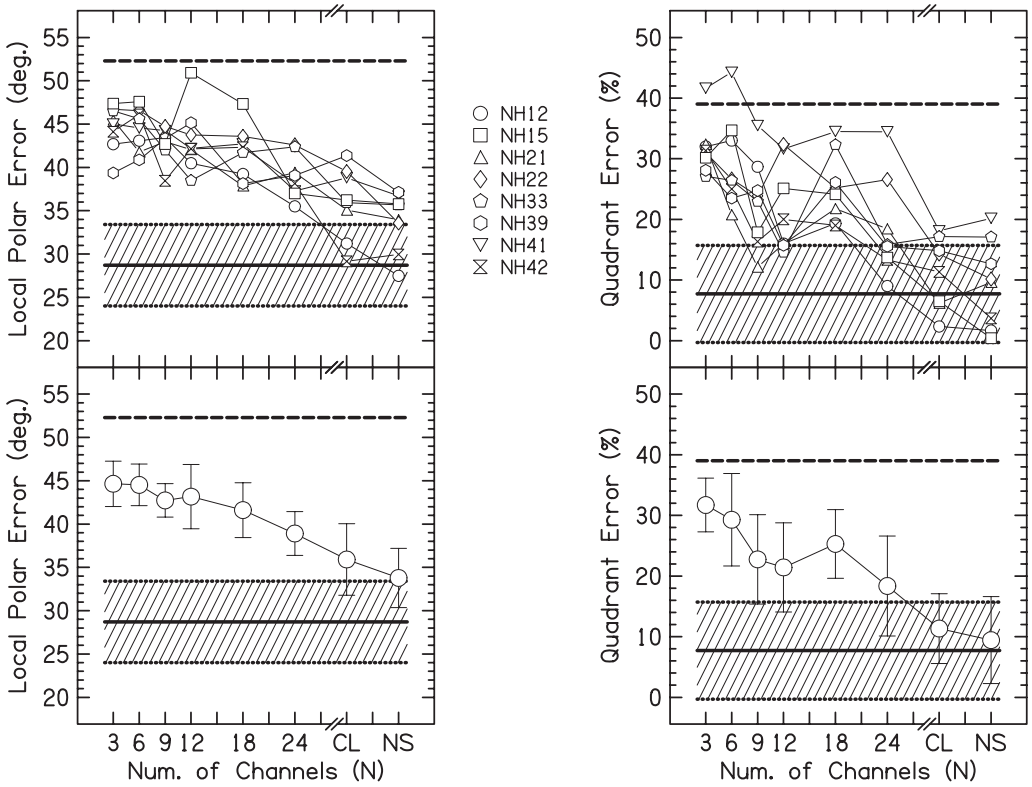


FIG. 3. Results of experiment 1. The upper panels show the individual data. The lower panels show the listener average and ±1 standard deviation. The left column shows the local polar error in degrees as a function of the number of channels. The right column shows the percentage of quadrant errors. Results for the WB clicks (CL) and WB noises (NS) are also included. The dashed lines show chance performance. The shaded area shows the average (solid line) and ±1 standard deviation (dotted lines) of the results from Middlebrooks (1999b) for virtual WB noise stimuli.

Tukey HSD *post-hoc* tests were performed for the local polar error and quadrant error rate. For the local polar error, there was no significant difference between the WB noises and WB clicks ($p=0.75$). There were significant differences between the WB conditions and the vocoded conditions ($p<0.007$ for all), with the exception of the difference between $N=24$ and the WB clicks ($p=0.33$). The differences between $N=24$ and $N=18$, 9, and 6 were not significant ($p>0.1$ for all), but the differences between $N=24$ and $N=12$, 6, and 3 were significant ($p<0.05$ for all). There were no significant differences between any pair of $N=18$, 12, 9, 6, and 3 at the 0.05 level.

For the quadrant error rate, there was no difference between the WB noises and WB clicks ($p=0.99$). There were significant differences between the WB conditions and the vocoded conditions ($p<0.011$ for all). The difference between $N=24$ and $N=18$ was not significant ($p=0.090$), but the differences between $N=24$ and $N=12$, 9, 6, and 3 were significant ($p<0.003$ for all). The differences between $N=18$ and $N=12$, 9, 6, and 3 were not significant ($p>0.098$ for all). The difference between $N=12$ and $N=9$ was not significant ($p=1$), but the differences between $N=12$ and $N=6$ and 3 were significant ($p<0.038$ for both). The difference between $N=9$ and $N=6$ was not significant ($p=0.14$), but the difference between $N=9$ and $N=3$ was significant ($p=0.011$). The difference between $N=6$ and $N=3$ was not significant ($p=0.97$).

F. Discussion

The localization of WB noises corresponded well to that measured by Middlebrooks (1999b), who tested well-practiced listeners with virtual acoustic WB noises in both the horizontal and vertical dimensions. As expected, the localization of our virtual sources is worse than that of the real sources measured in Middlebrooks (1999b). In our experiment, there was no significant difference between the WB noises and clicks, even though there was markedly less training for the clicks. Several studies have found that there are level, duration, and rate effects in vertical-plane localization (e.g., Vliegen and Van Opstal, 2004). However, few studies on vertical-plane localization have directly compared long-duration WB noises and click trains. Hartmann and Rakerd (1993) tested 880-ms WB noises and 10.4 pulses/s click trains in a speaker identification task. For levels comparable to the levels used in our experiment, localization performance of click trains appeared slightly worse than that of WB noises. Macpherson and Middlebrooks (2000) showed that the localization of click stimuli was qualitatively similar to 3-ms noise bursts. Hofman and Van Opstal (1998) tested 500-ms WB noises and 500-ms burst trains (3-ms noise bursts) at various rates, including 100 pulses/s. They found that the localization performance in the vertical direction for burst trains at 100 pulses/s was comparable to that for WB noises. Lower rates caused worse performance for burst trains compared to the WB noises. However, all three of these studies did not include a statistical comparison of the

localization of clicks and noises. Also, the localization training for these three studies was not extensive; some testing could be considered acute for some listeners. Hence, our result that there is no significant difference in the localization of WB noise and clicks is not inconsistent with these studies.

There was a marked increase in the local polar error and quadrant error rate when the GET vocoder was used, even for 24 channels. On one hand, performance was expected to be worse because the spectral cues were intentionally degraded and listeners received no training for the vocoded conditions, unlike the WB conditions. On the other hand, this is slightly surprising because it has been shown that the broadband spectral cues are sufficient for vertical-plane localization. For example, Kulkarni and Colburn (1998) showed that HRTF amplitude spectra smoothed to as few as 32 Fourier coefficients could be used before sound localization significantly deteriorated.[5] This corresponds to approximately five peaks and five notches in the sound spectrum in the frequency range from 4 and 16 kHz. Our 24-channel vocoded signal has approximately nine channels in the same frequency range, thus having nine potential critical points in the sound spectrum. However, it may be that the spectral contrast between channels in the vocoder is less than that of a HRTF with a smoothed sound spectrum. The GET vocoder had channels that were contiguous with respect to their corner frequencies, and the slopes of the channels were approximately 12 dB/oct; thus the channels have spectral profile information that overlaps. An extrapolation of this result to CI users, who typically have extensive channel interactions, which are analogous to overlapping channels, is that vertical-plane sound localization may be limited by the spectral resolution and contrast, even for 24 electrodes.

A recent study by Qian and Eddins (2008) explored the effect of varying the spectral modulation content of WB noises on virtual localization performance. In that study it was found that removing spectral modulation from 0.1 to 0.4 cycles/octave or from 0.35 to 0.65 cycles/octave resulted in significantly poorer localization (unsigned elevation difference between target and response angle) for low elevations (−30° to −20°) for three of six listeners. The vocoder used in the present experiment removed spectral modulation content, especially for a small number of channels. For example, for a stimulus at −30° elevation, the modulation spectrum was at least 10 dB down between 0.3 and 0.4 cycles/octave for $N=3$ compared to a WB noise. We analyzed a portion of our data as a function of angle by splitting the data into low elevation target (−30° to 0°) and high elevation target (0° to 30°) groups. Using a RM ANOVA (factors N and elevation), for the local polar error, there was a significant effect of N ($p<0.0001$), no significant effect of elevation ($p=0.26$), and no significant interaction between N and elevation ($p=0.083$). For the quadrant error, there was a significant main effect of N ($p<0.0001$), no significant effect of elevation ($p=0.97$), and no significant interaction between N and elevation ($p=0.84$). The RM ANOVA was repeated with just WB noises and $N=3$, the comparison that should show the largest contrast in spectral modulation content. The results did not change for the quadrant error rate, but did change for the local polar error because the interaction between N and elevation became significant ($p=0.028$). Therefore, this significant interaction may show modest support that removing spectral modulation information around 0.3–0.4 cycles/octave affects localization performance at low elevations. Note that there are several differences between this study and Qian and Eddins (2008), the most important being that we used individualized HRTFs while they used non-individualized HRTFs (although they did customize the HRTFs to each listener). As expected when using individualized vs non-individualized HRTFs, the overall performance of their listeners was much poorer than ours. For example, their average front-back confusion rate was $31\pm8\%$ compared to our $9.4\pm7.2\%$ average quadrant error rate for WB noises. Hence, it appears that the quality of the HRTFs used may affect the interpretation of the importance of spectral modulation cues in vertical plane sound localization.

Although localization performance was worse for the vocoder conditions, it was always better than chance. Said another way, CI listeners with their poorer frequency selectivity will be hindered in localizing sounds compared to NH listeners, but it seems possible to present salient vertical-plane localization cues to CI users. In the experiment, performance was roughly constant from 3 to 18 channels in local polar error and from 9 to 18 channels in quadrant error rate. Listeners also retained some sense of front-back directionality even for three channels. Median-plane sound localization with three channels may seem surprising at first. However, examining a typical set of DTFs [see Fig. 1(a)], the back positions can be approximated as a low-pass filtering of the front positions. Hence, one might predict some median-plane localization abilities with as few as two channels, a low-frequency channel to be used as a reference and a high-frequency channel for the front-back information. A similar view to median-plane localization was taken by Iida *et al.* (2007).

III. EXPERIMENT 2: VOCODER TRAINING

The results of experiment 1 showed that there is little difference in localization ability for NH listeners acutely tested using a CI simulation from 9 to 18 channels. This experiment will address two issues from experiment 1: long-term training with GET-vocoded stimuli and maintaining adequate resolution for the low-frequency channels associated with speech understanding. Majdak *et al.* (2010) showed that training on the order of several hundred trials is essential for saturation in the localization performance of virtual sound sources. It was hypothesized that listeners would need training for saturation in localization performance for GET-vocoded virtual sound sources, and that the results of experiment 1 underestimate the localization performance. It was also hypothesized that given the results of speech understanding tests from Goupell *et al.* (2008b), it would be possible to balance the competing demands of speech understanding and vertical-plane sound localization without significant loss of performance in either.

A. Stimuli

Stimuli were GET vocoded like those in experiment 1. The number of channels was fixed at $N=12$, which corresponds to the number of electrodes in the MED-EL Combi 40+, Pulsar, and Sonata implants. This condition showed localization performance better than chance in experiment 1, and little decrease in localization performance compared to $N=24$, the maximum number of electrodes available for any current CI.

Two spacings were used in this experiment. The first, called the "Log" spacing, corresponded to the spacing used for $N=12$ of experiment 1. The second, called the "speech-localization (SL)" spacing, represented an attempt to preserve speech information with a minimal sacrifice of spatial information. It is justified as follows.

Goupell *et al.* (2008b) low-pass filtered speech signals at cutoff frequencies of 8500, 4868, 2788, 1597, or 915 Hz for CI and NH listeners using a CI simulation while keeping the lower-frequency boundary fixed at 300 Hz. They found that there is almost no improvement in speech understanding performance beyond eight spectral channels (cutoff frequency of 2788 Hz). Keeping this in mind, the corner frequencies for the SL spacing were chosen such that the first eight low-frequency channels were logarithmically spaced from 0.3 to 2.8 kHz. This allowed for adequate resolution of the speech signal over the first two formants of speech. We refer to these eight channels as "speech" channels. The four higher-frequency channels were nearly logarithmically spaced from 2.8 to 16 kHz.[6] This was to include more of the high-frequency vertical-plane localization information that is typically omitted in CI processing of acoustic waveforms. We refer to these four channels as "spatial" channels. Visual inspection of the individualized HRTFs showed that this choice of corner frequencies yields reasonable contrast for the spatial channels so that vertical-plane sound localization might still be possible. Therefore, we chose the SL spacing in an attempt to allow for both good speech understanding and vertical-plane localization ability. As in experiment 1, the center frequency of each channel of the GET vocoder corresponded to the geometric mean of the corner frequencies. The information for each channel is given in Table I.

B. Procedure

The eight listeners of experiment 1 were split into two groups of four listeners each. The groups were chosen with respect to the rank-ordered quadrant error rate from experiment 1 for the condition $N=12$, which attempted to balance the average quadrant error rate for the groups. One group had listeners ranked 1, 4, 5, and 8 (average quadrant error rate was 22.0%). The other group had listeners ranked 2, 3, 6, and 7 (average quadrant error rate was 20.8%).

Both groups were provided acoustic training to Log-spaced or SL-spaced vocoded signals before testing. The training was similar to that in experiment 1. As before, listeners were trained until performance saturated. This was between 300 and 700 items depending on the listener. After the training, listeners were tested without feedback for 300 items. After data were taken for one type of spacing, the listeners were trained and tested on the other type of spacing in a similar fashion. Therefore, the results of experiment 2 show data from all eight listeners for each condition.

TABLE I. Stimulus information for the spacings used in experiment 2.

Mapping	Channel	f_{lower} (Hz)	f_{upper} (Hz)	f_c (Hz)	BW (Hz)
Log	1	300	418	354	118
	2	418	582	493	164
	3	582	811	687	229
	4	811	1 129	957	318
	5	1 129	1 573	1 333	444
	6	1 573	2 191	1 856	618
	7	2 191	3 052	2 586	861
	8	3 052	4 251	3 602	1199
	9	4 251	5 921	5 017	1670
	10	5 921	8 247	6 988	2326
	11	8 247	11 487	9 733	3240
	12	11 487	16 000	13 557	4513
SL	1	300	396	345	96
	2	396	524	456	128
	3	524	692	602	168
	4	692	915	796	223
	5	915	1 209	1 052	294
	6	1 209	1 597	1 390	388
	7	1 597	2 110	1 836	513
	8	2 110	2 788	2 425	678
	9	2 788	4 200	3 422	1412
	10	4 200	6 400	5 185	2200
	11	6 400	10 000	8 000	3600
	12	10 000	16 000	12 649	6000

C. Results

The results of the experiment are shown in Fig. 4. Note that the WB noise and WB click data were taken from experiment 1. Qualitatively, the results show that the local polar error and quadrant error rate were smallest for the WB noise, followed by the WB clicks, followed by the Log spacing, followed by the SL spacing. The errors for the SL spacing were slightly larger than the errors for the Log spacing. Results for all conditions were much better than chance performance. Comparing the results of the training on the Log spacing to the same condition with acute testing (experiment

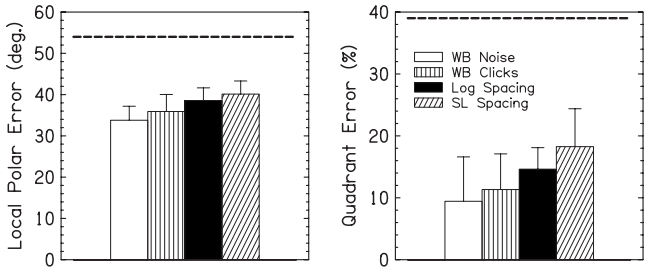


FIG. 4. Results of experiment 2 averaged over listeners. Error bars show one standard deviation of the mean. The dashed lines show chance performance. Listeners were trained to all conditions before testing. Data from the WB noise and WB click conditions are repeated from experiment 1.

TABLE II. Experiment 2: p-values for differences between conditions. Significant p-values (at the 0.05 level) are in bold.

Local polar error	WB clicks	Log	SL
WB noise	0.14	**0.0003**	**<0.0001**
WB clicks	···	**0.047**	**0.001**
Log	···	···	0.37
Quadrant error	**WB clicks**	**Log**	**SL**
WB noise	0.84	0.15	**0.005**
WB clicks	···	0.51	**0.034**
Log	···	···	0.42

1, condition $N=12$), on average, the local polar error decreased by 4.6° and the quadrant error rate decreased by 6.8%.

A RM ANOVA was performed over the factor condition for the local polar error and the quadrant error rate. There was a significant effect of condition for the local polar error and quadrant error rate ($p<0.0001$ and $p=0.006$, respectively). Table II shows p-values for *post-hoc* tests for both metrics. The WB conditions were significantly different from the vocoded conditions for the local polar error. The WB conditions were significantly different from the SL spacing for the quadrant error rate. The differences were not significant between WB conditions for both metrics. The differences were not significant between the Log and SL spacings for both metrics.

D. Discussion

This experiment aimed to test two hypotheses. The first hypothesis was that listeners could improve their localization performance of GET-vocoded stimuli with training when compared to experiment 1. Listeners' performance did improve significantly, but performance for the GET-vocoded stimuli remained slightly worse than that for the WB noise. Although there is a small but significant difference in localization performance between the Log spacing and the WB noise, listeners still performed much better than chance.

The second hypothesis was that listeners could localize using the SL spacing, which attempts to compromise between adequate spectral resolution of low frequencies (required for speech understanding) and the necessity of spectral information from 4 to 16 kHz (required for vertical-plane localization), without a marked decrease in localization performance. Listeners localized well below chance performance with the SL-spacing stimuli. This performance was significantly worse than that for the WB noise, but there was no significant difference from the Log-spacing stimuli.

IV. EXPERIMENT 3: SPEECH TEST

The previous experiment showed that NH listeners using a CI simulation could localize in the median plane relatively well using only 12 channels, 8 channels assigned to speech frequencies and 4 channels assigned to spectral localization frequencies. To ensure that the SL spacing does not degrade speech understanding, we performed a speech test using this spacing and a spacing similar to the clinical mapping used in current CI processing strategies.

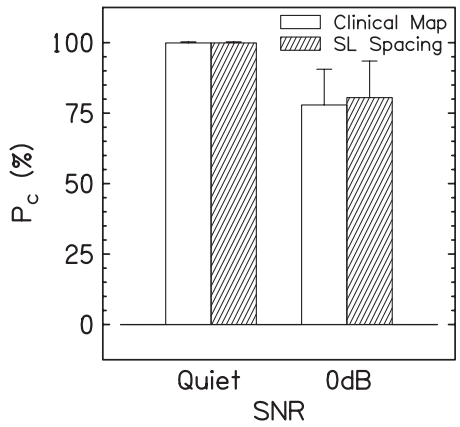


FIG. 5. Results of experiment 3 averaged over listeners, the percentage of correct words (P_c) for two spacings, and two SNRs in a speech understanding test. Error bars show one standard deviation of the mean.

A. Methods

Three sessions of feedback training and subsequent testing of speech understanding were performed using the Oldenburg Sentence Test. The procedure was similar to that in experiment 2 of Goupell *et al.* (2008b). The feedback training session consisted of listening to sentences with visual feedback on the computer screen. Eighty sentences were presented at four different signal-to-noise ratios (SNRs): +10, 5, and 0 dB, or in quiet. The testing session consisted of 40 sentences presented in a random order either at a 0-dB SNR or in quiet, 20 sentences for each SNR. Because of ceiling effects, it was not necessary to include the 10- and 5-dB conditions. At beginning of each session, ten sentences were presented in quiet to eliminate any short-term adaptation effects. The percentage of correct words was calculated from the third of the three sessions.

The training and testing were performed for two spacings, the SL spacing and a spacing similar to the processing of the Combi40/40+ and Pulsar CIs, called the "clinical map." The clinical map was the baseline condition in Goupell *et al.* (2008b). It had a lower-frequency boundary of 300 Hz, an upper-frequency boundary of 8500 Hz, and had 12 contiguous logarithmically-spaced frequency channels. A noise vocoder was used to simulate CI processing and the corner frequencies of the synthesis channels corresponded to that of the analysis channels. More details of the vocoder and procedure can be found in Goupell *et al.* (2008b).

The listeners were the same as the previous experiments, except that NH42 did not participate because of her limited availability. NH42 was replaced by NH10, the first author of the paper, who had audiometrically normal hearing and was 29 years old.

B. Results and discussion

Figure 5 shows the results of the experiment. There was very little difference between the two spacings, as expected.

In a RM ANOVA (factors: SNR and spacing), the difference was significant between SNRs ($p<0.0001$), but not between spacings ($p=0.58$). The interaction was also not significant ($p=0.58$). Therefore, the small change to the SL spacing did not decrease speech scores from the reference map for NH listeners. This was expected given the results of Goupell *et al.* (2008b), who tested speech understanding with a similar condition (called the extended-frequency range mapping or "$M_{14}N_{12}$") that had an upper-frequency boundary of 16 kHz. The difference between the SL spacing and the extended-frequency range mapping is that the SL spacing mapped frequencies to the appropriate tonotopic places, while the extended range mapping compressed the frequency information from 0.3 to 16 kHz to tonotopic places from 0.3 to 8.5 kHz. NH listeners did not show a decrease for the extended-frequency range mapping compared to the clinical map. CI listeners did show a significant decrease between these conditions. Thus, if CI listeners were tested with the SL spacing, it remains possible that they will have decreased speech understanding. Therefore, it is imperative that CI speech understanding is tested in future experiments for any novel spacing that includes a larger than typical frequency range.

V. GENERAL DISCUSSION

Experiments 1 and 2 tested NH listeners' ability to localize WB and GET-vocoded sound sources in the median plane. The experiments showed that while localization performance was worse for the spectrally degraded vocoded stimuli, localization performance was better than chance. This result validates results from previous studies, which maintain that the gross spectral structure is needed for vertical-plane sound localization (e.g., Kulkarni and Colburn, 1998; Langendijk and Bronkhorst, 2002; Qian and Eddins, 2008). In fact, our reasonably good localization performance was maintained for 9 or 12 channels logarithmically spaced from 0.3 to 16 kHz. Such a result is promising for the incorporation of spectral localization information in CI processing strategies.

Experiment 2 included a special 12-channel spacing, called the SL spacing, in the localization tests. The idea behind this spacing was a compromise between the competing demands of including adequate speech and median-plane localization information. Assuming that frequencies up to 2.4 kHz are essential for adequate vocoded-speech understanding (Goupell *et al.*, 2008b), the Log spacing provided seven speech channels and the SL spacing provided eight speech channels, thus a difference of approximately one channel assigned to spatial information. Experiment 2 showed that localization with the SL spacing was not different from localization with the Log spacing, which means that it is a viable option for future vertical-plane localization testing in CI listeners. Experiment 3 confirmed that speech understanding was not hindered by using the SL spacing.

Of course, localization of real sounds is better than virtual sounds. The local polar error is $22.7\pm5.1°$ for real sounds compared to $28.7\pm4.7°$ for virtual sounds, and the quadrant error rate is $4.6\pm5.9\%$ for real sounds compared to $7.7\pm8.0\%$ for virtual sounds (Middlebrooks, 1999b). For the 12-channel vocoder with the SL spacing of channels, there was an average degradation of 6.3° in local polar error and 8.8% in quadrant error rate for the virtual vocoded stimuli compared to the virtual WB noises. There was an average degradation of 17.4° in local polar error and 13.6% in quadrant error rate compared to the real WB sounds of Middlebrooks (1999b). Assuming no other deficits from using a CI than channelization, CI users will have less ability to distinguish elevation (increase in local polar error in the range from 6.3° to 17.4°) and will confuse front and back more often (increase in quadrant error rate in the range from 8.8% to 13.6%) than NH listeners.

All of the tests were performed with NH listeners using a CI simulation. Testing with CI listeners will also need to be done to determine the usefulness of the SL spacing for vertical-plane localization and speech understanding. Currently, CI listeners using behind-the-ear microphones and a clinical processing scheme (the highest frequency used was 8.5 kHz) cannot localize in the vertical planes (Majdak *et al.*, 2008). This study shows that the incorporation of high frequencies and spectral HRTF information, possibly by an in-the-ear microphone or other types of directional microphone, may provide salient cues for CI listeners, but perception of these cues is understandably worse than for NH listeners. Vertical-plane localization in CI users is expected to be even worse than that measured here for the NH listeners using a CI simulation for a myriad of reasons. One reason is that electrode arrays do not necessarily stimulate tonotopic places with matching frequency information. Ketten *et al.* (1998) reported the most apical electrode of some CIs, which often receive frequency information around 300 Hz, at a cochlear place tuned to greater than 1.4 kHz. Such a frequency-to-place mismatch is detrimental to speech understanding (e.g., Fu and Shannon, 1999). It could also be detrimental for vertical-plane sound localization, assuming that the CI user previously had hearing and developed an auditory map of space. The problem of tonotopic mismatch could be compounded by dead regions in the cochlea (Shannon *et al.*, 2001); important spectral features may be lost for some elevations, which would probably greatly hinder vertical-plane localization. Another reason is that the fidelity of current CI processing schemes cannot completely reconstruct a CI listener's individualized HRTF. It is well known that vertical-plane localization is worse with non-individualized than individualized HRTFs (e.g., Wenzel *et al.*, 1993). Despite all of these hurdles to overcome, it is possible that the plasticity of the auditory system would be able to cope with these potential problems. CI users may be able to learn a new auditory map of space, which was shown to be possible in NH listeners by Hofman *et al.* (1998).

Vertical-plane sound localization will also be affected by the poor spectral resolution in CI listeners. CIs typically produce relatively large excitation patterns in the cochlea (Nelson *et al.*, 2008). This large spread of excitation translates to a much poorer spectral resolution for CI listeners compared to NH listeners, which has been measured with a ripple-noise reversal test (Henry and Turner, 2003; Henry *et al.*, 2005). Related to this is the small number of perceptual channels available in CI users. Although we have assumed that 12

electrodes will translate to 12 perceptual channels for sound localization, many speech studies have shown that there are only about eight distinct perceptual channels in CI users (Dorman and Loizou, 1997; Fu and Shannon, 1999; Friesen *et al.*, 2001; Başkent and Shannon, 2005). Nevertheless, we have shown that median-plane sound localization remains quite good for as little as nine channels, nearly the same performance as 18 channels. Additionally, vertical-plane sound localization remains possible even with as few as three or six channels, although approaching chance performance. Even if the inclusion of spectral localization cues to CI processing simply gives the ability to distinguish front from back sources, this may well be worth it.

Lastly, Goupell *et al.* (2008a) showed that CI listeners depended mostly on intensity cues rather than spectral shape cues in several types of profile analysis tasks, tasks that are necessary to detect spectral vertical-plane sound localization cues. Detecting spectral shape cues may be the major challenge to cope with in addressing vertical-plane sound localization in CI users.

ACKNOWLEDGMENTS

We would like to thank our listeners for participating in this study and Michael Mihocic for running the experiments. We would also like to thank the Associate Editor John Middlebrooks, Fred Wightman, and an anonymous reviewer for comments about a previous version of this work. Portions of this study were presented at the Conference on Implantable Auditory Prostheses in Lake Tahoe in 2009. This study was funded by the Austrian Science Fund (FWF Project No. P18401-B15).

[1]The advantage of the first method of generating GET trains (generating a single Gaussian envelope, modulating the carrier with a fixed phase, and summing several GETs into a GET train) over the second method (generating the envelope for the entire train and modulating the carrier to obtain a GET train) is that the peak amplitude in all single GETs is constant for the first method but not necessarily constant for the second. Although the second method removes some higher-order modulations due to overlapping GETs, it cannot control the peak amplitude in each GET.

[2]The first channel for $N=24$, the smallest analysis bandwidth used in this experiment, had an unstable filter configuration. This perceptually translated into a strong pitch around 300 Hz. The entire stimulus was not perceptually distinct from the $N=18$ condition. Because of low frequency of this anomaly, we assumed that it would not affect median-plane sound localization.

[3]Middlebrooks (1999b) removed all *targets* that were outside a ±30° lateral region. We removed all *responses* that were outside a ±30° lateral region. For experiment 1, 8.2% of the data were removed because of this.

[4]Although the range of the experimental stimuli was from −30° to 210°, listeners were not restricted to this range in their responses. Hence, we used a slightly larger range in the calculation of chance performance to simulate this.

[5]Note that the smoothing performed in Kulkarni and Colburn (1998) was linear-frequency based, not log-frequency based. Although the high-frequency spectral localization features appear to have approximately linear-frequency scaling, since the auditory system has approximately log-frequency scaling, there was effectively less smoothing at higher frequencies than low frequencies.

[6]Impulse responses between $n=9$ and 10 canceled with logarithmic spacing, so we chose slightly different corner frequencies to avoid this.

Algazi, V. R., Avendano, C., and Duda, R. O. (**2001**). "Elevation localization and head-related transfer function analysis at low frequencies," J. Acoust. Soc. Am. **109**, 1110–1122.

Asano, F., Suzuki, Y., and Sone, T. (**1990**). "Role of spectral cues in median plane localization," J. Acoust. Soc. Am. **88**, 159–168.

Başkent, D., and Shannon, R. V. (**2005**). "Interactions between cochlear implant electrode insertion depth and frequency-place mapping," J. Acoust. Soc. Am. **117**, 1405–1416.

Blauert, J. (**1969**). "Sound localization in the median plane," Acustica **22**, 205–213.

Blauert, J. (**1997**). *Spatial Hearing* (MIT, Cambridge, MA).

Carlyon, R. P., and Deeks, J. M. (**2002**). "Limitations on rate discrimination," J. Acoust. Soc. Am. **112**, 1009–1025.

Carlyon, R. P., Long, C. J., and Deeks, J. M. (**2008**). "Pulse-rate discrimination by cochlear-implant and normal-hearing listeners with and without binaural cues," J. Acoust. Soc. Am. **123**, 2276–2286.

Dorman, M. F., and Loizou, P. C. (**1997**). "Speech intelligibility as a function of the number of channels of stimulation for normal-hearing listeners and patients with cochlear implants," Am. J. Otol. **18**, S113–114.

Friesen, L. M., Shannon, R. V., Başkent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Am. **110**, 1150–1163.

Fu, Q. J., and Shannon, R. V. (**1999**). "Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing," J. Acoust. Soc. Am. **105**, 1889–1900.

Fu, Q. J., Shannon, R. V., and Wang, X. (**1998**). "Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing," J. Acoust. Soc. Am. **104**, 3586–3596.

Gabor, D. (**1947**). "Acoustical quanta and the theory of hearing," Nature (London) **84**, 591–594.

Gardner, M. B., and Gardner, R. S. (**1973**). "Problem of localization in the median plane: Effect of pinnae cavity occlusion," J. Acoust. Soc. Am. **53**, 400–408.

Goupell, M. J., Laback, B., Majdak, P., and Baumgartner, W. D. (**2008a**). "Current-level discrimination and spectral profile analysis in multichannel electrical stimulation," J. Acoust. Soc. Am. **124**, 3142–3157.

Goupell, M. J., Laback, B., Majdak, P., and Baumgartner, W. D. (**2008b**). "Effects of upper-frequency boundary and spectral warping on speech intelligibility in electrical stimulation," J. Acoust. Soc. Am. **123**, 2295–2309.

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Haynes, D. S., and Labadie, R. F. (**2008**). "Interaural time and level difference thresholds for acoustically presented signals in post-lingually deafened adults fitted with bilateral cochlear implants using CIS+ processing," Ear Hear. **29**, 33–44.

Hartmann, W. M., and Rakerd, B. (**1993**). "Auditory spectral discrimination and the localization of clicks in the sagittal plane," J. Acoust. Soc. Am. **94**, 2083–2092.

Hebrank, J., and Wright, D. (**1974**). "Spectral cues used in the localization of sound sources on the median plane," J. Acoust. Soc. Am. **56**, 1829–1834.

Henry, B. A., and Turner, C. W. (**2003**). "The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 2861–2873.

Henry, B. A., Turner, C. W., and Behrens, A. (**2005**). "Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners," J. Acoust. Soc. Am. **118**, 1111–1121.

Hofman, P. M., and Van Opstal, A. J. (**1998**). "Spectro-temporal factors in two-dimensional human sound localization," J. Acoust. Soc. Am. **103**, 2634–2648.

Hofman, P. M., Van Riswick, J. G. A., and Van Opstal, A. J. (**1998**). "Relearning sound localization with new ears," Nat. Neurosci. **1**, 417–421.

Iida, K., Itoh, M., Atsue, I., and Morimoto, M. (**2007**). "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," Appl. Acoust. **68**, 835–850.

Ketten, D. R., Skinner, M. W., Wang, G., Vannier, M. W., Gates, G. A., and Neely, J. G. (**1998**). "In vivo measures of cochlear length and insertion depth of Nucleus cochlear implant electrode arrays," Ann. Otol. Rhinol. Laryngol. Suppl. **175**, 1–16.

Kistler, D. J., and Wightman, F. L. (**1992**). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," J. Acoust. Soc. Am. **91**, 1637–1647.

Kulkarni, A., and Colburn, H. S. (**1998**). "Role of spectral detail in sound-source localization," Nature (London) **396**, 747–749.

Kulkarni, A., and Colburn, H. S. (**2004**). "Infinite-impulse-response models of the head-related transfer function," J. Acoust. Soc. Am. **115**, 1714–1728.

Laback, B., and Majdak, P. (**2008**). "Binaural jitter improves interaural time-

difference sensitivity of cochlear implantees at high pulse rates," Proc. Natl. Acad. Sci. U.S.A. **105**, 814–817.
Laback, B., Majdak, P., and Baumgartner, W. D. (**2007**). "Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing," J. Acoust. Soc. Am. **121**, 2182–2191.
Laback, B., Pok, S. M., Baumgartner, W. D., Deutsch, W. A., and Schmid, K. (**2004**). "Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors," Ear Hear. **25**, 488–500.
Langendijk, E. H. A., and Bronkhorst, A. W. (**2002**). "Contribution of spectral cues to human sound localization," J. Acoust. Soc. Am. **112**, 1583–1596.
Lawson, D. T., Wilson, B. S., Zerbi, M., van den Honert, C., Finley, C. C., Farmer, J. C., Jr., McElveen, J. T., Jr., and Roush, P. A. (**1998**). "Bilateral cochlear implants controlled by a single speech processor," Am. J. Otol. **19**, 758–761.
Long, C. J., Eddington, D. K., Colburn, H. S., and Rabinowitz, W. M. (**2003**). "Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user," J. Acoust. Soc. Am. **114**, 1565–1574.
Lu, T., Carroll, J., and Zeng, F. G. (**2007**). "On acoustic simulations of cochlear implants," in *Conference on Implantable Auditory Prostheses*, Lake Tahoe, CA.
Macpherson, E. A., and Middlebrooks, J. C. (**2000**). "Localization of brief sounds: Effects of level and background noise," J. Acoust. Soc. Am. **108**, 1834–1849.
Macpherson, E. A., and Middlebrooks, J. C. (**2003**). "Vertical-plane sound localization probed with ripple-spectrum noise," J. Acoust. Soc. Am. **114**, 430–445.
Majdak, P., Balazs, P., and Laback, B. (**2007**). "Multiple exponential sweep method for fast measurement of head-related transfer functions," J. Audio Eng. Soc. **55**, 623–637.
Majdak, P., Goupell, M. J., and Laback, B. (**2010**). "3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training," Attention, Perception, and Psychophysics **72**, 454–469.
Majdak, P., Laback, B., and Goupell, M. J. (**2008**). "3-D localization of virtual sound sources in normal-hearing and cochlear-implant listeners (A)," J. Acoust. Soc. Am. **123**, 3562.
Majdak, P., Laback, B., and Baumgartner, W. D. (**2006**). "Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing," J. Acoust. Soc. Am. **120**, 2190–2201.
Middlebrooks, J. C. (**1992**). "Narrow-band sound localization related to external ear acoustics," J. Acoust. Soc. Am. **92**, 2607–2624.
Middlebrooks, J. C. (**1999a**). "Individual differences in external-ear transfer functions reduced by scaling in frequency," J. Acoust. Soc. Am. **106**, 1480–1492.
Middlebrooks, J. C. (**1999b**). "Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency," J. Acoust. Soc. Am. **106**, 1493–1510.
Møller, H., Sørensen, M. F., Hammershøi, D., and Jensen, C. B. (**1995**). "Head-related transfer functions of human subjects," J. Audio Eng. Soc. **43**, 300–321.
Morimoto, M., and Aokata, H. (**1984**). "Localization cues of sound sources in the upper hemisphere," J. Acoust. Soc. Jpn. (E) **5**, 165–173.
Musicant, A. D., and Butler, R. A. (**1984**). "The influence of pinnae-based spectral cues on sound localization," J. Acoust. Soc. Am. **75**, 1195–1200.
Nelson, D. A., Donaldson, G. S., and Kreft, H. (**2008**). "Forward-masked spatial tuning curves in cochlear implant users," J. Acoust. Soc. Am. **123**, 1522–1543.
Nopp, P., Schleich, P., and D'Haese, P. (**2004**). "Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants," Ear Hear. **25**, 205–214.
Oldfield, S. R., and Parker, S. P. (**1984**). "Acuity of sound localisation: A topography of auditory space. II. Pinna cues absent," Perception **13**, 601–617.
Pralong, D., and Carlile, S. (**1996**). "The role of individualized headphone calibration for the generation of high fidelity virtual auditory space," J. Acoust. Soc. Am. **100**, 3785–3793.
Qian, J., and Eddins, D. A. (**2008**). "The role of spectral modulation cues in virtual sound localization," J. Acoust. Soc. Am. **123**, 302–314.
Schoen, F., Mueller, J., Helms, J., and Nopp, P. (**2005**). "Sound localization and sensitivity to interaural cues in bilateral users of the Med-El Combi 40/40+cochlear implant system," Otol. Neurotol. **26**, 429–437.
Seeber, B. U., Baumann, U., and Fastl, H. (**2004**). "Localization ability with bimodal hearing aids and bilateral cochlear implants," J. Acoust. Soc. Am. **116**, 1698–1709.
Seeber, B. U., and Fastl, H. (**2008**). "Localization cues with bilateral cochlear implants," J. Acoust. Soc. Am. **123**, 1030–1042.
Senova, M. A., McAnally, K. I., and Marin, R. L. (**2002**). "Localization of virtual sound as a function of head-related impulse response duration," J. Audio Eng. Soc. **50**, 57–66.
Shannon, R. V., Galvin, J. J., III, and Baskent, D. (**2001**). "Holes in hearing," J. Assoc. Res. Otolaryngol. **3**, 185–199.
Shaw, E. A. (**1974**). "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," J. Acoust. Soc. Am. **56**, 1848–1861.
van den Brink, W. A. C., and Houtgast, T. (**1990**). "Spectro-temporal integration in signal detection," J. Acoust. Soc. Am. **88**, 1703–1711.
van Hoesel, R. J. M. (**2004**). "Exploring the benefits of bilateral cochlear implants," Audiol. Neuro-Otol. **9**, 234–246.
van Hoesel, R. J. M. (**2007**). "Sensitivity to binaural timing in bilateral cochlear implant users," J. Acoust. Soc. Am. **121**, 2192–2206.
van Hoesel, R. J. M. (**2008**). "Observer weighting of level and timing cues in bilateral cochlear implant users," J. Acoust. Soc. Am. **124**, 3861–3872.
van Hoesel, R. J. M., Bohm, M., Pesch, J., Vandali, A., Battmer, R. D., and Lenarz, T. (**2008**). "Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies," J. Acoust. Soc. Am. **123**, 2249–2263.
van Hoesel, R. J. M., and Tyler, R. S. (**2003**). "Speech perception, localization, and lateralization with bilateral cochlear implants," J. Acoust. Soc. Am. **113**, 1617–1630.
van Schijndel, N. H., Houtgast, T., and Festen, J. M. (**1999**). "Intensity discrimination of Gaussian-windowed tones: Indications for the shape of the auditory frequency-time window," J. Acoust. Soc. Am. **105**, 3425–3435.
van Wieringen, A., Carlyon, R. P., Long, C. J., and Wouters, J. (**2003**). "Pitch of amplitude-modulated irregular-rate stimuli in acoustic and electric hearing," J. Acoust. Soc. Am. **114**, 1516–1528.
Vliegen, J., and Van Opstal, A. J. (**2004**). "The influence of duration and level on human sound localization," J. Acoust. Soc. Am. **115**, 1705–1713.
Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). "Localization using nonindividualized head-related transfer functions," J. Acoust. Soc. Am. **94**, 111–123.
Wightman, F. L., and Kistler, D. J. (**1989**). "Headphone simulation of free-field listening. II: Psychophysical validation," J. Acoust. Soc. Am. **85**, 868–878.
Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., and Tam, C. (**2006**). "Perceptual recalibration in human sound localization: Learning to remediate front-back reversals," J. Acoust. Soc. Am. **120**, 343–359.
Zahorik, P., Wightman, F. L., and Kistler, D. J. (**1995**). "On the discriminability of virtual and real sound sources," in *Proceedings of the ASSP (IEEE) Workshop on Applications of Signal Processing on Audio and Acoustics* (IEEE, New York).

Effects of a curved vocal tract with grid-generated tongue profile on low-order formants

Paul H. Milenkovic[a)]
Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Engineering Drive, Madison, Wisconsin 53706

Srikanth Yaddanapudi
CD-Adapco, No. 1222, 1801 Warner Ranch Drive, Round Rock, Texas 78664

Houri K. Vorperian
Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, Wisconsin 53705

Raymond D. Kent
Department of Communicative Disorders, University of Wisconsin-Madison, 1875 Willow Drive, Madison, Wisconsin 53706

(Received 4 February 2008; revised 22 October 2009; accepted 30 November 2009)

A hyperbolic grid-generation algorithm allows investigation of the effect of vocal-tract curvature on low-order formants. A smooth two-dimensional (2D) curve represents the combined lower lip, tongue, and anterior pharyngeal wall profile as displacements from the combined upper lip, palate, and posterior pharyngeal wall outline. The algorithm is able to generate tongue displacements beyond the local radius of strongly curved sections of the palate. The 2D grid, along with transverse profiles of the lip, oral-pharyngeal, and epilarynx regions, specifies a vocal conduit from which an effective area function may be determined using corrections to acoustic parameters resulting from duct curvature; the effective area function in turn determines formant frequencies through an acoustic transmission-line calculation. Results of the corrected transmission line are compared with a three-dimensional finite element model. The observed effects of the curved vocal tract on formants F1 and F2 are in order of importance, as follows: (1) reduction in midline distances owing to curvature of the palate and the bend joining the palate to the pharynx, (2) the curvature correction to areas and section lengths, and (3) adjustments to the palate-tongue distance required to produce smooth tongue shapes at large displacements from the palate.
 [DOI: 10.1121/1.3277214]

PACS number(s): 43.70.Aj, 43.70.Bk, 43.70.Dn [CHS] Pages: 1002–1013

I. INTRODUCTION

Recent studies[1–4] have investigated vocal-tract shape changes during child development using an articulatory model to compute the formant working space for different sizes and proportions of vocal tract. Acoustic resonator theory predicts that the formant working space scales linearly with the reciprocal of overall vocal tract length, and this is the dominant effect seen. The changes in vocal-tract proportions occurring during development are also hypothesized to affect the formant frequencies. The aim of the present study is to gain insights into the acoustic mechanisms by which variations in the outer vocal-tract outline apart from length changes could shift formant frequencies. To achieve this aim, the method of hyperbolic grid generation is adapted to determining the tongue shape in relation to a measured outline.

A common procedure in the child-development studies is to derive an outer vocal-tract outline, a rectilinear-radial grid pattern, and a set of tongue-displacement basis functions from a reference adult subject. The outer vocal-tract outline refers to the contiguous outlines of upper lip, palate, and posterior pharyngeal wall in the midsagittal plane. These tongue-displacement basis functions can be derived from analysis of x-ray images[5,6] or be mathematically determined.[7] Manual identification of vocal-tract landmarks determines the grid pattern, and it is helpful to have a reference subject without a strongly curved palate outline so that the radial grid lines remain nearly normal to the palate and allow a sufficient range of displacement of the tongue before the radial lines converge. The oral and pharyngeal regions of this model may be differentially scaled to represent a child vocal tract at different stages of development by proportioning the rectilinear-radial grid pattern.

This procedure not only changes the overall vocal-tract length as well as the proportions between oral and pharyngeal regions but it also scales the basis functions controlling tongue shape in proportion to the scale changes in the respective vocal-tract regions. Boe *et al.*[1] ascribed the acoustic effect of changing vocal-tract proportions primarily to the scaling of the basis functions. The goal of the current investigation is to determine changes in formant frequency intrin-

[a)]Author to whom correspondence should be addressed. Electronic mail: phmilenk@wisc.edu

sic to changes in the curvature and proportions of the outer vocal-tract outline apart from basis function scaling and other influences.

An investigation into mechanisms by which vocal-tract proportions affect the formant frequencies depends on a model for the acoustic properties of a curved vocal tract. Morse and Ingard[8] along with Sundberg *et al.*[9] assumed that acoustic propagation at low frequencies follows a wave mode characterized by the conformal grid characteristic of streamline flow:[10] the wavefronts are defined by curved lines of constant-pressure in two dimensions (2Ds), curved surfaces in three dimensions (3Ds), that intersect the vocal-tract walls at right angles.[8,9,11,12] The flow in an acoustic wave is a local back-and-forth disturbance, not a bulk streaming flow, but it is assumed for the stated conditions that the acoustic particle velocities are aligned with streamlines under steady potential flow. Pressure measurements of cast models of the vocal tract[13] along with a finite-element model[14] (FEM) confirm those assumptions for frequencies up to 4 kHz for the shapes studied, although another FEM study showed a departure from such a plane-wave propagation mode at 1.5 kHz in the case of the large front cavity for /r/.[15]

A technique called hyperbolic grid generation[16–19] is commonly used to subdivide a region bounded on one side in 2D or 3D into grid elements along curved paths for the purpose of solving a partial differential equation (PDE) on that region.[20] The PDE could be for heat transfer, fluid flow, or as is the case here, acoustic wave propagation; grid generation is not the solution of that PDE but rather a preparatory step. The generated grid also happens to be the solution to a hyperbolic PDE, one describing the propagation of a grid-generating wavefront. Differing from a finite-difference PDE solver that locates wavefronts by interpolating from solution values at fixed grid points, the hyperbolic method directly expresses the next location of the wavefront by placing grid points along vectors directed away from points on the current location of the wavefront. This so-called marching algorithm computes the grid in one pass, and is thus favored over other PDE-derived methods, such as those solving for the potential field and requiring iterative updates of the entire grid.

The hyperbolic algorithm from Henshaw[18] has been adapted to locate the combined lower lip, tongue, and anterior pharyngeal wall outline by marching in steps away from outer vocal tract outline in the 2D midsagittal plane. The total marching distance is determined by lip opening, the weighted sum of tongue-displacement basis functions, or a constant distance in the region superior to the vocal folds within the respective regions of the vocal tract. The grid lines between the outer and inner outlines are generated with each intermediate marching step. The automatic generation of the grid allows separating scaling of those basis functions from changes in the outline. The transverse grid lines generated by this process provide curved paths nearly perpendicular to the palate and tongue that more closely represent acoustic wavefronts. Substituting straight grid lines that intersect the palate and tongue at oblique angles is known to result in errors in computed formant frequencies.[21]

A dissipation term evens the distribution of transverse grid lines, and a marching-speed term retards the grid wavefront when grid lines perpendicular to the wavefront are diverging and advances that wavefront when the perpendicular grid lines are converging. These two smoothing terms prevents the formation of grid shocks—crossings of the grid lines that result in sharp bends or cusps in the tongue shape for tongue displacements that exceed the radius of the local palate curvature. This allows the representation of a wide range of human-subject palate outlines rather than restricting the articulatory model to reference outlines with broad curvature.

The hyperbolic algorithm generates representations of acoustic streamline and wavefront paths as a byproduct of constructing the tongue shape. Keefe and Benade[11] calculated from streamline and wavefront paths a correction for the effective acoustic lengths and cross-sectional areas for a transmission-line model of strongly curved ducts. Resonance frequency shifts may be analytically determined for simple geometries, thus helping to explain the acoustic mechanism by which those shifts occur. This curvature correction is compared to the 3D wave mode solution from Sondhi's[22] analytical treatment of a curved duct. For this geometry as well as more complex ones, comparison is made to a finite-element analysis of the wave modes.

Hyperbolic grid generation coupled with a mathematical model of articulatory basis functions thus generates smooth tongue shapes representative of vowels in human speech. Given a model for the vocal-tract transverse cross section, grid lines connecting the palate to the tongue provide cross-distances for calculating the area function. This area function may in turn be corrected for vocal-tract curvature based on the streamline paths, and the intrinsic effect of curvature on formant frequencies may in this manner be quantified apart from the influence of other factors.

II. METHODS

A. Outer vocal-tract outline

Outer vocal-tract outlines were determined for two subjects, an adult male and an adult female, whose computed tomography (CT) x-ray images with superimposed tracings are shown in Fig. 1. These subjects received CT-scans for medical reasons not known to affect vocal-tract shape, but each subject was at rest and not speaking. As such, these data to not provide cross-distances or area functions, but distances and area functions representative of speech are supplied from published magnetic resonance imaging (MRI) data as described in Secs. II B–II F.

Each outline was traced from the midsagittal slice of a CT study using procedures described by Vorperian *et al.*[23] with changes noted below. These subjects received CT-scans for medical reasons not known to affect vocal-tract shape, and as such, the images are for a non-speech resting condition. A marker is drawn to locate a point on the anterior surface of the incisors at the level of the inferior surface of the upper lip. From that marker, a line is drawn to the lingual aspect of the central incisors. The reasons for interpolating through the incisors are (1) the teeth have gaps that permit acoustic wave propagation and (2) the volume displaced by

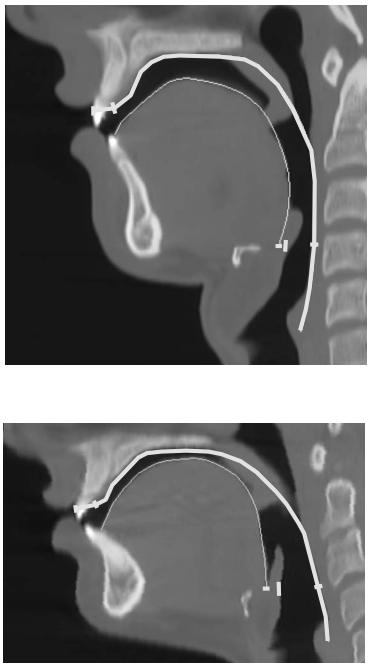


FIG. 1. Tracing of outer vocal-tract outline on CT-scan image.

the teeth represents a local perturbation in the area function affecting higher-order formants beyond those of interest.

A curvilinear path is continued by tracing along the length of the hard palate to the beginning of the soft palate. From there the curvilinear path crosses through the soft palate to the back of the pharyngeal wall down to the posterior aspect of the glottis. The path follows what is estimated to be the inferior surface of the soft palate when raised during vowel production. The demarcation between oral-pharyngeal and epilarynx regions was made with reference to the valleculae along with the superior surface of the hyoid. The epilarynx refers to the narrow conduit immediately superior to the vocal folds. The figure shows a mark for each subject denoting the separation of the anterior from the posterior boundaries of the epilarynx.

B. Determining area function from articulatory parameters

Articulatory parameters specify the displacement of the inner vocal-tract outline from the outer vocal-tract outline. A saturation function for that displacement, given in Eq. (3) below, prevents collision of the tongue with the palate. Displacements at points along the outer outline are in turn converted into areas using a model for the vocal-tract cross section.

A weighted sum of front-raising, back-raising, and jaw-opening basis functions controls the outer outline-tongue distance in the oral-pharyngeal portion of the vocal tract by

$$d_b(x) = d_{\text{avg}}(1 - w_{\text{fr}}h_{\text{fr}}(x) - w_{\text{br}}h_{\text{br}}(x) + w_{\text{jo}}h_{\text{jo}}(x)), \quad (1)$$

where average opening $d_{\text{avg}}=1.0$ cm, x is longitudinal distance along the central vocal-tract region normalized to $0 \leq x \leq 1$, the w coefficients vary with articulation, and setting tongue phase $\theta_t(x)=(1.75x+0.25)\pi$ and jaw phase $\theta_j(x)=(1.3x)\pi$ define the front-raising (fr), back-raising (br), and jaw-opening (jo) basis functions:

$$h_{\text{fr}} = \sin\theta_t, \quad h_{\text{br}} = -\cos\theta_t, \quad h_{\text{jo}} = 0.3 + 0.7\cos\theta_j. \quad (2)$$

Liljencrants[7] originally proposed the Fourier basis for the outline-tongue distance, but the first order Fourier terms produced shorter constriction regions than x-ray observations, requiring the addition of higher order Fourier terms. A stretching of the first order Fourier terms gives longer constriction regions. For example, basis function h_{fr} places the constriction of /i/ forward in the vocal-tract compared to the un-stretched Fourier basis function $\sin(2\pi x)$ while keeping the shape of the back cavity the same. A linear weighting of the front and back raising basis functions controls the place and depth of the tongue constriction after $A\cos x + B\sin x = \sqrt{A^2+B^2}\cos(x+\arctan(B/A))$.

The jaw-opening function has a longer wavelength than the front raising function and the combination of a positive jaw weight with a countervailing front raising weight produces the flared horn-like shape of /æ/. The jaw-opening function has a larger front opening relative to back narrowing; varying the jaw opening helps match the variation in the vocal-tract volume observed in MRI data, as will be discussed in Sec. II F.

A displacement saturation function imposes a minimum 0.15 cm vocal-tract opening according to

$$d_{\text{vt}} = 0.0505e^{4d_b} + 0.15, \quad d_b < 0.4,$$

$$d_{\text{vt}} = d_b, \quad d_b \geq 0.4, \quad (3)$$

where basis-function distance d_b and corrected vocal-tract distance d_{vt} are in centimeters. As the vocal tract is constricted, this saturation effect starts at a cross-distance of 0.4 cm, is continuous in both the function and its derivative between the linear and exponential portions of the curve, and has a limiting minimum distance of 0.15 cm for extreme negative undershoot of the value of d_b.

The longitudinal length of the lip region has been fixed at 1 cm for both the male and female subjects, and the lip-opening parameter controls the midsagittal cross-distance. The length and cross-distance of the epilarynx region are fixed by measurements of the CT image. The lip-vocal tract and vocal tract-epilarynx boundaries are given transition regions of 1 cm length.

The proposed transverse cross section for the central vocal tract is an inverted parabola with the apex at the palate, which is a simplified, flat-bottomed version of a profile attributed by Perrier *et al.*[24] to Maeda.[25] A parabola with palate-tongue distance 1 cm and width 3 cm is representative of CT transverse cross section images of the male subject, giving area from palate-tongue distance in cm^2 as

$$A = \alpha d_{\text{vt}}^{\beta}, \quad (4)$$

where $\alpha=2$ and $\beta=1.5$. This β is for a parabola whereas the α value is within reported ranges.[24,26] The model lip cross section is a pair of parabolas with bases meeting in the middle of the lip opening, scaled to give the same formula for area in terms of total lip opening to reduce interpolation artifact. The epilarynx cross section is an ellipse where the

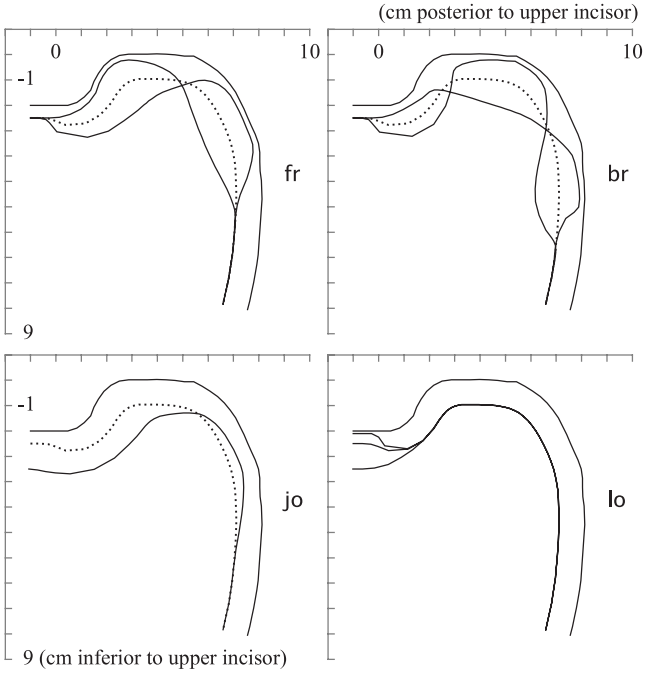


FIG. 2. Variation of the front-raising, back-raising, jaw-opening, and lip-opening basis-function weights for the male vocal-tract outline viewed in the midsagittal plane.

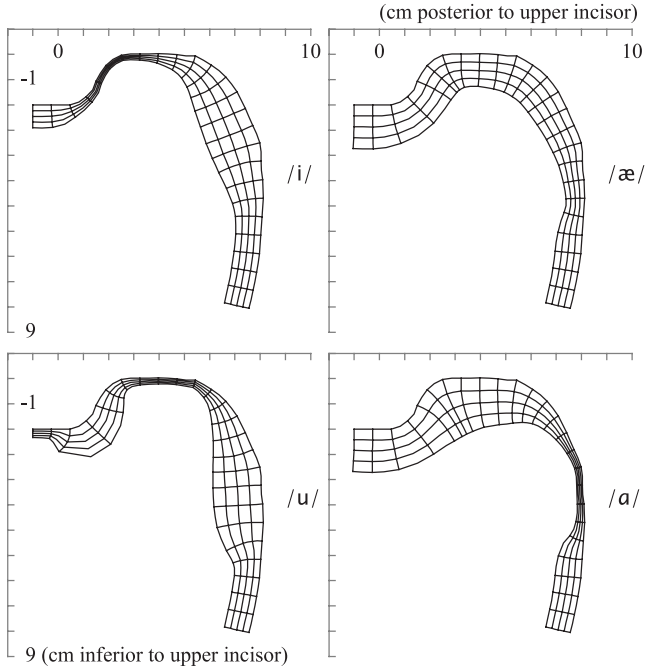


FIG. 3. Midsagittal vocal-tract shapes of vowels for the male vocal-tract outline generated by matching the articulatory model area function to MRI area functions from Story *et al.* (Ref. 35).

major axis is the anterior-posterior distance of the opening in that region and the minor axis is half that distance, giving area

$$A_{\text{epi}} = \frac{\pi}{8} d^2. \tag{5}$$

Cross-distance in the transition regions is determined by

$$d_{\text{trans}} = d_{\text{vt}} \cos \frac{\pi x_t}{2} + d\left(1 - \cos \frac{\pi x_t}{2}\right), \tag{6}$$

where x_t ranges from 1 to 0 in the lip-vocal tract transition and from 0 to 1 in the vocal tract-epilarynx transitions so as to bias the distances toward the values at the ends of the vocal tract. Area is calculated in the transition region by linear blending of area formulas for neighboring regions.

The tongue contour changes after adjusting individual basis-function weights are shown in Fig. 2 for the male vocal-tract outline. The tongue displacement is along a grid generated by the hyperbolic method presented in Sec. II C.

C. Hyperbolic grid generation

Grid $\mathbf{x}(s,t)=(x_x(s,t),x_y(s,t))$ denotes a collection of x-y points in the midsagittal plane. Integer index $0 \le s \le N$ selects placement along the outer vocal-tract outline ($s=0$ at the lips, $s=N=50$ at the glottis for a 50-section vocal tract), and $0 \le t \le 1$ denotes normalized distance between the outer and inner vocal-tract outlines. Curved paths connecting x-y points of constant t in this grid represent streamlines; paths of constant s represent acoustic wavefronts. The grid function $\mathbf{x}(s,t)$ thus maps a rectilinear grid in t-s space into a curved grid in x-y space; variable t also acts like time in the progression of the grid-generating wavefront from outer to inner vocal-tract outlines, propagating in a direction transverse to the acoustic wavefront. Figure 3 shows generated grids of cardinal vowels: only every other grid line is plotted for clarity.

The partial derivative of grid $\mathbf{x}(s,t)$ with respect to parameter t is given by

$$\mathbf{x}_t(s,t) = S(s,t)\mathbf{n}(s,t) + \varepsilon_d(\mathbf{x}_{ss}(s,t) - (\mathbf{x}_{ss}(s,t) \cdot \mathbf{n}(s,t))\mathbf{n}(s,t)), \tag{7}$$

where $S(s,t)$ is the local speed of the grid-generating wavefront, wavefront unit normal vectors $\mathbf{n}(s,t)$ are computed from tangent slopes derived from a fit of piecewise circular arcs to points on that wavefront,[27–29] second central difference $\mathbf{x}_{ss}(s,t)=\mathbf{x}(s-1,t)-2\mathbf{x}(s,t)+\mathbf{x}(s+1,t)$, and term $\mathbf{x}_{ss}(s,t)-(\mathbf{x}_{ss}(s,t)\cdot\mathbf{n}(s,t))\mathbf{n}(s,t)$ denotes that part of x-y vector $\mathbf{x}_{ss}(s,t)$ normal to $\mathbf{n}(s,t)$. A dissipation coefficient is scaled in relation to the time step according to $\varepsilon_d=0.4dt$, smoothing the distribution of points along the wavefront path to suppress crowding of the grid lines intersecting the wavefront by allowing those lines to deviate from being precisely normal to that wavefront.

Equation (7) is solved using the second-order Runge–Kutta numerical integration of Petersson with more steps[19] in place of the implicit integration of Henshaw[18] having more calculations per step. The integration divides the displacement between palate and tongue into steps dt, where $dt=1$ would apply the entire displacement from palate to tongue in one step. Employing a multi-step marching algorithm in place of a single-step procedure avoids crossing of grid lines, a problem when the palate-to-tongue distance exceeds the local radius of the curved palate. In addition to the aforementioned dissipation term, adjustments to the marching speed S can suppress crossing of the grid lines while providing control over the palate-tongue distance. Whereas the speed ad-

justments used to prevent grid-line intersection also prevent an exact match to a target palate-tongue distance, the following formula keeps the effect of distance mismatch small relative to other effects being investigated:

$$S(s,t) = d(s,t)\frac{\overline{\Delta a}}{\Delta a}(1 - \varepsilon_c(t)d(s,t)\kappa(s,t)). \tag{8}$$

The terms of this equation are as follows, starting with $d(s,t)$ computed for each marching step from an initial palate-tongue distance $d(s,0)$, supplied by the weighted vocal- tract basis functions, using the update

$$d(s,t_k) = d(s,t_{k-1})\left(1 + \frac{\alpha}{8}\right) - \alpha|\mathbf{x}(s,t_k) - \mathbf{x}(s,t_{k-1})|, \tag{9}$$

which compensates for changes in the palate-tongue distance brought about by the smoothing terms of $S(s,t)$, where $t_k = k/8$ and $\alpha = (1-\sqrt{t_{k-1}})/(1-t_k)$ for $0<k<8$. Choice of α controls the trade-off between a smooth tongue surface and a precise match to the basis function-determined palate-tongue distance.

Defining backward and forward differences $\underline{\mathbf{x}_{s-}} = \mathbf{x}(s,t) - \mathbf{x}(s-1,t)$ and $\mathbf{x}_{s+} = \mathbf{x}(s,t) - \mathbf{x}(s+1,t)$, speed term $\overline{\Delta a}/\Delta a$ adjusts marching distances to smooth out local values of wavefront arc length $\Delta a = 0.5(|\mathbf{x}_{s-}| + |\mathbf{x}_{s+}|)$ that deviate from averaged arc length $\overline{\Delta a}$ where the averaging is performed by eight applications of the difference equation,

$$\Delta a_1(s) = 0.8375\Delta a(s) + 0.08125(\Delta a(s-1) + \Delta a(s+1)). \tag{10}$$

Finally, speed term $1 - \varepsilon_c(t)d(s,t)\kappa(s,t)$ adjusts the marching steps to smooth local wavefront curvature

$$\kappa(s,t) = \frac{\mathbf{n}\cdot\mathbf{x}_{s-}}{|\mathbf{x}_{s-}|^2} + \frac{\mathbf{n}\cdot\mathbf{x}_{s+}}{|\mathbf{x}_{s+}|^2}. \tag{11}$$

Curvature is the reciprocal of the local radius of the wavefront; the product of curvature with the palate-tongue distance $d(s,t)$ is thus dimensionless as required for the formula. The coefficient of the curvature adjustment

$$\varepsilon_c(t) = 0.183\sqrt{t(1-t)} \tag{12}$$

has a maximum value midway between palate and tongue and minimum values at the ends to give maximum smoothing of the streamline paths midway between these two boundaries. This amount of curvature adjustment prevented the intersection of palate-tongue grid lines for the vocal-tract outlines and range of tongue displacements under consideration.

Too large a step size dt relative to the wavefront arc length Δa results in the growth of local disturbances. A bound on the step size that suppresses such disturbances is expressed as

$$dt = \min_s\left\{\left(\frac{1}{2\varepsilon_c}\right)\left(\frac{|\mathbf{x}_{s-}|^2 + |\mathbf{x}_{s+}|^2}{2d(s,t)^2}\right)\left(\Delta a/\overline{\Delta a}\right)\right\}. \tag{13}$$

D. Curvature correction of acoustic-tube lengths and areas

The vocal tract in vowel production is modeled by a non-uniform acoustic tube, where the cross sectional area is a function of distance along a curved path from lips to glottis. A curvature correction for a transmission-line analog after the method of Keefe and Benade[11] follows and is compared with the 3D wave mode solution of Sondhi and a 3D FEM solution.

A curvature correction to the lengths and areas of elements of the non-uniform acoustic tube may be understood by considering that a volume element of the duct that has length l in the direction of fluid motion, cross-sectional area A, and nearly constant-pressure and flow within that volume has acoustic inductance:

$$L = \frac{\rho l}{A}, \tag{14}$$

where ρ is the density of air and has capacitance

$$C = \frac{Al}{\rho c^2}, \tag{15}$$

and where c is the speed of sound. In turn, an effective acoustic length l and area A may be determined from specified values of L and C according to

$$l = c\sqrt{LC}, \quad A = \rho c\sqrt{\frac{C}{L}}. \tag{16}$$

A short section of a duct between a pair of constant-pressure surfaces may be divided into longitudinal sections representing stream tubes. Figure 3 of Keefe and Benade[11] illustrates a duct section and one of these stream tubes. The bundle of stream tubes sees a constant-pressure at each face, and the difference in pressure between the two faces acts on the fluid inertia to change the rate of flow in each stream tube. From circuit theory, the capacitances of these parallel circuit branches add, and the reciprocal inductance is the sum of the reciprocal inductances. The branch capacitance and inductance can be determined from the average area A and average length l of each stream tube, the capacitances and inductances of those branches can be combined, and the effective acoustic area A and length l of the duct section can be solved from the combined L and C values.

For a curved duct of constant cross section and bend radius, the constant-pressure surfaces are assumed to be radial slices, and the streamlines to follow circular paths between the inner and outer bend radius of the duct. The spacing of the streamlines between the inner and outer bend radius is not required; the only required assumption is that in a duct with a circular bend, geometric symmetry has the streamlines following circular paths. Wave mode solutions[22,30] show pressure variation in the radial direction, even for the plane-wave mode. Keefe and Benade[11] considered radial pressure variation of the form expected at low frequencies and did not find a significant change from constant-pressure. Keefe and Benade,[11] however, found that resonant frequencies of a bent tube measured in the laboratory changed in the same direction but in a lower amount

TABLE I. Formant frequencies in Hz for a curved vocal tract of uniform cross section, rigid walls, and zero wall and radiation loss. Curvature corrected: adjusts area and length by combining stream tubes; modal: acoustic wave mode solution of Sondhi (Ref. 22), FEM: 3D finite-element analysis.

	Rectangular cross section			Parabolic cross section	
	Curvature corrected	Modal	FEM	Curvature corrected	FEM
F1	502.9	502.7	503.0	521.8	522.0
F2	1508.7	1504.4	1505.4	1565.3	1563.3
F3	2514.6	2494.9	2497.6	2608.9	2596.9

than the prediction from integration over stream tubes. Nederveen[12] attributed this difference to small errors in measuring tube diameters in brasswind musical instruments; controlling for tube diameter gave close agreement between acoustic measurements and stream tube theory.

The curvature correction to the transmission-line analog may be compared to the 3D wave mode solution of the acoustic wave equation $\nabla^2 p=(1/c^2)(\partial^2 p/\partial t^2)$, where p is time and spatially varying pressure and c is the speed of sound. Separation of the time variable results in the Helmholtz equation $\nabla^2 P=-(\omega^2/c^2)P$, where P is the spatially varying coefficient of the sinusoidal time dependence of pressure. Sondhi[22] solved this equation analytically for a curved duct with hard walls, a radiation model of a simple acoustic short circuit, a constant bend radius, and a rectangular cross section.

The finite-element approximation to the wave mode is the solution to the matrix equation $KP=\lambda MP$.[31] Acoustic pressure in this approximation is assumed to vary linearly within tetrahedral volume elements between node values at the corners of those elements. For a lossless duct, this matrix equation is a generalized eigenvalue problem where matrices K and M are symmetric and positive definite, giving real-valued solutions for scalar $\lambda=\omega^2/c^2$ relating to mode frequency and vector P of node pressures relating to mode shape. A JAVA program calculates coefficients of the matrices K and M from the coordinates of the corners of the finite elements using the formulas from Huebner,[31] and the matrix equation is solved using the DSBGV double-precision symmetric-banded generalized eigenvalue, function from the JLAPACK software library.

Table I presents results for a curved duct with a 17 cm midline length, z-axis width 2.5 cm, forming a 90° arc having cross-distance of 4 cm and bend radii $r_1=8.82$ cm, $r_2=12.82$ cm. Using Sondhi's value for sound speed of 34 000 cm/s,[22] an equivalent straight duct has formants F1, F2, and F3 at 500, 1500, and 2500 Hz. Table I reports results for the preceding curvature correction, Sondhi's wave mode analysis, and a 3D FEM. The duct is assumed throughout to have rigid walls along with zero wall and radiation loss.

For constant-pressure across radial slices, integration of a duct of rectangular cross section and inner radius r_1 and outer radius r_2 determines an effective inductance,

$$L_{\text{eff}}=\frac{r_2-r_1}{\ln(r_2/r_1)(r_2+r_1)/2}L, \tag{17}$$

where L is the uncorrected inductance calculated from the midline distance; the capacitance C remains unchanged, and effective tube length and area may be calculated using Eq. (16). Table I also includes results for a parabolic cross section of height 4 cm and width 3.75 cm, having the same area as the rectangular cross section. The integrations for the curvature correction for this parabolic cross section were carried out with the MAPLE symbolic solver software package.

This finite-element analysis divides the duct into 50 hexahedral (six-sided box-like) sections along the longitudinal axis, eight sections in the radial direction, and one section along the z-axis normal to the plane of the bend. Increasing the number of hexahedra in the z-direction to 4 resulted in a formant shift of less than 0.3 Hz. Each hexahedral section is subdivided into six tetrahedral elements in such a way that neighboring pairs of tetrahedra share a face without overlap into a third tetrahedron: this condition ensures that pressure is piecewise linear throughout the entire duct volume without any step changes. Increase in the number of sections along the longitudinal and radial directions lowered the formant frequencies for the FEM in the direction of Sondhi's analytical modal results.

The curvature correction to the transmission-line model predicts a linear scaling of formant shift with increasing formant frequency. The modal and FEM values agree with the curvature correction at F1, but those values vary downward from linear scaling for higher formants. This decline may be related to the influence of the cross modes on the lower plane-wave mode resonances.[32] Sondhi's tables show the amount of decline to be reduced for a smaller bend radius difference, a change that raises the frequency of the first cross mode. The decline results in a net downward shift for F3 for the rectangular cross section not predicted by the curvature correction, but this effect is masked by a much larger upward shift in formants intrinsic to the curvature correction for a parabolic cross section

Applying the curvature correction to the area function of more complex vocal-tract shapes representative of vowel production requires numerical methods. A hyperbolic grid, combined with the simplified version of Maeda's parabolic model for the transverse cross section, divides the vocal tract into stream tube sections. The span between palate and tongue is once more divided into eight stream tubes, and the longitudinal axis from lips to glottis is divided into 50 sections. The cross section of each end of a stream tube section is approximated with a trapezoid of the same area. The distance between centroids of the trapezoid at each face determines the section length. The volume of each stream tube is determined by adding the volumes of a spatial tiling of tetrahedra as done for the FEM, and area of a stream tube section is calculated as volume divided by length.

The vocal tract is represented as a series of uniform-tube sections, and section values for $1/L$ and C, and hence the effective length and effective area, are determined by summing over $1/L$ and C values, respectively, for the discrete stream tubes marked by the grid. Numerical determination of

formant frequencies from a vocal tract represented by a series connection of uniform tubes of specified length and cross-sectional area is addressed in Sec. II E.

E. Determination of formant frequencies from area function

The method for calculating the acoustic frequency response of a series connection of acoustic tubes is based on the frequency-domain method of Atal *et al.*,[33] which takes into account the effects of viscous and heat conduction losses within the vocal-tract along with the yielding vocal-tract sidewall. The frequency-dependent lip radiation load is computed using a formula for a circular tube terminating in an infinite-plane baffle.[34]

Formant frequencies may be estimated by evaluating vocal-tract transfer function magnitude $|H|$ at different values of ω to determine peaks. This commonly used method encounters the problem of merged peaks, occurring when the spacing of a pair of formants is small relative to their bandwidths. Atal *et al.*[33] addressed this by evaluating $|H|$ at complex values away from the imaginary axis $s=j\omega$ and by finding zeroes of $1/H$ by a Newton–Raphson search procedure. An alternative formant-finding method follows from $1/H$ being real-valued in the limiting case of a lossless vocal tract. For a lossy vocal tract, the real part of $1/H$ is evaluated for zero crossings, which are empirically determined to be close to the frequencies of peaks of H but provide better resolution of closely spaced formants. Once a zero crossing is found by evaluating $1/H$ at discrete steps $\Delta\omega=2\pi 40$ Hz, interval bisection and linear interpolation refine the location of that zero.

The frequency-domain method was evaluated using MRI-derived area functions from Story *et al.*[35] That study computed formant frequencies by a different method, one employing a wave-analog synthesizer incorporating a side branch for the piriform sinuses. In place of this side branch, the frequency-domain method was modified by adding volume to the acoustic tube section immediately superior to the epilarynx. Dang and Honda[36] reported total piriform volume for an adult subject in excess of 2 cc. A frequency-domain calculation using a piriform volume correction of 1.25 cc, wall mass coefficient of 1 gm/cm^3, and sound speed of 35 000 cm/s gave the closest match (ΔF2 of /u/ $<$60 Hz, ΔF all others $<$20 Hz) to formant F1-F2 values reported by Story *et al.*[35]

F. Matching model to MRI area functions

Weights for lip opening along with three tongue basis functions were manually adjusted to obtain a best compromise between matching model-generated to the MRI-measured area functions from Story *et al.*[35] and matching the first two formant frequencies observed as peaks of the frequency response. This procedure determined basis function weights representative of the vowels /a/, /æ/, /i/, and /u/, obtaining a single set of articulatory parameter values to substitute into different vocal-tract profiles. These matches were obtained using a computer-generated display of the midsagittal plane vocal-tract profile, the area function derived from that profile, and the acoustic frequency response. Acoustic-tube lengths were determined from the longitudinal grid midline, without the curvature correction, to best represent the way the target MRI area functions had been measured.

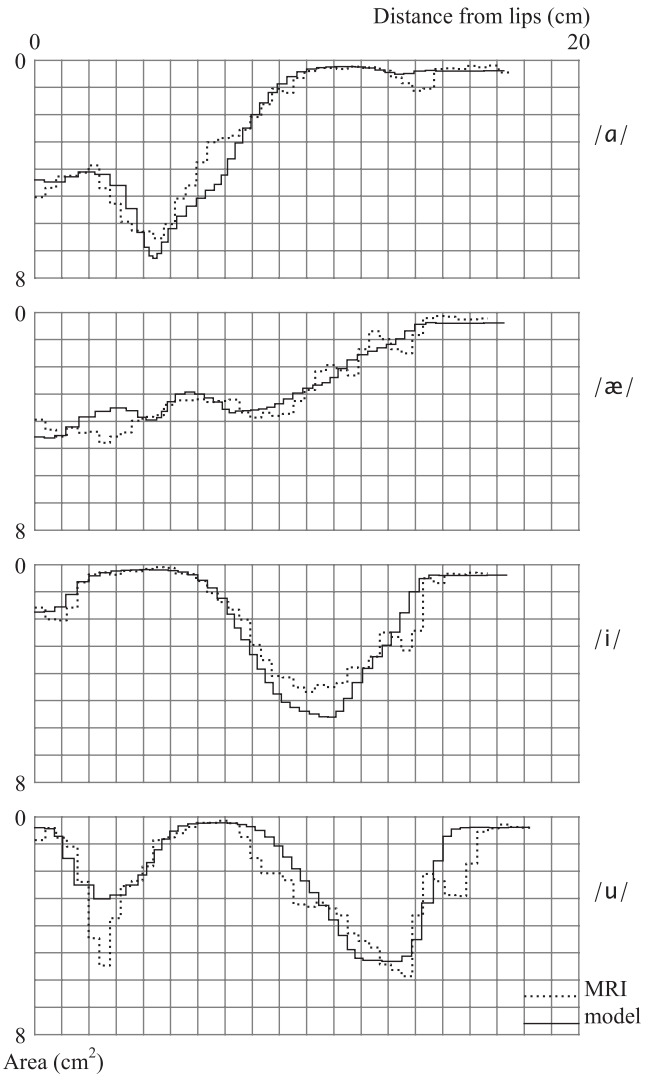


FIG. 4. Results of matching articulatory model area functions to MRI area functions from Story *et al.* (Ref. 35).

The resulting midsagittal outlines are shown in Fig. 3; Fig. 4 shows the match between area functions. These results proved to be sensitive to the length of the epilarynx tube. The male subject epilarynx tube was shortened by 0.35 cm from the CT-scan-derived area functions shown in Fig. 4 prior to calculating formants. This adjustment resulted in formant frequencies between model-generated and MRI area functions that were within ΔF1 of /a/ $<$25 Hz, ΔF all others $<$10 Hz.

The effects of varying the epilarynx length and adding piriform sinus volume are seen in Fig. 5. The reference condition is for the epilarynx length measured from the CT-scan without piriform sinuses. Shortening the epilarynx in the amount of 0.35 cm, the condition in the match to the MRI data, raised *F*2 for /i/. Keeping the measured epilarynx length and instead adding a 1.25 cc piriform volume lowered F1 and F2 for /a/ and /æ/.

MRI (Refs. 37 and 38) and CT (Ref. 24) studies of the

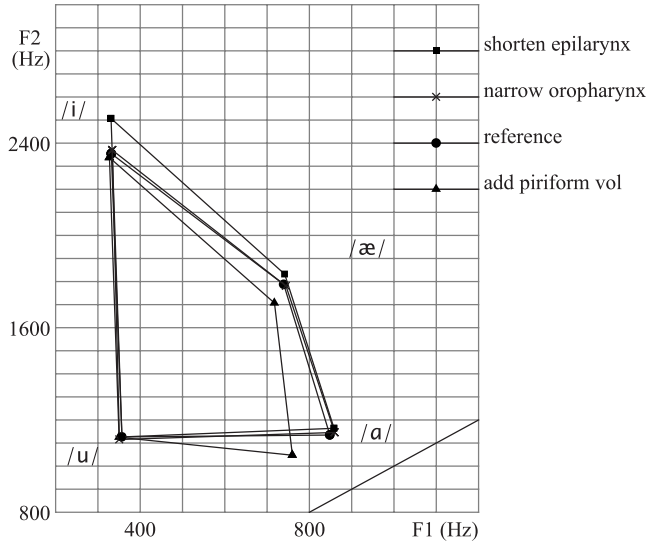


FIG. 5. Formant sensitivity to changes in the lower vocal tract: the reference condition is for the male outline with no piriform sinuses, the shortened epilarynx condition is for a length reduction of 0.35 cm, and the added piriform volume condition adds 1.25 cc to the piriform sinuses with the epilarynx length restored to the reference.

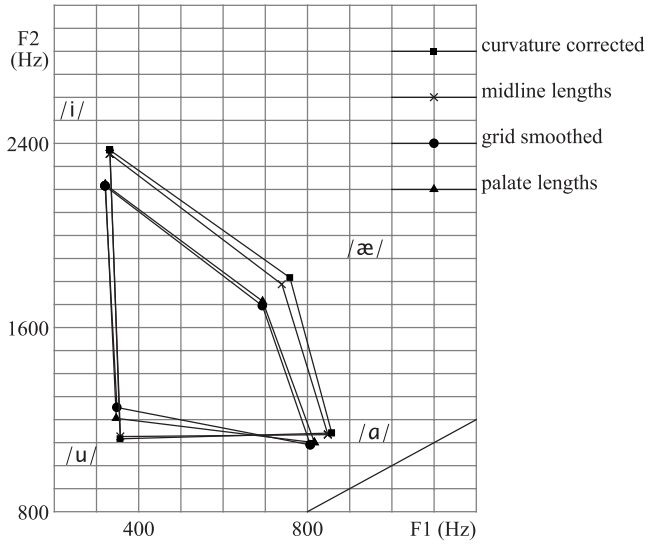


FIG. 6. Formant frequencies determined for the male vocal-tract outline using four methods of determining acoustic-tube lengths and areas.

coefficients of the αd^{β} model for computing area from cross-distance d show variations in these coefficients with place in the vocal tract. Two of these studies [24,38] show a consistent z-axis narrowing of the vocal tract in the oropharynx, related to a reduction in α or in β. The "narrow oropharynx" condition in Fig. 5 relates to a raised cosine function spanning 4 cm in length, located midway in the central vocal-tract region, scaled to give a peak reduction in α of 20%. The deviation from the reference condition resulting from this perturbation is small relative to the effects of variation in the epilarynx length and piriform volume.

III. RESULTS

A. Comparison of methods for determining area function

The effect of the curved vocal tract on formant frequencies is evaluated by comparing four methods for computing the vocal tract area function from the basis function model. This comparison is made on the male-subject outline with the CT-scan measured epilarynx length, using the basis function weights for the /a/, /æ/, /i/, and /u/ obtained by matching MRI data; the results are shown in Figs. 6–8. Numerical values of formants from the four methods are presented in Table II.

The "palate lengths" method refers to an unrolled vocal tract where the weighted combination of basis functions directly specifies palate-tongue cross-distances used to compute acoustic-tube area, where acoustic-tube lengths are determined with reference to distances along the outer vocal-tract outline. The "grid smoothed" method also uses palate-referenced acoustic-tube lengths, but the cross-distances are taken from the generated grids shown in Fig. 3. The "midline lengths" method uses the smoothed grid cross-distances for area, but distances along the grid midline extending from lips to vocal folds give acoustic-tube lengths. Finally, the "curvature corrected" method makes adjustments to acoustic-tube areas and lengths by combining stream tubes.

The formant shifts brought about by grid smoothing are small: this is a result of the choice of grid smoothing coefficients along with the adjustments made to grid marching speed to compensate for changes in cross-distance brought about by grid smoothing. Some shift in the cross-distances needs to be tolerated to prevent crossing of grid lines and the resulting sharp cusps in the tongue, but the grid-generation algorithm has been tuned so that the computed formant shifts that are small relative to the other effects under consideration.

Switching the acoustic-tube lengths from distances along the palate to the grid-generated midline brought about the greatest change in formants. Formant frequencies increased as a result of shorter midline distances and hence a shorter overall vocal-tract length relative to the palate distances. The curvature correction resulted in a small addi-

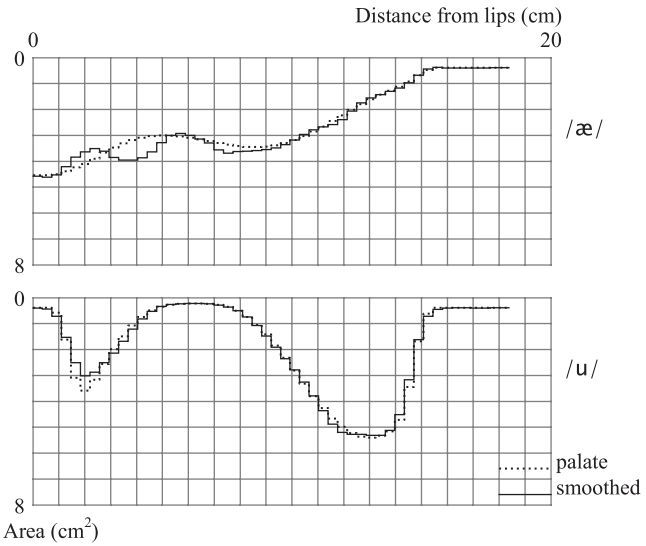


FIG. 7. Change to area function resulting from grid smoothing.

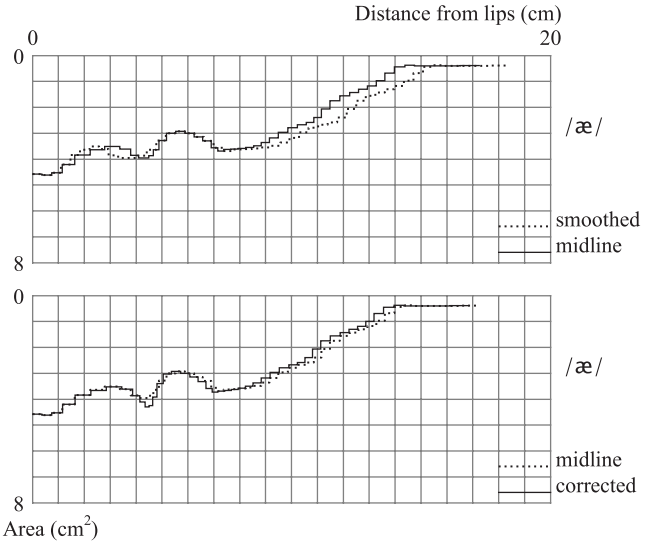


FIG. 8. Change to area function resulting from midline-distance correction, curvature correction.

tional increase in frequency relative to that midline-distance effect, an amount consistent with the analysis of simple ducts in an earlier section.

Consequences of the interaction of grid smoothing with vocal-tract curvature on the area function are seen in Fig. 7 for /æ/ and /u/. The vowel /æ/ has the simplest geometry of a horn-like flared tube while /u/ has the most complex geometry, with front and back cavities separated by a central constriction. Grid smoothing keeps the same tube lengths while producing area changes, which are greatest in the front part of the vocal tract having the reverse curve between the convex alveolar region and the concave palate vault. The area changes were larger for /æ/, where the grid algorithm had to march a greater distance from the palate, but the formant shifts were greater for /u/, where F2 is sensitive to changes in the front cavity.

The effect of changing from outer vocal-tract outline (palate) to midline-referenced acoustic-tube lengths for the grid-smoothed area function followed by changing to the curvature correction for /æ/ are seen in Fig. 8. The switch from palate to midline distances keeps the same acoustic-tube areas but changes tube lengths in proportion to the midline stretching in convex regions and shrinking in concave regions. The curvature correction results in changes to vocal-tract area, but the changes in length are small compared to the palate-to-midline shift, confirming the effect seen in simple ducts.

Table III compares the curvature correction of the transmission-line model with FEM analysis of the wave modes, where the hyperbolic grid gives the hexahedral sections. Values are reported as the change in formant frequency from the reference condition where acoustic-tube lengths in the transmission-line model come from midline lengths. This comparison is conducted for a lossless vocal tract; as a consequence, the values for the reference midline condition differ from those reported in Table II.

The FE1 condition uses a single layer of hexahedral sections in the z-direction normal to the midsagittal plane; the FE4 condition uses four layers. The curvature correction is in agreement with the FE1 condition to within 2 Hz for F1, 6 Hz for F2, and 9 Hz for F3. The FE4 condition results in much larger change from FE1 than the 0.3 Hz noted for the curved duct of uniform cross section. These shifts may relate to the convergence and divergence of the vocal tract in the z-direction resulting from variation in cross-sectional area along with the change in transverse profile between the elliptic epilarynx tube and the parabolic oral-pharynx region. Such variation appears to require consideration of the 3D compound curvature of the acoustic wavefront; the proposed curvature correction along with the FE1 condition only account for the curvature in the midsagittal plane.

The FE4 condition involves solution of an eigenvalue problem for 2295 node variables, a computation taking 64 s on a 1.2 GHz Pentium III processor. The transmission-line calculation of frequency response takes 20 ms, about a factor 3000 increase in speed, which becomes important when large numbers of vocal-tract shapes need to be computed.

TABLE II. Formant frequencies in Hz for the male vocal-tract outline using the four methods of determining acoustic-tube lengths and areas reported in Fig. 6.

		Palate lengths	Grid smoothed	Midline lengths	Curvature corrected
/a/	F1	817.3	806.8	849.1	857.8
	F2	1099.4	1090.2	1134.1	1141.1
	F3	2627.9	2624.6	2956.4	3029.0
/æ/	F1	694.2	693.6	740.1	758.6
	F2	1713.5	1694.0	1786.7	1815.8
	F3	2312.6	2298.7	2414.5	2461.1
/i/	F1	322.0	320.8	332.1	332.6
	F2	2220.3	2215.0	2353.1	2370.6
	F3	2879.8	2873.3	3109.5	3141.6
/u/	F1	347.3	349.1	357.3	356.5
	F2	1205.1	1251.4	1125.7	1115.7
	F3	2542.3	2572.6	2701.5	2720.4

TABLE III. Formant frequencies in Hz for the male vocal tract outline computed under lossless condition for comparison with the finite-element method. Comparisons are given as frequency differences with the midline lengths condition. Curvature corrected: adjusts area and length by combining stream tubes; FE1: finite-element method with one hexahedral (box) section in the z-direction; FE4: four hexahedral sections in the z-direction.

		Frequency	Frequency difference from midline lengths		
		Midline lengths	Curvature corrected	FE1	FE4
/a/	F1	829.7	6.9	6.8	3.6
	F2	1154.7	12.7	12.9	9.3
	F3	3082.5	69.6	72.6	52.5
/æ/	F1	729.9	23.4	22.9	21.5
	F2	1832.9	29.7	30.1	16.1
	F3	2521.1	51.8	50.8	35.2
/i/	F1	223.9	1.2	2.9	2.3
	F2	2337.1	17.3	18.3	−18.6
	F3	3155.6	38.8	47.5	27.5
/u/	F1	260.4	−0.3	1.4	0.3
	F2	1209.9	−12.8	−6.8	−12.6
	F3	2688.3	19.6	23.6	−17.9

B. Effects of oral-pharyngeal length ratio and bend radius

Results for changing the outer vocal-tract outline while keeping the same basis functions and weights to generate the tongue shape are presented in Figs. 9 and 10. Four vocal-tract outlines are compared. To control for vocal-tract length effects, these outlines are all normalized to have the male subject outer vocal-tract outline length, also matching lip length along with length and cross-distance of the epilarynx tube. The normalized female profile in Fig. 9 is the female profile from Fig. 1 after applying this process.

The effects of reduced bend radius are evaluated using stylized or simplified vocal tracts that have a convex alveolar surface and a concave palate vault of radius 1 cm, a 60° incline joining the alveolar and palate regions, a 2-cm rise to the top of the palate vault, and a 2-cm radius of the curve joining the oral and pharyngeal sections. The oral length is from anterior surface of the incisors to the posterior pharyngeal wall; the pharyngeal length is from the superior extent of the palate to the superior end of the epilarynx. The simplified female outline has an oral over pharyngeal length

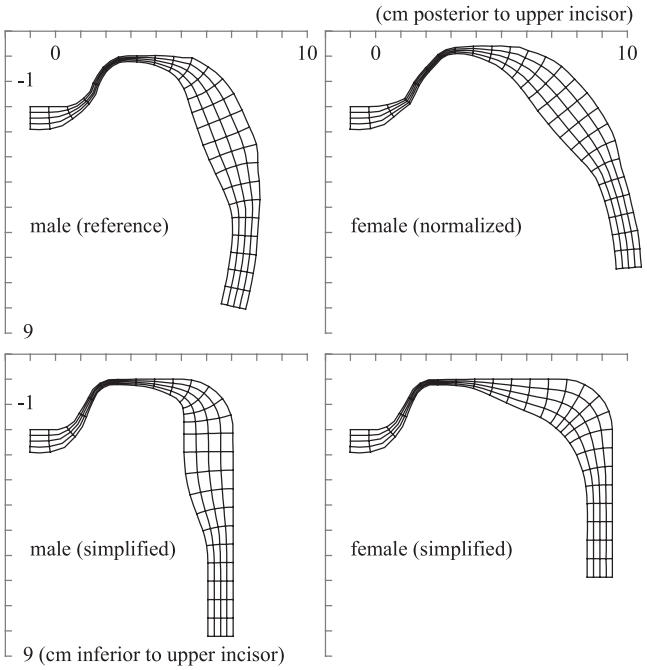


FIG. 9. Alternative vocal-tract shapes for vowel /i/ for different outer vocal-tract outlines. The natural male subject outline is the reference to which the other outlines are length normalized. The simplified shapes are stylized representations of the male and female shapes.

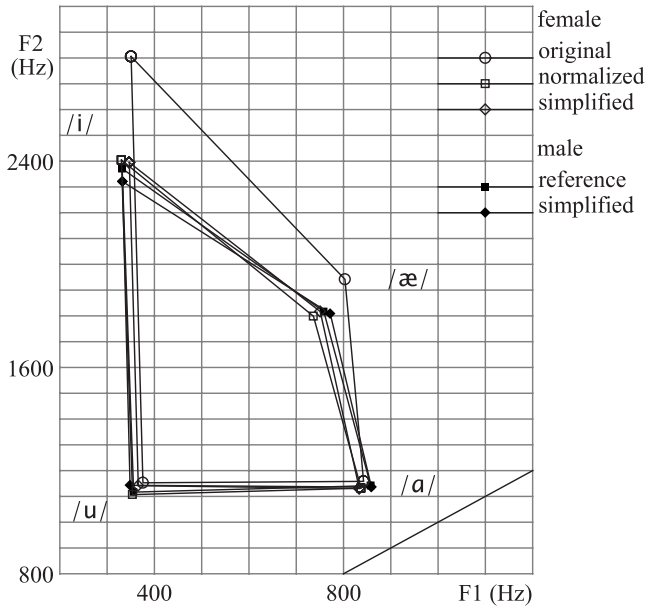


FIG. 10. Comparing model-derived formant frequencies from female outer vocal-tract outline with four outlines (female and male, natural and simplified) that are length normalized to the male outline.

ratio of 2.0 whereas the simplified male has a ratio of 1.0. Taking into account differences in degree and distribution of outline curvature, these simplified vocal tracts have similar proportions to the natural male and female outlines.

Figure 10 offers a comparison between formants computed using the curvature-correction method for the female subject before length normalization and the four length-normalized outlines. The four length-normalized outlines give tightly clustered formants compared to the male-female difference, confirming that length normalization can remove most of the difference in formant frequencies between the male and female subjects, and that differences in the placement and radii of vocal-tract bends provide a second-order effect.

IV. CONCLUSIONS

The goal of this work was to understand the mechanisms by which the vocal-tract bend causes shifts in formant frequencies. Methods based on manual placement of a rectilinear-radial grid have limitations relating to the fixed scaling of tongue displacement in relation to the grid, the requirement for a reference subject without a sharply curved palate, and the representation of acoustic wavefronts with straight lines oblique to the palate and tongue. The hyperbolic grid-generation algorithm addresses these limitations.

The hyperbolic method automatically generates a grid for positioning the tongue as a displacement from the palate that is adaptive to the vocal-tract shape, decoupling the scaling of the displacement profile from the location of fixed grid lines. The grid smoothing that takes place also allows tongue displacements that are large compared to the local radius of the palate, but that smoothing also contributes shifts to the area function that are influenced by vocal-tract curvature. Those shifts can be made small relative to other effects of the curved vocal tract by proper choice of smoothing coefficients. The resulting grid provides midline and cross-distances along curved paths for determining the lengths and areas of acoustic tubes for calculating formant frequencies; the grid also subdivides the space between palate and tongue in such a way as to allow a correction accounting for vocal tract curvature.

The dominant effect of the curved vocal tract on formant frequencies is the lengthening of the vocal-tract midline in the vicinity of convex curvature such as the alveolar region and the shortening of the midline for concave curvature such as the palate vault and the oral-pharyngeal bend. The curved vocal tract also produces a departure from linear formant scaling; vowels are affected differently because the midline length change is smaller for curvature in the neighborhood of a constriction and large when the separation between tongue surface and the outer vocal tract outline is large. The curved vocal tract results in pronounced outward expansion of the F1-F2 formant envelope for /æ/, where the primary effect of curvature is an overall shortening of the vocal tract. This pattern of envelope expansion stays consistent across changes in the vocal-tract cross sections.

The curvature correction for the transmission-line model is small relative to the changes attributable to the effect of curvature on midline distances. Comparison with FEM along with consideration of acoustic sensitivity suggests that a more detailed treatment of 3D wave propagation effects is only significant if the vocal tract dimensions are known to high accuracy.

After normalizing for vocal-tract length, the amount the midline-distance effect varied with change in oral-pharyngeal length ratio or change in vocal-tract bend radius was small. A much larger formant shift was observed with changes in the volume of the piriform sinuses as well as small change in the length of the epilarynx tube. The changes in these and other local structures merit further investigation regarding their contribution to non-linear scaling of formants during childhood development.

ACKNOWLEDGMENTS

This work was supported in part by NIH Research Grant Nos. R03 DC4362 (Anatomic Development of the Vocal Tract: MRI Procedures) and R01 DC6282 (MRI and CT Studies of the Developing Vocal Tract), from the National Institute of Deafness and other Communicative Disorders (NIDCD). Also, by a core Grant No. P-30 HD03352 to the Waisman Center from the National Institute of Child Health and Human Development (NICHHD). We thank Reid Durtschi for his assistance with anatomic measurements from the imaging studies. S.Y. is a former graduate student of University of Wisconsin-Madison.

[1]L. J. Boë, J. L. Heim, K. Honda, and S. Maeda, "The potential Neandertal vowel space was as large as that of modern humans," J. Phonetics **30**, 465–484 (2002).
[2]L. Menard, J. L. Schwartz, and L. J. Boe, "Role of vocal tract morphology in speech development: Perceptual targets and sensorimotor maps for synthesized French vowels from birth to adulthood," J. Speech Lang. Hear. Res. **47**, 1059–1080 (2004).
[3]L. Menard, J. L. Schwartz, and L. J. Boe, "Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood," J. Acoust. Soc. Am. **111**, 1892–1905 (2002).
[4]J. Serkhane, J. L. Schwartz, and P. Bessière, "Building a talking baby robot," Interaction Studies **6**, 253–286 (2005).
[5]R. Harshman, P. Ladefoged, and L. Goldstein, "Factor analysis of tongue shapes," J. Acoust. Soc. Am. **62**, 693–713 (1977).
[6]S. Maeda, "Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model," in *Speech Production and Speech Modelling*, edited by W. J. Hardcastle and A. Marchal (Kluwer, Netherlands, 1990), pp. 131–149.
[7]J. Liljencrants, "Fourier series description of the tongue profile," Speech Transmission Laboratory-Quarterly Progress Status Reports, Vol. **12**, No. 4, Royal Institute of Technology (KTH), Stockholm, 1971, pp. 9–18.
[8]P. M. C. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1987).
[9]J. Sundberg, B. Lindblom, and J. Liljencrants, "Formant frequency estimates for abruptly changing area functions: A comparison between calculations and measurements," J. Acoust. Soc. Am. **91**, 3478–3482 (1992).
[10]V. L. Streeter, *Fluid Dynamics* (McGraw-Hill, New York, 1948).
[11]D. H. Keefe and A. H. Benade, "Wave propagation in strongly curved ducts," J. Acoust. Soc. Am. **74**, 320–332 (1983).
[12]C. J. Nederveen, "Influence of a toroidal bend on wind instrument tuning," J. Acoust. Soc. Am. **104**, 1616–1626 (1998).
[13]K. Motoki, N. Miki, and N. Nagai, "Measurement of sound-pressure distribution in replicas of the oral cavity," J. Acoust. Soc. Am. **92**, 2577–2585 (1992).
[14]H. Matsuzaki, N. Miki, and Y. Ogawa, "3 D finite element analysis of Japanese vowels in elliptic sound tube model," Electron. Commun. Jpn. **83**, 43–51 (2000).
[15]X. Zhou, C. Y. Espy-Wilson, S. Boyce, M. Tiede, C. Holland, and A. Choe, "A magnetic resonance imaging-based articulatory and acoustic

study of "retroflex" and "bunched" American English /r/," J. Acoust. Soc. Am. **123**, 4466–4481 (2008).
[16]D. L. Brown, G. S. Chesshire, W. D. Henshaw, and D. J. Quinlan, "Overture: An object-oriented software system for solving partial differential equations in serial and parallel environments," in Conference: 8. SIAM Conference on Parallel Processing for Scientific Computing, Minneapolis, MN (1997), pp. 14–17.
[17]W. M. Chan and J. L. Steger, "Enhancements of a three-dimensional hyperbolic grid generation scheme," Appl. Math. Comput. **51**, 181–205 (1992).
[18]W. Henshaw, "The overture hyperbolic grid generator user guide, Version 1.0," Research Report No. UCRL-MA-134240, Lawrence Livermore National Laboratory, Livermore, CA, 2003.
[19]N. A. Petersson, "User's guide to Xcog version 2.0," Technical Report No. CHA/NAV/R-97/0048, Chalmers University of Tech., Gothenburg, Sweden, 1997.
[20]W. M. Chan, "Hyperbolic methods for surface and field grid generation," *Handbook of Grid Generation* (CRC, Boca Raton, FL, 1999).
[21]K. Mochizuki and T. Nakai, "Estimation of area function from 3-D magnetic resonance images of vocal tract using finite element method," Acoust. Sci. & Tech. **28**, 346–348 (2007).
[22]M. M. Sondhi, "Resonances of a bent vocal tract," J. Acoust. Soc. Am. **79**, 1113–1116 (1986).
[23]H. K. Vorperian, R. D. Kent, M. J. Lindstrom, C. M. Kalina, L. R. Gentry, and B. S. Yandell, "Development of vocal tract length during early childhood: A magnetic resonance imaging study," J. Acoust. Soc. Am. **117**, 338–350 (2005).
[24]P. Perrier, L. J. Boe, and R. Sock, "Vocal tract area function estimation from midsagittal dimensions with CT scans and a vocal tract cast: Modeling the transition with two sets of coefficients," J. Speech Hear. Res. **35**, 53–67 (1992).
[25]S. Maeda, "On the conversion of vocal tract x-ray data into formant frequencies," Bell Laboratories, Murray Hill, NJ, 1972.
[26]D. Beautemps, P. Badin, and R. Laboissière, "Deriving vocal-tract area functions from midsagittal profiles and formant frequencies: A new model for vowels and fricative consonants based on experimental data," Speech Commun. **16**, 27–47 (1995).
[27]V. Milenkovic, "Computer synthesis of continuous path robot motion," in *Proceedings of the Fifth World Congress Theory of Machines and Mechanisms* (ASME, New York, 1979), pp. 1332–1335.
[28]M. Loo and V. Milenkovic, "Multicircular curvilinear robot path generation," in *Robots 11 Conference Proceedings and 17th International Symposium Industrial Robots* (SME, Dearborn, MI, 1987), Vol. **18**, pp. 19–27.
[29]M. Loo, Y. A. Hamidieh, and V. Milenkovic, "Generic path control for robot applications," in *Robots 14 Conference Proceedings* (SME, Dearborn, MI, 1990), Vol. **10**, pp. 49–64.
[30]W. Rostafinski, "Monograph on propagation of sound waves in curved ducts," NASA Reference Publication No. 1248 (1991).
[31]K. H. Huebner, *The Finite Element Method for Engineers* (Wiley-Interscience, New York, 2001).
[32]K. Motoki, "Three-dimensional acoustic field in vocal-tract," Acoust. Sci. & Tech. **23**, 207–212 (2002).
[33]B. S. Atal, J. J. Chang, M. V. Mathews, and J. W. Tukey, "Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique," J. Acoust. Soc. Am. **63**, 1535–1553 (1978).
[34]L. L. Beranek, *Acoustics* (McGraw-Hill, New York, 1954).
[35]B. H. Story, I. R. Titze, and E. A. Hoffman, "Vocal tract area functions from magnetic resonance imaging," J. Acoust. Soc. Am. **100**, 537–554 (1996).
[36]J. Dang and K. Honda, "Acoustic characteristics of the piriform fossa in models and humans," J. Acoust. Soc. Am. **101**, 456–465 (1997).
[37]T. Baer, J. C. Gore, L. C. Gracco, and P. W. Nye, "Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels," J. Acoust. Soc. Am. **90**, 799–828 (1991).
[38]A. Soquet, V. Lecuit, T. Metens, and D. Demolin, "Mid-sagittal cut to area function transformations: Direct measurements of mid-sagittal distance and area with MRI," Speech Commun. **36**, 169–180 (2002).

Biomechanical modeling of the three-dimensional aspects of human vocal fold dynamics

Anxiong Yang[a)]
Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Medical School, Bohlenplatz 21, 91054 Erlangen, Germany

Jörg Lohscheller
Department of Medical Computer Science, Trier University of Applied Sciences, Schneidershof, 54293 Trier, Germany

David A. Berry
The Laryngeal Dynamics Laboratory Division of Head and Neck Surgery, David Geffen School of Medicine, University of California, Los Angeles, 31-24 Rehab Center, 1000 Veteran Avenue, Los Angeles, California 90095-1794

Stefan Becker
Institute of Fluid Mechanics, Friedrich-Alexander University of Erlangen-Nuremberg, Cauerstrasse 4, 91058 Erlangen, Germany

Ulrich Eysholdt, Daniel Voigt, and Michael Döllinger
Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Medical School, Bohlenplatz 21, 91054 Erlangen, Germany

(Received 14 May 2009; revised 15 October 2009; accepted 24 November 2009)

Human voice originates from the three-dimensional (3D) oscillations of the vocal folds. In previous studies, biomechanical properties of vocal fold tissues have been predicted by optimizing the parameters of simple two-mass-models to fit its dynamics to the high-speed imaging data from the clinic. However, only lateral and longitudinal displacements of the vocal folds were considered. To extend previous studies, a 3D mass-spring, cover-model is developed, which predicts the 3D vibrations of the entire medial surface of the vocal fold. The model consists of five mass planes arranged in vertical direction. Each plane contains five longitudinal, mass-spring, coupled oscillators. Feasibility of the model is assessed using a large body of dynamical data previously obtained from excised human larynx experiments, *in vivo* canine larynx experiments, physical models, and numerical models. Typical model output was found to be similar to existing findings. The resulting model enables visualization of the 3D dynamics of the human vocal folds during phonation for both symmetric and asymmetric vibrations.
 [DOI: 10.1121/1.3277165]

PACS number(s): 43.70.Aj, 43.70.Bk, 43.70.Gr, 43.72.Ar [AL] Pages: 1014–1031

I. INTRODUCTION

Human voice is generated by the oscillations of two lateral opposing vocal folds located in the larynx.[1] The vocal fold oscillations are excited by air flow through the trachea generated by the lung. Moreover, the characteristics of the acoustic signal, e.g., sound pressure level and pitch of tone, are influenced by vocal fold tensions.[2] A disturbed or hoarse voice arises from irregular vocal fold vibrations.[3] These irregularities are often caused by anatomical asymmetries between left and right vocal fold.[3–5] Asymmetries in the larynx such as unilateral vocal fold polyps or paralysis can be directly observed using endoscopes.[1] However, vocal fold pathology such as functional dysphonia, which presents no visual evidences for morphological laryngeal abnormalities, is only visible during vocal fold oscillation.[1] Hence, analyzing vocal dynamics is a critical component of voice research.

In the clinic, endoscopic high-speed (HS) digital imaging of the vocal fold vibration (100–300 Hz) is becoming increasingly more common, typically with sampling rates of 2000–4000 fps (frames per second). An appropriate approach to analyze endoscopically recorded vocal fold dynamics is the adaption of biomechanical models.[1] Therefore, one or more trajectories at specific locations on the vocal fold edges (see Fig. 1 in Ref. 6) are reproduced. Due to the increase in computational power, this method has been gradually improved:

First, only one lateral trajectory was recreated by using a two-mass-model (2MM).[7,8] The corresponding optimization was manually performed.[9] The first fully automatic optimization procedure was realized by Döllinger *et al.*[1] applying the Nelder–Mead algorithm. With aids of this parameter optimization procedure, the dynamics of healthy voices were simulated.[10] To differentiate normal from unilateral vocal

[a)]Author to whom correspondence should be addressed. Electronic mail: anxiong.yang@uk-erlangen.de.

fold paralysis, classification schemes were proposed.[2] Moreover, according to clinical diagnosis some specific laryngeal diseases exhibit their clinical picture only at certain vocal fold oscillation frequencies; therefore, parameter optimization for non-stationary vocal fold dynamics was performed.[11]

Later, vocal fold movements at three positions (see Fig. 1 in Ref. 6) located along the longitudinal direction were adapted by enhanced two-dimensional (2D) multi-mass-models[6,12] (2DMs). The optimization procedure for sustained phonation was successfully accomplished through verification on two normal and six pathological voices.[6] In addition, for the corresponding non-stationary phonation, a so-called block-based optimization procedure was developed and verified with clinical data sets.[12]

To date, most low-dimensional, biomechanical models of the vocal folds have considered vibration in only one or two dimensions.[13–18] Similarly, optimization has focused primarily on the analysis of superior dynamics of the vocal folds,[6] from a superior aspect. However, it is important to objectively quantify vertical movements[19] as well as medial surface dynamics of the vocal folds.[20,21] Moreover, mucosal wave propagation, which occurs primarily in vertical direction, establishes a continually mechanical transfer of energy from the glottal airstream to the oscillating tissue of vocal folds.[22–24] It is noteworthy that such models could be unduly restricted, if the most complicated aspects of vocal fold dynamics, which occur along the medial surface (including coupling with the glottal air, collision with the opposing vocal fold, etc.),[20,25] are not reasonably taken into account. Therefore, fully analyzing the vertical vibrations,[21] the medial surface dynamics of the folds, as well as the mucosal wave propagation,[19,26] is essential to quantify the vocal fold mechanical characteristics.[20]

In this work, we develop a 3D mass-spring, cover-model of vocal fold vibration. This model predicts the 3D vibrations of the entire medial surface of the vocal fold. In this work, we will assess the feasibility of the model using a large amount of dynamical data obtained from laboratory larynx experiments. Ultimately, one of the primary purposes of this model is to infer an appropriate approximation of biomechanical tissue properties across the entire medial surface of the vocal folds, as the model parameters are optimized to fit model dynamics to the experimental data of vocal fold vibrations. Therefore, it is hoped that the resultant vocal fold dynamics may be firmly interpreted in a physiological way. The model was developed to meet the following criteria.

A. 3D dynamics and mucosal wave

In order to simulate 3D vocal fold dynamics as realistic as possible, the findings of several 3D model templates were taken into account: human hemilarynx,[19,21,26–28] *in vivo* canine dynamics,[29,30] excised canine larynges,[31–34] physical models,[35] synthetic models[36,37] as well as numerical finite-element models.[38,39]

Experiments with excised canine larynges reflect that the amplitudes of the elliptical paths traversed by flesh-points arranged from superior to inferior are decreased.[32] The largest displacements and velocities were concentrated in the upper medial portion of *in vivo* canine vocal folds.[30]

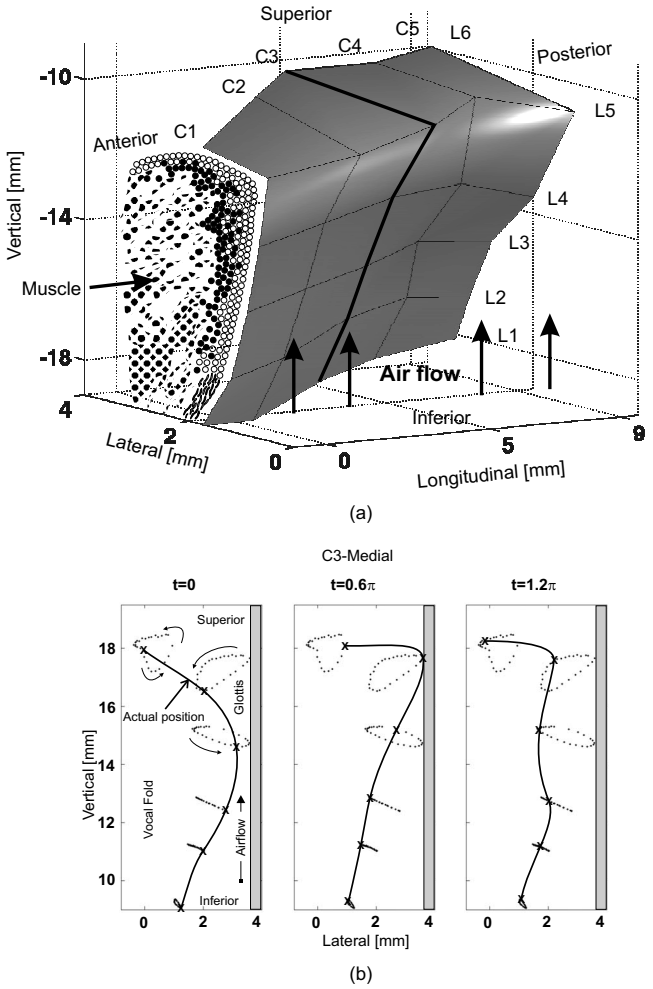


FIG. 1. (a) 3D geometry of one vocal fold surface obtained from hemilarynx experiments (Ref. 26). 30 surgical sutures serving as tracking markers are homogeneously distributed on the vocal fold surface. (Ref. 26). (b) Movements of the tracking markers along the vertical column C3 during one oscillation cycle. The antidromic vibration in lateral direction between inferior and superior sutures is clearly displayed (Ref. 26).

The human hemilarynx set-up was also considered. It allows to observe the dynamics of an entire vocal fold surface, since the other vocal fold is cut off.[21,27–29] For visualizing 3D trajectories extracted from performed HS recordings, six rows from inferior to superior and five vertical columns from ventral to dorsal of surgical micro-sutures were uniformly distributed over the vocal fold surface,[26] see Fig. 1(a). A typical oscillation cycle along the vertical suture column (C3) is shown in Fig. 1(b). The cycle started when the most inferior suture was at the most medial position and begins to move in lateral direction.[26] Within these studies, local 3D displacements, velocities, and accelerations were described and analyzed.[19,26]

During phonation the dynamics of human vocal fold mucosa occur and propagate wavelike in anterior-posterior, medio-lateral, and in caudal-cranial direction.[40] Bending and twisting may take place between the faster and the more slowly moving portions of the vocal folds in the lateral, longitudinal, and vertical dimensions, respectively.[40]

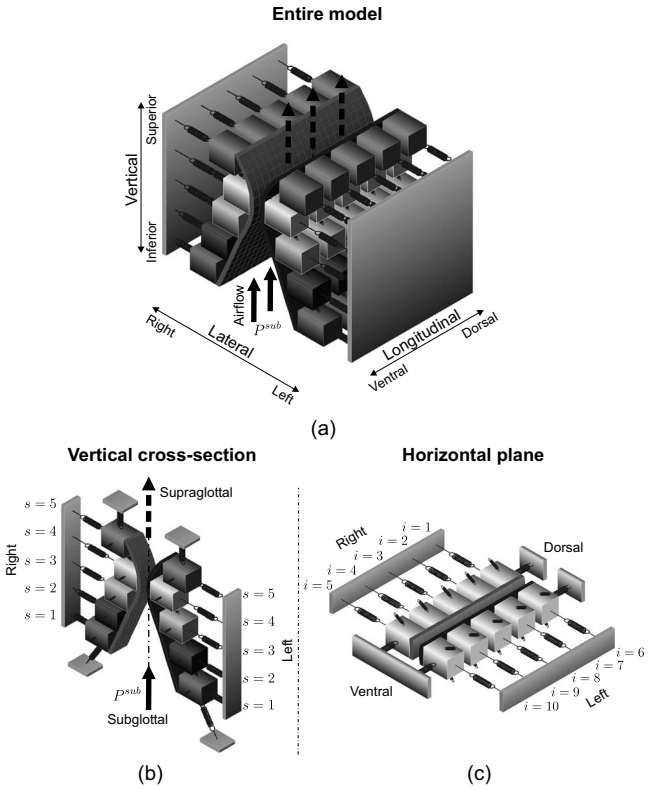


FIG. 2. (a) Schematic three-dimensional representation of the 3DM. Each of both vocal folds consists of five masses on each horizontal plane. The masses are elastically connected to a rigid body in lateral direction by using anchor springs. In addition, the masses are connected to each other through springs located in vertical and longitudinal directions. (b) 3D view of the individual cross-section of the 3DM with plane index s. (c) 3D view of one individual plane with column index i. The rigid bodies representing the anterior commissure and the fixed positions at the dorsal end of both vocal folds are displayed.

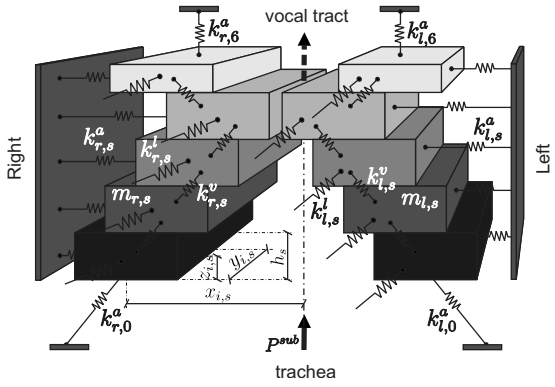


FIG. 3. Schematic representation of the cross-section of the 3DM of the vocal folds. The stiffness of anchor springs, longitudinal springs and vertical springs are indicated with $k^a_{r,s}$, $k^l_{r,s}$, $k^v_{r,s}$ with $s=1,\ldots,5$ from inferior to superior. The subscript r means the lateral right side of the vocal fold ($i=1,\ldots,5$). Likewise, the stiffness of different springs on the lateral left side are symbolized as $k^a_{l,s}$, $k^l_{l,s}$, $k^v_{l,s}$, with index $i=6,\ldots,10$. The thickness of the mass element located in plane s is labeled as h_s. The position of the mass element $m_{i,s}$ is denoted with 3D-components $x_{i,s}$, $y_{i,s}$, $z_{i,s}$. The damping $r_{i,s}$ for each mass element is not shown so as to keep clarity of the drawing.

B. Different glottis closure types

Endoscopic HS digital recordings[1,6,12] were taken as a basis to simulate the five common glottis closure types during normal phonation. The different glottis closure types have been described in detail before[41–44] and are, respectively, classified as *rectangle, hourglass, triangular-pointed dorsal, triangular-pointed ventral, and convex* (see Fig. 5 in Ref. 41).

II. METHODS

A. 3D-multi-mass-model

The vibrational function of human vocal folds depend on myoelastic and aerodynamic properties of the larynx,[1] which are simulated rationally as far as possible within the proposed 3D-multi-mass-model (3DM). The 3DM is on the basis of previous 2DM.[45,46] A corresponding three-dimensional view is sketched in Fig. 2.

Each vocal fold is supposed to consist of five horizontal layers (planes) arranged from inferior to superior. Corresponding to the hemilarynx experiments, each plane contains five longitudinal coupled mass-spring oscillators, see Fig. 2. The masses situated at different vertical planes are elastically connected by using springs in vertical direction. At the ventral end of the model, the left lateral and right lateral sides of the 3DM are connected together at an anchor being similar to the anterior commissure of the larynx, where both vocal folds merge together,[46] Fig. 2(c). Additionally, the dorsal extremities of both sides of the 3DM are coupled to their own fixed positions that are joined with the arytenoid cartilages, which do not vibrate during phonation.[46]

The viscous losses inside the glottis during vocal fold dynamics are neglected.[7,8,47] During phonation the driving aerodynamic force acting on left and right vocal fold is caused by the subglottal pressure P^{sub} which follows the flow dynamic Bernoulli equation.[1,48,49] Thus, under the action of subglottal pressure, the masses located on both sides are set into vibration. Because of simplicity of the 3DM, the nonlinearities[7] of the elastic forces are small enough to be negligible.[1]

The tissue properties of the vocal folds are simplified and represented with lumped elements, namely, masses $m_{i,s}$, dampings $d_{i,s}$, and different associated spring stiffness coefficients, e.g., anchor spring stiffness $k^a_{i,s}$, vertical stiffness $k^v_{i,s}$, and longitudinal stiffness $k^l_{i,s}$ (Fig. 3). The indices (i,s) label the different mass element ($i=1,\ldots,5$ on the right side, and $i=6,\ldots,10$ on the left side, from dorsal to ventral) at different planes with $s=1,\ldots,5$ from inferior to superior, see Fig. 2. The length of the model in the longitudinal direction [from the fixed position at the ventral end to the fixed position at the dorsal end, see Fig. 2(c)] is defined as the glottal length l_g (approximately 11–15 mm) of human vocal folds, based on the Su *et al.*[50] and Gray's anatomy.[51] The height of the model from the most inferior mass element to the most superior mass element is 10 mm. In the lateral direction the distance between the mass element and the glottal midline is 0.425 mm for minimum, and 2.425 mm for maximum. The more specific physical dimensions (e.g., vertical and longitudinal spacings between the mass elements) are referenced in Table I in Sec. II B.

TABLE I. Brief summary of initial model parameters of the 3DM for the simulation of 3D vocal fold dynamics. $i=1,\ldots,10$ is the index of each mass element at each plane. $s=1,\ldots,5$ denotes the plane from inferior to superior. $\tilde{k}^{v}_{i,s}$ occurs between the current plane s and the next upper plane $s+1$. Moreover, $\tilde{k}^{a}_{i,6}$, $\tilde{k}^{a}_{i,0}$ indicating the stiffness of the upper and lower vertical anchor springs illustrated in Fig. 3 are defined with the values 5 and 450 N/m respectively.

Plane s	$\tilde{m}_{i,s}$ (g)	$\tilde{k}^{a}_{i,s}$ (N/m)	$\tilde{k}^{v}_{i,s}$ (N/m)	h_s (mm)	$\lvert\tilde{x}^{r}_{i,s}\rvert$ (mm)	$\tilde{z}^{r}_{i,s}$ (mm)
1	0.02	50	450	2	1.675	0
2	0.02	20	80	2	1.075	$0.5(h_1+h_2)$
3	0.02	20	50	2	0.475	$\tilde{z}^{r}_{i,2}+0.5(h_2+h_3)$
4	0.02	8	50	2	0.425	$\tilde{z}^{r}_{i,3}+0.5(h_3+h_4)$
5	0.01	1.6	⋯	1	2.425	$\tilde{z}^{r}_{i,4}+0.5(h_4+h_5)$
Cross-section		$i=5,10$	$i=4,9$	$i=3,8$	$i=2,7$	$i=1,6$
$\tilde{y}^{r}_{i,s}$ (mm)		$0.17l_g$	$0.34l_g$	$0.5l_g$	$0.66l_g$	$0.83l_g$

In order to realize the above modeling, a system consisting of 50 ordinary differential equations is used for describing the 3D structure dynamics of the vocal folds. In other words, for the motion of each mass element $m_{i,s}$ within the 3DM the ordinary differential equation is defined as

$$m_{i,s}\ddot{\mathbf{x}}_{i,s}=\vec{\mathbf{F}}^{A}_{i,s}+\vec{\mathbf{F}}^{V}_{i,s}+\vec{\mathbf{F}}^{L}_{i,s}+\vec{\mathbf{F}}^{C}_{i,s}+\vec{\mathbf{F}}^{D}_{i,s},$$

$$\mathbf{x}_{i,s}=[x_{i,s},\quad y_{i,s},\quad z_{i,s}]^{T}. \tag{1}$$

$\mathbf{x}_{i,s}$ denotes the position of each mass element in the Cartesian coordinate system. $\dot{\mathbf{x}}_{i,s}$ and $\ddot{\mathbf{x}}_{i,s}$ denote the velocity and the acceleration, respectively. During phonation the acting forces which are considered in the 3DM are as follows:

(a) $\vec{\mathbf{F}}^{A}_{i,s}$: The anchor force acts as the function of the thyroarytenoid muscle in lateral direction,[45,52] see Eq. (A2).
(b) $\vec{\mathbf{F}}^{V}_{i,s}$: The vertical coupling force serves as internal force of the vocal fold tissue in vertical direction,[45,52] see Eq. (A4).
(c) $\vec{\mathbf{F}}^{L}_{i,s}$: Similarly, the longitudinal coupling force imitates the actions of the thyroarytenoid muscle and the vocal ligament in longitudinal direction, see Eq. (A5).
(d) $\vec{\mathbf{F}}^{C}_{i,s}$: During phonation the collision action between both vocal folds occurs, which causes the elastic structure deformation of the vocal folds, see Eq. (A8).
(e) $\vec{\mathbf{F}}^{D}_{i,s}$: Fluid driving force generated by glottal flow: As the driving origin of vocal fold dynamics, this force acts on the vocal fold surface from inferior to superior through the glottis, see Eqs. (A12)–(A14), (A20), and (A21).

All of the above-described forces act on the 3DM in three dimensions. The forces produced by damping elements are included in the different acting forces in detail, see Eqs. (A2), (A4), and (A5). The definitions and computation of these forces are based on the 2DM.[6,52] As an extension of the 2DM (Ref. 6) from 2D into 3D, the formulas of above-mentioned forces are illustrated in detail in the appendix with corresponding Figs. 12 and 13.

B. Model parameter initialization

The 3DM is initialized to vibrate in accordance with the physiological characteristics of 3D vocal fold dynamics as described in Sec. II A. Compared to the 2MM (Ref. 7) and 2DM (Ref. 6) the 3DM consists of high dimensionality on account of the spring stiffness ($50k^{a}_{i,s}$, $40k^{l}_{i,s}$, $40k^{v}_{i,s}$, $50k^{c}_{i,s}$), the mass $50m_{i,s}$, the damping coefficient ($50r^{a}_{i,s}$, $40r^{l}_{i,s}$, $40r^{v}_{i,s}$), the rest position $50\mathbf{x}^{r}_{i,s}$, as well as the pressure $5P_s$ yielding altogether 415 degrees of freedom.

As the initial configuration of the 3DM, a proper initialization for model parameters (i.e., stiffness, mass, thickness, and rest position) was chosen, as shown in Table I. It is based on the two parts: One is the 2MM presented by Ishizaka and Flanaga[7] and Schwarz *et al.*[2] The other is the diffusion properties of the vocal fold mucosal wave from inferior to superior as well as the phase delays obtained from excised larynx experiments,[26] see Table II. The initial stiffness $\tilde{k}^{l}_{i,s}$ of longitudinal coupling spring is obtained by using Eq. (A6). The stiffness $\tilde{k}^{c}_{i,s}$ is initialized after Eq. (A9). Moreover, in order to provide sufficient driving power for the 3DM, the initial subglottal pressure $\tilde{P}^{\mathrm{sub}}=2.4$ kPa has been proven to be adequate, which is located in normal range.[7,53] Additionally, within the 3DM, the glottal length l_g is set to 0.7 cm for females and 1.4 cm for males, which situates within the range of literature values and is identical to the lengths chosen in former works.[1,2,11,12]

TABLE II. Different averages of 3D maximum displacements (x,y,z), velocities (v), accelerations (a) and phase delays (ϕ_x,ϕ_y,ϕ_z) of 3D maximum displacements at six longitudinal lines (L1–L6, from inferior to superior) gathered from Boessenecker *et al.* (Ref. 26) (x,y,z) denote the averages of maximum displacements in lateral, longitudinal, and vertical directions, respectively. Likewise, (ϕ_x,ϕ_y,ϕ_z) are for the phase delays in three directions, respectively.

Suture lines	(x,y,z) (mm)	v (mm/ms)	a (mm/ms)	(ϕ_x,ϕ_y,ϕ_z) (deg)
L6	(1.29, 0.28, 0.91)	0.86	0.21	(170, 134, 51)
L5	(1.57, 0.18, 0.95)	0.98	0.24	(160, 82, 24)
L4	(1.52, 0.25, 0.5)	0.87	0.22	(84, 12, 17)
L3	(1.09, 0.17, 0.47)	0.59	0.15	(39, −4, 10)
L2	(0.64, 0.16, 0.37)	0.37	0.08	(18, −2, 6)
L1	(0.39, 0.15, 0.33)	0.29	0.08	(9, 3, 8)

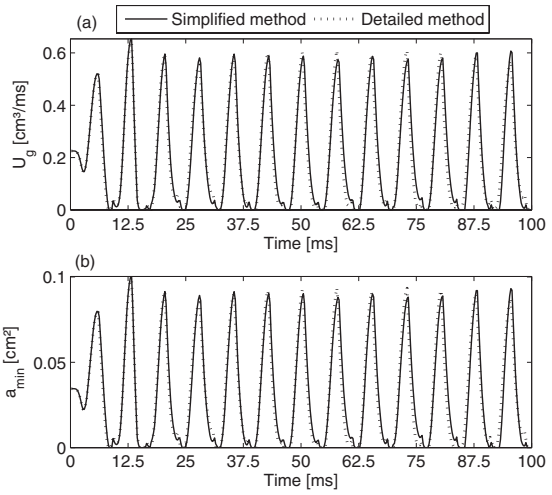


FIG. 4. Upper figure: Glottal volume flow rate U_g obtained by the simplified-method (solid line) and the detailed-method (dotted line). Lower figure: Minimum glottal area $a_{\min}$ obtained with the simplified-method and the detailed-method.

C. Pressure function

In order to compute the driving fluid forces, we discuss two methods to determine the pressure within the glottis. Based on the initial model parameters (i.e., standard model parameters, see Table I) the comparison of the volume flow rate U_g generated from both methods is performed. By the simplified-method the calculation of the pressure function is in accordance to Steinecke *et al.*[8] and Schwarz *et al.*,[6] represented in Eqs. (A10)–(A14). However, by the detailed-method the approximate pressure distribution presented by Ishizaka *et al.*[7] with the jet function for a divergent glottis presented by Titze[48] is used to compute the pressures and corresponding driving forces at different planes, as shown in Eqs. (A15)–(A21).

III. RESULTS

A. Volume flow rate U_g and minimal glottal area $a_{\min}$

Demonstrating the applicability of the model, computational results for volume flow rate, minimal glottal area, as well as pressure and glottal area at each planes are presented. The applied parameters are based on initial values in Table I and with $k^v_{i,3}=80$ N/m, $k^a_{i,3}=80$ N/m, $k^a_{i,4}=13$ N/m. Moreover, since the phonation threshold pressure (i.e., the minimum subglottal pressure required to initiate tissue oscillation, interpreted as a measure of ease of phonation)[24,54,55] is directly linearly related to the prephonatory glottal half-width, demonstrated in previous results and empirical observations.[23,24,55] To decrease phonation threshold pressure, the applied rest positions $\tilde{x}^r_{i,s}$ in the lateral direction are heuristically reduced to 32% of the initial values, so that the ease of phonation was resulted in. In Fig. 4(a) the volume flow rates, U_g in the simulation of vocal fold oscillation using the 3DM with new adjusted model parameters by the simplified- and detailed-methods are shown. The duration of the simulation time was 100 ms, starting from rest state $t=0$. Except the parts of launch phase of the simulation, the maximum volume flow rates U_g by using the simplified- and detailed-methods were, respectively, 0.607 and 0.609 $\mathrm{cm}^3/\mathrm{ms}$.

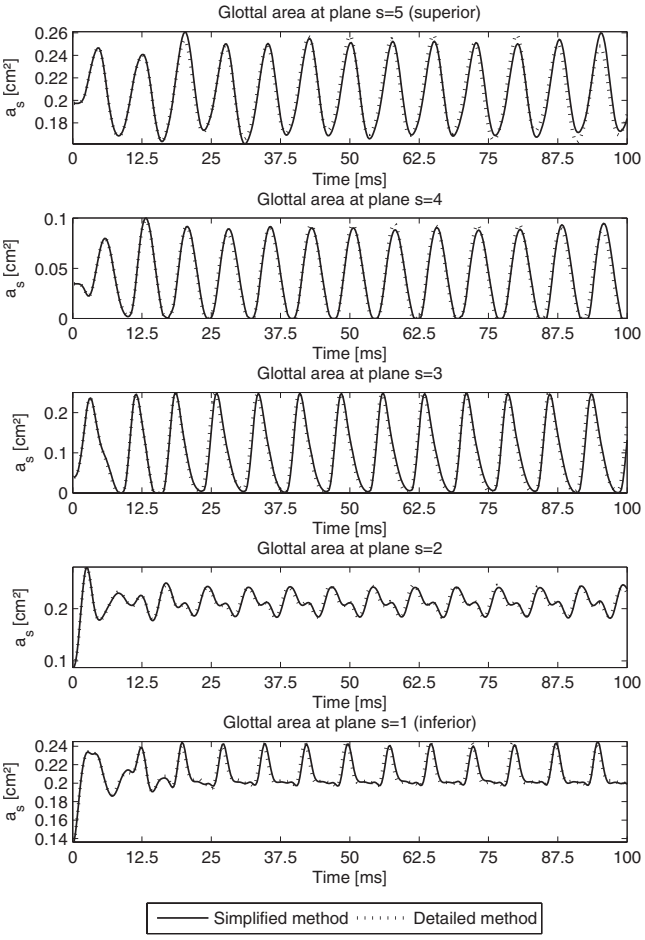


FIG. 5. Glottal area a_s for each plane ($s=1,\ldots,5$) by using the simplified-method (solid lines) and the detailed-method (dotted lines).

The volume flow rate U_g during the 3DM dynamics was periodically equal to zero due to sufficient glottal closure, see Fig. 4(a). The corresponding minimal glottal area $a_{\min}$ is shown in Fig. 4(b). In 3DM the minimal glottal area occurred at plane $s=4$ (corresponds to the edge of the vocal folds) and plane $s=3$ (right below the vocal fold edge).

Tiny secondary peaks of minimal glottal area appeared, Fig. 4(b). These secondary peaks corresponded to alternating minimal values of glottal area at planes $s=3,4$ (Fig. 5). This means mass elements at plane $s=4$ approached to glottal midline, meanwhile mass elements at plane $s=3$ were still pushed away from the glottis center and vice versa. Hence, minimal glottal area phase delays between plane $s=3$ and $s=4$ created the small secondary peaks.

B. Pressure P_s and corresponding glottal area a_s

For optimizing the computation time of the 3DM, the detailed and simplified pressure distribution functions are compared. As shown in Fig. 6, pressures P_s at different planes $s=1,\ldots,5$ within the 3DM by using the simplified- and detailed-methods are illustrated. The pressures at superior planes $s=4,5$ were derived zero. The shapes of the pres-

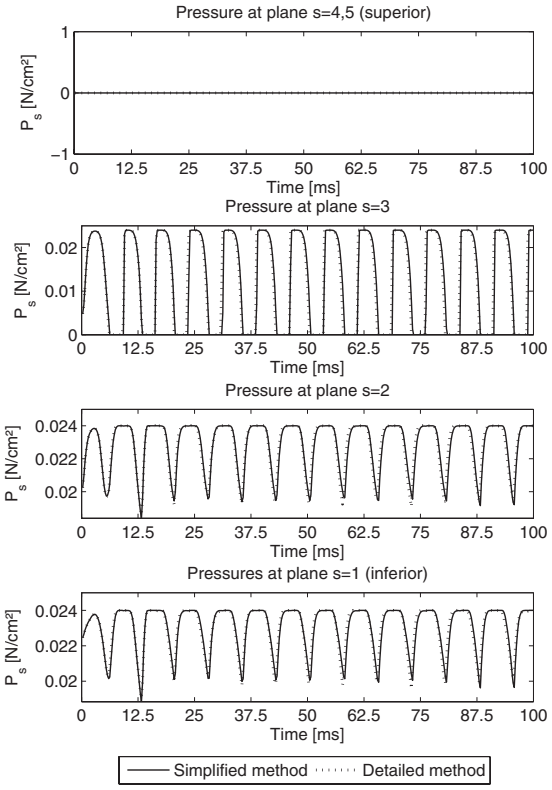


FIG. 6. Pressure P_s for different planes ($s=1,\ldots,5$) by using the simplified-method (solid lines) and the detailed-method (dotted lines).

sures indicate that the pressure at third plane had strong fluctuations. The pressure at the most inferior plane $s=1$ fluctuated the least. The corresponding frequencies for both pressure functions were 132 Hz and are in range of normal male voice. After the launch process of the 3DM, the maximum and minimum values of the pressures derived by the simplified- and the detailed-methods [Eqs. (A10)–(A21), see Appendix] are shown in Table III exhibiting similar values.

The corresponding glottal area a_s at each plane s by using the simplified- and detailed-methods are shown in Fig. 5. During stable vibration of the 3DM, the maximum and minimum values of the glottal area derived by the simplified- and detailed-methods can be seen in Table III. By means of the observation of variations for glottal area at each plane, we found that the amplitudes of the glottal area at planes $s=4,5$ were larger than others. The glottal area at the most inferior plane was the smallest in accordance to the results in Boessenecker *et al.*[26] The glottal pressure distributions for both methods (simplified and detailed) were equal. Hence, from now on, the applied pressure distribution based on the simplified-method as described in Eqs. (A10)–(A14) is applied.

C. Dynamics of the 3DM

In the following, important properties describing the structure dynamics of the 3DM are presented. The mucosal wave, the 3D trajectories, the displacements with corresponding fundamental frequencies, the velocities, as well as the accelerations are given for certain examples to show the model's applicability.

1. 3D movements and mucosal wave

To visualize the symmetric output dynamics generated with initial model parameters (Table I) within the 3DM, the dynamics of the mass elements situated on the right vocal fold during one oscillation cycle are displayed in Fig. 7. The model surface was interpolated using spline functions minimizing the flexion within curves.[56]

Correspondingly, movements of the mass elements on the right side during one oscillation cycle are described in Fig. 7 in detail. The dynamics of the 3DM start at rest position with rectangular glottis closure type; therefore, the trajectories for the launch phase of the 3DM were also displayed. The mucosal wave propagating on the vocal fold surface is clearly demonstrated. The corresponding duration of this cycle was 9 ms. In general, Fig. 7 emerges: In the beginning ($t=0$) the 3DM was closed, while the intraglottal pressure increased. The mass elements located at plane $s=3$ were still attached to the opposite side at midline. Due to the increase in intraglottal pressure, the mass elements located at planes $s=1,2$ on both sides began to deform measurably and were pushed away from midline. At $t=2/7\pi$, the mass elements at plane $s=3$ were pushed away from midline, while the mass elements at plane $s=2$ continued being pushed away from midline. However, the mass elements at plane 4 already began to approach the midline. At $t=4/7\pi$, the mass elements at plane $s=4$ further approached the midline and moved upward. The mass elements at plane $s=3$ were further

TABLE III. Maximum/minimum values of the pressures P_s and the glottal area a_s at each plane by using the simplified- and detailed-methods after the launch phase of the model.

Plane s		1	2	3	4	5
		Pressure				
Simplified	Max. P_s (N/cm^2)	0.024	0.024	0.024	0.0	0.0
	Min. P_s (N/cm^2)	0.02	0.0192	0.0	0.0	0.0
Detailed	Max. P_s (N/cm^2)	0.024	0.024	0.024	0.0	0.0
	Min. P_s (N/cm^2)	0.0196	0.0191	0.0	0.0	0.0
		Glottal area				
Simplified	Max. a_s (cm^2)	0.2434	0.2457	0.25	?0,0934	0.2578
	Min. a_s (cm^2)	0.1982	0.1824	0.0	0.0	0.1673
Detailed	Max. a_s (cm^2)	0.2447	0.2452	0.2480	0.0949	0.2582
	Min. a_s (cm^2)	0.1973	0.1844	0.0	0.0	0.1647

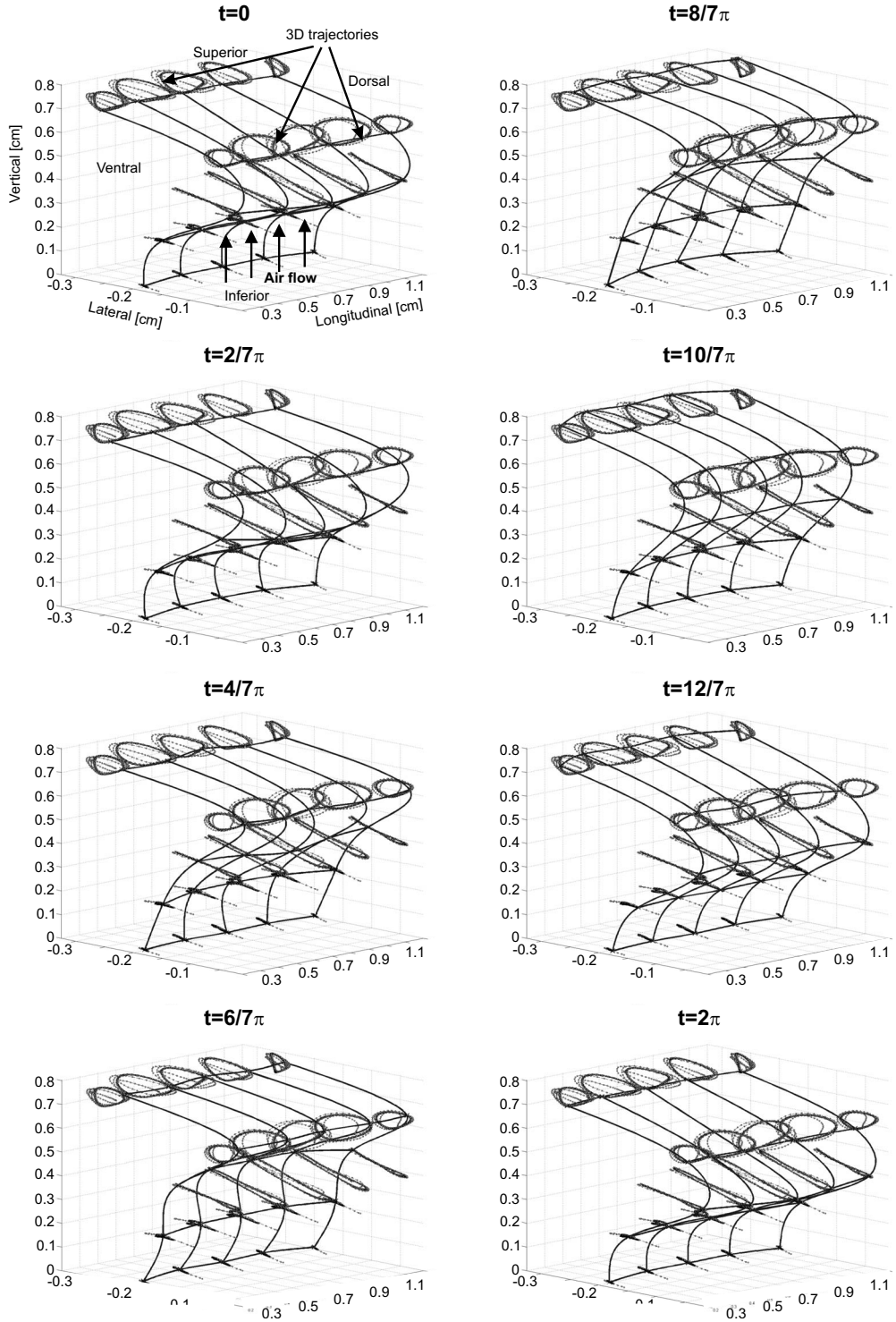


FIG. 7. Sequence of 3D dynamics for all mass elements located at one vocal fold during one oscillation cycle, with a spline-interpolated 3DM surface. The dotted lines represent the trajectories of each mass element.

pushed away from midline. The superior mass elements at plane $s=5$ were elevated. From time step $t=2/7\pi$ to $t=6/7\pi$, the mass elements at upper planes ($s=3,4,5$) moved upward due to high intraglottal pressure. From time step $t=8/7\pi$ to $t=12/7\pi$, the corresponding mass elements sank because of relaxation under the low intraglottal pressure. At $t=12/7\pi$, the model began to close, meanwhile the increase in intraglottal pressure started again. At $t=2\pi$, the model was closed again and the intraglottal pressure was gradually established. Glottis closure occurred at t planes $s=3,4$.

In order to inspect the quality of the 3DM, a comparison between the 3D trajectories generated from 3DM and the 3D trajectories extracted from excised human hemilarynges[26] at middle cross-section is depicted in Fig. 8. Owing to the tiny movements of the micro-suture located at the most inferior sagittal line [Fig. 8(b)], the dynamics at these micro-sutures

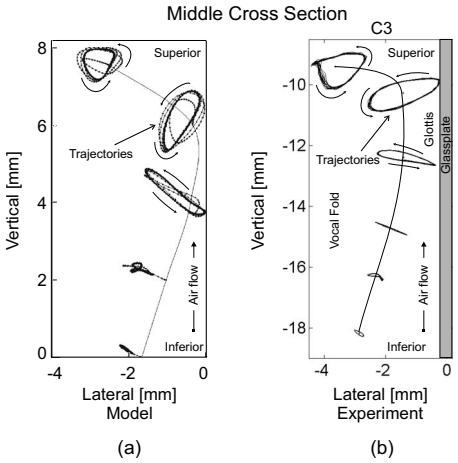


FIG. 8. (a) Five simulated 3D trajectories (dotted lines) along middle vertical cross section produced by 3DM. (b) Six experimental 3D trajectories (dotted lines) along middle vertical cross section extracted from an excised human hemilarynx (Ref. 26). The vertical contours in both images correspond to the rest positions of the mass elements as well as the vocal fold tissue, respectively. For all five mass elements (from inferior up to superior) in (a) the movement ranges can be roughly described as 0.5, 1, 2, 2, and 1.5 mm, respectively. By comparison, the movement ranges of the upper five suture points (from inferior up to the vocal fold edge) in (b) are almost 0.3, 1, 2, 2, and 1.7 mm, respectively. Hence, the trajectories show similar behavior.

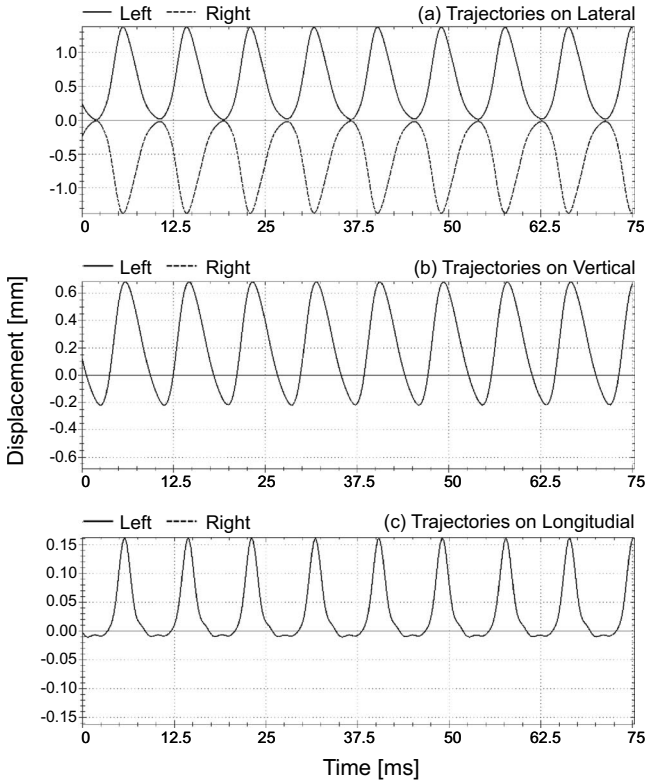


FIG. 9. Simulated symmetrical trajectories in lateral direction (upper figure), vertical direction (middle figure), as well as longitudinal direction (lower figure) of two opposing mass elements (dotted lines for the mass element $m_{2,3}$ on right side, solid lines for the mass element $m_{7,3}$ on left) which are situated at third plane and second cross-section in the symmetrical 3DM.

were neglected in 3DM. This means that only the movements of the micro-sutures located at the upper five sagittal lines (five planes) were taken into account. The shape of movement of the mass element at fifth plane (i.e., the most superior plane) was like a triangle, being similar to the experimental movement of the suture point with uppermost position of the excised vocal fold surface. In the fourth plane (i.e., around the vocal fold edge) the movement of the mass element in 3DM was much like an ellipse, which was similar to the experimental dynamics of the suture point at the corresponding plane, see Fig. 8. In third plane (i.e., right below the vocal fold edge) the movement of the mass element was like a flat ellipse which was similar to the experimental dynamics of the suture point at L4 (fourth sagittal line). The movement shapes of mass elements in lower planes ($s=1,2$) were like line segments, which were similar to the experimental movements of micro-suture at L2 and L3. Both models show a significant decrease in movements from medial to inferior.

2. 3D trajectories

Standard model parameters (Table I) were chosen now. The corresponding glottis closure type was rectangle. In order to show the model application, the absolute values $|\vec{x}^r_{i,s}|$ for lateral rest positions of the masses from inferior to superior were temporarily reset to 0.988, 0.634, 0.280, 0.251, 1.431 mm (i.e., convergent glottis). In Fig. 9 this example of 3D trajectories extracted from the symmetric 3DM is described. These 3D trajectories corresponded to mass elements $m_{2,3}$ and $m_{7,3}$ on both sides, respectively. Their corresponding movements in lateral direction were exactly symmetrical, see Fig. 9(a). Due to the symmetric characteristics of the healthy vocal folds as assumed in this work, the vertical and the longitudinal movements for mass elements which were located on both sides (left and right) at the same plane and same cross-section were identical, see Figs. 9(b) and 9(c).

3. 3D displacement maxima

Figure 10 shows the 3D displacements of mass elements within the 3DM, being generated by the symmetric initial model parameters (from now on, the rest positions for mass elements were defined as shown in Table I again). Due to symmetry, only one vocal fold side was depicted. The displacements of mass elements were symmetric, respectively, in the longitudinal direction and centered on the middle cross-section. The applied glottis closure type was rectangular. The maximum displacements of the mass elements can be described as follows.

a. Lateral displacement maxima. Figure 10(a) shows the comparison of the maximum lateral displacements belonging to the different mass elements $m_{i,s}$. The range of the maximum displacements in lateral direction was between 0.09 mm ($m_{1,1}$ and $m_{5,1}$) and 1.47 mm ($m_{3,3}$). This means that the highest lateral displacement occurred at plane $s=3$ on middle cross-section. Overall compared to the corresponding longitudinal displacements, the lateral displacement maxima were much higher. Additionally, for all mass elements the lateral displacements increased from the most inferior plane ($s=1$) up to the medial plane ($s=3$). Similarly,

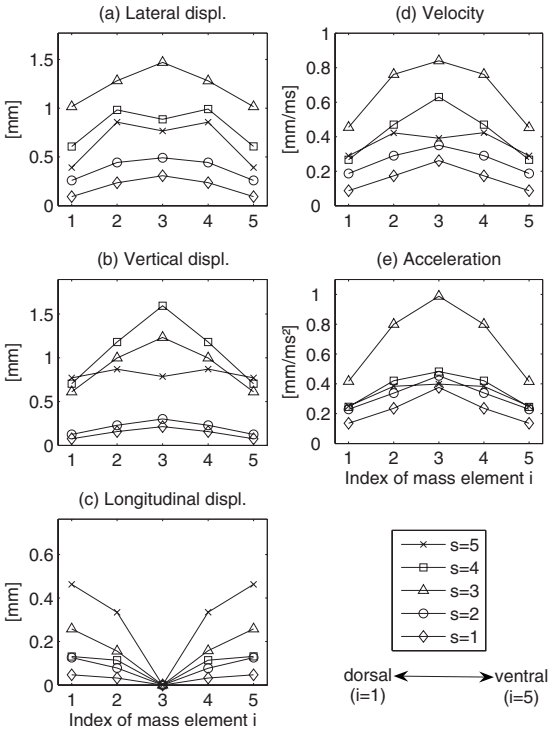


FIG. 10. [(a)–(c)] Maximum lateral, vertical, and longitudinal displacements of each mass element on one side of the 3DM. (d) Absolute maximum velocities of each mass elements on one side of the 3DM. (e) Absolute maximum accelerations of each mass elements on one side of the 3DM. $s=1,\ldots,5$ denote the planes from inferior up to superior.

an increase from the most superior plane ($s=5$) to the medial plane ($s=3$) occurred. Along longitudinal direction the lateral displacements decreased from middle cross-section to dorsal/ventral cross-section, except mass elements $m_{3,4}$ and $m_{3,5}$.

b. Vertical displacement maxima. The vertical displacement maxima ranged within 0.08 and 1.6 mm, Fig. 10(b). The highest vertical displacement maximum was situated at mass element $m_{3,4}$, which was located in middle cross-section at plane $s=4$. The second highest vertical displacement maximum occurred at medial plane ($s=3$) on the middle cross-section. The lowest vertical displacement maxima were at the ventral and dorsal mass elements ($i=1,5$) at the most inferior plane ($s=1$). The vertical displacement maxima at superior planes ($s=3,4,5$) were nearly identical on the ventral and dorsal cross-sections ($i=1,5$). At inferior planes ($s=1,2$) the vertical maximal displacement values were similar and significantly decreased compared to planes $s=3,4,5$. In general, the vertical displacement maxima increased from inferior up to superior and decreased from middle cross-section to dorsal/ventral cross-section.

c. Longitudinal displacement maxima. The appearance of longitudinal displacements is caused by longitudinal coupling forces with corresponding stiffnesses as described in Eqs. (A5)–(A7). Figure 10(c) shows the computed maximal longitudinal displacement values. The longitudinal displacement maxima ranged between 0 and 0.46 mm. The smallest longitudinal displacement occurred at the most inferior plane ($s=1$). The largest longitudinal displacement was found at the most superior plane ($s=5$). Compared with the corresponding lateral and vertical displacement maxima, the longitudinal displacement maxima were the smallest. Due to longitudinal symmetry of the initial model parameters, the longitudinal displacement maxima were equal to zero at middle cross-section. For all five planes, the longitudinal displacement maxima decreased from ventral ($i=5$) and dorsal ($i=1$) cross-sections to middle cross-section ($i=3$).

4. Velocity and acceleration maxima of dynamics

In order to get motion characteristics of the 3DM, the corresponding distributions of velocity and corresponding acceleration maxima values are shown in Figs. 10(d) and 10(e).

a. Velocity maxima. The maximal computed velocity values are shown in Fig. 10(d). Velocity maxima ranged from 0.09 up to 0.84 mm/ms. The highest velocity maximum was located on mass element $m_{3,3}$. The lowest velocity maxima were found on mass elements $m_{1,1}$ and $m_{5,1}$. In general, for all mass elements the velocity maxima increased from ventral/dorsal cross-section ($i=1,5$) up to middle cross-section ($i=3$), except mass element $m_{3,5}$. For all five cross-sections ($i=1,\ldots,5$), the velocity maxima increased from the most inferior plane ($s=1$) to the medial plane ($s=3$). Likewise, from the most superior plane ($s=5$) to the medial plane, there was also a significant increase in velocity maxima, except dorsal/ventral cross-section at plane $s=5$.

b. Acceleration maxima. Figure 10(e) shows the maximal computed accelerations. The maximal acceleration values ranged between 0.14 and 0.99 mm/ms^2. It was quite obvious that the acceleration maximum at medial plane ($s=3$) were the highest, especially at middle cross-section ($i=3$). The lowest acceleration maxima were at the most inferior plane ($s=1$), especially at ventral/dorsal cross-section ($i=1,5$). For other planes (except planes $s=1,3$), the acceleration maxima were similar and decreased significantly compared to those at plane $s=3$. For all mass elements the acceleration maxima increased from dorsal/ventral cross-section to middle cross-section.

D. Comparison of dynamics

In order to demonstrate the applicability and validation of the proposed model, the 3D structure dynamics are generated by the 3DM with different subglottal pressures P^{sub}. Since the range of the subglottal pressure is mostly described from 0 to 3.5 kPa,[7,57] we selected the subglottal pressures P^{sub} as 1.2, 1.6, 2, 2.4, 2.8, and 3 kPa. The longitudinal length l_g was set to 14 mm (male). For all pressure values, inital model parameters were chosen, Table I. The glottis closure type is always set as rectangular. Altogether six examples will be discussed.

Due to longitudinal and lateral symmetrical definition of the model parameters, the comparison of the six examples is performed for one vocal fold side along the vertical middle cross-section in the 3DM. Figure 11 shows the displacement, the velocity, and the acceleration maxima of the six examples. Owing to symmetries of predefinition for the examples, the maximal longitudinal displacement values for mass elements on the middle cross-section were derived

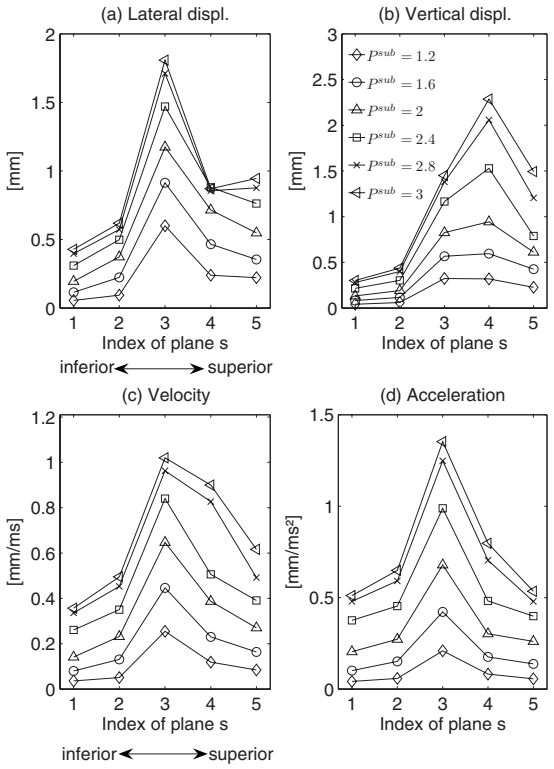


FIG. 11. Displacement, velocity and acceleration maxima of mass elements on the right side of the middle cross-section for different subglottal pressure configurations. The maximum displacement values are illustrated in lateral and vertical directions.

zero. In general the higher the subglottal pressure was set to, the greater the displacement and velocity as well as acceleration maxima were achieved. The fundamental frequencies by different subglottal pressures (1.2, 1.6, 2, 2.4, 2.8, and 3 kPa) were derived as 104, 108, 108, 108, 108, and 108 Hz.

1. Comparison of displacement maxima

Figure 11(a) shows the computed maximal lateral displacement values for the six different subglottal pressures. The range of maximal lateral displacement values was between 0.06 mm (P^{sub}=1.2 kPa, s=1) and 1.81 mm (P^{sub}=3 kPa, s=3). For all mass elements on middle cross-section, an increase in subglottal pressure resulted in a significant increase in lateral displacements. This increase of lateral displacement maxima was more explicit for the three upper planes (s=3,4,5), except for subglottal pressure P^{sub}=2.8, 3 kPa.

Figure 11(b) shows the maximal vertical displacement values by different subglottal pressures. Vertical displacement maxima ranged from 0.04 mm (P^{sub}=1.2 kPa, s=1) up to 2.29 mm (P^{sub}=3 kPa, s=4). As found for lateral displacement directions, an increase in subglottal pressure yielded a marked increase in vertical displacements for all mass elements in middle cross-section. Likewise, this increase in vertical displacement maxima was higher for the three upper planes (s=3,4,5).

2. Comparison of velocity and acceleration maxima

The maximal velocity values for different subglottal pressures are shown in Fig. 11(c). The range of maximal velocity values was between 0.03 mm/ms (P^{sub}=1.2 kPa, s=1) and 0.99 mm/ms (P^{sub}=3 kPa, s=3). An increase in subglottal pressure also yielded larger velocity values.

Figure 11(d) displays the computed maximal acceleration values for all mass elements at middle cross-section of one side. The range of values was from 0.02 mm/ms (P^{sub}=1.2 kPa, s=1) up to 1.44 mm/ms (P^{sub}=3 kPa, s=3). As found for velocity maxima of mass elements by different subglottal pressures, an increase in subglottal pressure resulted in a significant increase in acceleration values.

IV. DISCUSSION

In clinical practice, quantitative analysis of the data captured by HS digital imaging can be a daunting task. It can be realized by directly describing and analyzing laryngeal characteristics such as fundamental frequency, vocal fold length, and vocal fold displacements.[41,42,58,59] Using the instantaneous frequency and amplitude obtained from high-speed laryngeal imaging recordings, the glottal perturbation can be measured.[60] Also investigated were growth of vocal fold amplitudes, subglottal pressure as well as myoelastic forces reflecting the phonation onset dynamics.[61–63]

For further supporting clinical diagnosis of voice disorders, objective quantification of the vocal fold 3D dynamics is meaningful. In order to visualize and objectively quantify spatiotemporal characteristics in vocal fold 3D dynamics with reasonable computational costs, a new 3D biomechanical modeling approach was described. Therefore, by means of the 3DM (Fig. 2) as extension of the 2DM,[2] the 3D dynamics of the entire vocal folds are rationally reproduced. Through the modification of model parameters describing physiological properties (i.e., mass and stiffness distributed in the entire larynx), the model simulates vocal fold dynamics by using 25 control points distributed along the surface of each vocal fold.

The model parameters are appropriately initialized to comply with three aspects: (1) The model generated 3D dynamics are similar to 3D dynamics extracted from hemilarynx experiments,[21,26–28,32] (2) similar to synthetic larynx models[36,37] and finite-element models,[38,39] and (3) the 3DM inherits the biomechanical numerical characteristics of vocal fold dynamics, which have been proposed and justified in earlier works.[7,45,46,64] Its model parameters have the values which match the vocal fold physiology.

A. Generation of the volume flow rate, pressures, and glottal area

In this work, the volume flow rate U_g during the 3DM dynamics using the initial model parameters shown in Table I after phonation onset is periodically equal to zero. This result [see Fig. 4(a)] is in agreement with the volume flow rate presented in Refs. 65 and 66 as well as Ref. 67. In accordance with Eq. (A18), it can be explained that the minimum glottis area a_{min} in our case periodically amounts to zero [see Fig. 4(b)] due to regularly complete glottal closure

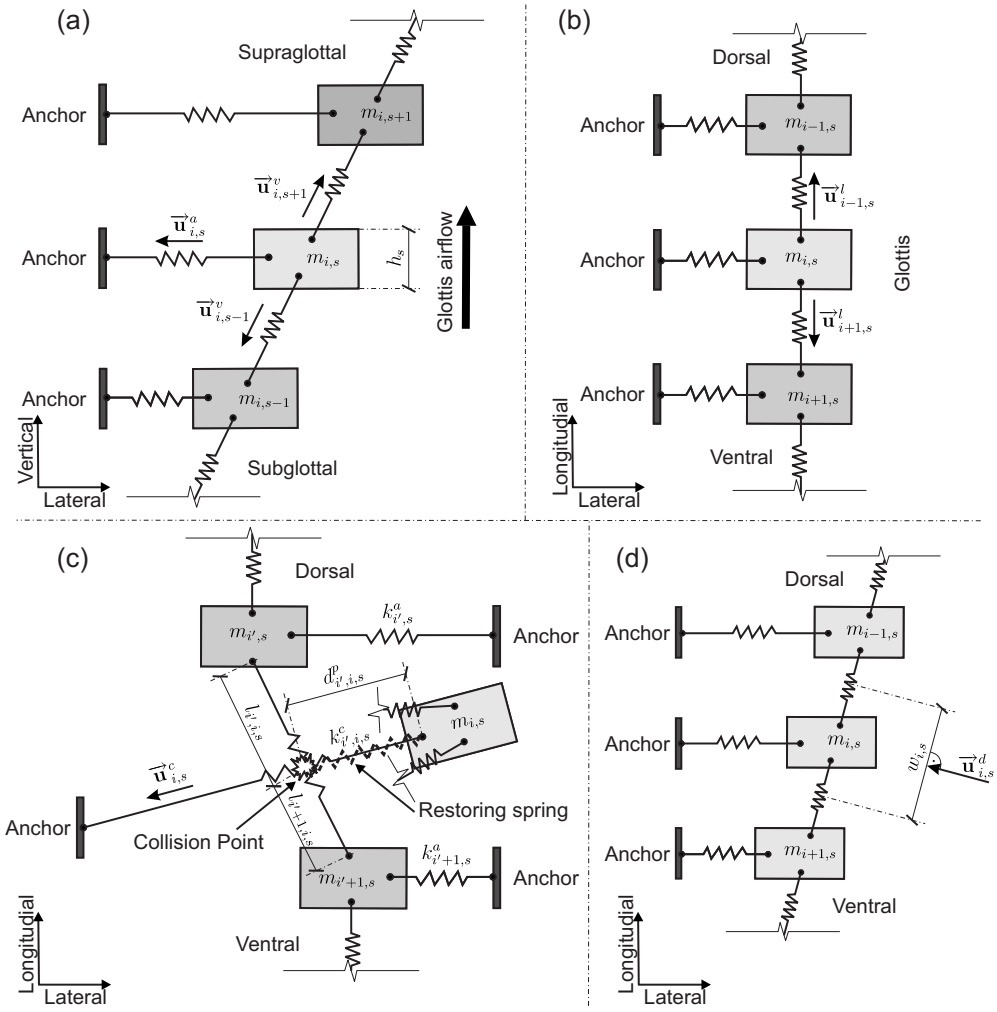


FIG. 12. (a) Graphical definition of the unit vector $\vec{\mathbf{u}}^{a}_{i,s}$ for mass element $m_{i,s}$ in the computation of the anchor force. Unit vector $\vec{\mathbf{u}}^{v}_{i,s-1}$, $\vec{\mathbf{u}}^{v}_{i,s+1}$ aiming at mass element $m_{i,s}$ for the vertical coupling force. h_s means thickness of the mass element $m_{i,s}$. (b) Unit vector $\vec{\mathbf{u}}^{l}_{i-1,s}$, $\vec{\mathbf{u}}^{l}_{i+1,s}$ for mass element $m_{i,s}$ in the computation of longitudinal coupling force. (c) Graphical definition of collision and corresponding unit vector $\vec{\mathbf{u}}^{c}_{i,s}$ for mass element $m_{i,s}$ in the calculation of collision force. When the collision on mass element $m_{i,s}$ occurs, the anchor spring of mass element $m_{i,s}$ crosses through the longitudinal coupling spring between two adjacent mass elements $m_{i's}$ and $m_{i'+1,s}$. $k^{c}_{i',i,s}$ indicates the equivalent stiffness of the restoring spring (Ref. 6) $d^{p}_{i',i,s}$ denotes the penetration depth during collision. (d) Graphical definition of the effective side length $w_{i,s}$ and the unit vector $\vec{\mathbf{u}}^{d}_{i,s}$ for the mass element $m_{i,s}$ in the computation of the driving force.

during vocal fold oscillations. It conforms to the phenomenon observed in the hemilarynx experiments.[26,30] In addition, glottal closure insufficiency can also occur in normal phonation, especially for women in the dorsal region.[42] However, the common case for normal phonation goes along with complete glottal closure.

The obtained pressures in the 3DM are constantly equal to zero for $s=4,5$. For $s=3$ the pressure is periodically equal to zero as presented by Lowell *et al.*[65] and Jiang *et al.*[68] as well as Story *et al.*[47] However, they derived also negative intraglottal pressure values, which is excluded due to Eqs. (A10) and (A19) by definition in our model. For $s=1,2$ the pressure is almost periodically zero. This can be explained by assumptions in Refs. 8 and 48 as well as Ref. 69. The net intraglottal pressure is positive in the case of the convergent orifice (since it has to decrease to a value of zero) and negative in the case of the divergent orifice (it has to increase to a value of zero). The airstream often detaches from the surface of the vocal folds and forms a jet[48] which keeps the pressure in the glottis closer to zero. Bernoulli's law is less

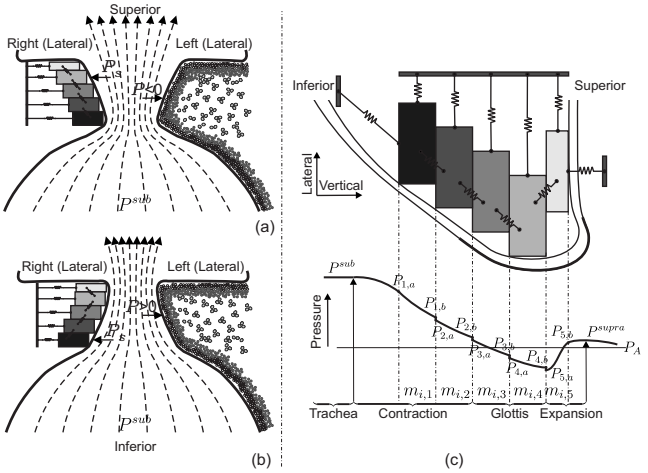


FIG. 13. (a) Airflow and driving pressure in the divergent glottis, including the cross-sectional structure of the 3DM. (b) Airflow and driving pressure in the convergent glottis where the driving pressure P acts on the tissue. Accordingly, in the 3DM, P_s acts on mass elements of each plane ($s=1,\ldots,5$). (c) Detailed pressure computation. P_A denotes the atmospheric pressure (Ref. 66). According to the assumption from Titze (Ref. 48) and conclusion in Ref. 69. P_A as well as P^{supra} is set to zero.

applicable to the divergent glottis. This assumption applied in Eqs. (A10) and (A20) yielded that pressures at upper planes $s=4,5$ were derived zero. This result supports the conclusion demonstrated in Spencer *et al.*:[36] *Flow studies indicate that fluid pressure acts mainly on the inferior surface*.

Moreover, secondary peaks of minimal glottal area illustrated in Fig. 4(b) explicitly demonstrate that the minimal values of glottal area alternatively arise at different planes ($s=3,4$, see Fig. 5). Similarly, the secondary peaks also occurred in the envelope of volume flow rate [Fig. 4(a)] dependent on $a_{\min}$ as found in Ref. 70. This phenomenon is in agreement with conclusions about glottal area waveforms presented in Koizumi *et al.*[66] and Wong *et al.*[52] as well as Ishizaka *et al.*[7] It also reveals that phase difference of glottal areas between different planes exists,[7] which is due to the displacement phase delays occurring during mucosal wave propagation as concluded in Ref. 26. However, no phase delays existed in pressures P_s ($s=1,\ldots,5$) due to the intraglottal pressure following the Bernoulli equation, see Fig. 6.

For the 3DM two pressure functions were compared, i.e., a simplified- and a detailed-method. The simplified-method originates from Schwarz *et al.*[6] and Steinecke *et al.*[8] In the detailed-method, pressure distribution along the glottal flow is taken into account. The main difference between both methods is the calculation of the glottal pressure. As a result derived from Fig. 4, the pressures, glottal area, and corresponding volume flow rate as well as minimum glottal area produced from both methods were identical. This means that the simplified-method yields similar pressure distribution as produced by the detailed-method. The reason that little difference was found in the results for the simplified- and detailed-methods might be that both of them assume no glottal viscous losses. It is consistent with the assumption that glottal viscous losses is neglected in the model. Due to simplicity of computation of the simplified-method in contrast to the detailed-method, we used and suggested the simplified-method.

B. Overall dynamics of the 3DM

In contrast to prior works,[7,45,46,64] the 3D model dynamics included the vertical component and simulated the mucosal wave, see Fig. 7. During the oscillation cycle the mass elements located at the superior planes $s=4,5$ moved generally lateral and downward during the open period, upward during the closed period, so as to finish the previous oscillation cycle. The movements of the mass elements at medial plane $s=3$ and inferior planes $s=1,2$ were predominantly performed in lateral direction (Fig. 7). These resulting vertical dynamics are comparable to experimental results from excised human vocal folds[26] and *in vivo* canine vocal folds.[30]

Additionally, Fig. 7 implicitly supports the conclusions presented by Döllinger *et al.*[19] and Berry *et al.*[32] that a low-pressure condition is created by a divergent glottis (supports the folds to approximate) and a high-pressure condition is created by a convergent glottis (support to push the vocal folds apart), see Fig. 8 in Ref. 32 and Fig. 3 in Ref. 19.

1. Mucosal wave propagation

In prior studies, with the aids of hemilarynx data[21,26] as well as experimental data from excised canine larynges[30–32] mucosal wave propagation in normal phonation was analyzed and quantified. It was confirmed that lateral phase delays were much higher than vertical or longitudinal phase delays. Results in Ref. 26 gave lead for displacement maxima in 3D: 0.26–1.90 mm (lateral), 0.21–1.15 mm (vertical), and 0.10–0.33 mm (longitudinal). Similarly, such phenomenon including vertical and lateral vibrations was reported in a study on aerodynamically and acoustically driven modes of the vocal fold vibration by using a physical model.[35]

In this work, the model dynamics are displayed along the vertical direction from inferior to superior through the entire glottis, see Fig. 8(a), which reflects the mucosal wave propagation of the human vocal folds. It shows the simulated 3D trajectories of each mass element at five planes being similar to the corresponding 3D trajectories extracted from previous experiments of excised vocal folds.[19,26] The simulated 3D trajectories of the mass element located at fifth plane, i.e., the most superior plane in the vertical direction, is like a triangle. Meanwhile, the experimental movement of the micro-sutures with the uppermost position at the surface of excised vocal folds is very much alike. At the fourth plane, i.e., around the vocal fold edge, the movement of the mass element is nearly elliptical, which is similar to the experimental dynamics of the micro-sutures at the corresponding plane. It moved upward and laterally to reach the most upper lateral position, then moved downward to reach the lowest position on the way back.[34] In fact, the trajectories of mass elements at the fourth plane suggest that the maximum vertical displacement corresponds to closed position and the minimum vertical displacement corresponds to maximum open position, which is in conformity with the observations in study of HS calibrated 3D imaging of human vocal fold vibration dynamics by George *et al.*[71] This shows that the 3D model reproduces successfully experimental investigated human 3D dynamics. It is also in agreement with results obtained from excised canine larynx experiments,[32] which revealed that the elliptical paths traversed by the fleshpoints of the left vocal fold over 1 cycle of vibration increased from inferior to superior (see Fig. 5 in Ref. 32). It supports the conclusion from Döllinger *et al.*[30] that during normal phonation mucosal wave propagated primarily from inferior to superior. Similar to 3D movements extracted from finite-element models,[38,39] this result can be used as an evidence to demonstrate the existence of increase of elliptical excursions along the vocal fold surface from inferior to superior. From a different perspective, the mucosal wave propagation might actually diminish the effects of impact force during the closing phase of the glottis.[40]

Additionally, the movements of mass elements within the 3DM are in the same range as the recorded displacements of human vocal folds in Baer[72] Specifically, the movement range of particle 3 (at the subglottal surface, approximately, 5 mm below the supraglottal surface, see Fig. 3(b) in Ref. 72) is about 1 mm. In the 3DM as well as the hemilarynx experiment[26] the lower mass element [at the plane $s=2$, ap-

proximately 5 mm below the most superior mass element, see Fig. 8(a)] and the lower suture point [at the layer L3, approximately 5 mm below the supraglottal surface in the hemilarynx experiments, see Fig. 8(b)] have the same movement range about 1 mm. The results can be used to reflect the rationality of the proposed 3DM relative to laryngeal experiments. Also, it confirmed the validity of earlier vocal fold physiological study in Baer.[72]

2. Displacement, velocity, and acceleration maxima during dynamics

In prior studies,[36,37,69,73–75] to determine mechanical fields on the superior surface of vocal folds as well as physical processes synthetic larynx models were proposed. As found in experiments of human hemilaryngeses, the largest inferior-superior vibratory displacements occurred in the center of the model fold [Figs. 4(a) and 4(b) in Ref. 36]. Additionally, Fig. 6 in Ref. 38 exhibited that the highest excursion of the vocal folds was at the upper parts.

In this work, the displacement maxima between the different mass elements were described in lateral, longitudinal, and vertical directions, respectively. Overall, the longitudinal displacement maxima were lower than the corresponding lateral and vertical components. The highest displacement maxima were mostly located at medial and superior planes (i.e., $s=3,4,5$). For each direction the highest values were, respectively, lateral 1.47 mm ($m_{3,3}$), longitudinal 0.46 mm ($m_{1,5}, m_{5,5}$), and vertical 1.6 mm ($m_{3,4}$). However, results obtained from human hemilarynx experiments in Refs. 26 and 19 and as well as excised canine larynx experiments in Refs. 30 and 32 and study of *in vivo* canine vocal fold dynamics in Ref. 30 indicate that lateral displacement maxima (medial-lateral) were the highest, and the longitudinal displacement maxima (anterior-posterior) were the smallest. Furthermore, previous work using a synthetic larynx model presented in Ref. 36 and excised canine larynx experiments in Ref. 34 show that vibratory displacements in the medial-lateral direction are significantly larger than in the vertical orientation (inferior-superior), see Figs. 4(c) and 4(d) in Ref. 36. Although similar to studies of finite-element models of vocal fold vibration,[38,39] in this work the vertical displacement maxima were higher than the displacement maxima in lateral and longitudinal directions. However, the vertical displacement maxima were only slightly higher than the displacement maxima in lateral direction.

Moreover, Figs. 10(a)–10(c) reveal that the largest elliptical dynamics are near the superior medial vocal fold and decrease within the inferior part. This result is in agreement with corresponding conclusions obtained from hemilarynx experiments in Ref. 19, determination of mechanical fields using a synthetic larynx model in Ref. 36, and a self-oscillating finite-element model in Ref. 76 as well as excised canine larynx experiments in Ref. 32. It can be explained by the combined effect of tissue deformation and airflow pressure generating a higher amplitude vibration than for the inferior part.

According to previous studies,[19,26,30,31] the velocity and acceleration maxima of the mucosal movements increased from inferior to superior. In our work, Fig. 10 reveals that the largest velocity and acceleration maxima were located at planes (i.e., $s=3,4$) which correspond to the region around the human vocal fold edge (Fig. 7). This result might be used to implicitly support the suggestion proposed by Spencer *et al.*[36] that *the medial surface is a free surface before collision onset, the stresses (depend on intraglottal pressure) orthogonal to that surface are small during the open phase*.

C. Applicability of the 3DM

The applicability of the 3DM was shown for six different subglottal pressure conditions. In order to demonstrate the model's applicability to various subglottal pressures P^{sub}, we made the emphasis on the comparison of the dynamic characteristics (i.e., displacement, velocity, and acceleration maxima). The discussed dynamic characteristics belong to mass elements located on the middle cross-section lying on the axis of the longitudinal symmetric 3DM predefined with the standard model parameters.

In the comparison of model dynamics in both directions (lateral/vertical) under different subglottal pressures, Figs. 11(a) and 11(b) reveal that larger subglottal pressures yield higher lateral and vertical displacements, especially within superior medial and superior part of the vocal fold. In accordance with Fig. 6 presented in Ref. 76 the vibration amplitude of the vocal fold is roughly the square root of the subglottal pressure, which was found by Titze.[77] This can be used to explain the higher displacements by increasing subglottal pressure.

The velocity and acceleration maxima are also significantly higher with larger subglottal pressure [Figs. 11(c) and 11(d)]. This result is in agreement with Spencer *et al.*,[36] which confirms that velocity is larger for higher subglottal pressure than that for lower subglottal pressure. For acceleration, a study by Jiang *et al.*[78] on canine larynges supports our findings.

Additionally, results in Fig. 11 associated with the behavior of fundamental frequencies imply that mucosal wave velocity v had no relation to fundamental frequency, unlike results obtained by Titze *et al.*[31,79] that mucosal wave velocity increases linearly with fundamental frequency. Moreover, it was found in Ref. 36 that mucosal wave velocity increased as subglottal pressure increased, while fundamental frequency was regarded as a function of subglottal pressure.[7] However, in this work the fundamental frequencies by different subglottal pressures were identical (108 Hz), except that at $P^{sub}=1.2$ kPa it was slightly smaller (104 Hz). Compared to the behaviors for $\eta=0$ in Fig. 6 in Ref. 57 and Fig. 11 in Ref. 7, our results exhibit the same kind of fundamental frequency dependence of pressure, where influence of subglottal pressure on fundamental frequency disappeared at higher pressures and small influence appeared at lower pressures. In accordance with demonstration and conclusion in prior studies,[7,57] causes of this phenomenon can be explained that the frequency of the linear oscillator depends only on the mass, the spring constant, and the damping constant, although the amplitude of the oscillation depends on driving force. The linear oscillator does not allow any variation in the frequency with the strength of the driving force, which is

directly proportional to the subglottal pressure.[57] To incorporate a variation in fundamental frequency with subglottal pressure into their model, Ishizaka *et al.*[7] introduced a nonlinear spring stiffness part depending on position with nonlinear coefficient η,[57] which is described in Eq. (A1). However, in this paper η is assumed as zero (that is linear springs) in order to neglect the influence of nonlinear spring force. This assumption might cause that the spring stiffness or fundamental frequency is independent on variation in subglottal pressure, even if the amplitude of oscillation is high. In general, this result implicitly supports the suggestion from Ishizaka *et al.*[7] that "the fundamental frequency variations with different subglottal pressure might be ascribed to two causes: One is the collision of the vocal folds at closure when the amplitude of vibration is not too large and the subglottal pressure is small. Another is the nonlinearity of the deflection of the muscles and ligaments at large amplitudes of vibration and at subglottal pressures more than several cm H_2O." In other words, the small-amplitude vocal fold oscillation at small subglottal pressure and the nonlinearity/linearity (i.e., $\eta=0$) of spring stiffness at higher subglottal pressure could cause different relationships between fundamental frequency and subglottal pressure.

Compared to the 3DM, 2D finite-element method (FEM) models[38,80] might indeed provide a more in detail representation concerning the physiological completeness.[12] Additionally, the detailed informations about the flow running through the human larynx can also be accurately simulated by FEM.[37] Nevertheless, the computation time for 1 cycle of vocal fold vibration using a fluid flow mechanical model presented in Ref. 81 takes about 15 h on one processor [AMD Opteron (TM), 1.8 MHz]. In studies on fluid-solid-acoustics interactions for human phonation,[82] the structure mechanical model of vocal folds takes about 20 days for 36 vibration cycles on four workstations (Intel Xeon 5160). The computational time for the 3DM takes currently 703–749 ms for 100 ms on a standard PC [Intel(R) Core(TM)2 Duo E8500, 3.16 GHz, 3.50 Gbyte RAM]. In comparison to FEM models, the 3DM purposefully has significantly less computational complexity, resulting in significantly higher computational speeds. As shown, by using the 3DM the mechanical properties not only from anterior to posterior but also in the vertical and lateral directions are simulated. Hence, it seems appropriate for extensive parameter studies and optimization on excised full/hemilarynx dynamics in future as performed with two or multi-mass-models so far.

V. SUMMARY

To enhance our understanding of both normal and disordered voice production, a more complete description of vocal fold dynamics may be necessary than is offered by the two-mass-model of Ishizaka and Flanagan (1972). Recent research has revealed the 3D dynamics of the medial surface of the vocal folds across a number of laboratory models, including the excised human larynx, the *in vivo* canine larynx, and physical models. The 3DM was developed to help conceptualize these data into one coherent model of vocal fold vibration. Preliminary data from the 3DM demonstrate its ability to oscillate in patterns which are similar to available experimental data. In future experiments, it is hoped that tissue parameters of experimental data (including tissue masses and tissue stiffnesses along the entire medial surface of the vocal folds) may be deduced as the parameters of the 3DM are optimized to fit its dynamics to dynamical experimental data. In this way, a firmer connection may be established between biomechanical tissue properties and the resultant vocal fold dynamics.

ACKNOWLEDGMENTS

This work was supported by DFG (*Deutsche Forschungsgemeinschaft*, German Research Foundation) Grant No. FOR 894/1 "Strömungsphysikalische Grundlagen der menschlichen Stimmgebung" (fluid physical basis of human vocalization). Dr. Berry's contributions to this investigation were supported by NIDCD (National Institute on Deafness and Other Communication Disorders) Grant No. R01 DC03072.

APPENDIX: DESCRIPTION OF APPLIED FORCES IN THE 3D MODEL

In following descriptions of the applied forces, the stiffness is defined as

$$k(x) = k(1 + \eta x^2), \tag{A1}$$

where k is assumed to be a linear constant. η is a nonlinear coefficient and can be set to 1 mm^{-2} for incorporating a nonlinear spring force into the model.[7,57] In this paper η is set to 0 mm^{-2}, which indicates that the spring stiffness is independent from displacements.

1. Anchor force

The anchor force is defined as

$$\vec{\mathbf{F}}^A_{i,s} := -k^a_{i,s} \cdot (\|\mathbf{a}_{i,s} - \mathbf{x}^r_{i,s}\|_2 - \|\mathbf{a}_{i,s} - \mathbf{x}_{i,s}\|_2) \cdot \vec{\mathbf{u}}^a_{i,s} - r^a_{i,s} \cdot (\vec{\mathbf{u}}^a_{i,s})^T \cdot \dot{\mathbf{x}}_{i,s} \cdot \vec{\mathbf{u}}^a_{i,s},$$

$$\vec{\mathbf{u}}^a_{i,s} = \frac{\mathbf{a}_{i,s} - \mathbf{x}_{i,s}}{\|\mathbf{a}_{i,s} - \mathbf{x}_{i,s}\|_2}, \tag{A2}$$

with $\mathbf{a}_{i,s}$ the 3D position of anchor; $\vec{\mathbf{u}}^a_{i,s}$ the unit vector (i.e., normalized vector) pointing from mass element position to adjacent anchor position in 3D, see Fig. 12(a); $k^a_{i,s}$ the stiffness of the vocal fold tissue; $r^a_{i,s}$ the anchor spring damping coefficient incorporating the viscous loss of the vocal fold tissue; with the damping ratio ζ_s, the damping coefficient $r^a_{i,s}$ is expressed as[6,7,83]

$$r^a_{i,s} = 2\zeta_s \sqrt{m_{i,s} k^a_{i,s}}; \tag{A3}$$

$\mathbf{x}^r_{i,s}$ the corresponding rest position of mass element $m_{i,s}$. Here, $(\cdot)^T$ denotes the transposed vector.

2. Vertical coupling force

Within the 3DM, due to the five planes arranged in vertical direction, we assume the vertical internal forces being represented by aid of vertical springs. According to the com-

putation of vertical spring force within 2DM proposed by Schwarz *et al.*,[6] the vertical spring force in the 3DM is derived after

$$\begin{aligned}\vec{\mathbf{F}}^{V}_{i,s} :=& -k^{v}_{i,s-1}\cdot(\|\mathbf{x}^{r}_{i,s-1}-\mathbf{x}^{r}_{i,s}\|_2-\|\mathbf{x}_{i,s-1}-\mathbf{x}_{i,s}\|_2)\cdot\vec{\mathbf{u}}^{v}_{i,s-1}\\ &-r^{v}_{i,s-1}\cdot(\vec{\mathbf{u}}^{v}_{i,s-1})^{T}\cdot(\dot{\mathbf{x}}_{i,s}-\dot{\mathbf{x}}_{i,s-1})\cdot\vec{\mathbf{u}}^{v}_{i,s-1}\\ &-k^{v}_{i,s}\cdot(\|\mathbf{x}^{r}_{i,s+1}-\mathbf{x}^{r}_{i,s}\|_2-\|\mathbf{x}_{i,s+1}-\mathbf{x}_{i,s}\|_2)\cdot\vec{\mathbf{u}}^{v}_{i,s+1}\\ &-r^{v}_{i,s}\cdot(\vec{\mathbf{u}}^{v}_{i,s+1})^{T}\cdot(\dot{\mathbf{x}}_{i,s}-\dot{\mathbf{x}}_{i,s+1})\cdot\vec{\mathbf{u}}^{v}_{i,s+1},\end{aligned}$$

$$\vec{\mathbf{u}}^{v}_{i,s-1}=\frac{\mathbf{x}_{i,s-1}-\mathbf{x}_{i,s}}{\|\mathbf{x}_{i,s-1}-\mathbf{x}_{i,s}\|_2},\quad \vec{\mathbf{u}}^{v}_{i,s+1}=\frac{\mathbf{x}_{i,s+1}-\mathbf{x}_{i,s}}{\|\mathbf{x}_{i,s+1}-\mathbf{x}_{i,s}\|_2},\tag{A4}$$

where $k^{v}_{i,s-1}$, $k^{v}_{i,s}$ and $r^{v}_{i,s-1}$, $r^{v}_{i,s}$ are the stiffness and damping coefficients of the vertical coupling springs connected with mass element $m_{i,s}$. The directions of unit vectors (i.e., normalized vector) $\vec{\mathbf{u}}^{v}_{i,s-1}$ and $\vec{\mathbf{u}}^{v}_{i,s+1}$ are, respectively, from position of mass element $m_{i,s}$ to position of mass element $m_{i,s-1}$ and $m_{i,s+1}$, see Fig. 12(a).

3. Longitudinal coupling force

With the stiffness $k^{l}_{i-1,s}$, $k^{l}_{i,s}$ and the damping coefficients $r^{v}_{i-1,s}$, $r^{v}_{i,s}$ of the longitudinal coupling springs connected to mass element $m_{i,s}$ the corresponding longitudinal spring force $\vec{\mathbf{F}}^{L}_{i,s}$ can be defined in accordance to Schwarz *et al.*[6] as

$$\begin{aligned}\vec{\mathbf{F}}^{L}_{i,s} :=& -k^{l}_{i-1,s}\cdot(\|\mathbf{x}^{r}_{i-1,s}-\mathbf{x}^{r}_{i,s}\|_2-\|\mathbf{x}_{i-1,s}-\mathbf{x}_{i,s}\|_2)\cdot\vec{\mathbf{u}}^{l}_{i-1,s}\\ &-r^{l}_{i-1,s}\cdot(\vec{\mathbf{u}}^{l}_{i-1,s})^{T}\cdot(\dot{\mathbf{x}}_{i,s}-\dot{\mathbf{x}}_{i-1,s})\cdot\dot{\mathbf{u}}^{l}_{i-1,s}\\ &-k^{l}_{i,s}\cdot(\|\mathbf{x}^{r}_{i+1,s}-\mathbf{x}^{r}_{i,s}\|_2-\|\mathbf{x}_{i+1,s}-\mathbf{x}_{i,s}\|_2)\cdot\vec{\mathbf{u}}^{l}_{i+1,s}\\ &-r^{l}_{i,s}\cdot(\vec{\mathbf{u}}^{l}_{i+1,s})^{T}(\dot{\mathbf{x}}_{i,s}-\dot{\mathbf{x}}_{i+1,s})\cdot\vec{\mathbf{u}}^{l}_{i+1,s},\end{aligned}$$

$$\vec{\mathbf{u}}^{l}_{i-1,s}=\frac{\mathbf{x}_{i-1,s}-\mathbf{x}_{i,s}}{\|\mathbf{x}_{i-1,s}-\mathbf{x}_{i,s}\|_2},\quad \vec{\mathbf{u}}^{l}_{i+1,s}=\frac{\mathbf{x}_{i+1,s}-\mathbf{x}_{i,s}}{\|\mathbf{x}_{i+1,s}-\mathbf{x}_{i,s}\|_2}.\tag{A5}$$

According to the assumption proposed by Wong[52] and Schwarz,[6] the stiffness $k^{l}_{i,s}$ and damping coefficient $r^{l}_{i,s}$ can be expressed in terms of stiffness and damping from the anchor spring

$$k^{l}_{i,s}=\xi_k\cdot(k^{a}_{i,s}+k^{a}_{i+1,s})\quad\text{with }\xi_k=0.2,\tag{A6}$$

$$r^{l}_{i,s}=\xi_r\cdot(r^{a}_{i,s}-r^{a}_{i+1,s})\quad\text{with }\xi_r=0.06.\tag{A7}$$

The directions of unit vectors (i.e. normalized vectors) $\vec{\mathbf{u}}^{l}_{i-1,s}$ and $\vec{\mathbf{u}}^{l}_{i+1,s}$ are from mass element $m_{i,s}$ toward mass element $m_{i-1,s}$ and $m_{i+1,s}$ in 3D, respectively, illustrated in Fig. 12(b).

4. Collision impact force

Based on the definition of collision proposed by Gunter[84] as displayed in Fig. 12(c), the collision force $\vec{\mathbf{F}}^{C}_{i,s}$ acting on mass $m_{i,s}$ in the 3DM is defined as[6]

$$\vec{\mathbf{F}}^{C}_{i,s} :=-\sum_{\forall i'}k^{c}_{i',i,s}d^{p}_{i',i,s}\vec{\mathbf{u}}^{c}_{i,s},\tag{A8}$$

$$k^{c}_{i',i,s} :=3\cdot\left(\frac{l_{i'+1,i,s}}{l_{i',i,s}+l_{i'+1,i,s}}k^{a}_{i',s}+\frac{l_{i',i,s}}{l_{i',i,s}+l_{i'+1,i,s}}k^{A}_{i'+1,s}\right),\tag{A9}$$

where $\vec{\mathbf{u}}^{c}_{i,s}$ is the unit vector for collision impact force in 3D, its corresponding direction is from the position of mass element $m_{i,s}$ to the position of the adjacent anchor, see Fig. 12(c). The stiffness $k^{c}_{i',i,s}$ is expressed according to the corresponding derivation in the 2DM.[6] i' denotes the index of mass element opposite $m_{i,s}$.

5. Driving force

The aerodynamic force, which causes the mass elements on both sides of the 3DM to vibrate, is called driving force.[6] It is produced by the glottal flow originating from the lung and complies with the Bernoulli law.[1,48,49,85,86] However, the driving force $\vec{\mathbf{F}}^{D}_{i,s}$ depends on not only the subglottal pressure P^{sub} but also the geometric dimensions, i.e., thickness h_s and vocal fold length l_g as well as the rest position $\mathbf{x}^{r}_{i,s}$.

In order to better simulate and compute the Bernoulli forces acting on the vocal folds from inferior to superior through the whole larynx, besides the simplified-method based on the derivation from the 2DM,[6,8] a so-called detailed-method is also used in the 3DM tentatively. Moreover, through the comparison between both methods the validity and corresponding accuracy of the computation for the Bernoulli forces are verified.

a. The simplified-method

In the 3DM, the simplified-method for computation of the Bernoulli driving force is presented as an extension from the 2DM. The main difference of the driving force between the 3DM and the 2DM is the Bernoulli forces that act on all five planes in 3D. Therefore, the driving pressure P_s at each plane is taken into account. Meanwhile, we assume that the subglottal pressure P^{sub} and the supraglottal pressure P^{supra} are constant, where $P^{\text{supra}}=0$ in accordance with the assumption of a jet function for no supraglottal loading.[8,69,85]

Figures 13(a) and 13(b) illustrate the convergent and divergent glottal shapes, respectively. Due to the function of the jet formed by airstream detaching from the vocal fold surface, the pressure inside the glottis is higher than zero for a convergent glottis and less than zero (or close to zero, between the flow separation point and the glottis exit) for a divergent glottis.[48,69] Therefore, in accordance with 2DM (Ref. 6) and Steinecke *et al.*,[8] the pressure P_s located at plane s within the 3DM can be defined as

$$P_s=\begin{cases}P^{\text{sub}}\cdot\left(1-\left(\dfrac{\max(0,a_{\min})}{a_s}\right)^2\right) & \text{for } s<\arg\min\limits_{s=1,\ldots,5}(a_s),\\ 0 & \text{for } s\geqslant\arg\min\limits_{s=1,\ldots,5}(a_s),\end{cases}\tag{A10}$$

$$a_{\min} = \min(a_1, a_2, a_3, a_4, a_5), \tag{A11}$$

where a_s is the glottal area with $s = 1, \ldots, 5$.

The driving force acting on every mass element can be resolved into two components: horizontal component $\vec{\mathbf{F}}_{i,s}^{D,\mathrm{hor}}$ and vertical component $\vec{\mathbf{F}}_{i,s}^{D,\mathrm{ver}}$. Following this analysis, the driving force within the 3DM is defined as

$$\vec{\mathbf{F}}_{i,s}^{D} = [\vec{\mathbf{F}}_{i,s}^{D,\mathrm{hor}}, \quad \vec{\mathbf{F}}_{i,s}^{D,\mathrm{ver}}]^T, \tag{A12}$$

$$\vec{\mathbf{F}}_{i,s}^{D,\mathrm{hor}} = P_s \cdot w_{i,s} \cdot h_s \cdot \vec{\mathbf{u}}_{i,s}^{d}, \tag{A13}$$

$$\vec{\mathbf{F}}_{i,s}^{D,\mathrm{ver}} = \begin{cases} P^{\mathrm{sub}} \cdot w_{i,s} \cdot \sqrt{(x_{i,s}^r - x_{i,s})^2 + (y_{i,s}^r - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s = 1, \quad |x_{i,s}^r| < |x_{i,s}|, \\ 0 & \text{if } s = 1, \quad |x_{i,s}^r| \geqslant |x_{i,s}|, \\ P_{s-1} \cdot w_{i,s} \cdot \sqrt{(x_{i,s-1} - x_{i,s})^2 + (y_{i,s-1} - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s \neq 1, \quad |x_{i,s-1}| > |x_{i,s}|, \\ 0 & \text{if } s \neq 1, \quad |x_{i,s-1}| \leqslant |x_{i,s}|, \end{cases} \tag{A14}$$

where $x_{i,s-1}$, $y_{i,s-1}$ and $x_{i,s}$, $y_{i,s}$ as well as $x_{i,s}^r$, $y_{i,s}^r$ are the x- and y-components of $\mathbf{x}_{i,s-1}$, $\mathbf{x}_{i,s}$, $\mathbf{x}_{i,s}^r$. Additionally, h_s indicates the thickness of mass element $m_{i,s}$, and $\vec{\mathbf{u}}_{i,s}^{d}$ is the unit vector for the horizontal component of driving force. The direction of the unit vectors is perpendicular to the vertical projection of $w_{i,s}$ at the horizontal plane. They point from inside the glottis towards outside the glottis, see Fig. 12(d). $\vec{\mathbf{u}}_z$ is the unit vector along the vertical axis.

In order to reasonably distribute the driving pressure acting on individual mass elements within the 3DM, an effective side length $w_{i,s}$ is used.[6] During phonation the position of collision point will be taken into account for the calculation of $w_{i,s}$, if the adjacent collision point occurs between mass element $m_{i,s}$ and mass element $m_{i-1,s}$ or $m_{i+1,s}$. Otherwise, $w_{i,s}$ is defined from the middle point of the adjacent longitudinal springs, i.e., $w_{i,s} = 0.5 \cdot \|\mathbf{x}_{i-1,s} - \mathbf{x}_{i+1,s}\|_2$, see Fig. 12(d).

b. The detailed-method

The detailed-method is based on an approximate pressure distribution along the glottal flow.[7,66,87] Owing to the small dimensions of the human glottis and the high velocity of glottal flow, we can assume the glottal flow to be quasi-steady.[8] By using the Bernoulli equation for one-dimensional flow, the driving pressure distribution along the glottal flow can be obtained,[7] which is schematically illustrated in Fig. 13(c).

Pressure $P_{1,a}$ locates at the inferior glottal entrance corresponding to the lower edge of mass element $m_{i,1}$. According to the Bernoulli equation,[7,8,18] the approximate pressure distribution is described by

$$P^{\mathrm{sub}} - P_{i,a} = \frac{\rho}{2} \cdot \left(\frac{U_g}{a_1} \right)^2, \tag{A15}$$

$$P_{s,a} - P_{s,b} = 12 \cdot \frac{\mu l_g^2 h_s}{a_s^3} \cdot U_g + \frac{\rho h_s}{a_s} \cdot \frac{dU_g}{dt}, \tag{A16}$$

$$P_{s+1,a} - P_{s,b} = \frac{\rho}{2} \cdot U_g^2 \cdot \left(\frac{1}{a_s^2} - \frac{1}{a_{s+1}^2} \right), \tag{A17}$$

where $P_{s,a}$, $P_{s,b}$ are the pressures at the lower and upper edges of mass element $m_{i,s}$ located at plane s. Additionally, ρ and μ are the density and viscosity coefficients, of air, respectively. Due to the simplification of multi-mass-model as well as the small value of the air viscous losses (approximately 1.87×10^{-7} g/cm ms), the air viscosity coefficient μ is predefined to zero in Eq. (A16). Hence, the consistency within the model development (in accordance with prior studies)[1,6,8,12] can be retained. l_g is the vocal fold length, and a_s is the cross-sectional glottal area at different planes ($s = 1, \ldots, 5$). Moreover, by using the above-proposed equations (A15)–(A17), the pressures $P_{s,a}$, $P_{s,b}$ at each plane can be derived. The volume flow rate of glottal airflow is U_g, which is derived in accordance with the assumption in Ref. 8 as follows:

$$U_g = \sqrt{\frac{2P^{\mathrm{sub}}}{\rho}} a_{\min}. \tag{A18}$$

The mean value of the pressure acting on the mass element located at plane s is given by

$$P_{s,m} = \tfrac{1}{2}(P_{s,a} + P_{s,b}). \tag{A19}$$

Bernoulli's law is less applicable to the divergent glottis in which the glottal pressure is less than zero.[48] In order to ensure the feasibility of Bernoulli's law pressure $P_{s,m}$ is set to zero, if $P_{s,m}$ derived in Eq. (A19) is negative.

Therefore, the horizontal component $\vec{\mathbf{F}}_{i,s}^{D,\mathrm{hor}}$ and vertical component $\vec{\mathbf{F}}_{i,s}^{D,\mathrm{ver}}$ of the driving force equation (A12) can be obtained in accordance to Lucero[18] and Koizumi[66] as

$$\vec{\mathbf{F}}_{i,s}^{D,\mathrm{hor}} = P_{s,m} \cdot w_{i,s} \cdot h_s \cdot \vec{\mathbf{u}}_{i,s}^{d}, \tag{A20}$$

$$\vec{\mathbf{F}}_{i,s}^{D,\mathrm{ver}} = \begin{cases} P^{\mathrm{sub}} \cdot w_{i,s} \cdot \sqrt{(x_{i,s}^{r} - x_{i,s})^2 + (y_{i,s}^{r} - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s = 1 \\ P_{s-1,b} \cdot w_{i,s} \cdot \sqrt{(x_{i,s-1} - x_{i,s})^2 + (y_{i,s-1} - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s \neq 1, \quad s < \arg \min_{s=1,\ldots,5} (a_s) \\ 0 & \text{if } s \neq 1, \quad s \geqslant \arg \min_{s=1,\ldots,5} (a_s), \end{cases} \tag{A21}$$

where $P_{s-1,b}$ indicates the pressure at the upper edge of mass element $m_{i,s-1}$ at plane $s-1$, and a_s denotes the cross sectional glottal area at different planes $(s=1,\ldots,5)$.

In order to simulate the 3D dynamics of human vocal folds, according to Eq. (1), the differential system of the new proposed model is solved by using the classical Runge–Kutta algorithm with the increment of 0.25 ms.[1,2]

[1] M. Döllinger, U. Hoppe, F. Hettlich, J. Lohscheller, S. Schubert, and U. Eysholdt, "Vibration parameter extraction from endoscopic image series of the vocal folds," IEEE Trans. Biomed. Eng. **49**, 773–781 (2002).
[2] R. Schwarz, U. Hoppe, M. Schuster, T. Wurzbacher, U. Eysholdt, and J. Lohscheller, "Classification of unilateral vocal fold paralysis by endoscopic digital high-speed recordings and inversion of a biomechanical model," IEEE Trans. Biomed. Eng. **53**, 1099–1108 (2006).
[3] U. Eysholdt, F. Rosanowski, and U. Hoppe, "Irregular vocal fold vibrations caused by different types of laryngeal asymmetry," Eur. Arch. Otorhinolaryngol. **260**, 412–417 (2003).
[4] L. Sulica and A. Blitzer, "Vocal fold paresis: Evidence and controversies," Curr. Opin. Otolaryngol. Head Neck Surg. **15**, 159–162 (2007).
[5] N. Isshiki, M. Tanabe, K. Ishizaka, and D. Broad, "Clinical significance of asymmetrical vocal cord tension," Ann. Otol. Rhinol. Laryngol. **86**, 58–66 (1977).
[6] R. Schwarz, M. Döllinger, T. Wurzbacher, U. Eysholdt, and J. Lohscheller, "Spatiotemporal quantification of vocal fold vibrations using high-speed videoendoscopy and a biomechanical model," J. Acoust. Soc. Am. **123**, 2717–2732 (2008).
[7] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Tech. J. **51**, 1233–1268 (1972).
[8] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal fold model," J. Acoust. Soc. Am. **97**, 1571–1578 (1995).
[9] P. Mergell, H. Herzel, and I. Titze, "Irregular vocal-fold vibration high-speed observation and modeling," J. Acoust. Soc. Am. **108**, 2996–3002 (2000).
[10] M. Döllinger, T. Braunschweig, J. Lohscheller, U. Eysholdt, and U. Hoppe, "Normal voice production: Computation of driving parameters from endoscopic digital high speed images," Methods Inf. Med. **42**, 271–276 (2003).
[11] T. Wurzbacher, R. Schwarz, M. Döllinger, U. Hoppe, U. Eysholdt, and J. Lohscheller, "Model-based classification of nonstationary vocal fold vibrations," J. Acoust. Soc. Am. **120**, 1012–1027 (2006).
[12] T. Wurzbacher, M. Döllinger, R. Schwarz, U. Hoppe, U. Eysholdt, and J. Lohscheller, "Spatiotemporal classification of vocal fold dynamics by a multi mass model comprising time-dependent parameters," J. Acoust. Soc. Am. **123**, 2324–2334 (2008).
[13] S. Adachi and J. Yu, "Two-dimensional model of vocal fold vibration for sound synthesis of voice and soprano singing," J. Acoust. Soc. Am. **117**, 3213–3224 (2005).
[14] Y. Zhang and J. J. Jiang, "Chaotic vibrations of a vocal fold model with a unilateral polyp," J. Acoust. Soc. Am. **115**, 1266–1269 (2004).
[15] I. R. Titze and B. H. Story, "Rules for controlling low-dimensional vocal fold models with muscle activation," J. Acoust. Soc. Am. **112**, 1064–1076 (2002).
[16] I. R. Titze, "Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation," J. Speech Lang. Hear. Res. **49**, 439–447 (2006).
[17] J. D. Sitt, A. Amador, F. Goller, and G. B. Mindlin, "Dynamical origin of spectrally rich vocalizations in birdsong," Phys. Rev. E **78**, 011905 (2008).
[18] J. C. Lucero and L. L. Koenig, "Simulations of temporal patterns of oral airflow in men and women using a two-mass model of the vocal folds under dynamic control," J. Acoust. Soc. Am. **117**, 1362–1372 (2005).
[19] M. Döllinger, D. A. Berry, and D. W. Montequin, "The influence of epilarynx area on vocal fold dynamics," Otolaryngol.-Head Neck Surg. **135**, 724–729 (2006).
[20] I. R. Titze, "Parametrization of the glottal area, glottal flow, and vocal fold contact area," J. Acoust. Soc. Am. **75**, 570–580 (1984).
[21] M. Döllinger and D. A. Berry, "Visualization and quantification of the medial surface dynamics of an excised human vocal fold during phonation," J. Voice **20**, 401–413 (2006).
[22] I. R. Titze, "The physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536–1552 (1988).
[23] I. R. Titze, S. S. Schmidt, and M. R. Titze, "Phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **97**, 3080–3084 (1995).
[24] R. W. Chan and I. R. Titze, "Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics," J. Acoust. Soc. Am. **119**, 2351–2362 (2006).
[25] D. A. Berry, M. J. O. Clark, D. W. Montequin, and I. R. Titze, "Characterization of the medial surface of the vocal folds," Ann. Otol. Rhinol. Laryngol. **110**, 470–477 (2001).
[26] A. Boessenecker, D. A. Berry, J. Lohscheller, U. Eysholdt, and M. Döllinger, "Mucosal wave properties of a human vocal fold," Acta. Acust. Acust. **93**, 815–823 (2007).
[27] M. Döllinger and D. A. Berry, "Computation of the three-dimensional medial surface dynamics of the vocal folds," J. Biomech. **39**, 369–374 (2006).
[28] M. Döllinger, N. Tayama, and D. A. Berry, "Empirical eigenfunctions and medial surface dynamics of a human vocal fold," Methods Inf. Med. **44**, 384–391 (2005).
[29] M. Döllinger, D. A. Berry, and G. S. Berke, "Medial surface dynamics of an in vivo canine vocal fold during phonation," J. Acoust. Soc. Am. **117**, 3174–3183 (2005).
[30] M. Döllinger, D. A. Berry, and G. S. Berke, "A quantitative study of the medial surface dynamics of an in vivo canine vocal fold during phonation," Laryngoscope **115**, 1646–1654 (2005).
[31] I. R. Titze, J. J. Jiang, and T. Y. Hsiao, "Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration," Ann. Otol. Rhinol. Laryngol. **102**, 58–63 (1993).
[32] D. A. Berry, D. W. Montequin, and N. Tayama, "High-speed digital imaging of the medial surface of the vocal folds," J. Acoust. Soc. Am. **110**, 2539–2547 (2001).
[33] C. G. Garrett, J. R. Coleman, and L. Reinisch, "Comparative histology and vibration of the vocal folds: Implications for experimental studies in microlaryngeal surgery," Laryngoscope **110**, 814–824 (2000).
[34] J. Kobayashi, E. Yumoto, M. Hyodo, and K. Gyo, "Two-dimensional analysis of vocal fold vibration in unilaterally atrophied larynges," Laryngoscope **110**, 440–446 (2000).
[35] Z. Zhang, J. Neubauer, and D. A. Berry, "Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds," J. Acoust. Soc. Am. **120**, 2841–2849 (2006).
[36] M. Spencer, T. Siegmund, and L. Mongeau, "Determination of superior surface strains and stresses, and vocal fold contact pressure in a synthetic larynx model using digital image correlation," J. Acoust. Soc. Am. **123**, 1089–1103 (2008).
[37] S. Becker, S. Kniesburges, S. Müller, A. Delgado, G. Link, M. Kaltenbacher, and M. Döllinger, "Flow-structure-acoustic interaction in a human voice model," J. Acoust. Soc. Am. **125**, 1351–1361 (2009).
[38] F. Alipour-Haghighi, D. A. Berry, and I. R. Titze, "A finite-element model of vocal fold vibration," J. Acoust. Soc. Am. **108**, 3003–3012 (2000).
[39] D. A. Berry, H. Herzel, I. R. Titze, and K. Krischer, "Interpretation of biomechanical simulations of normal and chaotic vocal fold vibrations with empirical eigenfunctions," J. Acoust. Soc. Am. **95**, 3595–3604 (1994).

[40]A. Sonninen and A.-M. Laukkanen, "Hypothesis of whiplike motion as a possible traumatizing mechanism in vocal fold vibration," Folia Phoniatr Logop **55**, 189–198 (2003).
[41]J. Lohscheller, U. Eysholdt, H. Toy, and M. Döllinger, "Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2-d diagrams for visualizing and analyzing the underlying laryngeal dynamics," IEEE Trans. Med. Imaging **27**, 300–309 (2008).
[42]O. Rasp, J. Lohscheller, M. Döllinger, U. Eysholdt, and U. Hoppe, "The pitch rise paradigm: A new task for real-time endoscopy of non-stationary phonation," Folia Phoniatr Logop **58**, 175–185 (2006).
[43]A. M. Sulter, H. K. Schutte, and D. G. Miller, "Standardized laryngeal videostroboscopic rating: Differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency and age," J. Voice **10**, 175–189 (1996).
[44]P. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. V. D. Heyning, M. Remacle, V. Woisard, and Committee on Phoniatrics of the European Laryngological Society (ELS), "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques guideline elaborated by ELS," Eur. Arch. Otorhinolaryngol. **258**, 77–82 (2001).
[45]R. Schwarz, "Model-based quantification of pathological voice production," in *Kommunikationsstörungen, Berichte aus Phoniatrie und Pädaudiologie (Dysphonia, Reports from Phoniatrics and Pedaudiology)*, edited by U. Eysholdt (Shaker, Aachen, 2007).
[46]T. Wurzbacher, "Vocal fold dynamics-quantification and model-based classification," in *Kommunikationsstörungen, Berichte aus Phoniatrie und Pädaudiologie (Dysphonia, Reports from Phoniatrics and Pedaudiology)*, edited by U. Eysholdt (Shaker, Aachen, 2007).
[47]B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**, 1249–1260 (1995).
[48]I. R. Titze, *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ, 1994).
[49]J. L. Flanagan and K. Ishizaka, "Computer model to characterize the air volume displaced by the vibrating vocal cords," J. Acoust. Soc. Am. **63**, 1559–1565 (1978).
[50]M. C. Su, T. H. Yeh, C. T. Tan, C. D. Lin, O. C. Linne, and S. Y. Lee, "Measurement of adult vocal fold length," J. Laryngol. Otol. **116**, 447–449 (2002).
[51]H. Gray, in *Anatomy of the Human Body*, 20th ed., edited by W. H. Lewis (Lea & Febiger, Philadelphia, 1918).
[52]D. Wong, M. Ito, N. B. Cox, and I. R. Titze, "Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases," J. Acoust. Soc. Am. **89**, 383–394 (1991).
[53]J. J. Jiang and I. R. Titze, "A methodological study of hemilaryngeal phonation," Laryngoscope **103**, 872–882 (1993).
[54]J. C. Lucero and L. L. Koenig, "On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds," J. Acoust. Soc. Am. **121**, 3280–3283 (2007).
[55]R. W. Chan, I. R. Titze, and M. R. Titze, "Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **101**, 3722–3727 (1997).
[56]G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide (Computer Science and Scientific Computing)* (Academic, San Francisco, CA, 1996).
[57]L. P. Fulcher, R. C. Scherer, A. Melnykov, V. Gateva, and M. E. Limes, "Negative coulomb damping, limit cycles, and self-oscillation of the vocal folds," Am. J. Phys. **74**, 386–393 (2006).
[58]U. Hoppe, F. Rosanowski, M. Döllinger, J. Lohscheller, and U. Eysholdt, "Visualization of the laryngeal motorics during a glissando," J. Voice **17**, 370–376 (2003).
[59]J. Lohscheller and U. Eysholdt, "Phonovibrogram visualization of entire vocal fold dynamics," Laryngoscope **118**, 753–758 (2008).
[60]Y. Yan, K. Ahmad, M. Kunduk, and D. Bless, "Analysis of vocal-fold vibrations from high-speed laryngeal images using a hilbert transform-based methodology," J. Voice **19**, 161–175 (2005).
[61]T. Braunschweig, J. Flaschka, P. Schelhorn-Neise, and M. Döllinger, "High-speed video analysis of the phonation onset, with an application to the diagnosis of functional dysphonias," Med. Eng. Phys. **30**, 59–66 (2008).
[62]T. Braunschweig, P. Schelhorn-Neise, and M. Döllinger, "Diagnosis of functional voice disorders by using the high speed recording technics," Laryngorhinootologie **87**, 323–330 (2008).
[63]D. D. Deliyski, P. P. Petrushev, H. S. Bonilha, T. T. Gerlach, B. Martin-Harris, and R. E. Hillman, "Clinical implementation of laryngeal high speed videoendoscopy: Challenges and evolution," Folia Phoniatr Logop **60**, 33–44 (2008).
[64]M. Döllinger, "Parameter estimation of vocal fold dynamics by inversion of a biomechanical model," in *Kommunikationsstörungen, Berichte aus Phoniatrie und Pädaudiologie, Band 10 (Dysphonia, Reports from Phoniatrics and Pedaudiology)*, edited by U. Eysholdt (Shaker, Aachen, 2002).
[65]S. Y. Lowell and B. H. Story, "Simulated effects of cricothyroid and thyroarytenoid muscle activation on adult-male vocal fold vibration," J. Acoust. Soc. Am. **120**, 386–397 (2006).
[66]T. Koizumi, S. Taniguchi, and S. Hiromitsu, "Two-mass models of the vocal cords for natural sounding voice synthesis," J. Acoust. Soc. Am. **82**, 1179–1192 (1987).
[67]X. Zheng, S. Bielamowicz, H. Luo, and R. Mittal, "A computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation," Ann. Biomed. Eng. **37**, 625–641 (2009).
[68]J. J. Jiang and I. R. Titze, "Measurement of vocal fold intraglottal pressure and impact stress," J. Voice **8**, 132–144 (1994).
[69]S. L. Thomson, L. Mongeau, and S. H. Frankel, "Aerodynamic transfer of energy to the vocal folds," J. Acoust. Soc. Am. **118**, 1689–1700 (2005).
[70]I. T. Tokuda, J. Horáček, J. G. Švec, and H. Herzel, "Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments," J. Acoust. Soc. Am. **122**, 519–531 (2007).
[71]N. A. George, F. F. M. de Mul, Q. Qiu, G. Rakhorst, and H. K. Schutte, "Depth-kymography: High-speed calibrated 3d imaging of human vocal fold vibration dynamics," Phys. Med. Biol. **53**, 2667–2675 (2008).
[72]T. Baer, "Observations of vocal fold vibration: Measurement of excised larynges," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo, 1981), Chap. 10, pp. 119–133.
[73]J. B. Park and L. Mongeau, "Experimental investigation of the influence of a posterior gap on glottal flow and sound," J. Acoust. Soc. Am. **124**, 1171–1179 (2008).
[74]J. B. Park and L. Mongeau, "Instantaneous orifice discharge coefficient of a physical, driven model of the human larynx," J. Acoust. Soc. Am. **121**, 442–455 (2007).
[75]J. S. Drechsel and S. L. Thomson, "Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model," J. Acoust. Soc. Am. **123**, 4434–4445 (2008).
[76]C. Tao, J. J. Jiang, and Y. Zhang, "Simulation of vocal fold impact pressures with a self-oscillating finite-element model," J. Acoust. Soc. Am. **119**, 3987–3994 (2006).
[77]I. R. Titze, "On the relation between subglottal pressure and fundamental frequency in phonation," J. Acoust. Soc. Am. **85**, 901–906 (1989).
[78]J. J. Jiang, A. G. Shah, M. M. Hess, K. Verdolini, F. M. Banzali, and D. G. Ganson, "Vocal fold impact stress analysis," J. Voice **15**, 4–14 (2001).
[79]I. R. Titze, "Phonation threshold pressure: A missing link in glottal aerodynamics," J. Acoust. Soc. Am. **91**, 2926–2935 (1992).
[80]H. E. Gunter, "A mechanical model of vocal-fold collision with high spatial and temporal resolution," J. Acoust. Soc. Am. **113**, 994–1000 (2003).
[81]H. Luo, R. Mittal, X. Zheng, S. A. Bielamowicz, R. J. Walsh, and J. K. Hahn, "An immersed-boundary method for flowstructure interaction in biological systems with application to phonation," J. Comput. Phys. **227**, 9303–9332 (2008).
[82]G. Link, M. Kaltenbacher, M. Breuer, and M. Döllinger, "A 2D finite-element scheme for fluid-solid-acoustic interactions and its application to human phonation," Comput. Methods Appl. Mech. Eng. **198**, 3321–3334 (2008).
[83]J. L. Flanagan and L. Landgraf, "Self-oscillating source for vocal-tract synthesizers," IEEE Trans. Audio Electroacoust. **16**, 57–64 (1968).
[84]H. E. Gunter, "Modeling mechanical stresses as a factor in the etiology of benign vocal fold lesions.," J. Biomech. **37**, 1119–1124 (2004).
[85]M. Little, P. McSharry, I. Moroz, and S. Roberts, "A simple, quasi-linear, discrete model of vocal fold dynamics," in *Nonlinear Analyses and Algorithms for Speech Processing*, edited by M. Faundez-Zanuy, L. Janer, A. Esposito, A. Satue-Villar, J. Roure, and V. Espinosa-Duro (Springer, Berlin, 2005), Vol. **3817**, pp. 348–356.
[86]M. P. de Vries, H. K. Schutte, A. E. P. Veldman, and G. J. Verkerke, "Glottal flow through a two-mass model: Comparison of Navier-Stokes solutions with simplified models," J. Acoust. Soc. Am. **111**, 1847–1853 (2002).
[87]T. Ikeda and Y. Matsuzaki, "A two-dimensional vocal cord model for heavy voice," Trans. Jpn. Soc. Mech. Eng., Ser. C **63**, 693–699 (1997).

Objective estimation of tracheoesophageal speech ratings using an auditory model[a)]

Robert McDonald and Vijay Parsa[b)]
Department of Electrical and Computer Engineering, and National Centre for Audiology, University of Western Ontario, London, N6A 4B8 Ontario, Canada

Philip C. Doyle
Voice Production and Perception Laboratory, Department of Otolaryngology and Doctoral Program in Rehabilitation Sciences, University of Western Ontario, London, N6A 4B8 Ontario, Canada

(Received 27 June 2008; revised 4 November 2009; accepted 11 November 2009)

Total laryngectomy is often the treatment of choice for many individuals diagnosed with advanced laryngeal cancer. This procedure alters the normal voice production mechanism, and tracheoesophageal (TE) speech is one alternative method of voicing postlaryngectomy. TE speech is created when pulmonary air is passed through the upper esophagus to create a vibratory source that is then articulated into speech. TE speech is often characterized by abnormal voice quality. Acoustic analysis of TE speech has the potential of quantifying the voice quality and assisting the speech language pathologist in facilitating rehabilitation. Motivated in part by the recent advances in telecommunication industry for speech quality estimation, this paper investigated the application of an auditory model in predicting the ratings of TE speech by normal hearing listeners. The Moore–Glasberg auditory model was employed to extract perceptually relevant features from the acoustic waveform, and these features were later combined to estimate the subjective ratings of TE speech. This approach was validated with a database of subjective ratings of speech samples recorded from 35 TE speakers. Results showed moderate correlations between the objective metrics and the subjective ratings, and these correlations were significantly better than those obtained with traditional methods used in the telecommunication applications.
 [DOI: 10.1121/1.3270396]

PACS number(s): 43.70.Dn, 43.70.Fq, 43.70.Jt, 43.72.Ar [AL] Pages: 1032–1041

I. INTRODUCTION

Tracheoesophageal (TE) speech is the outcome of a surgical-prosthetic method of voice restoration following surgical removal of the entire larynx. TE puncture voice restoration involves the surgical creation of a midline puncture in the common wall between the trachea and the esophagus. This puncture is then stented with a small one-way valved prosthesis. When the airway is sealed, the TE puncture prosthesis permits pulmonary (lung) air to flow from the airway into the esophageal reservoir.[1] The TE voice prosthesis serves only as a conduit for airflow; it does not create any voice source. Once air fills the esophagus, it sets the muscular tissue of the upper esophagus and lower pharynx into vibration, and this intrinsic, alaryngeal voice source is transmitted into the vocal tract where it is articulated into speech. Because the structures responsible for TE voice production are often anatomically asymmetrical and do not behave as an adductory-abductory system like that of the larynx, the generated sound source is non-normal in its quality. In particular, the fundamental frequency of TE voice signals is lower than the normal controls, especially for female talkers,[2] and the formants are generally higher for male TE speakers.[3] In addition, TE speech is characterized by normal or slightly greater than normal intensity, and slightly longer pause duration than normal speech.[2] Further more, the overall sound quality of TE speech is best described as highly aperiodic, rough, and noisy. Depending on the surgery and the length of rehabilitation, considerable variability across TE speakers does exist.[4,5] Thus the assessment of the perceived quality of TE speech is essential in judging the severity of the TE speech production and in monitoring the rehabilitation progress.

Perceptual evaluation of abnormal voice quality (both TE and laryngeal modes of production) has received considerable attention in the voice literature. Researchers have investigated and continue to investigate the effectiveness of different voice quality dimensions, the rating scales used, and the impact of listeners' background, experience, and exposure to abnormal voices. Kreiman *et al.*[6,7] systematically evaluated a number of issues related to perceptual measurements of pathological voice quality ratings. An interesting result from their work is the considerable disagreement among expert listeners on the ratings of voice quality. Kreiman *et al.*[7] attributed this to expert listeners' inherent biases shaped by their own experience with abnormal voices. Naive adult listeners, on the other hand, base their judgments of abnormal voice quality based on their internal standard of normal voice quality. Since this internal standard is relatively

[a)]Portions of this paper were presented at the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas, April 2–5, 2008.
[b)]Author to whom correspondence should be addressed. Electronic mail: parsa@nca.uwo.ca

stable and well established due to continuous exposure to normal voices, the quality ratings from naive listeners are more reliable. In addition, quality ratings from naive listeners are also indicative of how the abnormal voices are perceived in "real world" and thus deserve special attention.

While perceptual ratings of speech quality are the "gold standard," they require large groups of listeners and consume valuable time and resources. Instrumental measurements of voice are therefore preferable and these include acoustic, electroglottographic, aerodynamic, and videostroboscopic methods. Of these, acoustic measurements of voice are appealing due to their non-invasiveness and cost-effectiveness. Measurements of frequency and amplitude perturbations ("jitter" and "shimmer," respectively) and glottal noise have been computed from the acoustic waveform, and used to characterize the disordered voice. There is a substantial body of literature on the effectiveness of these acoustic measurements in analyzing voice signals collected from laryngeal pathologies. The reader is referred to a recent review article by Schoentgen[8] for a detailed discussion and analysis. The literature is sparse, however, on the potential of applying these acoustic measures for TE speech analysis, and a brief summary of relevant research articles is presented below.

Kazi *et al.*[3] applied the linear prediction (LP) analysis to sustained vowel samples collected from 50 TE speakers. LP analysis was used to identify the formant frequencies and to investigate the impact of such factors related to TE surgery as tumor stage, chemoradiotherapy, pharyngectomy, and postoperative complications. Results from this study indicated a significant shift in the formant frequencies for the TE vowels in comparison with the normal control group. Although it was surmised that this shift in formants has an impact on TE speech quality, no data were presented to support this assumption. In a different study, Kazi *et al.*[9] performed acoustic and perceptual analyses of sustained vowel and continuous speech samples collected for ten female TE speakers, ten male TE speakers, and ten female normal controls. Results from this study showed that there were significant differences in the jitter, shimmer, and glottal noise parameters between the female TE speakers and their normal cohort. However, the perceptual ratings reported in the study were carried out by only two expert otolaryngologists and no correlation between acoustic and perceptual measures was performed.

Van Gogh *et al.*[10] evaluated the perturbation and glottal noise measures with a sustained vowel database collected from 66 TE speakers. Their analysis revealed that only 52 out of the 66 voices were suitable for acoustic analysis due to the highly aperiodic nature in the discarded voice samples. For the selected voice samples, correlation analysis between the perceptual and acoustic measures revealed coefficients between 0.62 and 0.69 for rating scales directly related to voice quality. Most *et al.*[11] conducted an acoustic-perceptual study of three groups of speakers: laryngeal, tracheoesophageal, and esophageal. Acoustic analysis included measurement of jitter, shimmer, and maximum phonation time. Perceptual measurements included ratings of acceptability of the voice signal and speech intelligibility. Correlation analysis between the measures from these two modalities ranged between 0.74 and 0.82. However, there were only five TE speakers in this database and no meaningful conclusions can be drawn on the relationship between the acoustic and perceptual descriptors for the TE speech samples alone.

Using a database of 113 short voice and speech samples collected from 68 patients, Moerman *et al.*[12] investigated the relationship between acoustic and perceptual measurements. Acoustic parameters included jitter and measurements of voice regularity (percentage of voiced frames, average voicing evidence, and duration of voicing). Perceptual ratings of speech quality ("overall impression") were obtained from ten speech language pathology students. Correlation analysis between the objective and subjective data resulted in coefficients in the range 0.3–0.5.

Van As-Brooks *et al.*[13] evaluated sustained vowel samples collected from 39 TE speakers. The extracted acoustical parameters were median fundamental frequency, jitter, harmonics to noise ratio, and spectral tilt. Perceptual ratings were collected from four trained speech language pathologists where the TE speech sample was rated as either "good," "reasonable," or "poor." Thus this study forced the perceptual ratings to be a classification system rather than a continuous rating scale. As such, no correlation analysis was performed, but it was stated that the HNR and tilt parameters allow for similar classification objectively.

Schuster *et al.*[14] undertook an interestingly different approach to the evaluation of TE speech samples. These researchers investigated the application of automatic speech recognition systems in predicting the quality and intelligibility of TE speech. Using a database of 18 TE speech samples, they found a correlation of 0.84 between objective and subjective scores of intelligibility. It is not clear how these results will correspond to perceptual measures of TE speech quality. Moreover, only five raters were used to generate the subjective scores, which diminishes the statistical strength of the obtained results.

It is clear from the above literature review that a majority of previous investigations have utilized sustained vowel samples and extracted measures such as jitter, shimmer, and glottal noise level. However, judgments of the quality of a speaker's voice in real world are based on listeners' perception of running speech, and acoustic measures extracted from continuous speech samples may well be required for accurate prediction of perceptual judgments of abnormal voice quality that are relevant to the clients' life experience. The traditional jitter, shimmer, and harmonics-to-noise ratio measures are not valid for continuous speech samples, and an alternative approach to characterizing the voice quality of speech samples acquired from TE speakers is necessary.

Objective speech quality measurement is also an important topic in telecommunications and audio processing research. For example, the distortion introduced by the speech coding and decoding algorithms can be quantified using a set of distance metrics such as the Itakura–Saito (IS) measure, log-likelihood ratio (LLR) measure, and the cepstral distance (CD) measure.[15] These metrics are computed on a frame-by-frame basis using the following equations:

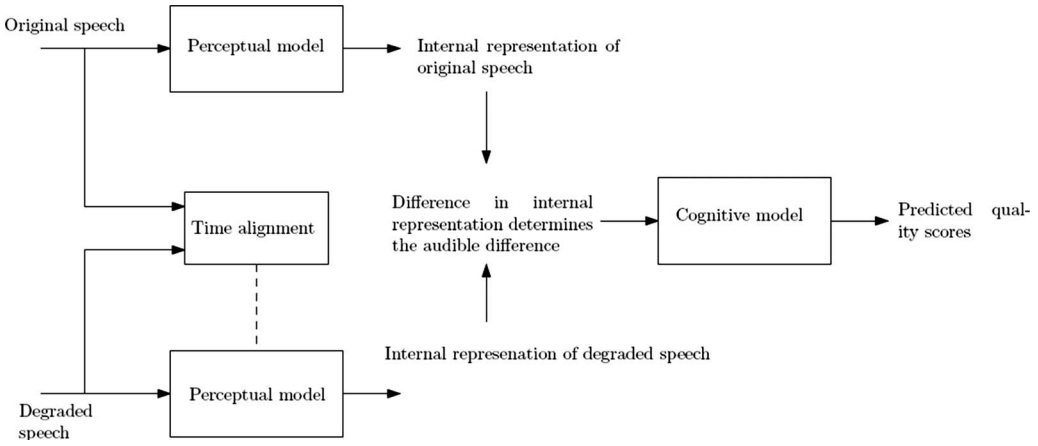


FIG. 1. This block diagrams depicts the conceptual underpinnings of the PESQ measurement procedure standardized by the (Refs. 18 and 19).

$$\mathrm{IS} = \frac{\sigma_o^2}{\sigma_d^2}\frac{\mathbf{a_d}\mathbf{R_o}\mathbf{a_d^T}}{\mathbf{a_o}\mathbf{R_o}\mathbf{a_o^T}} + \log\left(\frac{\sigma_o^2}{\sigma_d^2}\right) - 1, \tag{1}$$

$$\mathrm{LLR} = \log\left(\frac{\mathbf{a_d}\mathbf{R_o}\mathbf{a_d^T}}{\mathbf{a_o}\mathbf{R_o}\mathbf{a_o^T}}\right), \tag{2}$$

$$\mathrm{CD} = \frac{10\sqrt{2}}{\log 10}\sqrt{\sum_{k=1}^{P}[c_o(k) - c_d(k)]^2}, \tag{3}$$

where **a** is the linear prediction coefficient (LPC) vector, **R** is the auto-correlation matrix, σ^2 is the LPC gain, $c(k)$ are the cepstral coefficients, and P is the LPC order.[15] The subscripts o and d on these parameters represent the original and degraded speech frames, respectively. There is evidence that these distortion metrics are also useful in objective characterization of pathological voices.[16,17] For example, Gu *et al.*[17] evaluated the performance of IS and LLR metrics using a database of quality ratings of speech samples collected from ten speakers with Parkinson's disease. They found that the IS metric correlated highly with the subjective ratings and concluded that the IS metric is more suitable for disordered voice evaluation than the LLR metric.

However, recent advances in measuring the quality of speech processed by various speech and audio coding algorithms,[18–22] and the quality of speech transmitted over wireless and Internet channels,[23] have so far not transcended into the evaluation of the disordered voice quality. The newer speech quality evaluation procedures perform significantly better than the IS and LLR metrics due to the inclusion of models of normal audition. As an example, Fig. 1 depicts the block diagram of the perceptual evaluation of speech quality (PESQ) measurement procedure as standardized by the International Telecommunication Union (ITU)[19] for end-to-end speech quality assessment of narrowband telephone networks and speech codecs.

Central to the PESQ measurement procedure is the extraction of perceptually relevant features from both the original and processed speech samples using a psychoacoustic model. The differences between the features in the perceptual space are then mapped into a predicted quality score using the so-called "cognitive model," which applies appropriate weighting to the feature differences. There is substantial evidence that the predicted quality scores using this approach correlate better with subjective ratings of speech coder quality,[18,20,22] and noise reduction performance[15] than the IS and LLR measures.

The present study aims to exploit these recent advances in telecommunications for speech quality measurements by investigating the potential of employing the psychoacoustic models in extracting the perceptually relevant features from TE speech and subsequently mapping these features into a metric of TE speech quality. In addition to the ITU-standardized PESQ procedure, this study includes a measurement procedure based on an alternative psychoacoustic model, viz., the Moore–Glasberg's (MG) auditory model,[20] for feature extraction. The performance of the psychoacoustic model based objective metrics and the other telecommunication-specific distortion measures such as the IS and LLR is evaluated using a database of speech samples collected from 35 TE speakers.

II. MOORE–GLASBERG MODEL

The psychoacoustic model provides a way to link the physical acoustical data to the perceptual data by extracting perceptually relevant features.[24] Several models for auditory perception do exist but the MG model[20,25,26] has been recently shown to be a more accurate revision of the earlier auditory models. The MG model allows the computation of perceived loudness patterns which can be assimilated to form indices of voice quality. The computation of the loudness patterns of the reference and test signals is achieved through a series of steps, as shown in Fig. 2, and these are briefly explained below.

The level normalization step involved setting the sound pressure level (SPL) of the input signal to a fixed level of 65 dB SPL. In order to perform the loudness calculation only on frames that contain speech, a voice activity detector[23] (VAD) was used to exclude the inactive speech regions. The pre-processed signal was then transformed into the time-frequency domain using a Hanning window and the short-term Fourier transform (STFT). The speech signal was divided into frames of 32 ms length with 50% overlap. The

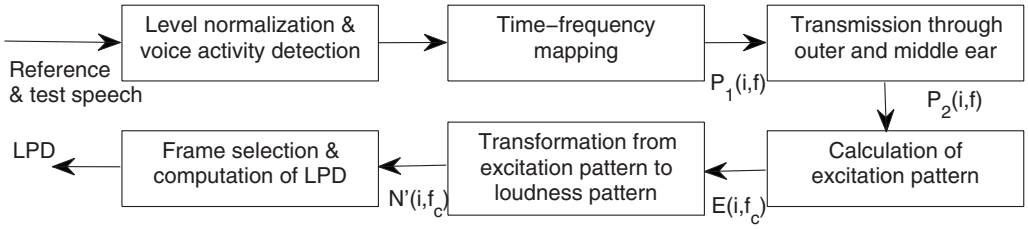


FIG. 2. This block diagrams displays the stepwise procedure for extracting the loudness pattern features and computing the LPD from the raw acoustic signal using the Moore–Glasberg auditory model.

power spectrum of each frame, $P_1(i,f)$, was calculated by taking the STFT and then squaring the real and imaginary components, where i is the frame index and f is the frequency in hertz. A frequency-dependent weighting was then applied to the short-term power spectrum to model the filtering characteristics of the outer- and middle-ear, as shown below:

$$P_2(i,f) = H(f)P_1(i,f) \tag{4}$$

The frequency-dependent weighting function $H(f)$ models the outer- and middle-ear filtering as[21,24]

$$I(f) = 6.5e^{-0.6((f/1000)-3.3)^2} - 10^{-3}\left(\frac{f}{1000}\right)^{3.6} - 2.184\left(\frac{f}{1000}\right)^{-0.8}, \tag{5}$$

$$H(f) = 10^{I(f)/20}. \tag{6}$$

The weighted power spectrum was then used to calculate the excitation pattern, as given by

$$E(i,f_c) = \int_0^\infty \phi(f,f_c,P_2)P_2(i,f)df, \tag{7}$$

where $\phi(f,f_c,P_2)$ is the auditory filter. The auditory filters take a "rounded exponential" shape of the form

$$\phi(f,f_c,P_2) = \left(1 + p\frac{|f-f_c|}{f_c}\right)\exp\left(-\frac{|f-f_c|}{f_c}\right). \tag{8}$$

The parameter p controls the slope of the auditory filters. A more precise modeling[26] assumes asymmetric auditory filters parametrized by upper and lower slope values p_u and p_l, respectively. Correspondingly, in calculating the output of a given filter arising from a given component, the p value is computed depending on whether the filter is centered above or below the frequency of that component, and the total input level of the auditory filter. The excitation pattern was transformed into the associated loudness pattern which was calculated separately for three different cases[20,24] (the frame index i is omitted for clarity):

Case 1: If $(10^9 \geq E(f_c) \geq E_{\mathrm{TH}}(f_c))$

$$N'(f_c) = C((G(f_c)E(f_c) + A(f_c))^{\alpha(f_c)} - A(f_c)^{\alpha(f_c)}); \tag{9}$$

Case 2: If $(E(f_c) > 10^9)$

$$N'(f_c) = C\left(\frac{E(f_c)}{1.115}\right)^{0.2}; \tag{10}$$

Case 3: If $(E(f_c) < E_{\mathrm{TH}}(f_c))$

$$N'(f_c) = C\left(\frac{2E(f_c)}{E(f_c)+E_{\mathrm{TH}}(f_c)}\right)^{1.5}[(G(f_c)E(f_c) + A(f_c))^{\alpha(f_c)} - A(f_c)^{\alpha(f_c)}], \tag{11}$$

where the description of various variables used in the MG model are given in Table I. The first case above shows that the loudness patterns are proportional to the compressed version of the excitation patterns. The frequency-dependent functions, $A(f_c)$ and $G(f_c)$, were necessary to model the decrease in cochlear amplification at low frequencies. The second case results from the fact that at high excitation levels the cochlear amplifier no longer provide any gain and all variables related to it are omitted.[26] The third case is a result of the loudness patterns decreasing more rapidly than is accounted for in the first case possibly because the neural activity of the signal becomes less than the auditory system neural noise at low sound pressure levels.[26]

There is evidence that these loudness patterns contain information related to perceived speech quality.[27,28] For example, using seven different databases of speech quality ratings of speech coding algorithms used in telecommunications, Chen *et al.*[28] showed that the distance between the loudness patterns extracted from the original speech signal and the distorted speech signal can be effectively used to predict the quality of the distorted speech signal. Examples of these loudness pattern distortion (LPD) metrics used in quantifying the speech quality include

$$w_1(i) = \frac{\sum_{u=1}^{M} N'_x(i,u)^2}{\sum_{i=1}^{N}\sum_{u=1}^{M} N'_x(i,u)^2},$$

TABLE I. Definition of various variables used in the computation of loudness patterns using the Moore–Glasberg model.

Variable	Meaning
N'	Loudness patterns
E	Excitaiton patterns produced by the signal
E_{TH}	Excitation patterns corresponding to the threshold in quiet
$G(f)$	Low level gain of the cochlear amplifier
$A(f)$	Frequency-dependent constant that determines the level Dependence of loudness compression
C	Scaling constant=0.047
$\alpha(f)$	Compression exponent

$$w_2(i) = \frac{\sum_{u=1}^{M} N'_y(i,u)^2}{\sum_{i=1}^{N}\sum_{u=1}^{M} N'_y(i,u)^2}, \tag{12}$$

$$d_1(i) = \sqrt{\sum_{u=1}^{M} (N'_x(i,u) - N'_y(i,u))^2}, \tag{13}$$

$$d_2(i) = \frac{\sqrt{\sum_{u=1}^{M} (N'_x(i,u) - N'_y(i,u))^2}}{\sum_{u=1}^{M} N'_y(i,u)^2}, \tag{14}$$

where the subscripts x and y stand for the reference signal and the test signal respectively, N' is the specific loudness per Equivalent Rectangular Bandwidth (ERB), N is the number of speech frames, and M is the number of auditory filters. The final results are based on distance metrics D1 to D4, derived as

$$\mathrm{D1} = \sum_{i=1}^{N} w_1(i) d_1(i), \tag{15}$$

$$\mathrm{D2} = \sum_{i=1}^{N} w_1(i) d_2(i), \tag{16}$$

$$\mathrm{D3} = \sum_{i=1}^{N} w_2(i) d_1(i), \tag{17}$$

$$\mathrm{D4} = \sum_{i=1}^{N} w_2(i) d_2(i). \tag{18}$$

III. DYNAMIC TIME-WARPING

As shown in the PESQ block diagram and from Eqs. (13) and (14), the objective metrics are calculated as the distance between the perceptual features extracted from a high quality reference signal and the test signal. While the high quality reference is readily available for telecommunication applications, it is not the case for the evaluation of disordered voices. Gu *et al.*[17] utilized a speech sample collected from a normal healthy speaker as the reference signal for measuring the quality of Parkinsonian speech. Since this study involves talkers with alaryngeal speech production, it was decided to employ the TE speech sample that received the highest rating as the reference template, and the metrics D1 to D4 were computed for every other TE speech sample relative to this reference template. This poses a challenge as different TE speakers have different speaking rates, which necessitates proper alignment of the reference and test signals in the feature space prior to the computation of the distance metrics. The alignment between the disparate speech samples was achieved by nonlinearly warping the time scale of the test signal to obtain the best registration with the reference signal. In this paper, we employ the dynamic programming (DP) technique to select a constrained warping function.[29] In the DP implementation, the reference and test signals are sampled at discrete points labeled $m=1,2,3,\ldots,M$ and $n=1,2,3,\ldots,N$ respectively. The warping function can then be defined as

$$n = w(m), \tag{19}$$

with the conditions

$$w(1) = 1, \tag{20}$$

$$w(M) = N. \tag{21}$$

The warping function must follow a path defined in the warping grid (m,n) which will be constrained to moving a maximum of 2 index points. It is pertinent to note that Gu *et al.*[17] employed a similar DTW procedure for aligning the speech samples obtained from patients with Parkinson's disease and the reference template recorded from a normal talker.

IV. METHOD

A. Stimuli

In order to evaluate the performance of different acoustic parameters, a database of listener evaluations of TE speech was utilized. The speech samples for the database were gathered from 35 adult male speakers between the ages of 45 and 65 years. All had undergone total laryngectomy and TE puncture voice restoration at least 1 year prior to the recording of the voice samples as part of their participation in the study. All speakers were judged to be highly intelligible despite variations in the overall quality of the signal. All recordings were gathered in a sound-treated environment using stereo recordings at 44.1 kHz sampling rate with 16-bit quantization. As part of the recording protocol, each speaker recorded the first paragraph of the Rainbow passage. The second sentence of the passage, "The rainbow is a division of white light into many beautiful colors" was then extracted from the larger sample and this signal was used for acoustic and perceptual measurements. With regard to the possible presence of stoma noise in these TE speakers, efforts were taken to eliminate this potential confound during the recording procedure. Specifically, all recordings were obtained using a directional microphone positioned at a distance of 15 cm from the speakers lips. The orientation of the microphone was in front of the speaker's mouth on an angle of approximately 30° such that the direct influence of airway noise is minimal. Thus, recorded signals were not influenced by stoma noise and this was confirmed at the time of each recording. Additionally, all speakers were highly proficient in digital occlusion of the tracheostoma; this eliminated any air turbulence during speech production and the gathering of voice signals. If any airway noise was observed to occur during the recording procedure, one of the experimenters reinstructed the speaker and a new recording was obtained to ensure the highest quality voice recordings with the absence of any stoma noise.

B. Subjective ratings

The recorded TE speech samples were played back to a group of 37 naive, normal-hearing, young adult listeners who

had no prior exposure to TE speech or postlaryngectomy verbal communication. The signals were played back to each listener in a random order and the listeners were instructed to rate their overall "comfort level" when listening to the sample using an equal appearing interval scale. The instructions provided to listeners asked them to rate each sample relative to "How comfortable would you feel listening to the persons speech in a social situation. Your rating should reflect your feelings about the way the person is speaking, not what the person is saying." Our decision to evaluate a perceptual dimension termed listener comfort rather than voice quality was based on several factors. In order to obtain perceptual judgments that would be representative of the typical population of listeners, we believed attempts at scaling voice quality offered considerable challenges specific to the population samples under study. One of the primary challenges confronted in such assessments of "voice" centers on the identification and use of a meaningful description of voice quality[30] and concerns of listener bias.[4,5,7] We believed that this definitional concern specific to quality was increased in the context of those signals of interest in the present study. Briefly, because TE voice/speech signals are the result of a highly aperiodic, non-normal voicing system (i.e., vibration of the upper esophagus in response to pulmonary airflow), they may create substantial challenges to the listener in the context of providing judgments. That is, TE voice signals clearly can be identified as unusual and they are always characterized by considerable differences relative to laryngeal-based pathological voice. Thus, because of the unusual nature of TE signals, we anticipated that judgments of listener comfort rather than voice quality would yield individualized responses from unbiased, naive listeners.[4,5,7] We consequently believed that the perceptual dimension of listener comfort could be used to assess the overall character of TE voice samples. In doing so, we anticipated that judgments of comfort would, in fact, serve as a direct surrogate for voice quality without encountering the longstanding definitional problems associated with this perceptual dimension (i.e., voice quality). To facilitate perceptual rating by listeners, descriptive anchors were provided on the scale; anchors ranged from "very comfortable" to "extremely uncomfortable." Listeners were asked to rate the listener comfort on a scale of 1 (extremely uncomfortable) to 9 (very comfortable). Thus, higher scores represented better voice quality and lower scores reflected poorer voice quality as perceived by the naive listeners. Data from all listeners were then collapsed for each of the 35 TE speakers and a mean voice quality rating was obtained. The average of the listener ratings for each speaker was then used to determine the speech sample with the best perceptual rating, and in the computation of correlation coefficients between objective acoustic and subjective perceptual ratings of these 35 TE speaker samples.

C. Objective metrics

The loudness patterns were extracted from all 35 TE speech samples using the MG model. As discussed in Sec. IV B, it is necessary to temporally align the reference and the test TE speech samples using the dynamic time warping (DTW) algorithm prior to the computation of the loudness pattern distortion. It must be noted here that TE speech samples are often characterized by long pauses,[2,31] and the pause duration is negatively correlated with perceived quality of TE speech.[31] Thus, in order to not overcompensate for the differences in the speaking rates of the TE speakers, two constraints were imposed: (a) the voice activity detector ignored pauses greater than 400 ms,[32] and (b) the search path through the DTW grid was restricted to two index points. Once the reference and test TE samples are aligned in the feature space, the distance metrics given in Eqs. (15)–(18) were computed.

In addition, three metrics commonly used in the telecommunication field were also computed. The IS, LLR, and PESQ values were computed using the MATLAB procedures described in Ref. 15. The same DTW algorithm was utilized to align the feature matrices prior to the computation of the IS and LLR metrics, while the PESQ standard includes a proprietary time alignment algorithm that ensures that the reference and test signals are aligned properly.

V. RESULTS

Figure 3 displays the spectrograms and loudness patterns extracted from good and poor quality TE speech samples. Visual inspection of the two spectrograms reveals contrasting information: the spectrogram associated with the good quality TE speaker exhibited well defined fundamental frequency and its harmonics, and smooth and coherent formant variations, while the spectrogram associated with the poorly rated TE speech sample has broader energy spread across the high frequency region and less defined coherent time-frequency structures. The frame-by-frame loudness patterns extracted from the time-frequency information are depicted in the right column of Fig. 3. The objective metrics described in Sec. IV quantify the dissimilarity between these two feature matrices.

Figure 4 displays the mean rank-ordered ratings and their standard error values obtained from naive listeners. As is the norm with subjective quality ratings, the standard deviations are lower at both ends of the ratings, with more variation in the middle regime. The reliability of the subjective ratings was assessed using the intraclass correlation coefficient under the framework of a two-way mixed model and computed using SPSS version 16.0.[33] The "single measures" intraclass correlation coefficient was 0.66 with a 95% confidence interval 0.55–0.77, and the "average measures" intraclass correlation coefficient was 0.99 with a 95% confidence interval 0.98–0.99. Thus the subjective ratings have a moderate degree of reliability.

Table II depicts the individual correlation coefficients between the objective metrics and the subjective ratings. It can be seen from this table that the better performing objective metrics based on the MG model achieve correlation coefficients of 0.69 with the subjective ratings. When combined using the step-wise linear regression function [computed in SPSS V16.0 (Ref. 33)], the overall correlation improved to 0.79 with the variables D1 and D4 contributing to the regres-

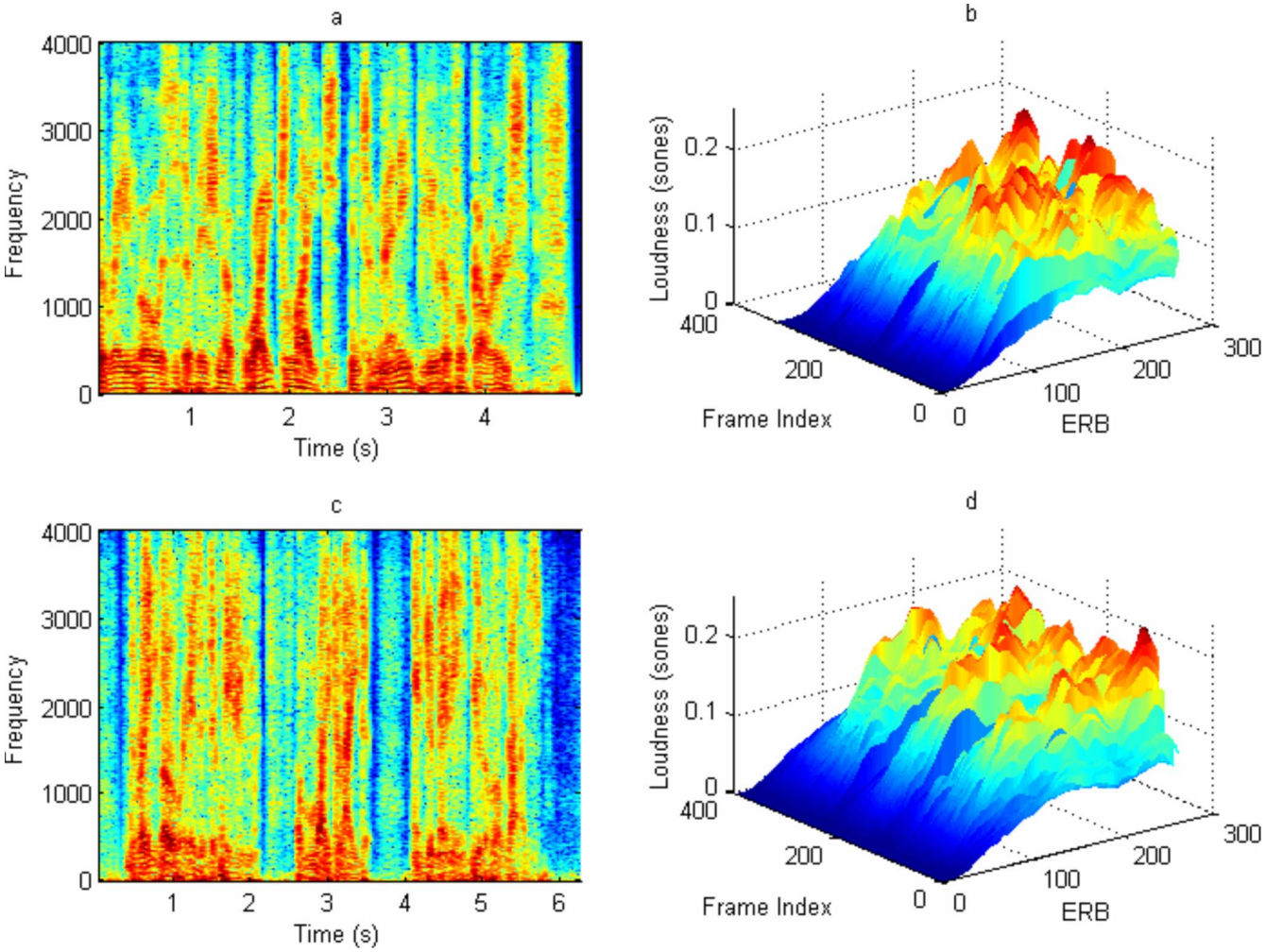


FIG. 3. (Color online) Spectrograms and loudness patterns computed from the sentence *The rainbow is a division of white light into many beautiful colors* spoken by good and poor quality TE speech samples. (a) Spectrogram of the speech sample from a subjectively well rated TE speaker, (b) frame-by-frame perceived loudness patterns computed from the spectrogram in (a), (c) spectrogram of the speech sample from a poorly rated TE speaker, and (d) frame-by-frame perceived loudness patterns computed from the spectrogram in (c).

sion function. The three metrics commonly used in the telecommunication applications performed poorer. It is worthwhile to note that the PESQ measure, which includes an auditory model in its front-end, performed the best among the three. Step-wise linear regression analysis using these three metrics did not result in a significant improvement of the correlation value.

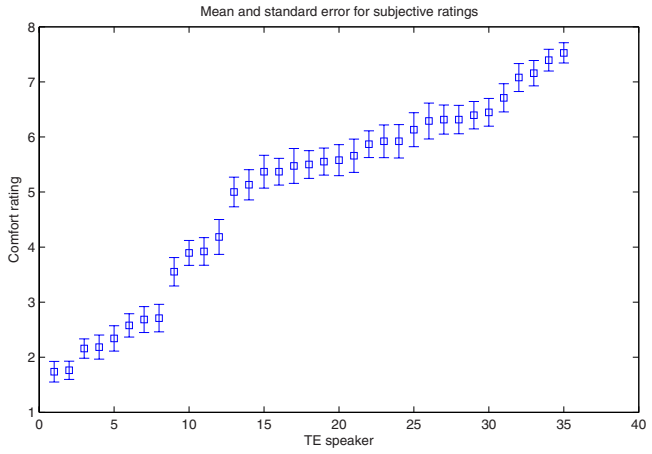


FIG. 4. (Color online) This graph portrays the mean and standard error values for the subjective ratings for all the TE speaker samples. The statistics were calculated from the data collected from a pool of 37 naive listeners.

TABLE II. This table displays the correlation coefficients between the objective metrics and the subjective ratings. The top half of the table reports the correlation coefficients obtained using the MG auditory model, while the bottom half reports the correlation values obtained using metrics routinely used in telecommunication applications.

Distance metric	Correlation, R	Variance accounted for, R^2
D1	0.69	0.48
D2	0.68	0.46
D3	0.41	0.17
D4	0.30	0.09
Combination	0.79[a]	0.62
Itakura–Saito	0.41	0.17
LLR	0.25	0.06
PESQ	0.52	0.27
Combination	0.54	0.29

[a]The linear regression equation is −157.648D1+7.228D4+6.472.

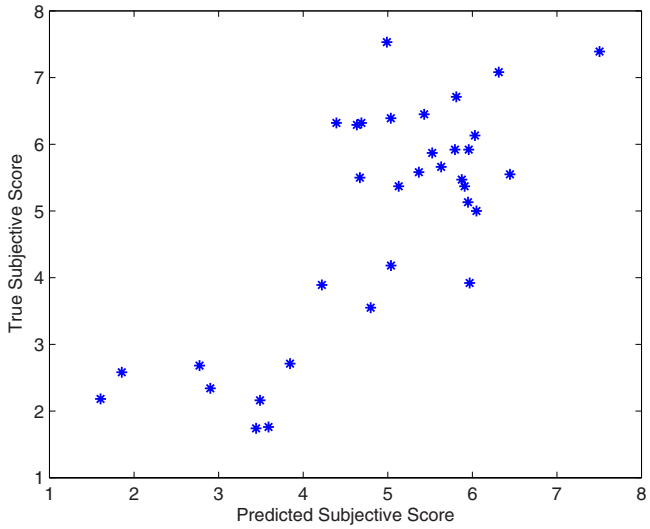


FIG. 5. (Color online) Scatter plot between the predicted and the actual subjective ratings. The predicted subjective ratings were computed using a linear combination of the LPD metrics D1 and D4.

Figure 5 depicts a scatter plot between the predicted and actual subjective ratings obtained from all 35 TE speech samples. The predicted subjective ratings is calculated using the linear regression function of the LPD metrics, D1 and D4 using the equation given at the bottom of Table II. It can be seen that a majority of the data points are clustered around the diagonal, underscoring the modest correlation between the predicted and actual subjective data. In contrast, Fig. 6 shows the scatter plot of the predicted scores using a linear combination of PESQ, IS, and LLR metrics. The dispersion of the data points away from the diagonal in the two-dimensional space is evident in this figure, indicating that these metrics are not as suitable for analyzing TE speech samples.

VI. DISCUSSION

Acoustic analysis of TE speech is a challenging problem due to the highly nonstationary and irregular nature of the speech signal. Conventional techniques of acoustical voice analysis, including estimators of fundamental frequency and amplitude perturbations, glottal noise levels, and linear prediction model based parameters, have not been successful in characterizing the TE speech. In this paper, we have undertaken an alternative approach to TE speech analysis with the ultimate goal of predicting the perceived quality of TE speech. This approach was inspired by the recent advances in telecommunications and audio processing industries, where sophisticated models of normal audition are used in speech and audio processing, and in the objective measurement of speech quality.

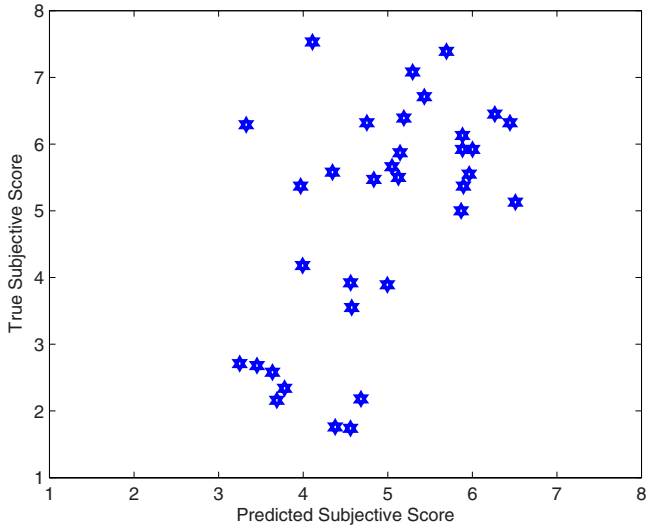


FIG. 6. (Color online) Scatter plot between the actual subjective ratings and the predicted subjective ratings. The predicted values here were computed using a linear combination of the PESQ, LLR, and IS metrics.

Early research in the telecommunications industry on quantifying the effects of distortion and background noise on speech quality concentrated on the so-called distance metrics. For example, the IS measure computes the weighted distance between the linear prediction coefficients extracted from the original, unprocessed speech sample and its distorted version. Only recently, this metric has been applied in the domain of disordered voice quality estimation. Using two databases of Parkinsonian speech samples, Gu *et al.*[17] showed that the IS metric exhibits high correlation with subjective ratings of quality. Correlation coefficients of 0.76 and 0.80 were obtained with subjective ratings of speech samples recorded from four and seven patients with Parkinson's disease, respectively. The present study also applied the IS metric for estimating the quality of TE speech, which resulted in a more modest correlation of 0.43. The differences are perhaps due to the larger number of disordered speech samples employed in our study, and the alaryngeal mode of speech production in TE subjects.

More recent work in the telecommunications and audio processing fields has concentrated on the inclusion of psychoacoustic models of human hearing to enhance the performance of signal processing algorithms. An example is the PESQ procedure standardized by the ITU for characterizing the speech coder/decoder performance. The PESQ measure employs the Zwicker's auditory model[34] for extracting perceptually relevant features and subsequently maps the feature differences between the original and distorted speech signals into a quality score. Using an extensive set of speech coder databases, the PESQ measure has been shown correlate very highly with subjective perception of speech quality, better than the earlier distance metrics such as the IS metric. Indeed, the experimental results obtained in this study demonstrate this as well, where the PESQ measure correlated better with subjective ratings of TE speech than the IS metric.

Recently, more accurate psycho-acoustical experiments have lead to a revised Zwickers model, viz., the MG model.[26] The MG model can be used to better explain how equal-loudness contours change as a function of level, why loudness remains constant as the bandwidth of a fixed-intensity sound increases up to the critical bandwidth, and the loudness of partially masked sounds. Using the MG model, the information content in a given speech sample can be compactly represented by frame-by-frame loudness patterns, and the difference between the loudness patterns of reference and test signals can be used to quantify the quality of the test speech sample. There is some evidence that MG

model can be useful in the acoustic analysis of disordered speech. For example, Shrivastav and Sapienza[35] applied the Moore–Glasberg model to predict the breathiness ratings with moderate success. Results from our study showed that the metrics obtained using the MG model performed the best in correlating with subjective ratings of TE speech, further highlighting the benefits of including a validated auditory model in disordered voice quality analysis.

The objective metrics discussed so far require that an original, clean version of the speech sample is available against which the test samples are compared in the feature space. In telecommunications and audio processing application, such a reference is typically available. Within the realm of disordered voice analysis, however, such a reference is not available. Gu *et al.*[17] solved this problem by utilizing a speech sample obtained from a normal, healthy speaker as the reference template, yielding good results. This approach has theoretical bases as it had been argued that naive listeners base their judgments of voice quality in relation to an internal "normal template."[6] Thus, if an appropriate feature space is constructed, then the features extracted from the disordered speech sample can be judged against the normal feature template. In this study, however, a normal speech sample was not used as the reference template; instead the objective metric was derived as the distance in the feature space between the TE signal that received the highest subjective score and the rest of TE signals in the database. Two factors contributed to this decision: (a) the alaryngeal mode of TE speech production may not compare well with typical laryngeal speech. Furthermore, the TE group is more homogeneous due to the removal of the vocal apparatus, which normalizes the differentiating factors related to speech production somewhat. For example, it is known that the fundamental frequency of male and female TE speakers is similar. This increased homogeneity facilitates for the relative comparison of the TE samples between themselves; (b) the speech recordings and speech analysis took place at two different time periods. A normal speech sample could have been acquired during the analysis phase; however, this idea was abandoned due to potential confounding variables arising from differences in recording setup and equipment. Thus, our objective methodology chose the one sample closest to the normal template (i.e., the one that received the best quality score) as the reference template and evaluated how far apart the rest of the database is with respect to this template. Results corroborate this approach where good agreement between the subjective ratings of speech quality and the objective metric was observed for a database of 35 TE speaker samples. This result may be database specific, and future research should investigate the potential of employing a normal speech sample as the reference template for the computation of objective metrics.

In order to ensure proper computation of the distance metrics, it was imperative to temporally align the reference and test TE speech samples. This was accomplished by using a voice activity detector to remove the pauses (with some restrictions) in the speech signal and then engaging the dynamic time warping algorithm. While the DTW technique has been mainly applied in the automatic speech recognition field, it has also been employed in the acoustic analysis of voice. For example, Gu *et al.*[17] employed the DTW technique prior to the computation of the IS metric. Similarly, Qi and Hillman[36] used the DTW method to align the individual glottal waveforms prior to the computation of harmonics-to-noise ratio (HNR). Within the temporal alignment procedure, additional constraints were imposed to ensure that the timing differences were not overly compensated. This was necessary because typical TE speech is characterized by longer than normal pause periods, and those TE speech samples with long pauses received poorer quality scores. The constraints imposed were as follows: (1) pause times greater than 400 ms (decided empirically) were not flagged by the voice activity detector, and (2) the search space for the warping path was restricted to two index points. For most cases in our database, the DTW provided very good alignment of the reference and test speech samples. For two of the samples, however, the alignment was not perfect due to a much slower speaking rate of these talkers. The resulting distance scores were therefore larger. It is interesting to note that these individuals also received poorer quality scores due to the increased effort needed to complete the target sentence, and thus the correlation between the objective and subjective scores was not drastically affected.

Finally, other models of auditory functioning can be evaluated within the context of objective quality prediction. The Moore–Glasberg model applied in this paper does not account for temporal masking effects, and an alternate model that incorporates this psychoacoustic phenomenon may lead to better correlations with the subjective data. In addition, further validation with a larger database of TE speech quality ratings would lend further support to these preliminary findings.

VII. CONCLUSIONS

Objective predictions of the perceptual ratings of TE speech samples were obtained using an auditory model. The objective speech quality metric for a given TE speech signal was derived as the distance from the highly rated TE speech sample in the auditory perceptual space. Results with a database of 35 TE speakers validated this approach, with an overall correlation coefficient of 0.79 between the objective predictions and the subjective ratings. This correlation, however modest, was significantly better than those achieved using metrics commonly used in the telecommunication industry. These results bode well for further exploration of the application of auditory models to predicting disordered voice quality.

ACKNOWLEDGMENTS

We thank the associate editor and the two anonymous reviewers for their valuable comments on the early version of the manuscript. We gratefully acknowledge the funding from NSERC Canada to Dr. Parsa through the "discovery grant" program. Additional funding was provided through an interdisciplinary research award through the Faculty of Health Sciences at the University of Western Ontario. We thank Guo Chen for providing the MATLAB implementation

of Moore–Glasberg model. We sincerely appreciate all the speakers and listeners who participated in this study for their time and efforts.

[1]M. Singer and E. Blom, "An endoscopic technique for restoration of voice after laryngectomy," Ann. Otol. Rhinol. Laryngol. **89**, 529–533 (1980).
[2]J. Robbins, H. Fisher, E. Blom, and M. Singer, "A comparative acoustic study or normal, esophageal, and tracheoesophageal speech production," J. Speech Hear Disord. **49**, 202–210 (1984).
[3]R. Kazi, V. Prasad, J. Kanagalingam, C. Nutting, P. Clarke, P. Rhys-Evans, and K. Harrington, "Assessment of the formant frequencies in normal and laryngectomized individuals using linear predictive coding," J. Voice **21**, 661–668 (2007).
[4]T. Eadie and P. Doyle, "Direct magnitude estimation and interval scaling of naturalness and severity in tracheoesophageal (TE) speakers," J. Speech Lang. Hear. Res. **45**, 1088–1096 (2002).
[5]T. Eadie and P. Doyle, "Scaling of voice pleasantness and acceptability in tracheoesophageal speakers," J. Voice **19**, 373–383 (2005).
[6]J. Kreiman, B. Gerratt, and M. Ito, "When and why listeners disagree in voice quality assessment tasks," J. Acoust. Soc. Am. **122**, 2354–2364 (2007).
[7]J. Kreiman, B. Gerratt, G. Kempster, A. Erman, and G. Berke, "Perceptual evaluation of voice quality—Review, tutorial and a framework for future-research," J. Speech Hear. Res. **36**, 21–40 (1993).
[8]J. Schoentgen, "Vocal cues of disordered voices: An overview," Acta Acust. **92**, 667–680 (2006).
[9]R. Kazi, E. Kiverniti, V. Prasad, R. Venkataraman, C. Nutting, P. Clarke, P. Rhys-Evans, and K. Harrington, "Multidimensional analysis of female tracheoesophageal prosthetic speech," Clin. Otolaryngol. **31**, 511–517 (2006).
[10]C. D. L. van Gogh, J. Festen, I. V. de Leeuw, A. Parker, L. Traissac, A. Cheesman, and H. Mahieu, "Acoustical analysis of tracheoesophageal speech," Speech Commun. **47**, 160–168 (2005).
[11]T. Most, Y. Tobin, and R. Minran, "Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production," J. Commun. Disord. **33**, 165–181 (2000).
[12]M. Moerman, G. Pieters, J. Martens, M. V. der Borgt, and P. Dejonckere, "Objective evaluation of the quality of substitution voices," Eur. Arch. Otorhinolaryngol. **261**, 541–547 (2004).
[13]C. J. van As-Brooks, F. K. van Beinum, L. Pols, and F. Hilgers, "Acoustical signal typing for evaluation of voice quality in tracheoesophageal speech," J. Voice **20**, 355–368 (2006).
[14]M. Schuster, T. Haderlein, E. Noth, J. Lorsscheller, U. Eysholdt, and F. Rosanowski, "Intelligibility of laryngectomees' substitute speech: Automatic speech recognition and subjective rating," Eur. Arch. Otorhinolaryngol. **263**, 188–193 (2006).
[15]P. Loizou, *Speech Enhancement: Theory and Practice* (CRC, Boca Raton, FL, 2007).
[16]L. Gu, J. G. Harris, R. Shrivastav, and C. Sapienza, "Disordered speech evaluation using objective quality measures," IEEE Trans. Acoust., Speech, Signal Process. **1**, 321–324 (2005).
[17]L. Gu, J. G. Harris, R. Shrivastav, and C. Sapienza, "Disordered speech assessment using automatic methods based on quantitative measures," EURASIP J. Appl. Signal Process. **2005**, 1400–1409 (2005).
[18]J. Beerends, A. Hesktra, A. Rix, and M. Hollier, "PESQ, the new ITU standard for objective measurement of speech quality," J. Audio Eng. Soc. **50**, 765–778 (2002).
[19]ITU, Perceptual speech quality measurement (International Telecommunication Union (ITU-T) Recommendation P.861).
[20]B. Moore and B. Glasberg, "A revised model of loudness perception applied to cochlear hearing loss," Hear. Res. **188**, 70–88 (2004).
[21]ITU, Method for objective measurements of perceived quality (International Telecommunication Union (ITU-R) Recommendation BS.1387-1) (2001).
[22]B. K. R. Huber, "PEMO-Q: A new method for objective audio quality assessment using a model of auditory perception," IEEE Trans. Audio, Speech, Lang. Process. **14**, 1902–1911 (2006).
[23]ITU, Single ended method for objective speech quality assessment in narrow-band telephony applications (International Telecommunication Union (ITU-T) Recommendation P.563) (2004).
[24]G. Chen, "Statistical model-based objective measures of speech quality," Ph.D. thesis, University of Western Ontario, London, ON, Canada (2007).
[25]B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103–138 (1990).
[26]B. Glasberg, B. Moore, and T. Baer, "A model for the prediction of thresholds, loudness and partial loudness," J. Audio Eng. Soc. **45**, 224–239 (1994).
[27]B. Moore and C. Tan, "Development and validation of a method for predicting the perceived naturalness of sounds subjected to spectral distortion," J. Audio Eng. Soc. **52**, 1228–1244 (2004).
[28]G. Chen and V. Parsa, "Loudness pattern-based, speech quality evaluation using Bayesian modeling and Markov chain Monte Carlo methods," J. Acoust. Soc. Am. **121**, EL77–EL83 (2007).
[29]M. Kadous, Dynamic Time Warping, http://www.cse.unsw.edu.au/~waleed/phd/html/node38.html (Last viewed 6/26/2008).
[30]W. H. Perkins, "Vocal function: A behavioral analysis," *Handbook of Speech Pathology and Audiology* (Appleton-Century-Crofts, New York, 1971).
[31]T. Haderlein, E. Noth, H. Toy, A. Batliner, M. Schuster, U. Eysholdt, J. Hornegger, and F. Rosanowski, "Automatic evaluation of prosodic features of tracheoesophageal substitute voice," Eur. Arch. Otorhinolaryngol. **264**, 1315–1321 (2007).
[32]J. Robbins *et al.* (Ref. 2) reported the mean pause times for normal and TE speakers as 624.7 and 891.2 ms, respectively. These times were calculated over the entire Rainbow passage. Since our study only analyzed a single sentence extracted from the Rainbow passage, the pause times were lower. The value of 400 ms was empirically chosen from our database.
[33]S. Inc., *SPSS for Windows, Rel. 17.0.1* (SPSS, Chicago, 2008).
[34]ITU, *Perceptual evaluation of speech quality* (International Telecommunication Union (ITU-T) Recommendation P.862).
[35]R. Shrivastav and C. M. Sapienza, "The use of an auditory model in predicting perceptual ratings of breathy voice quality," J. Acoust. Soc. Am. **43**, 469–485 (2003).
[36]Y. Qi, R. E. Hillman, and C. Milstein, "The estimation of signal to noise ratio in continuous speech for disordered voices," J. Acoust. Soc. Am. **105**, 2532–2535 (1999).

Age-related differences in vocal responses to pitch feedback perturbations: A preliminary study

Hanjun Liu
Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, People's Republic of China

Nicole M. Russo
Department of Pediatrics, Department of Behavioral Sciences, and RUSH Neurobehavioral Center, RUSH University Medical Center, 4711 West Golf Road, Skokie, Illinois 60076

Charles R. Larson[a)]
Department of Communication Science and Disorders, Northwestern University, 2240 Campus Drive, Evanston, Illinois 60208

(Received 24 June 2009; revised 12 November 2009; accepted 14 November 2009)

The present study tested the effect of age on voice fundamental frequency (F_0) responses to pitch-shifted feedback. Pitch-shift stimuli (−100 cents, 200 ms duration) were presented to 30 native-English speakers: 10 children (7–12 yrs), 10 younger adults (19–21 yrs), and 10 older adults (60–73 yrs). Significantly larger response magnitudes were found in the older group compared to the children and young adult groups, while the longest latencies were associated with the children group as compared to the two adult groups. These findings provide preliminary evidence of an age effect on the modulation of vocal responses to pitch-shifted feedback.
 [DOI: 10.1121/1.3273880]

PACS number(s): 43.70.Ep, 43.71.Ft, 43.71.Qr, 43.71.Rt [AL] Pages: 1042–1046

I. INTRODUCTION

The importance of auditory feedback during speech production has been well known (Lombard, 1911; Lane and Tranel, 1971) for many years. In recent years, the frequency perturbation technique was developed and applied to the investigation of the mechanisms underlying auditory feedback control of voice fundamental frequency (F_0) (Elman, 1981; Kawahara, 1994; Burnett *et al.*, 1997). In this protocol, voice pitch auditory feedback is unexpectedly shifted upward or downward and presented to the speakers during voice or speech production. Numerous studies have demonstrated that subjects compensate for these perturbations in pitch feedback by changing their voice F_0 in the opposite direction to the stimulus during ongoing production of speech and voice (Burnett *et al.*, 1998; Donath *et al.*, 2002; Natke *et al.*, 2003; Xu *et al.*, 2004; Chen *et al.*, 2007; Liu and Larson, 2007; Liu *et al.*, 2009). These studies suggest that auditory feedback can be used to stabilize voice F_0 around a desired level and facilitate accurate and timely adjustment in voice control during speech.

The task-dependent activity of the audio-vocal system has been demonstrated by observing the within-subject effect on the modulation of vocal F_0 responses such as vocal parameters (voice F_0 level) (Liu and Larson, 2007), auditory stimuli (duration, velocity, etc.) (Burnett *et al.*, 1998; Larson *et al.*, 2000), and experimental tasks (singing, speaking, etc.) (Natke *et al.*, 2003; Xu *et al.*, 2004; Chen *et al.*, 2007). However, only a few studies investigated the auditory feedback processing of voice F_0 from the perspective of different subject groups. For example, Kiran and Larson (2001) showed that the vocal responses to pitch-shifted feedback in individuals with Parkinson's disease had significantly longer peak times and end times than those of control subjects. It was reported that some children with autism spectrum disorders (ASDs) showed abnormal vocal responses to pitch perturbations as compared with typically developing children development (Russo *et al.*, 2008). During the exposure to the brief pitch perturbations in auditory feedback, singers compensated to a lesser degree than nonsingers (Jones and Keough, 2008). Moreover, singers produced significantly higher voice F_0 values during the initial test-trials that occurred after pitch-shift trials than those produced during baseline and control trials, which was not observed for nonsingers. These studies suggest that between-subject effects may contribute to the modulation of vocal F_0 responses.

One important remaining question that has not been previously addressed is the effect of age on the auditory feedback control of voice F_0. Age-related changes, as is well known, impact both structure and function of the voice and speech mechanisms. The effects of aging on speech production have been extensively investigated to identify the specific acoustic changes that occur in the speech of aging adults (Benjamin, 1981; Ramig and Ringel, 1983; Benjamin, 1997; Mueller, 1997; Sataloff *et al.*, 1997). It has been noted that, as compared to younger adults, acoustic characteristics associated with advanced age include slower speaking and reading rates, higher standard deviation for voice F_0, greater jitter and spectral noise, and lower vowel formants (Shipp and Hollien, 1969; Wilcox and Horii, 1980; Ramig and Ringel, 1983). These changes, generally, are attributed to the

[a)]Author to whom correspondence should be addressed. Electronic mail: clarson@northwestern.edu

physiological changes in speech mechanisms, reduced sensory feedback, decreased speed/accuracy of motor control, and diminished cognitive-linguistic function (Liss *et al.*, 1990; Torre and Barlow, 2009).

Schneider *et al.* (2009) noted that "many essential characteristics of vocal habits and several risk factors for voice disorders later in the professional life are already established in childhood." Special care has been given to the pediatric studies of voice-related disorders, but there have been fewer studies of the normal childhood voice. It is generally agreed that voice F_0 in normal children decreases with age (Bennett, 1983; Glaze *et al.*, 1988, 1990; Nicollas *et al.*, 2008). Conflicting findings, however, in measures of jitter and shimmer were reported. Nicollas *et al.* (2008) found no significant age effect on jitter and shimmer in normal children, which is in contrast with the findings reported by Glaze and colleagues (1988, 1990).

With respect to previous studies using the pitch-shift feedback paradigm, most studies were either done with younger adults or age-matched subjects. How other populations such as normal children respond to pitch-shifted voice feedback in comparison to older adults is unknown. Although all of these groups produce compensatory responses to voice pitch feedback perturbations like the younger adults (Kiran and Larson, 2001; Russo *et al.*, 2008), no between-subject comparisons have been performed to see if there are age-related differences in auditory feedback control of voice F_0. This comparison is important because it will provide information on voice development before laryngeal maturation and the aging-specific physiological changes that affect the processing of auditory feedback during ongoing vocalization. Ultimately, this information will help us understand the complex mechanisms underlying auditory feedback control of voice F_0.

The primary goal of this study was to compare vocal responses to voice pitch-shifted feedback between three groups of native-English speakers, children, young adults, and older adults. The data set for this study came from three previous studies conducted in our laboratories over recent years (Chen *et al.*, 2007; Liu *et al.*, 2008; Russo *et al.*, 2008). Because the methodology differed slightly between each study, we focused on one specific experimental paradigm that was common to all three studies. We tested the hypothesis that age would affect the latency and/or the magnitude of voice F_0 responses to pitch-shifted voice feedback.

II. METHOD

A. Subjects

Three age groups of native-English speaking subjects participated in the study: children ($N=10$, 7–12 years old), young adults ($N=10$, 19–21 years old), and older adults ($N=10$, 60–73 years old). Because this study was not restricted by gender and there were no known effects of gender on the pitch-shift reflexes, these subjects were not gender-matched. None of the children or the younger adults reported a history of speech, hearing, or neurological disorders. All of the older adults passed the hearing screening at 250, 500, 1000, and 2000 Hz at 25 dB hearing level (HL); at 4000 Hz, seven passed the screening at 25 dB HL and three at 50 dB HL. Hearing screening was also performed on children, and they all passed at the threshold of 20 dB HL for pure tone frequencies of 250–8000 Hz. All of the subjects signed informed consent approved by the Northwestern University Institutional Review Board. Additionally, all adults and both the parents of the children and the children themselves signed consent and assent for participation in the study.

B. Apparatus

The younger and the older adults were tested in a sound-treated room (IAC booth, model 1201), where they wore Sennheiser headphones with attached microphone (model HMD 280) throughout the testing. The children were tested in a laboratory room. Although the ambient noise level in the latter two rooms was not as tightly controlled as the sound treatment room, it was previously shown that the relative loudness of voice feedback or the addition of varying levels of masking noise had no significant effect on the amplitude or latencies of responses to pitch-shifted feedback (Burnett *et al.*, 1998). Before the testing, a Brüel and Kjar sound level meter (model 2250) and in-ear microphones (model 4100) were used to adjust the acoustical feedback pathway to a gain of 10 dB sound pressure level between the voice level measured 1 in. from the lips and the ear canal. For experimental testing, the vocal signal from the microphone was amplified with a Mackie mixer (model 1202) and shifted in pitch with an Eventide Eclipse Harmonizer, and then amplified with a Crown D75 amplifier and HP 350 dB attenuators. Max/MSP (v. 4.6 by Cycling 74) was used to control the harmonizer. The voice output, feedback, and transistor-transistor logic (TTL) control pulses were low-pass filtered at 5 kHz (finite impulse response filter), digitized at 10 kHz, and recorded on a Macintosh computer using CHART software (AD Instruments, Castle Hill, New South Wales, Australia).

C. Procedure

Subjects were asked to vocalize the vowel sound /a/ for about 5 s. For each vocalization, the pitch feedback was downward shifted a total of five times with a randomized inter-stimulus interval between 700 and 900 ms. The experimental procedure consisted of blocks of 8–16 vocalizations, producing 40–80 pitch-shift stimuli per subject. The stimulus duration was fixed at 200 ms and the magnitude was held constant at −100 cents (100 cents=1 semitone). Response magnitudes are calibrated in cents, which is a relative logarithmic scale that allows normalization of voice F_0 across subjects. The voice waveform was processed offline in PRAAT (Boersma, 2001) using an autocorrelation method to produce a train of pulses corresponding to the momentary fundamental period of the voice signal recorded via microphone. This pulse train was then converted into an analog wave in IGOR PRO (Wavemetrics, Inc., Lake Oswego, OR). The analog F_0 wave was then converted from Hz to cents with the formula

$$\text{cent} = 12 \times \frac{\log(F_0/196)}{\log 2} \times 100,$$

where F_0 denotes the fundamental frequency in Hz, 196 =arbitrary reference frequency.

For the purpose of obtaining the average F_0 response, voice signals for each trial were time-aligned with the trigger pulse, using 200 ms pre- and 500 ms post-stimulus windows, and averaged. Prior to the averaging, inspections of the individual trials were performed to eliminate those trials with unusually large amplitude, resulting from either signal processing errors or vocal interruption. Valid averaged vocal responses were defined by the averaged cent waveform exceeding a value of 2 standard deviations (SDs) of the pre-stimulus mean (baseline F_0) beginning at least 60 ms after the stimulus and lasting at least 50 ms (Chen *et al.*, 2007). The criterion for the end of a response is that it returns to within 2 SDs of the pre-stimulus mean for at least 30 ms. Latency of the averaged response was defined as the time from the stimulus onset at which the response exceeds 2 SDs of the pre-stimulus mean, and the response magnitude was defined as the difference between the pre-stimulus mean and the greatest value of the F_0 contour following the response onset. Measures of vocal response magnitude and latency were taken from the averaged cent waveforms and those values were analyzed in SPSS (v. 16.0) for significance tests with one-way analyses of variance (ANOVAs) (age).

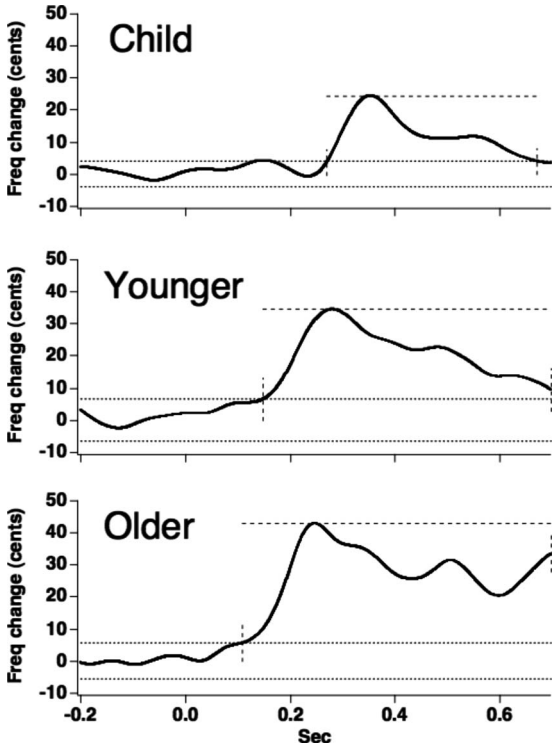


FIG. 1. Representative contours of voice F_0 responses to 100 cents downward stimuli for the children group (top), the younger adults group (middle), and the older adults group (bottom). Horizontal dotted lines indicate ±2 SDs of the pre-stimulus mean voice F_0. Vertical dashed lines affixed to averaged F_0 responses indicate onset and offset times of the response. Horizontal dashed line indicates response magnitude.

III. RESULTS

Figure 1 shows representative voice F_0 responses to 100 cents downward stimuli generated by the three groups of subjects. From top to bottom are the response contours for the children group, the younger adult group, and the older group, respectively. As can be seen from Fig. 1, these response contours were upward, as were all other responses in this study, and were in the opposite direction to the stimuli. While the response magnitudes appear to be roughly similar, the response of the child has a longer latency than the young or older adults.

Figure 2 shows boxplots of response magnitudes (top) and latencies (bottom) as a function of age. Table I shows the average values and SDs of response magnitudes and latencies across age. One-way factorial ANOVAs were performed on the response magnitude and latency as a function of age. For the response magnitude, significant differences were found among three age groups [$F(2,27)=5.822$, $p=0.014$]. *Post hoc* Bonferroni tests indicated that the older group (33 cents) produced significantly larger response magnitudes than the children group (21 cents) ($p=0.039$) and young adults group (20 cents) ($p=0.026$). Statistical analyses also revealed significant differences in the response latency among three groups [$F(2,27)=0.012$], and *post hoc* Bonferroni tests indicated that the children group (223 ms) produced significantly longer latencies than the young adults group (126 ms) ($p=0.033$) and older adults group (121 ms) ($p=0.024$).

IV. DISCUSSION

The present study compared responses to pitch-shifted voice feedback in English speakers across three different age groups to test whether age is a contributing factor to the modulation of voice F_0 responses. Significant differences in both the response magnitude and the latency across the age range demonstrate that the online control of voice F_0 by auditory feedback can be affected by the age of the subjects. Furthermore, the greatest response magnitudes for the older group and the longest latencies for the children group suggest age-specific characteristics of the audio-vocal system in voice F_0 control.

Previous research indicates that vocal response latency can be modulated as a function of stimulus magnitude (Larson *et al.*, 2001), voice F_0 level (Liu and Larson, 2007), and vocal task (Chen *et al.*, 2007). The present study provides evidence that the modulation of vocal response latency is age-dependent. The observed prolonged latency of vocal responses in the children group may be regarded as a reflection of typical neurophysiological maturational processes (Rojas *et al.*, 1998; Kotecha *et al.*, 2009). Although direct comparisons in response latency with previous pitch-shift studies are not available, similar results have been found in the studies of magnetoencephalography (MEG) responses in the auditory cortex to an auditory stimulus. It has been demonstrated that auditory evoked response latency (M100) is dependent on age in healthy developing children, with prolonged latencies for children and shorter latencies for those approaching

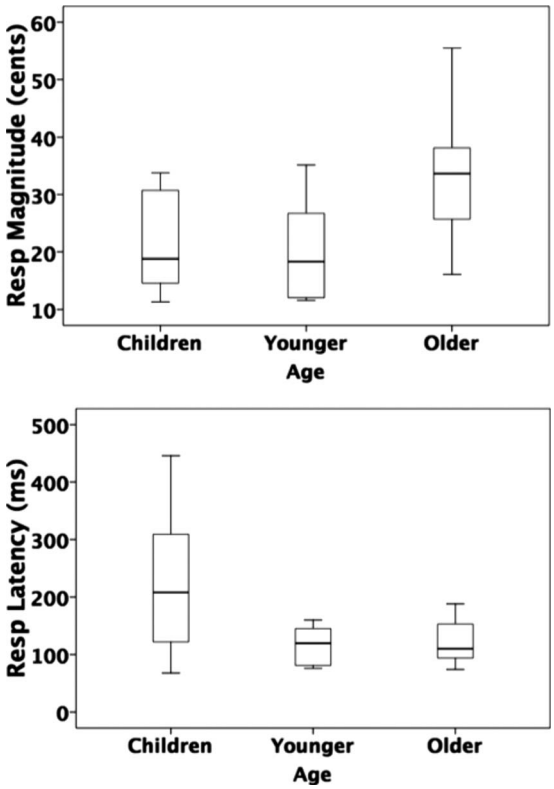


FIG. 2. Boxplots of averaged response magnitudes and latencies across age. Box plot definitions: middle line is median, top and bottom of boxes are 75th and 25th percentiles, whiskers extend to limits of main body of data defined as high hinge+1.5(high hinge−low hinge) and low hinge −1.5(high hinge−low hinge).

adulthood (Rojas *et al.*, 1998; Kotecha *et al.*, 2009). Moreover, no significant differences were observed in the response latencies between the younger and the older adults, indicating the absence of an age effect on the response latency in the adults. Therefore, the similar increase in response latency for the pitch-shift reflex and auditory evoked potentials suggests that similar mechanisms may be involved. Namely, if processing of auditory feedback at the brainstem and cortical level, as revealed through MEG studies, is slower in children compared with adults, this would very likely affect the pitch-shift reflex since it is dependent on auditory feedback. These results also suggest that a long latency or slower processing time of the pitch-shift reflex could serve as an effective indicator of auditory function in developing children.

As compared to the children and the younger adults, the older adults produced the largest response magnitudes. No significant differences were observed between the children and the younger adults. These findings suggest that the increase in the response magnitude from the children and the younger adults to the older adults may be attributed to the aging-specific physiological changes. These changes include immunological, neurological, respiratory, gastrointestinal, muscular, dermatological systems, etc. (Ramig *et al.*, 2001). Since the neural and kinesthetic mechanisms underlying auditory feedback control of voice F_0 are still unclear, there is no specific explanation for the greater response magnitudes in the older adults. According to our recent modeling work in auditory feedback control of voice F_0 (Larson *et al.*, 2008), however, it is possible that the age-related changes in neuromuscular, sensory feedback, and motor control might make major contributions for the response magnitudes generated by the older groups. Along these lines, it is possible that, as people age, they become more sensitive to changes in their voice auditory feedback and produce larger responses. If this is the case, it is surprising that such an effect would not also be evident in the transition from childhood to young adult status.

TABLE I. Averaged response magnitudes and latencies (SD) across age.

Age	Response magnitude (cents)	Response latency (ms)
Children	21 (8)	223 (122)
Younger	20 (9)	126 (54)
Older	33 (11)	121 (36)
Total	25 (11)	157 (90)

The primary limitation of the present study is that response magnitude and latency were measured from only one type of pitch-shift stimulus (100 cents, downward direction). Other parameters such as direction, magnitude, voice F_0, and so on should be undertaken in future studies to address the robust effect of age on voice F_0 responses. Another limitation of the current study is the lack of information concerning the gender effect on vocal response in magnitude and latency. Although there is no published report verifying the sex difference in auditory feedback control of voice, sex does impact the acoustic characteristics of voice in children (Nicollas *et al.*, 2008) and older adults (Mueller, 1997; Torre and Barlow, 2009). Thus, it is possible that sex might impact the control of voice F_0 in these two populations and needs to be further explored.

Although this preliminary study indicates an age effect on the auditory feedback control of voice F_0, there are several additional questions that need to be answered. For example, why did children produce similar response magnitudes to those for the younger adults despite their longer latencies? Similarly, why were only significantly larger response magnitudes observed in the older adults compared to the other two groups? These questions cannot be answered at this time because the functional network of the audio-vocal system underlying these responses is still unknown. Future studies using neural imaging and modeling, combined with perturbed auditory feedback, hold promise for defining this network in greater detail.

V. CONCLUSION

The primary purpose of this study was to determine if age has an effect on the regulation of voice F_0 response to pitch perturbation. The results showed significant age effects on both the response magnitude and the latency. Specifically, the children group produced the longest latencies compared with the younger and older adults, while the largest response magnitudes were generated by the older adults compared with the other two groups. This preliminary study provides evidence that age may play an important role in voice F_0

control, and the developing brain and the aging physiological changes may impact the processing of auditory feedback.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. 1R01DC006243 and Chinese NSF grant No. 30970965. The authors thank Chun Liang Chan for programming assistance.

Benjamin, B. J. (**1981**). "Frequency variability in the aged voice," J. Gerontol. **36**, 722–726.

Benjamin, B. J. (**1997**). "Speech production of normally aging adults," Semin Speech Lang **18**, 135–141.

Bennett, S. (**1983**). "A 3-year longitudinal study of school-aged children's fundamental frequencies," J. Speech Hear. Res. **26**, 137–141.

Boersma, P. (**2001**). "Praat, a system for doing phonetics by computer," Glot International **5**, 341–345.

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (**1998**). "Voice F0 responses to manipulations in pitch feedback," J. Acoust. Soc. Am. **103**, 3153–3161.

Burnett, T. A., Senner, J. E., and Larson, C. R. (**1997**). "Voice F0 responses to pitch-shifted auditory feedback; a preliminary study," J. Voice **11**, 202–211.

Chen, S. H., Liu, H., Xu, Y., and Larson, C. R. (**2007**). "Voice F0 responses to pitch-shifted voice feedback during English speech," J. Acoust. Soc. Am. **121**, 1157–1163.

Donath, T. M., Natke, U., and Kalveram, K. T. (**2002**). "Effects of frequency-shifted auditory feedback on voice F_0 contours in syllables," J. Acoust. Soc. Am. **111**, 357–366.

Elman, J. L. (**1981**). "Effects of frequency-shifted feedback on the pitch of vocal productions," J. Acoust. Soc. Am. **70**, 45–50.

Glaze, L., Bless, D., Milenkovic, P., and Susser, R. (**1988**). "Acoustic characteristics of children's voice," J. Voice **2**, 312–319.

Glaze, L., Bless, D., and Susser, R. (**1990**). "Acoustic analysis of vowel and loudness in children's voice," J. Voice **4**, 37–44.

Jones, J. A., and Keough, D. (**2008**). "Auditory-motor mapping for pitch control in singers and nonsingers," Exp. Brain Res. **190**, 279–287.

Kawahara, H. (**1994**). "Interactions between speech production and perception under auditory feedback perturbations on fundamental frequencies," J. Acoust. Soc. Jpn. **15**, 201–202.

Kiran, S., and Larson, C. R. (**2001**). "Effect of duration of pitch-shifted feedback on vocal responses in Parkinson's disease patients and normal controls," J. Speech Lang. Hear. Res. **44**, 975–987.

Kotecha, R., Pardos, M., Wang, Y., Wu, T., Horn, P., Brown, D., Rose, D., deGrauw, T., and Xiang, J. (**2009**). "Modeling the developmental patterns of auditory evoked magnetic fields in children," PLoS ONE **4**, e4811.

Lane, H., and Tranel, B. (**1971**). "The Lombard sign and the role of hearing in speech," J. Speech Hear. Res. **14**, 677–709.

Larson, C. R., Altman, K. W., Liu, H., and Hain, T. C. (**2008**). "Interactions between auditory and somatosensory feedback for voice F (0) control," Exp. Brain Res. **187**, 613–621.

Larson, C. R., Burnett, T. A., Bauer, J. J., Kiran, S., and Hain, T. C. (**2001**). "Comparisons of voice F_0 responses to pitch-shift onset and offset conditions," J. Acoust. Soc. Am. **110**, 2845–2848.

Larson, C. R., Burnett, T. A., Kiran, S., and Hain, T. C. (**2000**). "Effects of pitch-shift onset velocity on voice F0 responses," J. Acoust. Soc. Am. **107**, 559–564.

Liss, J. M., Weismer, G., and Rosenbek, J. C. (**1990**). "Selected acoustic characteristics of speech production in very old males," J. Gerontol. **45**, 35–45.

Liu, H., and Larson, C. R. (**2007**). "Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex," J. Acoust. Soc. Am. **122**, 3671–3677.

Liu, H., Wang, E. Q., Metman, L. V., and Larson, C. R. (**2008**). "Vocal responses to loudness- and pitch-shift perturbations in individuals with Parkinson's disease," in Proceedings of the Motor Speech Conference, Monterey, CA.

Liu, H., Xu, Y., and Larson, C. R. (**2009**). "Attenuation of vocal responses to pitch perturbations during Mandarin speech," J. Acoust. Soc. Am. **125**, 2299–2306.

Lombard, E. (**1911**). "Le signe de l'évélation de la voix (The sign of a rising voice)," Ann Mal Oreille Larynx **37**, 101–119.

Mueller, P. B. (**1997**). "The aging voice," Semin Speech Lang **18**, 159–169.

Natke, U., Donath, T. M., and Kalveram, K. T. (**2003**). "Control of voice fundamental frequency in speaking versus singing," J. Acoust. Soc. Am. **113**, 1587–1593.

Nicollas, R., Garrel, R., Ouaknine, M., Giovanni, A., Nazarian, B., and Triglia, J. M. (**2008**). "Normal voice in children between 6 and 12 years of age: Database and nonlinear analysis," J. Voice **22**, 671–675.

Ramig, L. A., and Ringel, R. L. (**1983**). "Effects of physiological aging on selected acoustic characteristics of voice," J. Speech Hear. Res. **26**, 22–30.

Ramig, L. O., Gray, S., Baker, K., Corbin-Lewis, K., Buder, E., Luschei, E., Coon, H., and Smith, M. (**2001**). "The aging voice: A review, treatment data and familial and genetic perspectives," Folia Phoniatr Logop **53**, 252–265.

Rojas, D. C., Walker, J. R., Sheeder, J. L., Teale, P. D., and Reite, M. L. (**1998**). "Developmental changes in refractoriness of the neuromagnetic M100 in children," NeuroReport **9**, 1543–1547.

Russo, N., Larson, C., and Kraus, N. (**2008**). "Audio-vocal system regulation in children with autism spectrum disorders," Exp. Brain Res. **188**, 111–124.

Sataloff, R. T., Rosen, D. C., Hawkshaw, M., and Spiegel, J. R. (**1997**). "The aging adult voice," J. Voice **11**, 156–160.

Schneider, B., Zumtobel, M., Prettenhofer, W., Aichstill, B., and Jocher, W. (**2009**). "Normative voice range profiles in vocally trained and untrained children aged between 7 and 10 years," J. Voice (in press).

Shipp, T., and Hollien, H. (**1969**). "Perception of the aging male voice," J. Speech Hear. Res. **12**, 703–710.

Torre, P., and Barlow, J. A. (**2009**). "Age-related changes in acoustic characteristics of adult speech," J. Commun. Disord. **42**, 324–333.

Wilcox, K. A., and Horii, Y. (**1980**). "Age and changes in vocal jitter," J. Gerontol. **35**, 194–198.

Xu, Y., Larson, C., Bauer, J., and Hain, T. (**2004**). "Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences," J. Acoust. Soc. Am. **116**, 1168–1178.

Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German[a)]

Christine Mooshammer[b)]
Haskins Laboratories, 300 George Street, Suite 900, New Haven, Connecticut 06511

(Received 30 December 2008; revised 23 November 2009; accepted 23 November 2009)

This study uses acoustic and physiological measures to compare laryngeal reflexes of global changes in vocal effort to the effects of modulating such aspects of linguistic prominence as sentence accent, induced by focus variation, and word stress. Seven speakers were recorded by using a laryngograph. The laryngographic pulses were preprocessed to normalize time and amplitude. The laryngographic pulse shape was quantified using open and skewness quotients and also by applying a functional version of the principal component analysis. Acoustic measures included the acoustic open quotient and spectral balance in the vowel /e/ during the test syllable. The open quotient and the laryngographic pulse shape indicated a significantly shorter open phase for loud speech than for soft speech. Similar results were found for lexical stress, suggesting that lexical stress and loud speech are produced with a similar voice source mechanism. Stressed syllables were distinguished from unstressed syllables by their open phase and pulse shape, even in the absence of sentence accent. Evidence for laryngeal involvement in signaling focus, independent of fundamental frequency changes, was not as consistent across speakers. Acoustic results on various spectral balance measures were generally much less consistent compared to results from laryngographic data. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277160]

PACS number(s): 43.70.Fq, 43.70.Gr, 43.70.Aj [DAB] Pages: 1047–1058

I. INTRODUCTION

In Germanic languages, prosodic variation is associated with changes in duration, fundamental frequency, intensity, and articulatory precision (e.g., Fry, 1955; Lehiste, 1970; Beckman, 1986). Most types of prosodic variation can be seen as a local enhancement of the prominence of syllables or words relative to their contexts. Different types of prominence, such as word stress and sentence accent, differ with respect to the contribution of individual parameters to the production of prominence: whereas intensity or vocal effort has been shown to be more closely associated with word stress, sentence accent is signaled by rapid *f*0 changes (e.g., Sluijter and van Heuven, 1996). However, a largely unresolved and controversial issue is whether lexical word stress and sentence accent changes are controlled physiologically by the same types of mechanisms. In this study, effects of linguistic prominence on acoustic and physiological measures are compared to effects of global vocal effort changes. The linguistic conditions studied here are lexical word stress with the two levels stressed and unstressed, and sentence accent with the two levels accented and unaccented. This latter variation was elicited by varying the focus.

This introduction is structured in the following way: first, general production mechanisms for vocal effort and linguistic prominence enhancements are reviewed. Then, laryngographic (henceforth Lx) and acoustic reflexes of these mechanisms are discussed with respect to different types of prominence enhancement. Finally, the aims of this study are stated.

A. Production of prominence enhancement

Raising the voice in order to speak more loudly can be achieved by two main mechanisms: increase in respiratory force yielding an increase in subglottal pressure (e.g., Ladefoged, 1967; Hixon, 1973) and increase in laryngeal activity, e.g., glottal adduction and adjustments of the length and stiffness of the vocal folds (e.g., Hirano *et al.*, 1969). The usage of these two mechanisms depends on several factors: As was shown by Finnegan *et al.* (2000), subglottal pressure makes a much larger contribution to vocal effort changes than laryngeal activity. Furthermore, a great deal of speaker-specific variability has been observed in the choice of strategies for effecting changes in vocal effort (Stathopoulos and Sapienza, 1993). Titze (1988) and Winkworth *et al.* (1995) suggested that the preference of one mechanism over the other might depend on the speech task: more global and longer-lasting increases in vocal effort are produced by the respiratory system, whereas shorter and more local changes such as word stress or focus are signaled by the much faster and more flexible laryngeal system. However, as reported by Finnegan *et al.* (2000), laryngeal muscle activity played a minor role to intensity regulation in sentence accent since the respiratory system was also capable of initiating more rapid changes in subglottal pressure. Stathopoulos and Sapienza (1993) showed that these two mechanisms are difficult to discern by airstream or Lx measurements since for higher levels of vocal effort, both laryngeal and respiratory mechanisms affected the shape of the Lx pulses with a quicker and more abrupt closing of the vocal folds and a longer closed phase.

[a)]Portions of this work were presented at the International Conference on Voice Physiology and Biomechanics, Marseille, August 2004, the Between Stress and Tone Conference, Leiden, June 2005 and the 7th International Seminar on Speech Production, Ubatuba, December 2006.

[b)]Electronic mail: tine.mooshammer@haskins.yale.edu

Several other parameters are also affected by increasing vocal effort: Fundamental frequency rises with subglottal pressure (e.g., Ladefoged, 1967; Stathopoulos and Sapienza, 1993). In addition, as found first by Schulman (1989), speakers changed the supraglottal articulation, e.g., using lower jaw positions and therefore increased first formant frequencies. These other manipulations can also affect loudness (by focusing spectral energy in regions where the ear is more sensitive).

Many studies have claimed that word stress is also primarily characterized by an increase in loudness or vocal effort. In his pioneering work in 1967, Ladefoged (1967) found that subglottal pressure not only contributed to global paralinguistic vocal effort changes but also to local variations in prominence, namely, word stress. His experiment showed that short contractions of the muscles activated during exhalation, measured by electromyography, cause an increase in subglottal pressure. However, apart from his later replication of this experiment (Ladefoged, 2005), none of the follow-up experiments directly measuring respiratory muscle activity could reproduce a significant stress effect. Marasek (1997) suggested in a modeling study that greater subglottal pressure alone underlies word stress, whereas sentence accent is primarily controlled by vocal fold tension. Indirect evidence for this position was provided in an acoustical study by Sluijter and van Heuven (1996), who found that the spectral slope was flatter in stressed vowels independent of their accentual status (see also Okobi, 2006 for a clear-cut lexical stress distinction in de-accented position). They attributed this difference in "spectral balance" to a different glottal configuration instead of different subglottal pressures.

There is, however, also accumulating counterevidence: Fant *et al.* (2000) found an increase in subglottal pressure, measured for one subject by tracheal puncture, comparable to vocal effort increases only for very high levels of emphatic stress but not for stressed vowels produced in more neutral environments. No clear-cut distinctions in acoustic parameter changes for stress and accent were found by Campbell and Beckman (1997) and Hanson (1997a). Furthermore, in a study by Heldner (2003) focal accent was also produced with a flatter spectral slope. The Finnegan *et al.* (2000) study also varied focal accent rather than word stress and still found higher subglottal pressure values for focused words. All these studies suggest that there is no clear physiological distinction between stress and accent.

In order to explore this issue further, changes due to word stress, sentence accent (by varying focus), and changes due to increased vocal effort (by asking a subject to vary loudness) were analyzed by using laryngographic techniques and spectral characteristics of the speech waveform. The assumption was that if word stress changes can be attributed to changes in vocal effort, then a similar pattern of Lx and spectral changes should be found in both word stress and raised loudness, and both are expected to be different from the types of Lx and spectral change that accompany sentence accent. In the literature a number of acoustic and Lx parameters have been identified as being affected by changes in vocal effort, sentence accent, and/or stress. Only the parameters used in the current study will be reviewed here: the Lx parameters open quotient (OQ), skewness quotient (SQ), slopes of glottal closing and opening, as well as the acoustic parameters *f*0, intensity, and the acoustic OQ, denoted as H1*-H2*, and spectral tilt, denoted as H1*-A3*. These will be defined precisely in Secs. I B and I C.

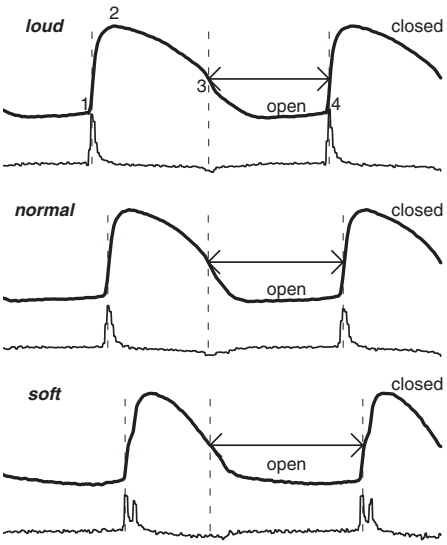


FIG. 1. Examples of Lx signals for loud, normal, and soft speech from speaker M01 at midvowel in stressed and accented position. Data printed as thin lines correspond to the first derivative of the Lx pulses (bold lines). In the upper panel landmarks are denoted by numbers: (1) onset of closing at maximum of the first derivative, (2) maximum contact, (3) moment of glottal opening, determined by a 3/7 threshold, and (4) next glottal closing. The length of the arrows corresponds to the open glottis interval.

B. Laryngeal reflexes

Laryngography or electroglottography has been very popular for recording phonatory behavior for the last 3 decades. This popularity can be attributed to the facts that laryngography is a completely non-invasive technique that does not interfere with normal articulation and to the simplicity of handling this instrument. However, since the laryngograph measures the time-varying impedance between the vocal folds and not the glottal area or the airflow, the signal is difficult to interpret which has led to several critical articles (e.g., Colton and Conture, 1990; Holmberg *et al.*, 1995; Sapienza *et al.*, 1998; Titze, 1990).

Figure 1 shows examples of Lx signals from the current study. The thicker lines represent the Lx pulses and the thin lines the first derivative for loud (top), normal (middle), and soft speech (bottom). In this figure the signal decreases with the amount of contact between the folds, i.e., low for no contact [numbers (1) and (4) in the upper signal] and high for closed glottis (2). By comparing endoscopic high speed images with Lx signals, Henrich *et al.* (2004) found that after the maximum (2) the glottis continues to be closed despite the decrease between (2) and (3). The moment of glottal closing (1) is usually labeled at the maximum of the derivative and most of the times is unambiguous and clearly defined. There are, however, occasionally cases with two peaks in the derivative, as shown in the signal derived from soft speech (bottom of Fig. 1) which are related to discontinuities in the glottal closing (Henrich *et al.*, 2004). Even more problematic is the detection of the moment of glottal opening

(e.g., Childers *et al.*, 1990). Whereas in airflow signals, glottal opening affects the measurements precisely and immediately, the change in Lx signals is much more gradual. Since clear negative peaks, indicating the glottal opening, are rarely found in the derivative (see Fig. 1), various thresholds are used instead. In Fig. 1, a threshold of 3/7 was applied as recommended by Howard *et al.* (1990) and Henrich *et al.* (2004). This difficulty comes about because mucous bridging and vertical phase differences between the lower and upper edges of the vocal folds often make the instance of glottal opening undetectable (see Sapienza *et al.*, 1998; Childers *et al.*, 1990; Colton and Conture, 1990).

Probably because of these problems, several parameters derived from airflow signals that showed variation with linguistic and non-linguistic prominence did not change when derived from the Lx signal. First, it will be considered how **vocal effort** affects different vibratory characteristics and their Lx correlates. The most obvious change due to vocal effort variation concerns the interval during which the glottis is open. The open quotient (ratio of the duration of open glottis to the entire period, henceforth OQ) derived from airflow data generally showed a significant decrease when speaking louder (e.g., Dromey *et al.*, 1992; Holmberg *et al.*, 1988 only for male speakers; Stathopoulos and Sapienza 1993; Sapienza *et al.*, 1998), because the open glottis interval decreases for loud speech. As shown in Fig. 1 by the length of the horizontal arrows, the no-contact interval in the Lx signal clearly increases going from loud to normal to soft speech. However, most studies using Lx data could not replicate this effect (Dromey *et al.* 1992; Sapienza *et al.*, 1998), presumably due to the noise introduced by the uncertainty for detecting the instant of glottal opening. Only Henrich *et al.* (2004) found a clear correlation between intensity and OQ based on the Lx signal, but this was for trained singers.

A second important characteristic varying with vocal effort is the symmetry of the pulse because the glottis closes more quickly at higher levels of vocal effort, which causes a less symmetric and more left-skewed pulse in the airflow signals. The symmetry of the glottal pulse has been quantified by the skewness or speed quotient SQ as the interval of the closing phase (see the raising slope of the pulse in Fig. 1) in relation to the opening phase (the falling slope of the pulse). Airflow SQ usually decreases for a higher vocal effort, indicating a more skewed pulse because of a faster glottal closing movement (Dromey *et al.*, 1992; Holmberg *et al.*, 1988; Sapienza and Stathopoulos, 1998). As for the OQ, this could not, however, be replicated with Lx data (see Dromey *et al.*, 1992; Sapienza *et al.*, 1998, who observed no significant changes for OQ and SQ based on Lx signals). It can be seen in Fig. 1 that Lx pulses are generally not symmetric but left skewed with a very steep raising slope which seems to limit the sensitivity to further changes. Apart from the OQ and SQ, the slopes of the closing and opening phases or the closing peak of the first derivative are also often successfully used to quantify the abruptness of the closure. To summarize, speaking louder decreases the OQ and increases the SQ, but consistent results could only be found for quotients based on airflow data.

The shape and time course of glottal vibrations might also be affected by prosodic variation. Following Beckman (1986) and others, Sluijter and van Heuven (1996) promoted the view that sentence accent and word stress are produced with different laryngeal mechanisms: lexical stress by skewing the glottal pulse and sentence accent by increasing the rate of pulses and thereby producing a higher pitch. Up to now, many acoustic studies have addressed this issue (see Sec. I C) but only a very few laryngographic and airflow investigations exist. To our knowledge, no airflow data are available for **word stress**, but in a study by Marasek (1996) on German, word stress (confounded with sentence accent) had a significant effect on the steepness of the closing and opening slopes in the Lx signal with steeper closing and shallower opening slopes. However, neither the OQ nor the SQ was affected.

A higher *f*0 increased the airflow OQ for global *f*0 changes (e.g., Holmberg *et al.*, 1989), singing (Henrich *et al.*, 2004), and also for **pitch accents** (Pierrehumbert, 1997), probably because when the vocal folds are stiffer they close only at the outer edges, as a result of which the closed phase is shortened (e.g., Titze, 1992). The airflow-based SQ also increases with *f*0, indicating a more symmetrical pulse at higher *f*0, in global tone changes (Holmberg *et al.*, 1989) and for more local pitch accents (Pierrehumbert, 1997). With Lx data, however, only the effects on the OQ could be confirmed (Marasek, 1997 for pitch accents in German, and Gendrot, 2003 for French for focalized vs non-focalized vowels) whereas SQ based on Lx data did not vary systematically with *f*0.

As mentioned above the discrepancy between measures based on airflow data and measures based on Lx data can probably be attributed to the problems with defining meaningful landmarks, especially the moment of glottal opening. In order to overcome these well-known difficulties with labeling landmarks in Lx signals, a more holistic approach was pursued in the current study by analyzing the shape of the Lx pulse as a whole. A similar approach was adopted by Mokhtari *et al.* (2003) who applied a principal component analysis (PCA) to the inverse filtered speech signals of Laver's (1980) recordings of several phonation types. The resulting components discriminated between a wide range of voice qualities. In the present study, function data analysis (henceforth FDA) was used to calculate functional versions of splines of time and amplitude normalized Lx pulses (for an introduction to FDA see, e.g., Ramsay and Silverman, 1997). After computing the spline functions, a PCA was applied to the data (for further details see Sec. II). The prediction was that Lx pulses taken from stressed syllables would exhibit a similar shape to pulses from loud speech and Lx pulses from unstressed syllables would pattern with softly spoken items.

C. Acoustic reflexes

As shown above, very few studies looked at effects of prosodic variation on voice quality using Lx or airflow data. There are, however, many studies addressing this question by means of acoustic data. As mentioned above, the vocal folds close more rapidly with rises of subglottal pressure or changes of the laryngeal settings. This change is not well

captured by the acoustic measure of intensity because it is highly dependent on the distance between the speaker's mouth and the microphone. The more abrupt glottal closure boosts the energy of the higher harmonics in the frequency region above 2000 Hz and therefore decreases the spectral slope. Due to this non-uniform effect on the spectrum, most acoustic studies on word stress or vocal effort are based on a measure of the power spectrum's slope. There are several different measures for quantifying this change, e.g., the balance between the sum of the amplitudes within certain frequency bands ("spectral balance," Sluijter and van Heuven, 1996), the difference between the overall intensity and the intensity of a low-pass filtered signal ("spectral emphasis," e.g., Traunmüller and Eriksson, 2000; Heldner, 2003) and the difference between the amplitude of the first harmonic and the third formant H1-A3, also termed rate of closure (see also Fig. 2, discussed further in Sec. II). Lower values for this latter measure indicate a flatter slope with more energy in the frequency region of the third formant. Holmberg *et al.* (1995) demonstrated for 20 female speakers that loud speech was produced with a lower H1-A3 than modal speech. A similar result would be expected in comparing stressed with unstressed vowels. A severe problem for all of these measures is that they reflect not only changes in the glottal source but are also affected by the vocal-tract resonances, that is, these measures vary with vowel quality. In order to compensate for these effects, a number of correction algorithms have been suggested. For example Hanson (1997b) and Hanson and Chuang (1999) corrected the amplitude of the first harmonic H1 by the location of the first formant frequency F1 and the amplitude of the F3 by the first two formant frequencies of the neutral vowel. However, as shown convincingly by Iseli *et al.* (2007), these algorithms can be improved by also correcting for the effect of bandwidths of the formants. Corrected amplitudes are usually denoted by an asterisk, e.g., H1* and A3*.

With respect to the conditions investigated in this study, i.e., vocal effort, word stress, and focal accent, we would expect that H1*-A3* should progressively increase from loud to normal to soft speech. Stressed vowels should exhibit lower values than unstressed vowels (as confirmed in Sluijter and van Heuven, 1996 for Dutch, although there were contradictory results in Claßen *et al.*, 1998 for German; Campbell and Beckman, 1997; and Hanson, 1997a for English), and sentence accent should not affect the slope of the spectrum (but see Heldner, 2003 for significant differences in spectral slope for focus variation in Swedish).

Speaking more loudly also increases the portion of the glottal cycle during which the glottis is closed because glottal closure is achieved more rapidly and more completely yielding a lower OQ of airflow. Differences in loudness and breathiness have also been found to be associated with amplitude differences between H1* and H2*. Louder and more modal voices are produced with lower values of the acoustic OQ, whereas softer and more breathy voices show increased values (Holmberg *et al.*, 1995). Again, this measure is strongly affected by the vocal-tract resonances, especially the position of F1. Therefore, we applied the correction algorithm by Iseli *et al.* (2007), which has the advantage that it continues to work if F1 approaches the second harmonic. Algorithms based on a correction for the F1 frequency (e.g., Hanson, 1997b) produce invalid data in these cases. Prior studies showed no significant effect for vocal effort (Holmberg *et al.*, 1995), word stress, and pitch accent (Sluijter, 1995; Okobi, 2006 for English; Claßen *et al.*, 1998), probably due to the prominent effect of vowel quality, even on corrected values (see Iseli *et al.*, 2007). However, it has been found that accented vowels are produced with less breathiness than unaccented vowels (Choi *et al.*, 2005). Since we wanted to compare the acoustic to the Lx OQ we included the acoustic OQ in our measurements.

D. Aims

To summarize so far, the first aim of this study is to compare effects of vocal effort, word stress, and sentence accent on the acoustic measures *f*0, H1*-H2*, and H1*-A3* and the Lx measures OQ, SQ, closing slope, opening slope, and FDA shape parameters. The hypothesis is that changes due to vocal effort and word stress affect the investigated parameters in a similar way whereas focus is produced mainly by changes in *f*0. In particular, the goal is to identify a set of parameters that distinguishes stressed from unstressed syllables independently of sentence accent. Earlier studies often failed because they confounded stress and accent. In this current study stress and focal accent are varied orthogonally, and, additionally, vocal effort is introduced for comparison. Consequently, stress could be seen as localized vocal effort change. Variation in sentence accent should not have an effect on the same parameters as stress and vocal effort. If sentence accent affects mainly stressed syllables, then parameters changed by accent variation should show significant differences between stressed focused items on the one hand and all others (stressed unfocused, unstressed focused, and unstressed unfocused). From the literature we expect *f*0, H1*-H2*, and OQ to vary in this way. Second, since the FDA shape analysis has not been applied to Lx data, the second aim of this study is to test whether FDA shape parameters are better suited for analyzing Lx pulses than conventional measures based on landmarks. Third, we are interested in the relationship between this new measure, the more conventional quotients, and the acoustic outcomes.

II. METHOD

A. Material and subjects

The test words were two-syllable words in German that varied in the position of the primary stress; for example, the words "Lena" (a woman's name) with primary lexical stress on the first syllable /ˈleː/ (henceforth *s* for strong) and "Lenor" (name of a washing powder) with primary lexical stress on the second syllable and hence a lexically unstressed, but full vowel /le/ on the first syllable (henceforth *w* for weak). These words were embedded in dialogues which elicited either an accented production in a focused context [F] associated with providing "new" information or an unaccented production because the information was already known [U]. The unfocused condition was constructed in such a way that no *f*0 movement during the test syllable was to be expected. Ques-

tions were pre-recorded and presented via headphones. The answers were presented on a computer screen. In order to get consistent realizations, the words that should be focused were printed in upper case letters. Examples of questions and answers are given below:

(a) Focussed and strong: [F, s]
Q: Wolltest Du Dir Friedas Buch ausleihen? (*Did you want to borrow Frieda's book?*)
A: Nein, ich wollte **LE**NAS Buch ausleihen. (*No, I wanted to borrow Lena's book.*)
(b) Unfocused and strong: [U, s]
Q: Wie findest Du Lena? (*How do you feel about Lena?*)
A: Ich HASSE **Le**na und ihre Schusseligkeit. (*I hate Lena and her absent-mindedness.*)
(c) Focussed and weak [F, w]:
Q: Kaufst Du Omo oder Lenor bei Schlecker? (*Do you buy Omo or Lenor at Schlecker's?*)
A: Ich kaufe **Le**nor bei Schlecker. (*I buy Lenor at Schlecker*)
(d) Unfocused and weak [U, w]:
Q: Wäschst Du nicht gern mit Lenor und Omo? (*Don't you like washing with Lenor and Omo?*)
A: Ich HASSE **Le**nor und Omo. (*I hate Lenor and Omo*)

As a result, we had four possible focal accent × lexical stress combinations: [F, s], [F, w], [U, s], and [U, w], which were spoken at a comfortable vocal effort level (henceforth N). The terms *strong* and *weak* are chosen here in order to avoid confusion of *unstressed* with *unfocused*. A further condition was that all of the [F, s] combinations were produced by the speakers in either a loud (L) or a soft voice (S). In the loud condition, speakers were instructed to speak loudly without shouting; for the soft conditions, the instruction was to speak softly without whispering. In order to ensure consistent loudness levels the questions that were presented over headphones were pre-recorded in the three different loudness conditions. These six (lexical stress × focus in normal loudness + soft and loud levels) possible combinations were repeated nine times for six of the speakers and eight times for the first speaker of this experiment. The items were in randomized order together with a second set of test items, which is not presented here.

Acoustic and Lx signals at a sampling rate of 16 kHz were obtained from seven male subjects between 20 and 35, speaking a northern variety of Standard German. Male speakers were preferred since due to their smaller thyroid angle and longer vocal folds Lx signals are more reliable (e.g., Colton and Conture, 1990) and since female speakers generally speak with a breathier voice quality which reduces the amplitude of the Lx signal (e.g., Holmberg *et al.*, 1995). A dynamic stand microphone (Sennheiser MD 421) was positioned at a distance of 50 cm from the speaker's mouth at an angle of about 45°. Since intensity is also measured in this experiment we instructed the speakers not to move their heads. For a 50 cm mouth-microphone distance a change of ±5 cm corresponds to +1.82/−1.66 dB for frequencies up to 117 Hz and +0.91/−0.83 dB for frequencies above.

All sentences were labeled according to the GToBI conventions (Grice *et al.*, 2005). Six items of the strong unfocused condition had to be excluded because speakers 4 and 5 failed to deaccent 2 and 4 items, respectively. All other items in the unfocused condition had a prominent pitch accent on the word *hasse*, preceding the test word. Some items had to be excluded because of mispronunciation (2) or because the Lx signal was distorted due to excessive vertical larynx movements during the test syllable (3). Altogether 361 items were analyzed.

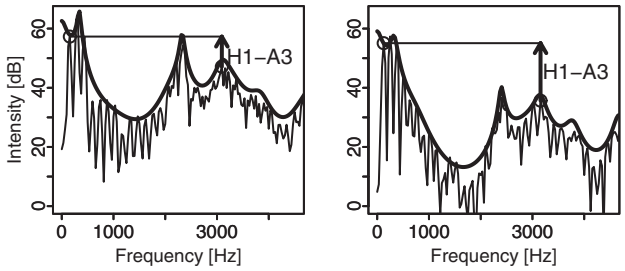


FIG. 2. Calculation of uncorrected spectral tilt H1-A3 for normal vocal effort (left panel) and soft vocal effort (right panel), based on data from speaker M01 during the vowel /e/ in strong and focused position. The solid bold lines display the LPC spectra, and the solid thin lines the narrow band DFT spectra. The length of the arrows corresponds to the difference between the amplitudes of H1 and F3.

B. Measurements

Acoustic measurements. From the speech signal the rms energy, *f*0, and formant frequencies for the first three formants were measured at the acoustical temporal midpoint of the vowel /e:/ from all the test words. The temporal midpoint was chosen in order to minimize coarticulatory effects of the consonant context. However, due to vowel reduction in unstressed position, the consonant context might affect weak vowels to a greater degree than strong ones (see Mooshammer and Geng, 2008 for German). The rms signal was calculated with a Hamming window of 50 ms and shifted by 5 ms. Formants were estimated by using 16 linear predictive coding (LPC) coefficients and a 25 ms Hamming window with 5 ms shift. For some items the number of coefficients had to be adjusted. The formant frequencies and *f*0 were used as reference values for the correction algorithms in order to calculate the acoustic OQ and the spectral tilt discussed below.

The acoustic OQ H1-H2 and the spectral tilt H1-A3 were obtained by applying the following steps: For the vowel interval two different kinds of spectra were calculated, a narrow-band discrete Fourier transform (DFT) spectrum with a frequency resolution of 40 Hz, a Hamming window of 32 ms and shifted by 5 ms, and a LPC spectrum with 22 coefficients and a pre-emphasis of −0.97. The frequencies and amplitudes of the first and second harmonics were detected by means of a peak-picking algorithm. The amplitude of the third formant, A3, was measured at the harmonics that was closest to the third peak in the LPC spectrum. Examples of the measurements before correction for changes in formant values are given in Fig. 2 for speaker M01 in normal speech (upper panel) and soft speech (lower panel). In order to compensate for the formant changes due to modifications of articulatory positions during the vowels, H1*-H2* and H1*-A3* were corrected by the procedure suggested by Iseli *et al.* (2007). Their approach is especially useful for correcting H1 of vowels with low frequency values of F1. Since in

TABLE I. Summary of analyzed parameters derived from the audio signal (upper part) and the laryngographic data (lower part).

Acoustic parameters	Description	Prediction for strong vs weak items	Prediction for focussed vs unfocused items
H1*-H2*	Acoustic open quotient	Lower	Higher
H1*-A3*	Spectral tilt	Lower	–
Lx parameters			
OQ	Open quotient	Lower	Higher
SQ	Skewness quotient	–	–
CSlope	Closing slope	Steeper (more positive)	–
OSlope	Opening slope	Steeper (more negative)	–

this study items with German /e/ with relatively low F1 values are analyzed, the correction suggested by Iseli *et al.* (2007) was applied:

$$H^*(\omega_0) = H(\omega_0) - \sum_{i=1}^{N} 10\log_{10} \times \frac{(1-2r_i\cos(\omega_i)+r_i^2)^2}{(1-2r_i\cos(\omega_0+\omega_i)+r_i^2)(1-2r_i\cos(\omega_0-\omega_i)+r_i^2)}, \quad (1)$$

where $r_i = \exp(-\pi B_i/F_s)$ and $\omega_i = 2\pi F_i/F_s$. The variables B_i and F_i are the bandwidths and frequency of the ith formant, F_s is the sampling rate, and N is the number of formants to be corrected for. In our case the amplitude of the first harmonic $H(\omega_0)$ is corrected by the first two formant frequencies yielding the corrected amplitude $H^*(\omega_0)$. As shown in Table I, a smaller spectral tilt value would be expected for stressed than for unstressed vowels. Likewise, smaller spectral tilt would be expected for loud than normal and normal than soft vocal effort levels. The acoustic OQ has been found to vary with sentence accent. Accordingly, a higher value for H1*-H2* is expected for unaccented items and also for sentences spoken in a soft voice because of an increase in breathiness.

Laryngographic measurements. From the Lx signal the two medial pitch periods during the vowel were extracted. As beginning and end of the extracted Lx pulses, the predetermined 3/7 threshold of glottal opening was used (see Howard *et al.*, 1990). Since we were not interested in shape differences induced only by different period lengths, the two pulses of all items were linearly time-normalized to a uniform length of 1000 samples by linear interpolation.

Since the Lx signal cannot be calibrated, the data were also amplitude normalized to an amplitude of 1 for the first glottal closure. In order to compensate for vertical larynx movements a line connecting the minima of the first and second periods was subtracted from all values. This effect of this amplitude normalization is illustrated in Fig. 3 showing the Lx pulses of about a quarter of all trials from all speakers. The left panel depicts the time-normalized z scores and the right panel the amplitude normalized pulses after subtraction of the line between the two minima. As can be seen the minima and maxima of the amplitude of the normalized data are aligned after this procedure and the amplitude induced variability is much smaller compared to the upper panel.

The following Lx parameters were computed for the medial pitch period of /e/ in all test conditions using EMU/R (Bombien *et al.*, 2006):

(1) For calculating the OQ, the 3/7 threshold was used as the instant of glottal opening (as suggested by Howard *et al.*, 1990) as well as the peak in the first derivative as the instant of glottal closing (see Fig. 4, left). The OQ was then calculated as the percentage of the open glottis interval to the pitch period duration.
(2) The SQ [using a 10% threshold as suggested by Marasek (1997), see Fig. 4, right] was computed as the ratio between the closing and the opening duration. This value decreases with a quicker closing movement.
(3) The slope of glottal adduction was also computed as the quotient between the amplitude of the closing movement and its duration, both defined with a 10% threshold.
(4) Similarly, the slope of the opening movement was defined as the quotient between the amplitude and the duration of the opening movement.

Table I shows the expectations on the way the laryngographic parameters should change based on simulations by Marasek (1997) and results from the literature.

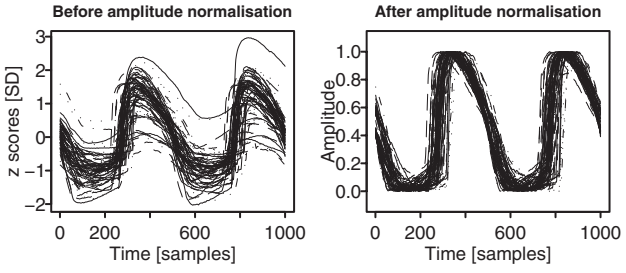


FIG. 3. Lx pulses after time normalization (left panel) and amplitude normalized pulses (right panel). For reasons of visibility only every fourth token is shown.

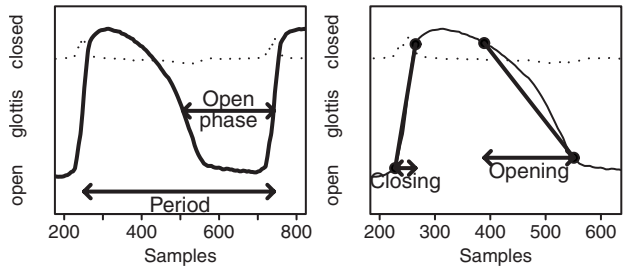


FIG. 4. Left panel: measurements of the open phase and the period duration for calculating the OQ, right panel: measurements of the closing and the opening movement, for calculating the SQ, bold lines indicate the closing and the opening slope. Since data were normalized, the time axis is in arbitrary samples as well as the y-axis in arbitrary amplitude units. Solid lines: normalized glottal pulses; dotted line: first derivative in arbitrary units.

In order to avoid the well-known difficulty with detecting specific landmarks in the Lx signal (see Sec. I), we analyzed the whole shape of the Lx pulse by means of the functional version of the principal component analysis (FPCA) using the R package FDA version 1.2.4 (for further details and formulas, see Ramsay and Silverman, 1997, 2002). Basis functions for the pre-processed medial two Lx pulses were computed by using the Fourier basis functions. Fourier basis functions are recommended for periodic data and involve the calculation of coefficients for the sine components of the waveform: in the current case the number of coefficients was set to 200. This number was necessary because a lower order modified the shape of the Lx pulses too much toward a sinusoidal wave. Smoothing of the resulting curves was obtained by a roughness penalty of the third-order time derivative with the smoothing parameter $\lambda = 10^{-12}$. The order of 200 and the smoothing values were determined by visual inspection of the results and used to ensure that important details of the original data are captured by the basis functions. In many studies FDA time registration is applied to the data prior to further analysis (e.g., Lee *et al.* 2006; Lucero and Koenig, 2000; Lucero and Löfqvist, 2005) for nonlinear time-warping. However, only the above-mentioned linear amplitude and time normalization were applied to the Lx pulses, because we were interested in the skewness of the Lx pulse and the steepness of the closing movement which also implies a shift of the maximum relative to the minimum. Dynamic time-warping tends to obscure these shape characteristics.

In order to identify the main sources of variability in the Lx pulses, a FPCA was applied after calculating the Fourier basis functions (Ramsay and Silverman 1997, 2002). In the case of Fourier basis functions, the FPCA is carried out on the covariance matrix of the smoothed Fourier coefficients. The resulting principal component weight functions are defined over the same range of time as the Fourier functions. Some smoothing was applied to the second derivative and only the first two factors were considered here. The resulting factor scores of the first two components PC1 and PC2 indicate at which time-stretches the Lx pulses show the largest variation as well as the extent of variation. The advantage of the PCA is that only two parameter values, the two factor scores, are needed for describing the range of Lx pulse shapes that occur. Standard statistics as described below were then carried out with these factor scores as dependent variables.

C. Statistics

The cell means of the acoustic measures H1*-H2* and H1*-A3*, the Lx parameters SQ, OQ, closing and opening slopes, and the derived factor scores were analyzed by repeated measures analysis of variance (ANOVA) with speaker as a random factor and with Greenhouse–Geisser correction for violations of the sphericity assumption using R. For significant effects, pairwise t-tests with Bonferroni adjustments were calculated. Since the design was not full-factorial, i.e., vocal effort was only varied for the stressed and focused items, two separate analyses were carried out with subsets of data: for linguistic prominence only data spoken in normal volume were used and tested for the two within-subject factors stress (levels: strong and weak) and focal accent (levels: focused and unfocused). For effects of vocal effort, only data from focused and stressed vowels were taken into account. The within-subject factor vocal effort had the levels loud, normal and soft. For testing speaker consistency, one-way ANOVAs were calculated for the two different data sets split by speaker.

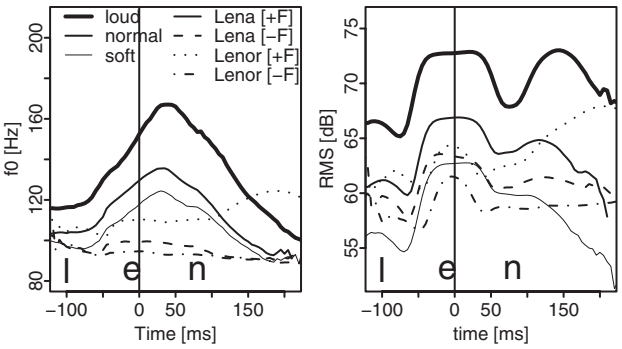


FIG. 5. Ensemble averages of *f*0 contours (left panel) and rms contours (right panel) during the test words *Lena* (initial stress) and *Lenor* (final stress). The vertical line indicates the mid-vowel measurement time-point. All contour are aligned to this time-point.

III. RESULTS

A. Acoustic parameters

Before analyzing several parameters at the midpoint of the vowel, it was verified that the speakers realized the conditions as expected. Therefore, ensemble averages of *f*0 and rms tracks for the entire test word were calculated. Figure 5 shows the averaged *f*0 contours in the left panel and the rms contours in the right. The contours are aligned at the midpoint of the vowel, indicated by the vertical lines. This is approximately the time-point at which measurements were taken for the analyses in Secs. III B–III D. Since the averaged items had different lengths, the averages further away from the midpoint are less reliable and therefore more "bumpy." Solid lines show variation with vocal effort with linewidth decreasing with effort, i.e., thick lines for loud speech and thin lines for soft speech and normal in-between.

As can be seen in the left figure, the items with varying vocal effort levels, which are all stressed on the first syllable and are focused, are produced with a steep rise on the first syllable followed by a fall toward a phrasal low tone. The range of *f*0 variation was largest for the loud condition ($\bar{x}$ = 67 Hz, s = 14 Hz), in-between for normal ($\bar{x}$ = 44, s = 12), and smallest in the soft condition ($\bar{x}$ = 38 Hz, s = 11 Hz). This order was consistent for all speakers. The increase in *f*0 with vocal effort is in agreement with findings from the literature (e.g., Ladefoged, 1967). Within the normal condition the *f*0 range was largest for stress on the initial syllable (i.e., *Lena*) in the focused condition ($\bar{x}$ = 44, s = 12). Pitch accents on the second syllable in *Lenor* in the focused condition were less extensive (dotted line in Fig. 5, $\bar{x}$ = 32 Hz, s = 9 Hz) and the whole word was produced with a somewhat higher *f*0 as compared to the unfocused items (see dashed and dashed dotted lines in Fig. 5). Only the focused word with initial

TABLE II. Results of repeated measures ANOVA and pairwise t-tests for the dependent acoustic variables H1*-H2* and H1*-A3*, the Lx variables OQ, SQ, CSlope, and OpSlope, and the factors vocal effort, stress, focal accent, and interactions between stress and accent. Greenhouse–Geisser corrected degrees of freedom are given for vocal effort, for stress, accent, and the interactions the degrees of freedom are always 1 and 6. Significant effects ($p<0.05$) are printed in bold.

Variable		Vocal effort	Stress	Focus	Interaction
H1*-H2*	F	1.3	**6.9**	2.5	0.2
	df	1.8, 10.9	s<w		
H1*-A3*	F	**11.8**	1.1	2.2	1.1
	df	1.2, 7.1 L<S			
OQ	F	**12.7**	**12.8**	>0.1	2.1
	df	1.3, 7.6 L<S	s<w		
SQ	F	0.97	0.6	**6.6**	0.8
	df	1.1, 6.8		F>U	
CSlope	F	0.63	0.5	>0.1	0.4
	df	1.6, 9.7			
OpSlope	F	0.32	**8.0**	0.3	2.4
	df	1, 6.5	*s>w*		

stress showed a prominent *f*0 movement in the vicinity of the measurement point; for the other three conditions the contour remained flat. This was consistent for all speakers after excluding some of the items (see Sec. II A.)

The right panel in Fig. 5 presents the rms contours. A comparison between the solid lines gives clear evidence that, overall, the speakers managed to produce the items with the required levels of vocal effort with a mean difference over the test items between loud and normal of 6 dB and between normal and soft of 5 dB. The minimal difference was 3 dB, produced between soft and normal speech by two speakers. Prosodic variation also affected the rms contour with weak and unfocused items having a similar rms as items spoken with a soft volume. This can be attributed to the fact that rms rises with *f*0. Therefore, the focused items, produced in a soft volume, had a higher rms during the *f*0 rise during the initial syllable, whereas during /l/ and for the final vowel rms was lower for the soft items than for all others. From these considerations, it can be concluded that the elicitation technique applied in this study was successful because speakers distinguished consistently between the three volume levels and generally produced the predicted *f*0 contours.

Table II gives the results of repeated measures ANOVAs for vocal effort and linguistic prominence (stress and focus) as independent variables and the dependent variables H1*-H2* and H1*-A3*, calculated at the vowel midpoint for the seven male speakers. Figure 6 shows the mean values and standard errors for the vocal effort and prosodic conditions. The grayscale of the bars denote our expectations on prominence measures, i.e., measures relevant for the stress distinction should pattern with vocal effort. Accordingly, darkness decreases in the following order: loud speech, stressed items in comfortable loudness condition, unstressed items in comfortable loudness condition and soft speech.

The acoustic OQ H1*-H2*, shown in the left panel of Fig. 6, was only affected by stress. On average, strong items had only a 1 dB lower acoustic OQ than weak items. Acoustic OQ was not significantly affected by focus or vocal effort.

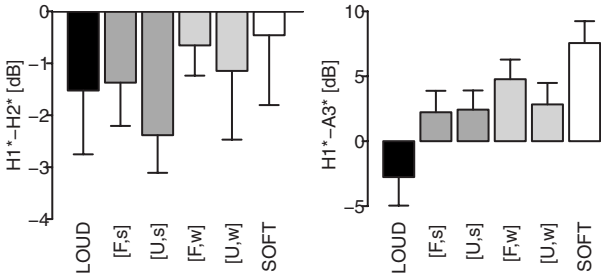


FIG. 6. Means and standard errors for the dependent acoustic variables H1*-H2* and H1*-A3* for vocal effort and linguistic prominence variations. F: focused; U: unfocused; s: strong (first syllable in *Lena*); w: weak (first syllable in *Lenor*).

One possible reason for the latter results might be the low F1 which was always close to the second harmonic for the vowel /e/. Furthermore, the correction algorithm reduced the OQ of the pitch-accented items by 6.3 dB (and most for the loud items) and items with a flat low *f*0 contour only by 3 dB. This non-uniform correction can be attributed to the fact that for the pitch-accented items *f*0 approaches the low first formant in German /e/. Therefore, the amplitude of *f*0, i.e., H1, is boosted by the first vocal-tract resonance. Since the Iseli *et al.* (2007) algorithm includes the bandwidth, it corrects for these *f*0 changes. This explains why no focus and vocal effort effects were found for the acoustic OQ as would have been expected from the literature.

Spectral tilt H1*-A3* (see right panel in Fig. 6) was affected only by vocal effort with substantially lower values of 10 dB for loud vs soft levels. This effect was significant for six of the seven speakers. Contrary to our expectations, stress did not affect this value significantly; however, for the focused items at least, the differences were in the right direction, i.e., spectral tilt for strong vowels was more similar to loud speech and for focused weak vowels [F, w] spectral tilt was more similar to soft speech. The applied correction algorithm reduced the amplitude difference between H1 and A3 by about 8 dB without obvious differential effects for different items.

B. Laryngographic measures

Results for the four Lx measures OQ, SQ, closing slope, and opening slope are given in Table II and Fig. 7. Table II indicates that OQ was significantly affected by vocal effort with the OQ for loud speech 6.6% smaller than for soft speech. There was a significant but smaller effect of stress: OQ was lower by 2.5% for strong items than for weak. Therefore, the glottis was closed for a longer duration for strong than for weak items and for loud than for soft items, which confirms the predictions stated in Table I.

Focus did not affect OQ, but it did affect SQ, resulting in significantly higher values (an average increase of 1.9%) for focused as compared to unfocused items. This means that Lx pulses in focused syllables were produced with a more symmetrical pulse, which is in agreement with data derived from aerodynamic measurements (Holmberg, *et al.* 1989, Pierrehumbert, 1997). According to Marasek's (1997) study, the slopes of the closing and opening movement should vary with stress. ClSlope showed no significant differences for

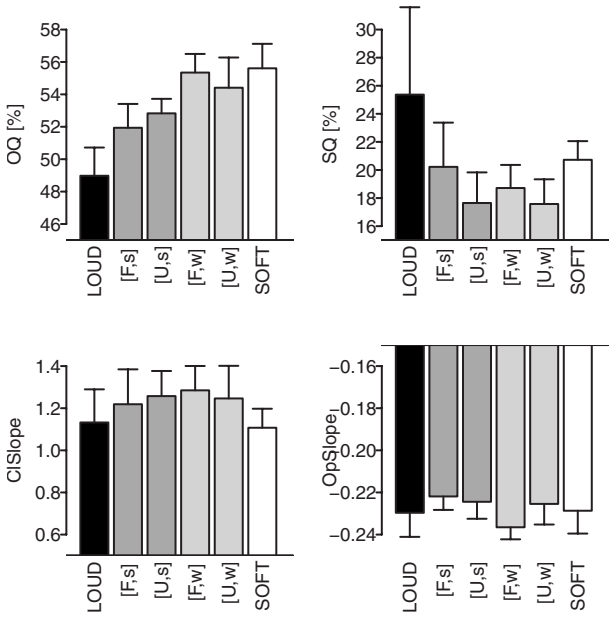


FIG. 7. Means and standard errors for the dependent Lx variables OQ, SQ, ClSlope, and OpSlope for vocal effort and prominence variations. F: focused; U: unfocused; s: strong (first syllable in *Lena*); w: weak (first syllable in *Lenor*).

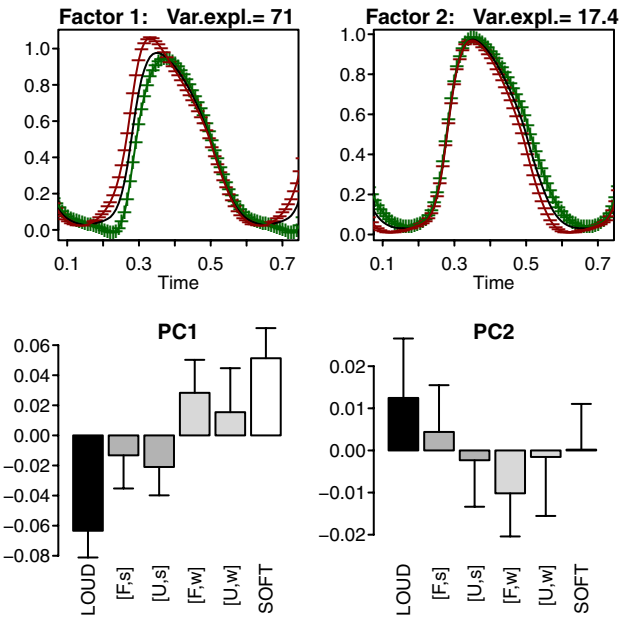


FIG. 8. (Color online) Upper panels: the shapes of the Lx pulse for negative factor values indicated by minus signs, positive factor values by plus signs, and the mean curve by a solid line for the first factor (left) and the second factor (right). Lower panels: means and standard errors of the factor scores for vocal effort and prominence variations. F: focused; U: unfocused; s: strong (first syllable in *Lena*); w: weak (first syllable in *Lenor*).

linguistic prominence or loudness. The steepness of glottal opening varied significantly for stress with steeper (negative) opening slopes for weak items. This result was contrary to our hypothesis, but as can be seen from Fig. 7 the range of variation for this parameter was very small (−0.28 to −0.19) and probably not very meaningful. Generally, the landmark-based parameters SQ, ClSlope, and OpSlope were not consistently affected by stress and vocal effort. Only the OQ varied in the expected direction, i.e., the OQ of weak syllables resembled the OQ of softly spoken syllables, and there was a similar relationship between strong and loud syllables.

C. Results from shape analysis of Lx pulses

Results of the FPCA are presented in Fig. 8. The upper two panels show the shapes of the Lx pulse for negative factor values indicated by minus signs, positive factor values by plus signs, and the mean curve by a dashed line. Shape differences as shown for the PC1 (left) and PC2 (right) explain about 71% and 17%, respectively, of the variance in Lx pulse shapes. Positive values of the first factor were characterized by a longer open phase (as can be seen from the downward hump before the glottis closes) compared to the mean curve. Negative factor values had a shorter open phase and a higher rise. The second factor showed variation in the opening slope, with curves for negative values exhibiting an earlier opening with a steeper opening slope and therefore a longer open phase and a more symmetrical Lx pulse. In gross terms 71% of the variation in Lx pulses can be attributed to the closing movement and the shape of the open phase and 17% to the opening movements. These changes, a later closing movement for positive values of factor 1 and an earlier opening for negative values of factor 2, increased the duration of the open glottis.

In the two panels below, mean and standard errors of PC1 and PC2 are displayed. The higher these values are, the greater the approximation to a "positive" Lx pulse shape. Hence, for the first factor, Lx pulses extracted from loud speech exhibited a "negative" pulse shape with a short open phase and a steep and short closing phase. Soft speech was more similar to the positive Lx pulse and involves a longer open phase. According to our predictions, Lx pulses for strong items should pattern with loud speech and for weak items with soft speech. Scores of the first principal component, shown in Fig. 8, and results from the repeated measures ANOVA in Table III, confirmed this prediction. PC1 differed significantly, both between loud and soft items and also between strong and weak items. One-way ANOVAs split by subjects showed significant stress effects for five of the seven speakers and significant vocal effort effects for six speakers.

For the second factor, there seemed to be a tendency for items bearing a pitch accent to pattern together with slower glottal opening and shorter open phases. The unfocused items and the focused, weak items with a flat *f*0 contour were produced with an earlier opening and a longer open phase. This was reflected in a significant interaction between

TABLE III. Results of repeated measures ANOVA and pairwise t-tests for PC1 and PC2 and the factors vocal effort, stress, focal accent, and interaction between stress and focus. The variances explained by the first two factors are reported in the second column. For vocal effort the degrees of freedom are adjusted for sphericity violations, in all other cases the degrees of freedom are 1 and 6. Significant F values ($p<0.05$) are printed in bold.

PC	Var. expl.	Vocal effort	Stress	Focus	F×S
1	71.0	**11.7**	**14.3**	1.0	0.1
		1.3 7.9 L<S	s<w		
2	17.7	2.6	5.7	0.5	**9.7**
		1.2 7.2			

TABLE IV. Correlation coefficients between factor scores and the glottal shape and acoustic parameters. Coefficients based on cell means were calculated for two subsets of data: for prominence the soft and loud items were excluded; for vocal effort the unfocused and the weak items were excluded. Only significant coefficients ($p<0.05$) are shown here. Highly significant coefficients are printed in bold.

	PC1 prominence	PC1 vocal effort	PC2 prominence	PC2 vocal effort
OQ	**0.67**	**0.81**	**−0.71**	**−0.77**
SQ			**0.61**	**0.51**
ClSlope			**−0.68**	**−0.68**
OpSlope	−0.56		**0.61**	**0.47**
*f*0		−0.51		
rms		**−0.61**		
H1*-H2*			**−0.64**	**−0.75**
H1*-A3*		**0.41**		

stress and focus. Pairwise *t*-tests, however, did not reach significance. The scores of PC2 were also not affected by loudness. Single speaker one-way ANOVAs were also highly inconsistent for the second factor.

In order to understand better the meaning of the factor scores, Pearson correlation coefficients were calculated for two subsets of the data (see Table IV): the prominence subset consists only of items in the normal loudness condition. The vocal effort subset contained data from strong and focused items in the three loudness conditions. For the prominence subset, only the OQ and the opening slope were significantly correlated with PC1. For the vocal effort variation, PC1 was correlated with OQ and the acoustic measures *f*0, rms, and H1*-A3*. The second principal component PC2 was significantly correlated with all of the parameters derived from the Lx signal and also with H1*-H2*. There were only some minor differences in the extent of the correlation coefficients for the two different data subsets, but not in direction. However, apart from the OQ, variables significantly related to PC2 did not contribute to distinguishing different levels of prominence or vocal effort. Therefore, the shape differences, captured by SQ, ClSlope, and OpSlope, and most of the variation in the second factor seemed to be irrelevant for the conditions varied in the current study.

D. Summary

In this study Lx pulse shapes, derived from Lx data, their parametrizations and acoustically derived parameters were compared for three levels of vocal effort, two levels of focal accent, and two levels of lexical stress. Based on the literature, it was predicted that strong items should be produced with more vocal effort than weak items.

The following parameters showed the same tendencies for lexical word stress and vocal effort: the Lx OQ and PC1. For vocal effort changes, there was a decrease in OQ from soft to loud speech. Word stress affected the OQ in a similar way, i.e., strong items had a lower OQ than weak. The OQ of the strong items was also more similar to the OQ of loud speech, and weak items resembled soft speech on this measure. This was also reflected by the global shape parameter PC1 (factor scores of the first principal component) that distinguished Lx pulse shapes of loud from soft items and strong from weak items. Again, the strong items resembled the loud condition and the weak items resembled the soft condition. Vocal effort and stress also affected *f*0 and rms in the same directions. However, these two parameters were also affected by focus. The acoustic parameter H1*-A3* was influenced by vocal effort but not by stress. The opposite was the case for the acoustic OQ (defined as H1*-H2*) and OpSlope which varied with stress but not with vocal effort.

Focus affected SQ with higher values in the focused condition than in the unfocused, meaning that the pulses were more symmetrical in the focused condition, independent of stress. PC2 (factor scores of the second principal component) indicated that strong focused items differed from the others by a shallower opening slope. Since the interaction was significant and only strong syllables showed a significant focus effect, this change might be linked to *f*0 variation.

IV. DISCUSSION

The hypothesis that lexical stress distinction is produced by the same voice source mechanisms as vocal effort was supported by two derived shape parameters of the Lx pulse, the OQ, and PC. For both parameters, the open phase was shortened for higher levels of stress and vocal effort. It is important to note that on the one hand these differences for strong and weak syllables were independent of focus, i.e., strong syllables were produced with a longer closed phase than weak syllables in both focused and unfocused positions. The Lx pulse for focused stimuli was on the other hand more symmetrical, as indicated by increase in skewness of the Lx pulse parametrized as SQ. Both loud and soft items were also produced with a more symmetrical pulse than the focused items, because the former were also produced with pitch accents. Since the Lx pulses were time-normalized before further processing, these effects of stress, focus, and vocal effort cannot be attributed to *f*0 differences. Therefore, vocal effort and stress on the one hand and focus on the other exert different influences on the shape of Lx pulses.

The relevant parameters here differ from Marasek's (1997) who found that stress influenced the steepness of the slopes. The slopes were not affected in the current study. Furthermore, in his study the realization of pitch accents increased the OQ whereas in our study the SQ was modified and not the OQ. Two explanations might account for these different empirical results. First, Marasek (1997) did not vary stress and accent orthogonally, i.e., the two factors were confounded because all strong items were produced with a pitch accent and the weak without. Related to this, the second difference is that Marasek (1997) did not time-normalize the data. As a consequence, the differences in the closing slopes he found for the stress distinction might be attributable to changes in period duration, e.g., steeper slopes for higher *f*0 in strong items.

Another striking difference that is less easy to explain comes from Marasek's (1996, 1997) modeling study. He attributed the unexpected OQ increase for a higher subglottal pressure to a more abrupt closing and opening. However, in our study we found an OQ decrease for higher levels of

stress and vocal effort, which is not only more plausible than his increase but also very well supported by results from airflow data (e.g., Dromey *et al.*, 1992; Holmberg *et al.*, 1988, Stathopoulos and Sapienza, 1993). Our results based on Lx data showing more symmetrical pulses for accented stimuli could be matched more closely to airflow data than those based on Marasek's (1997) more indirect modeling. Even though the data presented here are generally more in agreement with results based on airflow data, the very prominent and consistently affected parameter SQ did not significantly decrease for vocal effort increase in our Lx data. The lack of an effect on the symmetry of the Lx pulses has also been noted by others (e.g., Dromey *et al.*, 1992; Sapienza *et al.*, 1998). A change in symmetry in the direction of the prediction was clearly visible, however, from visual inspection of the extreme PCA factor scores in the present study. The shape of the negative factor value of the first PC in Fig. 8 was more asymmetrical than the positive. The difference, however, lies more in the relative duration of glottal opening than in the abruptness of glottal closing, which has been proposed as one of the major causes for the boost in energy in the higher frequency ranges (e.g., Stevens, 1977; Sluijter and van Heuven, 1996). Since for the modal voice of male speakers, the Lx pulse is already very left skewed compared to the airflow glottal pulse (e.g., Dromey *et al.*, 1992) or the glottal area (Childers *et al.*, 1990), an upper limit for the temporal sensitivity of the recording device might contribute to these negative results. A change in the opposite direction, i.e., more symmetry with shallower closing slopes, was associated with both an increase in *f*0 and variation in PC2.

In summary, the functional version of a PCA, applied to Lx pulses in this study for the first time, has provided a more holistic analysis of the shape of Lx pulses. The resulting factor scores can be seen as parameters of the pulse shape amendable to further exploration by traditional statistical methods. The major advantage of the FPCA is that it does not rely on the identification of often rather arbitrarily defined landmarks, such as the instance of glottal opening. This error-prone landmark definition for the SQ and the opening and closing slopes of the Lx pulse probably contribute to the absence of an effect of stress and vocal effort on these parameters in previous studies.

As argued above, the steepness of the glottal closing is assumed to be related to the spectral shape, i.e., the more abruptly the glottis is closed, the flatter the spectral slope. The measure for spectral slope used in the current study, H1*-A3* with the Iseli *et al.* (2007) corrections, was only affected by vocal effort, not by stress. There was only an insignificant tendency for a lower spectral tilt for strong syllables than weak syllables, but only when the latter were in focused position. This result is contrary to the very clear word stress effects found by Sluijter and van Heuven (1996) for Dutch and to Okobi (2006) for English. However, as was stated in the Introduction, there are several other studies that also did not replicate Sluijter and van Heuven (1995), e.g., Claßen *et al.*, 1998 for German; Campbell and Beckman, 1997; and Hanson, 1997b for English. However, across all these studies there is variation in the languages that were investigated, the vowels that were analyzed, and in the algorithms that were implemented. It is premature to conclude that spectral tilt is an independent correlate of word stress taking into account that some studies were not able to find a measure of spectral slope that could be generalized across different vowel types, vowel realizations, and *f*0. In German, as in many languages, the reduction in weak syllables comes about both because they are more strongly co-articulated with the adjacent sounds (e.g., Mooshammer and Geng, 2008 for German) and because they are more centralized in formant space. In our data, F3 was affected more by stress ($/e/_{strong}$=2713 Hz, $/e/_{weak}$=2504 Hz) than by vocal effort ($/e/_{loud}$=2761 Hz, $/e/_{normal}$=2713, $/e/_{soft}$=2726) (see also Traunmüller and Eriksson, 2000). It is not clear whether our negative results for stress has been brought about because the Iseli *et al.* (2007) method applied here has compensated for stress. This would imply that spectral tilt as a correlate of word stress in other studies can simply be attributed to modification in the vocal-tract and not to changes in subglottal pressure or glottal configurations. However, since the present laryngographic study leads to opposite conclusions, more modeling studies are needed for investigating the effect of glottal pulse shape changes on the shape of the power spectrum, independently of the formant structure.

V. CONCLUSIONS

In this study, effects of vocal effort, stress and focus on glottal and acoustic parameters were analyzed in order to identify one or more reliable correlates of stress, independently of other prosodic variations. The hypothesis was that changes due to word stress affect the same set of parameters as vocal effort changes. In order to test this hypothesis, seven speakers of German were recorded by means of a laryngograph processor. The most important finding was that strong syllables were produced with a longer closed phase and an Lx pulse shape that resembled the Lx pulses also observed during loud speech. Holistic Lx pulse shape differences were parametrized by applying a new method, the functional version of a principal component analysis. Only these two parameters, derived from the Lx signal, varied with stress and vocal effort in a similar direction independent of focal accent. Acoustic parameters were either affected by stress and focal accent together or by only one of stress or vocal effort. The negative results for spectral tilt, which have also been found in other studies, can probably be attributed to changes in formant frequencies due to vocal-tract modifications. In conclusion, the most reliable and consistent correlates of stress and vocal effort, OQ and glottal pulse shape, were derived from the Lx signal. Since they varied with stress in the absence of *f*0 changes, these glottal adjustments can be interpreted as an independent phonetic dimension for signaling lexical stress. Further research is needed in order to determine the actual causes of these changes, namely, subglottal pressure changes or modifications in the glottal configuration.

ACKNOWLEDGMENTS

I would like to thank Jelena Krivokapic, Mark Tiede, Christian Geng, Christine Shadle, and Arthur Abramson for

their valuable suggestions, Jonathan Harrington for help with the corpus design and with EMU/R as well as for very helpful comments, Jennifer Schneeberg for data labeling, and Herbert Fuchs for assisting the recordings. This work was partially supported by NIH Grant No. NIDCD DC008780.

Beckman, M. E. (**1986**). *Stress and Non-Stress Accent* (Foris, Dordrecht, The Netherlands).

Bombien, L., Cassidy, S., Harrington, J., John, T., and Palethorpe, S. (**2006**). "Recent developements in the EMU Speech Database System," in Proceedings of the 11th Australasian International Conference on Speech Science, pp. 313–316.

Campbell, N., and Beckman, M. E. (**1997**). "Stress, prominence and spectral tilt," in Proceedings of the ISCA Workshop on Intonation: Theory, Models and Applications, pp. 67–70.

Childers, D., Hicks, D., Moore, G., Eskenazi, L., and Lalwani, A. (**1990**). "Electroglottography and vocal fold physiology," J. Speech Hear. Res. **33**, 245–254.

Choi, J.-Y., Hasegawa-Johnson, M., and Cole, J. (**2005**). "Finding intonational boundaries using acoustic cues related to the voice source," J. Acoust. Soc. Am. **118**, 2579–2587.

Claßen, K., Dogil, G., Jessen, M., Marasek, K., and Wokurek, W. (**1998**). "Stimmqualität und Wortbetonung im Deutschen (Voice quality and word stress in German)," Linguistische Berichte **174**, 202–245.

Colton, R. H., and Conture, E. (**1990**). "Problems and pitfalls of electroglottography," J. Voice **4**, 10–24.

Dromey, C., Stathopoulos, E. T., and Sapienza, C. M. (**1992**). "Glottal airflow and electroglottographic measures of vocal function at multiple intensities," J. Voice **6**, 44–54.

Fant, G., Kruckenberg, A., and Liljencrants, J. (**2000**). "Acoustic-phonetic analysis of prominence in Swedish," In *Intonation: Analysis, Modelling and Technology*, edited by A. Botinis (Kluwer Academic, Dordrecht), pp. 55–86.

Finnegan, E. M., Luschei, E. S., and Hoffman, H. T. (**2000**). "Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech," J. Speech Lang. Hear. Res. **43**, 934–950.

Fry, D. B. (**1955**). "Duration and intensity as physical correlates of linguistic stress," J. Acoust. Soc. Am. **27**, 765–768.

Gendrot, C. (**2003**). "EGG and spectral investigations on final focalised positions in French," in Proceedings of the 15th ICPhS, pp. 547–550.

Grice, M., Baumann, S., and Benzmüller, R. (**2005**). "German intonation in autosegmental-metrical phonology," in *Prosodic Typology: The Phonology of Intonation and Phrasing*, edited by J. Sun-Ah (Oxford University Press, Oxford), pp. 55–83.

Hanson, H. M. (**1997a**). "Glottal characteristics of female speakers: acoustic correlates," J. Acoust. Soc. Am. **101**, 466–481.

Hanson, H. (**1997b**). "Vowel amplitude variation during sentence production," in Proceedings of the ICASSP, pp. 1627–1630.

Hanson, H. M., and Chuang, E. (**1999**). "Glottal characteristics of male speakers: Acoustic correlates and comparison with female data," J. Acoust. Soc. Am. **106**, 1064–1077.

Heldner, M. (**2003**). "On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish," J. Phonetics **31**, 39–62.

Henrich, N., d'Alessandro, C., Doval, B., and Castellengo, M. (**2004**). "On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation," J. Acoust. Soc. Am. **115**, 1321–1332.

Hirano, M., Ohala, J. J., and Vennard, W. (**1969**). "The function of laryngeal muscle in regulating fundamental frequency and intensity of phonation," J. Speech Hear. Res. **12**, 616–628.

Hixon, T. J. (**1973**). "Respiratory function in speech," in *Normal Aspects of Speech, Hearing, and Language*, edited by F. Minifie, T. J. Hixon, and F. Williams (Prentice-Hall, Englewood Cliffs, NJ), pp. 73–125.

Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**1988**). "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," J. Acoust. Soc. Am. **84**, 511–529.

Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**1989**). "Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch," J. Voice **3**, 294–305.

Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., and Goldman, S. L. (**1995**). "Comparison among aerodynamic, electroglottographic, and acoustic spectral measures of female voice," J. Speech Hear. Res. **38**, 1212–1223.

Howard, D. M., Lindsey, G. A., and Allen, B. (**1990**). "Toward the quantification of vocal efficiency," J. Voice **4**, 205–212.

Iseli, M., Shue, Y.-L., and Alwan, A. (**2007**). "Age, sex, and vowel dependencies of acoustic measures related to the voice source," J. Acoust. Soc. Am. **121**, 2283–2295.

Ladefoged, P. (**1967**). "Stress and respiratory activity," in *Three Areas of Experimental Phonetics*, edited by P. Ladefoged (Oxford University Press, London), pp. 1–49.

Ladefoged, P. (**2005**). "Speculations on the control of speech," in *A Figure of Speech: A Festschrift for John Laver*, edited by W. J. Hardcastle and J. MacKenzie Beck (Erlbaum, Mahwah, NJ), pp. 3–21.

Laver, J. (**1980**). *The Phonetic Description of Voice Quality* (Cambridge University Press, Cambridge).

Lee, S., Byrd, D., and Krivokapic, J. (**2006**). "Functional data analysis of prosodic effects on articulatory timing," J. Acoust. Soc. Am. **119**, 1666–1671.

Lehiste, I. (**1970**). *Suprasegmentals* (MIT, Cambridge, MA).

Lucero, J., and Koenig, L. (**2000**). "Time normalization of voice signals using functional data analysis," J. Acoust. Soc. Am. **108**, 1408–1420.

Lucero, J., and Löfqvist, A. (**2005**). "Measures of articulatory variability in VCV sequences," ARLO **6**, 80–84.

Marasek, K. (**1996**). "Glottal correlates of the word stress and the tense/lax opposition in German," in Proceedings of the ICSLP, University of Stuttgart, Stuttgart, Vol. **96**, pp. 1573–1576.

Marasek, K. (**1997**). "Electroglottographic description of voice quality." AIMS Working Papers Stuttgart, Vol. **3**.

Mokhtari, P., Pfitzinger, H. R., and Ishi, C. Toshinori (**2003**). "Principal components of glottal waveforms: towards parameterisation and manipulation of laryngeal voice quality," in Proceedings of the VOQUAL, pp. 133–138.

Mooshammer, C., and Geng, C. (**2008**). "Acoustic and articulatory manifestations of vowel reduction in German," J. Int. Phonetic Assoc. **38**, 117–136.

Okobi, A. (**2006**). "Acoustic correlates of word stress in American English," Ph.D. thesis, MIT, Cambridge.

Pierrehumbert, J. (**1997**). "Consequences of intonation for the voice source," in *Speech Production and Language: In honour of Osamu Fujimura*, edited by S. Kiritani, H. Hirose, and H. Fujisaki (Mouton de Gruyter, Berlin), pp. 111–130.

Ramsay, J. O., and Silverman, B. W. (**1997**). *Functional Data Analysis*. (Springer, New York).

Ramsay, J. O., and Silverman, B. W. (**2002**). *Applied Functional Data Analysis: Methods and Case Studies*. (Springer, New York).

Sapienza, C. M., Stathopoulos, E. T., and Dromey, C. (**1998**). "Approximation of open quotient and speed quotient from glottal airflow and EGG waveforms: Effects of measurement criteria and sound pressure level," J. Voice **12**, 31–43.

Schulmann, R. (**1989**). "Articulatory dynamics of loud and normal speech," J. Acoust. Soc. Am. **85**, 295–312.

Sluijter, A., and van Heuven, V. (**1996**). "Spectral balance as an acoustic correlate of linguistic stress," J. Acoust. Soc. Am. **100**, 2471–2485.

Stathopoulos, E. T., and Sapienza, C. M. (**1993**). "Respiratory and laryngeal function of women and men during vocal intensity variation," J. Speech Hear. Res. **36**, 64–75.

Stevens, K. N. (**1977**). "Physics of laryngeal behavior and larynx modes," Phonetica **34**, 264–279.

Titze, I. R. (**1988**). "Regulation of vocal power and efficiency by subglottal pressure and glottal width," in *Vocal Fold Physiology: Voice Production, Mechanisms, and Functions*, edited by O. Fujimura (Raven, New York), pp. 227–238.

Titze, I. R. (**1990**). "Interpretation of the electroglottographic signal," J. Voice **4**, 1–9.

Titze, I. R. (**1992**). "Phonation threshold pressure: a missing link in glottal aerodynamics," J. Acoust. Soc. Am. **91**, 2926–2935.

Traunmüller, H., and Eriksson, A. (**2000**). "Acoustic effects of variation in vocal effort by men, women and children," J. Acoust. Soc. Am. **107**, 3438–3451.

Winkworth, A. L., Davis, P. J., and Ellis, E. (**1995**). "Breathing patterns during spontaneous speech," J. Speech Hear. Res. **38**, 124–144.

Compensations in response to real-time formant perturbations of different magnitudes

Ewen N. MacDonald[a)]
Department of Psychology and Department of Electrical and Computer Engineering, Queen's University, Humphrey Hall, 62 Arch Street, Kingston, Ontario K7L 3N6, Canada

Robyn Goldberg
Department of Psychology, Queen's University, Humphrey Hall, 62 Arch Street, Kingston, Ontario K7L 3N6, Canada

Kevin G. Munhall
Department of Psychology and Department of Otolaryngology, Queen's University, Humphrey Hall, 62 Arch Street, Kingston, Ontario K7L 3N6, Canada

(Received 17 August 2009; revised 4 December 2009; accepted 6 December 2009)

Previous auditory perturbation studies have demonstrated that talkers spontaneously compensate for real-time formant-shifts by altering formant production in a manner opposite to the perturbation. Here, two experiments were conducted to examine the effect of amplitude of perturbation on the compensatory behavior for the vowel /ɛ/. In the first experiment, 20 male talkers received three step-changes in acoustic feedback: F1 was increased by 50, 100, and 200 Hz, while F2 was simultaneously decreased by 75, 125, and 250 Hz. In the second experiment, 21 male talkers received acoustic feedback in which the shifts in F1 and F2 were incremented by +4 and −5 Hz on each utterance to a maximum of +350 and −450 Hz, respectively. In both experiments, talkers altered production of F1 and F2 in a manner opposite to that of the formant-shift perturbation. Compensation was approximately 25%–30% of the perturbation magnitude for shifts in F1 and F2 up to 200 and 250 Hz, respectively. As larger shifts were applied, compensation reached a plateau and then decreased. The similarity of results across experiments suggests that the compensatory response is dependent on the perturbation magnitude but not on the rate at which the perturbation is introduced. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3278606]

PACS number(s): 43.70.Mn, 43.70.Bk [DAB] Pages: 1059–1068

I. INTRODUCTION

Auditory feedback plays an essential role in speech-motor control. Clinical studies have identified deficits in speech acquisition when hearing is impaired (Oller and Eilers, 1988) as well as deterioration of speech following post-lingual hearing loss (Cowie and Douglas-Cowie, 1992). Recent experimental work using perturbation techniques indicates that many aspects of speech—loudness control (Bauer *et al.*, 2006), timing (Kalveram and Jäncke, 1989), pitch (Burnett *et al.*, 1998), and formant frequency (Houde and Jordan, 1998)—are influenced by changes in the sound of the talker's voice. When feedback is unexpectedly changed, subjects alter their productions so as to compensate for the changes in the way they hear their own voice. These studies suggest that auditory feedback is part of a control system that actively influences the accuracy of articulator movement.

The nature of this control system for speech and other movements is not completely understood. It is generally believed that the controller must incorporate both feedforward (predictive) behavior and more direct feedback mechanisms (Wolpert and Kawato, 1998) and that both aspects of the controller are adaptive. The adaptive feedforward portion of this controller uses sensory information to learn a detailed representation or "internal model" of the articulator system and its environment and to improve the accuracy of the prediction through evaluating feedback errors (Kawato, 1989). The adaptive feedback portion of the controller is involved in rapid, immediate response to changes in sensory information (Churchland and Lisberger, 2009).

In speech production, auditory feedback plays a complex role in controlling the articulators. It supports control of both the source characteristics (e.g., the vocal pitch) and the vocal tract transfer function (e.g., the formant frequencies) and does so sometimes in rapid response to production changes (e.g., Burnett *et al.*, 1998; Purcell and Munhall, 2006a) as well as in a more predictive fashion (Jones and Munhall, 2005; Houde and Jordan, 1998). In this paper, we focus on the role of auditory feedback in the control of vowel quality and thus on how the sound of the voice is used for the predictive control of the vocal tract transfer function. In this role, sensory feedback is involved in planning the configuration of the oral articulators as well as the scale of their movements.

To use sensory feedback for these purposes, the speech planning system must have a mapping between the movements of the speech articulators and the acoustic spectral consequences of their actions. Thus, the motor planner must

a)Author to whom correspondence should be addressed. Electronic mail: ewen.macdonald@queensu.ca

"know" that movements of a given type will produce acoustic patterns of a certain kind. This input-output relationship must be a significant component of the speech-motor control system, just as a mapping between different spatial coordinate systems is important for visual-motor and vestibular-motor coordination (Beurze *et al.*, 2006). Many control schemes using such a mapping are possible (e.g., Tin and Poon, 2005), and there is need for systematic data on the input-output relationship for auditory feedback to begin to specify the class of controllers that best represents speech-motor control.

In this paper, we examine vowel formant feedback and the relationship between the magnitude of perturbation and the compensatory response. In general, when talkers hear real-time perturbations of their first or second formant, they compensate by altering the formant frequencies of their utterances in a direction opposite in frequency to the perturbation (Houde and Jordan, 1998; Purcell and Munhall, 2006b; Villacorta *et al.*, 2007; Munhall *et al.*, 2009). These compensations exhibit three consistent characteristics. First, compensation is only observed when perturbations have a magnitude greater than some threshold (Purcell and Munhall, 2006b). Second, compensation behavior exhibits a learning curve. Talkers require multiple trials to reach an asymptotic steady state both when the perturbation is initially applied and when it is removed (e.g., Munhall *et al.*, 2009). Third, average compensation is incomplete (Houde and Jordan, 1998; Purcell and Munhall, 2006b; Villacorta *et al.*, 2007; Munhall *et al.*, 2009). The average change in production is smaller in magnitude than the magnitude of the perturbation applied to the acoustic feedback.

The sensitivity to the scale of the perturbation and thus the gain parameter that governs compensatory behavior has been studied with vocal pitch (Burnett *et al.*, 1998; Liu and Larson, 2007) but has not been examined directly for formant frequency perturbations. Burnett *et al.* (1998) examined talkers' responses to pitch shifts ranging from 25 to 300 cents. Over this range, the magnitude of compensation did not vary with the magnitude of the pitch shift. However, the proportion of talkers found to compensate decreased (i.e., more talkers "followed" the perturbation) as the pitch shift was increased in magnitude. Liu and Larson (2007) examined talkers' responses over a smaller range of pitch shift magnitudes ranging from 10 to 50 cents. Over this range, the magnitude of compensation increased with the magnitude of the pitch shift. While complete compensation (i.e., the magnitude of compensation is equal to that of the pitch shift) was observed for the smallest pitch shift, 10 cents, partial compensation was observed for larger shifts. When measured as a proportion of the pitch shift, the compensation decreased as the perturbation was increased. For the largest pitch shift tested, 50 cents, compensation was approximately 30% that of the perturbation.

The gain in response to sensorimotor perturbations has been studied in eye movements as well as limb movements. Saccadic perturbation experiments also reveal an incomplete response to the perturbation of the perceived target location (Hopp and Fuchs, 2004). In smooth pursuit eye movements, the gain appears to be dynamically controlled (Churchland and Lisberger, 2009). Responses to brief perturbations are significantly enhanced during pursuit movements compared to fixations. Limb movements show complex gain modulation at different levels of the system. At the muscle level, the gain of the short-latency stretch response scales proportionally to the muscle activity prior to the perturbation (Pruszynski *et al.*, 2009). The response of hand movements to visual perturbations is largely proportional to the perturbation but is differentially influenced by motion and position information as well as the timing of the perturbation (Saunders and Knill, 2004). However, learning in such contexts has been shown to generalize to new gains between vision and movement (Krakauer *et al.*, 2000), indicating that gain may be an independent parameter even in complex sensorimotor contexts.

In order to begin to parametrize the formant feedback system, perturbations of different magnitudes must be introduced. Here we use two methods of formant perturbation described in the literature: the large abrupt step change in formant frequency (Munhall *et al.*, 2009) and the ramp of small formant frequency changes (e.g., Purcell and Munhall, 2006b). In Experiment 1, compensation was measured for formant-shifts of three different magnitudes using a within-subjects design. These step-changes in formant frequency involve feedback changes that are static for a number of trials to give subjects time to reach a maximum compensation for the given perturbation. In Experiment 2, the frequency shifts were increased in magnitude in small steps with each utterance. Thus, Experiment 2 provides a much broader range of perturbation magnitudes but involves a dynamic feedback environment that is continually changing away from the normal state. Together, these studies reveal the limits of short-term learning within this feedback system and specify the boundaries of the use of auditory feedback in the control of formant production.

II. GENERAL METHODS

A. Equipment

The equipment used was similar to that reported in Munhall *et al.*, 2009. Testing was conducted in an Industrial Acoustics Co. (IAC) sound booth. Talkers were instructed to say words that appeared on a computer monitor at a natural rate and speaking level. Each word prompt lasted 2.5 s and the inter-trial interval was approximately 1.5 s. Talkers spoke into a headset microphone (Shure WH20). This signal was amplified (Tucker-Davis Technologies MA3 microphone amplifier), low-pass filtered with a cut-off frequency of 4500 Hz (Frequency Devices 901 filter for Experiment 1 and Krohn-Hite 3384 filter for Experiment 2), digitized with a sampling rate of 10 kHz and filtered in real-time to produce formant-shifts (National Instruments PXI-8106 controller). The output was amplified and mixed with noise (Madsen Midimate 622 audiometer) and presented over headphones (Sennheiser HD 265) such that the speech and noise were presented at approximately 80 and 50 dBA, respectively.

B. Online formant-shifting and detection of voicing

Detection of voicing and formant-shifting was performed as previously described in Munhall *et al.*, 2009. Voic-

ing was detected using a statistical amplitude-threshold technique. The formant-shifting was achieved in real-time using an infinite impulse response filter. Formants were estimated every 900 μs using an iterative Burg algorithm (Orfandidis, 1988). Filter coefficients were computed based on these estimates such that a pair of spectral zeroes was placed at the location of the existing formant frequency and a pair of spectral poles was placed at the desired frequency of the new formant.

C. Estimating model order

The iterative Burg algorithm used to estimate formant frequencies requires a parameter, the model order, to determine the number of coefficients used in the auto-regressive (AR) analysis. Prior to data collection, talkers produced six utterances of seven English vowels in an /hVd/ context ("heed," "hid," "hayed," "head," "had," "hawed," and "who'd"). These utterances were analyzed with model orders ranging from 8 to 12. The best model order for each individual was selected using a heuristic based on minimum variance in formant frequency over a 25 ms segment midway through the vowel.

D. Offline formant analysis

The procedure used for offline formant analysis was the same as that used by Munhall *et al.* (2009). The boundaries of the vowel segment in each utterance were estimated using an automated process based on the harmonicity of the power spectrum. These boundaries were then inspected by hand and corrected if required.

The first three formant frequencies were estimated offline from the first 25 ms of a vowel segment using a similar algorithm to that used in online shifting. The formants were estimated again after shifting the window 1 ms and repeated until the end of the vowel segment was reached. For each vowel segment, a single "steady-state" value for each formant was calculated by averaging the estimates for that formant from 40% to 80% of the way through the vowel. While using the best model order reduced gross errors in tracking, occasionally one of the formants was incorrectly categorized as another (e.g., F2 being misinterpreted as F1, etc.). These incorrectly categorized estimates were found and corrected by examining a plot with all the "steady-state" F1, F2, and F3 estimates for each individual.

III. EXPERIMENT 1

In this experiment, talkers' compensation was measured for three different formant-shifts using a within-subjects design. Each of the three formant-shifts altered both F1 and F2 such that the vowel /ɛ/ was shifted toward /æ/. Both formants were shifted to follow the trajectory between /ɛ/ and /æ/ in the vowel space. The magnitudes of the three shifts were chosen so that the perturbations were small, medium, and large relative to the average difference in production between /ɛ/ and /æ/. The experiment tested two aspects of the input-output function: (a) whether the compensation magnitude varied linearly with perturbation magnitude across a range of formant values in the vowel space and 9b) whether the tendency toward partial compensation occurred for larger perturbations but not smaller variations in feedback (e.g., Liu and Larson, 2007).

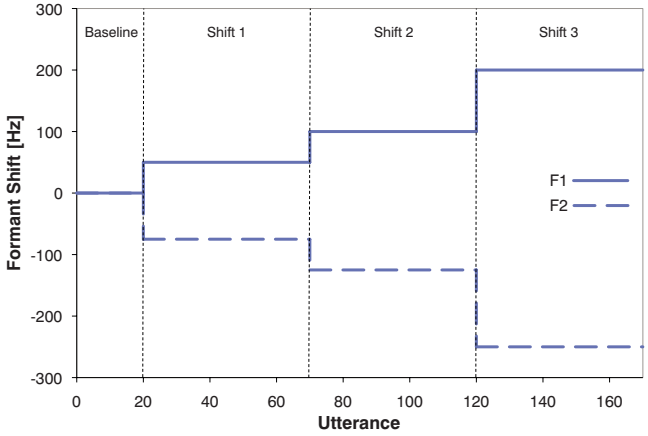


FIG. 1. (Color online) Feedback shift applied to the first formant (solid line) and second formant (dashed line) over for the course of Experiment 1. The vertical dashed lines denote the boundaries of the four phases: Baseline, Shift 1, Shift 2, and Shift 3.

A. Participants

The participants were 20 male undergraduate students from Queen's University ranging in age from 18 to 22 years old ($M=20.5$, $SD=1.24$). All spoke English as a first language and reported no history of hearing or language disorders. Hearing thresholds between 500 and 4000 Hz were assessed. One additional participant was excluded based on a hearing threshold greater than or equal to 25 dB hearing loss (HL) for at least one of the frequencies tested. All other participants had thresholds less than 25 dB HL.

B. Procedure

Over the course of Experiment 1, each talker was prompted to say the word "head" a total of 170 times. The experiment consisted of four phases (see Fig. 1). In the first phase, Baseline, 20 utterances were spoken with normal feedback (i.e., amplified and with noise added but no shift in formant frequency) to estimate Baseline F1 and Baseline F2 values. In each of the subsequent three phases, Shifts 1, 2, and 3, talkers produced 50 utterances with altered feedback with F1 increased and F2 decreased in frequency. F1 was increased in frequency by 50, 100, and 200 Hz for Shifts 1, 2, and 3, respectively. F2 was decreased in frequency by 75, 125, and 250 Hz for Shifts 1, 2, and 3, respectively.

C. Results

For each individual, the baseline average productions of F1 and F2 were calculated from the last 15 utterances of the Baseline phase (i.e., utterances 6–20) and the F1 and F2 results were then normalized by subtracting the subject's baseline average. The normalized results for each utterance, averaged across talkers, can be seen in Fig. 2. In all three phases with altered feedback, talkers, on average, compensated for the altered feedback by changing production of

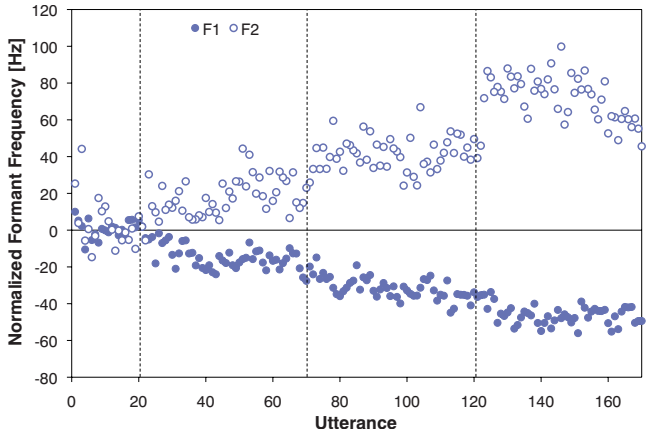


FIG. 2. (Color online) Average normalized F1 (solid circles) and F2 (open circles) frequencies for each utterance. The vertical dashed lines denote the boundaries between the four phases of the experiment: Baseline, Shift 1, Shift 2, and Shift 3.

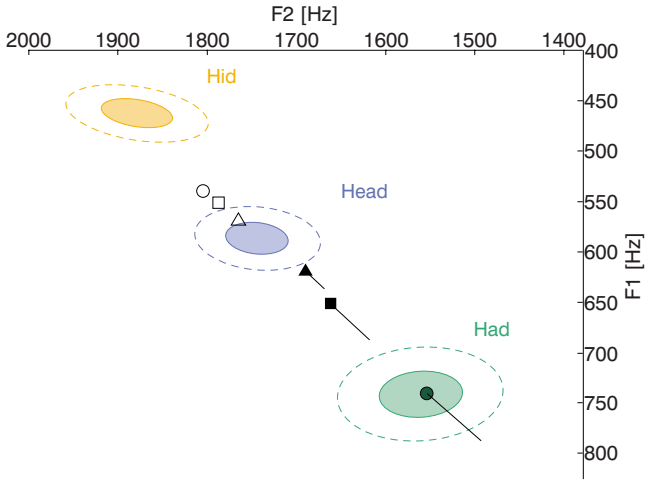


FIG. 3. (Color online) Experiment 1 results superimposed on an average talker's production of /ɪ/, /ɛ/, and /æ/. Each pair of concentric ellipses indicates the distribution of an average talker's production of /ɪ/, /ɛ/, and /æ/ in an /hVd/ context. The center of each pair of ellipses indicates the mean production of an average individual and the solid and dashed ellipses indicate 1 and 2 standard deviations, respectively. The average-steady state results for Shifts 1, 2, and 3 are indicated by the triangle, square, and circle symbols, respectively. Open symbols indicate the average production, while the filled symbols indicate the average acoustic feedback. The three lines indicate the effect of compensation on the acoustic feedback. The ends of the lines opposite the filled symbols indicate the acoustic feedback talkers would have heard if they did not compensate.

both F1 and F2 in a direction opposite that of the perturbation. Further, the magnitudes of the compensation increased over the three phases.

To quantify the change in production, four intervals were defined based on the last 15 utterances in each of the four phases (utterances 6–20 for baseline, 56–70 for Shift 1, 106–120 for Shift 2, and 156–170 for Shift 3). In each interval, it is assumed that formant production has reached a steady state. The F1 and F2 estimates can be seen in Table I. Repeated measures analyses of variance (ANOVAs) with interval as a within-subjects factor (and using Greenhouse–Geisser correction) confirmed a significant effect for both F1 $[F(2.62,57)=30.52,\ p<0.001]$ and F2 $[F(1.74,57)=13.51,\ p<0.001]$. Multiple pairwise comparisons using Bonferroni correction confirmed that the formant values in all four phases were significantly different from each other for both F1 and F2, $(p<0.05)$, except for the comparison between Shifts 2 and 3.

To directly compare the change in production for each formant for a given perturbation, a compensation measure was computed. Compensation was defined as the difference in production between the last utterances of each shift phase and those of the Baseline phase (i.e., the normalized formant values described above) with the sign of these measures harmonized for the two formants. The sign of the compensation was defined as positive if the change in production was opposed to that of the formant-shift or negative if the compensation followed the direction of the formant-shift. For the largest shift, most talkers compensated in a direction opposed to the formant-shift. However, two talkers exhibited following behavior for F2.

To examine the relationship of the vowel space to the compensation patterns, estimates for F1 and F2 were calculated for the utterances of /ɪ/, /ɛ/, and /æ/ ("hid," "head," and "had," respectively) from the vowels collected at the onset of the experiment to determine the model order for formant tracking. For each vowel, the F1 and F2 estimates from each utterance were considered as joint random variables whose means varied across individuals. The data were normalized by subtracting the individual's average and adding the population average formant frequency. For each vowel, the results were pooled to estimate the distribution of F1 and F2 for an average talker's utterance of that vowel.

Each concentric ellipse in Fig. 3 indicates the distribution of an average talker's production of /ɪ/, /ɛ/, and /æ/ in an /hVd/ context. The center of each pair of ellipses indicates the mean F1 and F2, while the solid and dashed ellipses indicate one and two standard deviations, respectively. The results from Experiment 1 have been superimposed on this portion of an average talker's vowel space. The open symbols indicate the average-steady-state production for Shifts 1, 2, and 3 (the triangle, square, and circle symbols, respectively). Similarly, the filled symbols indicate the average-steady-state acoustic feedback for Shifts 1, 2, and 3 (the triangle, square, and circle symbols, respectively). The three straight lines indicate the effect of compensation on the acoustic feedback. If talkers did not compensate, then the acoustic feedback that they would have heard is indicated by the ends of the lines opposite the filled symbols.

To examine the effect of individual vowel space differences on compensation, a correlation analysis was performed. For each individual, the average compensations in F1 and F2 for the largest shift condition were combined as components of a compensation vector in the vowel space.

TABLE I. Mean formant frequency in Hz for the last 15 utterances of each of the four phases in Experiment 1. One standard deviation is given in parentheses.

	Baseline	Shift 1	Shift 2	Shift 3
F1	582.6(29.9)	562.4(32.7)	547.8(28.9)	536.0(28.4)
F2	1705.8(80.4)	1727.1(80.1)	1749.1(89.7)	1766.5(88.8)

Similarly, vectors were computed based on the difference between average production of /ɛ/ and /æ/ for each individual. No significant correlation was found between the magnitudes of these two vectors. Further, no significant correlation was found when this analysis was repeated comparing compensation with individual differences between /ɛ/ and /æ/.

D. Discussion

In this experiment, the compensations to three pairs of formant perturbations were measured. As in our previous work on large, abrupt formant perturbations (Munhall *et al.*, 2009), the observed compensations exhibited learning curves that are approximately exponential and asymptote at a level of partial compensation. Subjects on average compensated 17.4, 34.9, and 46.6 Hz for perturbations in F1 of 50, 100 and 200 Hz. For F2, the compensations were similarly partial with compensations of 21.2, 43.3, and 60.7 Hz being observed for perturbation magnitudes of 75, 125, and 250 Hz. For both formants, the overall compensation function has significant non-linear components. While larger compensations are observed as shift magnitude is increased, the marginal increase in compensation is proportionally smaller than the marginal increase in shift magnitude. The general compensation patterns were not obviously related to the spacing of adjacent vowels in the vowel space.

For all perturbation magnitudes, the compensation response was partial. However, this is not a consequence of an inherent limitation in compensatory ability. As illustrated in Fig. 3, the average-steady-state compensation observed for Shift 3 is approximately the same magnitude as the perturbation applied in Shift 1. Indeed, the average compensations in Shift 3 are 46.6 and 60.7 Hz for F1 and F2, respectively. Thus, even though talkers were capable of fully compensating for small formant-shifts, only partial compensation was observed. Partial compensation has been consistently reported in previous studies (Houde and Jordan, 1998; Purcell and Munhall, 2006b; Villacorta *et al.*, 2007) in which different perturbation protocols were used (e.g., small perturbations on each trial), and thus partial compensation is not a function of the sudden large feedback changes used here. A number of possible explanations exist for this pattern. The sensory feedback used to control speech production is not limited to audition. Somatosensory feedback, for example, also plays a role in speech production (Tremblay *et al.*, 2003; Nasir and Ostry, 2008). When talkers compensated for the altered auditory feedback in our experiment, their change in production would produce a discrepancy in somatosensory feedback. One possibility is that the observed partial compensation is a result of the control system minimizing error across both sensory systems. Partial compensation has also been observed in saccadic adaptation, and Hopp and Fuchs (2004) suggested that the adaptation process might involve two components. The first is a rapid partial compensation followed by a very slow recalibration that is not completed in the limited time course of laboratory studies. Similar proposals have been made for two components in the adaptation process for arm movements in novel force fields (Smith *et al.*, 2006). This possibility exists for auditory feedback perturbation as well, since the exposure to altered feedback in an experiment is relatively brief.

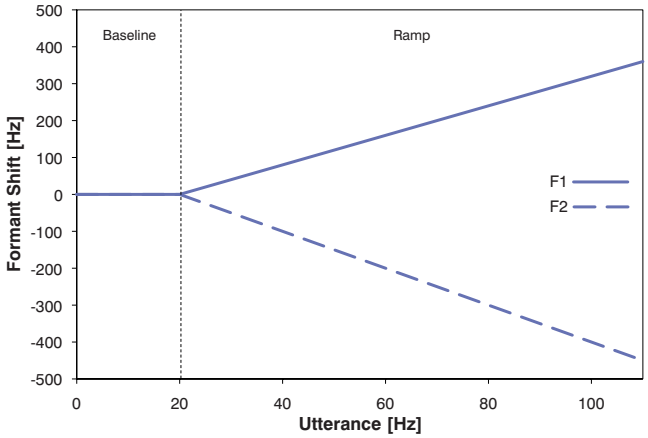


FIG. 4. (Color online) Feedback shift applied to the first formant (solid line) and second formant (dashed line) over the course of Experiment 2. The vertical dashed line denotes the boundary between the Baseline and Ramp phases.

In summary, the results from Experiment 1 suggest that compensation is partial for all perturbation magnitudes. While the function relating compensation to formant-shift magnitude may be linear for small perturbations, it is nonlinear and compressive for large perturbations similar to results for pitch perturbations (Liu and Larson, 2007).

IV. EXPERIMENT 2

In Experiment 1, the formant-shifts were applied as step-changes with each shift condition held constant for many utterances allowing talkers' compensation to reach steady state. Unfortunately, this methodology is limited as only a few formant-shift conditions can be observed in one session. Another approach is to progressively increase the magnitude of the perturbation of F1 and F2 with each utterance. Using this paradigm, the compensation to a larger number of different formant-shift magnitudes can be observed. Thus, Experiment 2 was conducted both to provide a more detailed estimate of the shape of the compensation function and to explore if the rate at which a perturbation is introduced affects compensation.

A. Participants

The participants were 21 male undergraduate students from Queen's University ranging in age from 17 to 24 years old ($M=18.5$, $SD=1.6$). All spoke English as a first language and reported no history of hearing or language disorders. All the participants had hearing thresholds less than 25 dB HL for frequencies ranging from 500 to 4000 Hz. None of the talkers participated in Experiment 1.

B. Procedure

Over the course of Experiment 2, each talker was prompted to say the word "head" a total of 110 times. The experiment consisted of two phases (see Fig. 4). In the first phase, Baseline, 20 utterances were spoken with normal

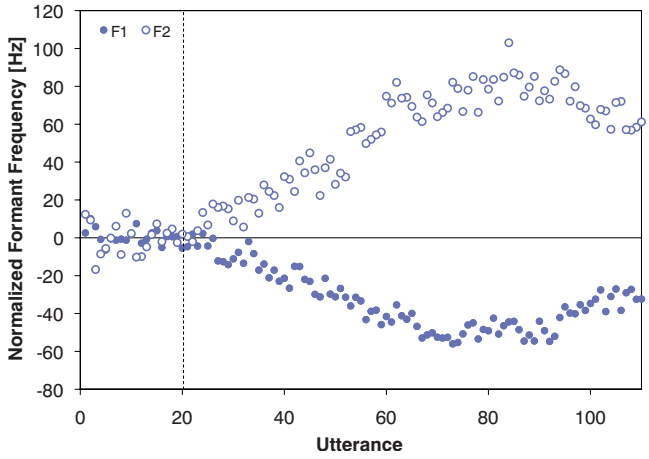


FIG. 5. (Color online) Average normalized F1 (solid circles) and F2 (open circles) frequencies for each utterance in Experiment 2. The vertical dashed line denotes the boundary between the Baseline and Ramp phases.

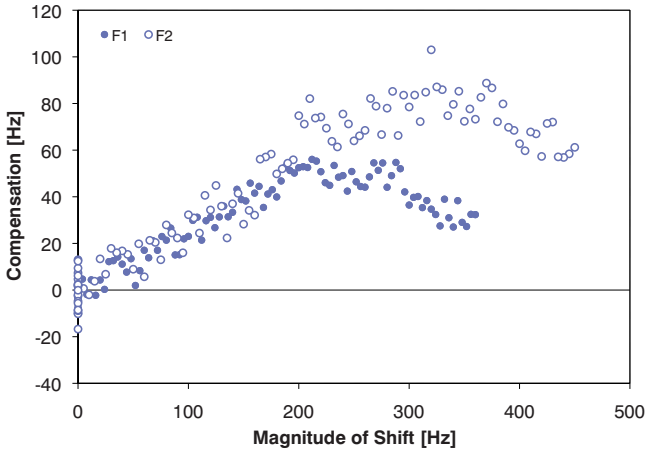


FIG. 6. (Color online) Compensation, averaged across all talkers, in F1 (solid circles) and F2 (open circles) as a function of the magnitude of formant-shift applied to acoustic feedback. Here, the magnitude of compensation is defined as the magnitude of the average normalized formant frequency. The sign of the compensation is defined as positive when the change in production is opposed to that of the formant-shift and negative when it follows that of the formant-shift.

feedback (i.e., amplified and with noise added but no shift in formant frequency) to estimate Baseline F1 and Baseline F2 values. In the second phase, Ramp, the utterances were spoken with altered feedback with F1 increased and F2 decreased in frequency. For each utterance, the magnitudes of the formant-shifts applied to F1 and F2 were increased by 4 and 5 Hz, respectively. Thus, for the 21st utterance, the shifts in frequency of F1 and F2 were +4 and −5 Hz; for the 110th utterance, the shifts in frequency of F1 and F2 were +360 and −450 Hz.

C. Results

The normalized formant frequencies for each talker's utterances were calculated in the same way as in Experiment 1 and the results averaged across talkers can be seen in Fig. 5. In the Ramp phase, talkers altered production of both F1 and F2 in a direction opposite to that of the perturbation. For the first 50 utterances of the Ramp phase (i.e., utterances 21–70), the normalized frequency for both F1 and F2 increased in magnitude approximately linearly. However, on subsequent utterances the magnitude reached a plateau and then began to decrease. To test if this decrease was statistically reliable, the average normalized formant frequency for utterances 66–70 was compared with that of utterances 101–110. For F1, the difference between the two intervals was significant ($p=0.01$) but for F2 the difference was not significant ($p>0.05$).

As in Experiment 1, we define compensation as the magnitude of the change in formant frequency from the baseline average with sign based on whether the change opposes (positive) or follows (negative) that of the perturbation. With the exception of one talker, all talkers exhibited positive compensation on average over the interval of utterances 61–70. The compensation, average across all talkers, as a function of frequency-shift magnitude is plotted in Fig. 6. From the figure, it can be seen that the functions for both F1 and F2 are similar in shape. Further, they are similar in magnitude for formant-shifts that are less than 200 Hz in magnitude. The compensations for the two formants, however, do not asymptote at the same level, nor do the compensations begin to decrease at the same point in frequency.

As can be seen in Figs. 5 and 6, production of both F1 and F2 changed approximately linearly over the first 50 utterances of the Ramp phase (utterances 21–70). The magnitudes of the frequency shifts applied during this interval ranged from 4–200 and 5–250 Hz for F1 and F2, respectively. For both F1 and F2, a linear regression was performed on the data from this interval and approximately 90% of the variance was accounted by this linear trend. The compensations for F1 and F2 were found to be approximately 25%–30% of the formant-shift magnitude (see regression values in Table II).

To test whether the small ramp perturbations used in Experiment 2 produce similar changes to those caused by the larger step perturbations used in Experiment 1, we compared the results from utterances in Experiment 2 with the matched utterances from Experiment 1. The compensations for perturbations of equivalent magnitude (100 and 200 Hz for F1 and 75, 125 and 250 for F2) were compared directly. However, in Experiment 2, the F1 perturbation was incremented by 4 Hz per utterance. Thus, there was no utterance for which the perturbation of F1 was 50 Hz (as used in Shift 1 of Experiment 1). A compensation response to a 50 Hz perturbation for F1 was approximated by averaging the compensations for perturbations of 48 and 52 Hz from the Experiment 2 data set and then compared to the result from Experiment 1. In Fig. 7, the average change in normalized formant frequency for

TABLE II. Linear regression of compensation as a function of frequency shift. The coefficients are based on the data from utterances 21–70 of Experiment 2. The 95% confidence intervals are given in parentheses.

	Slope	Intercept	R^2
F1 Compensation	0.260(±0.021)	−1.04(±2.47)	0.927
F2 Compensation	0.304(±0.032)	1.30(±4.67)	0.882

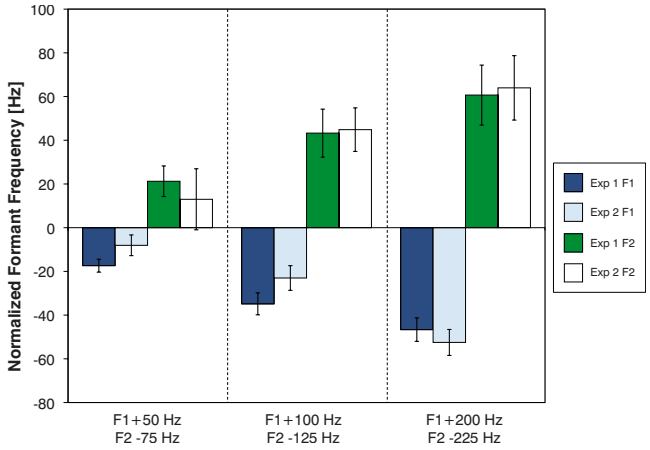


FIG. 7. (Color online) Comparison of compensation between Experiments 1 and 2. The compensations for perturbations of equivalent magnitude were compared directly. However, in Experiment 2, the F1 perturbation was incremented by 4 Hz per utterance. Thus, there was no utterance for which the perturbation of F1 was 50 Hz (as used in Experiment 1). A compensation response to a 50 Hz perturbation for F1 was approximated by averaging the compensations for perturbations of 48 and 52 Hz from the Experiment 2.

these utterances is plotted along with the results from Experiment 1. An examination of Fig. 7 suggests the results of both experiments are quite similar.

This similarity was confirmed with repeated measures ANOVAs, with magnitude of shift as a within-subjects factor and experiment as a between-subjects factor for both F1 and F2. For Experiment 2, four individuals had a tracking error in F2 for one of utterances 35, 34, and 70. Thus, their data have been omitted from the F2 ANOVA. No significant main effect of experiment was found for either F1 [$F(1,39)=797.8$, $p=0.35$) or F2 ($F(1,35)=1385.0$, $p=0.59$]. As well, no significant interaction of magnitude of shift × experiment was found for F1 [$F(2,78)=535.8$, $p=0.07$] or F2 [$F(2,70)=535.8$, $p=0.73$].

As in Experiment 1, utterances of seven different vowels were collected prior to conducting formant-shifting to determine the best AR model. Estimates for F1 and F2 were calculated for the utterances with vowels /ɪ/, /ɛ/, and /æ/ (utterances of "hid," "head," and "had," respectively). The mean and standard deviation of both F1 and F2 from talkers in Experiment 1 and 2 can be seen in Table III. The approximate end of the linear compensation occurs when perturbations lead to compensating productions of /ɛ/ that are 1.7 and 1.9 standard deviations away from normal production of F1 and F2, respectively.

D. Discussion

The purpose of Experiment 2 was to measure compensation to a large number of formant-shift magnitudes to provide a test of the sensorimotor system's maximum ability to compensate. For auditory feedback perturbations less than 200 Hz for F1 and 250 Hz for F2, subjects showed linear compensatory changes in response to the incremental changes in feedback, though with slopes less than 1. For larger feedback perturbations, subjects produced formant compensations that were increasingly less effective. The compensatory behavior in both formants approached an asymptote and then started to decrease with increasing perturbation magnitude. The maximum compensation observed in Experiment 2 was comparable to that observed in Experiment 1. These results suggest that the relationship plotted in Fig. 6 is a good estimate of the function relating compensation to the magnitude of formant-shift. Further, this response is independent of how the perturbation is introduced.

The compensation maxima observed for the male participants in this study suggest a fundamental constraint on the range of adaptive behavior based on auditory feedback. These maxima have not been apparent in the literature. Purcell and Munhall (2006b) and Munhall *et al.* (2009) used F1 perturbations of 200 Hz but did not go beyond this. Villacorta *et al.* (2007) used slightly smaller maximum perturbations. The present study also used a male population and previous studies tested females or mixed samples. Additional studies of the vowel space and compensation may help clarify the reasons for this ceiling in compensation. In particular, comparisons with data from females who have higher formant values and larger vowel spaces may be informative.

In comparing the maximum compensations of Experiments 1 and 2, it is interesting to note that the adaptation to the first shifts in Experiment 1 did not reduce the overall maximum compensation, indicating that the maximum is not easily modified. This suggests that the long term acoustic targets used by the control system were not altered during the brief time courses of these experiments.

The decrease in compensation for formant-shifts of large magnitude could be due to one or more factors. For some vowels and perturbations, there are physical constraints on the vocal tract that will limit maximum compensation. For example, decreasing F1 for the vowel /i/ would necessitate compensatory tongue movements higher than the position of the palate. However, for the vowel and perturbations tested here, physical constraints on the vowel space do not explain the observed limit on maximum compensation. As discussed

TABLE III. Average formant frequencies of three vowels for talkers in Experiments 1 and 2. One standard deviation is given in parentheses.

	/ɪ/		/ɛ/		/æ/	
	F1	F2	F1	F2	F1	F2
Exp. 1	463.0(15.4)	1878.8(43.2)	586.5(17.1)	1744.1(38.7)	741.0(24.9)	1560.9(50.9)
Exp. 2	454.1(16.4)	1865.6(44.3)	584.8(29.4)	1718.8(41.3)	728.7(20.2)	1556.4(44.7)
All	458.5(16.0)	1872.0(43.8)	585.7(24.2)	1731.1(40.1)	734.7(22.6)	1558.6(47.8)

previously in Experiment 1, the speech-motor control system also employs proprioceptive feedback. Under perturbed auditory feedback, as vowel production deviates from normal, the difference between proprioceptive and auditory feedback will also increase. Thus, the control system must manage a trade-off between the information provided by these two sensory feedback systems. When the difference between the expected and received acoustic feedback is very large, the control system may ignore or place little weighting on that feedback.

Another possibility is that the response to these large-magnitude formant-shifts is influenced by location of other vowels in the vowel space. From Table III we see that for the talkers in Experiment 2, the difference in frequency between an average production of /ɛ/ and /æ/ is approximately 144 and 162 Hz for F1 and F2, respectively. These values are less than the magnitude of the perturbation that results in maximum compensation. In Experiment 2, compensation does not begin to decrease until the acoustic feedback is shifted approximately 1 standard deviation beyond the average production of /æ/. Once again, comparison with data from female subjects would be informative.

One of the intriguing aspects of the present data is the degree to which F1 and F2 show the same feedback-scaling function. Even though the magnitude of the perturbations on any given trial differed between the formants, the compensatory behavior followed the same linear function (see Fig. 6). This is consistent with the idea that the auditory-motor error signal is processed in linear frequency for both formants. Since F1 and F2 frequencies and their perturbations span a broad region of frequency space, the similar behavior suggests that a unit change in frequency anywhere in that space is treated equally by the motor system. In contrast, vowel data from listening experiments are well accounted for by perceptual spaces that reflect the nonlinearities of pitch perception and the increasing bandwidth of peripheral auditory filters with increasing filter center frequency (see Rosner and Pickering, 1994). The linearity across the F1/F2 response corresponds with the findings of Purcell and Munhall (2006b) that raising and lowering F1 by 200 Hz produced compensations of the same size. One implication of these findings is that feedback perception may involve processes unique from standard speech perception. Zheng *et al.* (2009) showed that the neural correlates of feedback processing are distinct from listening to recordings of the same speech.

V. GENERAL DISCUSSION

Over the course of Experiments 1 and 2, talkers were exposed to altered auditory feedback in which the first and second formants were shifted in frequency. The frequency shifts increased over the course of each experiment: three step-changes in Experiment 1 and a continuously increasing ramp in Experiment 2. In both experiments, talkers compensated by changing the production of F1 and F2 in a manner opposite to that of the formant-shift perturbation. The compensation was incomplete in magnitude, approximately 25%–30% of that of the perturbation for shifts in F1 and F2 up to 200 and 250 Hz, respectively. As larger shifts were applied, compensation reached a plateau and then decreased.

The partial compensation shown in these studies has two aspects that may have different origins. First, the slope of the compensation/perturbation function is considerably less than 1, indicating that even for smaller feedback discrepancies that could be overcome, the speech-motor system tends to only make fractional adjustments in production. This was seen clearly in Experiment 1 where the compensation that was produced for the largest perturbation would have completely adjusted for the small initial perturbation in the experiment. The second aspect of the partial compensation phenomenon is that there is limit or maximum compensation that is observed. Both of these data patterns are observed for both F1 and for F2.

As indicated above, one possible account for the slope of the compensation function is that somatosensory feedback and auditory feedback jointly govern the speech-motor control system. Thus, sensorimotor control involves multisensory processing and some form of cue integration must take place. The degree to which the speech-motor control system weighs the errors from each modality is not known but evidence from audiovisual perception supports the idea that information from different sensory modalities can be flexibly combined to optimize perception (Burr and Alais, 2006; Larson *et al.*, 2008). In vowel production, the auditory and somatosensory contributions may vary for different vowels and speech contexts. For example, high vowels such as /i/ may have stronger somatosensory information and thus be less reliant on auditory feedback. Individual differences in sensory weightings might also explain the large variance in acoustic compensation observed across individuals (see Munhall *et al.*, 2009).

Previous studies on pitch compensation have demonstrated complete compensation for small perturbation magnitudes (Liu and Larson, 2007). In contrast, the function relating compensation magnitude to formant perturbation magnitude remains linear with a slope much less than 1 as perturbation magnitude is decreased. The laryngeal and oral systems differ greatly on many neural and biomechanical factors (innervation, somatosensory representation, movement range, mass of the articulators, etc.) as well as the difference in acoustic information being processed. The differences in the completeness of compensation between the two systems may have their origins in many of these factors.

While somatosensory information processing might explain the plateau and decrease in compensation observed for large perturbations, this behavior may also be due to the speech-motor control system rejecting the auditory feedback error as being spurious. For large perturbations, the auditory feedback may deviate far enough from the expected production that the feedback is attributed to some other environmental source or noise and not to the talker's own voice. Breakdowns in illusions such as the "rubber hand" illusion occur when the spatial orientation or position (Costantini and Haggard, 2007) or size of the visual stimulus (Pavani and Zampini, 2007) deviate beyond an acceptable range. For the creation of a perceived body image representation, there are limits for acceptable sensory information. The results of both

experiments suggest that if a rejection of unacceptable sensory information occurs for auditory feedback, it requires very large perturbations with magnitudes greater than the difference between adjacent vowels (Table III).

In both experiments, F1 and F2 were perturbed simultaneously. So far, our analysis has examined the compensation in both formants independently. However, the motor control system may or may not be able to compensate for perturbations in both formants independently. The system may preferentially compensate for perturbations of one formant over another. Alternatively, the system may attempt to compensate in a manner that requires coordinated changes in the two formants. For example, if the direction of compensation is opposite that of the perturbation in an F2-F1 vowel space plane, this will yield a specific relationship between F1 and F2.

This relationship between F1 and F2 can be shown more formally. Let ΔF1 and ΔF2 represent the magnitudes of formant-shifts applied to F1 and F2, respectively, and let $C_{F1}(\Delta F1)$ and $C_{F2}(\Delta F2)$ represent the compensations in F1 and F2 in response to the perturbation. If we assume that the motor control system produces compensation in a direction opposite that of the perturbation, then the following relationship holds:

$$\frac{C_{F1}(\Delta F1)}{C_{F2}(\Delta F2)} = \frac{\Delta F1}{\Delta F2}. \quad (1)$$

In Experiment 2, the ratio of ΔF1 to ΔF2 was held constant. Thus,

$$\Delta F2 = k\Delta F1. \quad (2)$$

Substituting Eq. (2) into Eq. (1) and rearranging yield the following relationship between C_{F1} and C_{F2}:

$$kC_{F1}(\Delta F1) = C_{F2}(k\Delta F1). \quad (3)$$

Thus, the assumption that the motor control system produces compensation in a direction opposite that of the perturbation implies a relationship between the compensation and perturbation components. As a result of Eq. (3), if C_{F1} is linear then C_{F2} is constrained to be linear with identical slope and with an intercept that is a factor of k larger than that of C_{F1}.

In both experiments, the function relating compensation to formant-shift magnitude is similar for F1 and F2; for the linear portion of the compensatory behavior (Fig. 6), the response slope is the same for both formants. In Fig. 8, the F1 and F2 compensation results from Experiment 2 are plotted in an F1-F2 vowel space context. The dashed line indicates the direction exactly opposite that of the perturbation in the F1-F2 space. The results from the first 70 utterances (circles) lie along the dashed line indicating that compensation was indeed opposite that of the perturbation. The results from the last 40 utterances (triangles) deviate from the dashed line and indicate that compensation is no longer in a direction opposite that of the perturbation. Thus, for the linear portion of the compensatory response (circles), the motor control system is able to compensate in a manner such that the compensations in F1 and F2 are linked. For large perturbations, the compensatory responses are both non-linear and no longer show the proportional linkages between formants. Thus, the proportional linkage between F1 and F2 observed for the first 70 utterances is a controlled response and not a physical necessity imposed by vocal tract geometry.

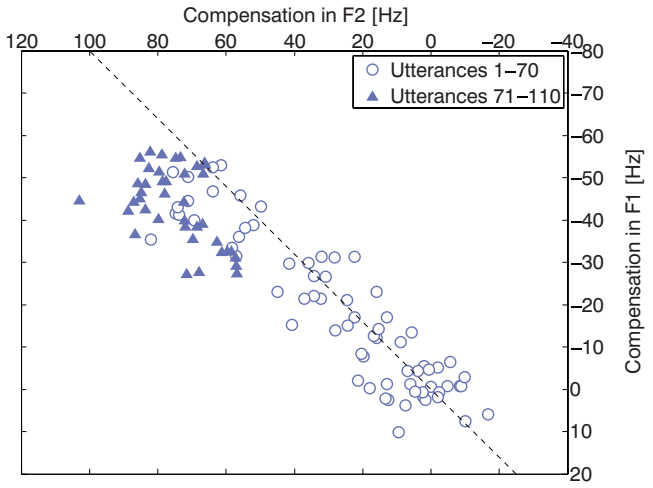


FIG. 8. (Color online) Scatter plot of F1 and F2 compensation results from Experiment 2 plotted in an F1-F2 vowel space context. The dashed line indicates the direction geometrically opposite that of the perturbation in the F1-F2 space. The results from the first 70 utterances (circles) lie along the dashed line indicating that compensation was directly opposite that of the perturbation. The results from the last 40 utterances (triangles) deviate from the dashed line and indicate that compensation is no longer in a direction opposite that of the perturbation.

Vowels in the English front vowel space vary along both F1 and F2 and as the vowels move from high to low vowel positions, they move progressively from front to back. Thus, front vowel differences on average form an angular trajectory in linear formant space. Our perturbations in this study are along this formant path as were the perturbations used by Houde and Jordan (2002). The degree to which the structure of the front vowel space constrains the direction and magnitude of compensatory behavior cannot be determined by the present data. However, the results of both experiments suggest that the error in auditory feedback used by the motor control system is not dependent on the spacing of vowels nor presumably by their category boundaries.

If the feedback error was processed post-categorically, then one would expect very little compensation for small formant-shift magnitudes and a sudden large increase in compensation once the formant-shift magnitudes crossed a categorical threshold. This behavior was not observed in Experiment 2. The results suggest that the function relating compensation to formant-shift magnitude is approximately linear for perturbations spanning the interval between two adjacent vowels. Thus, the control system is responding to a difference between the intended and received formant frequencies and not a difference between the intended and received vowel categories. This allows auditory feedback to be used for tighter control of formant frequency than would be possible with a system that only acted to maintain vowel category identity.

The emerging view is a system that maps linear formant frequency to movements over a restricted range surrounding a vowel. This mapping is involved in partially compensating for mismatches in auditory feedback within a few utterances

and thus acts as a rapid stabilizing system for speech motor control. The relative independence of the spatial movement dimensions in the vocal tract, and thus formant independence, is unknown and requires additional studies.

ACKNOWLEDGMENTS

This research was supported by the National Institute of Deafness and Communicative Disorders Grant No. DC-08092 and the Natural Sciences and Engineering Research Council of Canada. Experiment 1 was conducted by Robyn Goldberg as part of her Bachelor of Arts Honors thesis. We wish to thank Jeremy Gretton and Meighen Roes for their comments on the manuscript and Bryan Burt for his assistance in conducting Experiment 2.

Bauer, J. J., Mittal, J., Larson, C. R., and Hain, T. C. (**2006**). "Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude," J. Acoust. Soc. Am. **119**, 2363–2371.

Beurze, S. M., Van Pelt, S., and Medendorp, W. P. (**2006**). "Behavioral reference frames for planning human reaching movements," J. Neurophysiol. **96**, 352–362.

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (**1998**). "Voice F0 responses to manipulations in pitch feedback," J. Acoust. Soc. Am. **103**, 3153–3161.

Burr, D., and Alais, D. (**2006**). "Combining visual and auditory information," Prog. Brain Res. **155**, 243–258.

Churchland, A. K., and Lisberger, S. G. (**2009**). "Gain control in human smooth-pursuit eye movements," J. Neurophysiol. **87**, 2936–2945.

Costantini, M., and Haggard, P. (**2007**). "The rubber hand illusion: Sensitivity and reference frame for body ownership," Conscious. Cogn **16**, 229–240.

Cowie, R., and Douglas-Cowie, E. (**1992**). *Postlingually Acquired Deafness: Speech Deterioration and the Wider Consequences* (Mouton de Gruyter, New York).

Hopp, J. J., and Fuchs, A. F. (**2004**). "The characteristics and neuronal substrate of saccadic eye movement plasticity," Prog. Neurobiol. **72**, 27–53.

Houde, J. F., and Jordan, M. I. (**1998**). "Sensorimotor adaptation in speech production," Science **279**, 1213–1216.

Houde, J. F., and Jordan, M. I. (**2002**). "Sensorimotor adaptation of speech: I. Compensation and adaptation," J. Speech Lang. Hear. Res. **45**, 295–310.

Jones, J. A., and Munhall, K. G. (**2005**). "Remapping auditory-motor representations in voice production," Curr. Biol. **15**, 1768–1772.

Kalveram, K. T., and Jäncke, L. (**1989**). "Vowel duration and voice onset time for stressed and nonstressed syllables in stutterers under delayed auditory feedback condition," Folia Phoniatr (Basel) **41**, 30–42.

Kawato, M. (**1989**). "Motor theory of speech perception revisited from the minimum torque change neural network model," in Eighth Symposium on Future Electron Devices, Tokyo, Japan, pp. 141–150.

Krakauer, J. W., Pine, Z. M., Ghilardi, M. F., and Ghez, C. (**2000**), "Learning of visuomotor transformations for vectorial planning of reaching trajectories," J. Neurosci. **20**, 8916–8924.

Larson, C. R., Altman, K. W., Liu, H., and Hain, T. C. (**2008**). "Interactions between auditory and somatosensory feedback for voice F0 control," Exp. Brain Res. **187**, 613–621.

Liu, H., and Larson, C. R. (**2007**). "Effects of perturbation magnitude and voice F0 level on the pitch shift reflex," J. Acoust. Soc. Am. **122**, 3671–3677.

Munhall, K. G., MacDonald, E. N., Byrne, S. K., and Johnsrude, I. (**2009**). "Talkers alter vowel production in response to real-time formant perturbation even when instructed to resist compensation," J. Acoust. Soc. Am. **125**, 384–390.

Nasir, S. M., and Ostry, D. J. (**2008**). "Speech motor learning in profoundly deaf adults," Nat. Neurosci. **11**, 1217–1222.

Oller, D. K., and Eilers, R. E. (**1988**). "The role of audition in infant babbling," Child Dev. **59**, 441–449.

Orfandidis, S. J. (**1988**). *Optimum Signal Processing, An Introduction* (MacMillan, New York).

Pavani, F., and Zampini, M. (**2007**). "The role of hand size in the fake-hand illusion paradigm," Perception **36**, 1547–1554.

Pruszynski, J. A., Kurtzer, I. L., Lillicrap, T. P., and Scott, S. H. (**2009**). "Temporal evolution of "automatic gain-scaling"," J. Neurophysiol. **102**, 992–1003.

Purcell, D. W., and Munhall, K. G. (**2006a**). "Compensation following real-time manipulation of formants in isolated vowels," J. Acoust. Soc. Am. **119**, 2288–2297.

Purcell, D. W., and Munhall, K. G. (**2006b**). "Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation," J. Acoust. Soc. Am. **120**, 966–977.

Rosner, B. S., and Pickering, J. B. (**1994**). *Vowel Perception and Production* (Oxford University Press, Oxford).

Saunders, J. A., and Knill, D. C. (**2004**). "Visual feedback control of hand movements," J. Neurosci. **24**, 3223–3234.

Smith, M. A., Ghazizadeh, A., and Shadmehr, R. (**2006**), "Interacting adaptive processes with different timescales underlie short-term motor learning," PLoS Biology **4**, e179.

Tin, C., and Poon, C.-S. (**2005**). "Internal models in sensorimotor integration: Perspectives from adaptive control theory," J. Neural Eng. **2**, S147–S163.

Tremblay, S., Schiller, D. M., and Ostry, D. J. (**2003**). "Somatosensory basis of speech production," Nature (London) **423**, 866–869.

Villacorta, V. M., Perkell, J. S., and Guenther, F. H. (**2007**). "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception," J. Acoust. Soc. Am. **122**, 2306–2319.

Wolpert, D. M., and Kawato, M. (**1998**). "Multiple paired forward and inverse models for motor control," Neural Networks **11**, 1317–1329.

Zheng, Z. Z., Munhall, K. G., and Johnsrude, I. (**2009**). "Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production," J. Cogn Neurosci. In press.

A mathematical model of vowel identification by users of cochlear implants

Elad Sagi[a)]
Department of Otolaryngology, New York University School of Medicine, New York, New York 10016

Ted A. Meyer
Department of Otolaryngology–HNS, Medical University of South Carolina, Charleston, South Carolina 29425

Adam R. Kaiser and Su Wooi Teoh
Department of Otolaryngology, Head and Neck Surgery, DeVault Otologic Research Laboratory, Indiana University School of Medicine, Indianapolis, Indiana 46202

Mario A. Svirsky
Department of Otolaryngology, New York University School of Medicine, New York, New York 10016

(Received 1 January 2009; revised 25 November 2009; accepted 30 November 2009)

A simple mathematical model is presented that predicts vowel identification by cochlear implant users based on these listeners' resolving power for the mean locations of first, second, and/or third formant energies along the implanted electrode array. This psychophysically based model provides hypotheses about the mechanism cochlear implant users employ to encode and process the input auditory signal to extract information relevant for identifying steady-state vowels. Using one free parameter, the model predicts most of the patterns of vowel confusions made by users of different cochlear implant devices and stimulation strategies, and who show widely different levels of speech perception (from near chance to near perfect). Furthermore, the model can predict results from the literature, such as Skinner, *et al.* [(1995). Ann. Otol. Rhinol. Laryngol. **104**, 307–311] frequency mapping study, and the general trend in the vowel results of Zeng and Galvin's [(1999). Ear Hear. **20**, 60–74] studies of output electrical dynamic range reduction. The implementation of the model presented here is specific to vowel identification by cochlear implant users, but the framework of the model is more general. Computational models such as the one presented here can be useful for advancing knowledge about speech perception in hearing impaired populations, and for providing a guide for clinical research and clinical practice.
 [DOI: 10.1121/1.3277215]

PACS number(s): 43.71.An, 43.66.Ts, 43.71.Es, 43.71.Ky [MSS] Pages: 1069–1083

I. INTRODUCTION

Cochlear implants (CIs) represent the most successful example of a neural prosthesis that restores a human sense. The last two decades have been witness to systematic improvements in technology and clinical outcomes, yet substantial individual differences remain. The reference to the individual CI user is important because typical fitting procedures for CIs are guided primarily by the listener's preference, by what "sounds better," independent of their speech perception (which does not always correlate perfectly with subjective preference; Skinner *et al.*, 2002). Several researchers have suggested that one of the factors limiting performance in many CI users is precisely this lack of performance-based fitting. If CI users were fit according to their specific perceptual and physiological strengths and weaknesses clinical outcomes might improve significantly (Shannon, 1993). Yet, assessing the effect of all possible fitting parameters on a given CI user's speech perception is not feasible. In this regard, quantitative models may prove a useful aid to clinical practice. In the present study we propose a mathematical model that explains a CI user's vowel identification based on their ability to identify average formant center frequency values, and assess this model's ability to predict vowel identification performance under two CI device setting manipulations.

One example that demonstrates how such a model might guide clinical practice relates to the CI user's "frequency map," i.e., the frequency bands assigned to each stimulation channel. More than 20 years after the implantation of the first multichannel CIs the optimal frequency map remains unknown, either on average or for each specific CI user. The lack of evidence in this case is not total, however. Skinner *et al.* (1995) reported that a certain frequency map (frequency allocation table or FAT No. 7) used with the Nucleus-22 device resulted in better speech perception scores for a group of CI users than the frequency map that was the default for the clinical fitting software, and also the most widely used map at the time (FAT No. 9). Skinner *et al.*'s (1995) study resulted in a major shift and FAT No. 7 became much more commonly used by CI audiologists. Yet,

[a)]Author to whom correspondence should be addressed. Electronic mail: elad.sagi@nyumc.org

with the large number of possible combinations, testing the whole parametric space of frequency map manipulations is both time and cost prohibitive. A possible alternative would be to use a model that provides reasonable predictions of speech perception under each FAT, and test a listener's performance using only the subset of FATs that the model deems most promising.

Several acoustic cues have been shown to influence vowel perception by listeners with normal hearing, including steady-state formant center frequencies (Peterson and Barney, 1952), formant frequency ratios (Chistovich and Lublinskaya, 1979), fundamental frequency, formant trajectories during the vowel, and vowel duration (Hillenbrand *et al.*, 1995; Syrdal and Gopal, 1986; Zahorian and Jagharghi, 1993), as well as formant transitions from and into adjacent phonemes (Jenkins *et al.*, 1983). That is, listeners with normal hearing can utilize the more subtle, dynamic changes in formant content available in the acoustic signal. Supporting this notion is the observation that listeners with normal hearing are highly capable of discriminating small changes in formant frequency. Kewley-Port and Watson (1994) found that listeners with normal hearing could detect differences in formant frequency of about 14 Hz in the range of F1 and about 1.5% in the range of F2. Hence, when two vowels consist of similar steady-state formant values, listeners with normal hearing have sufficient acuity to differentiate between these vowels based on small differences in formant trajectories.

In contrast, due to device and/or sensory limitations, listeners with CIs may only be able to utilize a subset of these acoustic cues (Chatterjee and Peng, 2008; Fitzgerald *et al.*, 2007; Hood *et al.*, 1987; Iverson *et al.*, 2006; Kirk *et al.*, 1992; Teoh *et al.*, 2003). For example, in terms of formant frequency discrimination, Fitzgerald *et al.* (2007) found that users of the Nucleus-24 device could discriminate about 50–100 Hz in the F1 frequency range and about 10% in the F2 frequency range, i.e., roughly five times worse than the normal hearing data reported by Kewley-Port and Watson (1994). Hence, some of the smaller formant changes that help listeners with normal hearing identify vowels may not be perceptible to CI users. Indeed, Kirk *et al.* (1992) demonstrated that when static formant cues were removed from vowels, normal hearing listeners were able to identify these vowels at levels significantly above chance whereas CI users could not. Furthermore, little or no improvement in vowel scores was found for the CI users when dynamic formant cues were added to static formant cues. In more recently implanted CI users, Iverson *et al.* (2006) found that CI users could utilize the larger dynamic formant changes that occur in diphthongs in order to differentiate these vowels from monophthongs, but it was also found that normal hearing listeners could utilize this cue to a far greater extent than CI users.

CI users' limited access to these acoustic cues gives us the opportunity to test a very simple model of vowel identification that relies only on steady-state formant center frequencies. Clearly, such a simple model would be insufficient to explain vowel identification in listeners with normal hearing, but it may be adequate to explain vowel identification in current CI users. The model employed in the present study is an application of the multidimensional phoneme identification or MPI model (Svirsky, 2000, 2002), which was developed as a general framework to predict phoneme identification based on measures of a listener's resolving power for a given set of speech cues. In the present study, the model is tested on four experiments related to vowel identification by CI users. The first two were conducted by us and consist of vowel and first-formant identification data from CI listeners. The purpose of these two data sets was to test the model's ability to account for vowel identification by CI users, and to assess the model's account of relating vowel identification to listeners' ability to resolve steady-state formant center frequencies. The third and fourth data sets were extracted from Skinner *et al.*, 1995 and Zeng and Galvin, 1999, respectively. These two data sets were used to test the MPI model's ability to make predictions about how changes in two CI device fitting parameters (FAT and electrical dynamic range, respectively) affect vowel identification in these listeners.

II. GENERAL METHODS

A. MPI model

The mathematical framework of the MPI model is a multidimensional extension of Durlach and Braida's single-dimensional model of loudness perception (Durlach and Braida, 1969; Braida and Durlach, 1972), which is in turn based on earlier work by Thurstone (1927a, 1927b) among others. The MPI model is more general than the Durlach–Braida model not only due to the fact that it is multidimensional, but also because loudness need not be one of the model's dimensions. Let us first define some terms and assumptions that underlie the MPI model. We assume that a *phoneme* (vowel or consonant) is identified based on several *acoustic cues*. A given acoustic cue assumes characteristic values for each phoneme along the respective *perceptual dimension*. A subject's resolving power, or just-noticeable-difference (JND), along this perceptual dimension can be measured with appropriate psychophysical tests. The JNDs for all dimensions are subject-specific inputs to the MPI model. Because listeners have different JND values along any given dimension, the model's predictions can be different for each subject.

1. General implementation: Three steps

The implementation of the MPI model in the present study can be summarized in three steps. *First*, we must hypothesize what the relevant perceptual dimensions are. These hypotheses are informed by knowledge about acoustic-phonetic properties of speech, and about the auditory psychophysical capabilities of CI users (Teoh *et al.*, 2003). *Second*, we have to measure the mean location of each phoneme along each postulated perceptual dimension. These locations are uniquely determined by the physical characteristics of the stimuli and the selected perceptual dimensions. *Third*, we must measure the subjects' JNDs along each perceptual dimension using appropriate psychophysical tests, or leave the JNDs as free parameters to determine how well the model could fit the experimental data. Because there are several

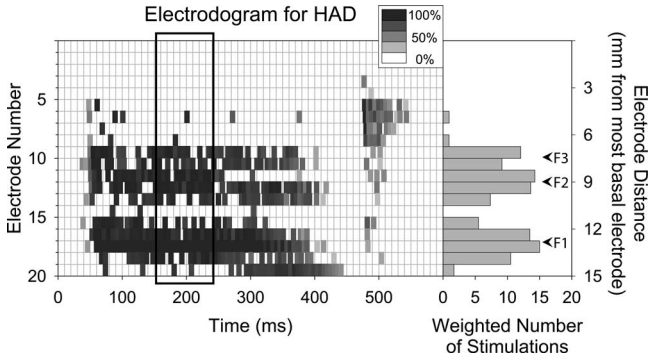


FIG. 1. Electrodogram of the vowel in "had" obtained with the Nucleus device. Higher electrode numbers refer to more apical or low-frequency encoding electrodes. Charge magnitude is depicted as a gray-scale from 0% (light) to 100% (dark) of dynamic range. Rectangle centered at 200 ms represents the time window used to compile histogram on the right, which represents a weighted count of the number of times each electrode was stimulated. Locations of mean formant energies (F1, F2, and F3 in millimeters from most basal electrode) extracted from histogram.

ways to measure JNDs, these two approaches could yield JND values that are related, but not necessarily the same.

Step 1. The proposed set of relevant perceptual dimensions for the present study of vowel identification by CI users is the mean locations along the implanted electrode array of stimulation pulses corresponding to the first three formant frequencies, i.e., F1, F2, and F3. These dimensions are measured in units of distance along the electrode array (e.g., mm from most basal electrode) rather than frequency (Hz). In experiment 1, different combinations of these dimensions are explored to determine a set of dimensions that best describe each CI subject's vowel confusion matrix. In experiments 3 and 4, the F1F2F3 combination is used exclusively.

Step 2. Locations of mean formant energy along the electrode array were obtained from "electrodograms" of vowel tokens. The details of how electrodograms were obtained are in Sec. II B. An electrodogram is a graph that includes information about which electrode is stimulated at a given time, and at what current amplitude and pulse duration. Depending on the allocation of frequency bands to electrodes, an electrodogram depicts how formant energy becomes distributed over a subset of electrodes. The left panel of Fig. 1 is an example of an electrodogram of the vowel "had" obtained with the Nucleus device where higher electrode numbers refer to more apical or low-frequency encoding electrodes. For each pulse, the amount of electrical charge (i.e., current times pulse duration) is depicted as a gray-scale from 0% (light) to 100% (dark) of the dynamic range, where 0% represents threshold stimulation level and 100% represents the maximum comfortable level. We are particularly concerned with how formant energies F1, F2, and F3 are distributed along the array over a time window centered at the middle portion of the vowel stimulus (rectangle in Fig. 1). The right panel of Fig. 1 is a histogram of the number of times each electrode was stimulated over this time window, weighted by the amount of electrical charge above threshold for each current pulse (measured with the percentage of the dynamic range described above). The histogram's vertical axis is in units of millimeters from the most basal electrode as measured along the length of the electrode array. These units are inferred from the inter-electrode distance of a given CI device (e.g., 0.75 mm for the Nucleus-22 and Nucleus-24 CIs and 2 mm for the Advanced Bionics Clarion 1.2 CI). To obtain the location of mean formant energy along the array for each formant, the histogram was first partitioned into regions of formant energies (one for each formant) and then the mean location for each formant was calculated from the portion of the histogram within each region. The frequency ranges selected to partition histograms into formant regions, based on the average formant measurements of Peterson and Barney (1952) for male speakers, were $F1 \leq 800$ Hz $< F2 \leq 2250$ Hz $< F3 \leq 3000$ Hz for all vowels except for "heard," for which $F1 \leq 800$ Hz $< F2 \leq 1700$ Hz $< F3 \leq 3000$ Hz. In Fig. 1, the locations of mean formant energies are indicated to the right of the histogram. Whereas each electrode is located at discrete points along the array, the mean location of formant energy varies continuously along the array.

Step 3. JND was varied as a free parameter with one degree of freedom until a predicted matrix was obtained that "best-fit" the observed experimental matrix. That is, in a given best-fit model matrix, JND was assumed to be equal for each perceptual dimension.

2. MPI model framework

Qualitative description. The MPI model is comprised of two sub-components, an *internal noise model* and a *decision model*. The internal noise model postulates that a phoneme produces percepts that are represented by a *Gaussian* probability distribution in a *multidimensional* perceptual space. For the sake of simplicity it is assumed that perceptual dimensions are *independent* (orthogonal) and distances are *Euclidean*. These distributions represent the assumption that successive presentations of the same stimulus result in somewhat different percepts, due to imperfections in the listener's internal representation of the stimulus (i.e., *sensory noise* and *memory noise*). The center of the Gaussian distribution corresponding to a given phoneme is determined by the physical characteristics of the stimulus along each dimension. The standard deviation along each dimension is equal to the listener's JND for the stimulus' physical characteristic along that dimension. Smaller JNDs produce narrower Gaussian distributions and can result in fewer confusions among different sounds.

The decision model employed in the present study is similar to the approach employed by Braida (1991) and Ronan *et al.* (2004), and describes how subjects categorize speech sounds based on the perceptual input. According to the decision model, the multidimensional perceptual space is subdivided into non-overlapping *response regions*, one for each phoneme. Within each response region there is a *response center*, which represents the listener's expectation about how a given phoneme should sound. One interpretation of the response center concept is that it reflects a subject's expected sensation in response to a stimulus (e.g., a prototype or "best exemplar" of the subject's phoneme category). When a percept (generated by the internal noise model) falls in the response region corresponding to a given phoneme (or, in other words, when the percept is closer to

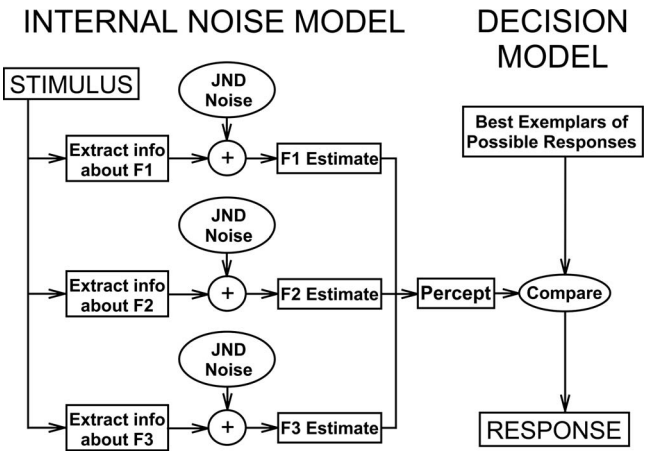


FIG. 2. Summary of the two-step iteration involved in a three-dimensional F1F2F3 MPI model for vowel identification. Internal noise model generates a percept by adding noise (proportional to input JNDs) to the formant locations of a given vowel. Decision model selects response center (i.e., best exemplar of a given vowel) with formant locations closest to those of percept.

the response center of that phoneme than to any other response center), then the decision model predicts that the subject will select that phoneme as the one that she/he heard. The ideal experienced listener would have response centers that are equal to the *stimulus centers*, which we define as the average location of tokens for a particular phoneme in the perceptual space. In other words, this listener's expectations match the actual physical stimuli. When this is not the case, one can implement a bias parameter to accommodate for differences between stimulus and response centers. In the present study, all listeners are treated as ideal experienced listeners so that stimulus and response centers are equal.

Using a Monte Carlo algorithm that implements each component of the MPI model, one can simulate vowel identifications to any desired number of iterations, and compile the results into a confusion matrix. Each iteration can be summarized as a two-step process. First, one uses the internal noise model to generate a sample percept for a given phoneme. Second, one uses the decision model to select the phoneme that has the response center closest to the percept. Figure 2 illustrates a block diagram of the two-step iteration involved in a three-dimensional MPI model for vowel identification, where the three dimensions are the average locations along the electrode array stimulated in response to the first three formants: F1, F2, and F3.

Mathematical formulation. The Gaussian distribution that underlies the internal noise model for the F1F2F3 perceptual dimension combination can be described as follows. Let E_i represent the ith vowel out of the nine possible vowels used in the present study. Let E_{ij} represent the jth token of E_i, out of the five possible tokens used for this vowel in the present study. Each token is described as a point in the three dimensional F1F2F3 perceptual space. Let this point T be described by the set $T=\{T_{\mathrm{F1}},T_{\mathrm{F2}},T_{\mathrm{F3}}\}$, so that $T_{\mathrm{F2}}(E_{ij})$ represents the F2 value of the vowel token E_{ij}. Let $J=\{J_{\mathrm{F1}},J_{\mathrm{F2}},J_{\mathrm{F3}}\}$ represent the subject's set of JNDs across perceptual dimensions so that J_{F2} represents the JND along the F2 dimension. Now let $X=\{x_{\mathrm{F1}},x_{\mathrm{F2}},x_{\mathrm{F3}}\}$ be a set of random variables across perceptual dimensions, so that x_{F2} is a random variable describing any possible location along the F2 dimension. Since perceptual dimensions are assumed to be independent, the normal probability density describing the likelihood of the location of a percept that arises from vowel token E_{ij} can be defined as $P(X|E_{ij})$ where

$$P(X|E_{ij}) = \frac{1}{J_{\mathrm{F1}}J_{\mathrm{F2}}J_{\mathrm{F3}}(\sqrt{2\pi})^3} e^{-(x_{\mathrm{F1}}-T_{\mathrm{F1}}(E_{ij}))^2/2J_{\mathrm{F1}}{}^2} \times e^{-(x_{\mathrm{F2}}-T_{\mathrm{F2}}(E_{ij}))^2/2J_{\mathrm{F2}}{}^2} e^{-(x_{\mathrm{F3}}-T_{\mathrm{F3}}(E_{ij}))^2/2J_{\mathrm{F3}}{}^2}. \quad (1)$$

Each presentation of E_{ij} results in a sensation that is modeled as a point that varies stochastically in the three dimensional F1F2F3 space following the Gaussian distribution $P(X|E_{ij})$. This point, or "percept," can be defined as $X'=\{x'_{\mathrm{F1}},x'_{\mathrm{F2}},x'_{\mathrm{F3}}\}$, where x'_{F2} is the coordinate of X' along the F2 dimension. The prime script is used here to distinguish X' as a point in X. The stochastic variation of X' arises from a combination of "sensation noise," which is a measure of the observer's sensitivity to stimulus differences along the relevant dimension, and "memory noise," which is related to uncertainty in the observer's internal representation of the phonemes within the experimental context.

In the decision model, the percept X' is categorized by finding the closest response center. Let $R(E_k)=\{R_{\mathrm{F1}}(E_k),R_{\mathrm{F2}}(E_k),R_{\mathrm{F3}}(E_k)\}$ be the location of the *response center* for the kth vowel so that $R_{\mathrm{F2}}(E_k)$ represents the location of the response center for this vowel along the F2 perceptual dimension. For vowel E_k, the stimulus center can be represented as $S(E_k)=\{S_{\mathrm{F1}}(E_k),S_{\mathrm{F2}}(E_k),S_{\mathrm{F3}}(E_k)\}$, where $S_{\mathrm{F2}}(E_k)$ is the location of the stimulus center for vowel E_k along the F2 perceptual dimension. $S_{\mathrm{F2}}(E_k)$ is equal to the average F2 value across the five tokens of E_k [i.e., the average of $T_{\mathrm{F2}}(E_{kj})$ for $j=1,\ldots,5$]. When a listener's expected sensation in response to a given phoneme is unbiased, then we say that the response center is equal to the stimulus center; i.e., $R(E_k)=S(E_k)$. Conversely, if the listener's expectations (represented by the response centers) are not in line with the physical characteristics of the stimulus (represented by the stimulus centers), then we say that the listener is a biased observer. In the present study, all listeners are treated as unbiased observers so that response centers are equal to stimulus centers.

The closest response center to the percept X' can be determined by comparing X' with all response centers $R(E_z)$ for $z=1,\ldots,n$ using the Euclidean measure

$$D_z = \sqrt{\left(\frac{x'_{\mathrm{F1}}-R_{\mathrm{F1}}(E_z)}{J_{\mathrm{F1}}}\right)^2+\left(\frac{x'_{\mathrm{F2}}-R_{\mathrm{F2}}(E_z)}{J_{\mathrm{F2}}}\right)^2+\left(\frac{x'_{\mathrm{F3}}-R_{\mathrm{F3}}(E_z)}{J_{\mathrm{F3}}}\right)^2}. \quad (2)$$

If $R(E_k)$ is the closest response center to the percept X' (in other words, if D_z is minimized when $z=k$), then the phoneme that gave rise to the percept (i.e., E_i) was identified as phoneme E_k and one can update Cell_{ik} in the confusion matrix accordingly. Using a Monte Carlo algorithm, the process of generating a percept with Eq. (1) and categorizing this percept using Eq. (2) can be continued for all vowel tokens to any desired number of iterations. It is important to note that the JNDs that appear in the denominator of Eq. (2) are

used to ensure that all distances are measured as multiples of the relevant just-noticeable-difference along each perceptual dimension.

B. Stimulus measurements

Electrodograms of the vowel tokens used in the present study were obtained for two types of Nucleus device and one type of Advanced Bionics device using specialized hardware and software. In both cases, vowel tokens were presented over loudspeaker to the device's external microphone in a sound attenuated room. The microphone was placed approximately 1 m from the loudspeaker and stimuli were presented at 70 dB C-weighted sound pressure level (SPL) as measured next to the speech processor's microphone.

Depending on the experiment conducted in the present study, measurements were obtained from either a standard Nucleus-22 device with a Spectra body-worn processor or a standard Nucleus-24 device with a Sprint body-worn processor. In either case, the radio frequency (RF) information transmitted by the processor (through its transmitter coil) was sent to a Nucleus dual-processor interface (DPI). The DPI, which was connected to a PC, captured and decoded the RF signal, which was then read by a software package called sCILab (Bögli *et al.*, 1995; Wai *et al.*, 2003). The speech processor was programmed with the spectral peak (SPEAK) stimulation strategy where the thresholds and maximum stimulation levels were fixed to 100 and 200 clinical units, respectively. Depending on the experiment, the frequency allocation table was set to FAT No. 7 and/or FAT No. 9.

For the Advanced Bionics device, electrodograms were obtained by measuring current amplitude and pulse duration directly from the electrode array of an eight-channel Clarion 1.2 "implant-in-a-box" connected to an external speech processor (provided by Advanced Bionics Corporation, Valencia, CA, USA). The processor was programmed with the continuous interleaved sampling (CIS) stimulation strategy and with the standard frequency-to-electrode assigned by the processor's programming software. For each electrode, the signal was passed through a resistor and recorded to PC by one channel of an eight-channel IOtech WaveBook/512H Data Acquisition System [12-bit analogue to digital (A/D) conversion sampled at 1 MHz].

C. Comparing predicted and observed confusion matrices

Two measures were used to assess the ability of the MPI model to generate a matrix that best predicted a listener's observed vowel confusion matrix. The first method provides a global measure of how a model matrix generated with the MPI model differs from an experimental matrix. The second method examines how the MPI model accounts for the specific error patterns observed in the experimental matrix. For both measures, matrix elements are expressed in units of percentage so that each row sums to 100%.

1. Root-mean-square difference

The first measure is the root-mean-square (rms) difference between the predicted and observed matrices. With this measure, the differences between each element of the observed matrix and each element of the predicted matrix are squared and summed. The sum is divided by the total number of elements in the matrix (e.g., $9\times9=81$) to give the mean-square, and its square-root the rms difference in units of percent. With this measure, the predicted matrix that minimized rms was defined as the best-fit to the observed matrix.

TABLE I. Example of a 2×2 comparison table comparing the vowel pairs confused more than a certain percentage of the time (5% in this case) by the subjects, to the vowel pairs that the model predicted would be confused.

Threshold=5%	Predicted≥5%	Predicted<5%
Observed≥5%	A=5	B=1
Observed<5%	C=1	D=29

2. Error patterns

The second measure examines the extent to which the MPI model predicts the pattern of vowel pairs that were confused (or not confused) more frequently than a predefined percentage of the time. Vowel pairs were analyzed without making a distinction as to the direction of the confusion within a pair, e.g., "had" confused with "head" vs "head" confused with "had." That is, in a given confusion matrix, the percentage of time the *i*th and *j*th vowel pair was confused is equal to $(\mathrm{Cell}_{ij}+\mathrm{Cell}_{ji})/2$. This approach was adopted to simplify the fitting criteria between observed and predicted matrices and should not be taken to mean that confusions within a vowel pair are assumed to be symmetric. In fact, there is considerable evidence that vowel confusion matrices are not symmetric either for normal hearing listeners (Phatak and Allen, 2007), or for the CI users in the present study.

After calculating the percentage of vowel pair confusions in both the observed and predicted matrices, a 2×2 contingency table can be constructed based on a threshold percentage. Table I shows an example of such a contingency table using a threshold of 5%. Out of 36 possible vowel pair confusions, cell A (upper left) is the number of true positives, i.e., confusions (≥5%) made by the subject and predicted by the model. Cell B (upper right) is the number of false negatives, i.e., confusions (≥5%) made by the subject but not predicted by the model. Cell C (lower left) is the number of false positives, i.e., confusions (≥5%) predicted to occur by the model but not made by the subject. Lastly, cell D (lower right) is the number of true negatives, i.e., confusions not made by the subject (<5%) and also predicted not to occur by the model (<5%). With this method of matching error patterns, a best-fit predicted matrix was defined as one that predicted as many of the vowel pairs that were either confused or not confused by a given listener as possible while minimizing false positives and false negatives. That is, best-fit 2×2 comparison matrices were selected so that the maximum value of B and C was minimized. Of these, the comparison matrix for which the value $2A-B-C$ was maximized was then selected. When more than one value for JND produced the same maximum, the JND that also yielded the lowest rms out of the group was selected. Best-fit 2×2 com-

TABLE II. Demographics of CI users tested for this study: 7 users of the Advanced Bionics device (C) and 18 users of the Nucleus device (N). Age at implantation and experience with implant are stated in years. Speech processing strategies are CIS, ACE (Advanced Combination Encoder), and SPEAK.

Subject	Implanted age	Implant experience	Implanted device	Strategy	No. of channels
C1	66	3.4	Clarion 1.2	CIS	8
C2	32	3.4	Clarion 1.2	CIS	8
C3	61	5.9	Clarion 1.2	CIS	8
C4	23	5.5	Clarion 1.2	CIS	8
C5	53	6.1	Clarion 1.2	CIS	5
C6	39	2.7	Clarion 1.2	CIS	6
C7	43	2.2	Clarion 1.2	CIS	8
N1	31	5.2	Nucleus CI22M	SPEAK	18
N2	59	11.2	Nucleus CI22M	SPEAK	13
N3	71	3	Nucleus CI22M	SPEAK	14
N4	67	2.9	Nucleus CI22M	SPEAK	19
N5	45	3.9	Nucleus CI22M	SPEAK	20
N6	48	9.1	Nucleus CI22M	SPEAK	16
N7	16	4.6	Nucleus CI22M	SPEAK	18
N8	66	2.3	Nucleus CI22M	SPEAK	18
N9	48	1.7	Nucleus CI24M	ACE	20
N10	42	2.3	Nucleus CI24M	SPEAK	16
N11	44	3.1	Nucleus CI24M	SPEAK	20
N12	75	1.7	Nucleus CI24M	SPEAK	19
N13	65	2.2	Nucleus CI24M	SPEAK	20
N14	53	1.9	Nucleus CI24M	SPEAK	20
N15	45	4.2	Nucleus CI24M	SPEAK	20
N16	45	3.2	Nucleus CI24M	SPEAK	20
N17	37	0.9	Nucleus CI24M	SPEAK	20
N18	68	1.2	Nucleus CI24M	SPEAK	20

parison matrices were obtained at three values for threshold: 3%, 5%, and 10%. Different thresholds were necessary to assess errors made by subjects with very different performance levels. A best-fit 2×2 comparison matrix was labeled "satisfactory" if both A and D were greater than (or at least equal to) B and C. According to this definition a satisfactory comparison matrix is one where the model was able to predict at least one-half of the vowel pairs confused by an individual listener, and do so with a number of false positives no greater than the number of true positives (vowel pairs accurately predicted to be confused by the individual).

III. EXPERIMENT 1: VOWEL IDENTIFICATION

A. Methods

1. CI listeners

Twenty-five postlingually deafened adult users of CIs were recruited for this study. Participants were compensated for their time and provided informed consent. All participants were over 18 years of age at the time of testing, and the mean age at implantation was 50 years ranging from 16 to 75 years. Participants were profoundly deaf (PTA $>$ 90 dB) and had at least 1 year of experience with their implant before testing, with the exception of N17 who had 11 months of post-implant experience when tested. The demographics for this group at time of testing are presented in Table II, including age at implantation, duration of post-implant experience, type of CI device and speech processing strategy, as well as number of active channels.

2. Stimuli and general procedures

Vowel stimuli consisted of nine vowels in /hVd/context, i.e., heed, hawed, heard, hood, who'd, hid, hud, had, and head. Stimuli included three tokens of each vowel recorded from the same male speaker. Vowel tokens would be presented over loudspeaker to CI subjects seated 1 m away in a sound attenuated room. The speaker was calibrated before each experimental session so that stimuli would register a value of 70 dB C-weighted SPL on a sound level meter placed at the approximate location of a user's ear-level microphone. In a given session listeners would be presented with one to three lists of the same 45 stimuli (i.e., up to 135 presentations) where each list comprised a different randomization of presentation order. In each list, two tokens of each vowel were presented twice and one token was presented once. Before the testing session, listeners were presented with each vowel token at least once knowing in advance the vowel to be presented for practice. During the testing session, no feedback was provided. All three lists were presented on the same day, and a listener was allowed a break between lists if required.

3. Application of the MPI model

Step 1. All seven possible combinations of one, two, or three dimensions consisting of mean locations of formant energies F1, F2, and F3 along the electrode array were tested.

Step 2. Mean locations of formant energies along the electrode array were obtained from electrodograms of each vowel token that was presented to CI subjects. A set of formant location measurements was obtained for each CI listener. Obtaining these measurements directly from each subject's external device would have been optimal, but time consuming. Instead, four generic sets of formant location measurements were obtained. One set was obtained for the Nucleus-24 spectra body-worn processor with the SPEAK stimulation strategy using FAT No. 9, and three sets were obtained for the Clarion 1.2 processor with the CIS stimulation strategy using the standard FAT imposed by the device's fitting software. The three sets of formant locations for Clarion users were obtained with the speech processor programmed using eight, six, and five channels. One Clarion subject had five active channels in his FAT, another one had six channels, and the remaining five had all eight channels activated. Two out of 18 of the Nucleus subjects and 4 out of 7 of the Clarion subjects used these standard FATs, whereas the other subjects used other FATs with slight modifications. For example, a Nucleus subject may have used FAT No. 7 instead of FAT No. 9, or one or more electrodes may have been turned off, or a Clarion subject may have used extended frequency boundaries for the lowest or the highest frequency channels. For these other subjects, each generic set of formant location measurements that we obtained was then modified to generate a unique set of measurements. Using

linear interpolation, the generic data set was first transformed into hertz using the generic set's frequency allocation table and then transformed back into millimeters from the most basal electrode using the frequency allocation table that was programmed into a given subject's speech processor at the time of testing. This method provided a unique set of formant location measurements even for those subjects with one or more electrodes shut off, typically to avoid facial twitch and/or dizziness.

Step 3. Using a CI listener's set of formant location measurements for a given perceptual dimension combination, MPI model-predicted matrices were generated while JND was varied using one degree of freedom from 0.03 to 6 mm in steps of 0.005 mm (i.e., a total of 1195 predicted matrices). The lower bound of 0.03 mm was selected as it represents a reasonable estimate of the lowest JND for place of stimulation in the cochlea achievable with present day CI devices (Firszt *et al.*, 2007; Kwon and van den Honert, 2006). Each predicted matrix (one for each value of JND) consisted of 5 000 iterations per vowel token, i.e., 225 000 entries in total. Predicted matrices were compared with the listener's observed vowel confusion matrix to obtain the JND that provided the best-fit between predicted matrices and the CI listener's observed vowel matrix. A best-fit JND value and predicted matrix was obtained for each CI listener, for each of the seven perceptual dimension combinations, both in terms of the lowest rms difference and in terms of the best 2×2 comparison matrix using thresholds of 3%, 5%, and 10%. The combination of perceptual dimensions that provided the best-fit to the data was then examined, both from the point of view of rms difference and of error patterns.

B. Results

Vowel identification percent correct scores for the CI listeners tested in the present study are listed in the second column of Table III. The scores ranged from near chance to near perfect.

TABLE III. Minimum rms difference between CI users' observed and predicted vowel confusion matrices for seven perceptual dimension combinations comprising F1, F2, and/or F3. The lowest rms values across perceptual dimensions are highlighted in bold and only values within 1% of this minimum were reported. The second and third columns list observed vowel percent correct and the rms difference between observed matrices and a purely random matrix.

CI User	Vowel (%)	rms							
		Random	F1F2F3	F1F2	F1F3	F2F3	F1	F2	F3
C1	72.6	25.2	**9.9**	10.0	⋯	10.1	⋯	⋯	⋯
C2	98.5	31.0	**5.2**	5.4	⋯	16.0	⋯	⋯	⋯
C3	94.1	29.7	**6.3**	6.7	⋯	⋯	⋯	⋯	⋯
C4	80.0	26.3	**9.1**	9.5	⋯	⋯	⋯	⋯	⋯
C5	21.5	11.0	14.9	15.0	**14.5**	⋯	⋯	⋯	15.5
C6	43.7	16.5	**10.8**	11.1	11.4	⋯	⋯	⋯	⋯
C7	83.7	27.0	**6.0**	6.1	⋯	⋯	⋯	⋯	⋯
N1	80.0	28.2	**14.9**	15.3	⋯	15.7	⋯	⋯	⋯
N2	22.2	11.5	⋯	**13.8**	⋯	⋯	⋯	14.1	14.7
N3	73.3	24.6	**8.0**	⋯	⋯	8.1	⋯	⋯	⋯
N4	70.4	26.7	13.3	⋯	⋯	13.3	⋯	**12.7**	⋯
N5	95.6	30.0	5.4	**4.4**	⋯	⋯	⋯	⋯	⋯
N6	81.7	27.2	**11.4**	12.0	⋯	12.4	⋯	⋯	⋯
N7	72.6	23.5	⋯	**10.4**	⋯	⋯	⋯	⋯	⋯
N8	26.1	11.6	11.9	**11.6**	⋯	12.2	⋯	12.4	⋯
N9	80.0	26.7	**9.0**	⋯	⋯	⋯	⋯	⋯	⋯
N10	81.5	26.3	10.7	**10.1**	⋯	⋯	⋯	⋯	⋯
N11	85.0	27.9	**10.2**	⋯	⋯	⋯	⋯	⋯	⋯
N12	42.2	16.4	**11.9**	12.7	⋯	12.1	⋯	12.5	⋯
N13	79.3	25.4	**8.4**	9.2	⋯	⋯	⋯	⋯	⋯
N14	81.5	26.9	**10.0**	⋯	⋯	⋯	⋯	⋯	⋯
N15	91.1	29.5	9.7	**9.2**	⋯	⋯	⋯	⋯	⋯
N16	59.3	24.7	15.3	⋯	⋯	15.8	⋯	**14.8**	⋯
N17	71.1	24.3	10.2	⋯	⋯	⋯	⋯	**9.8**	⋯
N18	66.7	24.2	**12.1**	⋯	⋯	13.0	⋯	⋯	⋯
Mean	70.1	24.1	**10.5**	11.1	14.9	12.7	17.7	13.7	19.7
		No. of best rms	15	6	1	0	0	3	0

1. rms differences between observed and predicted matrices

Also listed in Table III are the minimum rms differences between predicted and observed matrices as a function of seven possible perceptual dimension combinations. The perceptual dimension combination that produced the lowest minimum rms is highlighted in bold, and rms values greater than 1% above the lowest minimum rms have been omitted. As one can observe, the perceptual dimension combination that produced the lowest minimum rms was F1F2F3 for 15 out of 25 listeners. For eight of the remaining ten listeners, the F1F2F3 perceptual dimension combination provided a fit that was not the best, but was within 1% of the best-fit. Of these remaining ten listeners, six were best fitted by the F1F2 combination, three by the F2 combination, and one by the F1F3 combination.

The third column of Table III contains the rms difference between each listener's observed vowel confusion matrix and a purely random matrix, i.e., one where all matrix elements are equal. Any good model should yield a rms difference that is much smaller than the values that appear in this column. Indeed, this is true for 20 out of 25 CI users for which the lowest minimum rms values achieved with the MPI model (highlighted in bold) are at least 10% lower than those for a purely random matrix (i.e., third column of Table III). The remaining five CI users (C5, C6, N2, N8, and N12) had the lowest vowel identification scores in the group (between 21% and 44% correct). For these subjects, the MPI model does not do much better than a purely random matrix, especially for the three subjects whose scores were only about twice chance levels.

A repeated measures analysis of variance (ANOVA) on ranks was conducted on the rms values we obtained for all subjects. Perceptual dimension combinations, as well as the random matrix comparison, were considered as different treatment groups applied to the same CI subjects. A significant difference was found across treatment groups ($p<0.001$). Using the Student–Newman–Keuls method for multiple post-hoc comparisons, the following significant group differences were found at $p<0.01$: F1F2F3 rms

TABLE IV. Best 2×2 comparison matrices between observed vowel confusion matrices from CI users and those predicted from MPI model. Key for best comparison matrices is on bottom: dim=perceptual dimension combination, thr=threshold at which best comparison matrix was obtained, and *p*-value =result of Fisher exact test; A, B, C, and D, as in Table I. Bottom right, average best 2×2 comparison matrix.

C1	F1F2F3
5%	<0.001
7	0
1	28

C2	F1F2F3
5%	1.00
0	0
2	34

C3	F1F2F3
10%	0.024
3	2
3	28

C4	F1F2F3
5%	0.003
4	2
2	28

C5	F1F2F3
5%	0.026
23	4
4	5

C6	F1F2F3
5%	0.002
12	3
5	16

C7	F1F2F3
5%	0.013
3	2
2	29

N1	F2
10%	0.027
2	1
2	31

N2	F1F2F3
10%	0.015
11	7
3	15

N3	F1F2F3
5%	0.003
4	2
2	28

N4	F1F2F3
5%	<0.001
4	2
0	30

N5	F1F2
3%	0.005
4	1
4	27

N6	F2F3
10%	0.027
2	1
2	31

N7	F1F2F3
10%	0.013
3	2
2	29

N8	F1F2F3
5%	0.041
16	5
6	9

N9	F1F2F3
5%	0.010
3	1
3	29

N10	F1F2F3
3%	0.024
5	5
3	23

N11	F1F2F3
10%	0.010
2	0
2	32

N12	F1F2F3
5%	<0.001
14	4
2	16

N13	F1F2F3
5%	0.030
4	4
3	25

N14	F1F2F3
10%	0.027
2	1
2	31

N15	F1F2F3
10%	0.010
2	0
2	32

N16	F1F2F3
5%	0.026
5	4
4	23

N17	F1F2F3
3%	0.002
11	4
4	17

N18	F1F2F3
5%	0.003
9	4
4	19

Key	
Subject	Dim
thr	*p*-value
A	B
C	D

Average	
6.20	2.44
2.76	24.60

<F1F2 rms<F2F3 rms<F2 rms<F1F3 rms<F1, F3 and random rms. No significant differences were found between F1, F3, and the random case.

2. Prediction of error patterns

Table IV shows the extent to which the MPI model can fit the patterns of vowel confusions made by individual CI users. The table lists one example of a best 2×2 comparison matrix for each subject. At the bottom of Table IV is a key that identifies where to find for each comparison matrix the subject identifier, the perceptual dimension from which the best comparison matrix was selected, the threshold (3%, 5%, or 10%), the *p*-value obtained from a Fisher exact test, and elements A–D of the comparison matrix as outlined in Table I of Sec. II. The following criteria were used for selecting the matrices listed in Table IV: (1) a satisfactory 2×2 comparison matrix with F1F2F3 at the 5% threshold, (2) a satisfactory matrix with F1F2F3 at any threshold, and (3) a satisfactory matrix at any perceptual dimension. Under these criteria, satisfactory matrices were obtained for 24 out of 25 subjects. The only exception was subject C2 who confused very few vowel pairs and for whom a satisfactory comparison matrix could not be obtained. On the lower right of Table IV is an average of elements A–D for all 25 exemplars listed in Table IV. On average, the MPI model predicted the pattern of vowel confusions in 31 out of 36 possible vowel pair confusions. As for the Fisher exact tests, the comparison matrices in Table IV were significant at $p<0.05$ for 24 out of 25 subjects (again subject C2 was the exception), half of which were significant at $p\leq 0.01$.

Table V shows the number of satisfactory best-fit 2×2 comparison matrices obtained for each listener at each perceptual dimension combination. As comparison matrices were obtained at thresholds of 3%, 5%, and 10%, the maximum number of satisfactory comparison matrices at each perceptual dimension combination is 3. The bottom row of Table V lists the total number of satisfactory comparison matrices at each perceptual dimension combination. As one can observe, the F1F2F3 combination produced the largest number of satisfactory best-fit 2×2 comparison matrices, corroborating the result obtained with the best-fit rms criteria.

C. Discussion

It is not surprising that a model based on the ability to discriminate formant center frequencies can explain at least some aspects of vowel *identification*. Rather, what is novel about the results of the present study is that the MPI model produced *confusion matrices* that closely matched CI users' vowel confusion matrices, including the general pattern of errors between vowels, despite differences in age at implantation, implant experience, device and simulation strategy

TABLE V. Number of "satisfactory" 2×2 comparison matrices at thresholds of 3%, 5%, and 10% for each perceptual dimension.

Subject	F1F2F3	F1F2	F1F3	F2F3	F1	F2	F3
C1	3	3	0	3	0	3	0
C2	0	0	0	0	0	0	0
C3	1	0	0	0	0	0	0
C4	2	2	0	2	0	2	0
C5	1	1	1	0	1	0	0
C6	3	3	3	3	3	3	0
C7	3	2	0	1	0	1	0
N1	0	0	0	0	0	1	0
N2	2	3	1	3	1	3	2
N3	2	2	0	2	0	2	0
N4	3	3	0	3	0	3	0
N5	0	1	0	0	0	0	0
N6	0	0	0	1	0	0	0
N7	2	3	0	2	0	3	0
N8	2	1	1	2	1	2	0
N9	3	0	0	2	0	1	0
N10	1	2	0	1	0	2	0
N11	1	0	0	1	0	0	0
N12	3	3	3	3	3	3	3
N13	2	2	0	2	0	2	0
N14	2	0	0	1	0	0	0
N15	1	0	0	1	0	0	0
N16	3	3	0	2	0	3	0
N17	1	3	1	2	0	3	1
N18	3	2	2	3	0	3	0
Total	44	39	12	40	9	40	6

used (Table II), as well as overall vowel identification level (Table III). It is important to stress that these results were achieved with only one degree of freedom. The ability to demonstrate how a model accounts for experimental data is strengthened when the model can capture the general trend of the data while using fewer instead of more degrees of freedom (Pitt and Navarro, 2005). With one degree of freedom, when a model with F1F2F3 does better than a model with F1F2, or when a model with F1F2 does better than a model with F2 alone, one can interpret the value of an added perceptual dimension without having to account for the possibility that the improvement was due to an added fitting parameter.

Whether in terms of rms differences (Table III) or prediction of error patterns (Table V) it is clear that F1F2F3 was the most successful formant combination in accounting for CI users' vowel identification. Upon inspection of the other formant dimension combinations, both Tables III and V suggest that models that included the F2 dimension tended to do better than models without F2, and Table III suggests that the F1F2 combination was a close second to the F1F2F3 combination. The implication may be that F2, and perhaps F1, are important for identifying vowels in most listeners, whereas F3 may be an important cue for some implanted listeners, particularly for r-colored vowels such as heard, but perhaps not for others (Skinner *et al.*, 1996).

The model was able to explain most of the confusions made by most of the individual listeners, while making few false positive predictions. This is an important result because one degree of freedom is always sufficient to fit one independent variable, such as percent correct, but it is not sufficient to predict a data set that includes 36 pairs of vowels. It should come as no surprise that percent correct scores in a predicted vowel matrix drop as the JND parameter is increased. Any model that employs a parameter to move data away from the main diagonal would accomplish the same result. However, the MPI model succeeds in the sense that increasing the JND moves data away from the main diagonal toward a specific vowel confusion pattern determined by the set of perceptual dimensions proposed. Although the fit between predicted and observed data was not perfect, it was strong enough to suggest that the proposed model captures some of the mechanisms CI users employ to identify vowels.

IV. EXPERIMENT 2: F1 IDENTIFICATION

A. Methods

One of the premises underlying the MPI model of vowel identification by CI users in the present study is that a relationship exists between these listeners' ability to identify vowels and their ability to identify steady-state formant frequencies. To test this premise, 18 of the 25 CI users tested for our vowel identification task were also tested for first-formant (F1) identification.

1. Stimuli and general procedures

The testing conditions for this experiment were the same as for the vowel identification experiment in Sec. III A 2, differing only in the type and number of stimuli to identify. For F1 identification, stimuli were seven synthetic three-formant steady-state vowels created with the Klatt 88 speech synthesizer (Klatt and Klatt, 1990). The synthetic vowels differed from each other only in steady-state first-formant center frequencies, which ranged between 250 and 850 Hz in increments of 100 Hz. The fundamental, second, and third formant frequencies were fixed at 100, 1500, and 2500 Hz, respectively. Steady-state F1 values were verified with an acoustic waveform editor. The spectral envelope was obtained from the middle portion of each stimulus, and the frequency value of the F1 spectral peak was confirmed. Each stimulus was 1 s in duration and the onset and offset of the vowel envelope occurred over a 10 ms interval, this transition being linear in dB. The stimuli were digitally stored using a sampling rate of 11 025 Hz at 16 bits of resolution. Listeners were tested using a seven-alternative, one interval forced choice absolute identification task. During each block of testing stimuli were presented ten times in random order (i.e., 70 presentations per block). Prior to testing, participants would familiarize themselves with each stimulus (numbered 1–7) using an interactive software interface. During testing, participants would cue the interface to play a stimulus and then select the most appropriate stimulus number. After each selection, feedback about the correct response was displayed on the computer monitor before moving on to the next stimulus. Subjects completed seven to ten testing blocks (with the exception of listeners N6 and N7 who completed six and five testing blocks, respectively). This number of testing blocks was chosen as it was typically sufficient for most listeners to

provide at least two runs representative of asymptotic, or best, performance.

2. Cumulative-d′ (Δ′) analysis

For each block of testing a sensitivity index d' (Durlach and Braida, 1969) was calculated for each pair of adjacent stimuli (1 vs 2, 2 vs 3, …, 6 vs 7) and then summed to obtain the total sensitivity, i.e., Δ′, which is the *cumulative-d′* across the range of first-formant frequencies between 250 and 850 Hz (i.e., from stimuli 1 to 7). For a given pair of adjacent stimuli, d' was calculated by subtracting the mean responses for the two stimuli and dividing by the average standard deviation of the responses to the two stimuli. For each CI user, the two highest Δ′ among all testing blocks were averaged to arrive at the final score for this task. The average of the highest two Δ′ scores represents an estimate of asymptotic performance, i.e., failure to improve Δ′. Asymptotic performance was sought as it provides a measure of sensory discrimination performance after factoring in learning effects and factoring out fatigue. As is customary for Δ′ calculations, any d' score greater than 3 was set to $d'=3$ (Tong and Clark, 1985). We defined the JND as occurring at $d'=1$, so that Δ′ equals the number of JNDs across the range of first-formant frequencies between 250 and 850 Hz. We then divided this range (i.e., 600 Hz) by Δ′ to obtain the average JND in Hz.

To test the premise that a relationship exists between CI listeners' ability to identify vowels and their ability to discriminate steady-state formant frequencies, two correlation analyses were made using the average JNDs (in hertz) measured in the F1 identification task. One comparison was between JNDs (in hertz) and vowel identification percent correct scores. The other comparison was between JNDs (in hertz) and the F1F2F3 MPI model input JNDs (in millimeters) that yielded best-fit predicted matrices in terms of lowest rms difference.

B. Results

Listed in Table VI are CI subjects' observed percent correct scores for vowel identification and observed average JNDs (in hertz) for first-formant identification (F1 ID). Also listed in Table VI are CI subjects' predicted vowel identification percent correct and input JNDs (in millimeters) that provided best-fit model matrices using the F1F2F3 MPI model. Comparing the observed scores, a scatter plot of vowel scores and JNDs for the 18 CI users tested on both tasks (Fig. 3, top panel) yields a correlation of $r=-0.654$ ($p=0.003$). This result suggests that in our group of CI users, the ability to correctly identify vowels was significantly correlated with the ability to identify first-formant frequency. Furthermore, for the same 18 CI users, a scatter plot of the MPI model input JNDs in millimeters against the observed JNDs in hertz from F1 identification (Fig. 3, bottom panel) yields a correlation of $r=0.635$, $p=0.005$ (without the data point with the highest predicted JND in millimeters, $r=0.576$ and $p=0.016$). Hence, a significant correlation exists between the JNDs obtained from first-formant identification and the JNDs obtained indirectly by optimizing model matrices to fit the vowel identification matrices obtained from the same listeners. That is, fitting the MPI model to one data set (vowel identification) produced JNDs that are consistent with JNDs obtained with the same listeners from a completely independent data set (F1 identification).

TABLE VI. Observed percent correct scores for vowel identification and average JNDs (in hertz) for first-formant identification, and F1F2F3 MPI model-predicted vowel percent correct scores and input JNDs that minimized rms difference between predicted and observed vowel confusion matrices for CI users tested in this study (NA=not available).

	Observed		Predicted (F1F2F3)	
Subject	Vowel (%)	JND (Hz)	Vowel (%)	JND (mm)
C1	72.6	279	72.6	0.095
C2	98.5	144	91.6	0.040
C3	94.1	138	89.5	0.040
C4	80.0	NA	77.8	0.080
C5	21.5	359	24.1	0.685
C6	43.7	111	45.9	0.125
C7	83.7	88	84.9	0.060
N1	80.0	NA	70.9	0.280
N2	22.2	NA	28.8	1.575
N3	73.3	141	71.6	0.230
N4	70.4	247	70.6	0.280
N5	95.6	NA	91.8	0.070
N6	81.7	131	75.5	0.225
N7	72.6	123	80.7	0.150
N8	26.1	324	29.0	1.725
N9	80.0	NA	76.9	0.270
N10	81.5	NA	72.6	0.175
N11	85.0	159	80.8	0.220
N12	42.2	224	45.8	0.820
N13	79.3	116	80.4	0.225
N14	81.5	138	79.4	0.235
N15	91.1	NA	87.3	0.140
N16	59.3	185	52.8	0.645
N17	71.1	141	72.7	0.315
N18	66.7	311	64.1	0.430

C. Discussion

The significant correlations in Fig. 3 lend support to the hypothesis that CI users' ability to discriminate the locations of steady-state mean formant energies along the electrode array contributes to vowel identification, and also provides a degree of validation for the manner in which the MPI model of the present study connects these two variables. Nevertheless, the correlations were not very large, accounting for approximately 40% of the variability observed in the scatter plots. One important difference between identification of vowels and identification of formant center frequencies is that the former involves the assignment of lexically meaningful labels stored in long-term memory whereas the latter does not. Hence, if a CI user has very good formant center frequency discrimination, their ability to identify vowels could still be poor if their vowel labels are not sufficiently resolved in long-term memory. That is, good formant center frequency discrimination is necessary but not sufficient for good vowel identification.

As a side note, the observed JNDs in Table VI were larger than those reported by Fitzgerald *et al.* (2007).

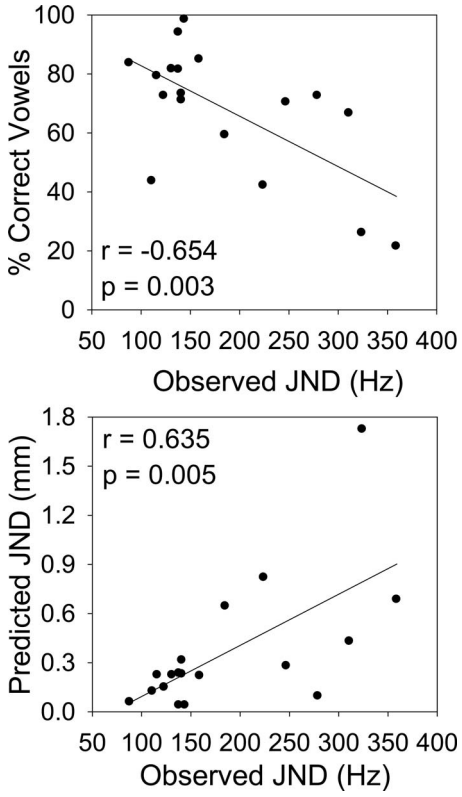


FIG. 3. Top panel: scatter plot of vowel identification percent correct scores against observed JND (in hertz) from first-formant identification obtained from 18 CI users (r=−0.654, p=0.003). Bottom panel: scatter plot of F1F2F3 MPI model's input JNDs (in millimeters) that produced best-fit to subjects' observed vowel matrices (minimized rms) against these subjects' observed JND (in hertz) from first-formant identification (r=0.635 and p =0.005).

However, this is to be expected as their F1 discrimination task measured the JND above an F1 center frequency of 250 Hz, whereas our measure represented the average JND for F1 center frequencies between 250 and 850 Hz.

V. EXPERIMENT 3: FREQUENCY ALLOCATION TABLES

A. Methods

Skinner *et al.* (1995) examined the effect of FAT Nos. 7 and 9 on speech perception with seven postlingually deafened adult users of the Nucleus-22 device and SPEAK stimulation strategy. Although FAT No. 9 was the default clinical map, Skinner *et al.* (1995) found that their listeners' speech perception improved with FAT No. 7. The speech battery they used included a vowel identification task with 19 medial vowels in /hVd/context, 3 tokens each, comprising 9 pure vowels, 5 *r*-colored vowels, and 5 diphthongs. The vowel confusion matrices they obtained (and recordings of the stimuli they used) were provided to us for the present study.

1. Application of MPI model

The MPI model was applied to the vowel identification data of Skinner *et al.* (1995) in order to test the model's ability to explain the improvement in performance that occurred when listener's used FAT No. 7 instead of FAT No. 9. As a demonstration of how the MPI model can be used to explore the vast number of possible settings for a given CI fitting parameter in a very short amount of time, the MPI model was also used to provide a projection of vowel percent correct scores as a function of ten different frequency allocation tables and JND.

Step 1. One perceptual dimension combination was used to model the data of Skinner *et al.* (1995) and to generate predictions at other FATs. Namely, mean locations of formant energies along the electrode array for the first three formants combined, i.e., F1F2F3, in units of millimeters from the most basal electrode.

Step 2. Because our MPI model predicts identification of and confusions among vowels based on CI users' discrimination of mean formant energy locations, only ten of the vowels used by Skinner *et al.* (1995) were used in our model; i.e., the nine purely monophthongal vowels and the *r*-colored vowel heard. Using the original vowel recordings used by Skinner *et al.* (1995) and sCILab software (Bögli *et al.*, 1995; Wai *et al.*, 2003), two sets of formant location measurements were obtained from a Nucleus-22 spectra body-worn processor programmed with the SPEAK stimulation strategy. One set of measurements was obtained while the processor was programmed with FAT No. 7, and the other while the processor was programmed with FAT No. 9. Both sets of measurements were used for fitting Skinner *et al.*'s (1995) data, and for the MPI model's projection of vowel percent correct as a function of JND. For the model's projection at other FATs, formant location measurements were obtained using linear interpolation from FAT No. 9. The other frequency allocation tables explored in this projection were FAT Nos. 1, 2, and 6–13.

Step 3. For Skinner *et al.*'s (1995) data, the MPI model was run while allowing JND to vary as a free parameter until model matrices were obtained that best-fit the observed group vowel confusion matrices at FAT Nos. 7 and 9. The JND parameter was varied from 0.1 to 1 mm of electrode distance in increments of 0.01 mm using one degree of freedom; i.e., JND was the same for each perceptual dimension. Only one value of JND was used to find a best-fit to both sets of observed matrices in terms of minimum rms combined for both matrices. For the MPI model's projection of vowel identification as a function of the various FATs, model matrices were obtained for JND values of 0.1, 0.2, 0.4, 0.8, and 1.0 mm of electrode distance, where JND was assumed to be the same for each perceptual dimension. Percent correct scores were then calculated from the resulting model matrices. In all of the above simulations, the MPI model was run using 5000 iterations per vowel token.

B. Results

1. Application of MPI model to Skinner et al. (1995)

For the ten vowels we included in our modeling, the average vowel identification percent correct scores for the group of listeners tested by Skinner *et al.* (1995) were 84.9% with FAT No. 7 and 77.5% with FAT No. 9. For the MPI model of Skinner *et al.*'s (1995) data, a JND of 0.24 mm produced best-fit model matrices. The rms differences between observed and predicted matrices were 4.3% for FAT

TABLE VII. 2×2 comparison matrices for MPI model matrices produced with JND=0.24 mm and Skinner *et al.*'s (1995) vowel matrices obtained with FAT Nos. 7 and 9. The data follow the key at the bottom of Table IV.

FAT No. 7	F1F2F3	FAT No. 9	F1F2F3
3%	$p<0.001$	3%	$p<0.001$
6	3	6	5
2	34	1	33

No. 7 and 6.2% for FAT No. 9. The predicted matrices had percent correct scores equal to 85.1% with FAT No. 7 and 79.4% with FAT No. 9. Thus, the model predicted that FAT No. 7 should result in better vowel identification (which was true for all JND values between 0.1 and 1 mm) and it also predicted the size of the improvement. The 2×2 comparison matrices that demonstrate the extent to which model matrices account for the error pattern in Skinner *et al.*'s (1995) matrices are presented in Table VII. The comparison matrices were compiled using a threshold of 3%. With one degree of freedom, the MPI model produced model matrices that account for 40 out of 45 vowel pair confusions in the case of FAT No. 7 and 39 out of 45 vowel pair confusions in the case of FAT No. 9. For both comparison matrices, a Fisher's exact test yields $p<0.001$.

2. MPI model projection at various FATs

The FAT determines the frequency band assigned to a given electrode. The ten FATs used to produce MPI model projections of vowel percent correct scores are summarized in Table VIII, which depicts the FAT number (1, 2, and 6–13), channel number (starting from the most apically stimulating electrode), and the lower frequency boundary (in hertz) assigned to a given channel (the upper frequency boundary for a given channel is equal to the lower frequency boundary of the next highest channel number, and the upper boundary for the highest channel number is provided in the bottom row). The percent correct scores obtained from MPI model matrices at each FAT, and as a function of JND are summarized in Fig. 4. Two observations are worth noting. First, a lower JND for a given frequency map results in a higher predicted percent correct score. That is, a lower JND would provide better discrimination between formant values and hence a smaller chance of confusing formant values belonging to different vowels. Second, for a fixed JND, percent correct scores begin to gradually decrease as the FAT number is increased to higher FAT numbers beyond FAT No. 7, with the exception of JND=0.1 mm where a ceiling effect is observed. As FAT number increases from No. 1 to No. 9, a larger frequency range is assigned to the same set of

TABLE VIII. Frequency allocation table numbers (FAT No.) 1, 2, and 6–13 for the Nucleus-22 device. Channel numbers begin with the most apically stimulated electrode and indicate the lower frequency boundary (in hertz) assigned to a given electrode. Bottom row indicates upper frequency boundary for highest frequency channel. Approximate range of formant frequency regions indicated by text in bold: F1 (300–1000 Hz), F2 (1000–2000 Hz), and F3 (2000–3000 Hz).

	FAT No.									
Channel	1	2	6	7	8	9	10	11	12	13
1	75	80	109	120	**133**	**150**	**171**	**200**	**240**	**150**
2	175	186	**254**	**280**	**311**	**350**	**400**	**466**	**560**	**300**
3	**275**	**293**	**400**	**440**	**488**	**550**	**628**	**733**	**880**	**700**
4	**375**	**400**	**545**	**600**	**666**	**750**	**857**	**1 000**	**1 200**	**1100**
5	**475**	**506**	**690**	**760**	**844**	**950**	**1 085**	**1 266**	**1 520**	**1500**
6	**575**	**613**	**836**	**920**	**1022**	**1 150**	**1 314**	**1 533**	**1 840**	**1900**
7	**675**	**720**	**981**	**1080**	**1200**	**1 350**	**1 542**	**1 800**	**2 160**	**2300**
8	**775**	**826**	**1127**	**1240**	**1377**	**1 550**	**1 771**	**2 066**	**2 480**	**2700**
9	**884**	**942**	**1285**	**1414**	**1571**	**1 768**	**2 020**	**2 357**	**2 828**	3100
10	**1015**	**1083**	**1477**	**1624**	**1805**	**2 031**	**2 321**	**2 708**	3 249	3536
11	**1166**	**1244**	**1696**	**1866**	**2073**	**2 333**	**2 666**	3 110	3 732	4062
12	**1340**	**1429**	**1949**	**2144**	**2382**	**2 680**	3 062	3 573	4 288	4666
13	**1539**	**1642**	**2239**	**2463**	**2736**	3 079	3 518	4 105	4 926	5360
14	**1785**	**1904**	**2597**	**2856**	3174	3 571	4 081	4 761	5 713	6158
15	**2092**	**2231**	3042	3347	3719	4 184	4 781	5 578	6 694	7142
16	**2451**	**2614**	3565	3922	4358	4 903	5 603	6 537	7 844	8368
17	**2872**	3063	4177	4595	5105	5 744	6 564	7 658	9 190	⋯
18	3365	3589	4894	5384	5982	6 730	7 691	8 973	⋯	⋯
19	3942	4205	5734	6308	7008	7 885	9 011	⋯	⋯	⋯
20	4619	4926	6718	7390	8211	9 238	⋯	⋯	⋯	⋯
Upper	5411	5772	7871	8658	9620	10 823	10 557	10 513	10 768	9806

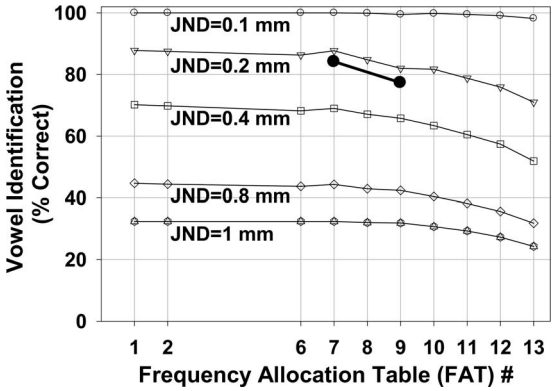


FIG. 4. F1F2F3 MPI model prediction of vowel identification percent correct scores as a function of FAT No. and JND (in millimeters). Filled circles: Skinner *et al.*'s (1995) mean group data when CI subjects' used FAT Nos. 7 and 9.

electrodes. For FAT Nos. 10–13, the relatively large frequency span is maintained while the number of electrodes assigned is gradually reduced. Hence, the MPI model predicts that vowel identification will be deleteriously affected by assigning too large of a frequency span to the CI electrodes. In Fig. 4, the two filled circles joined by a solid line represent the vowel identification percent correct scores obtained by Skinner *et al.* (1995) for the ten vowel tokens we included in our modeling.

C. Discussion

The very first thing to point out is the economy with which the MPI model can be used to project estimates of CI users' performance. The simulation routine implementing the MPI model produced all of the outputs in Fig. 4 in a matter of minutes. Contrast this with the time and resources required to obtain data such as that of Skinner *et al.* (1995), which amounts to two data points in Fig. 4. It would be financially and practically impossible to obtain these data experimentally for all the frequency maps available with a given cochlear implant, let alone for the theoretically infinite number of possible frequency maps.

Without altering any model assumptions, the model predicts the increase in percent correct vowel identification attributable to changing the frequency map from FAT No. 9 to FAT No. 7 with the Nucleus-22 device. In retrospect, Skinner *et al.* (1995) hypothesized that FAT No. 7 might result in improved speech perception because it encodes a more restricted frequency range onto the electrodes of the implanted array. Encoding a larger frequency range onto the array involves a tradeoff: The locations of mean formant energies for different vowels are squeezed closer together. With less space between mean formant energies, the vowels become more difficult to discriminate, at least in terms of this particular set of perceptual dimensions, resulting in a lower percent correct score.

How does this concept apply to the MPI model projections at different FATs displayed in Fig. 4? The effect of different FAT frequency ranges on mean formant locations along the electrode array is depicted in Table VIII where approximate formant regions are indicated in bold. The frequency boundaries defined for each formant are 300–1000 Hz for F1, 1000–2000 Hz for F2, and 2000–3000 Hz for F3. Under this definition of formant regions, five or more electrodes are available for each of F1 and F2 for all maps up to FAT No. 8, and progressively decrease for higher map numbers. In Fig. 4, percent correct changes very little between FAT Nos. 1 and 8, suggesting that F1 and F2 are sufficiently resolved, and then drops progressively for higher map numbers. Indeed, FAT No. 9 has one less electrode available for F2 in comparison to FAT No. 7, which may explain the small but significant drop in percent correct scores with FAT No. 9 observed by Skinner *et al.* (1995).

Apparently, the changes in the span of electrodes for mean formant energies in FAT Nos. 7 and 9 are of a magnitude that will not contribute to large differences in vowel percent correct score for JND values that are very small (less than 0.2 mm) or very high (more than 0.8 mm), but are relevant for JND values that are in between these two extremes.

Although the prediction of the MPI model in Fig. 4 suggests that there is not much to be gained (or lost, for that matter) by shifting the frequency map from FAT No. 7 to FAT No. 1, there is strong evidence to suggest that such a change could be detrimental. Fu *et al.* (2002) found a significant drop in vowel identification scores in three postlingually deafened subjects tested with FAT No. 1 in comparison to their clinically assigned maps (FAT Nos. 7 and 9), even after these subjects used FAT No. 1 continuously for three months. Out of all the maps in Table VIII, FAT No. 1 encodes the lowest frequency range to the electrode array, and potentially has the largest frequency mismatch to the characteristic frequency of the neurons stimulated by the implanted electrodes; particularly for postlingually deafened adults who retained the tonotopic organization of the cochlea before they lost their hearing. The results of Fu *et al.* (2002) suggest that the use of FAT No. 1 in postlingually deafened adults results in an excessive amount of frequency shift, i.e., an amount of frequency mismatch that precludes complete adaptation. In Fig. 4, response bias was assumed to be zero (see Sec. IIA2) so that no mismatch occurred between percepts elicited by stimuli and the expected locations of those percepts. The contribution of a nonzero response bias to lowering vowel percent correct scores for the type of frequency mismatch imposed by FAT No. 1 is addressed in Sagi *et al.*, (2010) wherein the MPI model was applied to the vowel data of Fu *et al.* (2002).

VI. EXPERIMENT 4: ELECTRICAL DYNAMIC RANGE REDUCTION

A. Methods

The electrical dynamic range is the range between the minimum stimulation level for a given channel, typically set at threshold, and the maximum stimulation level, typically set at the maximum comfortable loudness. Zeng and Galvin (1999) systematically decreased the electrical dynamic range of four adult users of the Nucleus-22 device with SPEAK stimulation strategy from 100% to 25% and then to 1% of the original dynamic range. In the 25% condition, dynamic range was set from 75% to 100% of the original dynamic range. In the 1% condition, dynamic range was set from 75% to 76% of the original dynamic range. CI users were then

tested on several speech perception tasks including vowel identification in quiet. One result of Zeng and Galvin (1999) was that even though the electrical dynamic range was reduced to almost zero, the average percent correct score for identification of vowels in quiet dropped by only 9%. We sought to determine if the MPI model could explain this result by assessing the effect of dynamic range reduction on formant location measurements. If reducing the dynamic range has a small effect on formant location measurements, then the MPI model would predict a small change in vowel percent correct scores.

1. Application of MPI model

Step 1. One perceptual dimension combination was used to model the data of Zeng and Galvin (1999). Namely, mean locations of formant energies along the electrode array for the first three formants, i.e., F1F2F3, in units of millimeters from the most basal electrode.

Step 2. Three sets of formant location measurements were obtained, one for each dynamic range condition. For the 100% dynamic range condition, sCILab recordings were obtained for the vowel tokens used in experiment 1 of the present study, using a Nucleus-22 spectra body-worn processor programmed with the SPEAK stimulation strategy and FAT No. 9. The minimum and maximum stimulation levels in the output of the speech processor were set to 100 and 200 clinical units, respectively, for each electrode. For the other two dynamic range conditions, the stimulation levels in these sCILab recordings were adjusted in proportion to the desired dynamic range. That is, the charge amplitude of stimulation pulses, which spanned from 100 to 200 clinical units in the original recordings, was proportionally mapped to 175–200 clinical units for the 25% dynamic range condition, and to 175–176 clinical units for the 1% dynamic range condition. Formant locations were then obtained from electrodograms of the original and modified sCILab recordings.

Step 3. In Zeng and Galvin, 1999, the average vowel identification score in quiet for the 25% dynamic range condition was 69% correct. Using the formant measurements for this condition, the MPI model was run while varying JND, until a JND was found that produced a model matrix with percent correct equal to 69%. This value of JND was then used to run the MPI model with the other two sets of formant measurements for the 100% and 1% dynamic range conditions. In each case, the MPI model was run with 5000 iterations per vowel token, and the percent correct of the resulting model matrices was compared with the scores observed in Zeng and Galvin, 1999.

B. Results

With the MPI model, a JND of 0.27 mm provided a vowel percent correct score of 69% using the formant measurements obtained for the 25% dynamic range condition. With the same value of JND, the formant measurements obtained for the 100% and 1% dynamic range conditions yielded vowel matrices with 71% and 68% correct, i.e., a drop of 3%. The observed scores obtained by Zeng and Galvin (1999) for these two conditions were 76% and 67%, respectively, i.e., a drop of 9%. On one hand, the MPI model employed here explains how a large reduction in electrical dynamic range results in a small drop in the identification of vowels under quiet listening conditions. On the other hand, the MPI model underestimated the magnitude of the drop observed by Zeng and Galvin (1999).

C. Discussion

It should not come as a surprise that the F1F2F3 MPI model employed here predicts that a large reduction in the output dynamic range would have a negligible effect on vowel identification scores in quiet. After all, reducing the output dynamic range (even 100-fold) causes a negligible shift in the location of mean formant energy along the electrode array. More importantly, why did this model underestimate the observed results of Zeng and Galvin (1999)? One explanation may be that the model does not account for the relative amplitudes of formant energies, which can affect percepts arising from F1 and F2 center frequencies in close proximity (Chistovich and Lublinskaya, 1979). Reducing the output dynamic range can affect the relative amplitudes of formant energies without changing their locations along the electrode array. This effect may explain why Zeng and Galvin (1999) found a larger drop in vowel identification scores than those predicted by the MPI model. Hence, the MPI model employed in the present study may be sufficient to explain the vowel identification data of experiments 1 and 3, but may need to be modified to more accurately predict the data of Zeng and Galvin (1999).

Of course, the prediction that reducing the dynamic range will not largely affect vowel identification scores in quiet only applies to users of stimulation strategies such as SPEAK, ACE, and n-of-m. This effect would be completely different for a stimulation strategy like CIS, where all electrodes are activated in cycles, and the magnitude of each stimulation pulse is determined in proportion to the electric dynamic range. For example, in a CI user with CIS, the 1% dynamic range condition used by Zeng and Galvin (1999) would result in continuous activation of all electrodes at the same level regardless of input, thus obliterating all spectral information about vowel identity.

VII. CONCLUSIONS

A very simple model predicts most of the patterns of vowel confusions made by users of different cochlear implant devices (Nucleus and Clarion) who use different stimulation strategies (CIS or SPEAK), who show widely different levels of speech perception (from near chance to near perfect), and who vary widely in age of implantation and implant experience (Tables II and III). The model's accuracy in predicting confusion patterns for an individual listener is surprisingly robust to these variations despite the use of a single degree of freedom. Furthermore, the model can predict some important results from the literature, such as Skinner *et al.*'s (1995) frequency mapping study, and the general trend (but not the size of the effect) in the vowel results of Zeng and Galvin's (1999) studies of output electrical dynamic range reduction.

The implementation of the model presented here is specific to vowel identification by CI users, dependent on

discrimination of mean formant energy along the electrode array. However, the framework of the model is general. Alternative models of vowel identification within the MPI framework could use dynamic measures of formant frequency (i.e., formant trajectories and co-articulation), or other perceptual dimensions such as formant amplitude or vowel duration. One alternative to the MPI framework might involve the comparison of phonemes based on time-averaged electrode activation across the implanted array, treated as a single object rather than breaking it down into specific "cues" or perceptual dimensions (cf. Green and Birdsall, 1958; Müsch and Buus, 2001). Regardless of the specific form they might take, computational models like the one presented here can be useful for advancing our understanding about speech perception in hearing impaired populations, and for providing a guide for clinical research and clinical practice.

ACKNOWLEDGMENTS

Norbert Dillier from ETH (Zurich) provided us with his sCILab computer program, which we used to record stimulation patterns generated by the Nucleus speech processors. Advanced Bionics Corporation provided an implant-in-a-box so we could monitor stimulation patterns generated by their implant. Margo Skinner (may she rest in peace) provided the original vowel tokens used in her study as well as the confusion matrices from that study. This study was supported by NIH-NIDCD Grant Nos. R01-DC03937 (P.I.: Mario Svirsky) and T32-DC00012 (PI: David B. Pisoni) as well as by grants from the Deafness Research Foundation and the National Organization for Hearing Research.

Bögli, H., Dillier, N., Lai, W. K., Rohner, M., and Zillus, B. A. (**1995**). Swiss Cochlear Implant Laboratory (Version 1.4) ([computer software]), Zürich, Switzerland.

Braida, L. D. (**1991**). "Crossmodal integration in the identification of consonant segments," Q. J. Exp. Psychol. **43A**, 647–677.

Braida, L. D., and Durlach, N. I. (**1972**). "Intensity perception. II. Resolution in one-interval paradigms," J. Acoust. Soc. Am. **51**, 483–502.

Chatterjee, M., and Peng, S. C. (**2008**). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition," Hear. Res. **235**, 143–156.

Chistovich, L. A., and Lublinskaya, V. V. (**1979**). "The 'center of gravity' effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli," Hear. Res. **1**, 185–195.

Durlach, N. I., and Braida, L. D. (**1969**). "Intensity perception. I. Preliminary theory of intensity resolution," J. Acoust. Soc. Am. **46**, 372–383.

Firszt, J. B., Koch, D. B., Downing, M., and Litvak, L. (**2007**). "Current steering creates additional pitch percepts in adult cochlear implant recipients," Otol. Neurotol. **28**, 629–636.

Fitzgerald, M. B., Shapiro, W. H., McDonald, P. D., Neuburger, H. S., Ashburn-Reed, S., Immerman, S., Jethanamest, D., Roland, J. T., and Svirsky, M. A. (**2007**). "The effect of perimodiolar placement on speech perception and frequency discrimination by cochlear implant users," Acta Oto-Laryngol. **127**, 378–383.

Fu, Q. J., Shannon, R. V., and Galvin, J. J., III (**2002**). "Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant," J. Acoust. Soc. Am. **112**, 1664–1674.

Green, D. M., and Birdsall, T. G. (**1958**). "The effect of vocabulary size on articulation score," Technical Memorandum No. 81 and Technical Note No. AFCRC-TR-57-58, University of Michigan, Electronic Defense Group.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (**1995**). "Acoustic characteristics of American English vowels," J. Acoust. Soc. Am. **97**, 3099–3111.

Hood, L. J., Svirsky, M. A., and Cullen, J. K. (**1987**). "Discrimination of complex speech-related signals with a multichannel electronic cochlear implant as measured by adaptive procedures," Ann. Otol. Rhinol. Laryngol. **96**, 38–41.

Iverson, P., Smith, C. A., and Evans, B. G. (**2006**). "Vowel recognition via cochlear implants and noise vocoders: Effects of formant movement and duration," J. Acoust. Soc. Am. **120**, 3998–4006.

Jenkins, J. J., Strange, W., and Edman, T. R. (**1983**). "Identification of vowels in 'vowelless' syllables," Percept. Psychophys. **34**, 441–450.

Kewley-Port, D., and Watson, C. S. (**1994**). "Formant-frequency discrimination for isolated English vowels," J. Acoust. Soc. Am. **95**, 485–496.

Kirk, K. I., Tye-Murray, N., and Hurtig, R. R. (**1992**). "The use of static and dynamic vowel cues by multichannel cochlear implant users," J. Acoust. Soc. Am. **91**, 3487–3497.

Klatt, D. H., and Klatt, L. C. (**1990**). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," J. Acoust. Soc. Am. **87**, 820–857.

Kwon, B. J., and van den Honert, C. (**2006**). "Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users," J. Acoust. Soc. Am. **120**, EL1–EL6.

Müsch, H., and Buus, S. (**2001**). "Using statistical decision theory to predict speech intelligibility. I. Model structure," J. Acoust. Soc. Am. **109**, 2896–2909.

Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**, 175–184.

Phatak, S. A., and Allen, J. B. (**2007**). "Consonant and vowel confusions in speech-weighted noise," J. Acoust. Soc. Am. **121**, 2312–2326.

Pitt, M. A., and Navarro, D. J. (**2005**). in *Twenty-First Century Psycholinguistics: Four Cornerstones*, edited by A. Cutler (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 347–362.

Ronan, D., Dix, A. K., Shah, P., and Braida, L. D. (**2004**). "Integration across frequency bands for consonant identification," J. Acoust. Soc. Am. **116**, 1749–1762.

Sagi, E., Fu, Q.-J., Galvin, J. J., III, and Svirsky, M. A. (**2010**). "A model of incomplete adaptation to a severely shifted frequency-to-electrode mapping by cochlear implant users," J. Assoc. Res. Otolaryngol. (in press).

Shannon, R. V. (**1993**). in *Cochlear Implants: Audiological Foundations*, edited by R. S. Tyler (Singular, San Diego, CA), pp. 357–388.

Skinner, M. W., Arndt, P. L., and Staller, S. J. (**2002**). "Nucleus 24 advanced encoder conversion study: Performance versus preference," Ear Hear. **23**, 2S–17S.

Skinner, M. W., Fourakis, M. S., Holden, T. A., Holden, L. K., and Demorest, M. E. (**1996**). "Identification of speech by cochlear implant recipients with the multipeak (MPEAK) and spectral peak (SPEAK) speech coding strategies I. vowels," Ear Hear. **17**, 182–197.

Skinner, M. W., Holden, L. K., and Holden, T. A. (**1995**). "Effect of frequency boundary assignment on speech recognition with the SPEAK speech-coding strategy," Ann. Otol. Rhinol. Laryngol. **104**, (Suppl. 166), 307–311.

Svirsky, M. A. (**2000**). "Mathematical modeling of vowel perception by users of analog multichannel cochlear implants: Temporal and channel-amplitude cues," J. Acoust. Soc. Am. **107**, 1521–1529.

Svirsky, M. A. (**2002**). in *Etudes et Travaux*, edited by W. Serniclaes (Institut de Phonetique et des Langues Vivantes of the ULB, Brussels), Vol. **5**, pp. 143–186.

Syrdal, A. K., and Gopal, H. S. (**1986**). "A perceptual model of vowel recognition based on the auditory representation of American English vowels," J. Acoust. Soc. Am. **79**, 1086–1100.

Teoh, S. W., Neuburger, H. S., and Svirsky, M. A. (**2003**). "Acoustic and electrical pattern analysis of consonant perceptual cues used by cochlear implant users," Audiol. Neuro-Otol. **8**, 269–285.

Thurstone, L. L. (**1927a**). "A law of comparative judgment," Psychol. Rev. **34**, 273–286.

Thurstone, L. L. (**1927b**). "Psychophysical analysis," Am. J. Psychol. **38**, 368–389.

Tong, Y. C., and Clark, G. M. (**1985**). "Absolute identification of electric pulse rates and electrode positions by cochlear implant subjects," J. Acoust. Soc. Am. **77**, 1881–1888.

Wai, K. L., Bögli, H., and Dillier, N. (**2003**). "A software tool for analyzing multichannel cochlear implant signals," Ear Hear. **24**, 380–391.

Zahorian, S. A., and Jagharghi, A. J. (**1993**). "Spectral-shape features versus formants as acoustic correlates for vowels," J. Acoust. Soc. Am. **94**, 1966–1982.

Zeng, F. G., and Galvin, J. J., III (**1999**). "Amplitude mapping and phoneme recognition in cochlear implant listeners," Ear Hear. **20**, 60–74.

Unsupervised speech segmentation: An analysis of the hypothesized phone boundaries

Odette Scharenborg[a)]
Centre for Language and Speech Technology, Radboud University Nijmegen, Erasmusplein 1, 6525 HT Nijmegen, The Netherlands

Vincent Wan
Department of Computer Science, Speech and Hearing Research Group, University of Sheffield, Regent Court, 211 Portobello Street, Sheffield S1 4DP, United Kingdom

Mirjam Ernestus
Center for Language Studies, Radboud University Nijmegen, and Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands

(Received 27 May 2009; revised 19 October 2009; accepted 26 November 2009)

Despite using different algorithms, most unsupervised automatic phone segmentation methods achieve similar performance in terms of percentage correct boundary detection. Nevertheless, unsupervised segmentation algorithms are not able to perfectly reproduce manually obtained reference transcriptions. This paper investigates fundamental problems for unsupervised segmentation algorithms by comparing a phone segmentation obtained using only the acoustic information present in the signal with a reference segmentation created by human transcribers. The analyses of the output of an unsupervised speech segmentation method that uses acoustic change to hypothesize boundaries showed that acoustic change is a fairly good indicator of segment boundaries: over two-thirds of the hypothesized boundaries coincide with segment boundaries. Statistical analyses showed that the errors are related to segment duration, sequences of similar segments, and inherently dynamic phones. In order to improve unsupervised automatic speech segmentation, current one-stage bottom-up segmentation methods should be expanded into two-stage segmentation methods that are able to use a mix of bottom-up information extracted from the speech signal and automatically derived top-down information. In this way, unsupervised methods can be improved while remaining flexible and language-independent.
 [DOI: 10.1121/1.3277194]

PACS number(s): 43.72.Ar, 43.72.Ne [SSN] Pages: 1084–1095

I. INTRODUCTION

Over the past few years, interest in the automatic segmentation of speech has increased. In the fields of automatic speech recognition and text-to-speech, there is a need for large amounts of reliably segmented speech data, for instance, for improving recognition and synthesis performance. Furthermore, automatic speech segmentation methods are used for the automatic phonetic analysis of large amounts of speech data (e.g., Kuperman *et al.*, 2007). In the past, speech data were segmented by hand, but with the need for and availability of ever increasing amounts of speech data the task of manual speech segmentation becomes too time-consuming and expensive. Furthermore, manual labeling and segmentation are subjective, resulting in significant differences in the transcriptions created by different expert listeners (Cucchiarini, 1993). Automatic systems, on the other hand, are consistent.

Automatic speech segmentation is the partitioning of a continuous speech signal into discrete, non-overlapping units. Generally, automatic speech segmentation methods are divided into two types. Supervised methods require *a priori* knowledge (e.g., Brugnara *et al.*, 1993; Kim and Conkie, 2002; Pellom and Hanson, 1998). Most of the supervised methods are based on forced alignment techniques starting from an orthographic transcription of the speech material. This means that the representation of the word or utterance in terms of discrete units is known (from a lexicon which includes the words' pronunciations) and pre-trained acoustic models of these units are needed for the forced alignment. The task of the segmentation algorithm is then to optimally locate the unit boundaries (Sharma and Mammone, 1996). Unsupervised methods, on the other hand, require no training data for segmenting the speech signal. Instead, they use sets of rules derived from or encoding human knowledge to segment speech. Acoustic (rate of) change (e.g., Sharma and Mammone, 1996; see for early work on unsupervised automatic speech segmentation, Bridle and Sedgwick, 1977; for more recent work, see below) is an example of prior human knowledge that is used to solve the speech segmentation task. The task for an unsupervised segmentation algorithm then is two-fold; the number of segments in the speech signal needs to be determined (this is usually determined by a parameter such as the parameter δ described in Sec. III) and the position of the boundaries on the basis of the acoustic signal needs to be determined (Sharma and Mammone, 1996).

[a)]Author to whom correspondence should be addressed. Electronic mail: o.scharenborg@let.ru.nl

There are some good reasons for using unsupervised methods. First of all, supervised methods require (extensive) training on (carefully prepared) speech material. The training material needs to be transcribed in terms of the units the algorithm is supposed to segment the speech signal into, usually phones. Furthermore, usually large amounts of training data are needed to train the supervised algorithms; however, large amounts of training data are not always easily obtained and neither are transcriptions. Unsupervised methods, on the other hand, do not require training; so, obviously no training material is needed. For each new language, speech style, dialect or accent, supervised algorithms may need to be retrained, whereas unsupervised methods are based on human knowledge and understanding of the nature of speech and are therefore language and speech style independent. Furthermore, supervised methods require the units to be defined beforehand, e.g., phones, diphones, syllables, and words, in order to be able to train models for them, whereas unsupervised methods, in principle, do not. Thus unsupervised methods yield a desirable and more flexible framework for the automatic segmentation of speech. Finally, unsupervised segmentation methods are generally simpler algorithms than supervised methods (Dusan and Rabiner, 2006).

This paper focuses on unsupervised speech segmentation. A review of the current approaches for unsupervised speech segmentation shows that although very different approaches are used, the results obtained are remarkably similar (see Sec. II). Nevertheless, unsupervised speech segmentation algorithms are not yet able to perfectly reproduce manually obtained reference transcriptions (see also Sec. II). This paper compares a phone segmentation obtained using only the acoustic information present in the signal with a reference segmentation created by human transcribers. We present an in-depth analysis (Sec. V) of the output of our unsupervised speech segmentation algorithm (see Sec. III) on the task of segmenting speech from the TIMIT database (Garofolo, 1988; Sec. IV). This unsupervised speech segmentation algorithm uses acoustic change as the criterion to segment the speech signal into phones.

Naturally, the choice of the automatic segmentation method, and the assumptions underlying the method, will have an impact on the segmentation results. We chose to analyze the results of an automatic segmentation algorithm that only uses acoustic change as the criterion for hypothesizing a segment boundary, as we believe that other criteria, such as heuristics or sophisticated signal processing, will add additional complexity to the underlying system, which will have an impact on the results. If we then perform analysis on such a highly complex automatic segmentation algorithm, the results will be highly specific to that segmentation algorithm. Additionally, more convoluted results will mean that the effects of different parts of the segmentation algorithm may become difficult to tease apart. Hence we use an automatic segmentation method that only uses acoustic change as a criterion in order to ensure that the analysis is clean. Furthermore, we believe that the assumptions underlying our method have implications for other unsupervised automatic speech segmentation methods. We will address this issue in Sec. VI.

This enterprise will indicate where acoustic change is indeed a correct indicator of a segment boundary, and where it is not, thus revealing weaknesses in the criterion of acoustic change for unsupervised automatic speech segmentation. Furthermore, since most unsupervised speech segmentation algorithms use acoustic change as the means to decide when to hypothesize a boundary we believe that the analysis presented here is of interest for unsupervised speech segmentation, in general, and will reveal weaknesses in automatic speech segmentation technologies. This paper ends with suggestions on how to develop unsupervised speech segmentation algorithms that are able to create segmentations that are closer to those created by human transcribers (Sec. VI).

II. PERFORMANCES OF UNSUPERVISED SPEECH SEGMENTATION APPROACHES: A BRIEF OVERVIEW

It is not straightforward to compare the performances of different unsupervised speech segmentation algorithms described in the literature as algorithms are often tested on different sets of speech material. To make the comparison here as fair as possible, we will only discuss those methods that have been tested on TIMIT as the majority of the reported algorithms have been tested on this speech corpus.

The performance of speech segmentation algorithms is usually assessed by comparing their segmentation to a "ground truth," which usually consists of manually placed boundaries, such as those provided with the TIMIT database. A hypothesized boundary is judged to be correctly placed if it falls within a "tolerance window" from the segment boundary in the ground truth segmentation. Generally, a tolerance window of 20 ms is used (although some researchers report performances for a range of distances). This distance of 20 ms is somewhat arbitrarily chosen; however, it is backed-up by evidence from manual segmentation by Wesenick and Kipp (1996). They found that on a set of 64 read German sentences hand segmented by three humans, the mean deviation in placement of the segment boundaries could be as large as 16 ms, while 93% of the manual segmentations were inside a 15 ms time interval and 96% within 20 ms. So, 20 ms also seems to be a window within which human segmentations are in agreement with one another. The performance of a segmentation algorithm is generally presented in terms of the correct detection rate (CDR), which is expressed as

$$\mathrm{CDR} = \frac{\#\mathrm{boundaries\_correct}}{\#\mathrm{boundaries\_truth}} \times 100, \quad (1)$$

where boundaries_correct are the hypothesized boundaries that fell within the tolerance window distance from the ground truth boundaries, and boundaries_truth are the boundaries in the ground truth segmentation.

A second important issue when comparing the performances of different segmentation methods is the number of boundaries hypothesized: with an equal number of correctly placed boundaries, the method that has a number of hypoth-

esized boundaries closest to the number of actual boundaries in the ground truth is better. This is expressed as a percentage over- (or under-)segmentation (OS):

$$\mathrm{OS} = \left(\frac{\#\text{boundaries\_found}}{\#\text{boundaries\_truth}} - 1 \right) \times 100, \quad (2)$$

where boundaries_found are the boundaries hypothesized by the segmentation algorithm, and boundaries_truth are the boundaries in the ground truth segmentation.

Pereiro Estevan *et al.* (2007) presented an unsupervised speech segmentation method based on maximum margin clustering (see Sec. III for more details). At a tolerance window of 20 ms, and an over-segmentation of −1.4% (i.e., an under-segmentation), the method obtained a CDR of 67.9% on the TIMIT test data. Aversano *et al.* (2001) obtained a CDR of 73.6% correct on 480 utterances produced by 48 speakers from the TIMIT database, at an over-segmentation of 0% and a 20 ms tolerance window, using a method that captures the changes in speech signals defined as a "jump function" and subsequently hypothesizes boundaries at the peaks of the jump function. Qiao *et al.* (2008) tried to solve the segmentation problem by searching for the "optimal segmentation" using a probabilistic framework. Their "rate distortion estimated by a full covariance matrix" method obtained a CDR of 76.7% on the training set of TIMIT, using a 20 ms tolerance window; they, however, did not report the over-segmentation rate. Dusan and Rabiner (2006) presented a method that detects boundaries by searching for the peaks in a spectral transition measure. They obtained a performance of 84.6% of correctly detected boundaries at a 20 ms tolerance window on the training set of TIMIT; they did not report over-segmentation rates.

III. THE SPEECH SEGMENTATION ALGORITHM

As explained above, we investigated the fundamental problems for automatic speech segmentation algorithms. To that end, we used a segmentation algorithm based on acoustic change as a criterion in order to ensure that the analysis is clean. We opted for the unsupervised speech segmentation algorithm presented in Pereiro Estevan *et al.*, 2007. The reasons for this choice are two-fold. First of all, in order to investigate the fundamental problems for automatic speech segmentation algorithms, patterns in the errors are more easily detected when there are more errors. Second, the algorithm by Pereiro Estevan *et al.* (2007) was easily available.

The speech segmentation algorithm relies on a method called maximum margin clustering (MMC) (Xu *et al.*, 2004). MMC is a promising kernel method. It is an unsupervised form of support vector machine (SVM) (Burges, 1998): the two are related by the maximum margin criterion for finding the optimum solution. The objective of MMC is to split a set of unlabeled feature vectors (represented by the dots in Fig. 1) such that the margin separation between the two resulting sets or clusters is maximal. Figures 1(a) and 1(b) are examples of a *non-optimal* dichotomy. The empty region bounded by the two lines is called the margin and should have maximal width, that is, it should be as wide as possible while remaining empty. Figure 1(c) is an example of an optimal dichotomy.

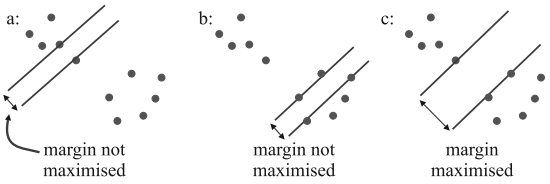


FIG. 1. The maximum margin criterion: (a) and (b) are examples of a non-optimal dichotomy, and (c) is an example of an optimal dichotomy.

The speech was parameterised by 12 mel frequency cepstral coefficients (MFCCs) and log energy and augmented with their first and second derivatives resulting in 39-dimensional MFCC vectors. The MFCCs were computed on windows of 15 ms with a 5 ms frame shift, and cepstral mean and variance normalization was applied.

A sliding window 18 MFCC vectors wide (90 ms) is used to isolate a small section of the signal for analysis. The MMC algorithm is applied to the MFCC vectors inside the window. Each MFCC vector is assigned to either the left or right of the margin based on the maximum margin criterion, resulting in two clusters such that the MFCC vectors on one side of the margin are more similar to one another than the MFCC vectors on the other side of the margin, and ensuring that the margin between the two clusters is maximized. Note that, in the present study, clustering ignores the time ordering of the MFCCs. The cluster assignment is plotted in the first column of Fig. 2(a) with a different shading in each element to indicate the frames' assignments, thus each element in the first column of Fig. 2(a) indicates the cluster assignment of one of the 18 MFCC vectors. The sliding window is shifted by one frame (5 ms) and the process is repeated producing the subsequent columns of Fig. 2(a). Thus, each MFCC vector is part of the analysis window multiple times (i.e., 18 times, as the window is 18 frames wide). When a boundary (a change in the shading) is hypothesized in one column, then in a subsequent column the same boundary should occur one frame earlier (i.e., lower) in the window, as the MFCC

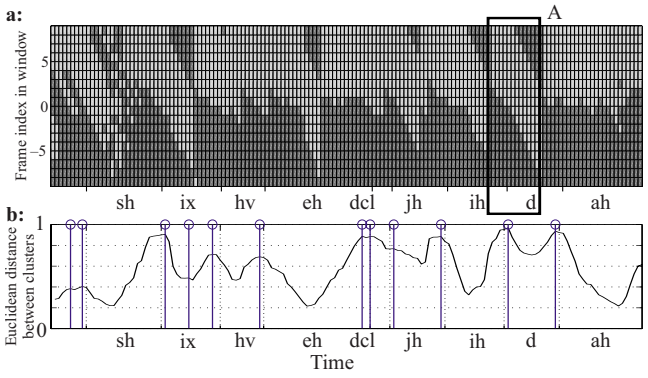


FIG. 2. (Color online) (a) Sliding window frame assignment representation; each column shows the cluster label assignments (light/dark gray) at a given time. (b) Euclidean distance between the means of the clusters; the detected boundaries are indicated by the solid vertical lines. Horizontally, the TIMIT segments and boundaries are indicated with the dashed vertical lines of the phrase "She had your dar(k)" (the [k] is not present in the figure; in IPA: /ʃɪ hɛd jɔr dɑ/).

vector's position has shifted one frame compared to its position in the previous analysis window. Hypothesized boundaries thus manifest themselves as diagonal structures in Fig. 2(a) such as shown in rectangle *A* of Fig. 2(a).

The method used to find potential boundaries consists of a combination of two approaches. The first method detects the diagonal structures in Fig. 2(a) using a mask, which is divided along its diagonal into two: each element in the upper right triangle must match the lighter shaded elements of the graph while the elements in the lower left triangle must match the darker shade. The total number of matching elements in the mask is counted each time. When all of the mask's elements are matched, then a segment boundary is marked at the middle vector. Tuning experiments (on the TIMIT test set) described in Pereiro Estevan *et al.* (2007) showed that the optimal size of the mask is a 4×3 matrix. The second approach plots the Euclidean distance between the centers (means) of the two clusters in the sliding analysis window in MFCC space. When the Euclidean distance between the cluster means shows a peak [vertical lines in Fig. 2(b)] then a boundary is hypothesized. A parameter δ controls the sensitivity of the peak detector. It specifies the minimum amount that the curve must decrease and then increase before another peak is detected. A smaller δ results in more hypothesized boundaries, whereas a greater δ results in fewer hypothesized boundaries. Tuning experiments (on the TIMIT test set) described in Pereiro Estevan *et al.* (2007) showed that $\delta = 0.001$ yielded the best results.

The detected boundaries from both approaches are combined (in a left-to-right fashion) such that the resulting set of detected boundaries consists of all boundaries hypothesized by either approach. However, if both methods hypothesize boundaries within two vectors (10 ms) of each other, then they are replaced by a single boundary located at the frame halfway between the two. We refer to this as "smoothing."

IV. THE SPEECH DATA

This study used the TIMIT speech corpus. TIMIT consists of sentences read by 630 native speakers of eight major dialect regions of American English. It is labeled and segmented by hand in terms of 59 segments; i.e., the 50 phones listed in Table II, six labels for the closure parts of stop consonants (one for each stop; the stops listed in Table II only refer to the release parts of the stop consonants), and three labels for silence, which were collapsed onto one [sil] segment (see last row Table II). Of the 630 speakers in the corpus, 438 (70%) were male. For the analyses, TIMIT's standard test set was used (excluding the sa utterances). The test set used consists of 1344 utterances; 168 speakers each produced eight utterances from a set of 624 different utterances. The average number of boundaries per utterance is 36.5.

V. ANALYZING THE HYPOTHESIZED AND MISSED BOUNDARIES

The number of boundaries hypothesized by the MMC algorithm was similar to the number of boundaries in the transcriptions of the test set of TIMIT. The 1344 utterances of TIMIT's test set contain 45 514 boundaries, while the speech segmentation algorithm hypothesized 44 885 boundaries (resulting in an under-segmentation of 1.4%). Taking the TIMIT manual segmentations as ground truth, 30 926 (67.9%) of the hypothesized boundaries were correctly hypothesized, i.e., they appeared within a distance of 20 ms of the manually placed boundaries in TIMIT. Thus, 14 588 boundaries were missed. The algorithm also hypothesized 13 959 boundaries that do not coincide with segment boundaries in TIMIT; we refer to these as additionally hypothesized boundaries. In these cases, there is apparently a difference between clusters of frames inside the sliding window that is big enough to warrant hypothesizing a boundary, even though there is no segment boundary according to the TIMIT segmentation.

We carried out an in-depth analysis investigating when these additional boundaries occur and when boundaries are missed taking the TIMIT manual segmentations as ground truth. In previous analyses (Scharenborg *et al.*, 2007), we found that, like for automatic speech recognition systems, silence is problematic for our unsupervised speech segmentation algorithm. More specifically, the end of a silence tended to be hypothesized poorly due to problems with the endpointing algorithm. To avoid the endpointing problem, in the current study, we only investigated those boundaries that occurred at least 45 ms from the start or before the end of the reference TIMIT file. This resulted in 10 884 additionally hypothesized boundaries and 13 385 missed boundaries. Note that there are still silence frames, for instance, due to silences between words.

TABLE I. Specification of the AFs and their respective values.

AF	Values
Manner	Approximant, retroflex, fricative, nasal, stop, vowel, silence
Place	Bilabial, labiodental, dental, alveolar, velar, nil, silence
Voice	+voice, −voice
Height	High, mid, low, nil, silence
Backness	Front, central, back, nil
Roundness	+round, −round, nil
Staticity	Static, dynamic

A. Set-up of the analyses

In order to be able to generalize over different segments, we characterized segments by "articulatory features" (AFs). AFs are the acoustic correlates of articulatory properties of speech sounds. We used the set of seven articulatory features shown in Table I. The names of the AFs are self-explanatory, except maybe for *staticity*, which states whether an acoustic change occurs (as, e.g., is the case for diphthongs: [dynamic]) or not ([static]). It might seem that the AF *staticity* will correlate almost perfectly with the hypothesized boundaries; however, this is not the case as *staticity* is related to the bigger manifestations of acoustic change, whereas our segmentation algorithm is also sensitive to smaller acoustic changes. [nil] is used when an AF is not applicable for a segment, for instance, consonants do not have a value for

TABLE II. Feature value specification of each TIMIT phone label in the test set. Note that the affricates /tʃ, dʒ/ are classified as fricatives since the largest parts of these segments are continuous.

Phone	Manner	Place	Voice	Height	Backness	Roundness	Staticity
æ	Vowel	Nil	+voice	Low	Front	−round	Static
ɛ	Vowel	Nil	+voice	Mid	Front	−round	Static
aʊ	Vowel	Nil	+voice	Low	Front	−round	Dynamic
eɪ	Vowel	Nil	+voice	Mid	Front	−round	Dynamic
aɪ	Vowel	Nil	+voice	Low	Front	−round	Dynamic
ɨ	Vowel	Nil	+voice	High	Front	−round	Static
i	Vowel	Nil	+voice	High	Front	−round	Dynamic
ɪ	Vowel	Nil	+voice	High	Front	−round	Static
ə	Vowel	Nil	+voice	Mid	Central	−round	Static
ʌ	Vowel	Nil	+voice	Mid	Central	−round	Static
ɔ	Vowel	Nil	+voice	Low	Back	+round	Static
oʊ	Vowel	Nil	+voice	Mid	Back	+round	Dynamic
oɪ	Vowel	Nil	+voice	Low	Back	+round	Dynamic
ʊ	Vowel	Nil	+voice	High	Back	+round	Static
u	Vowel	Nil	+voice	High	Back	+round	Dynamic
ɑ	Vowel	Nil	+voice	Low	Back	+round	Static
ʉ	Vowel	Nil	+voice	High	Back	+round	Dynamic
l	Approximant	Alveolar	+voice	Nil	Nil	Nil	Dynamic
ɭ	Approximant	Alveolar	+voice	Nil	Nil	Nil	Dynamic
w	Approximant	Velar	+voice	Nil	Nil	Nil	Dynamic
j	Approximant	Velar	+voice	Nil	Nil	Nil	Dynamic
p	Stop	Bilabial	−voice	Nil	Nil	Nil	Dynamic
b	Stop	Bilabial	+voice	Nil	Nil	Nil	Dynamic
t	Stop	Alveolar	−voice	Nil	Nil	Nil	Dynamic
d	Stop	Alveolar	+voice	Nil	Nil	Nil	Dynamic
k	Stop	Velar	−voice	Nil	Nil	Nil	Dynamic
g	Stop	Velar	+voice	Nil	Nil	Nil	Dynamic
ɾ	Stop	Alveolar	+voice	Nil	Nil	Nil	Dynamic
f	Fricative	Labiodental	−voice	Nil	Nil	Nil	Static
v	Fricative	Labiodental	+voice	Nil	Nil	Nil	Static
θ	Fricative	Dental	−voice	Nil	Nil	Nil	Static
ð	Fricative	Dental	+voice	Nil	Nil	Nil	Dynamic
s	Fricative	Alveolar	−voice	Nil	Nil	Nil	Static
ʃ	Fricative	Alveolar	−voice	Nil	Nil	Nil	Static
z	Fricative	Alveolar	+voice	Nil	Nil	Nil	Static
ʒ	Fricative	Alveolar	+voice	Nil	Nil	Nil	Static
h	Fricative	Velar	−voice	Nil	Nil	Nil	Static
ɦ	Fricative	Velar	+voice	Nil	Nil	Nil	Static
tʃ	Fricative	Alveolar	−voice	Nil	Nil	Nil	Dynamic
dʒ	Fricative	Alveolar	+voice	Nil	Nil	Nil	Dynamic
m	Nasal	Bilabial	+voice	Nil	Nil	Nil	Static
n	Nasal	Alveolar	+voice	Nil	Nil	Nil	Static
ŋ	Nasal	Velar	+voice	Nil	Nil	Nil	Static
ɱ	Nasal	Bilabial	+voice	Nil	Nil	Nil	Dynamic
ɳ	Nasal	Alveolar	+voice	Nil	Nil	Nil	Dynamic
ŋ	Nasal	Velar	+voice	Nil	Nil	Nil	Dynamic
ɾ̃	Nasal	Alveolar	+voice	Nil	Nil	Nil	Static
ɚ	Retroflex	Alveolar	+voice	Nil	Nil	Nil	Dynamic
ɹ	Retroflex	Alveolar	+voice	Nil	Nil	Nil	Dynamic
ɝ	Retroflex	Nil	+voice	Nil	Nil	Nil	Dynamic
Sil	Sil	Sil	−voice	Sil	Nil	Nil	Static

front-back (as their place of articulation is defined with the AF *place*) or *roundness*. [sil] is used for silent frames. Note that for *voice*, silence is marked as [-voiced], and for *staticity* as [static]. A final observation: in TIMIT, the silence (i.e., the closure) and release (i.e., the burst) parts of stops have been annotated separately, but in our study the transcription of a sequence of a silence part followed by a release part is changed to represent a single segment. Table II presents an overview of the feature value specification of each of the phone labels in the TIMIT set.

In the first series of analyses (Sec. V B), we investigated the presence of additional boundaries. In the second series of

analyses (Sec. V C), we investigated the segment contexts in which boundaries were missed. Section V D summarizes the most salient results of the analyses. The analyses were carried out using generalized linear mixed-effect models, thus containing both fixed and random predictors, using the logit link function. The fixed predictors are the AFs we defined (see Table I) and the duration of a segment (see below). The random predictors are described below. The parameters of the generalized linear models are set using maximum likelihood estimation. We used contrast coding.[1] A generalized model, with the logit link function, has the form

$$\text{logit } p = c\beta_I \text{AF}_I + \beta_2 \text{AF}_2 + \beta_3 \text{AF}_3 + \cdots + \beta_N \text{ duration},$$

where logit p represents $\log [p(1-p)]$. In our case, p is the probability that a boundary is inserted inside a segment or the probability of missing a boundary (logit p is the "dependent variable"). The constant c is the intercept. The different β's (Chatterjee *et al.*, 2000) represent the relevance (effect size) of the different AFs and duration for the estimation of the logit p: a larger absolute β corresponds to a larger effect of the corresponding predictor.

In the following analyses, only statistically significant (calculated using F-tests) effects are part of the final statistical model and reported. In addition, we report the absolute estimated values of the different β's, with an explanation of whether the likelihood of an additional or missed boundary increases or decreases with the associated feature.

B. Additionally hypothesized boundaries

We investigated whether the presence of an additional boundary within a segment can be predicted based on its context. As indicated above, the algorithm hypothesized 10 884 additional boundaries. When a boundary falls within a segment of which the initial boundary or final boundary is missing, it might be the case that the additional boundary is, in fact, a misallocated initial/final boundary instead of a "true" additionally hypothesized boundary. We, therefore, only investigated those 19 033 segments for which the initial and final boundaries were correctly hypothesized. Of these 19 033 segments, 4079 segments had an additional boundary; these were compared to the 14 954 segments that did not have an additional boundary. Thus, 37.5% (4079 of 10 884) of the additional boundaries hypothesized by the algorithm were true additional boundaries. The number of segments with additional boundaries thus is high enough to be able to detect patterns in the errors made by the unsupervised segmentation algorithm.

For each target segment, as well as its preceding and succeeding segment, the phone label in the TIMIT transcription was determined. Additionally, the segment was rewritten in terms of its AF values (see Table II, for an overview). The duration and the AF values of the target segment, in addition to all two-way interactions, were tested as fixed predictors. Segment durations were calculated from the hand-segmented TIMIT data. The mean segment duration over all analyzed segments was 80.8 ms. As random predictors, speaker identity and the phone identities of the target segment itself and of the preceding and succeeding segments were tested.

We ensured none of the variables correlated with one another in the analyses. In case a variable correlated with another, we either removed the variable from the analysis or reorganized the levels of the variable, as described below. However, for duration we followed a different path. The duration of a segment turned out to correlate with some of the AF values of the segment itself. We orthogonalized the duration of a segment with these AF values. We let the duration be predicted by the AF values with which it correlates, and used the residuals of this model as the "duration" predictor in the analyses. This residual duration thus is the duration that cannot be predicted on the basis of the AF values. This procedure ensures that if we find an effect for duration, it indeed can be attributed to duration and not to the AF values. The residual duration is calculated for every analysis separately.

Since only *manner* can be meaningfully specified for all segments, we first investigated whether the presence of an additional boundary within a segment can be predicted by the *manners* of articulation of the phone sequence. Of the random predictors, speaker identity as well as the phone identities of the target and the preceding segment appeared to contribute to the explanation of the variation ($p<0.05$). Furthermore, we observed robust effects of the residual duration of the segment [$\beta=0.0220$, $F(1,19025)=1399.883$, $p<0.0001$]: additional boundaries are more likely in longer segments. Second, we observed robust effects of the manner of articulation of the segment under analysis [$F(6,19025)=11.646$, $p<0.0001$].

The primary reason that longer segments are more likely to have additional boundaries is because of the applied segmentation algorithm. MMC analyzes a fixed number of frames equivalent to 90 ms of speech. If a segment is longer than 90 ms, the analysis window will contain frames coming from one phone only. MMC will always hypothesize a boundary, even if the acoustic change within the segment is very small as the MMC algorithm is sensitive to the tiniest of changes in the MFCC feature space. If the acoustic change is big enough, it is picked up by the peak detector, or if the MMC algorithm places the hypothesized boundaries consistently over time, the boundary is picked up by the mask detection method (see Sec. III). Both mechanisms result in the segment boundary being hypothesized. In the case of shorter segments whose durations match the size of the analysis window, the boundary hypothesized by MMC is more likely to indeed be a segment boundary because the frames in the analysis window will be from two segments, thus resulting in fewer additional boundaries.

Further statistical analyses allowed us to establish four groups of manner values with the following general trends. As is shown in Fig. 3, stops appeared to have most additional boundaries (32.8% of the stops had an additional boundary), followed by fricatives (25.2%) and silence (23.6%), followed by vowels (20.7%), while retroflexes (14.0%), nasals (7.8%), and approximants (3.4%) appeared to have the smallest number of additional boundaries. It is not surprising that stops get more additional boundaries than any other type of segments. The algorithm is designed to group together frames that are similar. Since stops consist of two distinct parts (remember that the closure and release part of stops are labeled as one

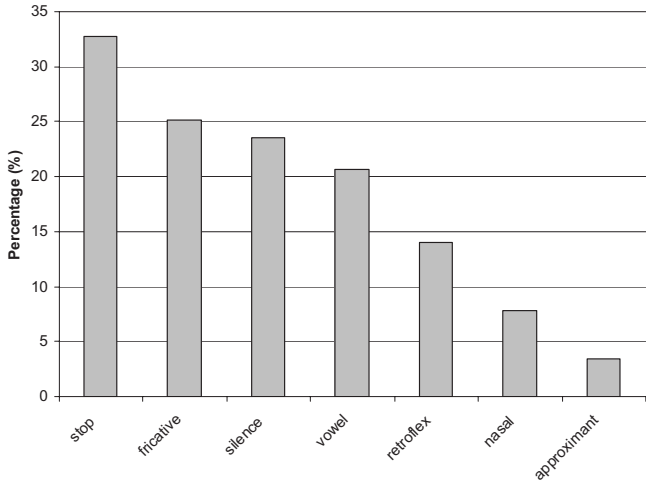


FIG. 3. The percentage of segments containing an additional boundary per *manner* class.

segment—contrary to the standard TIMIT labeling), the algorithm tends to divide the stop into two separate segments resulting in the hypothesis of additional boundaries. Fricatives might show more variation than sonorant consonants resulting in more additional boundaries for fricatives. In case of silence, as the MMC algorithm is sensitive to the tiniest of acoustic change, it is likely that it picks up small background noises, for instance, due to speaker noises or the microphone, resulting in additional boundaries.

The acoustic realizations of vowels, retroflexes, nasals, and approximants are very similar due to the lack of a stricture of the vocal tract sufficient to cause audible turbulence during the production of these segments. It might therefore be surprising that vowels behave differently from retroflexes, nasals, and approximants in that vowels get more additional boundaries. This difference can most likely be attributed to the fact that 29.0% of the [vowel] segments are diphthongs; during the realization of a diphthong vowel the articulators move from one position to the next, resulting in acoustic change. Since the algorithm is designed to group together frames that are similar, the acoustic change results in the hypothesis of (additional) boundaries, and the algorithm divides the vowel into two separate segments. Retroflexes, nasals, and approximants, on the other hand, are more or less static sounds, and this staticity results in fewer additionally hypothesized boundaries.

In order to test the role of the other AFs in the hypothesizing of additional boundaries, we analyzed obstruents, nasal consonants, and vowels separately, and investigated which of their characteristics predict the presence of an additional boundary. For the obstruents, *voice*, *manner* (thus either [stop] or [fricative]), *staticity*, and *place* of articulation are meaningful AFs. The predictors *manner* and *staticity* are correlated, and in order to avoid collinearity, *staticity* was taken out as a predictor. Furthermore, *place* and *manner* of articulation are correlated. To remove this correlation, we relabeled the values of *place* into three levels: [front], [middle], and [back], see Table III for an overview.

We orthogonalized duration with the remaining predictors. The resulting model showed that of the random predictors only the identities of the target and following segments contributed to the explanation of the variation ($ps < 0.0001$). Furthermore, we found a main effect for residual duration [$\beta = 0.0197$, $F(1, 19025) = 626.105$, $p = 0$] and an interaction for residual duration and *manner*, while the main effect of *manner* was not significant: additional boundaries are more often hypothesized in longer stops than in fricatives [$\beta = 0.0155$, $F(1, 19025) = 50.374$, $p < 0.0001$]. Residual duration also interacted with *voice* [$\beta = 0.0106$, $F(1, 19025) = 14.029$, $p < 0.001$], which also showed a main effect [$\beta = 1.3723$, $F(1, 19025) = 30.187$, $p < 0.0001$]: a [+voice] segment has a smaller likelihood for getting additional boundaries than a [−voice] segment; however, this difference is attenuated with increasing residual duration of the segment.

As explained above, stops are more likely to get additional boundaries than fricatives. In longer stops, it is to be expected that the additional boundary is placed further away from the end boundary of the stop, and this boundary is therefore maintained instead of being "smoothed" away (see Sec. III), as may be expected for shorter stops. This will result in more additional boundaries especially for longer stops than for fricatives or shorter stops.

We expected the acoustic change occurring in [+voice] segments to be less than the acoustic change occurring in [−voice] segments for two reasons. First, the closure part of voiced stops may contain voicing, resulting in less acoustic change from the closure to the burst compared to going from a silent closure to a burst in voiceless stops. Second, the burst in voiced stops are not as pronounced as the bursts of voiceless stops, which also results in less acoustic change for the voiced segments. This difference between voiced and voiceless stops attenuates for longer segments, because of the way MMC performs segmentation. As explained above, if a segment is longer than 90 ms, the analysis window will contain frames coming from one phone only, resulting in

TABLE III. Overview of the relabelling of the place AF values for the obstruent analysis.

		#fricatives		#stops	
AF value		Old	New	Old	New
[front]	[bilabial]	0	1284	990	990
	[labiodental]	940		0	
	[dental]	344		0	
[middle]	[alveolar]	2507	2507	1084	1084
[back]	[velar]	204	204	878	878

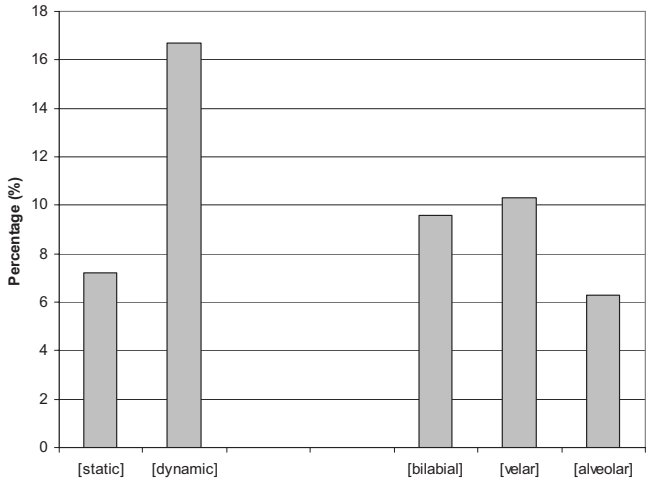


FIG. 4. Results for the analysis of the [nasal] segments: The percentage of segments containing an additional boundary per *staticity* feature value (left) and per *place* feature value (right).

more additional boundaries irrespective of the type of segment. This attenuating effect of duration was also found in other analyses. As it is explained here and above, we will not come back to it below.

As *voice* and *manner* are not meaningful predictors for [nasal] segments, we only studied the role of *staticity* and *place* of articulation in addition to the random predictors speaker identity and the phone identities of the target, preceding, and succeeding segment, for [nasal] segments. In TIMIT, four static and three dynamic nasals occur. The [static] nasals are /m, n, ŋ, ɾ̃/ (/ɾ̃/ is the nasal flap in American English as in wintergreen); the [dynamic] nasals are syllabic /m̩, n̩, ŋ̩/. Duration was again orthogonalized. *Place* only consisted of three AF values, i.e., [bilabial], [alveolar], and [velar], as [labiodental] and [dental] nasals do not occur in English.

Of the random predictors, only phone identity of the succeeding segment appeared to contribute to the explanation of the variation. We found main effects for the residual duration of the segment: additional boundaries were more likely for increasing segment duration [$\beta=0.0410$, $F(1,19025)=149.264$, $p<0.0001$]. *Staticity* also showed a main effect [$\beta=1.2210$, $F(1,19025)=24.174$, $p<0.0001$]: additional boundaries were less often hypothesized in [static] segments than in [dynamic] segments (7.2% of the [static] nasals had additional boundaries compared to 16.7% of the [dynamic] nasals; see also Fig. 4). Finally, *place* of articulation showed a main effect [$F(2,19025)=4.9334$, $p=0.01$].

Like the effect of duration (see the manner analysis for an explanation), the effect of *staticity* is as expected. *Staticity* indicates whether an acoustic change occurs in the segment or not, and as explained above, dynamic change is the basis for hypothesizing boundaries (as is for most unsupervised speech segmentation algorithms). It is thus to be expected that [dynamic] segments more often have an additional boundary than [static] segments.

Further statistical analyses showed that additional boundaries were more likely for [bilabial] (9.6%) and [velar] (10.3%) nasals than for [alveolar] nasals (6.3%), which is also shown in Fig. 4. The movements of the articulators are comparatively bigger when producing a constriction at the lips (for bilabial nasals) or in the back of the oral cavity (for velar nasals) compared to producing a constriction close to the alveolar ridge (for alveolar nasals). The transitions are thus comparatively greater for [bilabial] and [velar] nasals, which implies more acoustic change and thus more additionally hypothesized boundaries for [bilabial] and [velar] nasals.

Vowels differ in their specification for *height, backness, roundness*, and *staticity*. With respect to the AF values for staticity, all diphthong vowels are marked as [dynamic], while all monothongs are marked as [static]. *Backness* correlated with *roundness*, *height*, and *staticity* and was thus removed from the analysis. Duration was then orthogonalized based the remaining three predictors.

Two random predictors appeared to contribute: speaker identity and the phone identity of the target segment ($ps<0.05$). We again found a main effect for residual duration [$\beta=0.0153$, $F(1,19025)=364.728$, $p<0.0001$]: additional boundaries were more likely with increasing duration. The second predictor that showed a main effect was *staticity* [$\beta=0.8915$, $F(1,19025)=38.122$, $p<0.0001$]: as before, the likelihood for additional boundaries was smaller for [static] segments than for [dynamic] segments; however, this difference attenuated with increasing residual segment duration [$\beta=0.0067$, $F(1,19025)=19.334$, $p<0.0001$]. Third, *height* [$F(2,19025)=40.446$, $p<0.0001$] showed a main effect: additional boundaries were more likely to be hypothesized in [low] vowels than in [high] and [mid] vowels; however, also this difference disappears for increasing residual duration of the segment [$\beta=0.0072$, $F(2,19025)=6.377$, $p<0.005$]. Finally, *roundness* [$\beta=0.3490$, $F(1,19025)=10.883$, $p<0.0001$] showed a main effect: additional boundaries were less often positioned in [−round] vowels than in [+round] vowels ([−round]: 19.4% vs [+round]: 29.1%).

The higher number of additional boundaries for [dynamic] and for longer segments is again in agreement with the results presented above. During the production of a [low] vowel, the mouth is more open than during the production of a [high] or [mid] vowel, while for the production of the constriction of preceding and following consonants, the mouth needs to be fairly, or even entirely, closed. As a consequence, the formant transitions are comparatively greater in [low] vowels, which implies more acoustic change and thus more additionally hypothesized boundaries than in [high] and [mid] vowels. A similar explanation holds for the difference in the likelihood of additional boundaries for [+round] and [−round] segments. In order to produce a round vowel, there is more lip movement involved compared to the production of an unround vowel, which results in more acoustic change, and thus an increase in additional boundaries for round vowels.

C. Missed boundaries

We subsequently investigated whether a missing boundary can be predicted based on its segment context. For these analyses, we followed the general procedure of the analysis of the additional boundaries. As indicated above, the algo-

rithm missed 13 385 boundaries. To ensure a boundary was indeed missed and not merely shifted in time, we restricted our analyses to the 7733 missed boundaries that were not preceded or followed by additional boundaries. These were compared to 18 105 boundaries that were not missed and also were not preceded or followed by additional boundaries. Thus, 57.8% (7733 of 13 385) of the missed boundaries were true missed boundaries. Again, the number of missed segments is high enough to be able to detect patterns in the errors made by the unsupervised segmentation algorithm.

Acoustic change leads to the hypothesizing of boundaries. So, if the frames on either side of the boundary are similar, boundaries are more likely to be missed. We therefore expect that only the agreement in *manner* class is a meaningful predictor. This is easily illustrated with a few examples. Imagine a segment [f] followed by [æ]. [f] and [æ] are both [static] segments (see Table II). Even though there is an agreement in *staticity*, the acoustic change between the two segments is rather big, since this is a transition from a fricative to a vowel. Likewise, imagine the transition from an [l] to a [d] sound. Even though there is agreement in *voice* and *place*, the acoustic change occurring when going from the first to the second segment will be quite big again, as this is a transition from an approximant to a stop. We hypothesize that the frames on either side of a boundary are similar when the *manner* of articulation is similar, and thus that boundaries are more likely to be missed when there is an agreement in *manner* AF value for the two segments on either side of the boundary. We tested this in our analysis.

For each (missed or present) boundary, we determined the phone label of the preceding and following segment from the TIMIT transcription and their manner AF values (following Table II). The durations of the segments turned out to correlate with their manners of articulation; therefore, the duration of the preceding segment was orthogonalized with the *manner* of the preceding segment; likewise, the duration of the succeeding segment was orthogonalized with the *manner* of the succeeding segment.

We created a new variable indicating whether the two *manner* AF values of the surrounding segments were similar or dissimilar. We grouped together approximants, vowels, retroflexes, and nasals since for all these sounds the constriction of the vocal tract is minimal, this in contrast to fricatives and stops where there is a clear closure or audible turbulence. [silence] segments were grouped together with fricatives and stops, since stops also have silent portions in them.

The residual durations, the recoded manner AF values of the preceding and succeeding segments, and the two-way interactions were tested as fixed predictors. As crossed random factors, the phone identities of the two segments, as well as the speaker identity were tested. Only those predictors that proved to be significant were kept in the model.

The resulting model showed that of the random predictors only identities of the preceding and succeeding segments contribute to the explanation of the variation. We found a main effect for the agreement in *manner* between the two segments [$\beta=1.154$, $F(1,25832)=853.697$, $p<0.0001$]. As expected, boundaries are more likely to be missed when the segments on either side of the boundary have the same *manner* class. The acoustic changes occurring when going from a fricative, silence, or stop to a retroflex, nasal, vowel, or approximant, or vice versa, are far greater than when going from a segment from one class to a segment from the same class. Furthermore, we found a main effect for the residual duration of the preceding segment [$\beta=0.0279$, $F(1,25832)=681.684$, $p<0.0001$]: the likelihood of missing a boundary decreases with increasing duration of the preceding segment; however, this is less so when both segments have the same *manner* class [$\beta=0.0190$, $F(1,25832)=235.010$, $p<0.0001$]. The residual duration of the succeeding segment also showed a main effect [$\beta=0.0102$, $F(1,25832)=218.029$, $p<0.0001$]: again, an increasing duration reduces the likelihood of missing a boundary. The effect of the duration of the succeeding segment attenuates with increasing duration of the preceding segment, and vice versa [$\beta=0.00004$, $F(1,25832)=11.768$, $p<0.05$].

As explained above, the MMC analysis window is exactly 90 ms long, and MMC only hypothesizes one boundary per analysis window. Thus, in case segments are much shorter than 90 ms, especially when there are multiple segments embedded in the 90 ms analysis window, MMC will miss some of the boundaries between the segments. With increasing segment duration, the analysis window will no longer contain more than two segments, but two or only one segment; this thus results in a smaller likelihood of missing boundaries. Nevertheless, when the segments on either side of the boundary are similar, there is little acoustic change and thus the likelihood of missing that boundary increases compared to when the segments on either side are dissimilar, as the boundary detection method will be less likely to detect the acoustic change.

D. Summary

As is clear from the above analyses of the additional boundaries, the duration of the segments plays a major role in predicting the presence of an additional boundary. In all analyses, duration showed a main effect. If segments are longer, it is more likely that the frames in the 90 ms analysis window all belong to the same segment. MMC will always hypothesize a boundary, even if the acoustic change within the segment is very small as the MMC algorithm is sensitive to the tiniest of changes in MFCC feature space. If the acoustic change is big enough, it is picked up by the peak detector, or if the boundaries were consistently placed over time by the MMC algorithm it is picked up by the mask detection method (see Sec. III), resulting in an additional boundary. The second major finding is that stops get more additional boundaries than any other type of segment. MMC is designed to group together frames that are similar. Since stops typically consist of two distinct parts, whereas our transcription of the [stop] segment puts both parts into one segment, the algorithm often divides the stop into two separate segments (so the boundary detection method is able to detect the boundary), resulting in the hypothesis of additional boundaries. This effect of *manner* is attenuated in longer obstruents, as all long segments are equally prone to additional boundaries due to the characteristics of the applied segmen-

tation algorithm. The fourth major finding is the effect of *staticity*. *Staticity* indicates whether an acoustic change occurs in the segment or not. As explained above, dynamic change is the basis for hypothesizing boundaries. It was thus to be expected that [dynamic] segments more often have an additional boundary than [static] segments.

The analyses of missed boundaries showed that duration also plays a major role in predicting the presence or absence of a boundary: the likelihood of hypothesizing a boundary increases with increasing duration of the preceding and succeeding segment. As explained above, MMC only hypothesizes one boundary per analysis window. In case segments are much shorter than 90 ms, especially when there are multiple segments in the analysis window, MMC will miss some of the boundaries. With increasing segment duration, the analysis window will contain fewer segments thus reducing the likelihood of missing boundaries. The second interesting finding is that boundaries are more likely to be missed when the segments on either side of the boundary have the same *manner* class, which is to be expected since when frames on either side of the boundary are similar, the acoustic change is smaller, thus increasing the likelihood of missing the boundary.

VI. GENERAL DISCUSSION

Despite using different algorithmic implementations, most unsupervised speech segmentation methods achieve similar performance in terms of percentage correct boundary detection (at similar under- or over-segmentation rates; see Sec. II). Nevertheless, unsupervised speech segmentation algorithms are not able to perfectly reproduce manually obtained reference transcriptions. We are interested in trying to unearth the fundamental problems for unsupervised automatic speech segmentation algorithms. To that end, we compared a phone segmentation obtained using only the acoustic information present in the signal with a reference segmentation created by human transcribers, and analyzed the boundaries that were additionally hypothesized and those that were missed.

The comparison of the different unsupervised speech segmentation algorithms and their performances in Sec. II showed that more sophisticated automatic segmentation algorithm will likely result in an improved segmentation algorithm. However, in such more sophisticated systems, it is more difficult to tease apart the effects of different parts, e.g., related to the heuristics, assumptions, or signal processing, of the segmentation algorithm; furthermore, the results become specific to that segmentation algorithm. The criterion that often underlies the decision process for the hypothesis of boundaries in unsupervised automatic speech segmentation algorithms is acoustic change. This let us to the question how good an indicator of a segment boundary acoustic change is. To answer this question, we chose to analyze the results of an automatic segmentation algorithm that only uses acoustic change as the criterion for hypothesizing a segment boundary, and that produced enough errors that made it possible to detect patterns in the errors. This enterprise indicated where acoustic change is indeed a correct indicator of a segment boundary, and where it is not, thus revealing fundamental problems for automatic speech segmentation algorithms.

It is important to note though that the criterion of acoustic change can be applied in several ways with different underlying assumptions of which ours is just one. As a consequence, the results and thus the analyses presented in this paper are still somewhat tied to the speech segmentation algorithm that was used. Nevertheless, we believe that the results found are directly related to the acoustic change criterion and will hold for most other speech segmentation algorithms based on this criterion.

The analyses showed that acoustic change indeed is a fairly good indicator of segment boundaries: 67.9% of the boundaries hypothesized by the MMC algorithm coincide with segment boundaries when there is no over-segmentation. The analyses showed that the MMC algorithm is sensitive to very subtle changes within a segment: it was able to pick up acoustic changes as great as the transition from the closure to the burst of a plosive, but also as subtle as the formant transitions in [low] vowels.

So, why are some boundaries erroneously inserted or missed? The analyses showed that the errors made by the unsupervised speech segmentation algorithm can be split into three groups. Here, we will address these three groups of errors made by the segmentation algorithm. Below, we will give suggestions as to how we believe these errors can (partially) be dealt with in order to improve unsupervised speech segmentation. First of all, the MMC algorithm has problems related to segment duration. Second, the MMC algorithm has problems dealing with adjacent segments that are similar; boundaries are likely to be missed if both segments have the same or a similar manner class, e.g., a nasal followed by a vowel. We should note, however, that this is also problematic for human transcribers. Third, the algorithm has no means of dealing with inherently dynamic phones; the algorithm, for instance, often hypothesizes boundaries between the closure and the burst in stops resulting in two separate segments instead of one stop segment. We believe that most, if not all, of these problems will occur in most current algorithms of unsupervised speech segmentation that use acoustic (rate of) change as the only, or the most important, means to segment speech. In order to develop unsupervised speech segmentation algorithms that are able to create segmentations that are closer to those created by human transcribers, these three main problems need to be dealt with.

A partial solution to the issue of segment duration is related to the way MMC performs speech segmentation. In the analysis, a fairly long window of 90 ms was used which means that we may miss some boundaries in a series of short segments, as the MMC algorithm can only hypothesize one boundary per window. However, a short(er) analysis window is more likely to contain frames belonging to one segment only, resulting in more additional boundaries. One way to improve the algorithm's performance for short segments is to use a multi-class MMC (Zhao *et al.*, 2008), instead of the two-class MMC we used so far; this will enable the detection of several clusters within one analysis window. Another way of improving the algorithm is to use an analysis window that changes in size dynamically or performs the analysis using

multiple window sizes; this could also help reduce the number of additional boundaries in longer segments. This will also make it possible to deal with segments that inherently differ in duration, e.g., consonants are in principle shorter in duration than vowel. A third method is an approach or model that is able to keep track of the duration (as a function of speech rate) and knows when the right conditions are met to hypothesize a boundary. One way of doing this is to make the threshold δ of the Euclidean distance method dependent on the distance to the last hypothesized boundary. This might, however, result in an over-segmentation. Future research will shed light on these possible solutions.

One way of dealing with boundaries that are missed between adjacent segments where acoustic change between the segments is small is to allow the MMC algorithm to oversegment and then apply rules to filter away the additional hypotheses. An example of such a rule might be the removal of the additional boundary that is hypothesized in stops. For instance, if a segment preceding a boundary contains silence or murmur and the segment following the boundary contains a burst, the boundary between the two segments should be removed. These rules may be formulated based on the statistical regularities in large datasets. Possibly, some of these rules may need to be language-dependent. This line of reasoning suggests that we need to know what the sound is in order to be able to segment the speech more accurately.

Unsupervised speech segmentation algorithms are necessarily bottom-up approaches, since they have no prior knowledge of the material (other than the acoustic signal). They hypothesize boundaries without knowledge of possible acoustic phenomena related to the transition from one segment to the next. They have no knowledge about the number of segments there are in the speech material they are supposed to segment, nor about the types of labels of the segments. This thus implies that rules based on durational information, information about the (dis)similarity of adjacent segments, or inherently dynamic phones cannot be obtained in a bottom-up fashion. So perhaps unsupervised speech segmentation can only be improved by including so-called top-down information, i.e., information about segment labels, as is used by supervised speech segmentation algorithms. However, this would be undesirable for reasons listed in Sec. I.

In keeping with the wish to build a system that is flexible and language-independent, we propose a system that, rather than taking the fully supervised approach of training a predefined set of phone models, automatically clusters the hypothesized segments that are derived from acoustic change into broad classes, such as voiced or voiceless, classes for different durations, and/or classes associated with some of the other AF values based on the (dis)similarities in the acoustic signal, and then reconsider the hypotheses accordingly. We refer to this type of information as automatically derived top-down information. This approach would yield a multi-stage system, consisting of a mix of bottom-up and automatically derived top-down information used for the task of speech segmentation. In such a system, the first stage would consist of the original bottom-up unsupervised speech segmentation algorithm; its output would subsequently be smoothed by a model based on the labels of the broad classes and the acoustic signal in a separate step.

To summarize, the analyses showed that acoustic change indeed is a fairly good indicator of segment boundaries: over two-thirds of the boundaries hypothesized by the MMC algorithm coincide with segment boundaries when there is no over-segmentation. The remaining errors highlighted the fundamental problems for unsupervised automatic speech segmentation methods; these are related to segment duration, sequences of similar segments, and inherently dynamic phones. In order to improve unsupervised automatic speech segmentation, we suggest current one-stage bottom-up segmentation methods to be expanded into two-stage segmentation methods that are able to use top-down information based on automatically derived broad classes and rules. In this way unsupervised methods can be improved while remaining flexible and language-independent. Note, however, that some rules might be language-dependent. Obviously, when including these rules, the resulting two-stage segmentation method will no longer be language-independent.

To conclude, it is difficult to hypothesize what would happen to our analyses when they would have been carried out with a more sophisticated speech segmentation algorithm. We therefore would like to encourage future papers on new algorithms of unsupervised automatic speech segmentation to also include an analysis of the errors, along the lines of our analyses of additionally hypothesized and missed boundaries, so it will become clear where the improvement of the more sophisticated segmentation algorithm originates.

ACKNOWLEDGMENTS

The results presented in this article supersede earlier presented results published in the Proceedings of ICASSP 2007, Honolulu, HI, and the Proceedings of Interspeech 2007, Antwerp, Belgium. This research was supported by a Veni-grant from the Netherlands Organisation for Scientific Research (NWO) to Odette Scharenborg and by a EURYI-award from the European Science Foundation to Mirjam Ernestus.

[1]One segment or combination of AF values is used as the "Intercept," i.e., the default, to which all other segments or combinations of AF values are compared.

Aversano, G., Esposito, A., Esposito, A., and Marinaro, M. (**2001**). "A new text-independent method for phoneme segmentation," in Proceedings of the 44th IEEE Midwest Symposium on Circuits and Systems, Fairborn, OH, Vol. **2**, pp. 516–519.

Bridle, J., and Sedgwick, N. (**1977**). "A method for segmenting acoustic patterns, with applications to automatic speech recognition," in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 656–659.

Brugnara, F., Falavigna, D., and Omologo, M. (**1993**). "Automatic segmentation and labeling of speech based on hidden Markov models," Speech Commun. **12**, 357–370.

Burges, C. J. C. (**1998**). "A tutorial on support vector machines for pattern recognition," Data Min. Knowl. Discov. **2**, 1–47.

Chatterjee, S., Hadi, A. S., and Price, B. (**2000**). *Regression Analysis by Example* (Wiley, New York).

Cucchiarini, C. (**1993**). "Phonetic transcription: A methodological and em-

pirical study," Ph.D. thesis, University of Nijmegen, Nijmegen, The Netherlands.
Dusan, S., and Rabiner, L. (**2006**). "On the relation between maximum spectral transition positions and phone boundaries," in Proceedings of Interspeech, Pittsburgh, PA.
Garofolo, J. S. (**1988**). "Getting started with the DARPA TIMIT CD-ROM: An acoustic phonetic continuous speech database," National Institute of Standards and Technology (NIS), Gaithersburgh, MD.
Kim, Y.-J., and Conkie, A. (**2002**). "Automatic segmentation combining an HMM-based approach and spectral boundary correction," in Proceedings of International Conference on Spoken Language Processing, Denver, CO, pp. 145–148.
Kuperman, V., Pluymaekers, M., Ernestus, M., and Baayen, R. H. (**2007**). "Morphological predictability and acoustic duration of interfixes in Dutch compounds," J. Acoust. Soc. Am. **121**, 2261–2271.
Pellom, B. L., and Hansen, J. H. L. (**1998**). "Automatic segmentation of speech recorded in unknown noisy channel characteristics," Speech Commun. **25**, 97–116.
Pereiro Estevan, Y., Wan, V., and Scharenborg, O. (**2007**). "Finding maximum margin segments in speech," in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Honolulu, HI.
Qiao, Y., Shimomura, N., and Minematsu, N. (**2008**). "Unsupervised optimal phoneme segmentation: Objectives, algorithm and comparisons," in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, NV.
Scharenborg, O., Ernestus, M., and Wan, V. (**2007**). "Segmentation of speech: Child's play?" in Proceedings of Interspeech, Antwerp, Belgium, pp. 1953–1956.
Sharma, M., and Mammone, R. (**1996**). ""Blind" speech segmentation: Automatic segmentation of speech without linguistic knowledge," in Proceedings of International Conference on Spoken Language Processing, Philadelpia, PA, pp. 1237–1240.
Wesenick, M.-B., and Kipp, A. (**1996**). "Estimating the quality of phonetic transcriptions and segmentations of speech signals," in Proceedings of International Conference on Spoken Language Processing, Philadelpia, PA, pp. 129–132.
Xu, L., Neufeld, J., Larson, B., and Schuurmans, D. (**2004**). "Maximum margin clustering," in Proceedings of NIPS, Vancouver, Canada.
Zhao, B., Wang, F., and Zhang, C. (**2008**). "Efficient multiclass maximum margin clustering," in Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, pp. 1248–1255.

Nonlinear progressive waves in a slide trombone resonator

Pablo L. Rendón,[a] Felipe Orduña-Bustamante, Daniela Narezo, and Antonio Pérez-López
Grupo de Acústica y Vibraciones, Centro de Ciencias Aplicadas y Desarrollo Tecnológico, Universidad Nacional Autónoma de México, Ciudad Universitaria–México, D.F. 04510 México

Jacques Sorrentini
Institut Fresnel, UMR 6133 CNRS, Campus de Saint-Jerôme, Marseille, France

(Received 17 July 2009; revised 24 November 2009; accepted 3 December 2009)

The propagation of finite-amplitude waves inside a slide trombone is studied through direct pressure measurements corresponding to dynamic extremes. A two-microphone method is used to separate left-moving and right-moving waves inside the trombone, permitting the detection of nonlinear effects associated with progressive waves. It is found that a redistribution of energy across the spectrum toward the higher-frequencies occurs for large distances and high initial pressure levels. These results are consistent with the theory of weakly nonlinear acoustics and also with those reported in this same context by other authors, but which have been obtained mostly through examination of standing-waves. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277221]

PACS number(s): 43.75.Fg [NHF] Pages: 1096–1103

I. INTRODUCTION

The behavior of most wind instruments can be, in general, adequately described by a feedback loop consisting of a nonlinear acoustic generator, such as a reed or the player's lips and a linear resonator, which is the air column inside the horn of the instrument.[1] The nonlinearity of the source is well-established, and is described by Backus and Hundley[2] in the case of a trumpet, and by Elliott and Bowsher[3] for a trombone. Backus and Hundley also obtained results which indicate that the trumpet resonator behaves in an essentially linear manner. Beauchamp,[4] however, observed that in the case of a trombone the radiated harmonics depend on the amplitude of the initial signal and thus correctly questioned the validity of using the linear approximation in this case. These results need not be contradictory, as will be discussed later, due to the great difference in length of the resonators of these instruments. It has since become clear that the brightness of sound associated with loud playing of certain brass instruments is due mainly to nonlinear propagation of waves inside the instrument pipes, as has been shown in experiments by Hirschberg *et al.*[5] Attempts have been made by Msallam *et al.*[6] to give an account of the nonlinear character of intense radiated tones produced by trombones by modeling the whole instrument through a series of sources, resonators, and appropriate filters. Their results are sound but are not compared to the behavior of an actual instrument. Thus, the main difficulty in understanding the manner in which these "brassy sounds" are produced has to do with the description of the nonlinear wave regime inside the instrument resonator. Msallam *et al.* used a time-domain model that incorporates both wave steepening in the slide and thermoviscous losses at the walls of the slide, with good results in the sense that for sound synthesis applications their results improved the quality of reproduction of the aforementioned brassy sounds. Nonlinear distortion effects are, however, underestimated by their model. Shortly after, Thompson and Strong[7] developed a frequency-domain model that again incorporates the effects of non-linear steepening and thermoviscous wall losses and, further, Gilbert *et al.*[8] recently produced a frequency-domain simulation tool to predict the brassiness behavior of brass instruments. The frequency domain is better suited to dealing with the thermoviscous wall attenuation than the time-domain, and it also allows the authors to break up the slide into a series of small elements where the effect of nonlinear distortion can then be applied locally. A comparison between measurements of the radiated spectrum in front of the bell and the results of using both linear and nonlinear models for propagation along the slide yielded the expected result that at loud dynamics the nonlinear model is much more accurate than the linear one.

As pointed out by Crighton,[9] nonlinear propagation is associated with the large-scale distortion of a sound wave through the cumulative effect over a long distance or time of locally small nonlinearities. This nonlinear distortion is observed mainly through a steepening of the wave profile, the scale of which is dependent on both the maximum slope of the initial pressure profile and the distance traveled. The waveform distortion may eventually give rise to thin regions where pressure perturbations become almost discontinuous. The structure of these regions, as is described by Blackstock and Hamilton,[10] depends strongly on whether or not nonlinearity is balanced to some degree by attenuation; in any event, these thin regions are normally called thermoviscous shocks. The structure and dynamics of shocks in this context has been extensively studied by, amongst others, Whitham,[11] and these results have led to what is now called weak-shock theory, which is the study of thermoviscous shocks when the velocities of the waves involved are very small compared to the local speed of sound. Shock wave formation in trombones was first documented by Hirschberg *et al.*[5] through both pressure measurements at the bell of a trombone and flow-visualization techniques. These results established con-

[a]Author to whom correspondence should be addressed. Electronic mail: pablo.rendon@ccadet.unam.mx

clusively that nonlinear wave-steepening must be taken into account at high dynamic levels to describe propagation inside the trombone resonator; at levels equivalent in musical terms to *fortissimo* (ff), shock-formation was observed. The foundations of the theory of nonlinear wave propagation in ducts where wall losses are involved is given by Chester,[12,13] who used a perturbation scheme to give approximate solutions to the sets of nonlinear equations he obtained for both open and closed tubes in near-resonance conditions. The formation of nonlinear standing-waves in an acoustical resonator has been studied by both Van Buren[14] (numerically) and Ilinskii *et al.*[15] (analytically and numerically); in the rather general conditions considered by Ilinskii (the resonator is assumed to have an arbitrary shape), the resulting equations prove to be extremely complicated. Menguy and Gilbert,[16] on the other hand, developed a model for propagation of finite-amplitude plane waves in a cylindrical air-filled tube and obtained uncoupled Burgers-type equations to describe left-moving and right-moving waves valid even for distances beyond shock-formation. Their model is valid for any combination of boundary conditions, and will be discussed further in section II A.

The fact that Menguy and Gilbert's model allows for separation of progressive waves going in opposite directions is rather surprising because it seems to indicate that a linear approximation may be used locally when considering the interaction between these waves. It would thus appear not to be a cumulative phenomenon, but it is in fact a cumulative phenomenon within the framework of weakly nonlinear propagation, allowing for a locally linear approximation. This result is extremely important because the cumulative effects of nonlinearity on propagation of sound waves can only be properly isolated for progressive waves, and not for standing-waves. The main contribution of this paper will be to obtain measurements for the separated right-moving (or incident, as they will be called from here on) and left-moving (or reflected) waves, and to analyze the evolution of the waveforms in order to determine whether nonlinear distortion has taken place for different distances and signal amplitudes. These pressure measurements are not obtained directly, as it were, but using a two-microphone scheme described in Sec. II B.

II. THEORETICAL CONSIDERATIONS

A. Governing equations

The evolution equation for nonlinear acoustic waves propagating unidirectionally in a long tube of uniform cross section has been obtained both by Sugimoto[17] and by Menguy and Gilbert,[16] as well as, in a slightly different form, by Chester.[12] A thorough review of the fundamental equations of nonlinear acoustics is given by Hamilton and Blackstock.[18] Of particular interest in the context of propagation inside the trombone slide is the form of the Burgers equation obtained by Menguy and Gilbert through a multiple scales approach. If we denote by u_i and u_r the velocities of the incident and reflected waves, respectively, the governing equations are written as follows:

$$\frac{\partial u_i}{\partial x} = u_i \frac{\partial u_i}{\partial \tau_i} - \delta \frac{\partial^{1/2} u_i}{\partial \tau_i^{1/2}}, \tag{1}$$

$$\frac{\partial u_r}{\partial x} = -u_i \frac{\partial u_r}{\partial \tau_r} + \delta \frac{\partial^{1/2} u_r}{\partial \tau_r^{1/2}}, \tag{2}$$

where x is the long space coordinate associated with the order of nonlinearity, $\tau_{i,r}$ are the retarded times measured in a frame moving with the local sound speed c in the appropriate direction, and δ is a constant which incorporates the effects of wall friction. The definition of the fractional derivative of order $1/2$ is[19]

$$\frac{\partial^{1/2} u}{\partial \tau^{1/2}} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\tau} \frac{1}{\sqrt{\tau - \tau'}} \frac{\partial u}{\partial \tau'} d\tau'. \tag{3}$$

It has been shown by Gilbert *et al.*[8] that in the vicinity of the shock, we should add to the right hand side of Eqs. (1) and (2) a term akin to that which appears in the classical plane Burgers equation to account for volumetric losses. However, since—as we will show in Secs. III and IV—in this paper we deal with distances of propagation shorter than the shock-formation distance, we do not include these terms. A shock wave propagating in a long, air-filled tube is subjected to wall friction through the formation of a thin boundary layer adjacent to the tube walls. The dissipation produced by this boundary layer is very different in nature to that associated with the diffusivity of sound since it proves to be of a cumulative nature as the wave propagates along the tube. The fractional derivative term in Eqs. (1) and (2) is associated then with a "hereditary" dissipation effect and replaces the more conventional sound diffusivity term in the classical Burgers equations. The key feature, as we discussed in Sec. I, of these equations is that they are uncoupled, and thus a stationary wave inside the tube can be considered as the sum of two progressive waves, much as it would be in the linear case. In Sec. II B, a method is described which will allow us to separate these two progressive waves having measured the acoustic pressure fields at different points along the trombone slide.

B. The two-microphone method

The sound pressure pulses measured in the trombone (reported in Sec. III), in spite of initially having a short duration relative to the acoustic travel time along the instrument, very soon start to broaden, and combine with reflections at the ends, making it very difficult to distinguish between wave components traveling in different directions. A wave separation technique, developed by Seybert and Ross[20] and fine-tuned by Schultz *et al.*[21] and Orduña *et al.*,[22] was used in order to obtain an approximation to these traveling wave components, as explained next.

The sound pressure measured at two closely spaced microphone positions x_1, x_2 along the instrument, is assumed as a linear superposition of waves traveling in both directions (we argue that for sufficiently small separations, this representation should remain approximately valid also for nonlinear disturbances):

$$P(k,x_1) = A\exp(-ikx_1) + B\exp(+ikx_1), \tag{4}$$

$$P(k,x_2) = A\exp(-ikx_2) + B\exp(+ikx_2), \tag{5}$$

where $k=\omega/c$ is the wavenumber, and A and B are complex Fourier amplitudes corresponding to incident and reflected wave components traveling toward, and from the end (bell) of the instrument, respectively. With reference to the middle point between the two microphones $x=(x_1+x_2)/2$, we can set $x_1=-\delta$, $x_2=+\delta$, where $x_2-x_1=2\delta$ is the microphone separation. Solving for the Fourier amplitudes of the traveling waves yields

$$A(k,x) = \frac{P_1\exp(-ik\delta) - P_2\exp(-i3k\delta)}{1-\exp(-i4k\delta)}, \tag{6}$$

$$B(k,x) = \frac{P_2\exp(-ik\delta) - P_1\exp(-i3k\delta)}{1-\exp(-i4k\delta)}, \tag{7}$$

which, in principle, should achieve the wave separation objective. Note that we have abbreviated $P_1=P(k,x_1)$ and $P_2=P(k,x_2)$. Time signals of the traveling wave components are obtained by inverse Fourier transformation.

In practice, however, we must contend with the following complications: (a) numerical singularities around the first two zeros of the denominator $1-\exp(-i4k\delta)$, which occur at frequencies $f=0$ and $f=c/4\delta$; (b) instrumentation errors in the microphone signals; more especially, phase errors at low frequencies, where microphone separation is much smaller than the acoustic wavelength, upsetting precise calculation of the numerator when phase errors in the microphones become comparable to $|\Delta\phi|\approx k\delta$. By choosing a measurement frequency span of 6400 Hz, the denominator zero at $f=c/4\delta \approx 14\,350$ Hz (with $\delta=6$ mm) is directly avoided. The zero at $f=0$ is handled by simply setting $A=B=0$. Finally, microphone phase errors are handled by high-pass filtering the frequency spectra of the measured sound pressure signals, using a Butterworth highpass power transfer response, as follows:

$$|H(f)|^2 = \frac{f^{2N}}{f_c^{2N}+f^{2N}}, \tag{8}$$

where $f_c=100$ Hz is the filter cut-off (half-power) frequency, and $N=2$ is the filter order (choosen relatively low, in order to obtain a mild frequency cut-off, and short time-domain ringing).

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. Experimental setup

Measurements were done on a Silvertone slide trombone, bought new, and considered among some professional Mexican brass players as a good-quality student instrument. The trombone was prepared, as shown in Fig. 1, with the slide fully extended, somewhat beyond of what is known in trombone playing practice as the *seventh position*. Sound excitation was provided by an 8 in. loudspeaker driver unit, closely coupled to the mouthpiece-end tube of the trombone, but without a mouthpiece attached. The loudspeaker front face was covered by a 1 in. thick piece of conglomerate

FIG. 1. Picture showing some details of the experimental setup used for sound pressure pulse measurements in a slide trombone.

wood composite, with a small circular hole and a small cylindrical plastic (PVC) fitting part, coupling to the mouthpiece-end of the trombone. Only a small volume, corresponding to the recessed loudspeaker conic surface, was left enclosed between the loudspeaker cone, the wood cover, the PVC fitting, and the trombone tube. Pairs of small 1.5 mm diameter holes were drilled, $2\delta=12$ mm apart for each pair (except $2\delta=11$ mm for the first pair at Pos. 1), at the locations shown as Pos. 1–3 in Fig. 2 along the trombone. The holes will be labeled from now on as 1.1, 1.2, 2.1, 2.2, 3.1, and 3.2, with the first digit corresponding to the position, as shown in Fig. 2, and the second digit indicating whether it is the first or second hole in the position, with the positive direction going from the mouthpiece to the bell of the trombone. Henceforth, by referring simply to position 1, etc., we effectively mean hole 1.1, etc. These holes were used to insert a sound pressure probe microphone Brüel&Kjaer type 4182. Measurements were taken one hole at a time, sequentially, with all other holes kept covered by small plasticine blobs, in order to maintain a good experimental approximation to a continuous closed bore condition. Some plasticine was also used to seal the very thin gap around the probe microphone inserted in the measuring hole. Ten measurements were made in each hole, re-inserting the probe microphone anew every time. There will be a trade-off balancing uncertainty with regard to repeatability of the experiment and errors derived from the usage of several microphones, but in the end we have chosen to utilize one unique microphone to make all measurements.

Transient sound pressure pulse measurements were done using a dual-channel analyzer Brüel&Kjaer type 2034, and transferred to a personal computer through a GPIB/IEEE-488.1 instrumentation interface. The internal signal generator

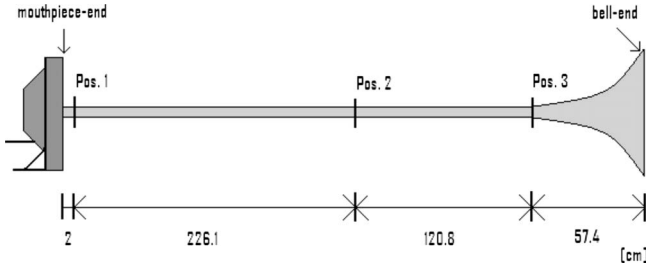


FIG. 2. Diagram illustrating the three measurement positions along the body of the trombone.

of the analyzer was used to feed a power amplifier, and then the loudspeaker, with an impulse signal repeated every 125 ms. The interval between pulses was sufficient for sound to decay considerably inside the trombone. The analyzer's *time (signal) enhancement* mode was used to average 100 pulse-synchronized sound pressure transients, consisting of 2048 digital samples captured at a 16 384 Hz sampling rate (125 ms time duration). Similar measurements were done ten times, and further averaged; producing a single, 1000 times averaged, sound pressure pulse for each hole, statistically taking into account variations in repeated insertion and positioning of the probe microphone. The input signal level was varied via the power amplifier, in order to achieve two different peak driving amplitudes (measured as the sound pressure pulse peak amplitude at the first hole, closest to the loudspeaker) of 90 and 148 dB, the higher second level being the maximum safely and consistently achievable with the equipment used. In this way, 12 averaged sound pressure pulses were obtained, at each of the six holes and for the two driving amplitudes. Measurements were carried out in a relatively quiet, thermally stable laboratory room. Sound pressure peak levels measured inside the trombone, similarly as above, but without excitation from the loudspeaker, were registered at about 50 dB, while the ambient room temperature was monitored at a constant 23 °C during all of the measurement sessions.

B. Results

As we have discussed previously, measuring progressive waves inside the trombone slide permits us to test two important features of nonlinear propagation which cannot be checked if the measurements obtained are of a standing-wave nature. Namely, we expect to see the scale of nonlinear effects increases with distance of propagation, but also with the initial amplitude of the pulse produced by the loudspeaker. Thus, we consider two rather different pressure levels to compare in order to establish whether the effects of nonlinear propagation are at all present, and if they are, whether they are more pronounced for the louder pulse than for the softer pulse, as expected. To establish a correspondence between a given pressure level inside the trombone and a qualitative measure of the loudness of the trombone played normally as a musical instrument, we recorded the pressure level at measurement position 2 on the trombone, playing at musical dynamic levels *piano* and fortissimo. The results are in agreement with the results obtained by Hirschberg *et al.*[5] for the pipe pressure: the rms pressure levels in this case are 1056 Pa (154 dB) for piano and 2530 Pa (162 dB) for fortissimo. The sounded note is approximately F♯2 in both cases.

Because of the distances between measurement positions as described in Sec. III A, it is extremely important to generate very short pulses so that at some given time it is possible to spatially confine it between positions 1 and 3; see Fig. 2. Once the pulse is reflected on the bell-side of the trombone, the pulse will broaden as a consequence of the geometry of the bell, as described by Pyle,[23] thus the necessity of making the pulses as short as possible. There is a

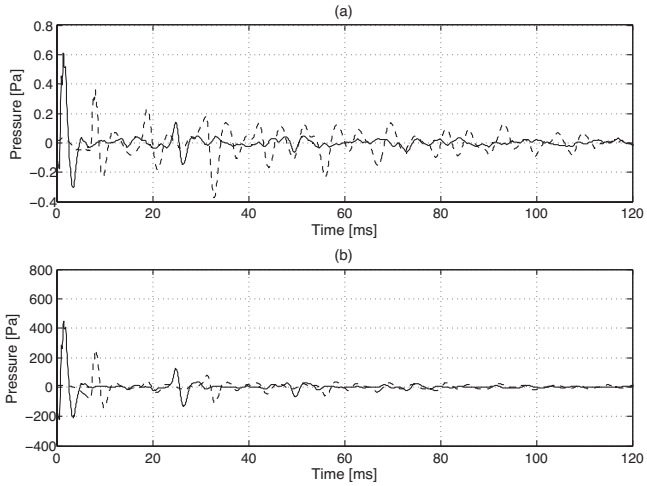


FIG. 3. Incident waves resulting from the separation method obtained at positions 1 (solid line) and 2 (dotted line) for a duration of 120 ms for the (a) 90 dB pulse and the (b) 148 dB pulse.

trade-off, however, associated with generating very short pulses through a loudspeaker, and it is that very large amplitudes of the pressure perturbation then become difficult to achieve. Nonetheless, the scale of nonlinear perturbation effects depends strongly not on the average amplitudes of the time-signals but on the maximum slope of the pressure pulses, which, since the shape of the pulses is essentially the same at different amplitudes, is also related to the maximum amplitude.[9] Thus we have chosen pulses with maximum amplitudes of 0.63 Pa (90 dB) and 502 Pa (148 dB) to illustrate different dynamics; the difference between these pressure levels is, in effect, of the order of 10^3. We must note, though, that the 90 dB pulse would then correspond to extremely soft dynamics, and the 148 dB pulse would still represent a dynamic level below piano. These pulses are shown in Fig. 3. Notice that in both cases the duration of the pulses is roughly 0.008 s, which would correspond to a distance within the tube of roughly 2.76 m, which is less than the distance between positions 1 and 3 on the trombone. In order to appreciate more clearly the details of the waveforms, the time signal will be truncated so as to follow the evolution of the pulses through only two reflections—one at the bell-end of the trombone and one at the mouthpiece-end.

The two-microphone separation method described in Sec. II B is used to calculate the incident and reflected waves at all positions after several reflections have ocurred both at the mothpiece-end and at the bell-end of the trombone. At position 3 this separation method was found to be less reliable than at the first two positions due to its proximity to the bell-end of the trombone. The results of this separation for positions 1 and 2 are shown, for the incident pulse, in Fig. 3. The time delay between pulses is of the order of 10 ms, which is as expected since the full length of the extended trombone is 4.06 m. It is, however, clearly not possible to identify nonlinear propagation effects solely by examining these wave profiles. This is due to the fact that the pressure amplitudes involved in our experiment are not large enough because of the method used, where an electro-acoustic source is chosen as a sound source, and this represents an

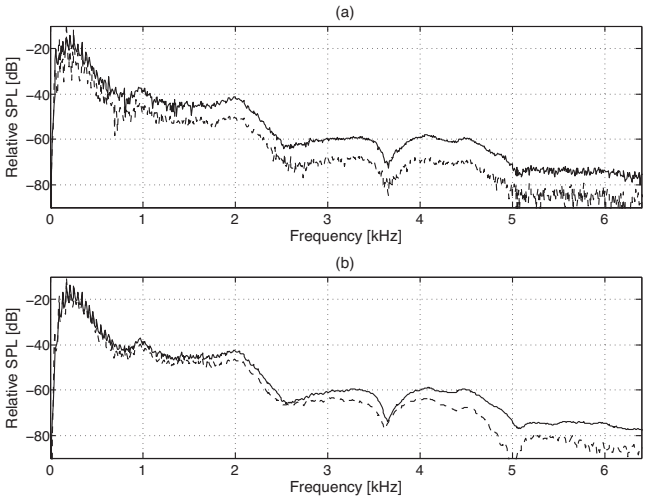


FIG. 4. Spectra of the incident waves resulting from the separation method obtained at positions 1 (solid line) and 2 (dotted line) for a duration of 120 ms for the (a) 90 dB pulse and the (b) 148 dB pulse.

important limitation of this type of scheme. Using a trombone player as a source, it is possible to achieve significantly louder sound levels, thus making it possible to observe directly in the time-domain the distortion of the wave profiles, as shown by Hirschberg *et al.*[5] Examination of the spectrum of these waves, portrayed in Fig. 4, yields the result that the high-frequency content of the spectrum is more significant for position 1 than it is for position 2. This result is, in principle, surprising, but we must consider that attenuation will be more efficient for high-frequencies, as explained by Pierce,[24] and will here outweigh the effects of energy pumping due to nonlinear distortion. The effects of attenuation are, however, still linear, so the fact that the difference between the spectra measured at positions 1 and 2 is less marked for the pulse of greater amplitude can only be accounted for by nonlinear propagation effects, which will be more prominent precisely for the pulse of greater amplitude.

Hirschberg *et al.*[5] successfully detected the effects of nonlinear propagation through observation of the distortion of the pressure waveforms, whereas Thompson and Strong[7] examined the frequency spectrum to look for evidence of nonlinear propagation. The waveforms obtained at high dynamic levels by Hirschberg *et al.* match the results obtained for nonlinear standing-waves by Van Buren[14] and Ilinskii *et al.*[15] quite well. There is a problem in following this approach when producing pulses rather than standing-waves, and it has to do with the fact that attenuation inside the tube, as described in Eqs. (1) and (2), will reduce the overall amplitude of the pulses and thus make pronounced slopes less noticeable. Taking this difficulty in mind, the approach followed by Thompson and Strong is to be preferred in this particular case as energy spectral densities (ESDs) can be normalized in order to permit comparison of the high-frequency spectra. In Sec. IV, we describe the criteria which we will use in order to establish the scale of nonlinear effects associated with the incident pulses for the two dynamic levels under consideration.

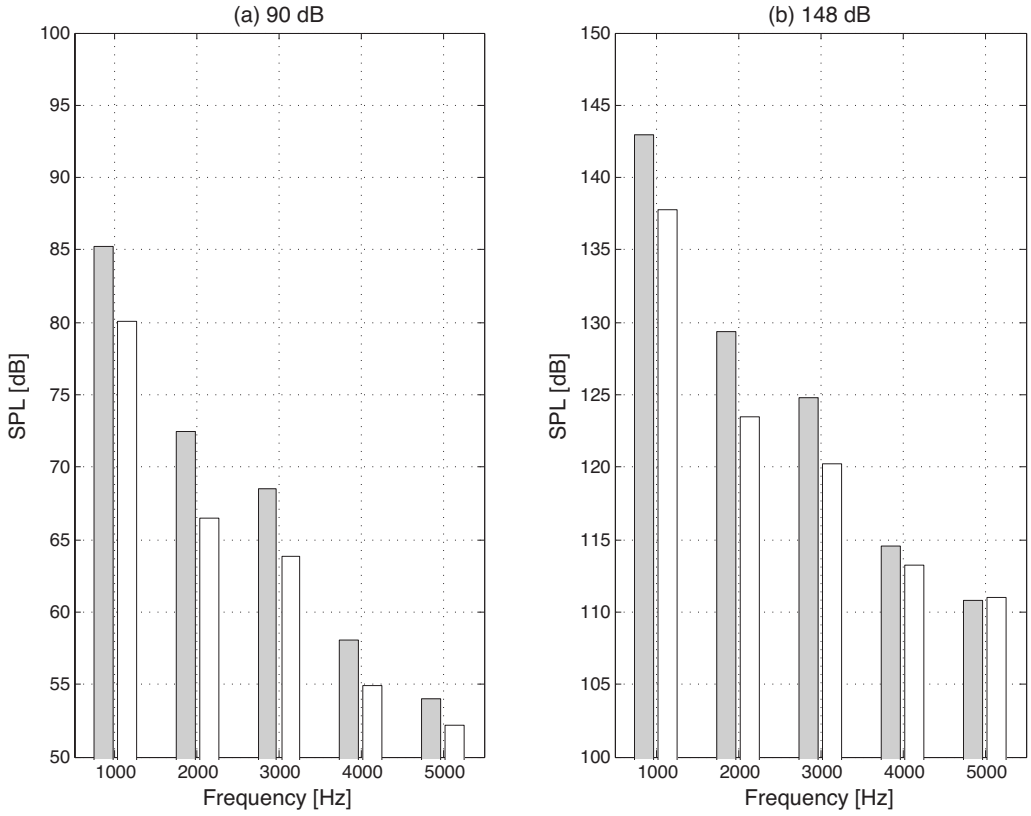


FIG. 5. Relative energy transfers of the incident waves between positions 1 and 2 for five frequency bands of width 1000 Hz, and for both the (a) 90 dB pulse and the (b) 148 dB pulse. The gray bars correspond to position 1, and the white bars correspond to position 2.

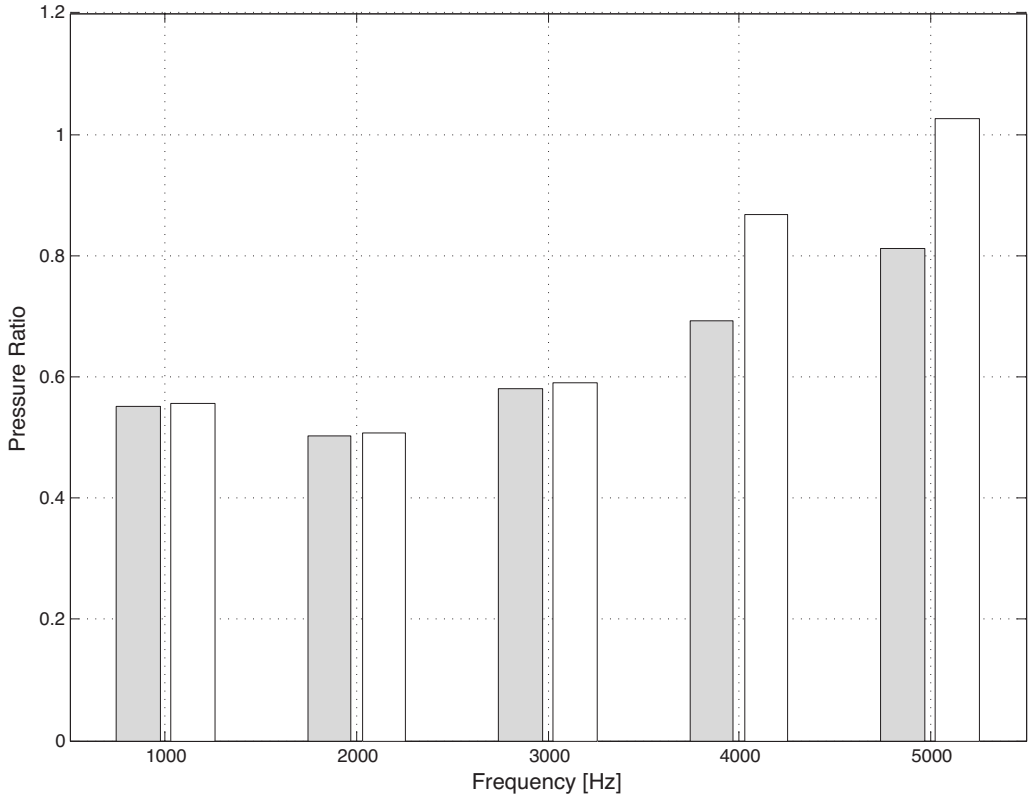


FIG. 6. Ratio between pressure levels of the incident wave at positions 2 and 1 for five frequency bands of width 1000 Hz and initial amplitudes of 90 dB (gray) and 148 dB (white).

IV. SPECTRAL ANALYSIS OF THE PRESSURE MEASUREMENTS

In the time domain, the nonlinear propagation of signals in nondispersive media is characterized above all by a deformation of the wave profiles which causes steep gradients to form, eventually resulting in the traveling regions of finite-width which we have already mentioned, known as shock fronts. From the point of view of spectral behavior, all spectral components interact with similar efficiency, since they are all in resonance with one another, and as a result we observe a broadening of the spectrum allied to a multiplication of spectral lines in the manner of a cascade. When an intense noise wave propagates, the different spectral components that make it up interact in such a way that energy is redistributed over the spectrum. The wave steepening that occurs previous to shock-formation leads to an energy transfer to high-frequencies.[25] As we have previously discussed, at the end of Sec. III B, when the pulses propagate for long distances (several lengths of the trombone), attenuation becomes the dominant influence on the wave profile, and nonlinearity merely mitigates its effect—any pulse propagating inside the trombone would eventually die away. The effects of attenuation are already clearly visible as the pulse arrives at the second position, but the effects of nonlinear propagation would not be easily appreciated by simply examining the slopes of the time signals. In order to detect the aforementioned redistribution of energy which characterizes nonlinear propagation, we have divided the spectrum, ranging from 500 to 5500 Hz, in five bands of equal width (1000 Hz)—the resolution of the frequency spectrum is accounted for when delimiting the bands. Then, the ESD of both pulses has been numerically calculated for each interval at different measurement positions. The ESD is known to be useful in the study of broadband noise propagation,[26] and the idea to use it here comes precisely from the fact that the energy cascade which characterizes nonlinear propagation can be detected more reliably through comparison of the ESD over suitably sized intervals of the spectra than through comparison of the spectra themselves.

Plots constructed in this manner for the incident pulse are shown in Fig. 5 (in a decibel scale, for clarity) of the relative change between positions 1 and 2. From these plots we may observe that for the 90 dB pulse there is practically no transfer of energy to the high-frequency spectrum, whereas in the case of the 148 dB pulse the pumping of energy to the higher-frequencies is quite noticeable, as evidenced by the rightmost bars in Fig. 5. Since thermoviscous losses occur as the pulses propagate, it is not possible to simply compare the values, for a given pulse, of the ESD for the different intervals between any two positions, but if instead we consider the quotient of the ESDs corresponding to two positions, we may obtain what is, in fact, a plot of the relative change in power for five frequency bands between these two positions. In order to compare the rate at which energy is transferred—if at all—toward the high-frequency spectrum for different initial amplitudes, and at the same distances, we have plotted, in Fig. 6, the ratio between pressures at positions 2 and 1 for each frequency band and for each initial amplitude. It is seen that for both initial amplitudes the transfer of energy is more marked for higher-

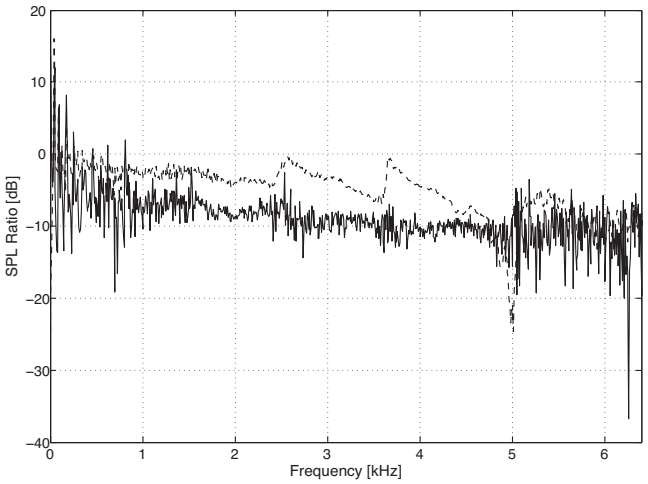


FIG. 7. Ratio between SPLs of the incident wave at positions 2 and 1 for initial amplitudes of 90 dB (solid line) and 148 dB (dotted line).

frequencies, but also that at these frequencies the energy transfer is more significant for the 148 dB pulse, as expected.

In fact, if we plot (Fig. 7) the ratio of sound pressure levels (SPLs) between the pulses at positions 2 and 1 we observe that the rate of decay of the spectrum at high-frequencies is less marked for the 148 dB pulse than it is for the 90 dB pulse, an effect which could already be observed in Fig. 4. That the effect of nonlinear steepening is not more marked is not only due to the dominant effect of attenuation but also to the fact that the shock-formation distance is not attained by any of the pulses before they are reflected. For a plane pressure wave given by $p=p_0 \sin(2\pi ft)$, the shock-formation distance x_s is $x_s=(2\pi f\beta M)^{-1}$, where β is the parameter of nonlinearity, f is the frequency, and M is the acoustic Mach number.[10] Substituting the values of these parameters for air, we obtain $x_s \sim (6.5\times 10^6)/(f*p_0)$. For frequencies and initial amplitudes of the range, we have considered in our analysis, say, $f=1000$ Hz and $p_0=500$ Pa, the shock-formation distance is of the order of 13 m, so that even without considering the effect of thermoviscous dissipation it is clear that no shock waves will be formed inside the trombone. Hirschberg *et al.*[5] did observe shock waves because they produce signals with $p_0 \sim 4$ kPa, so that in their case the shock-formation distance is approximately 1.5 m, well within the trombone resonator. These sound levels, which represent an increase of about 30 dB with respect to the levels we achieve are made possible, as has been mentioned previously, by their use of a trombonist as a sound source. Thus, all the results obtained for the incident pulse are consistent with nonlinear propagation theory: the scale of nonlinearity increases both with maximum pressure amplitude and with propagation distance. Our results are also all consistent with the classical theoretical predictions on the subject of propagation of broadband waves, first proposed by Rudenko and Soluyan.[26]

V. CONCLUSIONS

The effects of nonlinear cumulative distortion on the propagation of finite-amplitude waves inside a trombone have been classically studied by either producing or modeling standing-waves inside the pipe. A more natural and convenient way to examine nonlinear wave propagation is to assume weakly nonlinear propagation inside the trombone in order to be able to isolate progressive waves. The balance between nonlinearity and thermoviscous attenuation that leads to the steepening of wave profiles becomes quite apparent in this manner. A scheme involving quotients of the power spectral density of narrow frequency bands has been devised in order to confirm that the energy cascade toward the high end of the spectrum that accompanies the wave steepening does indeed occur. Furthermore, it is established that the steepening takes place at a rate directly related to both the initial wave amplitude and the distance traveled. The increase in power of the higher harmonics radiated by a trombone, leading to what is known as brassy sound, is explained in its majority then by nonlinear propagation effects, quite apart from any nonlinearity observed in the feedback system established when standing-waves are produced.

As for the reflected pulse, it has been established by Msallam *et al.*[6] among others that the trombone bell acts in some sense as a high-pass filter, so that we should expect the high-frequency spectrum of the reflected signal to be impoverished. Our results indicate that the diminished amplitude of the reflected pulse affects its propagation in such a way that the effect of attenuation gains precedence over that of nonlinearity and the same type of analysis which proved successful for the incident waves could not be replicated for the reflected waves because of this. Of interest from a musical point of view, however, would be the study of the transitory signals produced by a trombone player upon attacking a note, as their evolution and distortion could be tracked along the length of the instrument.

ACKNOWLEDGMENTS

The authors whish to thank DGAPA, UNAM for its financial support through Project No. PAPIIT IN120008. They also wish to acknowledge Sebastián Contreras Osorio, who played the trombone, and César Velázquez Hernández, who gave us advice on acquiring and taking care of the trombone used.

[1] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments* (Springer-Verlag, New York, 1998).
[2] J. Backus and T. C. Hundley, "Harmonic generation in the trumpet," J. Acoust. Soc. Am. **49**, 509–519 (1971).
[3] S. J. Elliott and J. M. Bowsher, "Regeneration in brass wind instruments," J. Sound Vib. **83**, 181–217 (1982).
[4] J. W. Beauchamp, "Analysis of simultaneous mouthpiece and output waveforms of wind instruments," Technical Report, Audio Engineering Society Preprint No. 1626, New York, NY, 1980.
[5] A. Hirschberg, J. Gilbert, R. Msallam, and A. Wijnands, "Shock waves in trombones," J. Acoust. Soc. Am. **99**, 1754–1758 (1996).
[6] R. Msallam, S. Dequidt, R. Caussé, and S. Tassart, "Physical model of the trombone including nonlinear effects. Application to the sound synthesis of loud tones," Acust. Acta Acust. **86**, 725–736 (2000).
[7] M. Thompson and W. Strong, "Inclusion of wave steepening in a frequency-domain model of trombone sound production," J. Acoust. Soc. Am. **110**, 556–562 (2001).
[8] J. Gilbert, L. Menguy, and M. Campbell, "A simulation tool for brassiness studies," J. Acoust. Soc. Am. **123**, 1854–1857 (2008).
[9] D. G. Crighton, "Nonlinear acoustics," in *Modern Methods in Analytical Acoustics* (Springer-Verlag, New York, 1992), pp. 648–670.
[10] D. T. Blackstock, M. F. Hamilton, and A. D. Pierce, "Progressive waves in

lossless and lossy fluids," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, London, 1998), pp. 65–150.
[11]G. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1999).
[12]W. Chester, "Resonant oscillations in closed tubes," J. Acoust. Soc. Am. **18**, 44–66 (1964).
[13]W. Chester, "Resonant oscillations of a gas in an open-ended tube," Proc. R. Soc. London, Ser. A **377**, 449–467 (1981).
[14]A. L. Van Buren, "Mathematical model for non-linear standing waves in a tube," J. Sound Vib. **42**, 273–280 (1975).
[15]Y. A. Ilinskii, B. Lipkens, T. S. Lucas, T. W. Van Doren, and E. A. Zabolotskaya, "Nonlinear standing waves in an acoustical resonator," J. Acoust. Soc. Am. **104**, 2664–2674 (1998).
[16]L. Menguy and J. Gilbert, "Weakly nonlinear gas oscillations in air-filled tubes; solutions and experiments," Acust. Acta Acust. **86**, 798–810 (2000).
[17]N. Sugimoto, "Burgers equation with a fractional derivative; hereditary effects on nonlinear acoustic waves," J. Fluid Mech. **225**, 631–651 (1991).
[18]M. F. Hamilton and D. T. Blackstock, eds., *Nonlinear Acoustics* (Academic, New York, 1998).
[19]N. Sugimoto, "Generalized burgers equations and fractional calculus," *Nonlinear Wave Motion* (Longman Scientific & Technical, Harlow, England, 1989), pp. 162–169.
[20]A. Seybert and D. Ross, "Experimental determination of acoustic properties using a two-microphone random-excitation technique," J. Acoust. Soc. Am. **61**, 1362–1370 (1977).
[21]T. Schultz, M. Sheplak, and L. N. Cattafesta, "Uncertainty analysis of the two-microphone method," J. Sound Vib. **304**, 91–109 (2007).
[22]F. Orduña-Bustamante, M. A. Escobar-Reina, and J. I. Cervantes-Cruz, "The acoustic multimeter: A sensor for sound waves in ducts," in Proceedings of the 16th Congress of the Mexican Society of Instrumentation (2001).
[23]R. W. Pyle, "Effective length of horns," J. Acoust. Soc. Am. **57**, 1309–1317 (1975).
[24]A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, Melville, NY, 1989).
[25]S. N. Gurbatov and O. V. Rudenko, "Statistical phenomena," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock, (Academic, London, 1998), pp. 377–398.
[26]O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Consultants Bureau, New York, 1977).

Ultrasound characterization of red blood cell aggregation with intervening attenuating tissue-mimicking phantoms

Emilie Franceschini,[a] François T. H. Yu, and François Destrempes
Laboratory of Biorheology and Medical Ultrasonics, University of Montreal Hospital Research Center (CRCHUM), Pavillon J. A. de Sève (Room Y-1619), 2099 Alexandre de Sève, Montréal, Québec H2L 2W5, Canada

Guy Cloutier
Laboratory of Biorheology and Medical Ultrasonics, University of Montreal Hospital Research Center (CRCHUM), Pavillon J. A. de Sève (Room Y-1619), 2099 Alexandre de Sève, Montréal, Québec, H2L 2W5, Canada and Department of Radiology, Radio-Oncology and Nuclear Medicine and Institute of Biomedical Engineering, University of Montreal, Montreal, Québec, H3T IJ4, Canada

(Received 10 June 2009; revised 22 November 2009; accepted 27 November 2009)

The analysis of the ultrasonic frequency-dependent backscatter coefficient of aggregating red blood cells reveals information about blood structural properties. The difficulty in applying this technique *in vivo* is due to the frequency-dependent attenuation caused by intervening tissue layers that distorts the spectral content of signals backscattered by blood. An optimization method is proposed to simultaneously estimate tissue attenuation and blood structure properties, and was termed the structure factor size and attenuation estimator (SFSAE). An ultrasound scanner equipped with a wide-band 25 MHz probe was used to insonify porcine blood sheared in both Couette and tubular flow devices. Since skin is one of the most attenuating tissue layers during *in vivo* scanning, four skin-mimicking phantoms with different attenuation coefficients were introduced between the transducer and the blood flow. The SFSAE gave estimates with relative errors below 25% for attenuations between 0.115 and 0.411 dB/MHz and $kR<2.08$ (k being the wave number and R the aggregate radius). The SFSAE can be useful to examine *in vivo* and *in situ* abnormal blood conditions suspected to promote pathophysiological cardiovascular consequences.
 [DOI: 10.1121/1.3277200]

PACS number(s): 43.80.Cs, 43.35.Bf, 43.35.Yb, 43.80.Ev [FD] Pages: 1104–1115

I. INTRODUCTION

Ultrasonic tissue characterization techniques using the radio frequency (rf) backscattered signals have received broad interest for the past 25 years. One approach is to use the magnitude and frequency dependence of the rf backscatter spectrum in order to quantify the tissue structures such as the size, acoustic impedance, and concentration of the scatterers. Many *in vitro* and *in vivo* experiments have been performed to demonstrate the utility of this approach for characterizing the eye,[1,2] liver,[3] kidney,[4] prostate,[5] and breast.[6] Recently, the frequency dependence of the ultrasound (US) backscatter coefficient was studied to assess the level of red blood cell (RBC) aggregation.[7]

It is well known that when RBCs are under low shear rates ($<10\ s^{-1}$), they interact strongly with each other and form complex three-dimensional (3D) rouleaux structures. When the shear rate increases, rouleaux structures disaggregate. This phenomenon is normal and occurs in the circulation of many mammalian species. However, RBC hyperaggregation, an abnormal increase in RBC aggregation, is a pathological state associated with several circulatory diseases such as deep venous thrombosis, atherosclerosis, and diabetes mellitus. These pathologies inflict particular sites (inferior members for thrombosis, arterial bifurcations for atherosclerosis, and the foot and eye for diabetes). It would thus be of great interest to elucidate the role of flow-dependent rheological parameters, such as RBC aggregation, in the etiology of these pathologies *in vivo* and *in situ* with US techniques. To achieve this goal, the backscattering coefficient from blood was parametrized: two indices describing RBC aggregation, the packing factor and mean aggregate diameter, were extracted from the structure factor size estimator (SFSE).[7] The SFSE is a second-order data reduction model based on the structure factor and adapted to a dense medium such as blood. This approach is based on the analysis of the backscattered power spectrum that contains information about the size, spatial organization, concentration, and mechanical properties of scatterers (i.e., RBCs). The difficulty in using the SFSE *in vivo* is that the spectral content of backscattered echoes is also affected by attenuation caused by intervening tissue layers between the probe and the blood flow. More generally, US scatterer size estimation techniques for tissue characterization are facing similar challenges and several approaches to this problem have been developed.[8–23]

In vitro attenuation measurement methods are difficult to transfer *in vivo* since most of the clinical studies can be performed only in a single-transducer backscatter configuration. Indeed, *in vitro* backscatter coefficient measurements are generally compensated with an estimation of the attenuation using transmission mode[8,9] or reflection mode with a reflector on the opposite side of the examined tissue,[10,11] which is

[a]Present address: Laboratoire de Mécanique et d'Acoustique LMA, CNRS UPR 7051, 31 chemin Joseph Aiguier, 13402 Marseille Cedex 20, France.

difficult to implement in clinical practice. Many researchers have then attempted to improve the *in vivo* backscatter power spectrum estimates by assuming *a priori* attenuation values for the different intervening tissue layers.[12–14] Thicknesses of the intervening tissues are evaluated from the backscattered signals and typical attenuation values are assigned to each tissue, based on the results found in the relevant literature. Unfortunately, the tissue attenuation coefficients vary among patients (see, for example, the detailed review of the literature on tissue attenuation by Goss *et al.*[15]) and must be determined on a patient-specific basis. Several research groups have thus developed *in vivo* measurement techniques to evaluate the frequency-dependent attenuation for the compensation of the backscatter power spectrum.[16–19] Recently, Bigelow and co-workers[20–22] introduced a new algorithm that has the advantage to estimate simultaneously the effective radius of the tissue microstructure and the total attenuation. These two parameters were determined by using a single minimization method that fits the spectrum of the backscattered rf echoes from the region of interest (ROI) to an estimated spectrum by an appropriate model. In our previous work, this last strategy has been adapted for estimation of RBC scatterer sizes.[23] Blood structural parameters and total attenuation were determined simultaneously by using an optimization method, termed the structure factor size and attenuation estimator (SFSAE).

The goal of this paper was to further develop the SFSAE and assess its ability to evaluate three parameters (the packing factor, mean aggregate diameter, and total attenuation) with *in vitro* experiments mimicking *in vivo* conditions. To ensure that the global minimum of the optimization strategy is found, the algorithm was modified by employing an exhaustive search on the attenuation, combined with an analytical solution for the two blood structure parameters given for a fixed value of the attenuation, rather than a trust-region technique[24] used in our previous paper.[23] Porcine RBCs were first sheared in a Couette flow system, and ultrasonic rf echoes were obtained using a 25 MHz center-frequency transducer. Four skin-mimicking phantoms with different attenuation coefficients were introduced between the transducer and the blood flow. Other experiments were conducted in a tubular flow allowing to be closer to the *in vivo* situation.

The theoretical framework describing the SFSAE development is given in Sec. II. Section III explains the two experimental setups. Section IV presents results and compares the new SFSAE with the SFSE with compensation for attenuation. The validation of the method is finally discussed in Sec. V.

II. STRUCTURE FACTOR SIZE AND ATTENUATION ESTIMATOR (SFSAE)

Ultrasonic scattering from blood is mainly caused by the RBCs. Indeed, blood can be mechanically described as a colloidal suspension of RBCs in plasma. RBCs constitute the vast majority (97%) of the cellular content of blood and occupy a large volume fraction (hematocrit) of 35%–45% under normal conditions. These RBCs cannot be treated as independent scatterers since particle interactions (collision, attraction, deformation, and flow dependent motions) are strong. The theoretical model of US backscattering by blood that we developped[7] is based on the particle approach,[25,26] which consists of summing contributions from individual RBCs and modeling the RBC interaction by a particle pair-correlation function. Assuming that all the RBCs in the insonified blood are identical and using the Born approximation (weak scattering), the model predicts that the theoretical backscatter coefficient of blood is given by[23]

$$BSC_{theor}(k) = m\sigma_b(k)S(k)A(k), \tag{1}$$

where k is the wavenumber, m is the number density of RBCs in blood, σ_b is the backscattering cross section of a single RBC, S is the structure factor describing the spatial organization of RBCs, and A is the frequency-dependent attenuation function. The number density of RBCs m can be estimated by measuring the hematocrit H by microcentrifugation; m is then given by $m=H/V_s$, where V_s is the volume of a RBC (typically 87 μm^3). The backscattering cross-section σ_b of a weak scatterer small compared to the wavelength (Rayleigh scatterer) can be determined analytically as follows:

$$\sigma_b(k) = \frac{1}{4\pi^2}k^4V_s^2\gamma_z^2, \tag{2}$$

where $\gamma_z=(Z_{\mathrm{RBC}}-Z_{\mathrm{plasma}})/Z_{\mathrm{plasma}}$ is the fractional variation of impedance between the RBC and its suspending medium (i.e., the plasma). The structure factor S is by definition the Fourier transform of the pair-correlation function[26] g and can be approximated by its second-order Taylor expansion[7] in k as

$$S(k) = 1 + m\int (g(r)-1)e^{-2jkr}dr \approx W - \frac{12}{5}(kR)^2. \tag{3}$$

In this expression, $g(r)$ represents the probability of finding two particles separated by a distance r. W is the low-frequency limit of the structure factor $[\,S(k)|_{k\to 0}]$ called the packing factor.[26,27] R is the radius of 3D RBC aggregates assumed to be isotropic. We introduce $D=R/a$ as the isotropic diameter of an aggregate (expressed in number of RBCs) with a the radius of one RBC sphere-shaped model of volume V_s. The attenuation function A is given by

$$A(k) = e^{-4\alpha_0 k/8.68(c/2\pi)} = e^{-4\alpha_0 f/8.68}, \tag{4}$$

where c is the mean speed of sound in the intervening tissue layers, f is the frequency in MHz, and α_0 is the attenuation coefficient (in dB/MHz) defined by $\alpha_0=\Sigma_i\alpha_i e_i$, where α_i and e_i are, respectively, the intervening tissue layer attenuations (in dB/cm MHz) and thicknesses. One can note in Eq. (4) the coefficient 8.68 that expresses unit conversion from decibels to neper: $\alpha_0[\mathrm{Np/MHz}]=\alpha_0[\mathrm{dB/MHz}]/8.68$. According to the above equation, we thus assume, as a first approximation, that the attenuation increases linearly with the frequency: $\alpha(f)=\alpha_0 f/8.68$. Altogether, we obtain the following expression for the theoretical backscatter coefficient from blood:

$$BSC_{theor}(k) = \frac{1}{4\pi^2} mk^4 V_s^2 \gamma_z^2 \left(W - \frac{12}{5}(ka)^2 D^2 \right) e^{-4\alpha_0 k/8.68(c/2\pi)}. \tag{5}$$

In particular, for a given value of the wavenumber k (equivalently, a given value of the frequency f) and a given value of the attenuation coefficient α_0, $BSC_{theor}(k)$ is a linear function of the variables W and D^2.

The measured backscatter coefficient reported in this study was computed as

$$BSC_{meas}(k) = BSC_{ref}(k) \frac{\overline{P_{\text{meas}}(k)}}{\overline{P_{\text{ref}}(k)}}. \tag{6}$$

In Eq. (6), the mean backscattered power spectrum $\overline{P_{\text{meas}}}$ was obtained by averaging the power spectra of 20 backscattered echoes from blood. The mean power spectrum $\overline{P_{\text{ref}}}$ was obtained from a reference sample of non-aggregated RBCs at a low hematocrit of 6% (i.e., Rayleigh scatterers).[28] In this case, 20 echoes were also averaged. The backscatter coefficient of this reference sample BSC_{ref} was estimated by using the expression of the Perkus–Yevick packing factor for spheres W_{PYs} (that is a function of the hematocrit H) as follows:[26,27]

$$BSC_{ref}(k) = m\sigma_b(k) W_{\text{PYs}} = m\sigma_b(k) \frac{(1-H)^4}{(1+2H)^2}. \tag{7}$$

This reference sample was used to compensate the backscattered power spectrum $\overline{P_{\text{meas}}}$ for the electromechanical system response, and the depth-dependent diffraction and focusing effects caused by the US beam.

The packing factor W, aggregate diameter D, and total attenuation along the propagation path α_0 were determined by matching the measured BSC_{meas} given by Eq. (6) with the theoretical BSC_{theor} given by Eq. (5). For this purpose, we searched values of $(W, D^2, \alpha_0) \in (0,\infty) \times (0,\infty) \times [0,1]$, W and D being variables without dimension and α_0 expressed in dB/MHz, minimizing the cost function which synthesizes all of the wavenumbers k_i $(i=1 \cdots N)$ within the −20 dB bandwidth of $\overline{P_{\text{meas}}}$:

$$\begin{aligned} F(W, D^2, \alpha_0) &= \sum_i \| BSC_{meas}(k_i) - BSC_{theor}(k_i) \|^2, \\ &= \sum_i \left\| BSC_{meas}(k_i) - \frac{1}{4\pi^2} mk_i^4 V_s^2 \right. \\ &\quad \left. \times \gamma_z^2 \left(W - \frac{12}{5}(k_i a)^2 D^2 \right) e^{-4\alpha_0 k/8.68(c/2\pi)} \right\|^2. \end{aligned} \tag{8}$$

The cost function had in the majority of cases reported in this study one minimum, as was observed by plotting the cost function surface $F(W, D^2, \alpha_0)$ with varying values of α_0. An example is given in Fig. 1. That is why a simple minimization routine *lsqnonlin* in MATLAB (The MathWorks, Inc., Natick, MA), i.e., a trust-region method based on the interior-reflective Newton method,[24] was first employed in our previous paper.[23] Nevertheless, on a possibility of the appearance of several minima, an exhaustive search on the value of α_0 was employed in this paper in order to ensure

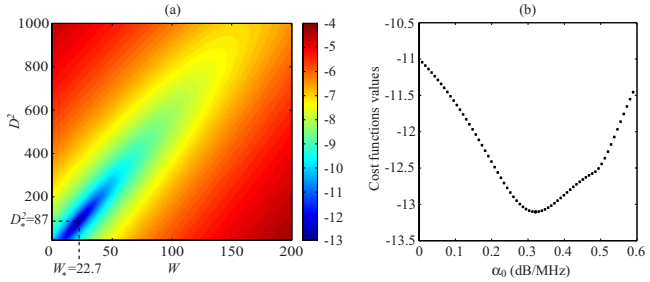


FIG. 1. (Color online) (a) Typical aspect of the logarithm of the cost function $F(W, D^2, \alpha_0)$ for a fixed value of α_0. The logarithm is shown here in order to enhance the visual contrast. This cost function has one minimum denoted (W_*, D^2_*) that depends on α_0. (b) Typical aspect of the function $\log(F(W_*, D^2_*, \alpha_0))$ for varying values of α_0 (W_* and D^2_* being calculated for each α_0). This cost function has a single minimum.

that the global minimum is found. If the value of α_0 is fixed, a quadratic function of the two variables W and D^2 is obtained, as can be seen from Eq. (8), and by calculating its critical points, a system of two linear equations of two unknowns can be easily solved. One can show using the Cauchy–Schwartz' inequality that the determinant of this system never vanishes (unless all frequencies are equal) and thus that the solution is unique. If the unique solution of this system denoted $(W_*(\alpha_0), D^2_*(\alpha_0))$ is such that $W_*(\alpha_0) \leq 0$ or $D^2_*(\alpha_0) \leq 0$, then the cost function $F(W, D^2, \alpha_0)$ has no global minimum on $(0,\infty) \times (0,\infty)$ (with the value of α_0 fixed). Also, this case means that the value of α_0 has to be rejected, since it is meaningless. Otherwise, the global minimum of $F(W, D^2, \alpha_0)$ on $(0,\infty) \times (0,\infty)$ occurs at $(W_*(\alpha_0), D^2_*(\alpha_0))$. In practice, there is a maximal value α_* of α_0 between 0 and 1 dB/MHz, for which $W_*(\alpha_0) \geq 0$ and $D^2_*(\alpha_0) \geq 0$. Thus, the global minimum of $F(W, D^2, \alpha_0)$ on the domain $(0,\infty) \times (0,\infty) \times [0, \alpha_*]$ is the global minimum of the function $F(W_*(\alpha_0), D^2_*(\alpha_0), \alpha_0)$ on the interval $[0, \alpha_*]$. The optimal value of α_0 can thus be obtained upon sweeping the interval $[0, \alpha_*]$. In our tests, we started the exhaustive search with a step of 10^{-5} dB/MHz. Then, the step was iteratively decreased by a factor of 10^{-1} dB/MHz, while the search is performed on the interval of length twice the preceeding step around the best value obtained so far. The iterative (multiresolution) exhaustive search was stopped after a step of 10^{-20} dB/MHz. At the end of the SFSAE algorithm, a simple test allowed to verify the validity of the solution. If the solution (W, D^2, α_0) of the optimization problem was such that W or D^2 was on the inferior boundary of the domain (i.e., W or D^2 are equal to zero), the solution was rejected since such values of W or D^2 are unrealistic. This last point will be discussed later in Sec. V D.

III. METHODS

A. Blood sample preparation

Blood from two different pigs were used for the two experiments in a Couette device and in a tube. This fresh porcine whole blood was obtained from a local slaughter house and anticoagulated with 3 g/l of ethylene diamine tetra acetic acid (EDTA). Then the whole blood was centri-

fuged and the plasma and buffy coat were removed. Two blood samples were then prepared for each experiment: (i) a H6 reference sample, which was a 6% hematocrit non-aggregating RBCs resuspended in physiological saline solution, and (ii) a 40% hematocrit T40 test sample, which consisted of RBCs resuspended in plasma to promote aggregation.

B. *In vitro* experiments

1. In vitro experiments in a Couette flow system (Couette device)

US measurements were first performed in a Couette device to produce a linear blood velocity gradient at a given shear rate (see Fig. 1 in Ref. 29). The system consists of a rotating inner cylinder with a diameter of 160 mm surrounded by a fixed concentric cylinder of diameter 164 mm. A 60 ml blood sample was sheared in the 2 mm annular space between both coaxial cylinders. An US scanner (Vevo 660, Visualsonics, Toronto, Canada) equipped with the RMV 710 probe was used in B-mode. The oscillating single-element focused circular transducer had a center frequency of 25 MHz, a diameter of 7.1 mm, and a focal depth of 15 mm. We acquired RF data from this scanner at a sampling frequency of 250 MHz with 8 bit resolution (Gagescope, model 8500CS, Montreal, Canada). The probe was mounted in the side wall of the fixed outer cylinder and was positioned to have its focal zone at the center of the gap between both cylinders. To ensure ultrasonic coupling, the hole within the outer stationary cylinder (containing the probe) was filled with a liquid agar gel based mixture. When solidified, this gel was cut to match the curvature of the cylinder to avoid any flow disturbance. The gel was a mixture of distilled water, 3% (w/w) agar powder (A9799, Sigma Chemical, Saint-Louis, MO), 10% (w/w) glycerol, and a specific concentration of 50 μm cellulose scattering particles (S5504 Sigmacell, Sigma Chemical, Saint-Louis, MO) that determined the attenuation coefficient. Five experiments were performed with five mixtures having SigmaCell (SC) concentrations varying from 0% to 1% (w/w). The 0% concentration constituted the non-attenuating gel and the four other mixtures mimicked skin attenuations.

All US measurements were made at room temperature. Prior to each measurement, the T40 blood was sheared at 200 s^{-1} during 30 s to disrupt RBC aggregates. The shear rate was then reduced to residual values of 5, 10, 20, 30, and 50 s^{-1} for 90 s until an equilibrium in the state of aggregation was reached. For each shear rate, 20 B-mode images were constructed from acquired rf echoes each 4 s for a total period of analysis of 80 s. Each image contained 384 vertical lines. For 180 vertical lines at the center of the B-mode images, echoes were selected with a rectangular window of axial length 0.4 mm at 20 depths every 0.031 mm (i.e., with 92% overlap between windows). For each depth, the power spectra of the backscattered rf echoes were averaged over 20 acquisitions (corresponding to the 20 acquired B-mode images) to provide $\overline{P_{\text{meas}}}$. This protocol was repeated five times with the same blood and with each of the five agar-based phantoms.

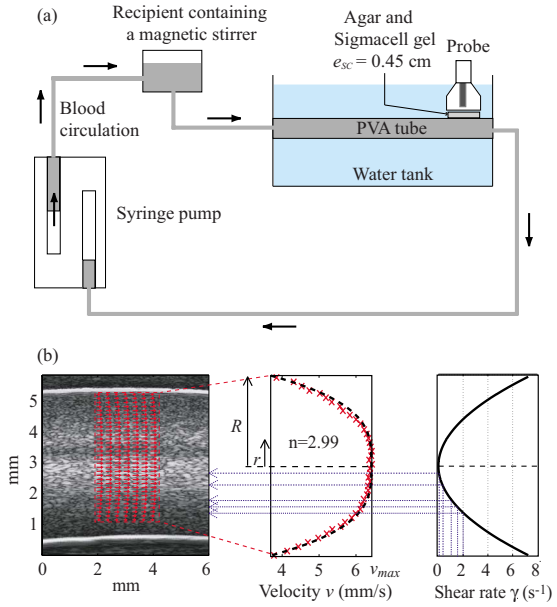


FIG. 2. (Color online) (a) Diagram of the experimental apparatus used in the tubular flow experiments. (b) Echographic image of the T40 blood in the tube (obtained with the US scanner Visualsonics) and estimates of particle displacement fields by B-mode speckle tracking and corresponding mean velocity and shear rate as a function of depths in the tube.

Then, the T40 blood was removed and the H6 sample was introduced in the Couette device. The H6 sample was sheared at 50 s^{-1} and coupled with the 0% SC concentration agar gel. Echoes were windowed as for the H40 sample at the same depths and their power spectra were averaged over 20 acquisitions to obtain $\overline{P_{\text{ref}}}$. This reference power spectrum allowed to normalize the average power spectrum $\overline{P_{\text{meas}}}$ [as described in Eq. (6), Sec. II].

2. In vitro experiments in a tube

In order to be closer to *in vivo* conditions, US measurements were also performed in a tubular flow device. The experimental test section was an horizontal tube made of polyvinyl alcohol (PVA) cryogel. This tube had an internal diameter of 4.9 mm and a length of 10 cm, and was immersed in a tank filled with degassed water at room temperature. The US measurements were made at a position 7 cm from the tube entrance with the Vevo 660 US scanner. The transducer focal point at 15 mm was positioned at the middle of the tube, imaged in the longitudinal plane. The T40 blood sample was circulated in the PVA tube using a double syringe pump (Harvard PHD 2000) at a constant flow rate from the outflow syringe, to the tube and the second inflow syringe. In order to ensure that the RBCs are disaggregated at the tube entrance, the T40 sample blood passed through a small recipient containing a magnetic stirrer just before entering the PVA tube, as indicated in Fig. 2(a). Similarly to the Couette experiments (Sec. III B 1), a layer of gel was placed between the probe and the tubular flow. Five experiments were successively performed on the same blood with five mixtures having SC concentrations: 0%, 1%, 1.5%, 2%, and 2.5% (w/w).

The flow rate was chosen equal to 5 ml/min to promote

TABLE I. Values of the average attenuations $\overline{\alpha_{SC}}$ of skin-mimicking phantoms expressed in dB/cm MHz and corresponding attenuation coefficients $\overline{\alpha_{SC}}e_{SC}$ expressed in dB/MHz for the Couette and tube experiments.

Sigmacell (SC) concentration (%)	Attenuation $\overline{\alpha_{SC}}$ (dB/cm/MHz)	Attenuation coefficient for the Couette flow device $\overline{\alpha_{SC}}e_{SC}$ (e_{SC}=1.3 cm) (dB/MHz)	Attenuation coefficient for the tube flow device $\overline{\alpha_{SC}}e_{SC}$ (e_{SC}=0.45 cm) (dB/MHz)
0	0.0054 ± 0.0015	0.0070 ± 0.0019	0.0024 ± 0.0006
0.25	0.088 ± 0.018	0.115 ± 0.024	⋯
0.5	0.168 ± 0.023	0.219 ± 0.030	⋯
0.75	0.246 ± 0.027	0.320 ± 0.035	⋯
1	0.316 ± 0.031	0.411 ± 0.040	0.142 ± 0.014
1.5	0.475 ± 0.053	⋯	0.214 ± 0.024
2	0.633 ± 0.051	⋯	0.285 ± 0.023
2.5	0.768 ± 0.060	⋯	0.346 ± 0.027

RBC aggregation and when 30 ml of blood had circulated in the tube, 20 B-mode images were constructed for 80 s. The data were processed as described for the *in vitro* experiment in the Couette device, except that echoes were selected with a rectangular window of length 0.4 mm at 65 depths every 0.031 mm (again with 92% overlaps). A reference measurement was done by using the H6 sample circulating in the tube at a flow rate of 5 ml/min. Echoes were windowed as for the H40 sample at the same 65 depths and their power spectra were averaged over 20 acquisitions to obtain $\overline{P_{ref}}$. One can note here that $\overline{P_{ref}}$ allow to compensate $\overline{P_{meas}}$ not only for the electromechanical system response, and the depth-dependent diffraction and focusing effects caused by the US beam, but also for the US response of the PVA cryogel tube.

3. Attenuation measurements for comparison with the SFSAE attenuation estimates

The attenuation coefficients of the reference (0% SC) and of the skin-mimicking phantoms were determined by using a standard substitution method. The Vevo 660 US scanner equipped with the same RMV 710 25 MHz transducer was used in M-mode and in transmission/reception with a reflector on the opposite side of the phantom for reflection measurements. Reflected signals were recorded both with and without the agar gel sample in the acoustic path. The attenuation coefficient was then estimated using a log spectral difference technique.[30] For a given concentration of SC, measurements were obtained from two different sample thicknesses, and for each, six regions were scanned for averaging purpose. In our experimental devices, thicknesses of skin-mimicking phantoms e_{SC} were fixed to 1.3 cm in the case of the Couette device and 0.45 cm for the tubular flow experiments. Values of the average attenuations $\overline{\alpha_{SC}}$ expressed in dB/cm MHz and of the attenuation coefficients $\overline{\alpha_{SC}}e_{SC}$ expressed in dB/MHz for the Couette and tube experiments are reported in Table I. As it can be observed in this table, attenuation coefficients of skin-mimicking phantoms were in the same range as the human dermis (which is 0.21 dB/MHz at 14–50 MHz considering a 1 mm dermis thickness[31]).

For the Couette device configuration, blood attenuation was also measured at different shear rates in the reflection mode. The gel had a 0% SC concentration, the probe was used in M-mode and the rotating inner cylinder was used as the reflector. For each shear rate, acquisitions of 20 rf lines were performed both with blood and with water in the Couette device. The water acquisition was used for normalization. For each acquisition, blood attenuation was estimated using the log spectral difference technique. These values were then averaged to provide $\overline{\alpha_{blood}}$. Values obtained were 0.053 ± 0.011, 0.036 ± 0.008, 0.024 ± 0.005, 0.016 ± 0.003, and 0.015 ± 0.003 dB/mm MHz for shear rates of 5, 10, 20, 30, and 50 s^{-1}, respectively.

4. Reference measurements for blood structural parameters with the 0% SC concentration phantom for in vitro experiments

For both *in vitro* devices, experiments with the 0% SC phantom were realized in order to have reference results on packing factors W_{ref} and aggregate diameters D_{ref}. For these reference measurements, the measured backscatter coefficient was compensated for predetermined values of blood and 0% SC phantom attenuations and thus computed as follows:

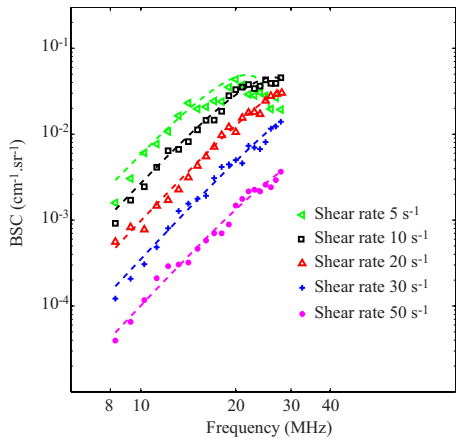


FIG. 3. (Color online) Backscatter coefficients for blood sheared at different residual shear rates and measured with the 0% SC concentration phantom, and corresponding fitting with the classical SFSE with compensation for blood attenuation (in dashed lines).

TABLE II. Reference values of the packing factor $\overline{W_{\text{ref}}}$ and diameter $\overline{D_{\text{ref}}}$ obtained with the SFSE at different residual shear rates. The parameter r^2 represents the correlation coefficient of the fitted SFSE model on experimental data BSC_{meas}.

Shear rate (s^{-1})	$\overline{W_{\text{ref}}}$	$\overline{D_{\text{ref}}}$	r^2
5	24.56 ± 2.82	10.11 ± 0.72	0.67 ± 0.02
10	9.14 ± 1.89	4.59 ± 0.78	0.95 ± 0.02
20	3.90 ± 0.79	2.68 ± 0.54	0.93 ± 0.03
30	1.39 ± 0.34	1.35 ± 0.38	0.95 ± 0.02
50	0.44 ± 0.040	0.82 ± 0.23	0.95 ± 0.03

$$BSC_{meas}(k) = BSC_{ref}(k)\frac{\overline{P_{\text{meas}}(k)}}{\overline{P_{\text{ref}}(k)}} \times e^{-4/8.68(\alpha_{\text{blood}}e_{\text{blood}}+\alpha_{\text{SC}}e_{\text{SC}})k(c/2\pi)}, \quad (9)$$

and the parameters W_{ref} and D_{ref} were deduced by fitting the measured backscatter coefficient to the theoretical backscatter coefficient using the classical SFSE:[7]

$$BSC_{theor}(k) = \frac{1}{4\pi^2}mk^4V_s^2\gamma_z^2\left(W_{\text{ref}} - \frac{12}{5}(ka)^2D_{\text{ref}}^2\right). \quad (10)$$

The parameters W_{ref} and D_{ref} were assumed to be true values of packing factors and aggregate diameters at all shear rates, and will be compared in Sec. IV with packing factors and diameters estimated by the SFSAE when skin-mimicking phantoms were used.

IV. RESULTS

A. *In vitro* experiments in the Couette device

Figure 3 shows the typical BSC_{meas} as a function of frequency for different residual shear rates in the case of 0% SC (i.e., reference measurements). Also represented are corresponding fitted curves obtained with the SFSE after compensation for predetermined values of blood attenuation. The parameters W_{ref} and D_{ref} were estimated over 20 depths and 180 rf lines of the B-mode image (i.e., 3600 estimations) in a chosen ROI. Their mean values $\overline{W_{\text{ref}}}$, $\overline{D_{\text{ref}}}$ as well as their standard deviations $\sigma_{W_{\text{ref}}}$, $\sigma_{D_{\text{ref}}}$ are reported in Table II. It can be noticed that the BSC_{meas} amplitude as well as the estimation of the parameters $\overline{W_{\text{ref}}}$ and $\overline{D_{\text{ref}}}$ decrease when the shear rate increases (i.e. when the level of aggregation becomes smaller). Also given in Table II is the correlation coefficient r^2 in order to assess the goodness of fit between the model and the measured data. It reveals that the highest level of aggregation (at 5 s^{-1}) had the worst fit, i.e., $r^2=0.67$.

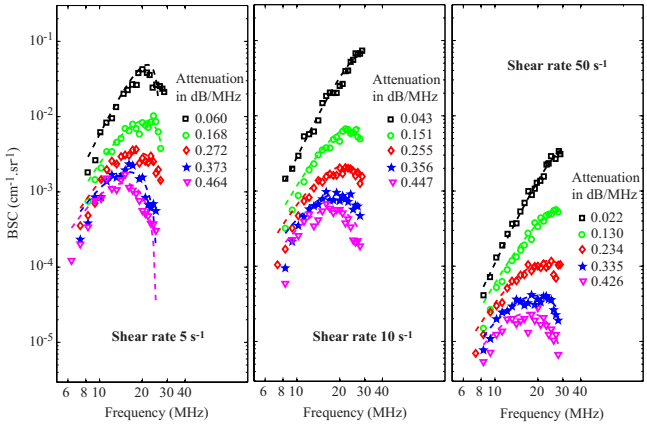


FIG. 4. (Color online) Measured backscatter coefficients for blood sheared at 5, 10, and 50 s^{-1} and with each of the five phantoms. The corresponding fitted models (in dashed lines) are the SFSE for the 0% SC phantom with compensation for blood attenuation (black squares) and the SFSAE for the four skin-mimicking phantoms (0.25%, 0.5%, 0.75%, and 1% SC).

Typical results of the SFSAE minimization procedure for the different agar phantoms at shear rates of 5, 10, and 50 s^{-1} are given in Fig. 4. The attenuations given in Fig. 4 correspond to the reference total attenuations $\overline{\alpha_{\text{ref}}}$ i.e., the sum of $\overline{\alpha_{\text{SC}}}$ and $\overline{\alpha_{\text{blood}}}$ estimated in the reflection mode as shown in Sec. III B 3. For all shear rates, an increase in the total attenuation has the effect of decreasing the amplitude of the BSC_{meas} at all frequencies and also has an effect of moving the peak of the BSC_{meas} at lower frequencies. It is very interesting to observe how attenuation influences the well-known quasi-Rayleigh backscattering behavior of disaggregated RBCs (i.e., at a shear rate of 50 s^{-1}). Indeed, for the 50 s^{-1} shear rate combined with the 0% SC concentration (reference measurement), the frequency dependence of the BSC_{meas} is close to f^4 and is thus quasi-linear on our log compressed scale, as expected. In the presence of attenuation, the frequency dependencies of the BSC_{meas} are drastically modified.

TABLE III. Reference values of the packing factor $\overline{W_{\text{ref}}}$ and diameter $\overline{D_{\text{ref}}}$ obtained with the SFSE (0% SC), and values of the packing factor $\bar{W}$, diameter $\bar{D}$, and attenuation $\overline{\alpha_0}$ obtained with the SFSAE (0.25%, 0.5%, 0.75%, and 1% SC) at the shear rate 5 s^{-1} in the Couette flow device. The parameter ϵ indicates the relative error and r^2 the correlation coefficient of the fitted SFSAE model on experimental data BSC_{meas}.

SC (%)	0	0.25	0.5	0.75	1
$\overline{W_{\text{ref}}}$	24.56 ± 2.82	⋯	⋯	⋯	⋯
$\bar{W}$	⋯	36.97 ± 12.40	23.33 ± 8.96	26.83 ± 8.65	23.01 ± 6.83
ϵ_W (%)	⋯	50.53	−0.65	9.24	−6.31
$\overline{D_{\text{ref}}}$	10.11 ± 0.72	⋯	⋯	⋯	⋯
$\bar{D}$	⋯	10.87 ± 1.74	9.05 ± 1.61	11.11 ± 1.68	11.08 ± 1.68
ϵ_D (%)	⋯	7.52	−10.48	9.89	9.59
$\overline{\alpha_{\text{ref}}}$ (dB/MHz)	0.060 ± 0.018	0.168 ± 0.035	0.272 ± 0.041	0.373 ± 0.046	0.464 ± 0.051
$\overline{\alpha_0}$ (dB/MHz)	⋯	0.246 ± 0.045	0.314 ± 0.046	0.409 ± 0.048	0.445 ± 0.051
ϵ_{α_0} (%)	⋯	46.43	15.44	9.65	−4.09
r^2	0.67 ± 0.02	0.81 ± 0.06	0.79 ± 0.05	0.79 ± 0.07	0.74 ± 0.05

TABLE IV. Reference values of the packing factor $\overline{W_{\text{ref}}}$ and diameter $\overline{D_{\text{ref}}}$ obtained with the SFSE (0% SC), and values of the packing factor $\bar{W}$, diameter $\bar{D}$, and attenuation $\overline{\alpha_0}$ obtained with the SFSAE (0.25%, 0.5%, 0.75%, and 1% SC) at the shear rate 10 s^{-1} in the Couette flow device. The parameter ϵ indicates the relative error and r^2 the correlation coefficient of the fitted SFSAE model on experimental data BSC_{meas}.

SC (%)	0	0.25	0.5	0.75	1
$\overline{W_{\text{ref}}}$	9.14 ± 1.89	⋯	⋯	⋯	⋯
$\bar{W}$	⋯	10.14 ± 3.03	9.35 ± 2.73	9.08 ± 1.93	10.86 ± 3.88
ϵ_W (%)	⋯	10.94	2.29	−0.65	18.81
$\overline{D_{\text{ref}}}$	4.59 ± 0.78	⋯	⋯	⋯	⋯
$\bar{D}$	⋯	5.25 ± 1.01	4.83 ± 0.81	4.84 ± 0.55	5.07 ± 1.26
ϵ_D (%)	⋯	14.38	5.23	5.45	10.45
$\overline{\alpha_{\text{ref}}}$ (dB/MHz)	0.043 ± 0.010	0.151 ± 0.032	0.255 ± 0.038	0.356 ± 0.043	0.447 ± 0.048
$\overline{\alpha_0}$ (dB/MHz)	⋯	0.172 ± 0.061	0.283 ± 0.053	0.367 ± 0.033	0.410 ± 0.057
ϵ_{α_0} (%)	⋯	13.91	10.98	3.09	−8.28
r^2	0.95 ± 0.02	0.86 ± 0.06	0.84 ± 0.08	0.86 ± 0.08	0.81 ± 0.06

The parameters W, D, and α_0 were estimated over 20 depths and 180 rf lines of the B-mode image in the same ROI used for the reference measurements. Their mean values $\bar{W}$, $\bar{D}$, $\overline{\alpha_0}$ and corresponding r^2 are reported in Tables III, IV and V for shear rates of 5, 10, and 50 s^{-1}, respectively. The values of $\bar{W}$ and $\bar{D}$ from the SFSAE are quite similar to the reference values of $\overline{W_{\text{ref}}}$ and $\overline{D_{\text{ref}}}$, as well as the total attenuation $\overline{\alpha_0}$ from the SFSAE are similar to the reference total attenuation $\overline{\alpha_{\text{ref}}}$, except for the shear rate of 5 s^{-1} and the 0.25% SC skin-mimicking phantom. In these tables, the relative errors for each parameter ϵ_W, ϵ_D, and ϵ_{α_0} correspond, respectively, to $(\bar{W}-\overline{W_{\text{ref}}})/\overline{W_{\text{ref}}}$, $(\bar{D}-\overline{D_{\text{ref}}})/\overline{D_{\text{ref}}}$, and $(\overline{\alpha_0}-\overline{\alpha_{\text{ref}}})/\overline{\alpha_{\text{ref}}}$.

Figure 5 shows quantitative images superimposed on the gray-scale B-mode images of the blood sheared at 10 s^{-1} in the Couette device. The color of each pixel was assigned based on the value of the parameters estimated by the SFSE with the 0% SC phantom and by the SFSAE with the four skin-mimicking phantoms. The color bars relate the pixels in the quantitative images to the estimated parameters (W, D, and α_0) in the chosen ROI. The black pixels in the ROI correspond to rejected solutions of the optimization method (when the estimated packing factor W or diameter D^2 was found equal to zero, which is unrealistic). A clear distinction between images of the estimated total attenuation with the four skin-mimicking phantoms is apparent, which was expected, whereas images of the estimated parameters W and D are quite similar, which was also wished.

For each residual shear rate, quantitative images for the three parameters W, D, and α_0 were constructed. The average estimates are summarized in Fig. 6. The lower and upper error bars for relative errors correspond to

$$\epsilon_{\text{lower}} = \frac{(\bar{X}-\sigma_X)-(\overline{X_{\text{ref}}}-\sigma_{X_{\text{ref}}})}{\overline{X_{\text{ref}}}-\sigma_{X_{\text{ref}}}},$$

$$\epsilon_{\text{upper}} = \frac{(\bar{X}+\sigma_X)-(\overline{X_{\text{ref}}}+\sigma_{X_{\text{ref}}})}{\bar{X}+\sigma_{X_{\text{ref}}}}, \tag{11}$$

where X represents one of the estimated parameters W, D, or α_0. For all shear rates and all skin-mimicking phantoms, the SFSAE gave quantitatively satisfactory estimates of $\bar{W}$ and $\bar{D}$ with relative absolute errors below 25%, except for the shear rate of 5 s^{-1} and the 0.25% SC skin-mimicking phantom.

The use of the SFSAE when there is little attenuation (i.e., the 0% SC phantom) was also investigated. Figure 7 presents the averaged parameters $\bar{W}$, $\bar{D}$, and $\overline{\alpha_0}$ obtained from the SFSAE with the 0% SC phantom. In this experiment, the total attenuation came mainly from the blood

TABLE V. Reference values of the packing factor $\overline{W_{\text{ref}}}$ and diameter $\overline{D_{\text{ref}}}$ obtained with the SFSE (0% SC), and values of the packing factor $\bar{W}$, diameter $\bar{D}$, and attenuation $\overline{\alpha_0}$ obtained with the SFSAE (0.25%, 0.5%, 0.75%, and 1% SC) at the shear rate 50 s^{-1} in the Couette flow device. The parameter ϵ indicates the relative error and r^2 the correlation coefficient of the fitted SFSAE model on experimental data BSC_{meas}.

SC (%)	0	0.25	0.5	0.75	1
$\overline{W_{\text{ref}}}$	0.44 ± 0.10	⋯	⋯	⋯	⋯
$\bar{W}$	⋯	0.46 ± 0.19	0.53 ± 0.21	0.38 ± 0.15	0.35 ± 0.11
ϵ_W (%)	⋯	4.54	20.45	−13.64	−20.45
$\overline{D_{\text{ref}}}$	0.82 ± 0.23	⋯	⋯	⋯	⋯
$\bar{D}$	⋯	0.89 ± 0.24	0.98 ± 0.28	0.90 ± 0.21	1.02 ± 0.19
ϵ_D (%)	⋯	8.54	19.51	9.76	24.39
$\overline{\alpha_{\text{ref}}}$ (dB/MHz)	0.022 ± 0.005	0.130 ± 0.027	0.234 ± 0.033	0.335 ± 0.038	0.426 ± 0.043
$\overline{\alpha_0}$ (dB/MHz)	⋯	0.148 ± 0.059	0.287 ± 0.060	0.370 ± 0.061	0.409 ± 0.063
ϵ_{α_0} (%)	⋯	13.85	22.65	10.45	−3.99
r^2	0.95 ± 0.03	0.91 ± 0.04	0.90 ± 0.06	0.86 ± 0.05	0.86 ± 0.06

$(0.015 \leq \alpha_{blood} \leq 0.053$ dB/MHz) and a little part from the SC phantom (0.007 dB/MHz). Also represented in Fig. 7 are the SFSAE results with the 0.25% SC phantom in order to compare the SFSAE performance in the presence of higher attenuation. When there is little attenuation, estimates of $\bar{W}$ were less accurate for shear rates of 5, 10, and 20 s^{-1} (in comparison with $\bar{W}$ obtained with the four other SC phantoms) with relative errors between 25% and 42%. Estimates of $\overline{\alpha_0}$ were between 0.041 and 0.133 dB/MHz, so the total attenuation were overestimated giving the largest relative errors between 124% and 211%.

Finally, to support the discussion reported later, the packing factor W_{comp} and the diameter of the aggregates D_{comp} were also evaluated by compensating the BSC_{meas} in the SFSE with the predetermined values measured in reflection mode (Sec. III B 3). Values of W_{comp} and D_{comp} were averaged over 20 depths and 180 lines of the B-mode image in the same ROI used for the reference measurements to provide $\overline{W_{comp}}$ and $\overline{D_{comp}}$. Results are presented in Fig. 8. The relative errors are below 25% for all shear rates and all skin-mimicking phantoms.

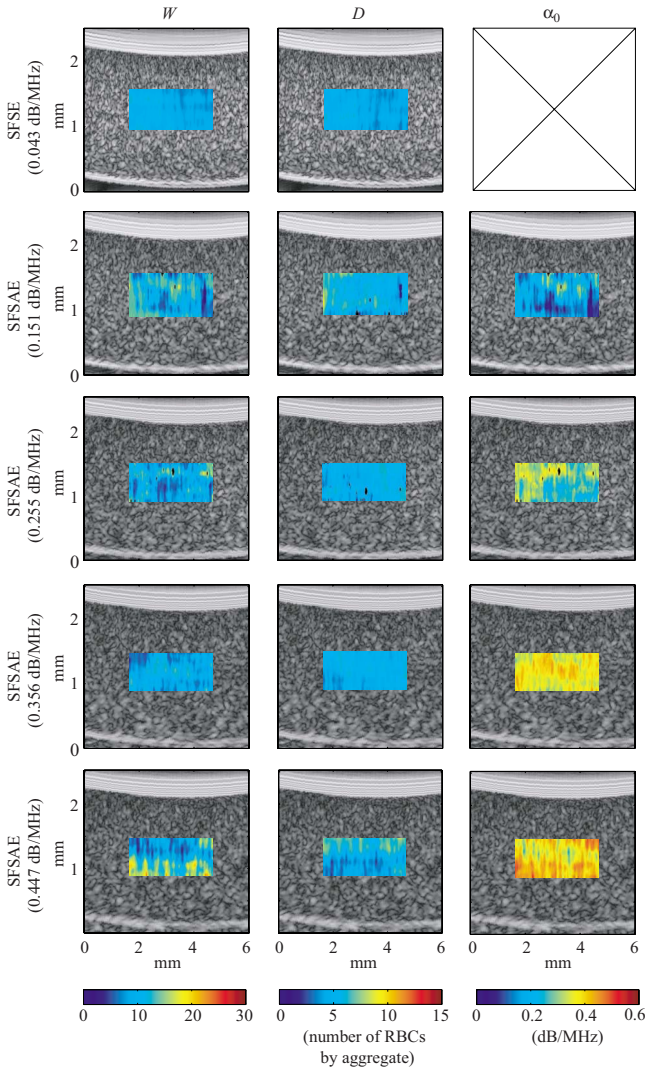


FIG. 5. (Color online) Quantitative images of blood sheared at 10 s^{-1} in the Couette device superimposed on the gray-scale B-mode images. Parameters were estimated by the classical SFSE with compensation for blood attenuation for the 0% SC concentration (top panel) and by the SFSAE for the four skin-mimicking phantoms.

B. *In vitro* experiments in the tubular flow device

Figure 9 shows quantitative images superimposed on the gray-scale B-mode images of tubular blood flow for each skin-mimicking phantom. As observed in the Couette experiment, a clear distinction between images of the estimated total attenuation is apparent, whereas images of the estimated parameters W and D are quite similar.

In order to know the shear rate profile in the tube, a speckle tracking method[32] was used to obtain the displacement fields of RBCs, as shown in Fig. 2(b). The displacement of the speckle pattern between each frame is then related to the velocity through the time elapsed between the images. The mean velocity profile across the tube was fitted to the following simple power-law model:

$$v(r) = v_{max}(1 - (r/R)^n), \tag{12}$$

where $v(r)$ is the mean velocity at radius r, R is the radius of the tube, v_{max} is the maximum velocity, and n is the power-law exponent ($n=2$ for parabolic flow). As expected, the velocity profile was slightly blunted due to RBC aggregation with $n=2.99$ [see Fig. 2(b)]. The shear rate profile, $\gamma(r)$, was deducted from this velocity profile by computing the derivative of $v(r)$:

$$\gamma(r) = -\frac{\partial v(r)}{\partial r} = n v_{max} \frac{r^{n-1}}{R^n}. \tag{13}$$

Five shear rates were selected on the shear rate profile (i.e., 0.1, 0.3, 1, 1.5, and 2 s^{-1}) corresponding to specific depths [see Fig. 2(b)] and thus corresponding to specific blood thicknesses (i.e., 2.7, 2.9, 3.4, 3.6, and 3.8 mm). For each

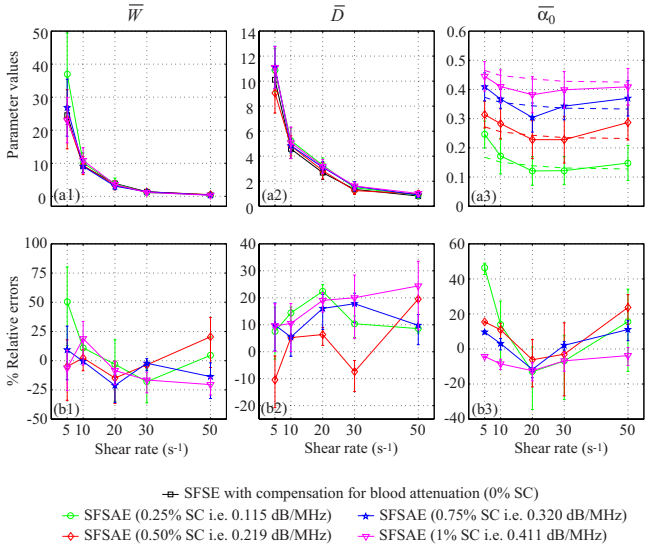


FIG. 6. (Color online) (a) Values of $\bar{W}$, $\bar{D}$, and $\overline{\alpha_0}$ (in dB/MHz) for different residual shear rates estimated by the SFSAE for the four skin-mimicking phantoms in the Couette experiments (mean ± standard deviation). Also represented are values of $\overline{W_{ref}}$ and $\overline{D_{ref}}$ estimated by the classical SFSE with compensation for blood attenuation for the 0% SC concentration and values of $\overline{\alpha_{ref}}$ in dashed lines. (b) Corresponding relative errors of $\bar{W}$, $\bar{D}$, and $\overline{\alpha_0}$.

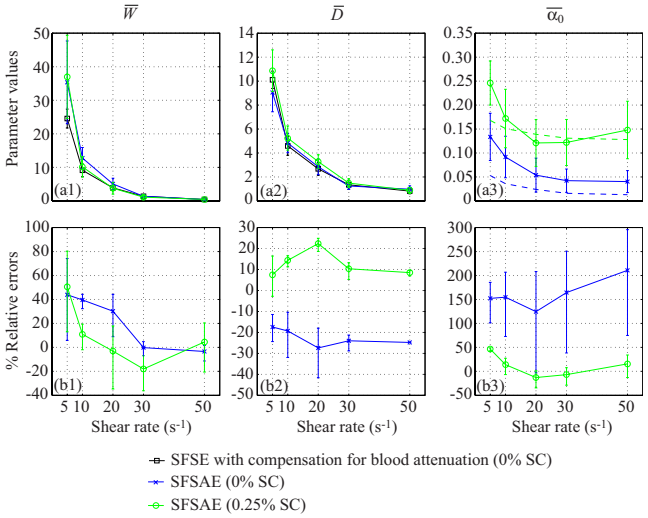


FIG. 7. (Color online) (a) Values of $\overline{W}$, $\overline{D}$, and $\overline{\alpha_0}$ (in dB/MHz) for different residual shear rates estimated by the SFSAE for the 0% and 0.25% SC concentrations. Also represented are values of $\overline{W_{\text{ref}}}$ and $\overline{D_{\text{ref}}}$ estimated by the classical SFSE with compensation for blood attenuation for the 0% SC concentration and values of $\overline{\alpha_{\text{ref}}}$ in dashed lines. (b) Corresponding relative errors of $\overline{W}$, $\overline{D}$, and $\overline{\alpha_0}$.

shear rate, values of W, D, and α_0 were simultaneously estimated by the SFSAE and averaged over three depths (one specific depth and their two adjacent depths, i.e., ± 0.031 mm) and 180 lines of the B-mode image. These average values of $\overline{W}$ and $\overline{D}$ obtained with the SFSAE were compared with average values of $\overline{W_{\text{ref}}}$ and $\overline{D_{\text{ref}}}$ evaluated by the SFSE with attenuation-compensation. The reference total attenuation $\overline{\alpha_{\text{ref}}}$ was the sum of $\overline{\alpha_{\text{SC}}} e_{\text{SC}}$ estimated in the reflection mode (see Sec. III B 3) and of $\overline{\alpha_{\text{blood}}} e_{\text{blood}}$, with e_{blood} varying for each shear rate. The blood attenuation $\overline{\alpha_{\text{blood}}}$ was

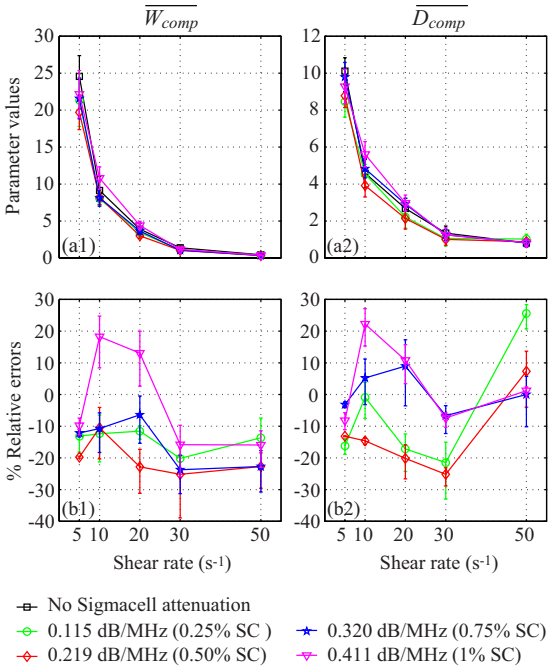


FIG. 8. (Color online) (a) Values of $\overline{W_{\text{comp}}}$ and $\overline{D_{\text{comp}}}$ for different residual shear rates estimated by the classical SFSE with compensation for attenuation by taking predetermined values measured in reflection mode for the four skin-mimicking phantoms (mean $\pm$ standard deviation). (b) Corresponding relative errors of $\overline{W_{\text{comp}}}$ and $\overline{D_{\text{comp}}}$.

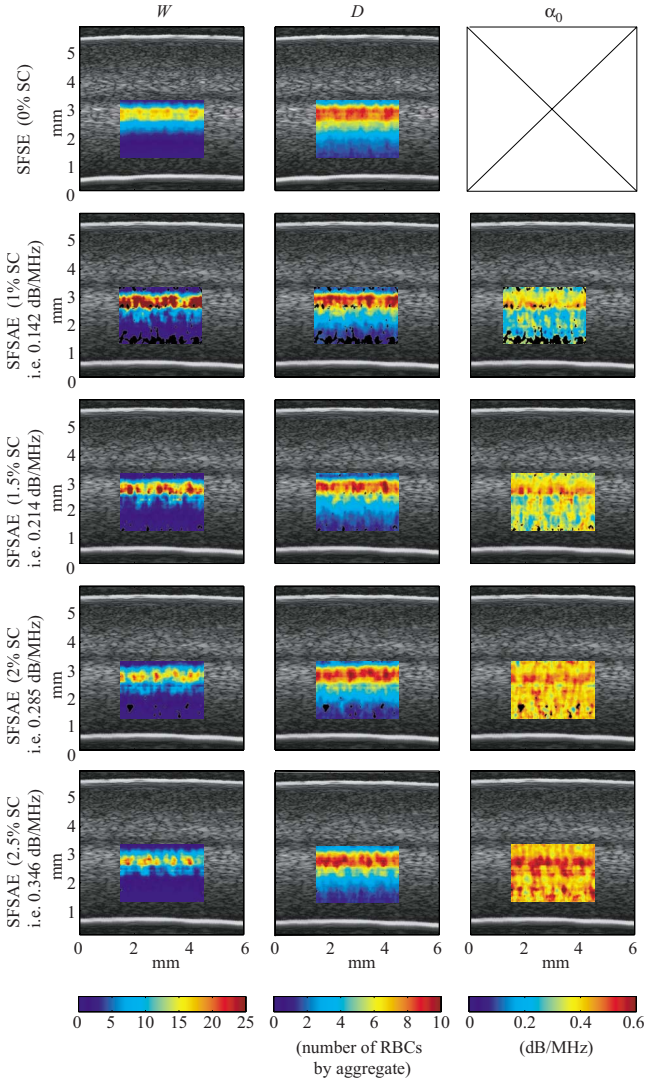


FIG. 9. (Color online) Quantitative images of blood sheared in a tube superimposed on the gray-scale B-mode images. Parameters were estimated by the classical SFSE with compensation for blood attenuation for the 0% SC concentration (top panel) and by the SFSAE for the four skin-mimicking phantoms.

arbitrarily chosen to be equal to 0.028 dB/mm MHz. This point will be discussed later in Sec. V A. Figure 10 summarizes these results. The SFSAE gave estimates of $\overline{W}$, $\overline{D}$, and $\overline{\alpha_0}$ with relative errors below 25% for all skin-mimicking phantoms at shear rates between 0.3 and 2 s^{-1}. Worse estimates were obtained for the smallest shear rate of 0.1 s^{-1}: for 1% and 1.5% SC concentrations, relative errors for $\overline{W}$ were 61% and 32%, respectively, those for $\overline{\alpha_0}$ were 56% and 47%, respectively, whereas those for $\overline{D}$ remained below 20%.

V. DISCUSSION

A. The difficulty to measure the blood attenuation in a tube flow device

The following discussion intends to explain why blood attenuation values in function of the shear rate cannot be obtained directly by measurements in a tube by a standard substitution method. The ideal protocol would be to realize blood attenuation measurements on a Couette apparatus (as

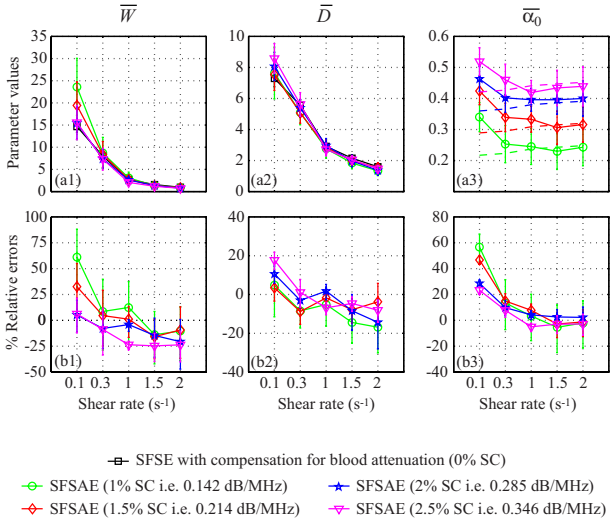


FIG. 10. (Color online) (a) Values of $\overline{W}$, $\overline{D}$, and $\overline{\alpha_0}$ (in dB/MHz) for different shear rates estimated by the SFSAE for the four skin-mimicking phantoms in the tube experiments (mean ± standard deviation). Also represented are values of $\overline{W_{\text{ref}}}$ and $\overline{D_{\text{ref}}}$ estimated by the classical SFSE with compensation for blood attenuation for the 0% SC concentration and values of $\overline{\alpha_{\text{ref}}}$ in dashed lines. (b) Corresponding relative errors of $\overline{W}$, $\overline{D}$ and $\overline{\alpha_0}$.

performed in Sec. III B 3) for the different shear rates present in the tube. In our experiments, the tube was not long enough (kinetic time) to obtain the same aggregate sizes as in the Couette device, that is why we chose a mean blood attenuation equal to 0.028 dB/mm MHz corresponding to the average of blood attenuations estimated with the Couette device for the different shear rates (see Sec. III B 3).

B. Limitations of the SFSAE with respect to *kR* and attenuation

For both Couette and tubular flow devices, the SFSAE gave good estimates of W, D, and α_0 with relative errors below 25% for all skin-mimicking phantoms and for all shear rates, except for the smallest shear rates and the smallest attenuations (0.115 dB/MHz for the Couette and 0.142 dB/MHz for the tube): relative errors for W were around 50% (Couette) and 61% (tube), for α_0 around 48% (Couette) and 56% (tube), for D below 25% as indicated in Figs. 6 and 10. In the tube experiment, one can also notice that relative errors for $\overline{\alpha_0}$ were always worse at the smallest shear rate (0.1 s^{-1}) for all skin-mimicking phantoms. For the peculiar case when there is little attenuation (i.e., ≤0.06 dB/MHz), the SFSAE did not give accurate results (absolute relative errors reaching 42% for W and 211% for α_0).

For the Couette device, the correlation coefficients given in Tables III–V reveal that the SFSAE produced quite good fits to data for all skin-mimicking phantoms and all shear rates, i.e., $0.74 \leq r^2 \leq 0.91$. Nevertheless, for both SFSE and SFSAE models, the smallest r^2 values occurred at the lowest shear rate of 5 s^{-1}. It is illustrated by fits of SFSE and SFSAE models in Fig. 4 that are worse at 5 s^{-1} (in comparison with fits at 10 and 50 s^{-1}), especially for the largest attenuation.

Considering the correlation coefficients and the accuracy of the estimates, both SFSE and SFSAE seem to reach their limit of applicability for large aggregate sizes: for the Couette experiments, typically D_{ref}=10.11 in Table II (i.e., kR =2.88 with f=25 MHz) and for the tube experiments, typically D_{ref}=7.29 in Fig. 10 (i.e., kR=2.08). To conclude, the SFSAE performed well for $kR<2.08$ and for total attenuations between 0.115 and 0.411 dB/MHz. In the following, these bounds will define the SFSAE validity domain, i.e., the domain where the SFSAE gives accurate estimates of W, D, and α_0.

It is interesting to note that the limitations of the SFSAE can be used to interpret the attenuation quantitative images obtained in the tubular flow device (Fig. 9). Indeed, these images are less homogeneous compared to those obtained in the Couette device since the standard deviations for attenuation were between 0.052 and 0.082 dB MHz for the tube and between 0.033 and 0.063 dB/MHz for the Couette. It is clearly linked with the presence of shear rates (<0.3 s^{-1}) in the tube experiments that result in large aggregate sizes outside the validity domain of the SFSAE.

C. Comparison of the SFSAE and SFSE with compensation for attenuation

For the Couette experiments, relative errors for $\overline{W}$ and $\overline{D}$ were generally below 25% (except for one value) with the SFSAE (Fig. 6) and with the SFSE with attenuation-compensation (Fig. 8). For W at 5 s^{-1} and a 0.25% SC, the SFSAE was not in its validity domain. In the SFSAE validity domain, the accuracy of the estimates obtained with the SFSAE was thus as satisfactory as those obtained with the SFSE with attenuation-compensation. The SFSAE has the major advantage to be easily applicable *in vivo* contrary to the SFSE attenuation-compensation method, needing the attenuation and thickness of the tissue intervening layers to be known.

D. Rejected solutions

In Figs. 5 and 9, one can notice the rejected solutions represented in black pixels. These solutions were rejected because their estimated diameters were equal to 0 (whereas the estimated packing factors never vanish). One might think that these rejected solutions come from cost functions having several local minima. Several cases of rejected solutions were studied and their corresponding cost functions had always a unique global minimum (data not shown).

To better understand what happened in these special cases, Fig. 11(a) represents two backscatter coefficients for blood sheared at 20 s^{-1} in the Couette flow device and measured with the 0.25% SC concentration phantom. Although both experimental backscatter coefficient were quite similar, the corresponding fitted models with the SFSAE gave two very different estimates of W, D, and α_0. One estimation corresponds to a kept solution and the other to a rejected solution with an estimate of D equal to 0. In Fig. 12 the mean values of $\overline{W}$, $\overline{D}$, and $\overline{\alpha_0}$ are summarized for the kept and rejected solutions estimated with the SFSAE for all shear rates and for the 0.25% SC concentration phantom. It can be observed, for the rejected solutions, that both $\overline{W}$ and $\overline{\alpha_0}$ were overestimated. This phenomenon can be explained by the

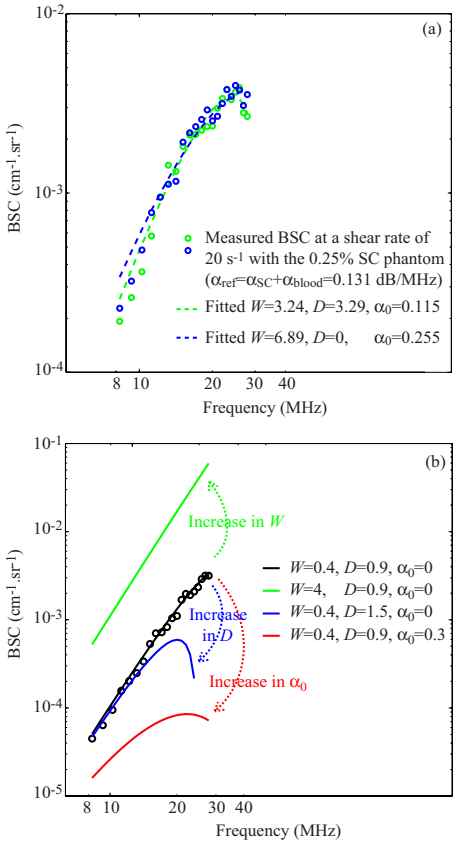


FIG. 11. (Color online) (a) Two backscatter coefficients for blood sheared at 20 s^{-1} in the Couette flow device and measured with the 0.25% SC phantom, and corresponding fitted models with the SFSAE. (b) Effect of increasing W, D, and α_0 on the backscattering coefficient on the SFSAE model.

difficulty to evaluate simultaneously three parameters that have the same effect on the backscatter coefficient as explained below. Figure 11(b) illustrates the respective effect of varying W, D, and α_0 on the backscatter coefficient. We arbitrarily selected a shear rate of 50 s^{-1} (i.e., a quasi-disaggregated case) and no SC attenuation. An increase in W has the effect of increasing the amplitude of the backscatter

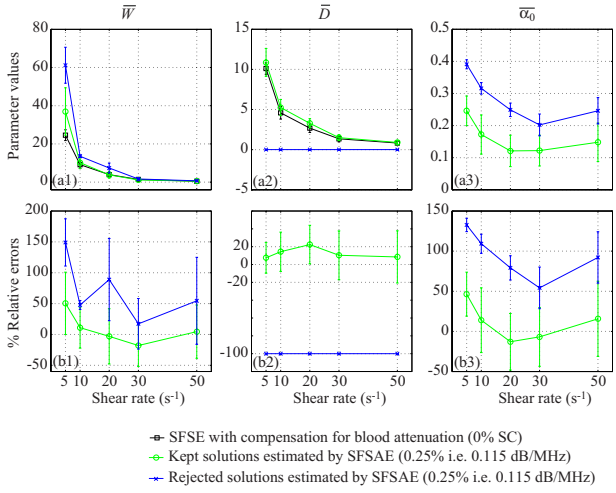


FIG. 12. (Color online) (a) Values of $\bar{W}$, $\bar{D}$, and $\overline{\alpha_0}$ (in dB/MHz) for the kept and rejected solutions of the SFSAE for different residual shear rates and for the 0.25% SC concentration phantom in the Couette experiment (mean ± standard deviation). (b) Corresponding relative errors of $\bar{W}$, $\bar{D}$, and $\overline{\alpha_0}$.

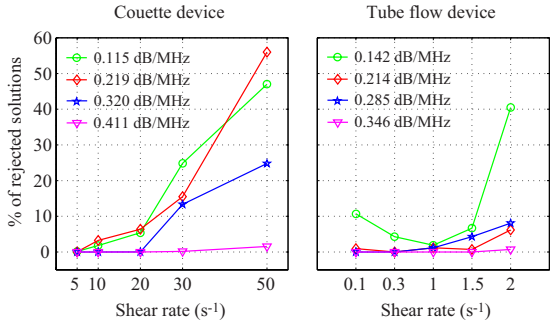


FIG. 13. (Color online) Percentage of rejected solutions in the ROIs: 20 depths and 180 RF lines for the Couette and 3 depths and 180 RF lines for the tube.

coefficient at all frequencies; whereas increasing D has an effect on the frequency dependence. But more interesting is to observe that an increase in α_0 has the effect of decreasing the amplitude and also of modifying the frequency dependence of the backscatter coefficient. It is thus difficult to estimate simultaneously these three parameters, because a modification in α_0 plays the same role (i.e., effects on the amplitude and the frequency dependence) that modifications in W (i.e., amplitude) and D (i.e., frequency dependence). The typical example given in Fig. 11(a) and parameter values presented in Fig. 12 can now be better understood. When the optimization method gives an unrealistic estimate of D equal to 0, the frequency dependence of the backscatter coefficient is only given by α_0, which is overestimated. But since increasing α_0 has also the effect of decreasing the amplitude of the backscatter coefficient, the overestimation of α_0 goes with the overestimation of W.

Figure 13 gives the percentage of rejected solutions in the ROIs (i.e., 20 depths and 180 rf lines for the Couette, as shown in Fig. 5, and three depths and 180 rf lines for the tube) at each shear rate for both Couette and tubular flow devices. For both devices, it can be clearly observed that percentages of rejected solutions are highest for experiments with the smallest attenuation (0.25% SC for the Couette and 1% SC for the tube) and with the highest shear rate, i.e., with quasi-disaggregated RBCs (50 s^{-1} for the Couette and 2 s^{-1} for the tube). The rejected solutions were thus more important when the frequency dependence was nearly Rayleigh. In our previous study,[23] we did not reject solution, that is why bad estimates were obtained for quasi-disaggregated RBCs. The rejection of some solutions was vital for having accurate results in the cases of quasi-disaggregated RBCs and of small total attenuation.

E. On the use of the SFSAE *in vivo*

The accuracy of the SFSAE was quantitatively demonstrated in a validity domain. One might question the use of this method *in vivo*. Although the SFSAE gave estimates with large errors for $kR > 2.08$ and $\alpha_0 \leq 0.06$ dB/MHz, the estimated parameters presented in Figs. 6 and 10 show that the SFSAE gave qualitatively satisfactory estimates for all SC phantoms at all shear rates, since the estimates of $\bar{W}$ and $\bar{D}$ versus shear rates had the same behaviors as $\overline{W_{\text{ref}}}$ and $\overline{D_{\text{ref}}}$. It means that the validity of estimates is possible for *in vivo*

experiments, since the SFSAE is able to predict the aggregation state and attenuation tendencies.

VI. CONCLUSION

Aggregating porcine RBCs were examined using US via *in vitro* experiments in Couette and tubular flow devices. Estimates of blood structural parameters and total attenuation were made from the frequency dependence of the US backscatter using the SFSAE model. This study revealed that the SFSAE provided accurate quantitative estimates of blood microstructure parameters W and D for $kR < 2.08$ and for total attenuations between 0.115 and 0.411 dB/MHz. In this validity domain, the results obtained with the SFSAE was quantitatively as satisfactory as those obtained with the SFSE with attenuation-compensation (i.e., when prior information on the attenuation is available). Outside this validity domain, the SFSAE was able to give the aggregation state tendency.

The SFSAE has been shown to be able of estimating blood backscattering properties in the presence of tissue intervening layers. This method is easily applicable *in vivo* because of the simultaneous estimation of the blood structural properties and total attenuation, contrary to the SFSE attenuation-compensation method, needing the attenuation and thickness of the tissue intervening layers to be known. Future works should focus on *in vivo* and *in situ* assessment of the pathophysiological impact of abnormal RBC aggregation on the cardiovascular system.

ACKNOWLEDGMENTS

This work was supported by the Canadian Institutes of Health Research (Grant Nos. MOP-84358 and CMI-72323), by the Heart and Stroke Foundation of Canada (Grant No. PG-05-0313), and by the National Institutes of Health of USA (Grant No. RO1HL078655).

[1]F. L. Lizzi, M. Greenebaum, E. J. Feleppa, and M. Elbaum, "Theoretical framework for spectrum analysis in ultrasonic tissue characterization," J. Acoust. Soc. Am. **73**, 1366–1373 (1983).
[2]E. J. Feleppa, F. L. Lizzi, D. J. Coleman, and M. M. Yaremko, "Diagnostic spectrum analysis in ophthalmology: A physical perspective," Ultrasound Med. Biol. **12**, 623–631 (1986).
[3]F. L. Lizzi, M. Ostromogilsky, E. J. Feleppa, M. C. Rorke, and M. M. Yaremko, "Relationship of ultrasonic spectral parameters to features of tissue microstructure," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **33**, 319–329 (1986).
[4]M. F. Insana, J. G. Wood, and T. J. Hall, "Identifying acoustic scattering sources in normal renal parenchyma in vivo by varying arterial and ureteral pressures," Ultrasound Med. Biol. **18**, 587–599 (1992).
[5]E. J. Feleppa, T. Liu, A. Kalisz, M. C. Shao, N. Fleshner, and V. Reuter, "Ultrasonic spectral-parameter imaging of the prostate," Int. J. Imaging Syst. Technol. **8**, 11–25 (1997).
[6]M. L. Oelze, W. D. O'Brien, J. P. Blue, and J. F. Zachary, "Differentiation and characterization of rat mammary fibroadenomas and 4T1 mouse carcinomas using quantitative ultrasound imaging," IEEE Trans. Med. Imaging **23**, 764–771 (2004).
[7]F. T. H. Yu and G. Cloutier, "Experimental ultrasound characterization of red blood cell aggregation using the structure factor size estimator," J. Acoust. Soc. Am. **122**, 645–656 (2007).
[8]M. O'Donnell, J. W. Mimbs, and J. G. Miller, "Relationship between collagen and ultrasonic backscatter in myocardial tissue," J. Acoust. Soc. Am. **69**, 580–588 (1981).
[9]M. F. Insana, T. J. Hall, and J. L. Fishback, "Identifying acoustic scattering sources in normal renal parenchyma from the anisotropy in acoustic properties," Ultrasound Med. Biol. **17**, 613–626 (1991).
[10]K. A. Wear, M. R. Milunski, S. A. Wickline, J. E. Perez, B. E. Sobel, and J. G. Miller, "Differentiation between acutely ischemic myocardium and zones of completed infarction in dogs on the basis of frequency-dependent backscatter," J. Acoust. Soc. Am. **85**, 2634–2641 (1989).
[11]A. F. W. Van Der Steen, J. M. Thijssen, J. A. W. M. Van Der Laak, G. P. J. Ebben, and P. C. M. de Wilde, "Correlation of histology and acoustic parameters of liver tissue on a microscopic scale," Ultrasound Med. Biol. **20**, 177–186 (1994).
[12]K. A. Wear, B. S. Garra, and T. J. Hall, "Measurements of ultrasonic backscatter coefficients in human liver and kidney *in vivo*," J. Acoust. Soc. Am. **98**, 1852–1857 (1995).
[13]T. J. Hall, M. F. Insana, L. A. Harrison, and G. G. Cox, "Ultrasonic measurement of glomerular diameters in normal adult humans," Ultrasound Med. Biol. **22**, 987–997 (1996).
[14]Z. F. Lu, J. A. Zagzebski, and F. T. Lee, "Ultrasound backscatter and attenuation in human liver with diffuse disease," Ultrasound Med. Biol. **25**, 1047–1054 (1999).
[15]S. A. Goss, R. L. Johnston, and F. Dunn, "Comprehensive compilation of empirical ultrasonic properties of mammalian tissues," J. Acoust. Soc. Am. **64**, 423–457 (1978).
[16]P. He and J. F. Greenleaf, "Application of stochastic analysis to ultrasonic echoes—Estimation of attenuation and tissue heterogeneity from peaks of echo envelope," J. Acoust. Soc. Am. **79**, 526–534 (1986).
[17]L. X. Yao, J. A. Zagzebski, and E. L. Madsen, "Backscatter coefficient measurements using a reference phantom to extract depth-dependent instrumentation factors," Ultrason. Imaging **12**, 58–70 (1990).
[18]B. J. Oosterveld, J. M. Thijssen, P. C. Hartman, R. L. Romijn, and G. J. E. Rosenbusch, "Ultrasound attenuation and texture analysis of diffuse liver disease: Methods and preliminary results," Phys. Med. Biol. **36**, 1039–1064 (1991).
[19]V. Roberjot, S. L. Bridal, P. Laugier, and G. Berger, "Absolute backscatter coefficient over a wide range of frequencies in a tissue-mimicking phantom containing two populations of scatterers," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 970–978 (1996).
[20]T. A. Bigelow, M. L. Oelze, and W. D. O'Brien, "Estimation of total attenuation and scatterer size from backscatter ultrasound waveforms," J. Acoust. Soc. Am. **117**, 1431–1439 (2005).
[21]T. A. Bigelow and W. D. O'Brien, "Signal processing strategies that improve performance and understanding of the quantitative ultrasound SPECTRAL FIT algorithm," J. Acoust. Soc. Am. **118**, 1808–1819 (2005).
[22]T. A. Bigelow and W. D. O'Brien, "Evaluation of the Spectral Fit algorithm as functions of frequency range and Δka_{eff}," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **52**, 2003–2010 (2005).
[23]E. Franceschini, F. T. H. Yu, and G. Cloutier, "Simultaneous estimation of attenuation and structure parameters of aggregated red blood cells from backscatter measurements," J. Acoust. Soc. Am. **123**, EL85–EL91 (2008).
[24]T. F. Coleman and Y. Li, "A reective Newton method for minimizing a quadratic function subject to bounds on some of the variables," SIAM J. Optim. **6**, 10401058 (1996).
[25]L. Y. L. Mo and R. S. C. Cobbold, "Theoretical models of ultrasonic scattering in blood," in *Ultrasonic Scattering in Biological Tissues*, edited by K. K. Shung and G. A. Thieme (CRC, Boca Raton, FL, 1993), Chap. 5, pp. 125–170.
[26]V. Twersky, "Low-frequency scattering by correlated distributions of randomly oriented particles," J. Acoust. Soc. Am. **81**, 1609–1618 (1987).
[27]K. K. Shung, "On the ultrasound scattering from blood as a function of hematocrit," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **SU-26**, 327–331 (1982).
[28]S. H. Wang and K. K. Shung, "An approach for measuring ultrasonic backscattering from biological tissues with focused transducers," IEEE Trans. Biomed. Eng. **44**, 549–554 (1997).
[29]L. C. Nguyen, F. T. H. Yu, and G. Cloutier, "Cyclic changes in blood echogenicity under pulsatile flow are frequency dependent," Ultrasound Med. Biol. **34**, 664–673 (2008).
[30]R. Kuc and M. Schwartz, "Estimating the acoustic attenuation coefficient slope for liver from reflected ultrasound signals," IEEE Trans. Sonics Ultrason. **SU-26**, 353–362 (1979).
[31]I. Raju and M. A. Srinivasan, "High-frequency ultrasonic attenuation and backscatter coefficients of *in vivo* normal human dermis and subcutaneous fat," Ultrasound Med. Biol. **27**, 1543–1556 (2001).
[32]R. E. N. Shehada and R. S. C. Cobbold, "Ultrasound methods for investigating the non-newtonian characteristics of whole blood," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 96–104 (1994).

Black-capped (*Poecile atricapillus*) and mountain chickadee (*Poecile gambeli*) contact call contains species, sex, and individual identity features

Lauren M. Guillette
Department of Psychology, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Laurie L. Bloomfield
Department of Psychology, Algoma University, Sault Sainte Marie, Ontario P6A 2G4, Canada

Emily R. Batty and Michael R. W. Dawson
Department of Psychology, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Christopher B. Sturdy[a)]
Department of Psychology and Centre for Neuroscience, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

(Received 15 May 2009; revised 4 November 2009; accepted 4 December 2009)

The *tseet* contact call, common to both black-capped and mountain chickadees, is among the most frequently produced call of each species, but has remained little studied until now. In the current study, the authors characterized the *tseet* call of adult allopatric and sympatric black-capped and mountain chickadees in terms of nine acoustic features in a fashion similar to descriptive accounts of both species' *chick-a-dee* calls. Summary statistics, the potential for individual coding, and classification by linear discriminant analysis were used to describe the *tseet* call. The authors were able to correctly classify *tseet* calls in terms of which group or individual produced it with high accuracy. Furthermore, several acoustic features are highly individualized, indicating that the chickadees may use these features to identify signalers as individuals or members of a particular group. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3277247]

PACS number(s): 43.80.Ka [JAS] Pages: 1116–1123

I. INTRODUCTION

Bird calls are typically produced year round by both sexes and serve specific functional roles such as to deter predators, raise alarm within a flock, coordinate flock movement (Smith, 1991), announce or exchange food, and maintain group cohesion (Marler and Slabbekoorn, 2004). Previous studies have described the *chick-a-dee* call of black-capped chickadees (*Poecile atricapillus*), mountain chickadees (*P. gambeli*), Carolina chickadees (*P. Carolinensis*), and the closely related tufted titmice (*Baeolophus bicolor*). Quantitative analyses reveal that the note types within the *chick-a-dee* calls of all species are individualized, with significant differences between individuals (within each species) (Charrier *et al.*, 2004; Bloomfield *et al.*, 2004, 2005; Owens and Freeberg, 2007). However, whether other vocalizations common to both black-capped and mountain chickadees also contain features that convey information about the signalers' species, sex, or individual identity remain untested.

In addition to the *chick-a-dee* call, the *tseet* call is a contact call in the vocal repertoires of both black-capped and mountain chickadees. *Tseet* calls are the most frequently produced vocalization by black-capped chickadees (Odum, 1942) and, although not yet quantified, possibly by mountain chickadee as well (pers. obs.). Despite being the most commonly produced chickadee call-type, the *tseet* call has received little attention from researchers. In the present study, we examine the *tseet* calls of black-capped and mountain chickadees, the most closely related of the seven species of North American chickadees (Gill *et al.*, 2005).

The aim of the current study is to describe the *tseet* call of black-capped and mountain chickadees in a manner similar to previous work on the *chick-a-dee* call of these species. Specifically, we measure and examine nine acoustic features of the *tseet* calls of these species for evidence of individual identity features, and we classify calls according to the individual, species, and geographic origin for the birds that produced each call using linear discriminant analysis (LDA). We examine the *tseet* calls of (1) black-capped chickadees originating from regions that are also inhabited by mountain chickadees (sympatric group), (2) mountain chickadees from this same region, and (3) black-capped chickadees originating from regions where there are no mountain chickadees (allopatric group).

II. METHODS

A. Subjects

A total of 30 birds of at least 1 year of age (determined by the shape and coloring of outer tail retrices, Pyle, 1997) were captured between March 2002 and March 2004. Ten black-capped chickadees (5 male, 5 female) originated from several locations in Edmonton, Alberta, Canada (53° 06′N, 113° 04′W) and had no experience with mountain chickadees since both species are relatively non-migratory, and

[a)]Author to whom correspondence should be addressed. Electronic mail: csturdy@ualberta.ca

mountain chickadees do not inhabit the Edmonton area. Thus, black-capped chickadees originating from Edmonton are referred to as the allopatric group throughout. Ten mountain chickadees (5 male, 5 female) and ten black-capped chickadees (5 male, 5 female) were captured from several location in Kananaskis Valley (Alberta, Canada, 51° 02′N, 115° 03′W), approximately 300 km southwest of Edmonton. In Kananaskis, mountain and black-capped chickadees co-occur and can be observed foraging in close proximity of each other (pers. obs.), and hence, black-capped chickadees originating from Kananaskis are referred to as the sympatric group. Sex identification was conducted by DNA analysis (Griffiths and Double, 1998).

Each species was housed separately at the University of Alberta in individual Jupiter Parakeet cages (0.3×0.4×0.4 m; Rolf C. Hagen, Inc., Montreal, Canada). Housing conditions allowed for auditory and visual contacts, but not physical contact with conspecifics. Birds had food (Mazuri Small Bird Maintenance Diet; Mazuri, St. Louis, MO), water (vitamin supplemented on alternate days; Hagen, Rolf C. Hagen, Inc., Montreal, Canada), grit, and cuttle bone *ad libitum*. Birds were given three to five sunflower seeds daily. Birds also received one mealworm three times a week and a mixture of eggs and greens twice a week. Birds were maintained on a light-dark cycle that mimicked the natural cycle for Edmonton, Alberta.

B. Recordings

Birds were individually recorded in a sound attenuating chamber (1.83×1.83×1.83 m, Industrial Acoustics Corporation, Bronx, NY) using an AKG C 1000S condenser microphone connected to a Marantz PMD 670 digital recorder (frequency response: 10–20 000 Hz; Marantz Electronics, Eindhoven, The Netherlands). Digitized files (44 100 Hz) were downloaded to an Intel based PC and analyzed with SIGNAL 4.0 (Engineering Design, Berkeley, CA).

C. Acoustic analysis

Ten high quality (e.g., not distorted from cage noise), clearly classifiable (i.e., not another note type produced in isolation from a call) *tseet* notes were randomly selected from several recording sessions for each bird. To standardize all analyses with similar resolution, each individual call was saved as a separate file with a duration of 300 ms by adding leading and trailing silence to each file. Three temporal and three spectral measures were taken from the sound spectrograms (cut-off amplitude of −35 dB relative to peak amplitude) and fast Fourier transforms (FFTs, settings for temporal measures: Hanning window=256 points, precision=5.8 ms; FFT settings for spectral measures: Hanning window=1024 points, precision=43 Hz).

The three temporal measures are (1) the ascending duration (AD), which is measured from the start of the call to the point where the rapid frequency modulation ends; (2) the descending duration (DD), which is measured from the highest frequency of the call where the call begins to descend in frequency, until the end of the call; and (3) the total duration (TD) of the call. The three frequency measures are (1) the start frequency (SF), (2) the peak frequency (PF), and (3) end frequency (EF) of the call. Three additional features of the calls are (1) the maximum (i.e., loudest) frequency (Fmax), which was measured in a spectrum window (modal window size=4096 points; (2) the rate of rise for the ascending slope (FMasc), which was calculated using the formula (PF −SF)/AD; and (3) the rate of decay in the descending slope (FMdesc), which was calculated using the formula (EF −PF)/DD (see Fig. 1).

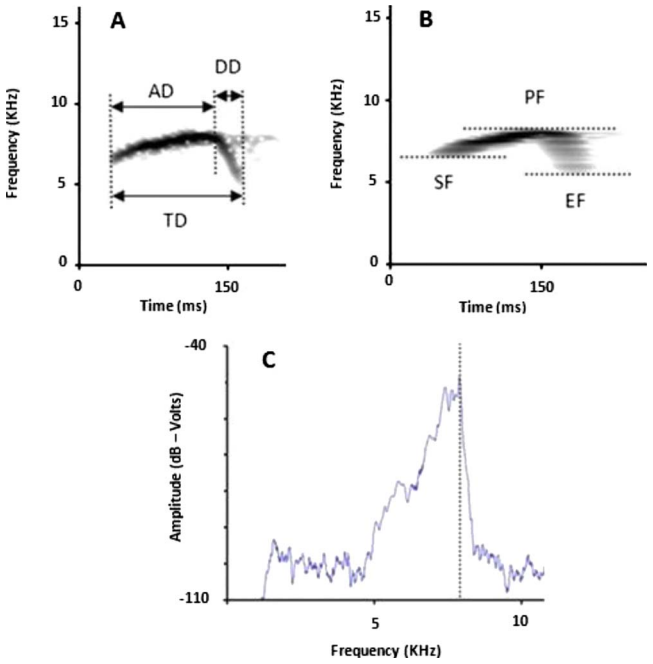


FIG. 1. (Color online) Sound spectrogram and spectrum illustrating note-type features measured in *tseet* calls. Panel A: Frequency (*y*-axis) by time (*x*-axis) sound spectrogram (FFT window=256 points) of a *tseet* call. Vertical lines represent approximate boundaries for ascending duration (AD), descending duration (DD), and total note duration (TD). Panel B: Frequency (*y*-axis) by time (*x*-axis) sound spectrogram (FFT window=1024 points) of a *tseet* call. Horizontal lines illustrate the start frequency (SF), peak frequency (PF), and end frequency (EF). Panel C: Relative amplitude (*y*-axis) by frequency (*x*-axis) sound spectrum (window size=4096 points) of a *tseet* call. The vertical line illustrates the frequency at max amplitude (Fmax).

III. RESULTS

A. Classification of *tseet* calls by individuals

Bloomfield and Sturdy (2008) demonstrated that black-capped chickadees can memorize up to 15 individual black-capped and mountain chickadee *chick-a-dee* calls, each from a unique individual. For birds to be able to discriminate among individuals based on their *tseet* calls, one or more of the nine acoustic features outlined above may contain information specific to each individual within a group. One or more acoustic parameter(s) must have high individual stereotypy and thus may be used by conspecifics for individual identification. If this is the case, then acoustic variation in a single feature among all *tseet* calls in a group (e.g., SF for all mountain chickadees) must be greater than the acoustic variation in that same feature within an individual's calls (e.g., average SF for an individual mountain chickadee). To identify the acoustic parameters that may contain individual

identity in the *tseet* call, we examined the potential for individual coding (PIC) (Charrier *et al.*, 2004). The PIC value is the ratio of the coefficient of variation between individuals (CV_b) to the average of the coefficients of variation within an individual (CV_w) for a particular group (e.g., allopatric chickadees). In this formulation, $CV_b=(SD/X)\times 100$, where SD is the standard deviation of the group (e.g., allopatric chickadees), X is the group mean for an acoustic feature (e.g., PF), and CV_w is the average of the coefficients of variation for each individual in that group. Acoustic features that yield a PIC value greater than 1 are potentially used for individual recognition since the inter-individual variability (for that specific acoustic feature) is greater than the intra-individual variability (Sokal and Rohlf, 1995).

Table I shows the means, standard deviations, coefficients of variations within and between individuals, and the PIC. While all PIC values are greater than 1 (i.e., for each species, allopatric or sympatric and each sex), the *tseet* call of males tended to have more individualized acoustic features than the *tseet* call of their female counterparts. Across all sympatric birds the acoustic features with the highest PIC values are the spectral features of start and end frequencies. For allopatric chickadees, the acoustic feature with the highest PIC value is the total duration.

We employed stepwise LDAs (SPSS 15.0 for Windows) to classify each *tseet* call in terms of which individual or group of chickadees produced it. If it is possible for chickadees to use a feature, or a combination of features from *tseet* calls to identify individuals as belonging to a particular group, then it follows that the LDA should be able to predict group membership with a high degree of accuracy (see Dawson *et al.*, 2006). Several different LDAs were preformed; each LDA could use the nine different acoustic features, outlined in the acoustic analysis section above, as independent variables to classify *tseet* calls into groups according to which chickadee produced the call.

Analyzing errors in the LDA predictions (i.e., when the LDA identifies a *tseet* call as being produced by the incorrect individual or group of chickadees) is a useful tool to understand subtle similarities and differences between the *tseet* calls produced by the different individuals, groups, and sexes. In LDA the squared canonical correlation (R_c^2) can be interpreted as the proportion of variation that is accounted for by the independent variables. The standardized discriminant function coefficient is used to assess the relative importance of each independent variable through its unique contribution to a particular discriminant function (Betz, 1987; Klecka, 1980). The limitation of interpreting the standardized discriminant function coefficient is that variables that are highly correlated share contributions to the discriminant score. Therefore we have also included the structure coefficients, the bivariate correlation between the variable and discriminant function which therefore are not influenced by covariances with other variables. Structure coefficients identify the aspect of the discriminant function that discriminates between groups (Klecka, 1980).

Nine stepwise LDAs were conducted to classify *tseet* calls in terms of which individual produced it. A separate LDA was conducted for each group of birds (male allopatric, female allopatric, male sympatric, female sympatric, male mountain, female mountain, all allopatric, all sympatric, and all mountain chickadees). The purpose of these LDAs is twofold: (1) to determine the ability of the model to correctly classify each *tseet* call in terms of which individual produced it and (2) to determine the correspondence between the feature with the highest PIC and the feature that was loaded most heavily on (standardized coefficient) or correlated with (structure coefficient) the first discriminant function for each group of birds. In this way, we used two complimentary statistical techniques for assessing individualization of *tseet* calls.

Table II shows a summary of the LDAs conducted to classify *tseet* calls in terms of which individual in a distinct group produced the call. The original and cross-validation of correctly identified cases, along with the eigenvalue and the canonical correlation, the standardized discriminant function coefficient, and structure coefficient for the first discriminant functions are listed. For four groups of birds, the feature with the largest standardized discriminant function coefficient and largest structure coefficient for the first discriminant function matches the feature which yielded the highest PIC: TD for female allopatric black-capped chickadees, SF for female sympatric chickadees, SF for male and female sympatric chickadees combined, and EF for male and female mountain chickadees combined, although the combined group is likely driven by the males in the mixed sex mountain group. There was almost a perfect match, with the exception of female mountain chickadees, for the type of feature (temporal vs spectral) with the highest PIC value and highest structure coefficient, suggesting that these features account for the majority of the individual differences in *tseet* calls within each group.

B. Classification of *tseet* calls into groups

Another LDA classified *tseet* calls into three groups, namely, calls produced by allopatric chickadees, calls produced by sympatric chickadees, and calls produced by mountain chickadees. The purpose of this LDA was to classify each call in terms of the bird (allopatric, sympatric, and mountain chickadees) that produced it using the nine acoustic measures. Two discriminant functions (function 1 eigenvalue=3.716, $R_c^2=0.889$; function 2 eigenvalue =1.216, $R_c^2=0.6741$) correctly classified *tseet* calls in terms of species and geographic origin in 87.7% of cases (see Table III). A z-test for the differences between two proportions shows that the predictions yielded by the LDA are significantly better than predictions expected by chance for allopatric black-capped ($z=9.01$, $p<0.05$), sympatric black-capped ($z=6.07$, $p<0.05$), and mountain chickadees ($z=9.28$, $p<0.05$, Betz, 1987; Glass and Stanely, 1970). All but one of the nine acoustic measures (AD) contributed to the discriminant functions (see Table IV). From examining group centroid loadings on the discriminant functions, we see that function 1 separates allopatric black-capped from mountain chickadees and the acoustic measure with the largest standardized coefficient for function 1 is end frequency. The second discriminant function separated the sympatric black-

TABLE I. Means, standard deviations, between-bird coefficients of variation (CV_b), within-bird coefficients of variation (CV_w), and PIC for each acoustic feature measured in *tseet* calls. ΣPIC is the sum of the PIC values for that group.

		TD	AD	DD	SF	PF	EF	Fmax	FMasc	FMdesc
					Male allopatric					
ΣPIC=15.34	*X*	54.67	36.70	12.04	6009.12	7590.27	4264.49	7127.07	46.43	−284.01
	SD	7.66	10.38	3.57	434.67	564.05	651.31	478.49	17.45	91.99
	CV_b	14.01	28.29	29.63	7.23	7.43	15.27	6.71	37.58	−32.39
	CV_w	4.84	15.18	20.58	5.35	3.69	11.16	4.32	28.07	−21.37
	PIC	2.89	1.86	1.44	1.35	2.01	1.37	1.55	1.34	1.52
					Female allopatric					
ΣPIC=11.44	*X*	48.61	36.43	10.85	5907.02	7619.89	4678.73	7153.96	48.20	−275.79
	SD	7.37	6.34	1.78	367.27	299.35	600.00	371.97	12.05	63.64
	CV_b	15.17	17.41	16.43	6.22	3.93	12.82	5.20	24.99	−23.08
	CV_w	10.42	12.05	13.03	5.53	3.19	9.36	3.94	23.19	−19.96
	PIC	1.46	1.44	1.26	1.12	1.23	1.37	1.32	1.08	1.16
					Male sympatric					
ΣPIC=16.20	*X*	52.56	19.62	11.08	5959.33	7736.13	6785.75	7734.68	68.43	−148.26
	SD	6.80	9.20	4.55	836.86	334.06	394.38	263.52	20.26	88.76
	CV_b	12.94	46.91	41.08	14.04	4.32	5.81	3.41	29.61	−59.87
	CV_w	6.85	32.09	30.69	5.17	2.07	4.00	1.30	27.41	−38.43
	PIC	1.89	1.46	1.34	2.72	2.08	1.45	2.62	1.08	1.56
					Female sympatric					
ΣPIC=15.85	*X*	47.97	33.87	12.95	5006.43	7306.42	5559.13	7371.12	47.18	−228.28
	SD	4.61	5.41	1.86	831.48	405.82	630.32	303.79	9.52	35.80
	CV_b	9.62	15.97	14.34	16.61	5.55	11.34	4.12	20.17	−15.68
	CV_w	8.10	13.68	9.96	4.65	2.60	5.73	2.11	19.08	−11.62
	PIC	1.19	1.17	1.44	3.57	2.14	1.98	1.96	1.06	1.35
					Male mountain					
ΣPIC=17.33	*X*	79.72	39.67	17.92	5600.30	6901.08	5621.67	6705.55	40.09	−68.75
	SD	18.46	13.99	6.69	1137.85	511.96	1133.88	482.63	32.62	21.14
	CV_b	23.15	35.27	37.35	20.32	7.42	20.17	7.20	81.37	−30.75
	CV_w	12.84	29.88	22.02	9.44	3.67	6.58	3.58	43.37	−20.23
	PIC	1.80	1.18	1.70	2.15	2.02	3.06	2.01	1.88	1.52
					Female mountain					
ΣPIC=11.79	*X*	61.95	30.03	15.04	6416.82	7364.21	6527.00	7125.56	32.08	−58.41
	SD	9.61	15.04	4.98	493.66	348.72	506.42	329.86	16.41	17.47
	CV_b	15.52	50.07	33.12	7.69	4.74	7.76	4.63	51.16	−29.91
	CV_w	12.41	37.64	24.66	3.94	2.64	4.25	2.49	54.44	−30.32
	PIC	1.25	1.33	1.34	1.95	1.79	1.83	1.86	0.94	0.99
					Allopatric					
ΣPIC=13.70	*X*	51.64	36.57	11.45	5958.07	7605.08	4471.61	7140.52	47.32	−279.90
	SD	8.08	8.56	2.87	403.62	449.49	656.87	426.60	14.94	78.81
	CV_b	15.64	23.41	25.06	6.77	5.91	14.69	5.97	31.58	−28.16
	CV_w	7.63	13.61	16.80	5.44	3.44	10.26	4.13	25.63	−20.67
	PIC	2.05	1.72	1.49	1.25	1.72	1.43	1.45	1.23	1.36
					Sympatric					
ΣPIC=19.21	*X*	50.26	26.74	12.02	5482.88	7521.27	6172.44	7552.90	57.80	−188.27
	SD	6.23	10.38	3.58	958.18	428.22	808.44	336.79	19.03	78.43
	CV_b	12.39	38.80	29.83	17.48	5.69	13.10	4.46	32.92	−41.66
	CV_w	7.48	22.88	20.32	4.91	2.33	4.87	1.70	23.25	−25.02
	PIC	1.66	1.70	1.47	3.56	2.44	2.69	2.62	1.42	1.66
					Mountain					
ΣPIC=16.47	*X*	70.83	34.85	16.48	6008.56	7132.65	6074.34	6915.55	36.08	−63.58
	SD	17.15	15.24	6.05	964.26	494.04	985.02	462.27	26.00	19.98
	CV_b	24.21	43.73	36.69	16.05	6.93	16.22	6.68	72.06	−31.43
	CV_w	12.67	33.70	23.50	7.19	3.53	5.89	3.41	47.28	−24.80
	PIC	1.91	1.30	1.56	2.23	1.96	2.75	1.96	1.52	1.27

TABLE II. The original (Orig) and cross-validated (Cross) percent of cases correctly classified by individual identity by stepwise linear discriminant analysis for all groups of birds, M=male, F=female, allop =allopatric black-capped chickadee, sym=sympatric black-capped chickadee, mo=mountain chickadee. The eigenvalue (Eigen) and canonical correlation (R_c^2), acoustic feature with the highest standardized canonical function coefficient (Std coeff), and highest structure coefficient (Str coeff) for the first discriminant function. PIC represents the acoustic feature, which yielded the highest PIC value. TD=total duration, AD=ascending duration, DD=descending duration, Fmax=frequency at loudest amplitude, SF=start frequency, EF=end frequency, and PF=peak frequency.

Group	Orig	Cross	Eigen	R_c^2	Std coeff	Str coeff	PIC
M allop	86	82	8.888	0.948	PF	TD	TD
F allop	60	56	1.211	0.74	TD	TD	TD
M sym	92	88	11.92	0.961	Fmax	Fmax	SF
F sym	98	96	36.021	0.986	SF	SF	SF
M mo	84	70	12.25	0.962	Fmax	EF	EF
F mo	88	70	5.316	0.917	Fmax	AD	SF
Allop	63	48	3.448	0.88	PF	DD	TD
Sym	95	88	18.895	0.975	SF	SF	SF
Mo	75	58	10.097	0.954	EF	EF	EF

capped chickadees from allopatric black-capped and mountain chickadees. The feature with the largest standardized coefficient for function 2 is Fmax. Both functions combined yielded Wilks's lambda=0.095, distributed as $\chi^2(16, N=299)=691.514$ and $p<0.001$, indicating that the group centroids differed significantly from each other and indicating that the model is successfully discriminating among the groups.

The last LDA classified *tseet* calls into six groups, namely, calls produced by either male allopatric, female allopatric, male sympatric, female sympatric, male mountain, and female mountain chickadees, also based on the nine acoustic measures obtained from our sample of *tseet* calls. Five discriminant functions correctly classified *tseet* calls in terms of sex, species, and geographic origin in 67.7% of all cases (see Table V). While all five functions combined yielded Wilks's lambda=0.033, distributed as $\chi^2(35, N=299)=996.358$ and $p<0.001$, the first two discriminant functions accounted for the majority (88.7%) of the variance accounted for overall. In the first discriminant function both EF (2.396) has large standardized function coefficients, while Fmax (−0.641) has the largest standardized function coefficients in the second discriminant function (see Table VI). A z-test for the differences between two proportions show that the predictions yielded by the LDA are significantly better ($p<0.05$) than predictions expected by chance for all groups (male allopatric, $z=20.48$; female allopatric, $z=22.33$; male sympatric, $z=32.86$; female sympatric, z

TABLE III. Matrix of classification by group membership of the actual groups of chickadees *tseet* calls and the LDA predicted group classification based on nine measured acoustic features. Correct LDA classification are presented (in percentages) along the diagonal in bold. Misclassifications are presented (in percentages) in corresponding rows and columns. Allopatric =allopatric black-capped chickadees, Sympatric=sympatric black-capped chickadees, and Mountain=mountain chickadees. Overall, 89.3% of original cases are classified correctly and 87.7% of cross-validated cases are classified correctly.

	LDA predicted group classification		
Actual group	Allopatric	Sympatric	Mountain
Allopatric	**91**	4	5
Sympatric	18	**76**	6
Mountain	0	4	**96**

TABLE IV. Group centroids, discriminant structure matrix, and the standardized canonical discriminant function coefficients for discriminant analysis of the acoustic measures used to classify *tseet* calls in terms of the group of birds producing each call. BC=black-capped chickadee, Allop =allopatric, Sym=sympatric, MO=mountain chickadee, TD=total duration, DD=descending duration, PF=peak frequency, EF=end frequency, FMasc =rate of rise in ascending slope, FMdesc=rate of fall in descending slope, and Fmax=frequency at loudest amplitude.

	Discriminant function	
	1	2
Group	Group centroids	
BC Allop	−2.404	−0.734
BC Sym	0.084	1.551
MO	2.320	−0.817
Variable	Discriminant structure matrix	
FMdesc	0.699	−0.147
DD	0.239	−0.204
PF	−0.216	0.155
Fmax	−0.106	−0.553
EF	0.418	−0.445
TD	0.345	−0.428
FMasc	−0.110	0.344
SF	0.008	−0.263
Variable	Standardized coefficients	
FMdesc	0.302	−0.124
DD	0.500	−0.136
PF	0.051	−0.714
Fmax	−0.388	1.059
EF	1.150	0.393
TD	0.402	−0.125
FMasc	0.302	−0.124
SF	−0.327	−0.482

TABLE V. Matrix of classification by group membership and sex of the actual groups of chickadee *tseet* calls and the LDA predicted group classification based on nine measured acoustic features. Correct LDA classifications are presented (in percentages) along the diagonal in bold. Misclassifications are presented (in percentages) in corresponding rows and columns. Allopatric=allopatric black-capped chickadees, sympatric=sympatric black-capped chickadees, and mountain=mountain chickadees. Overall, 70.7% of original cases are classified correctly and 67.7% of cross-validated cases are classified correctly.

	LDA predicted group classification					
Actual group	Male allopatric	Female allopatric	Male sympatric	Female sympatric	Male mountain	Female mountain
Male allopatric	**56**	42	2	0	0	0
Female allopatric	28	**60**	0	10	2	0
Male sympatric	0	0	**82**	10	0	8
Female sympatric	2	36	0	**62**	0	0
Male mountain	0	0	2	0	**58**	40
Female mountain	0	0	4	0	8	**88**

=23.24; male mountain, z=21.40; and female mountain, z =35.75).

IV. DISCUSSION

Bioacoustic analyses can provide insight into the particular features of vocal signals that may potentially convey species, sex, geographic origin, and individual identity. Here we use bioacoustic analyses as a first step toward characterizing temporal and spectral regularities found in chickadee *tseet* calls produced by different species, sexes, and individuals from different geographic regions. These differences in signal characters are potentially useful for discrimination

TABLE VI. Eigenvalues, canonical correlations (R_c^2), group centroids, discriminate structure matrix, and standardized canonical discriminant function coefficients for discriminant analysis of the acoustic measures used to classify *tseet* calls in terms of the group of birds producing each call. M=male, F=female, allop =allopatric black-capped chickadee, sym=sympatric chickadee, mo=mountain chickadee, FMdesc=rate of fall in descending slope, AD=ascending duration, DD=descending duration, TD=total duration, SF=start frequency, EF=end frequency, PF=peak frequency, and FMasc=rate of rise in ascending slope.

	Discriminant function				
	1	2	3	4	5
Eigenvalue	5.069	1.46	0.491	0.31	0.034
R_c^2	0.914	0.77	0.574	0.487	0.181
Group			Group centroids		
M allop	−2.931	0.629	0.526	0.145	−0.279
F allop	−2.598	0.179	0.264	−0.179	0.334
M sym	1.667	−2.056	0.647	0.432	0.003
F symp	−0.827	−0.908	−1.415	−0.084	−0.059
M mo	2.291	1.703	−0.257	0.717	0.054
F mo	2.398	0.453	0.236	−1.031	−0.053
Variable			Discriminant structure matrix		
EF	0.513	−0.496	0.047	−0.358	0.163
TD	0.307	0.637	0.061	0.606	−0.330
DD	0.168	0.335	−0.246	0.109	−0.346
AD	−0.109	0.433	−0.307	0.098	0.257
SF	0.048	0.076	0.734	−0.471	−0.131
FMasc	−0.030	−0.375	0.211	0.544	0.266
Fmax	−0.020	−0.668	0.148	−0.091	−0.350
Variable			Standardized coefficients		
EF	1.396	−0.211	−0.150	−0.173	0.502
TD	0.569	0.242	0.222	0.819	−0.528
DD	0.530	0.179	−0.199	−0.498	−0.117
AD	−0.145	0.326	−0.073	0.313	0.728
SF	0.018	0.390	1.085	−0.366	0.218
FMasc	0.154	−0.268	0.636	0.616	0.846
Fmax	−0.406	−0.614	−0.058	0.474	−0.913

among black-capped and mountain chickadees, male and female chickadees, and individual chickadees within each group.

A. Classification by individuals

The *tseet* call seems to be a reliable indicator of individual identity for most but not all of the groups in our study. For allopatric black-capped chickadees, who live in areas devoid of mountain chickadees, males *tseet* calls are more highly individualized and easier to classify than female *tseet* calls. However, for both sympatric black-capped and mountain chickadees, *tseet* calls of both males and females can be classified with high accuracy, suggesting that the *tseet* call is highly individualized in both sexes in these populations. This could perhaps be due to the different habitats of sympatric and mountain chickadees compared to allopatric chickadees. The population of allopatric chickadees were captured in the North Saskatchewan River valley in Edmonton, Alberta, while the sympatric and mountain chickadees originated from Kananaskis, Alberta, which is a more densely wooded, expansive forest compared to the Edmonton location. Perhaps individual identification of vocalizations has a higher benefit in the latter; thus the *tseet* call has evolved to be more highly individualized in both sexes in both groups of chickadees that reside in Kananaskis. An alternative theory for the high individuality of *tseet* calls in the Kananaskis population could be explained by acoustic character displacement. Acoustic character displacement is the results of a process by which features (morphological, acoustic, etc.) diverge over time so that two or more closely related species which cohabitat becomes increasingly dissimilar (Brown and Wilson, 1956). Although we would need to examine the *tseet* calls of allopatric mountain chickadees to confirm this claim, it does appear that the *tseet* calls of closely related heterospecifics are more highly individualized compared to a population of one species that live in an area of allopatry.

Overall, we see a correspondence between LDA and PIC with regard to which feature (temporal or spectral) is more unique to each individual. Of course, the next step is to test the birds in a categorization experiment, manipulating individual features or sets of acoustic features of within the *tseet* calls. Charrier and colleagues (Charrier *et al.*, 2005; Charrier and Sturdy, 2005) conducted a similar classification study on black-capped chickadees by manipulating sets of acoustic features for different note types of the *chick-a-dee* call. The performance of the birds on an acoustic discrimination task demonstrated an increase in misclassification of certain note types as they were shifted into the frequency space occupied by other note types.

Predictions yielded from the current data set would suggest that shifting the peak, end and loudest frequencies higher in mountain chickadee *tseet* calls would result in these calls being classified as *tseet* calls produced by sympatric black-capped chickadees. These types of studies could provide insight into featural evolution that could be the result of acoustic character displacement between sympatric black-capped and mountain chickadees.

B. Classification by species and geography

We were able to use the acoustic features contained in *tseet* calls to correctly, statistically, and significantly classify birds on the basis of both species and geographic region of origin with a high degree of accuracy. The best classified groups were allopatric and mountain chickadees. When allopatric black-capped chickadees were misclassified in this analysis, they were equally misclassified as mountain and sympatric black-capped chickadees. Although this is unexpected initially that allopatric black-capped chickadees would be classified as mountain chickadees, it follows if (1) both groups are thought to descend from a common chickadee ancestor and (2) neither are under any current pressure to modify their vocalizations, thus allowing the *tseet* call to remain in a more common ancestral form.

Allopatric chickadees from our study population reside in areas with limited contact with congeners—there are no mountain chickadees and few, widely dispersed boreal chickadees. In contrast to allopatric and mountain chickadees, sympatric black-capped chickadees are classified correctly with less accuracy than the other groups and are misclassified three times as often as allopatric black-capped chickadees compared to mountain chickadees. Of all three groups, sympatric chickadees are in the most constant contact with congeners; sympatric chickadees are in contact with altitudinal-migrating mountain chickadees both in the winter and also during the spring and summer, as well as boreal chickadees. We hypothesize that these differences in the amount of contact with congeners may have given rise to the differences observed in *tseet* call production.

C. Classification by species, geography, and sex

Classification of *tseet* calls according to more specific groupings of male and female for allopatric, sympatric, and mountain chickadees was correct in only 68% of cases overall. Nonetheless, the relatively poorer level of classification by the model allows us to look at misclassifications to examine subtle similarities of *tseet* calls produced by the different sexes within and between populations. Male allopatric black-capped and male mountain chickadees were most frequently misclassified as in-group (i.e., within the same species) females and rarely misclassified as out-group birds. In contrast, male sympatric chickadees were nearly equally misclassified as in-group females and female mountain chickadees. Although the classification results at the level of sex within group are less clear than those observed for individuals within group or species by location, it is too early at this point to determine whether this differentiation is simply too difficult to achieve using the current methodology or, alternatively, whether more accurately classifying by species, geography, and sex requires a larger sample of birds and their calls.

Although all nine features we measured were initially included in the LDA, a stepwise analysis revealed that only seven features contributed significantly to the model. Thus, although a linear analysis can use all acoustic features to classify calls, all features are not required. This supports the notion that some acoustic features may be more critical than

others not only in a statistical classification procedure but in the field when birds are discriminating on the basis of this vocalization.

ACKNOWLEDGMENTS

This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, an Alberta Ingenuity Fund (AIF) New Faculty Grant, a Canada Foundation for Innovation (CFI) New Opportunities Grant, along with start-up funding and CFI partner funding from the University of Alberta, Edmonton, Alberta, Canada to C.B.S., and a NSERC Discovery Grant to M.R.W.D. L.L.B. was supported by an Alberta Ingenuity Fund Studentship and E.R.B. was supported by a NSERC PGS-D scholarship. All animal studies were conducted in accordance with the Canadian Council on Animal Care Guidelines and Policies with approval from the Animal Care and Use Committee for Biosciences for the University of Alberta and the Animal Care Committee and the University of Calgary Life and Environmental Sciences Animal Care Committee. Chickadees were captured and research was carried under an Environment Canada Canadian Wildlife Service Scientific permit, Alberta Fish and Wildlife Capture and Research permits, and City of Edmonton Parks permit.

Betz, N. E. (**1987**). "Use of discriminant analysis in counselling psychology research," J. Couns. Psychol. **34**, 393–403.

Bloomfield, L. L., Charrier, I., and Sturdy, C. B. (**2004**). "Note types and coding in parid vocalizations II: The chick-a-dee call of the mountain chickadee (*Poecile gambeli*)," Can. J. Zool. **82**, 780–793.

Bloomfield, L. L., Phillmore, L. S., Weisman, R. G., and Sturdy, C. B. (**2005**). "Note types and coding in parid vocalizations III: The chick-a-dee call of the Carolina chickadee (*Poecile carolinensis*)," Can. J. Zool. **83**, 820–833.

Bloomfield, L. L., and Sturdy, C. B. (**2008**). "All 'chick-a-dee' calls are not created equally. Part I: Open ended categorization of chick-a-dee calls by sympatric and allopatric chickadees," Behav. Processes **77**, 73–86.

Brown, W. L., and Wilson, E. O. (**1956**). "Character displacement," Syst. Zool. **5**, 49–65.

Charrier, I., Bloomfield, L. L., and Sturdy, C. B. (**2004**). "Note types and coding in parid vocalizations I: The chick-a-dee call of the black-capped chickadee (*Poecile atricapillus*)," Can. J. Zool. **82**, 769–779.

Charrier, I., Lee, T. T.-Y., Bloomfield, L. L., and Sturdy, C. B. (**2005**). "Acoustic mechanisms of note-type perception in black-capped chickadee calls," J. Comp. Psychol. **119**, 371–380.

Charrier, I., and Sturdy, C. B. (**2005**). "Call-based species recognition in black-capped chickadees," Behav. Processes **70**, 271–281.

Dawson, M. R. W., Bloomfield, L. L., Charrier, I., and Sturdy, C. B. (**2006**). "Statistical classification of black-capped (*Poecile atricapillus*) and mountain chickadee (*P. Gambeli*) call notes," J. Comp. Psychol. **120**, 147–153.

Gill, F. B., Silkas, B., and Sheldon, F. H. (**2005**). "Phylogeny of titmice (Paridae) II: Species relationships based on sequences of the mitochondrial cytochrome-β gene," Auk **122**, 121–143.

Glass, G. V., and Stanely, J. C. (**1970**). *Statistical Methods in Education and Psychology* (Prentice-Hall, Englewood Cliffs, NJ).

Griffiths, R., and Double, M. C. (**1998**). "A DNA test to sex most birds," Mol. Ecol. **7**, 1071–1075.

Klecka, W. R. (**1980**). *Discriminant Analysis* (SAGE, Newbury Park, CA).

Marler, P. J., and Slabbekoorn, H. (**2004**). *Nature's Music: The Science of Birdsong* (Elsevier, Amsterdam).

Odum, E. P. (**1942**). "Annual cycle of the black-capped chickadee," Auk **59**, 499–531.

Owens, J. L., and Freeberg, T. M. (**2007**). "Variation in chick-a-dee calls of tufted titmice, *Baeolphus bicolor*: Note type and individual distinctiveness," J. Acoust. Soc. Am. **122**, 1216–1226.

Pyle, P. (**1997**). *Identification Guide to North American Birds: Part I Columbidae to Ploceidae* (Slate Creek, Bolinas, CA).

Smith, S. M. (**1991**). *The Black-Capped Chickadee: Behavioural Ecology and Natural History* (Cornell University Press, Ithaca, NY).

Sokal, R. R., and Rohlf, F. J. (**1995**). *Biometry*, 3rd ed. (Freeman, New York).

Time-frequency and advanced frequency estimation techniques for the investigation of bat echolocation calls

Yannis Kopsinis[a)]
Institute for Digital Communications, School of Engineering and Electronics, The University of Edinburgh, Alexander Graham Bell Building, King's Buildings, Edinburgh EH9 3JL, United Kingdom

Elias Aboutanios
School of Electrical Engineering and Telecommunications, The University of New South Wales Sydney, New South Wales 2052, Australia

Dean A. Waters
Institute of Integrative and Comparative Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom

Steve McLaughlin
Institute for Digital Communications, School of Engineering and Electronics, The University of Edinburgh, Alexander Graham Bell Building, King's Buildings, Edinburgh EH9 3JL, United Kingdom

(Received 31 July 2009; accepted 10 December 2009)

In this paper, techniques for time-frequency analysis and investigation of bat echolocation calls are studied. Particularly, enhanced resolution techniques are developed and/or used in this specific context for the first time. When compared to traditional time-frequency representation methods, the proposed techniques are more capable of showing previously unseen features in the structure of bat echolocation calls. It should be emphasized that although the study is focused on bat echolocation recordings, the results are more general and applicable to many other types of signal.
 [DOI: 10.1121/1.3283017]

PACS number(s): 43.80.Ka, 43.60.Hj [JAS] Pages: 1124–1134

I. INTRODUCTION

Since the first description of the ultrasonic calls made by bats[1] and evidence that they are used for echolocation,[2] various analytical techniques have been applied in order to characterize and investigate them. While the temporal characteristics of echolocation calls could be established readily even on simple oscilloscopes, the frequency characteristics proved more problematic for early technology. The fact that some bats used frequency-modulated signals going from high to low frequencies was recognized early on[3,4] and was subsequently followed by the discovery of constant frequency calls, such as those used by Rhinolophid bats.[5,6] High-speed tape recorders and zero-crossing techniques (e.g., Ref. 7) allowed basic frequency parameters, such as maximum and minimum frequencies to be determined, and permitted the classification of echolocation calls into different types.[8] Analog techniques, such as the "sound spectrograph,"[9] which was developed commercially as the Kay Electric Company sound spectrograph and first released in 1951, were then used to analyze signals in the joint time-frequency (TF) domain (e.g., Ref. 10). Detailed descriptions of call structures, however, had to wait until technology was developed to acquire signals for digital analysis along with the development of more computationally efficient techniques. Fourier techniques became more widely available with the development of the Cooley–Tukey algorithm,[11] allowing the use of the computationally more efficient fast Fourier transform. Such techniques could then be applied to the analysis of bat echolocation calls using digital computers for the purposes of understanding the structure of the call from the point of view of biosonar[12] as well as the relationship between the call structure and foraging task.[13]

Signal decomposition techniques are applied to bat echolocation calls for a number of reasons. First, they allow the analysis and interpretation of echolocation call structure with reference to species identification,[14] taxonomy,[15] task,[16] environment,[17] and target.[18] This is usually done by the visual representation of the signal, either as a power spectrum or sonogram via a short time Fourier transform (STFT). Second, the parameters extracted from the signal via decomposition may be used to understand the mechanisms by which bats may determine target range,[19] characteristics,[20] and azimuth and elevation.[21]

The Fourier spectrogram based on the STFT is the most widely used technique for analyzing time-varying signals. With respect to the inspection of echolocation bat calls particularly, apart from limited exceptions (see, for example, Refs. 22 and 23), it is the technique of choice. However, it is not the best method for TF analysis for the investigation of the frequency content of biologically produced signals.

The aim of this paper is two-fold. It first serves as a guide for those involved in the analysis of bat echolocation calls to facilitate the correct interpretation of the frequency content of such signals. Second, traditional techniques for

a)Author to whom correspondence should be addressed. Electronic mail: kopsinis@ieee.org

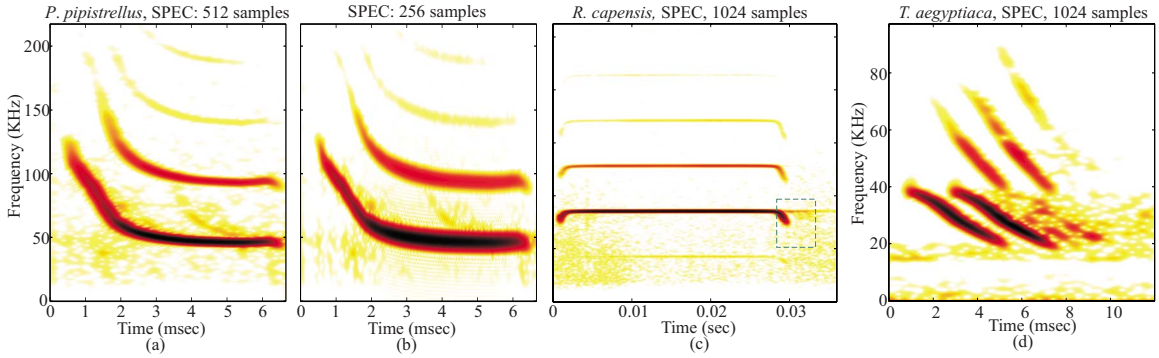


FIG. 1. (Color online) [(a) and (b)] Spectrograms of echolocation calls produced by a *P. pipistrellus* bat using 512 and 256 samples long Gaussian analysis windows, respectively. [(c) and (d)] Spectrograms of a *R. capensis* and a *T. aegyptiaca* bat call, respectively.

the analysis of bat echolocation calls are compared with some less well known methods offering certain advantages. It should be emphasized that although the study is focused on bat echolocation types of signal, the results are more general and applicable to other types of signals. Matlab scripts of the methods developed in this paper are freely available at [http://www.see.ed.ac.uk/~ykopsini/tf.htm (last viewed Jan. 14, 2010)].

II. DATA COLLECTION METHOD

The echolocation calls analyzed and discussed in this paper are field recordings collected during 2007 and 2008. The recordings of *Pipistrellus pipistrellus* and *Myotis daubentonii* took place in the UK, while those of *Rhinolophus capensis* and *Tadarida aegyptiaca* were made in South Africa. In all cases, the bats were recorded in free flight while foraging with the microphone mounted on a pole with a maximum height of 3 m to better align it with the foraging bat. Calls were obtained from bats flying free from clutter to remove interfering echoes, and with the heterodyned audio output of the bat detector switched off to eliminate any low frequency feedback through the microphone and consequent low frequency artifacts. In all cases, a Pettersson D-1000x bat detector with a solid dielectric microphone was used, and the calls were sampled at 750 kHz and digitized with 16 bit resolution, streamed to a compact flash card in the device, and saved in the ''.wav'' file format standard. These calls were then sorted visually based on the sonograms generated in BATSOUND PRO (Pettersson Electronic, Sweden). Species identification was made with reference to distributions and call information in the literature and in the event of any ambiguity in identification, or any evidence of clipping, the affected call was not selected for analysis.

III. SPECTROGRAM

The spectrogram, which is essentially the squared modulus of the STFT,[24,25] is one of the most widespread tools for time-frequency analysis of time-varying signals.

The STFT of a signal $x(t)$ is given by

$$S_x^{(h)}(t,f) = \int x(u)h^*(u-t)e^{-j2\pi fu}du, \tag{1}$$

where $h^*(t)$ is the complex conjugate of an analysis window localized around $t=0$. The superscript (h) indicates the dependence of the STFT analysis on the characteristics, especially the length, of the analysis window used. If $h(t)$ were omitted, or equivalently set to 1 $\forall t$, then the STFT reduces to the Fourier transform. Otherwise, $h(u-t)$ suppresses the signal outside an area around the time instance t. In simple words, allowing the parameter t to take on all the values within the time interval where the signal evolves centers the analysis window on t and permits us to estimate via the Fourier transform the frequency content of the signal in the neighborhood of t. Finally, the spectrogram, $\mathrm{SPEC}_x^{(h)}(t,f) = |S_x^{(h)}(t,f)|^2$, can be loosely interpreted as the energy distribution of the signal over the TF plane.

A. Spectrogram time-frequency resolution

According to Eq. (1) if a short analysis window, $h(t)$, is used, then only a short interval of the signal centered on t enters into the calculation of the STFT and the result of $S_x^{(h)}(t,f)$ is well localized around t. Otherwise, if $h(t)$ is large, then the outcome of the STFT is influenced by a larger portion of the signal, i.e., a longer time interval around the analysis time point t, leading to low time resolution in the sense that all signal components having time separation shorter than the analysis window duration will contribute simultaneously to the estimate of the signal energy at the associated TF points, rendering them indistinguishable.

On the other hand, the bandwidth of the analysis window is responsible for the frequency resolution achieved. Specifically, the smaller the bandwidth of the analysis window is, the better the frequency localization becomes providing improved frequency resolution.

Unfortunately, analysis windows having arbitrarily short length and small bandwidth at the same time cannot be constructed due to the Heisenberg uncertainty principle, leading to a fundamental resolution trade-off: Using a window with a short duration gives a high time resolution at the expense of a compromised frequency resolution and vice versa.

B. Spectrograms of echolocation bat calls

Figures 1(a) and 1(b) show the spectrograms[26] of echolocation calls belonging to *Pipistrellus pipistrellus* bat. We observe that this is a multiharmonic signal with the fundamental frequency component evolving from about 130 to 50 kHz. Moreover, it has a characteristic nonlinear chirp shape consisting of two distinct parts. The first one is a fast time-varying down-chirp (from 0.5 to 2 ms) and the second one, corresponding to the rest of the signal, has a nearly constant frequency profile. In Fig. 1(a) the STFT uses a Gaussian analysis window that is 512 samples long, corresponding to about 0.68 ms time duration. This is a relatively long window that provides high frequency but low time resolution. The result is a good concentration of the spectrogram energy around the frequencies that correspond to the constant frequency part of the signal which appears thin compared to the down-chirp part. Indeed, along the signal sections where the frequency content is not changing rapidly with time, the requirement for high time resolution becomes a luxury allowing the use of a long analysis window. On the other hand, when dealing with fast time-varying signals (or signal sections), the adoption of long analysis windows can be problematic particularly if the signal changes significantly during the window duration. In particular, improvement in the down-chirp frequency concentration, both in the fundamental component and the harmonics, can be achieved using a sorter, 256 samples long, analysis window, as is shown in Fig. 1(b). Inevitably, the shorter window has a larger bandwidth leading to poor frequency localization and consequently to a wide spectrogram around the actual frequencies of the constant frequency part of the signal.

The second echolocation call examined here belongs to a *Rhinolophus capensis* bat and is a multiharmonic constant frequency signal having characteristic short down-chirp tails. Its spectrogram, using a 1024 samples long analysis window, is shown in Fig. 1(c). In this case, a long analysis window might be preferred since it provides more accurate estimates of the actual frequencies of the larger part of the signal. However, the exact time instances that the signal appears and disappears cannot be accurately determined by the spectrogram due to the corresponding low time resolution. In a similar fashion, the down-chirp tails diverge from their actual shape.

The ability of STFT to analyze a signal that is corrupted by noise is investigated next using an echolocation call signal from *Tadarida aegyptiaca* [illustrated in Fig. 1(d)]. The actual recording used here presents an additional challenge as it contains the echolocation calls of two bats emitted with a time difference of just 1 ms. The energy concentration of the specific signal cannot be improved by using a different window. This is due to the slope that the chirps have.

Based on these examples, we can infer that the spectrogram via the STFT cannot serve as a generalised analysis tool for all types of bat calls. Moreover, as we saw in the case of the nonlinear chirp, the results are limited since a particular analysis window is not appropriate for the entire signal. However, there are signal shapes that lend themselves to a sufficiently accurate interpretation of their spectrograms provided a suitable window length is used.

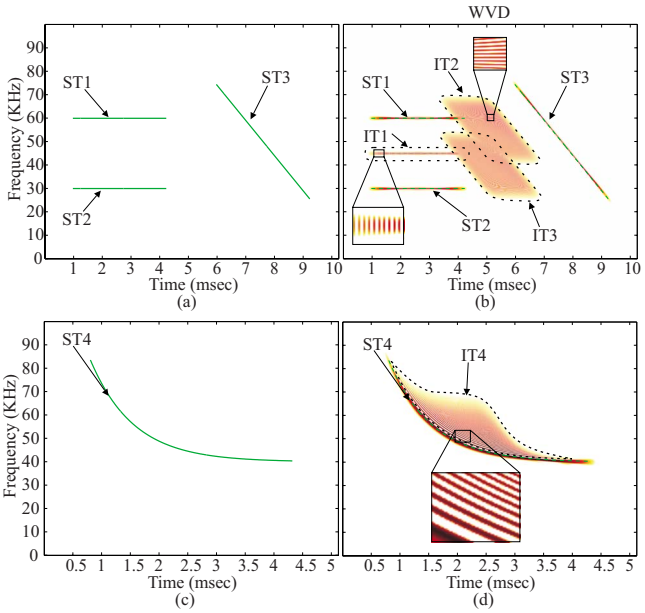


FIG. 2. (Color online) (a) Test signal comprising two constant frequency harmonics and a linear chirp. (b) The corresponding WVD. [(c) and (d)] Nonlinear chirp test signal and the corresponding WVD, respectively.

IV. A BROADER CLASS OF QUADRATIC TIME-FREQUENCY REPRESENTATIONS

The spectrogram, discussed above, belongs to the class of energy TF distributions given by the square modulus of the respective linear transformation, namely the STFT. An alternative road to TF representations, which provides greater flexibility and in many cases better performance, is to directly estimate the signal energy on the TF plane.[27,28] A prominent example of this is the Wigner–Ville distribution (WVD):

$$\mathrm{WV}_x(t,f) = \int x\left(t+\frac{u}{2}\right)x^*\left(t-\frac{u}{2}\right)e^{-j2\pi fu}du. \qquad (2)$$

For the sake of clarity and simplicity, the WVD study is initially based on two artificial signals shown in Figs. 2(a) and 2(c). The first one is a combination of a constant frequency (CF) and a linear chirp signal, whereas the second one is a nonlinear chirp signal. The WVD estimate corresponding to Fig. 2(a) is shown in Fig. 2(b). Two immediate observations can be made. First, the energy around each of the actual signal components, indicated by ST1, ST2, and ST3, is extremely well concentrated both in time and frequency, significantly outperforming the spectrogram.[29] Indeed, the WVD offers exact localization of linear chirps.[30] Second, there are areas, clustered into three groups IT1, IT2, and IT3, that falsely show high energy. These are the result of interference terms (ITs) that the WVD inherently produces and are misleading with respect to the readability of the TF representation. It is pivotal to comprehend the generation mechanism and characteristics of the WVD artifacts both for enhancing our ability to correctly interpret the TF representation of an unknown signal and for introducing techniques for interference reduction. As an example consider a signal x

that is the sum of two distinct components x_1 and x_2, having energy concentrations around the points (t_1,f_1) and (t_2,f_2), respectively. Then, it is straightforward to show that $\mathrm{WV}_x(t,f)=\mathrm{WV}_{x_1}(t,f)+\mathrm{WV}_{x_2}(t,f)+2\,\mathrm{Re}\{\mathrm{WV}_{x_1x_2}(t,f)\}$, where the operator Re indicates the real part. The first two terms are the signal terms (STs) that accurately depict the energies of the two signal components on the TF plane, whereas the third term is a cross-product IT that appears as an additional, false, component. The IT is due to the interaction of the two signal components and is located exactly at the center of the line segment connecting the points (t_1,f_1) and (t_2,f_2).

When the signal under consideration is more complex, such as in Fig. 2(a), then all of the signal components interact with one another to produce numerous ITs. Consequently, the interference term IT1 is the result of the interaction between the CF harmonics ST1 and ST2 and is found midway between these two signal terms. Similarly, IT2 is the result of the interaction between ST3 and ST1. The width of the areas of these ITs is related to the relative distance between the points of ST3 and ST1. For example, the top-right (top-left) part of IT2 is due to the interaction of the rightmost (leftmost) part of ST1 with the topmost part of ST3. Similarly, IT3 is generated from the interactions between ST2 and ST3.

Figure 2(d) shows the WVD of the nonlinear chirp of Fig. 2(c). The localization of the nonlinear chirp is worse than that of the linear chirp but is superior to the spectrogram. Also, the observed ITs are not the result of interactions between signal parts belonging to different components, but are caused by different parts of the same signal component. For this reason they are often referred to as *inner interferences* to differentiate them from the *outer interferences* of Fig. 2(b). However, a clear distinction between the two types of interference terms does not exist since they are both generated by the same mechanism.

In the case of digital signals (which is true here), interferences between the signal components lying at the negative and positive frequencies might appear. For that reason, the WVD is applied on the analytic signal version of the signal under consideration.[30]

A. WVD interference reduction

An important feature of the ITs is their sinusoidal oscillatory structure, evident in the magnified insets shown in Figs. 2(b) and 2(d). We observe that the direction of the oscillations is perpendicular to the line connecting the interacting components [e.g., ST1 and ST2 or ST1 and ST3 in Fig. 2(b)]. Moreover, the frequency of the oscillations is distance dependent and the further apart the signal components are, the faster the oscillations become. This can be seen in the magnified inset in Fig. 2(d). Indeed, we see that the frequency of the oscillations changes as the ITs move away from the signal term, with the reason being that the distance between the signal parts that produce the distant ITs is greater.

The oscillatory nature of ITs can be exploited in order to eliminate them by means of smoothing. According to this framework, the IT smoothing is realized along the time and the frequency axis separately using two analysis window functions $g(t)$ and $h(t)$, respectively. The resulting TF representation is referred to as smoothed pseudo-WV distribution (SPWVD) which is given by

$$\mathrm{SPWV}_x(t,f)=\int h'(u)\left\{\int g(\tau-t)x\left(\tau+\frac{u}{2}\right)\times x^*\left(\tau-\frac{u}{2}\right)d\tau\right\}e^{-j2\pi fu}du,$$

where $h'(u)=h(\frac{1}{2}u)h^*(-\frac{1}{2}u)$.

The window $g(t)$ can reduce ITs evolving parallel to the time axis. The required length depends on the oscillation frequency and should cover at least a full oscillation period. In other words, the further apart two interacting components are the shorter $g(t)$ should be. On the other hand, $h(t)$ is capable of reducing ITs evolving parallel to the frequency axis. In contrast to $g(t)$, the shorter $h(t)$ is, the more intense the smoothing along frequency axis becomes. The reason is that $h(t)$ does not contribute directly to the overall smoothing function but via the Fourier transform of $h'(t)$.

Unfortunately, apart from IT elimination, the smoothing operation reduces the concentration of the actual signal components. As a result, there is always a trade-off between the time-frequency resolution achieved and the amount of ITs that will be eliminated through smoothing. In any case, the shorter $g(t)$ and the longer $h(t)$ that effectively reduces the ITs to an accepted for interpretation degree.

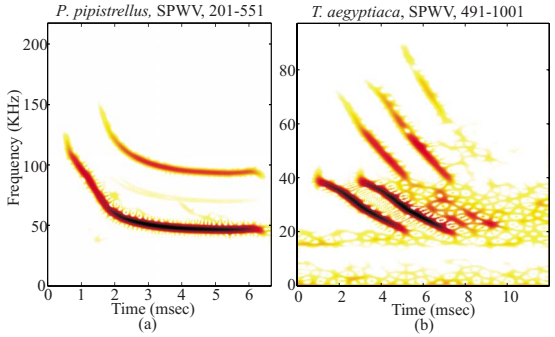


FIG. 3. (Color online) SPWVD TF representations of echolocation calls produced by *P. pipistrellus* and *T. aegyptiaca* bats.

Figure 3 shows the TF representation of the SPWVD of bat signals analyzed in Sec. III. In the case of *P. pipistrellus* echolocation call [Fig. 3(a)], the SPWVD yields a well-concentrated TF representation of the nonlinear chirp, especially in the down-chirp parts where the signal energy is strong enough. At the same time, the low amplitude signal harmonics are missed. The window functions $g(t)$ and $h(t)$ are Gaussian with lengths 201 and 551 samples, respectively. Observe that there are some ITs between the harmonic components. These could be completely removed by adopting a longer window $g(t)$, say, 301 samples long. However, this would lead to reduced concentration of the actual signal components.

Figure 3(b) shows the SPWVD of *T. aegyptiaca* using $g(t)$ and $h(t)$ 491 and 1001 long, respectively. We observe that the resulting TF representation is significantly more concentrated around the actual frequencies than the spectrogram

depicted in Fig. 1(d). Moreover, both methods have similar sensitivity to noise but in the case of the SPWVD, the noise is shown in a less compact way, forming cell-like structures. This, in some cases, can favor better readability.

With respect to the *R. capensis* echolocation call whose spectrogram was shown in Fig. 1(c), the corresponding SPWVD does not offer any advantage so it is not shown.

V. HIGH RESOLUTION SPECTRUM TECHNIQUES

In this section we present an alternative adaptive filterbank approach[31] called the minimum variance estimator (MVE). The adaptation of a bank of filters to the signal leads to a significantly higher resolution compared to the STFT, albeit at a higher computational load.

Assuming that the signal is stationary or almost stationary over a segment of length L then the MVE aims to estimate the energy contribution that a specific frequency f makes to that signal. This is done by designing a finite impulse response (FIR) filter of length M that passes the signal component at f undistorted while attenuating the rest of the frequencies as much as possible. Let us denote by $\mathbf{h}_f$ the coefficients vector of such a FIR filter, then the output of the convolution of the filter with the L samples long signal segment yields $K=L-M+1$ samples given by

$$y_k(f) = \mathbf{h}_f^H \mathbf{x}_k, \quad \text{for } k = 1, \ldots, K, \tag{3}$$

where $\mathbf{x}_k=[x_k, x_{k+1}, \ldots, x_{k+M-1}]^T$. The mean power over the K output samples is given by

$$\bar{P}(f) = \frac{1}{K}\sum_{k=1}^{K} |y_k(f)|^2 = \mathbf{h}_f^H \mathbf{R} \mathbf{h}_f, \tag{4}$$

where $\mathbf{R}=(1/K)\Sigma_{k=1}^{k}\mathbf{x}_k\mathbf{x}_k^H$ is the $M\times M$ sample covariance matrix of the input data. The design problem can be expressed as the following constrained minimization:

$$\min_{\mathbf{h}_f} \mathbf{h}_f^H \mathbf{R} \mathbf{h}_f \quad \text{such that } \mathbf{h}_f^H \mathbf{s}(f) = 1, \tag{5}$$

with $\mathbf{s}(f)=[1, e^{j2\pi f}, e^{j4\pi f}, \ldots, e^{j2(M-1)\pi f}]^T$ being the template at the frequency of interest. It can be shown that the solution to this problem is a filter of the form[31]

$$\mathbf{h}_f = \frac{\mathbf{R}^{-1}\mathbf{s}(f)}{\mathbf{s}^H(f)\mathbf{R}^{-1}\mathbf{s}(f)}. \tag{6}$$

The MVE filter then strikes the best balance between the attenuation and the main lobe width.[32] Using a bank of these filters, with each one tuned to a different frequency on a specified grid, we can obtain a sampling of the overall frequency content of the signal on that grid.

The frequency resolution achieved by the MVE is determined by the filter length M. A longer filter would exhibit a narrower main lobe and higher sidelobe attenuation, yielding a higher frequency resolution. On the other hand, the reliability of the sample covariance matrix of the input data and consequently the $\bar{P}(f)$ estimates is given by the amount of averaging used (which is set by K). Therefore, for a fixed L, there is an inherent compromise between the resolution and reliability of the spectral estimate. It should be emphasized that, in order to ensure that the matrix $\mathbf{R}$ is invertible, the minimum allowable value of K is $M+1$. Equivalently, this translates to a total data segment $L\geq 2M$. Note that since $L=M+K-1$, a large M and/or K requires the availability of a long enough signal segment.

A. Implementation of the MVE and performance trade-offs

Consider the general case where we have available N data samples of a non-stationary signal, and we seek to estimate its TF spectral distribution. Then, similar to the STFT, the signal can be partitioned into a number of segments of length L, over which the signal can be assumed to be stationary or nearly stationary, and the MVE applied to each of those signal segments.

As pointed out in Sec. III, the implementation of the MVE spectrum estimator presents a number of compromises in setting the various inter-related parameters. At this point we introduce an extra parameter that facilitates the discussion. Let α specify the number of samples that L exceeds its minimum allowable value, that is, $L=2M+\alpha$. Equivalently, we have $K=M+1+\alpha$. In order to specify the MVE algorithm, any two parameters among L, M, K, and α need to be set. Here we propose fixing the parameters M and α.

The larger that α is, the smoother the TF representation looks. Thus, α should be large enough in order for the MVE estimates to be reliable. We have found that a value in the interval [30, 100] is acceptable. With respect to M, it is desirable to make it as large as possible in order to achieve high frequency resolution. However, this necessitates a large L which may violate the stationarity requirement. In fact, L must be kept small enough for the stationarity assumption to be, at least approximately, maintained. In the case of a chirp, for instance, the faster the rate of change in the chirp frequency, the more severe the limit on L (and therefore on M) is. In other words, although a longer filter would lead to improved spectral resolution, the signal non-stationarity would negate this advantage as it blurs the spectrum itself. Having set M and α, the resulting length of the signal segments can be computed from $L=2M+\alpha$.

When segmenting a data record of N samples into the blocks of length L, the same approach to the STFT can be used. That is, one can either use non-overlapping blocks or alternatively allow them to overlap. In the first case, the total number of segments will be $N_t=\lfloor N/L \rfloor$, where $\lfloor\ \rfloor$ denotes the floor function. Thus, N_t determines the time points of the TF plane where the spectral estimates are computed. In practice, however, the total signal length N is usually not very large, and the readability of the TF representation is enhanced by overlapping the blocks. If the segments overlap by Δ samples, the number of blocks N_t becomes

$$N_t = \left\lfloor \frac{N-L}{L-\Delta} \right\rfloor. \tag{7}$$

From this, the value of the overlap Δ can then be calculated to give the desired number of time points N_t. Note that whereas the sampling of the time axis is set by N_t, that of the frequency axis can be made arbitrarily fine by choosing a denser frequency grid. This is akin to the zero padding of the

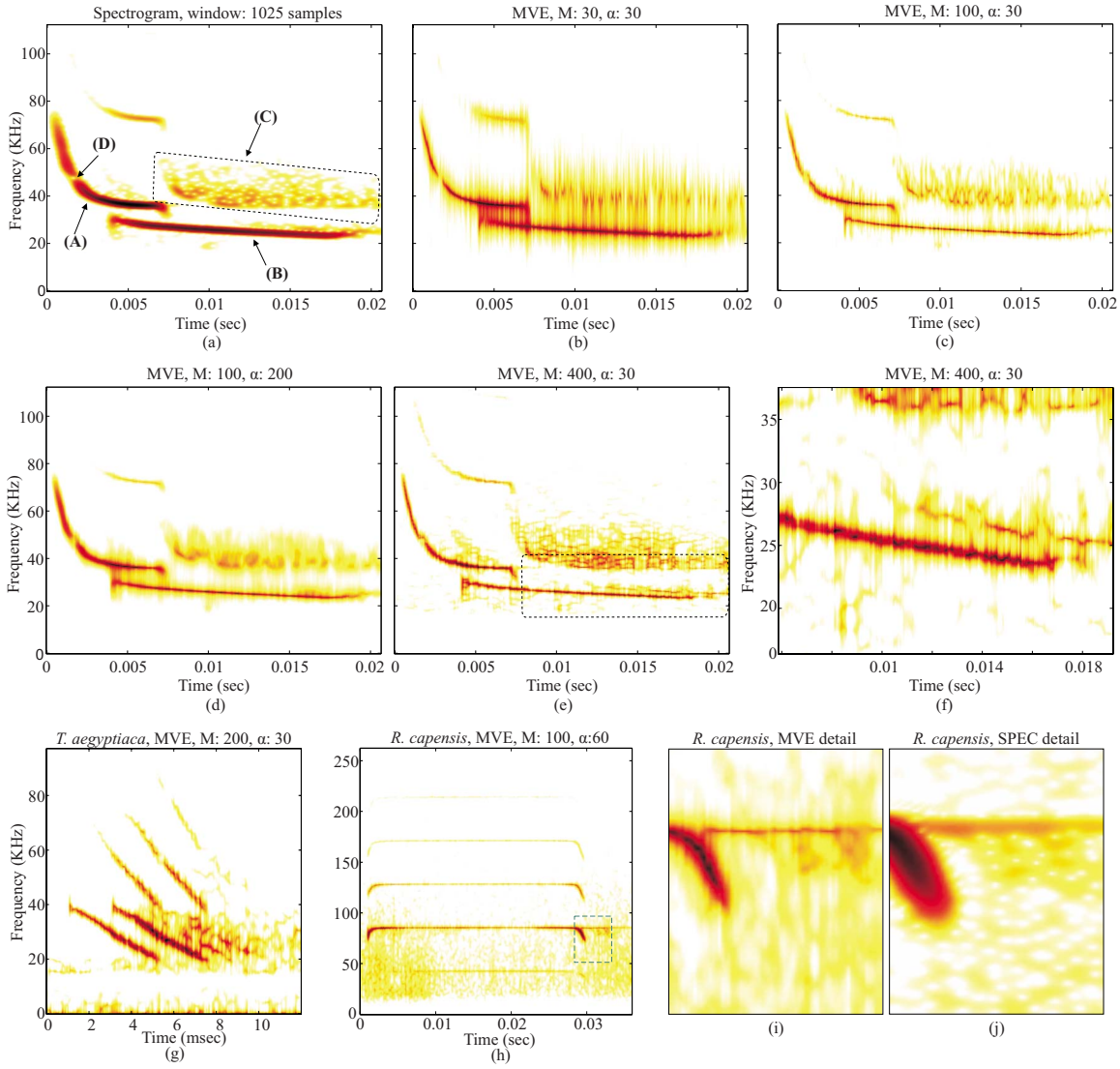


FIG. 4. (Color online) Study of MVE on several echolocation calls.

Discrete Fourier Transform (DFT) and comes at the expense of an increased computational load.

B. MVE application on actual echolocation calls

A number of examples of MVE-based TF representations are shown in Fig. 4. For comparison purposes, the spectrogram of the first echolocation call studied with MVE is given in Fig. 4(a). This signal consists of three distinct components that originated from different sources. First, present is a call denoted by (A) together with one relatively weak harmonic belonging to a *P. capensis*. Second, another bat call from *T. aegyptiaca* is denoted by (B). And finally, the return of the strong harmonic of call (A) is designated by (C). Moreover, we observe that the down-chirp part of (A) exhibits a frequency null (D) that is not well localized by the spectrogram in either the time or frequency. Figures 4(b) and 4(c) present the outputs of the MVE with two filters of lengths M equal to 30 and 100, respectively, with $\alpha=30$ in both cases. Note that the filter with $M=30$ is not long enough to adequately attenuate the interference. As a result the TF representation looks spiky with many disturbing interference lines between strong signal components. However, the frequency null is accurately reproduced in its full extent. Setting M to 100 enhances the frequency resolution and leads to a much clearer and well-concentrated TF representation. Although there are still some artifacts especially between the two strong components (A) and (B), the accurate interpretation of the frequency content of the signal is not difficult. Additionally, the echolocation call return (C) is shown with high detail. However, the visibility of the frequency null is somewhat degraded by the longer filter used which then spans the duration of the null. Now recall that α determines the reliability (smoothness) of the TF representation. Using $\alpha=30$ guarantees enough reliability (small variance of the magnitude estimates). However, if a still smoother TF representation is desired, α can be increased, to say 200, as shown in Fig. 4(d). Increasing the filter length to $M=400$, a much greater frequency resolution can be achieved, as is evident in Fig. 4(e). The consequences, however, are threefold. First, the ability of the filter to attenuate interfering frequencies is enhanced, leading to a clear distinction between strong neighboring signal components, e.g., (A) and (B). Also, com-

ponents that are much closer together can now be distinguished. An example of this is given in the area marked by the dotted rectangle, which is magnified and displayed in Fig. 4(f). An additional signal component slightly above component (B) is now visible. In fact, this is a delayed version of (B) resulting from its reflection by an object. Its delay with respect to (B) is similar to the delay of the return call (C) from its originating signal (A). This leads us to infer that the two bats were likely close together and their calls reflected from the same object. A third, negative consequence of the use of a longer filter is a relative loss of time resolution especially in the return (C), leading to fairly elongated signal parts in the low amplitude components. However, this does not have a significant effect on the strong components [such as the down-chirp part of (A)] apart from the frequency null area which seems quite distorted. For completeness, the MVE TF representations of some of the echolocation calls investigated before are also given in Figs. 4(g) and 4(h). The MVE TF distribution of the *T. aegyptiaca* echolocation call appears somewhat sharper than those obtained from the spectrogram and SPWVD. In the *R. capensis* call case, we see a more concentrated localization of the signal components at the expense of a relatively poor accuracy in the power evaluation of all the harmonics but the strongest one. This is a disadvantage of MVE since it is known[33] that it is a biased estimator of the amplitude. However, the high resolution achieved can be helpful in interpreting and extracting valuable information from the TF representation. Indeed, the detail revealed by the close up, shown in Fig. 4(i), of the echolocation call "tail" [marked by the dotted rectangle in Fig. 4(h)] indicates that a return of the strong harmonic has been captured in the specific recording. Although this conclusion could be drawn from the corresponding spectrogram shown in Fig. 4(j), the MVE provides more details, allowing more accurate statements to be made about the nature and number of reflecting objects.

VI. SPECTRAL ANALYSIS BASED ON INSTANTANEOUS FREQUENCY ESTIMATES

Although the notion of instantaneous frequency (IF) was introduced decades ago (see, e.g., Ref. 27) there is still a lot of controversy regarding a proper definition as well as a physical interpretation for it.[34–36] Perhaps, the most widely accepted definition of the IF is as the derivative of the phase of the signal. More specifically, for a signal $x(t) = a(t)\cos(\phi(t))$ the IF in hertz is given by

$$\mathrm{IF}(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt}, \tag{8}$$

where $\phi(t)$ is the phase and $a(t)$ the instantaneous amplitude (IA) of the signal. Now it is impossible to compute two unknowns, namely, the amplitude and phase, from the observed real signal $x(t)$.[37] Thus, there have been numerous attempts to solve this issue and for a review and comparison of these corresponding methods, we refer the interested reader to Refs. 38–42.

In this study, we adopt the analytic signal approach[37] according to which the signal, $x(t)$, under examination can be written in complex form as

$$w(t) = x(t) + i\mathcal{H}[x(t)] = x(t) + iy(t), \tag{9}$$

where $\mathcal{H}[x(t)]$ is the Hilbert transform[43] of $x(t)$ and i is the imaginary unit. The specific complex representation $w(t)$ is analytic in the sense that it contains the full spectral content of $x(t)$ solely in the positive frequencies. Accordingly, the IA and phase can unambiguously be derived as

$$a(t) = \sqrt{x^2(t) + y^2(t)} \tag{10}$$

and

$$\phi(t) = \arctan\left(\frac{y(t)}{x(t)}\right). \tag{11}$$

In order to efficiently use the IF for the study of echolocation calls we need first to understand its properties when dealing with different signals. It should be stated that IF estimation does not offer an alternative technique for constructing a TF representation of a signal. In other words, there is only a limited number of cases that the IF coincides with the actual frequency content of a signal, and in general, the IF at a time t assumes values that do not correspond to any of the frequencies of the signal components revealed by the TF representation to constitute the signal at t. Nonetheless, the IF can be a useful tool for echolocation call study provided that the ability to correctly interpret it is first developed.

A. Interpretation and properties of IF

The behavior of the IF of a signal depends strongly on whether the signal is monocomponent or multicomponent. By monocomponent we refer to a signal that, at any time instant, contains only a single frequency. In a broader, more relaxed sense, a signal can be characterized as monocomponent if it is highly narrow-band locally. With respect to echolocation calls, a monocomponent signal does not contain harmonics. In contrast, an echolocation call comprised of a number of harmonics is a multicomponent signal since it contains at any time instant more than one frequency. Moreover, multiple overlapped echolocation calls would make up a multicomponent signal regardless of whether they contain harmonics or not. Note that the characterization of a signal as multicomponent or monocomponent applies locally in time and does not refer to full length of the data record. In other words a signal can be monocomponent or multicomponent at various times as components of different frequencies appear and/or disappear.

Once the IF and IA of a signal have been computed, they can be combined in order to form a spectrogram-like plot $H(t,f)$ which, for any time instant t, is the IA, $a(t)$, of the signal at the frequency f which is active, determined by IF(t), and zero at the rest of the frequencies. $H(t,f)$ is usually referred to as Hilbert spectrum.[44]

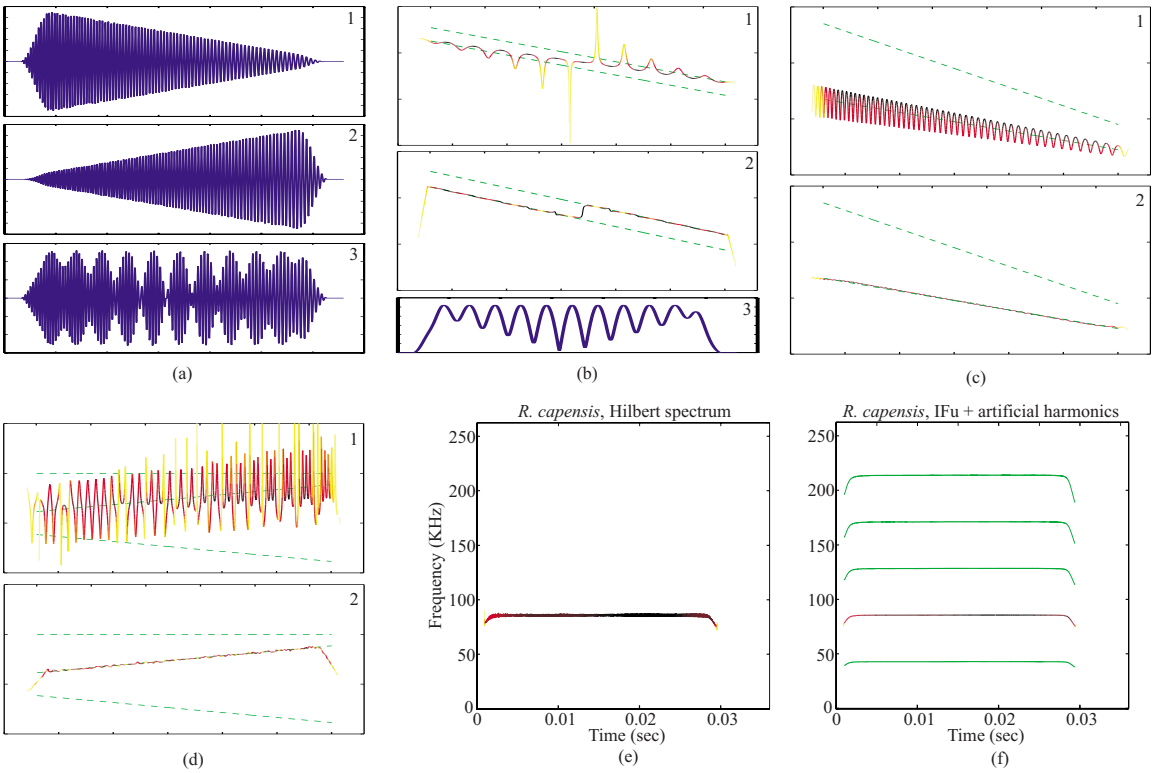


FIG. 5. (Color online) [(a)–(d)] Analysis of multicomponent signal-based IF estimates. [(e) and (f)] IF-based analysis of a *R. capensis* echolocation call.

B. IF of monocomponent and multicomponent signals

When applied to a noiseless monocomponent signal, the IF provides exact estimation of the frequency content of the signal. This is a very useful property in cases of echolocation calls that have no, or extremely attenuated, harmonics. Unfortunately, the situation is much more complicated in the case of multicomponent signals.

The behavior of the IF of multicomponent signals is investigated first using the two-component signal depicted in Fig. 5(a). In particular, subfigures 1 and 2 show these two components individually while subfigure 3 displays their sum. They are chosen so that the amplitude of the first component decreases with time whereas that of the second component increases. The Hilbert spectrum of the signal is shown in Fig. 5(b1), together with the actual frequency tracks of the two components (indicated by dashed lines). The tracks represent parallel linear chirps, the lower of which belongs to the first component [shown in Fig. 5(a1)]. The resulting IF plot exhibits a spiky form that clearly does not match the actual frequency content of the signal. That is to be expected since, at any time instant, the rate of change in the phase (which gives the IF) includes the effects of both instantaneous frequencies and it is not possible to decouple those two effects (in order to resolve the two tracks) using a single function estimating the IF. We will now briefly discuss some properties of the IF that are useful for the types of signals that are of interest to us in this study. Proofs and/or further details can be found in Refs. 34 and 45–48. The following are a number of properties of the IF that can be observed in Fig. 5(b).

(1) The IF is a non-symmetric oscillation exhibiting spikes that point downward or upward. More specifically, these spikes point toward the component with the larger amplitude. Therefore, in the first half of the plot, the lower frequency chirp has the larger amplitude [see Fig. 5(a1)] and consequently the spikes are directed down. The opposite is true for the second half of the signal.

(2) Smoothing of the IF by applying moving average with a long enough analysis frame T allows improved estimates of the frequency of the component with the larger amplitude to be obtained. Let the unweighted average[49] of the IF(t) resulting from the smoothing procedure be denoted by $\mathrm{IF}_u(t)$. Then

$$\mathrm{IF}_u(t) = \frac{1}{T}\int_{t-T/2}^{t+T/2} \mathrm{IF}(u)du. \tag{12}$$

The frequency estimate, $\mathrm{IF}_u(t)$, is shown in Fig. 5(b2). As expected, it locks on the first component during the first half of the signal duration (since that component has the larger amplitude) and then picks up the higher frequency component during the second half (as it now stronger). This property is quite useful and enables the estimation of the fundamental frequency of the echolocation bat calls as it usually is the stronger component of the signal.

(3) The strength of the IF oscillations (that is the height of the spikes) depends on the relative amplitudes of the components. The closer the amplitudes are to each other, the larger the spikes become. On the other hand, when one component dominates the other (see, for instance, the beginning and end sections of the signal), then the

spikes are small.

Besides the properties of the IF listed above, we mention that the instantaneous amplitude estimate, Eq. (10), matches the envelope of the signal, as shown in Fig. 5(b3).

Some additional properties of the IF can also be gleaned from another example, shown in Fig. 5(c1), of a multicomponent signal consisting of a linear down-chirp with a lower amplitude harmonic. Unlike the previous example, the amplitudes of both the fundamental and harmonic are constant in time, whereas the difference (or separation) between their frequencies changes with time, which is due to the fact that they are harmonics. Now we make the following observations.

(4) The period of the oscillations of the IF depends on the difference between the frequencies of the two components. In fact, the larger this difference is, the faster the oscillations become.
(5) As pointed out in property 3, the amplitude of the oscillations (hight of the spikes) depends on the relative amplitudes of the components. Additionally, in the current example we see that the amplitude of the oscillations depends also on the difference in the frequencies, becoming larger as this difference increases.

When more than two signal components come into play, a more complicated Hilbert spectrum is generated, as shown in Fig. 5(d1). Once again the actual frequencies of the components are represented by the dashed lines, and their amplitudes are constant and equal to 0.4, 1, and 0.5, respectively, starting from the higher to the lower frequency. Most of the properties discussed above can be adapted to the case of three or more components. Note that the unweighted average is still valid and can be directly applied, as shown in Fig. 5(d2). Indeed, the middle component is the dominant one here. However, we should emphasize that as the number of components increases, the dominant component of the signal has to be much stronger relative to the rest of the components in order for the $\mathrm{IF}_u(t)$ estimate to be accurate.

C. Echolocation data study

The first real echolocation call that we study with the IF method is an *R. capensis* signal. The spectrogram and MVE TF representations of this signal were presented in Figs. 1(d) and 4(h), respectively, and the Hilbert spectrum was given in Fig. 5(e). The best the IF can achieve is to estimate a single frequency component. The result is a well-concentrated estimate of the highest amplitude component, which is the second lowest frequency component. In fact, the width of the curve is the result of the oscillations emanating from the presence of the other harmonics. However, as the amplitudes of the remaining components of the signal are very low relative to the dominant one, the resulting estimate is quite accurate. This is a common scenario in bat echolocation calls. When the unweighted average is applied, the frequency estimate is forced to lock accurately on the dominant frequency component, as shown by the second lowest component in Fig. 5(f). Moreover, if it were known that the estimated component is the second harmonic, we can use the unweighted average estimate to obtain the frequency tracks of other potential harmonics that may or may not be contained in the signal.

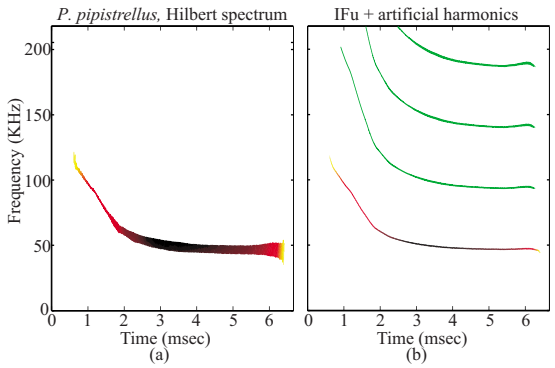


FIG. 6. (Color online) IF-based analysis of a *P. pipistrellus* echolocation call.

Next, we consider the application of the IF method to the *P. pipistrellus* echolocation call studied with previous methods. The resulting Hilbert spectrum is illustrated in Fig. 6(a). We observe that the frequency track is well concentrated in the down-chirp part but exhibits some very fast oscillation in the region of constant frequency. Again we stress that the degree of concentration of the IF estimates around a particular frequency is independent of the time-frequency resolution of the method. Instead, it depends only on the amplitudes of other signal components present at each time instant. Therefore, we infer from the observed frequency trajectory that the harmonics are largely attenuated during the down-chirp part, something that is confirmed by the related spectrogram shown in Fig. 1. This implies that the down-chirp part of the signal is effectively monocomponent. In the section of constant frequency, on the other hand, the harmonics are stronger leading to a coarser IF estimate. Again a more accurate estimate of the highest amplitude signal component can be computed using the unweighted average (especially as the harmonics are significantly weaker than the strongest component). This result, together with its potential harmonics, is depicted in Fig. 6(b).

Apart from harmonics estimation, IF-based analysis can also prove useful in cases that other TF representation methods do not provide results that are easy to interpret. Take as an example an echolocation call produced by a *Myotis daubentoni* bat shown in Fig. 7(a). These specific bat calls often appear to have a special amplitude modulation. The spectrogram of this call is displayed in Fig. 7(b) and consists of a slightly nonlinear chirp that looks somewhat different than usual. It is wider than expected and has energy nulls along its length. With the aid of the Hilbert spectrum, useful information about both the spectral shape and the amplitude modulation can be drawn. As can be seen in Fig. 7(c) the IF exhibits spikes similar to those appearing when two signal components having closely spaced frequencies are active simultaneously [e.g., Fig. 5(b)]. Moreover, due to the fact that the spikes point downward, we can conclude that the signal component having the higher amplitude also has the lower frequency. This combination of two closed in frequency

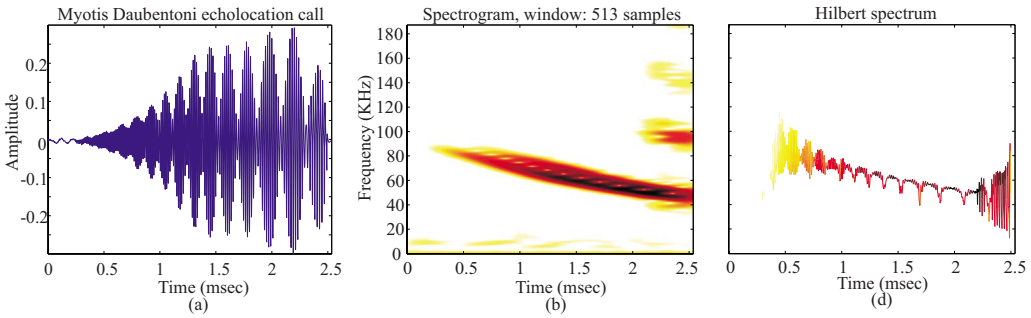


FIG. 7. (Color online) IF-based analysis of *Myotis Daubentoni* echolocation call.

components is in agreement with the amplitude modulation that the signal exhibits.

VII. GENERAL DISCUSSION

Based on the above study, we can conclude that despite having a performance that is usually worse than its competitors, the STFT is often a reliable tool for analysis and investigation of bat echolocation calls especially in cases that the bat signals do not contain both constant frequency (or nearly constant frequency) and fast down-chirp parts. However, one must always be mindful of the trade-off in the time-frequency resolution, which can make it hard to accurately determine the actual extent along the time or the frequency axis of the signal being investigated. Also a clear advantage of the STFT is its low computational cost.

The SPWVD is more flexible in dealing with signals comprising different types of echolocation calls. With a proper selection of the smoothing windows, better performance than the STFT can be attained, albeit at a higher computational cost. A deep understanding of the generation mechanism of the interference terms is essential for two reasons. First, it assists in the proper tuning of the smoothing filters, and second it allows for a correct interpretation of the resulting TF representation. It should be noted that if the interference terms can be recognized, and therefore ignored, less smoothing would be needed in order to produce much better concentration and time-frequency resolution than the STFT.

To the best of our knowledge, the MVE has never before been applied to the study of bat echolocation calls. It appears to be a useful technique capable of revealing details "missed" by other methods. However, MVE time resolution is limited when it is necessary to use quite long filters (large values of M parameter). This is the case when the studied signal contains components closely spaced in frequency. Otherwise, somewhat disturbing interferences appear between the neighboring (in frequency) components. Moreover, it comes at a significant computational cost that is comparable or even larger than that of the SPWVD.

A final comparison of the above methods is shown in Figs. 8(a)–8(c) concerning two successive echolocation calls of a *P. pipistrellus* feeding-buzz. At the buzz stage the echolocation calls exhibit abrupt frequency changes that limit the flexibility of the analysis window selection for the STFT. Indeed, the other methods offer more accurate TF representations.

IF analysis and Hilbert spectrum cannot replace the TF representation methods. However, they can often provide extra, useful information and insight into the signal under consideration. The Hilbert spectrum of the unweighted average of the IF corresponding to the feeding-buzz is depicted in Fig. 8(d). Based on the results of Sec. VI we would expect the unweighted average of the IF to coincide with the actual frequency of the harmonic having the higher amplitude. However, Fig. 8(d) does not hold in this specific case since the IF estimate exhibits an abrupt "jump." This can be explained by the fact that the higher amplitude harmonic (the lower frequency one) has a spectral null, which can be noted from the TF representations [Figs. 8(a)–8(c)]. The IF estimate at the point of the frequency null, where the amplitude of the harmonic fades abruptly, tends to track the higher frequency harmonic. At the time instant of the frequency null, the signal component with the larger amplitude is no longer the lower frequency harmonic, but the higher frequency one. Thus, the IF-based analysis is capable of revealing the frequency null and its temporal location.

VIII. CONCLUSION

In this paper, advanced frequency-content estimation techniques were exposed and applied in the context of bio-

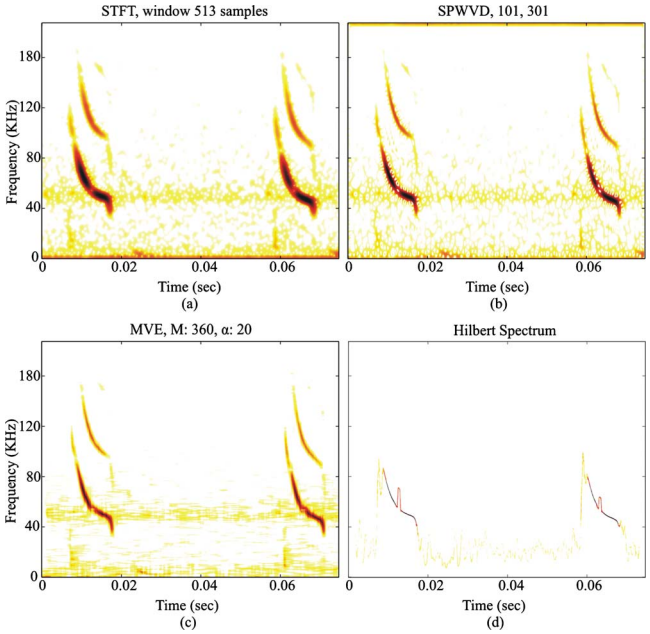


FIG. 8. (Color online) Comparison of different methods when analyzing two successive echolocation calls of a *P. pipistrellus* feeding-buzz.

logical signal analysis. It was shown that the proposed techniques are capable of showing previously unseen features in bat echolocation call structure.

ACKNOWLEDGMENT

This work was performed as part of the BIAS consortium under a grant funded by the EPSRC under their Basic Technology Programme.

[1]G. W. Pierce and D. R. Griffin, "Experimental determination of supersonic notes emitted by bats," J. Mammal. **19**, 454–455 (1938).
[2]D. R. Griffin and R. Galambos, "The sensory basis of obstacle avoidance by flying bats," J. Exp. Zool. **86**, 481–506 (1941).
[3]D. R. Griffin, "Measurements of the ultrasonic cries of bats," J. Acoust. Soc. Am. **22**, 247–255 (1950).
[4]D. R. Griffin, "Ultrasonic sounds of bats, their acoustical dimensions and biological significance," Anat. Rec. **108**, 540–540 (1950).
[5]F. P. Mohres, "Die ultraschall-orientierung der fledermause (The ultrasound orientation of bats)," Naturwissenschaften **39**, 273–279 (1952).
[6]F. P. Mohres, "Uber die ultraschallorientierung der hufeisennasen (chiroptera-rhinolophinae) (On the ultrasonic orientation of the horseshoe nosed bats (chiroptera-rhinolophinae))," Zeitschrift Fur Vergleichende Physiologie **34**, 547–588 (1953).
[7]D. A. Cahlander, J. J. G. McCue, and F. A. Webster, "The determination of distance by echolocating bats," Nature (London) **201**, 544–546 (1964).
[8]D. R. Griffin and A. Novick, "Acoustic orientation of neotropical bats," J. Exp. Zool. **130**, 251–300 (1955).
[9]W. Koenig, H. K. Dunn, and L. Y. Lacy, "Sound spectrograph," J. Acoust. Soc. Am. **18**, 19–49 (1946).
[10]A. Novick and J. R. Vaisnys, "Echolocation of flying insects by bat chilonycteris parnellii," Biol. Bull. **127**, 478–488 (1964).
[11]J. W. Cooley and J. W. Tukey, "An algorithm for machine calculation of complex Fourier series," Math. Comput. **19**, 297–301 (1965).
[12]J. A. Simmons and M. J. O'Farrell, "Echolocation by long-eared bat, plecotus-phyllotis," J. Comp. Physiol. **122**, 201–214 (1977).
[13]M. B. Fenton and G. P. Bell, "Echolocation and feeding-behavior in 4 species of myotis (chiroptera)," Can. J. Zool. **57**, 1271–1277 (1979).
[14]S. Parsons and G. Jones, "Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks," J. Exp. Biol. **203**, 2641–2656 (2000).
[15]G. Jones, M. Morton, P. M. Hughes, and R. M. Budden, "Echolocation, flight morphology and foraging strategies of some West-African hipposiderid bats," J. Zool. **230**, 385–400 (1993).
[16]O. Berger-Tal, R. Berger-Tal, C. Korine, M. W. Holderied, and M. B. Fenton, "Echolocation calls produced by kuhl's pipistrelles in different flight situations," J. Zool. **274**, 59–64 (2008).
[17]H. G. Broders, C. S. Findlay, and L. G. Zheng, "Effects of clutter on echolocation call structure of myotis septentrionalis and m-lucifugus," J. Mammal. **85**, 273–281 (2004).
[18]J. H. Fullard, "Echolocation assemblages and their effects on moth auditory systems," Can. J. Zool. **60**, 2572–2576 (1982).
[19]I. Matsuo, K. Kunugiyama, and M. Yano, "An echolocation model for range discrimination of multiple closely spaced objects: Transformation of spectrogram into the reflected intensity distribution," J. Acoust. Soc. Am. **115**, 920–928 (2004).
[20]D. Genzel and L. Wiegrebe, "Time-variant spectral peak and notch detection in echolocation-call sequences in bats," J. Exp. Biol. **211**, 9–14 (2008).
[21]U. Firzlaff and G. Schuller, "Directionality of hearing in two cf/fm bats, pteronotus parnellii and rhinolophus rouxi," Hear. Res. **197**, 74–86 (2004).
[22]R. A. Altes, "Wide-band, proportional-bandwidth Wigner-Ville analysis," IEEE Trans. Acoust., Speech, Signal Process. **38**, 1005–1012 (1990).
[23]R. A. Altes, "Signal processing for target recognition in biosonar," Neural Networks **8**, 1275–1295 (1995).
[24]S. N. Nawab and T. F. Quatieri, in *Short-Time Fourier Transform*, edited by J. S. Lim and A. V. Oppenheim (Prentice-Hall, Englewood Cliffs, NJ, 1988).
[25]M. R. Portnoff, "Time-frequency representations of digital signals and systems based on short-time Fourier analysis," IEEE Trans. Acoust., Speech, Signal Process. **28**, 55–69 (1980).
[26]The TF representations in the current section and Sec. IV are mainly produced with the aid of the time-frequency toolbox for MATLAB (Ref. 50).
[27]L. Cohen, *Time-Frequency Analysis: Theory and Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1995).
[28]F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," IEEE Signal Process. Mag. **9**, 21–67 (1992).
[29]L. Cohen, "Time-frequency distributions—A review," Proc. IEEE **77**, 941–981 (1989).
[30]P. Flandrin, *Time-Frequency/Time-Scale Analysis*, 1st ed. (Academic, New York, 1999).
[31]J. Capon, "High resolution frequency-wavenumber spectrum analysis," Proc. IEEE **57**, 1408–1418 (1969).
[32]An equivalent equation that directly gives the power estimate at the filter output bypassing the explicit filter impulse response computation is $\bar{P}_{\min}(\omega)=1/\mathbf{s}^H(\omega)\mathbf{R}^{-1}\mathbf{s}(\omega)$.
[33]Y. Jiang, P. Stoica, and J. Li, "Array signal processing in the known waveform and steering vector case," IEEE Trans. Signal Process. **52**, 23–35 (2004).
[34]L. Mandel, "Interpretation of instantaneous frequencies," Am. J. Phys. **42**, 840–846 (1973).
[35]B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals," Proc. IEEE **80**, 520–538 (1992).
[36]P. M. Oliveira and V. Barroso, "On the concept of instantaneous frequency," in IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 1998 (1998).
[37]D. E. Vakman, *Signals, Oscillations, and Waves: A Modern Approach*, 1st ed. (Artech House, Boston, MA, 1998).
[38]B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 2: Algorithms and applications," Proc. IEEE **80**, 540–568 (1992).
[39]L. B. Fertig and J. H. McClellan, "Instantaneous frequency estimation using linear prediction with comparisons to the DESAs," IEEE Signal Process. Lett. **3**, 54–56 (1996).
[40]S. Mukhopadhyay and P. Sircar, "Parametric modelling of non-stationary signals: A unified approach," Signal Process. **60**, 135–152 (1997).
[41]P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On amplitude and frequency demodulation using energy operators," IEEE Trans. Signal Process. **41**, 1532–1550 (1993).
[42]A. Potamianos and P. Maragos, "A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation," Signal Process. **37**, 95–120 (1994).
[43]S. Lawrence and J. Marple, "Computing the discrete-time analytic signal via fft," IEEE Trans. Signal Process. **47**, 2600–2603 (1999).
[44]It should be noted that Hilbert spectrum plots are usually related to a decomposition technique called empirical mode decomposition (EMD) (Ref. 51) that aims to break a multicomponent signal into a number of monocomponent signals (Ref. 52) in order to obtain accurate IF estimates. Unfortunately, EMD is rarely efficient for echolocation bat calls due to the fact that the high frequency harmonics are usually of low amplitude compared to the lower frequency components.
[45]A. Potamianos and P. Maragos, "Speech formant frequency and bandwidth tracking using multiband energy demodulation," J. Acoust. Soc. Am. **99**, 3795–3806 (1996).
[46]P. J. Loughlin, "Time-varying frequencies of a signal," Proc. SPIE **3162**, 109–122 (1997).
[47]A. Potamianos and P. Maragos, "Speech analysis and synthesis using an am-fm modulation model," Signal Process. **28**, 195–209 (1999).
[48]L. Rossi and G. Girolami, "Instantaneous frequency and short term Fourier transforms: Application to piano sounds," J. Acoust. Soc. Am. **110**, 2412–2420 (2001).
[49]The term unweighted IF average is used to make a distinction with respect to the weighted IF average that appears in IF related literature (see, for example, Ref. 47). Note that the weighted average is not suitable for the applications studied in this paper and is consequently not discussed here.
[50]F. Auger, P. Flandrin, P. Gonçalvès, and O. Lemoine, "Time-frequency toolbox, ver. 0.2," http://tftb.nongnu.org/ (Last viewed July 31, 2009).
[51]N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, W. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. R. Soc. London, Ser. A **454**, 903–995 (1998).
[52]Y. Kopsinis and S. McLauglin, "Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach," IEEE Trans. Signal Process. **56** (Jan), 1–13 (2008).

The effect of signal duration on the underwater hearing thresholds of two harbor seals (*Phoca vitulina*) for single tonal signals between 0.2 and 40 kHz

Ronald A. Kastelein,[a] Lean Hoek, and Paul J. Wensveen
Sea Mammal Research Company (SEAMARCO), Julianalaan 46, Harderwijk 3843 CC, The Netherlands

John M. Terhune
Department of Biology, University of New Brunswick, P.O. Box 5050, Saint John, New Brunswick E2L 4L5, Canada

Christ A. F. de Jong
MON/Acoustics, TNO Science and Industry, P.O. Box 155, Delft 2600 AD, The Netherlands

(Received 8 July 2009; revised 8 December 2009; accepted 10 December 2009)

The underwater hearing sensitivities of two 2-year-old female harbor seals were quantified in a pool built for acoustic research by using a behavioral psycho-acoustic technique. The animals were trained only to respond when they detected an acoustic signal ("go/no-go" response). Detection thresholds were obtained for pure tone signals (frequencies: 0.2–40 kHz; durations: 0.5–5000 ms, depending on the frequency; 59 frequency-duration combinations). Detection thresholds were quantified by varying the signal amplitude by the 1-up, 1-down staircase method, and were defined as the stimulus levels, resulting in a 50% detection rate. The hearing thresholds of the two seals were similar for all frequencies except for 40 kHz, for which the thresholds differed by, on average, 3.7 dB. There was an inverse relationship between the time constant (τ), derived from an exponential model of temporal integration, and the frequency [$\log(\tau)=2.86-0.94\log(f)$; τ in ms and f in kHz]. Similarly, the thresholds increased when the pulse was shorter than ~780 cycles (independent of the frequency). For pulses shorter than the integration time, the thresholds increased by 9–16 dB per decade reduction in the duration or number of cycles in the pulse. The results of this study suggest that most published hearing thresholds ≤1 kHz for harbor seals are probably not absolute, as they were derived from signals with durations shorter than the time constants for those frequencies.
 [DOI: 10.1121/1.3283019]

PACS number(s): 43.80.Lb, 43.80.Nd [WWA] Pages: 1135–1145

I. INTRODUCTION

The underwater hearing of harbor seals (*Phoca vitulina*) has been quantified by measuring responses to pure tones (Møhl, 1968; Terhune, 1988; Turnbull and Terhune, 1993; Kastak and Schusterman, 1998; Southall *et al.*, 2005), narrow-band frequency-modulated (FM) signals (Kastelein *et al.*, 2009a), and 1/3-octave noise bands (Kastelein *et al.*, 2009b). Relatively long signals were used in these studies (200–900 ms). In mammals, hearing thresholds depend on signal duration once the signal is shorter than the integration time of the hearing system for that frequency. The amplitude of a signal required for detection increases as duration decreases below a certain time constant, as has been shown for humans (Hughes, 1946; Plomp and Bouman, 1959; Watson and Gengel, 1969) and various other mammals (Fay, 1992; Gleich *et al.*, 2007) including two marine mammal species: the bottlenose dolphin *(Tursiops truncatus*; Johnson, 1968) and the harbor seal (Terhune, 1988). It is suggested that the detection process in the auditory system uses some kind of "temporal integration." The time constant associated with this process is the "integration time" of the auditory system.

Many underwater sounds of ecological importance for seals, such as their own vocalizations, are of short duration (Sterling, 1973; Møhl *et al.*, 1975; Thomas and Kuechle, 1982; Van Parijs and Kovacs, 2002). For example, many harp seal (*Pagophilus groenlandicus*) underwater calls are tones as brief as 0.1 s, and one type of broadband click has a duration of 0.1 ms (Møhl *et al.*, 1975). The distances at which short-duration vocalizations can be used for communication can only be assessed with any accuracy if the hearing threshold levels of seals for short signals are known.

The harbor seal has a wide distribution in the coastal waters of the northern hemisphere. In coastal waters, human activities producing underwater noise occur more often than in the open ocean. Anthropogenic noises of short duration may have high or low duty cycles. To determine the distance at which anthropogenic noises are audible to harbor seals (and may disturb them during activities such as foraging, predator avoidance, communication, and navigation), information on the harbor seal's underwater hearing sensitivity for short-duration signals is needed.

Some information on the effects of signal duration on the hearing thresholds of harbor seals has already been collected. Terhune (1988) tested the ability of a harbor seal to detect short-duration sounds, but presented the subject with a

[a]Author to whom correspondence should be addressed. Electronic mail: researchteam@zonnet.nl

number of pulses per second. In the 4–32 kHz range, the underwater detection threshold levels of the harbor seal for multiple short-duration pure tones increased when the total duration of the pulses fell below ~400 cycles.

The aim of the present study was to determine the hearing threshold for pulses in order to assess the effect of pulse duration on pulse audibility. The present study involves two subjects, a wider frequency range than used by Terhune (1988), and single sounds per trial. The results enable more accurate estimation of the ranges at which seals can detect single short-duration vocalizations and anthropogenic noises such as those created by pile driving, seismic surveys (airguns: multiple sounds, but with >1 s pulse interval), explosions, and some acoustic harassment devices. Lacking information on loudness growth as a function of level and duration, the results also allow a relative estimation of how anthropogenic noises of different signal durations are perceived by harbor seals at different distances (received levels usually decrease as distance increases) and under different transmission situations. The measurement results are compared to data from existing models of temporal integration in perception, in order to determine the extent to which these models can be used to predict the audibility of sounds of frequencies and durations relevant to particular situations at sea. The results will also indicate optimal signal durations (i.e., > integration time for each test frequency) for future hearing studies with harbor seals.

II. MATERIALS AND METHODS

A. Study animals

The study animals were two female harbor seals (identified as 1 and 2). During the study they aged from 1.5 to 2.5 years old and their body weight was approximately 42 kg. The seals had participated in two similar psychophysical hearing studies just before the present study (Kastelein *et al.*, 2009a, 2009b), and so were well accustomed to the daily hearing test routine.

B. Study area

The study was conducted at the SEAMARCO Research Institute (Goes, The Netherlands), which is in a remote area specifically selected for acoustic research. The research was conducted in an outdoor pool (8×7 m^2, 2 m deep), with an adjacent haul-out space (Fig. 1). The pool was designed and constructed to be as quiet as possible and to reduce reflection of sounds above 25 kHz (see Kastelein *et al.*, 2009a).

During test sessions, the seals were tested alternately. The seal being tested positioned itself at the listening station (an L-shaped, 32 mm in diameter, water-filled polyvinylchloride tube with an end cap). The animal not being tested was trained to keep very still and quiet for 15 min in the water next to the haul-out area (this was quieter than staying on land, where a scratch of a flipper nail could trigger a prestimulus response in the animal being tested). The operator and the equipment used to produce the stimuli were in a research cabin next to the pool, out of sight of both animals (Fig. 1).

C. Test stimuli

A schematic of the equipment used to configure and emit outgoing signals is shown in Fig. 2. The seals' hearing sensitivities were tested for pure tones of 0.2, 1, 4, 16, 32, and 40 kHz, produced by a waveform generator (Hewlett Packard, Houston, Texas, model 33120A). The signal duration was set by using the "burst" option of the waveform generator, which allowed the number of cycles produced to be defined. Signal duration varied between 0.5 and 5000 ms, depending on the frequency; 59 frequency-duration combinations were tested in all (Table I). The signals always started and ended at the zero crossing point. To prevent the production of unwanted high frequency on-/offset energies (transients), the tone pulses were filtered by a Butterworth band-pass filter (Krohn-Hite, Avon, Massachusetts, model 3362). The filter was set at 1/1-octave bandwidth around each test frequency. The minimum signal duration was set at 8 cycles (or 0.5 ms at 16 kHz and higher), because filtering slightly influenced the amplitude of the first and last cycles of the tone burst, and the proportion of cycles in a pulse, which were influenced, became relatively high below 8 cycles. The longest duration tested was that at which the thresholds stopped decreasing (showing that the integration time had been exceeded), which depended on the frequency.

A modified audiometer for testing human aerial hearing (Midimate, Taastrup, Denmark, model 602 with extended frequency range) was used to control the amplitude of signals. The free field sound pressure level (SPL) (in dB re 1 μPa, rms) at the seal's head while it was at the listening station could be varied in 5 dB increments. This step size was determined by the audiometer; 5 dB steps are generally used in human audiometry. The 0.2 kHz signals from the audiometer had to be amplified by an audio amplifier (Sony, Tokyo, Japan, model TA-F335 R) to reach audible received levels for the seals.

A directional transducer (Ocean Engineering Enterprise, North Canton, Ohio, DRS-12; 30 cm in diameter) was used to project the signals into the water. Multi-path arrivals can introduce both temporal and spatial variations in the observed SPL at the listening station. To minimize this, the transducer was placed in a corner of the pool in a protective wooden box, lined with rubber with an irregular surface. The transducer was hung with four nylon cords from the cover of the box and made no contact with the box. A stainless steel weight was attached to the bottom of the transducer to compensate for its buoyancy. The transducer was 1.85 m from the tip of the listening station (Fig. 1), and was positioned so that the acoustic axis of the projected sound beam pointed at the center of the listening station (i.e., at the center of the study animal's head while it was at the listening station). To reduce reflections from the bottom of the tank and water surface reaching the listening station, a baffle board was placed half way between the transducer and the animal. The board consisted of 2.4 m high, 1.2 m wide, 4 cm thick plywood, covered with a 2 cm thick closed cell rubber mat on the side facing the transducer. A 30 cm diameter hole was made in the board with its center at the same level as the seal's head and the transducer (1 m below the water surface).

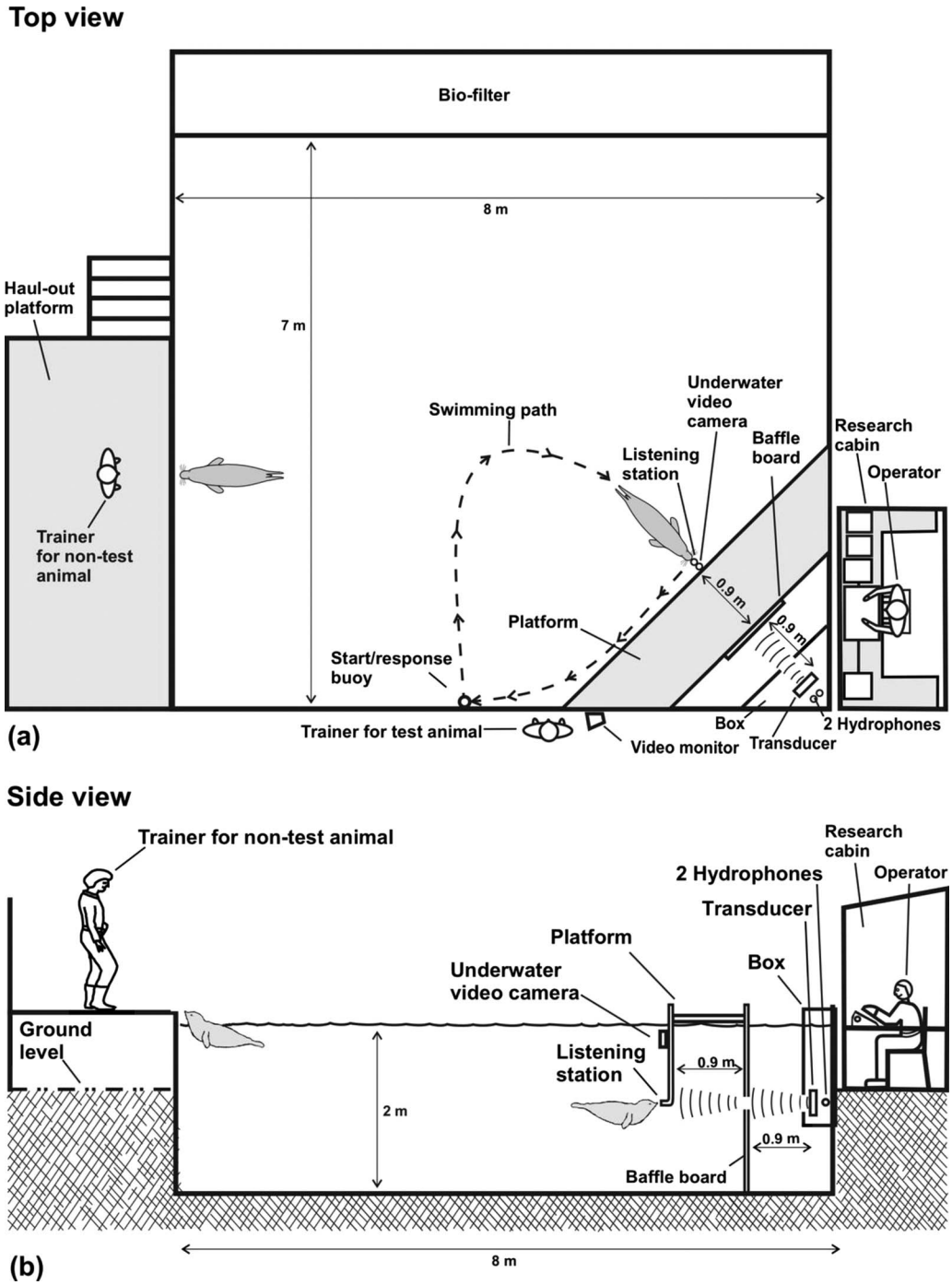


FIG. 1. The study area, showing one harbor seal in position for testing at the underwater listening station, and the non-test animal waiting with its trainer: (a) top view and (b) side view (both to scale).

As an indicator of the condition of the transducer, its capacity was checked once a week with a capacity meter (SkyTronic, Nijverdal, The Netherlands, model 600.103). Throughout the study period the capacity remained constant.

D. Background noise and stimuli level calibration measurements

The background noise and received levels of the signals were measured by TNO (Netherlands Organization for Applied Scientific Research, Delft, The Netherlands). The recording and analysis equipment consisted of a B&K 8101 hydrophone with a TNO (custom-built) power supply, a B&K PULSE 3560 D multichannel high frequency analyzer, and a laptop computer with B&K PULSE software LABSHOP version 12.1. The system was calibrated with a B&K 4223 pistonphone. The sound pressure level of each pulse (frequency and duration) was derived from the received 90% energy flux density, divided by the corresponding 90% time duration (Madsen, 2005).

Great care was taken to make the seal's listening environment as quiet as possible. Nobody was allowed to move within 15 m of the pool during sessions. Underwater background noise levels were measured three times during the

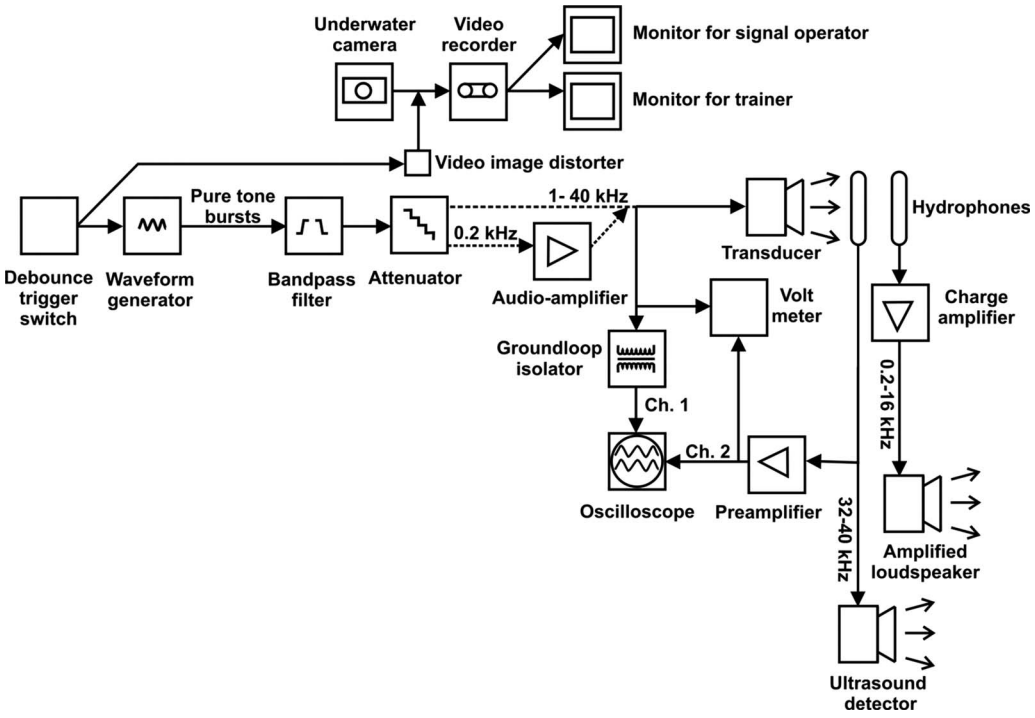


FIG. 2. Block diagram of the signal generation and listening systems.

10-month study period, under the same weather conditions as during the test sessions (wind speed Beaufort 4 or below). 1/3-octave band background noise sound pressure levels were determined in the range 25 Hz–100 kHz. The mean ($n=3$) background noise in the pool was very low (in the low frequency range, it was below sea state 0; see Kastelein *et al.*, 2009a, 2009b).

The received SPL (dB re 1 μPa, rms) of each frequency-duration combination was measured approximately once each month at the seal's head position during the hearing tests (Fig. 1). The measured received levels varied by 1–4 dB, depending on the frequency. During trials, the seal's head position (at the listening station) was carefully monitored, and was consistent to within a few centimeters. The SPL measured at various locations in the $20\times20\times20$ cm area occupied by the seal's head varied by a maximum of 3 dB. The received SPLs were calibrated at levels of 40–60 dB (depending on frequency and signal duration) above the hearing threshold levels found in the present study for each frequency. The linearity of the transmitter system was checked several times during the study; it was consistent to within 1 dB.

E. Experimental procedure

The experimental procedure was as in the previous tonal and noise band audiogram studies conducted with the same seals (Kastelein *et al.*, 2009a and 2009b). The seals were trained to respond ("go") in the presence of a signal and to withhold the response ("no-go") in the absence of a signal. A trial began when the seal not being tested was near the platform with a trainer, and the seal being tested was positioned with its head at the start/response buoy at the edge of the pool next to the trainer for the test animal [Fig. 1(a)]. When the trainer for the test animal gave a vocal command accompanied by a gesture (pointing downwards), the test animal descended to the listening station, so that its external auditory meatus was 200 cm from the sound source and 100 cm below the water surface [i.e., mid-water; Fig. 1(b)]. Each animal was trained to position its nose against the listening station so that its head axis was in line with the projected beam axis of the transducer. The listening station was not connected to the sound box, and the transducer was suspended within the box by wires, so the animals were not able to use vibration via contact conduction to the nose to detect the signals. The animals' positions could be viewed from above by means of an underwater camera (Mariscope, Micro, Kiel, Germany), which was attached to the listening station. The images were visible to the trainer for the test animal (who was near the start/response buoy out of the animal's view when it was at the listening station) and to the operator in the research cabin.

Two trial types were conducted during each experimental session: signal-present and signal-absent trials. In signal-present trials, the stimulus was presented unpredictably between 4 and 10 s after the animal was positioned correctly at the listening station. A minimum waiting time of 4 s was chosen because it took about 4 s for the waves created by the animal's descent to dissipate. If the animal detected the sound, it responded by leaving the listening station ("go" response) at any time during the signal's duration and returning to the start/response buoy [Fig. 1(a)]. The operator then indicated to the trainer for the test animal that the response was correct (a hit), after which the trainer gave a vocal signal and the seal received a fish reward. If the animal did not respond to the signal, the operator indicated to the trainer that the animal had failed to detect the signal (a miss). The trainer then indicated to the animal (by tapping softly on the

TABLE I. The mean 50% detection thresholds (± SD, n=5–7 sessions) and prestimulus response levels (for both signal-present and signal-absent trials) of the two 2-year-old female harbor seals for pure tones of 0.5–5000 ms duration (depending on the frequency).

			Harbor seal 1			Harbor seal 2		
Frequency (kHz)	No. of cycles	Duration (ms)	Mean 50% detection threshold level (dB re 1 μPa, rms) ± SD	No. of reversal pairs	Prestimulus response level (%)	Mean 50% detection threshold level (dB re 1 μPa, rms) ± SD	No. of reversal pairs	Prestimulus response level (%)
0.2	1 000	5 000	63 ± 4	40	4	60 ± 6	42	1
0.2	500	2 500	64 ± 4	56	7	63 ± 4	50	10
0.2	400	2 000	66 ± 5	48	2	64 ± 5	59	3
0.2	300	1 500	66 ± 4	57	6	64 ± 4	54	7
0.2	250	1 250	66 ± 4	49	11	65 ± 4	59	7
0.2	200	1 000	70 ± 4	91	7	72 ± 3	74	3
0.2	100	500	73 ± 4	67	9	74 ± 4	70	4
0.2	60	300	73 ± 4	54	7	74 ± 3	69	4
0.2	40	200	73 ± 4	70	8	74 ± 4	61	3
0.2	20	100	76 ± 4	57	10	76 ± 4	100	2
0.2	10	50	80 ± 3	52	7	83 ± 4	77	5
1	2 500	2 500	54 ± 4	59	6	53 ± 4	47	8
1	2 000	2 000	53 ± 4	72	3	55 ± 4	54	7
1	1 500	1 500	55 ± 4	79	5	57 ± 3	62	4
1	1 000	1 000	55 ± 4	67	4	56 ± 3	97	6
1	500	500	57 ± 4	86	6	56 ± 3	71	3
1	300	300	56 ± 4	64	6	57 ± 4	76	2
1	200	200	58 ± 4	78	5	58 ± 3	77	1
1	100	100	59 ± 4	75	5	61 ± 4	70	7
1	50	50	66 ± 4	66	9	64 ± 4	60	6
1	10	10	77 ± 4	64	10	78 ± 5	66	3
4	4 000	1 000	52 ± 4	75	4	53 ± 3	64	3
4	2 000	500	52 ± 3	57	8	54 ± 3	60	5
4	1 200	300	52 ± 4	67	7	53 ± 3	73	3
4	800	200	53 ± 4	65	7	53 ± 3	70	2
4	400	100	56 ± 4	69	5	58 ± 4	71	2
4	200	50	60 ± 4	56	7	62 ± 4	63	3
4	40	10	72 ± 6	51	8	70 ± 4	61	9
4	20	5	79 ± 5	51	5	83 ± 5	62	2
16	16 000	1 000	61 ± 3	77	2	62 ± 4	68	2
16	8 000	500	61 ± 3	64	7	63 ± 4	68	2
16	4 800	300	61 ± 4	55	3	63 ± 4	62	1
16	3 200	200	64 ± 3	76	5	63 ± 4	53	1
16	1 600	100	66 ± 4	63	5	62 ± 4	107	1
16	800	50	65 ± 4	57	7	64 ± 3	65	4
16	160	10	71 ± 3	57	7	73 ± 3	74	3
16	80	5	78 ± 4	55	9	80 ± 4	59	2
16	16	1	87 ± 4	55	7	90 ± 4	67	1
16	8	0.5	90 ± 4	62	5	93 ± 4	52	6
32	32 000	1 000	65 ± 3	90	5	67 ± 3	71	5
32	16 000	500	67 ± 4	89	4	67 ± 4	60	3
32	9 600	300	66 ± 3	89	6	68 ± 3	58	1
32	6 400	200	66 ± 4	109	5	67 ± 4	58	1
32	3 200	100	69 ± 4	60	4	67 ± 4	66	2
32	1 600	50	68 ± 3	89	7	66 ± 4	42	4
32	320	10	71 ± 3	72	2	67 ± 3	48	4
32	160	5	74 ± 3	56	5	70 ± 3	44	5
32	32	1	79 ± 4	60	8	78 ± 4	61	2
32	16	0.5	83 ± 4	62	8	82 ± 4	89	2
40	40 000	1 000	57 ± 4	73	4	63 ± 3	84	4
40	20 000	500	57 ± 4	82	8	61 ± 4	61	5
40	12 000	300	59 ± 3	61	7	63 ± 4	69	4
40	8 000	200	60 ± 4	77	3	63 ± 3	100	2
40	4 000	100	59 ± 4	92	4	61 ± 3	58	4
40	2 000	50	60 ± 3	54	3	64 ± 3	71	3

TABLE I. (*Continued.*)

			Harbor seal 1			Harbor seal 2		
Frequency (kHz)	No. of cycles	Duration (ms)	Mean 50% detection threshold level (dB re 1 μPa, rms) ± SD	No. of reversal pairs	Prestimulus response level (%)	Mean 50% detection threshold level (dB re 1 μPa, rms) ± SD	No. of reversal pairs	Prestimulus response level (%)
40	400	10	64 ± 3	54	8	67 ± 3	65	2
40	200	5	65 ± 4	50	6	68 ± 4	52	4
40	40	1	78 ± 4	58	10	81 ± 4	54	6
40	20	0.5	83 ± 4	62	4	88 ± 4	56	4

side of the pool) that the trial had ended, thus calling the animal back to the start/response buoy. No reward was given following a miss. If the animal moved away from the listening station to the start/response buoy before a signal was produced (a prestimulus response), the operator indicated to the trainer to end the trial without rewarding the animal. After a prestimulus response, the animal was ignored for 8–10 s by the trainer.

In signal-absent, or catch, trials, the operator signaled to the trainer to end the trial after a predetermined random interval of 4–10 s from when the seal had stationed. The trial was terminated when the trainer blew very softly on a whistle. The tapping on the pool wall and whistle blowing were done softly, to minimize the difference in the level of sounds heard by the seals, in order to help them focus on very faint sounds throughout the sessions. If the animal being tested responded correctly by remaining at the listening station until the whistle was blown (a correct rejection), it then returned to the start/response buoy and received a fish reward. If the seal left the listening station before the whistle was blown (a prestimulus response), the operator indicated to the trainer to end the trial without rewarding the animal. The same amount of fish was given as a reward for correct responses in signal-present and signal-absent trials. In both signal-present and signal-absent trials, the trainer was unaware of the trial type when she sent the animal to the listening station. After sending the animal to the listening station, the trainer stepped out of the seal's view.

A session generally consisted of 30 trials per animal and lasted for about 15 min per animal. The seals were always tested in the same order, animal 2 immediately after animal 1. Sessions consisted of 70% signal-present and 30% signal-absent trials presented in random order. For each session, one of four data collection sheets was used. Each sheet comprised a different random series of trial types. Each seal had its own set of four data collection sheets. In each session, the signal amplitude was varied according to the 1-up, 1-down adaptive staircase method. This conventional psychometric technique (Robinson and Watson, 1973) results in a 50% correct detection threshold (Levitt, 1971). During preliminary sessions a rough threshold per test frequency was determined. During subsequent experimental sessions, the starting SPL of the signal was 10–15 dB above the estimated threshold for each frequency-duration combination. Following each hit, the signal amplitude on the next signal-present trial was reduced by 5 dB. Following each miss, the signal level was increased on the next signal-present trial by 5 dB. Prestimulus responses did not lead to a change in signal amplitude for the next trial.

Thresholds were determined for each of the 59 frequency-duration combinations. To prevent the animals' learning process from affecting the threshold levels, a different frequency was tested each day. Adjacent frequencies were tested on successive days (going from low to high or from high to low frequencies), so that the difference in frequency between days was limited, thus reducing the animals' potential needs to adapt to a frequency. For each session, a signal duration was selected randomly from the potential signal durations for the selected frequency for that day (Table I). During a session, the two seals were presented with the same frequency-duration combination. Usually three experimental sessions per day with each animal were conducted 5 days/week (at 0900, 1200, and 1500 h). Data were collected between February 2008 and January 2009.

Before each session, the acoustic equipment producing the stimuli was checked to ensure that it was functioning properly and that the stimuli were being produced accurately. Also, the background underwater noise level was checked up to 20 kHz to make sure it was low enough. The equipment used to do this is shown in Fig. 2, and the methods are described by Kastelein *et al.* (2009a).

F. Analysis

Switches in the seal's response from a detected signal (a hit) to an undetected signal (a miss), or vice versa, called reversals, were used to calculate detection thresholds. A detected level and successive undetected level, or *vice versa*, are called a reversal pair. No warm-up trials were needed, as the thresholds were usually stable within sessions. Sessions with more than 20% prestimulus responses (i.e., in 6 or more of the 30 trials per session), which would have been omitted from the analysis, did not occur during the entire study period.

Thresholds for each animal for each frequency-duration combination were calculated by taking the mean of all reversal pairs (on average, 66 reversal pairs per frequency-duration combination), obtained in 5–7 sessions (Table I). The minimum number of reversal pairs was set at 40; but in some cases, the last session needed to reach this number contained many reversal pairs, so that more reversal pairs were used for analysis. Also, the equipment settings re-

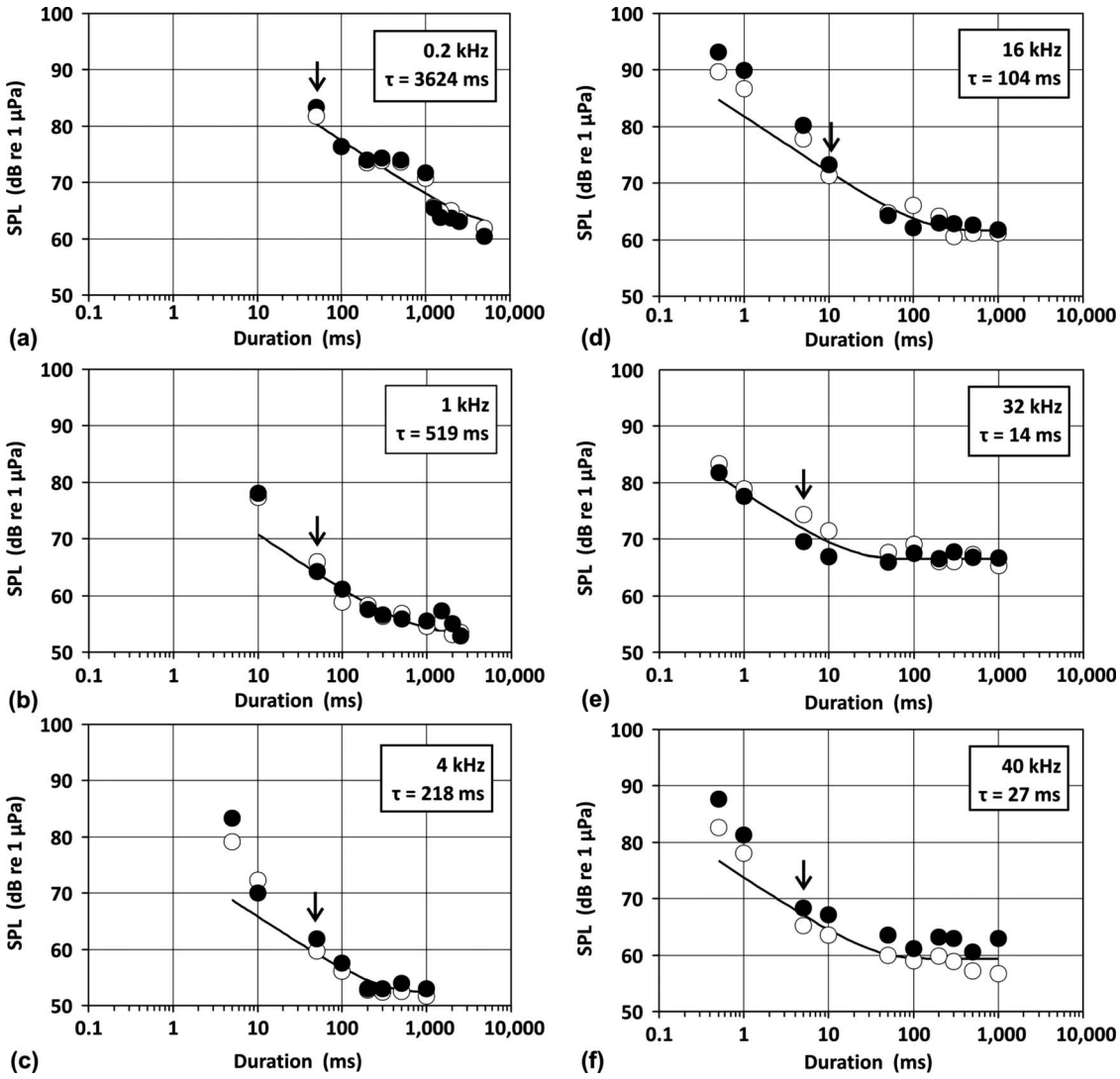


FIG. 3. The relationship between signal duration and mean 50% detection thresholds (dB re 1 μPa, rms) in the two female harbor seals [(○) seal 1 and (●) seal 2] for pure tones of (a) 0.2, (b) 1, (c) 4, (d) 16, (e) 32, and (f) 40 kHz (for details see Table I). The curves give the least-squares fits of the mean data to the exponential model [Eq. 5 in Plomp and Bouman, 1959]; the arrows indicate the shortest tones that were used to fit the model. For each frequency the time constant of integration (τ) is calculated.

mained the same in a session. In each session both seals were tested, so if one of the seals needed more reversals to reach the 40 reversal pair criterion, the other seal would have more reversal pairs. The hearing thresholds of the two seals for each frequency were compared by using a paired t-test.

To compare the hearing abilities of the harbor seals with other species, the exponential model of temporal integration (Plomp and Bouman, 1959) was fitted to the mean hearing thresholds for each signal frequency using least-squares regression. This model can be described by the following equation:

$$I/I_\infty = 1/(1 - e^{-t/\tau}), \quad (1)$$

where I is the intensity of a signal with duration t at the detection threshold, I_∞ is the intensity of a long signal, and τ is a time constant. Thresholds for very short tones are known to diverge from the model's fixed slope of 3 dB per doubling of duration, because of spectral splatter outside the critical band, so some were excluded from the analysis ($<5-20$ ms, depending on the frequency, arrows in Fig. 3). The intensity of the longest signal was set as the reference intensity I_∞, but if it decreased the sum of squared errors, the mean sound pressure level over the longest two signals was used.

Based on hearing thresholds of a harbor seal for multiple short-duration pure tones, Terhune (1988) proposed a model in which the threshold increase was related to the number of cycles in the signals. The mean threshold increases, relative to the threshold of the longest signal tested at each frequency in the present study, were examined with respect to the number of cycles in the tone pulses. The data for all frequencies at which the tones presented fewer than 1000 cycles per pulse were fitted to a linear equation based on the log of the number of cycles in the pulse. The Spearman r product-moment correlations were calculated between the log of the number of cycles in the pulse and the threshold shift, and also the log of the pulse duration and the threshold shift (independent of frequency in both instances).

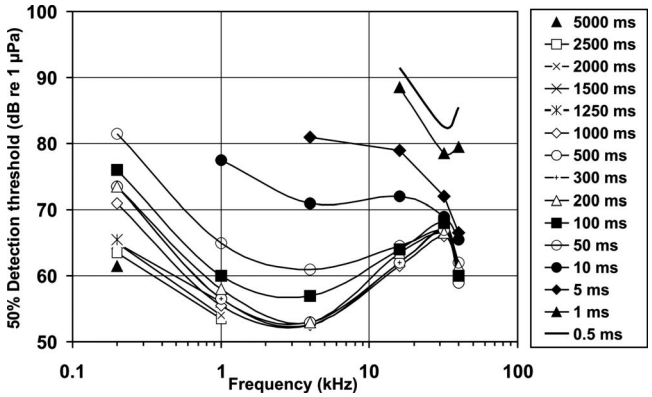


FIG. 4. The average hearing thresholds of the two seals in the present study for 0.2–40 kHz signals with durations between 0.5 and 5000 ms (depending on the frequency).

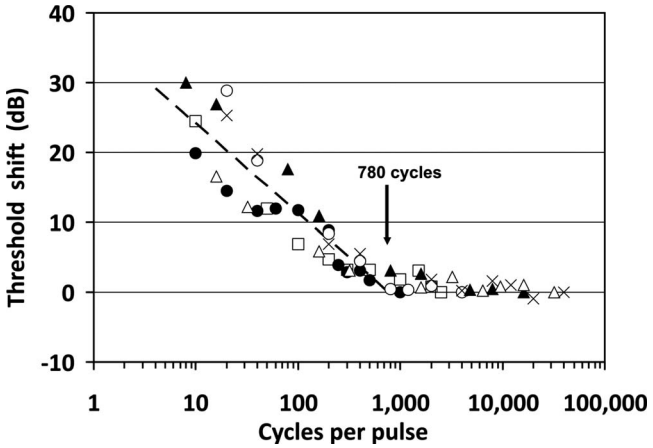


FIG. 5. The relationship between the number of cycles in a pulse and the mean increase in the hearing thresholds in the two female harbor seals for pure tones of 0.2 (●), 1 (□), 4 (○), 16 (▲), 32 (△), and 40 kHz (×). The dashed line represents the mean trend of the threshold increase with decreasing number of cycles [threshold increase (dB) = 36.90 − 12.77 log(cycles)].

III. RESULTS

The mean prestimulus response rate (for both signal-present and signal-absent trials) was 6% (range 2%–11%) for animal 1, and 4% (range 1%–10%) for animal 2 (Table I).

The relationships between signal duration and the 50% hearing threshold level for both harbor seals and for the six frequencies are shown in Fig. 3 and Table I. The hearing thresholds of the two seals did not differ significantly for the 0.2, 1, 4, 16, and 32 kHz signals (paired t-tests: 0.2 kHz: $t=0.16$, $P=0.875$, $n=11$; 1 kHz: $t=-1.10$, $P=0.299$, $n=10$; 4 kHz: $t=-2.02$, $P=0.083$, $n=8$; 16 kHz: $t=-1.27$, $P=0.235$, $n=10$; 32 kHz: $t=1.30$, $P=0.225$, $n=10$), but differed for the 40 kHz signals ($t=-10.09$, $P<0.001$, and $n=10$). At 40 kHz the thresholds of seal 2 were consistently higher (by, on average, 3.7 dB) than those of seal 1. If 3.7 dB is subtracted from the thresholds of seal 2, the thresholds of the two seals become similar (paired t-test, $t=-0.0002$, $P=0.99$, and $n=10$), indicating that even for this frequency both seals had the same pattern of threshold increase with decreasing signal duration, and the differences between the two seals did not vary with the frequency.

The mean audiograms of the two seals for the tested signal durations are shown in Fig. 4. This presentation shows the general trend: The threshold increases as the pulse duration shortens. For signals shorter than the integration time, the thresholds increased by 9–16 dB/decade reduction in the pulse duration or number of cycles in the pulse (Table I).

For signal durations with fewer than 1000 cycles/pulse (independent of the frequency: 0.2–40 kHz), the mean threshold increases fit the equation: Threshold increase (dB) = 36.90 − 12.77 log(cycles) (Fig. 5). This equation suggests that the detection thresholds of the seals increase when the pulse duration is shorter than ~780 cycles. The Spearman r product-moment correlation between the log(cycles) and the threshold shift is $r^2=0.83$ ($t=-12.84$, $n=36$, and $P<0.0001$). The Spearman r product-moment correlation between the log(duration) and the threshold shift is $r^2=0.46$ ($t=-5.42$, $n=36$, and $P<0.0001$).

IV. DISCUSSION AND CONCLUSIONS

A. Evaluation of the data

Because both seals were tested within the same session, the differences between the 40 kHz thresholds of signals of the same duration, obtained for the two animals, were probably due to differences in the seals' hearing sensitivities, and were not caused by differences in equipment, equipment settings, methodologies, personnel, background noises, or motivations. Because, apart from the absolute threshold values of the 40 kHz thresholds (which differed consistently by on average only 3.7 dB), the hearing thresholds of both seals were similar, the results of the present study can be assumed to be representative of young harbor seals with good hearing.

The gradual increase in hearing thresholds with decreasing signal duration was clear for all frequencies except 0.2 kHz. For this frequency the thresholds kept decreasing up to the maximum signal duration of 5 s, and the seals showed an irregular increase in threshold levels for signals between 200 and 1250 ms durations [Fig. 3(a)]. Integration time is usually thought of as a physiological property of the hearing system with a value of no more than a few hundreds of milliseconds. Therefore, in hindsight, the unusual pattern in the 0.2 kHz data could be an artifact. While the background noise levels were sufficiently low on average, it is possible that during some research sessions, the 0.2 kHz thresholds were influenced, especially at such a low frequency where background noise is known to fluctuate. It is also possible that near field effects (due to the seals' heads being close to the projector), such as particle motion, played a role, and that the 0.2 kHz signals were detected partly through mechano-reception via the seals' mystacial vibrissae (Dehnhardt *et al.*, 1998). Further research may be needed to explain the observed phenomenon.

The thresholds found for the 1000 ms duration pure tone signals in the present study were similar to those found for 900 ms narrow-band FM tonal signals in a previous study with the same seals in the same environment, which was conducted in the year immediately before the start of the

present study (Kastelein *et al.*, 2009a; average difference between the 900 and 1000 ms thresholds of both seals was 1.9 dB; range 0–4 dB). Thus, pure tones and narrow-band FM signals (with a modulation range of ±1% of the center frequency) of similar duration result in similar hearing thresholds in harbor seals.

B. Models of perception

The data obtained in the present study have been compared against existing models of the relationship between mammalian hearing thresholds and pulse duration. Such models would allow a generalization of the findings of this study to frequencies and pulse durations that have not been tested.

In the 4–32 kHz range, the underwater detection threshold levels of a harbor seal for multiple short-duration pure tones increased once signals presented fewer than ~410 cycles/pulse. The equation

$$\text{threshold increase (dB)} = 40.5 - 15.5 \log(\text{cycles})$$

fits the data for the seal closely (r^2=0.96) (Terhune, 1988). An analysis (Terhune, 1988) of bottlenose dolphin data reported by Johnson (1968) for short-duration tone thresholds for pulses at 4, 20, 45, and 100 kHz and pulse lengths from 1000 to 0.1 ms showed that the best fit equation was

$$\text{threshold increase} = 29.8 - 8.2 \log(\text{cycles}) \quad (r^2 = 0.94).$$

For these two species at high frequencies, the linear relationship between the log of the number of cycles in the pulse has a very high correlation with the threshold increase (dB). In the current study, the linear relationship between the log of the number of cycles in the pulse and the threshold increase (dB) for the two harbor seals is not as strong (r^2=0.83) but it covers a wider frequency range than was examined by Johnson (1968) and Terhune (1988). The correlation between the threshold increase and the log of the pulse duration is much weaker (r^2=0.46). This comparison is confounded somewhat by the relationship between the pulse duration and the number of cycles in the pulse. The "cycles" model, however, has the advantage that an explanation for the shortening of the integration time with increasing frequency does not have to be addressed.

For human hearing, two basic models have been used to describe temporal integration functions in perception: The power function model (Green *et al.*, 1957), and the exponential model (Plomp and Bouman, 1959). In both, the ear is assumed to be integrating sound intensity over time as an energy-detector, although, in reality, it is probably neural activity that is being integrated (Zwislocki, 1960). The long constants of temporal integration (100–300 ms) needed under the energy-detector hypothesis are in contrast with the very short constants of temporal integration of neurons in the auditory periphery, and are also difficult to reconcile with the high temporal resolution of the auditory system (Green, 1985). Alternatives to this "resolution-integration paradox" (deBoer, 1985) have been proposed, including a "multiple looks" model (Viemeister and Wakefield, 1991), in which a short integration time in the first stages of auditory processing creates short samples or "looks," which are processed somewhere within the central auditory system.

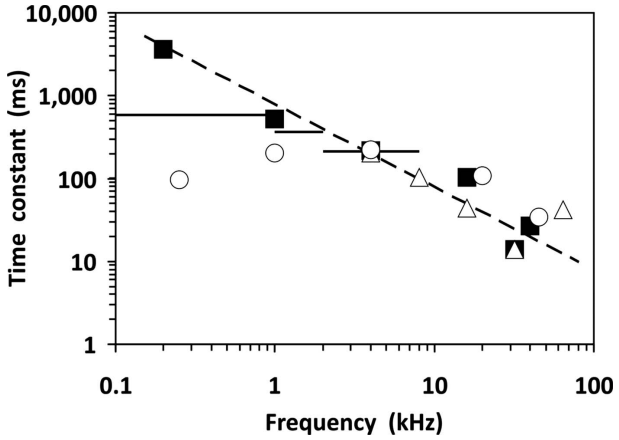


FIG. 6. Integration time constants (τ) based on the mean hearing thresholds of the two seals in the present study (■), another harbor seal (data reanalyzed in the same way as the data from the present study, from Terhune, 1988; △), and a bottlenose dolphin (Johnson, 1968; ○), and the average value from 20 studies on humans and other mammals in three frequency areas indicated by horizontal lines [up to 1, 1–2, and 2–8 kHz; reanalyzed by Gerken *et al.* (1990)]. The dashed line represents the duration of a pulse containing 780 cycles: The minimum number of cycles for which the seals' hearing thresholds were found to be independent of pulse duration.

The time constants of the exponential models ranged from 14 ms at 32 kHz to 3624 ms at 0.2 kHz (Fig. 3). For comparison, the time constants (τ) of the seals in the present study and of other mammals including humans (Fay, 1988; Terhune, 1988; Johnson, 1968; Watson and Gengel, 1969; Gerken *et al.*, 1990) are plotted against frequency (Fig. 6). There was an inverse relationship between the time constant (τ), derived from an exponential model of temporal integration, and the frequency [$\log(\tau)$=2.86-0.94$\log(f)$; τ in ms and f in kHz (the dashed line in Fig. 6)]. The dashed line in Fig. 6 represents the duration of a pulse, which contains 780 cycles, the minimum number of cycles for which the hearing threshold appeared to be independent of the pulse duration. This line follows the trend of the integration time constants. Hence, the present study thus suggests that the integration time constant of the seals' hearing correlates with a threshold in the number of cycles in the pulse. At frequencies ≤1 kHz, however, the time constants of the bottlenose dolphin and human data are lower than for the harbor seal (Fig. 6).

V. CONCLUSIONS

The present study suggests that most published harbor seal hearing thresholds for sounds ≤1 kHz ($\tau \geq 519$ ms) are not absolute hearing thresholds, as they were derived with signals shorter than the integration time constant of the species for those frequencies [Møhl, 1968 (±500 ms); Terhune, 1988 (500 ms); Turnbull and Terhune, 1993 (repeated signals, 50 ms, 10/s); Kastak and Schusterman, 1998 (500 ms); Southall *et al.*, 2005 (500 ms); Kastelein *et al.*, 2009a, 2009b (900 ms)]. Taking the results of the present study into consideration, hearing thresholds of harbor seals for signals ≤1 kHz are lower than those previously reported. The thresholds of the two seals at 0.2 kHz dropped, on average,

11 dB when the signal duration was increased from 900 to 5000 ms. Published low frequency hearing thresholds of other pinniped species (Phocids, Otariids, and Odobenids) may also be too high, if the tests were conducted with signal durations shorter than the species' integration times for those frequencies.

Our results indicate that ecologically important signals (such as the seals' own vocalizations) of shorter duration than the integration time are audible over shorter distances than estimated based on published hearing thresholds. A detection threshold increase of 6 dB would result in a reduction in the seal-to-seal communication range of 50% or more. Unmasked sounds, however, with frequencies $\leq$1 kHz and of durations longer than those used in previous harbor seal hearing studies, can be heard by the seals over longer distances than formerly believed. The same applies to anthropogenic signals: Tonal signals with fewer than ~780 cycles are audible to harbor seals at shorter distances from the sound source than estimated from published hearing thresholds, which were based on longer signals. The harbor seal detection thresholds for 900 ms narrow-band noises are essentially the same as for pure tones (Kastelein *et al.*, 2009a, 2009b). However, the majority of anthropogenic noises are not tonal. Therefore research is needed to determine if the threshold increases in short-duration, broadband sounds are similar to those of tonal signals described here.

ACKNOWLEDGMENTS

We thank the following students: Krista Krijger, Tess van der Drift, Janna Loot, Hanna Lobbezo, Niek Boonman, Marieke Priester, Marijke Oris, Remona Kerssies, Djolien de Wit, Linda Dekker, Eline Berrevoets, Jessica Schop and Marjan Viergever, and volunteers Jesse Dijkhuizen, Cathy Philipse, Saskia Roose, Joke de Lange, Wijnand Wolf, Sandra Wisse, and Jaya de Jonge, for their help in collecting the data, and Rob Triesscheijn for making the graphs. We thank Veenhuis Medical Audio (Marco Veenhuis and Herman Walstra) for donating and modifying the audiometer, and Arie Smink for the design, construction, and maintenance of the electronic equipment. We thank Bert Meijering (Director of Topsy Baits) for providing space for the SEAMARCO Research Institute. Willem Verboom (JunoBioacoustics), Erwin Jansen, and Pieter van Beek (both from TNO-Delft) conducted the acoustic calibration measurements. We thank Nancy Jennings (Dotmoth.co.uk, Bristol, UK), Willem Verboom, and two anonymous reviewers for their valuable constructive comments on this manuscript. Funding for this project was obtained from The Netherlands Ministry of Defense (Contract No. WO geluidsduur zeezoogdieren 235-08-6006-01). We thank Vincent Gales, Ronald de Rooij (DRMV, Netherlands Ministry of Defense), and René Dekeling (Royal Netherlands Navy), for their guidance on behalf of the commissioner during the project. We thank director Just van den Broek and curator of animals Henk Brugge (both from Ecomare) for making the harbor seals available for this project. The seals' training and testing were conducted under authorization of The Netherlands Ministry of Agriculture, Nature and Food Quality, and Department of Nature Management, with Endangered Species Permit No. FF/75A/2005/048.

deBoer, E. (**1985**). "Auditory time constants: A paradox?," in *Time Resolution in the Auditory System*, edited by A. Michelsen (Springer, Heidelberg), pp. 141–158.

Dehnhardt, G., Mauck, B., and Bleckmann, H. (**1998**). "Seal whiskers detect water movement," Nature (London) **394**, 235–236.

Fay, R. R. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay Associates, Winnetka, IL).

Fay, R. R. (**1992**). "Structure and function in sound discrimination among vertebrates," in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer, New York), pp. 229–263.

Gerken, G. M., Bhat, V. K. H., and Hutchison-Clutter, M. (**1990**). "Auditory temporal integration and the power function model," J. Acoust. Soc. Am. **88**, 767–778.

Gleich, O., Kittel, M. C., Klump, G. M., and Strutz, J. (**2007**). "Temporal integration in the gerbil: The effects of age, hearing loss and temporally unmodulated and modulated speech-like masker noises," Hear. Res. **224**, 101–114.

Green, D. M. (**1985**). "Temporal factors in psychoacoustics," in *Time Resolution in the Auditory System*, edited by A. Michelsen (Springer, Heidelberg), pp. 122–140.

Green, D. M., Birdsall, T. G., and Tanner, W. P. (**1957**). "Signal detection as function of signal intensity and duration," J. Acoust. Soc. Am. **29**, 523–531.

Hughes, J. W. (**1946**). "The threshold of audition for short periods of stimulation," Philos. Trans. R. Soc. London, Ser. B **133**, 486–490.

Johnson, C. S. (**1968**). "Relation between absolute threshold and duration-of-tone pulses in the bottlenosed dolphin," J. Acoust. Soc. Am. **43**, 757–763.

Kastak, D., and Schusterman, R. J. (**1998**). "Low-frequency amphibious hearing in pinnipeds: Methods, measurements, noise, and ecology," J. Acoust. Soc. Am. **103**, 2216–2228.

Kastelein, R. A., Wensveen, P. J., Hoek, L., Verboom, W. C., and Terhune, J. M. (**2009a**). "Underwater detection of tonal signals between 0.125 and 100 kHz by harbor seals (*Phoca vitulina*)," J. Acoust. Soc. Am. **125**, 1222–1229.

Kastelein, R. A., Wensveen, P. J., Hoek, L., and Terhune, J. M. (**2009b**). "Underwater detection of noise bands between 0.125 and 100 kHz by harbor seals (*Phoca vitulina*)," J. Acoust. Soc. Am. **126**, 476–483.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Madsen, P. T. (**2005**). "Marine mammals and noise: Problems with root mean square sound pressure levels for transients," J. Acoust. Soc. Am. **117**, 3952–3957.

Møhl, B. (**1968**). "Auditory sensitivity of the common seal in air and water," J. Aud. Res. **8**, 27–38.

Møhl, B., Terhune, J. M., and Ronald, K. (**1975**). "Underwater calls of the harp seal. *Pagophilus groenlandicus*," Rapp. P.-V. Reun.-Cons. Int. Explor. Mer **169**, 533–543.

Plomp, R., and Bouman, M. A. (**1959**). "Relation between hearing threshold and duration for tone pulses," J. Acoust. Soc. Am. **31**, 749–758.

Robinson, D. E., and Watson, C. S. (**1973**). "Pyschophysical methods in modern pyschoacoustics," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. **2**, pp. 99–131.

Southall, B. L., Schusterman, R. J., Kastak, D., and Reichmuth Kastak, C. (**2005**). "Reliability of underwater hearing thresholds in pinnipeds," ARLO **6**, 243–249.

Sterling, I. (**1973**). "Vocalizations in the ringed seal (*Phoca hispida*)," J. Fish. Res. Board Can. **30**, 1592–1594.

Terhune, J. M. (**1988**). "Detection thresholds of a harbor seal to repeated underwater high-frequency, short duration sinusoidal pulses," Can. J. Zool. **66**, 1578–1582.

Thomas, J. A., and Kuechle, V. B. (**1982**). "Quantitative analysis of Weddell seal (*Leptonychotes weddelli)* underwater vocalizations at McMurdo Sound, Antarctica," J. Acoust. Soc. Am. **72**, 1730–1738.

Turnbull, S. D., and Terhune, J. M. (**1993**). "Repetition enhances hearing detection thresholds in a harbour seal (*Phoca vitulina*)," Can. J. Zool. **71**, 926–932.

Van Parijs, S. M., and Kovacs, K. M. (**2002**). "In-air and underwater vocalizations of eastern Canadian harbour seals, *Phoca vitulina*," Can. J. Zool. **80**, 1173–1179.
Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am. **90**, 858–865.
Watson, Ch. S., and Gengel, R. W. (**1969**). "Signal duration and signal frequency in relation to auditory sensitivity," J. Acoust. Soc. Am. **46**, 989–997.
Zwislocki, J. J. (**1960**). "Theory of temporal auditory summation," J. Acoust. Soc. Am. **32**, 1046–1060.

Quantitative test for sensory hand symptoms based on mechanoreceptor-specific vibrotactile thresholds

Anthony J. Brammer
Ergonomic Technology Center, University of Connecticut Health Center, Farmington, Connecticut 06030 and Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

Paivi Sutinen
Department of Physical Medicine and Rehabilitation, North Karelia Central Hospital, FIN-80210 Joensuu, Finland

Sourish Das[a)]
Department of Statistics, University of Connecticut, Storrs, Connecticut 06269

Ilmari Pyykkö
Department of Otorhinolaryngology, Tampere University Hospital, FIN-33521 Tampere, Finland

Esko Toppila
Finnish Institute of Occupational Health, FIN-00250 Helsinki, Finland and Department of Otorhinolaryngology, Tampere University Hospital, FIN-33521 Tampere, Finland

Jukka Starck
Finnish Institute of Occupational Health, FIN-00250 Helsinki, Finland

(Received 18 February 2009; revised 6 November 2009; accepted 10 November 2009)

A vibrotactile test for assessing the presence or absence of sensory symptoms in the hand has been developed from thresholds believed mediated by Merkel disks and Meissner corpuscles at the fingertips. It is constructed from the summed differences between the thresholds recorded at the fingertip of an individual and the mean values of the threshold for healthy persons at the same stimulation frequencies. The summed normalized threshold shift, $TS_{Sum(SD)}$, is shown to be related to reports by subjects of numbness and pain using three statistical tests for evaluating the significance of associations in 2×2 contingency tables. The small number of subjects (15) restricts direct calculation of a fence value for $TS_{Sum(SD)}$, t, between the presence and absence of symptoms: accordingly, interpolation between calculated t values has been performed graphically. A common range of t values can be identified that is judged significantly by each statistical test ($3.3\leqslant t<3.9$ for numbness; $3.6\leqslant t<4.1$ for pain). The range encompasses the boundary previously identified between "normal" and "abnormal" vibrotactile thresholds, that is, thresholds for which the (two-sided) probability of occurrence in the hands of healthy persons is $p\approx0.05$.
 [DOI: 10.1121/1.3270395]

PACS number(s): 43.80.Qf, 43.80.Vj, 43.66.Wv [FD] Pages: 1146–1155

I. INTRODUCTION

Quantitative sensory testing (QST) has been used in clinical medicine since the 19th century for investigating peripheral sensory neuropathies (Dyck *et al.*, 2005). The tests were originally primitive—pin prick (for pain), hot/cold objects (for temperature sense), or a cotton wool swab (for fine touch) applied to the skin surface. While there have been many attempts to devise instruments to quantify some aspect of touch, pain, or the temperature sense, none has achieved broad acceptance by the medical community (Yarnitsky, 1997).

There are undoubtedly several reasons for the apparent limited confidence in QST, not the least of which is the systematic difference in the results of measurements of the same modality performed using different apparatus. Also, the reproducibility of tests conducted using some instruments is poorly documented, as are normal values for healthy persons. Furthermore, differences in test outcomes and the relationships with symptoms have been common in the many studies that have attempted to apply some QST modality to individuals or to a population group. These limitations have been cited in a consensus report on QST to the American Academy of Neurology (Shy *et al.*, 2003), which concluded, nevertheless, that such tests are potentially useful tools for measuring sensory impairment in clinical or research studies.

The issues documented by the consensus report may be addressed in several ways. For vibrotactile perception, which is potentially one of the most useful QST modalities, the first two issues have been addressed, at least in part, by the development of standardized methods for the measurement and

[a)]Present address: The Statistics and Applied Mathematical Sciences Institute, Duke University, Durham, NC.

assessment of vibrotactile perception thresholds in the hands (ISO 13091-1, 2001; ISO 13091-2, 2003). Here we address the third issue, namely, the specification of a vibrotactile metric that is robustly associated with a common sensory symptom reported by individuals (numbness). The derivation of such metrics has received limited attention in the literature (Lundborg *et al.*, 1992; Lundström *et al.*, 1992). A combination of sensory tests, including vibrotactile perception, has, however, been used effectively as an aid to diagnosis (Dyck *et al.*, 2005), and for ongoing medical monitoring of persons at risk of developing sensorineural disturbances in their hands (Jetzer *et al.*, 2003).

This paper focuses on the relationship between a metric derived from mechanoreceptor-specific vibrotactile thresholds and a subject's report of numbness or pain in his/her hands. The psychophysical and physiological bases for the metric have been described elsewhere (Brammer *et al.*, 2007a). It is constructed from the threshold shift, which is defined as the difference between the threshold recorded at the fingertip of an individual and the mean value of the threshold at the same stimulation frequency for healthy persons. Thresholds believed mediated by the slowly adapting type I (SAI) receptors and fast adapting type I (FAI) receptors within a fingertip are employed, as these are known both to be associated with symptoms of deteriorating tactile acuity in the hands (Coutu-Wakulczyk *et al.*, 1997), and change little with age in healthy persons (Brammer *et al.*, 1993; 2007a). The metric is then used to predict the presence or absence of sensory symptoms reported by subjects in the upper extremities. A group of 18 subjects, many of whom were known to have reported sensory hand symptoms (Brammer *et al.*, 2007b), were available for the analysis. Their selection was one of convenience. It was known that the threshold shifts for the SAI and FAI receptor populations within a fingertip were highly correlated in these subjects, a condition that may be related to their occupation but cannot be presumed generally to occur. Accordingly, the threshold shifts recorded from the SAI and FAI populations have been summed for each fingertip, and then combined first by hand and finally by subject to form a single value that is taken as the value of the quantitative sensory test for that subject. The essence of this work is the application of statistically based procedures to the vibrotactile metric, a continuous variable, to derive a robust estimate of a "fence" to distinguish between the reported presence or absence of a symptom, a dichotomous variable.

II. SUBJECTS, APPARATUS, AND METHODS

A. Study group

A small group of manual workers was available for the present study, many of whom were known to have reported sensory hand symptoms in the past. These workers have served as subjects in previous studies (Brammer *et al.*, 2007a, 2007b). Their participation in the study was voluntary, and subjects gave their informed consent. The ethics committees at the participating organizations approved the study protocol. The general health and, in particular, the hand and arm function of these and other workers performing similar work in the same region of Finland have been followed prospectively (Brammer *et al.*, 2007b; Sutinen *et al.*, 2006; Koskimies *et al.*, 1992). None of the subjects had operated power tools on the day of the vibrotactile measurements, and, in consequence, the perception thresholds recorded are believed to be free of any temporary loss in sensation resulting from acute exposure of the hand to intense vibration (Brammer *et al.*, 2006).

Each subject completed a questionnaire, the contents of which were unknown to the examiner conducting vibrotactile tests. Questions concerned: general health, use of medications, and work and medical history; numbness in the upper extremities, and its influence on the subject's sleep and productivity at work; pain in the hands, arms, and neck; and eating and smoking habits. The responses to the questionnaire were confirmed by a physician, who was unaware of the results of the vibrotactile measurements.

B. Measurement of vibrotactile perception thresholds

1. Apparatus

The apparatus conforms to method A of international standard ISO 13091-1 (2001) and has been described elsewhere (Brammer and Piercy, 1991; Brammer *et al.*, 2007a). It consists, briefly, of (1) a vibration stimulator suspended from a beam balance, the fulcrum of which is mounted on a vertically adjustable track; (2) an arm rest on which the hand and forearm are placed in supination; (3) a 3 mm diameter cylindrical probe through which the stimulus is applied to the skin with a controlled contact force of 0.05 N; (4) an accelerometer and electronics to record the stimulus at the surface of the skin; and (5) a computer to apply the psychophysical algorithm and calculate perception thresholds. The psychophysical procedure employed a version of the up-down algorithm and incorporated adaptive step sizes, determined by the subject's performance, to reduce the time required to administer the test (Brammer *et al.*, 2007a). The minimum step size was 2 dB. The stimuli were sinusoidal bursts of 800 ms duration at frequencies of 20 Hz, and above, and 1.6 s duration at frequencies of 6.3 Hz, and below. The quiescent interval between stimuli was 0.6 s. Successive bursts initially increased in intensity until the subject signaled, by pressing a switch, that the stimulus had been detected. Successive bursts then decreased in intensity until the subject signaled, by releasing the switch, that the stimulus could no longer be felt, whereupon the burst intensity was, once again, increased. This cycling of burst intensity was repeated at least four times, and the mean threshold acceleration was calculated from the arithmetic sum of the sequence of ascending and descending thresholds when each were expressed in dB re 10^{-6} m/s^2 (ISO 13091-1, 2001). The performance of the apparatus was confirmed daily, before commencing the measurement of vibrotactile thresholds.

An algorithm designed to identify inconsistencies in the subject's performance during the threshold-tracking task operated in the background searching for changes in ascending and/or descending thresholds as the task proceeded. The algorithm flags inconsistent responses and, also, both increases and decreases in threshold with time. As the calculations are

performed in real time, the examiner may choose to repeat or reject the results while the subject is present. It should be noted that a trend to increasing threshold with time during the tracking task would be consistent with sensory adaptation caused by the stimulus. Within the limit in resolution imposed by the step size, we were unable to detect systematic changes in the thresholds accepted for the analysis.

2. Measurement procedure

The subject was seated with forearm and hand supported horizontally, below shoulder level, and positioned the armrest for maximum comfort. Once comfortably seated, the stimulator was positioned so that the probe could be lowered onto a fingertip. Vibrotactile stimuli were then applied to the skin at amplitudes close to the threshold of perception, according to the algorithm. With the measurement conditions so defined, vibrotactile thresholds believed mediated by the SAI receptors could be obtained by stimulating at 4 and 6.3 Hz, and by the FAI receptors by stimulating at 20 and 32 Hz (Brammer *et al.*, 2007a).

The measurement procedure was first explained to the subject, together with the sensations he was likely to feel. After positioning the stimulator, several practice runs were performed to familiarize the subject with the stimuli and measurement procedure. When the examiner was confident the subject was adequately prepared, thresholds were determined at the fingertips of digits 3 and 5 of both hands.

C. Methods of analysis

1. Vibrotactile threshold shifts

The reference threshold shift is defined as the difference between the threshold recorded at the fingertip of an individual and the mean value of the threshold at the same stimulation frequency derived from measurements conducted on healthy persons (ISO 13091-2, 2003). The threshold shift at a fingertip is written TS_4, $TS_{6.3}$, TS_{20}, and TS_{32}, for stimulation frequencies of 4, 6.3, 20, and 32 Hz, respectively, and is expressed in decibels. The metric that has proved useful for assessing the status of hands in which the threshold shifts recorded from the SAI and FAI receptors in a fingertip are highly correlated is the summed normalized threshold shift, $TS_{Sum(SD)}$. It is constructed as follows (Brammer *et al.*, 2007a, 2007b): (1) the threshold shifts at frequencies mediated by the same receptor population are arithmetically averaged; (2) the mean threshold shift is normalized by a parameter expressing the distribution of thresholds observed in the hands of healthy persons for that receptor population (in this case the standard deviation), to form $TS_{SAI(SD)}$ and $TS_{FAI(SD)}$; (3) the standard deviation of $TS_{Sum(SD)}$ is then calculated, taking into account the observed correlation between thresholds mediated by the SAI and FAI receptor populations in the same finger; and (4) the largest value of $TS_{Sum(SD)}$ observed on any finger of a subject is taken as the value of the metric for that subject.

The summed normalized threshold shift at a fingertip when all thresholds are expressed in decibels (dB) is then

TABLE I. 2×2 contingency table for test metric, $TS_{Sum(SD)}$, with fence, t.

Test result	Symptom present (gold standard)	Symptom absent (gold standard)	Total
Positive	True positive	False positive	
$TS_{Sum(SD)} > t$	a	b	$a+b$
Negative	False negative	True negative	
$TS_{Sum(SD)} \leqslant t$	c	d	$c+d$
Total	$a+c$	$b+d$	$n=a+b+c+d$

$$TS_{Sum(SD)} = TS_{SAI(SD)} + TS_{FAI(SD)}. \quad (1)$$

In this equation,

$$TS_{SAI(SD)} = [TS_4 + TS_{6.3}]/2SD_{SAI} \quad (2)$$

and $TS_{FAI(SD)}$ has been constructed similarly. As the thresholds recorded from healthy hands appear to approximate a Gaussian distribution (Brammer *et al.*, 1993; ISO 13091-2, 2003), the range of the distribution may be expressed in terms of the standard deviation of its mean, namely, SD_{SAI} and SD_{FAI} for the SAI and FAI distributions, respectively.

Since the response of the two receptor populations in a fingertip are known to be highly correlated in these subjects, the standard deviation of $TS_{Sum(SD)}$ will involve the covariance of $TS_{SAI(SD)}$ and $TS_{FAI(SD)}$, $Cov(TS_{SAI(SD)}, TS_{FAI(SD)})$. Thus, when expressed in terms of variances (Brammer *et al.*, 2007b):

$$Var(TS_{Sum(SD)}) = 2[1 + Cov(TS_{SAI(SD)}, TS_{FAI(SD)})], \quad (3)$$

where the covariance may be approximated by the correlation between $TS_{SAI(SD)}$ and $TS_{FAI(SD)}$, as $SD_{SAI} \approx SD_{FAI}$. Application of Eq. (3) to the data permits an appropriate value for the standard deviation of $TS_{Sum(SD)}$ to be constructed. The acuity of a finger is then considered "normal" (N) if its value of $TS_{Sum(SD)}$ falls within 1.96 standard deviations of the mean value recorded in the hands of healthy persons (i.e., $p<0.05$), and "abnormal" (A) if the value of $TS_{Sum(SD)}$ falls outside this range.

In order to explore the relationships between $TS_{Sum(SD)}$ and symptoms, it is necessary to construct metrics for each hand and for each person, rather than for each finger. For each hand, the larger threshold shifts recorded from digits, 3 and 5 has been taken as the appropriate metric. Similarly, the larger threshold shifts for each hand has been taken as the appropriate single metric for the subject.

2. Statistical tests

A 2×2 contingency table was first constructed for the categorical variables (Table I). The symptom reported by the subject as confirmed by the examining physician, was taken as the "gold standard" (i.e., presumed correct) and the ability of $TS_{Sum(SD)}$ to predict the presence or absence of symptoms was examined using the chi-squared test with Yates' correction for small samples (Altman, 1991):

$$X_Y^2 = \frac{n[|ad-bc| - (n/2)]^2}{(a+b)(c+d)(a+c)(b+d)}, \quad (4)$$

where a is the number of true positives (i.e., subjects reporting symptom when test result indicates presence of symptom), b is the number of false positives, c is the number of false negatives, d is the number of true negatives, and n is the total number of subjects (see Table I). The test statistic has a chi-squared distribution with one degree of freedom. The value of $TS_{Sum(SD)}$ selected as the boundary between the presence or absence of symptoms was set as a numerical fence, t (in units of $SD_{SAI} \approx SD_{FAI}$), which was varied to examine the strength of the associations.

Now the test statistic for association fails if any of the factors in the denominator of Eq. (4) is zero, which becomes increasing likely as the sample size decreases, and the magnitude of $TS_{Sum(SD)}$ deviates from t. Accordingly, a Bayesian estimate of the association has been computed based on 20 000 samples generated from a multinomial distribution of the contingency table (four-dimensions) (Gelman *et al.*, 2004). The algorithm developed for this purpose is described in the Appendix.

Tests for the statistical significance of the associations were conducted to identify the most significant values of the fence to employ. For this purpose, Fisher's exact test for 2×2 contingency tables was performed in addition to the chi-squared test (Altman, 1991). The null hypothesis was that there was no association between the test statistic and the target symptom, and the alternative hypothesis was that there was a positive association (one-sided test): a probability value of $p=0.05$ was considered significant. The 95% credible interval (CI) for the Bayesian estimate of the measure of association was also computed and taken to indicate a statistically significant association when the CI did not include zero (Gelman *et al.*, 2004). The overall evaluation of the significance of the association was based on the combined results of the chi-squared test, Fisher's exact test, and the Bayes 95% CI.

The performance of the metric as a QST for a symptom for which a statistically significant association with $TS_{Sum(SD)}$ could be demonstrated was explored by calculating the corresponding sensitivity $(a/a+c)$ and specificity $(d/b+d)$ for a given fence value. Since the proportion of subjects with a positive test result who are correctly diagnosed $(a/a+b)$ and the proportion of subjects with a negative test result who are correctly diagnosed $(d/c+d)$ both depend on the prevalence of the symptom in the group at risk, a receiver operating characteristic (ROC) curve was constructed (Altman, 1991). The gradient of the ROC curve (i.e., sensitivity/[1—specificity]) gives the likelihood ratio, that is, the ratio of the probability of a positive test result in a subject with "disease" compared to the probability of a positive test result in a subject without disease. The likelihood ratio thus measures the change in certainty of the diagnosis obtained by conducting the test.

III. RESULTS

A. Subjects and symptoms

A summary of symptoms is given for each subject in Table II. Detailed symptom reports were unavailable for three subjects who did not participate in the heavy manual work (4, 12, and 16) and were, in consequence, excluded from the statistical tests of association. 39% of the workers reported numbness in the hands, which in two cases (11%) was described as being sufficiently incapacitating to influence the person's ability to work. 17% of subjects reported nocturnal numbness in the arms that caused awakening at night, in one case every night (subject 18). Six subjects (33%) reported pain in the hands, forearm, (upper) arm, or throughout the upper extremity, and 7 subjects (39%) reported neck pain.

TABLE II. Summary of symptoms, hand and arm functions and work history for each subject.

Subject	Age (years)	Numbness			Pain	Neck pain
		Status	Effect	At Night		
1	43	2	0	1	1	Yes
2	46	1	0	0	0	No
3	49	1	0	1	0	No
4	41	⋯	⋯	⋯	⋯	⋯
5	46	2	0	2	4	Yes
6	38	0	0	0	0	Yes
7	47	0	0	0	0	No
8	30	0	0	0	0	No
9	30	0	0	0	0	No
10	49	0	0	0	⋯	No
11	55	0	0	0	0	Yes
12	39	⋯	⋯	⋯	⋯	⋯
13	40	1	0	⋯	2,3 (L)	Yes
14	41	0	0	0	3 (R)	No
15	53	2	1	2	2,3 (R)	Yes
16	42	⋯	⋯	⋯	⋯	⋯
17	43	0	0	0	0	No
18	56	2	1	3	2 (R)	Yes

Legend:
Numbness status: 0—none; 1—a little; 2—a lot
Numbness effect: 0—none; 1—loss of earnings; 2—stops work
Numbness at night, causing awakening: 0—never; 1—once/week; 2— 2–6 times/week; 3—nightly
Pain: 0—none; 1— in hand; 2—in forearm; 3—in arm; 4—in upper extremity

B. Summed normalized threshold shifts

The summed normalized threshold shifts for the left and right hands of each subject are shown in Table III. Note that these results are expressed in units of standard deviation [see Eq. (1)]. Reference to this table confirms that the majority of summed normalized threshold shifts are positive rather than negative, implying that the thresholds are, on average, less sensitive in this group of subjects than those to be expected from healthy persons of the same age. For the group, the mean values of $TS_{Sum(SD)}$ are 3.44 for the left hand and 3.34 for the right hand. While the averaged results indicate that substantial changes in acuity are being detected, this analysis will focus on the status of *individual* hands in order to establish the utility of the metric as a QST. Inspection of Table III also reveals that large threshold shifts on one hand are commonly confirmed on the other hand.

The status of each hand with sensitivity that deviates from the range associated with the hands of healthy persons

TABLE III. Summed normalized shifts in threshold from expected values (expressed in units of SDs).

Subject	Left hand	Right hand	Both hands
1	8.0	7.1	8.0
2	3.5	1.9	3.5
3	2.9	5.2	5.2
4	3.8	3.2	3.8
5	2.9	4.3	4.3
6	1.3	2.4	2.4
7	2.3	−1.8	2.3
8	3.4	2.7	3.4
9	−1.1	−0.2	−0.2
10	2.5	−0.4	2.5
11	3.3	2.4	3.3
12	−2.3	3.5	3.5
13	9.6	7.9	9.6
14	4.8	4.2	4.8
15	3.9	4.8	4.8
16	2.1	3.1	3.1
17	6.5	4.6	6.5
18	4.6	5.2	5.2

is summarized in Table IV. The subject's reported numbness status is also listed for comparison. To obtain these results, the boundary between the classification of threshold shift as normal or abnormal has been established from the variance (or SD) of $TS_{Sum(SD)}$. The magnitude of the SD of $TS_{Sum(SD)}$ has been calculated from Eq. (3) and the measured values of the correlation between $TS_{SAI(SD)}$ and $TS_{FAI(SD)}$, and is listed for each finger in Table V. It can be seen from Table IV that significant shifts in threshold were observed in the hands of 9 of the 18 subjects, while a total of 15 hands (41.7%) are assessed by the threshold shift metric to possess abnormal sensitivity. It is also evident from the table that most of the subjects with values of $TS_{Sum(SD)}$ that deviated from the range associated with healthy persons also reported numbness in their hands.

C. Association between threshold shift metric and numbness

The association between the summed normalized threshold shift and the presence or absence of numbness in the hands reported by the subjects ($n=15$) has been analyzed

TABLE IV. Summary of thresholds for hands deviating from "normal."

Subject	Numbness	Left hand	Right hand
1	Y	A	A
3	Y	N	A
4	···	A	N
5	Y	N	A
13	Y	A	A
14	N	A	A
15	Y	A	A
17	N	A	A
18	Y	A	A
Abnormal Hands (A):		7 (38.9%)	8 (44.4%)

TABLE V. Boundaries for summed normalized threshold shifts.

Digit	Correlation	Summed normalized threshold shift: SD of $TS_{Sum(SD)}$ [a]	Summed normalized threshold shift: Boundary N-A
LH3	0.87	1.93	3.78
LH5	0.97	1.98	3.88
RH3	0.94	1.97	3.86
RH5	0.87	1.93	3.78

[a]See Eq. (3).

from 2×2 contingency tables using the statistical methods described. The association between the metric constructed for both hands (column 4 of Table III) and reports of numbness (column entitled "numbness status" in Table II) is shown in Table VI. Overall, each test statistic identifies a range of fence values that produce a statistically significant association between summed normalized threshold shift and numbness status, which are shown underlined in Table VI. While the test statistics differ somewhat in the range of fence values that they associate with numbness, values of $3.0 \leqslant TS_{Sum(SD)} < 4.0$ can be seen to be the most significant, with probabilities reaching $p<0.01$ for Fisher's exact test when the fence is set to $TS_{Sum(SD)}=3.5$. Note that the Bayes CI becomes statistically significant when the 2.5% credibility value exceeds zero.

Since the statistical tests can be expected to become continuous functions of the fence value in the limit when $n \rightarrow \infty$ (corresponding to an infinitely large number of subjects), it is instructive to summarize this information graphically to assist interpretation. For this reason, the results of these statistical tests are shown in Fig. 1, where the p-values for Yates' chi-squared test and for Fisher's exact test (left-hand scale), and the values for Bayes CI (right-hand scale) have been plotted as functions of the fence value. Curves have been fitted to the data from Table VI, and are shown by the continuous line for Yates' chi-squared test p-value, long dashed line for Fisher's exact test p-value, and short dashed lines for the Bayes CI. The data points from our subjects are shown in Fig. 1, as open squares, filled circles, and filled squares and triangles, respectively. A horizontal line corresponding to a Bayes CI and a p-value of zero, and a line corresponding to a p-value of 0.05 complete the presentation. Inspection of this graph shows that a unique region can readily be identified between fence values of 3.3 and 3.8 where the three statistical tests all indicate a significant association exists (viz., Yates' chi-squared and Fisher's p-value less than 0.05, and the Bayes 97.5% and 2.5% CIs greater than zero). Hence, a robust association is confirmed between $TS_{Sum(SD)}$ and each subject's reports of numbness status.

An expanded plot of the same statistical test parameters is shown in Fig. 2, where the presentation follows the scheme described for Fig. 1, to examine for any association between $TS_{Sum(SD)}$ and reports of numbness at night (see Table II). Expanding the scales renders more visible the region where statistical significance may occur. By inspection of this diagram, it is evident that the three statistical tests all

TABLE VI. Sensitivity, specificity, and tests of association for numbness status ($n=15$).

$TS_{Sum(SD)}$ fence, t	Sensitivity (%)	Specificity (%)	Yates' chi-squared test		Fisher's test	Bayes 95% CI
			X_Y	p-value	p-value	(2.5%, 97.5%)
2	100.0	12.5	0.005	0.945	1.000	−0.26, 0.42
2.5	100.0	37.5	1.356	0.244	0.200	−0.01, 0.57
3	100.0	50.0	2.558	0.110	0.077	0.08, 0.62
3.4	100.0	62.5	4.051	0.044	0.026	0.17, 0.67
3.5	100.0	75.0	5.904	0.015	0.007	0.24, 0.72
4	85.7	75.0	3.359	0.067	0.040	0.10, 0.65
4.5	71.4	75.0	1.637	0.201	0.132	−0.03, 0.59
5	57.1	87.5	1.641	0.200	0.119	−0.03, 0.58
5.5	28.6	87.5	0.017	0.897	0.569	−0.30, 0.45
7	28.6	100.0	0.744	0.388	0.200	−0.07, 0.52

indicate that a significant association exists for an overlapping range of fence values between 3.7 and 4.4. Hence, a statistically significant association is confirmed between $TS_{Sum(SD)}$ and reports of numbness at night. No attempt has been made to explore a relationship between $TS_{Sum(SD)}$ and "numbness effect" in view of the small number of subjects reporting this symptom (see Table II).

D. Association between threshold shift metric and pain and neck pain

A plot of the statistical test parameters to examine for any association between $TS_{Sum(SD)}$ and reports of "pain in the hands or arms" in Table II is shown in Fig. 3, where the presentation follows the scheme described for Figs. 1 and 2. By inspection of this diagram, it is evident that the three statistical tests all identify the same region of fence values close to 4 where, perhaps unexpectedly, a statistically significant association between $TS_{Sum(SD)}$ and subjects' reports of pain in the hands or arms exists.

A plot of the statistical test parameters to examine for any association between $TS_{Sum(SD)}$ and reports of "neck pain" in Table I is shown in Fig. 4, where the presentation again follows the scheme described previously. In this case, however, it is evident that the three statistical tests do not identify any fence values where a statistically significant association exists.

E. Diagnostic potential of metric for numbness and pain in the hands

While the summed normalized threshold shift has clearly been shown to be associated with the reports of numbness and pain in the hands of these subjects, it remains to be demonstrated whether the metric can usefully serve as a clinical test.

The performance of the metric as a test for numbness status is shown for different fence values in Table VI. Overall, the sensitivity of the test, that is, the proportion of subjects with numbness correctly identified, decreases while the specificity, that is, the proportion of subjects without numbness correctly identified, increases as the fence is increased from 1 to 7. However, it is known from the statistical tests for association between $TS_{Sum(SD)}$ and reports of numbness that only fence values in the range $3.3 \leqslant t < 3.9$ may be reliably used for a clinical test. It can thus be seen from Table VI

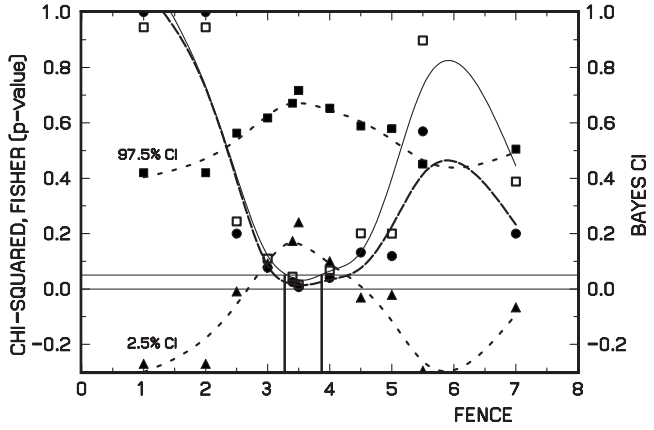


FIG. 1. Statistical tests for association between $TS_{Sum(SD)}$ fence values and numbness status. The Yates' chi-squared test p-value is shown by the continuous line, the Fisher's exact test p-value by the long-dashed line, and the Bayes CI (97.5% and 2.5%) by the short-dashed lines. The subject's data are plotted as open squares, filled circles, and filled squares and triangles, respectively ($n=15$).

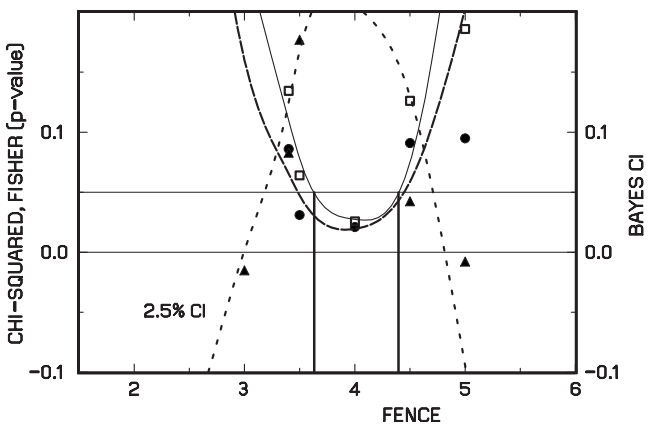


FIG. 2. Statistical tests for association between $TS_{Sum(SD)}$ fence values and numbness reported at night. The Yates' chi-squared test p-value is shown by the continuous line, the Fisher's exact test p-value by the long-dashed line, and the Bayes CI (2.5%) by the short-dashed line. The subject's data are plotted as open squares, filled circles, and filled triangles, respectively ($n=15$).

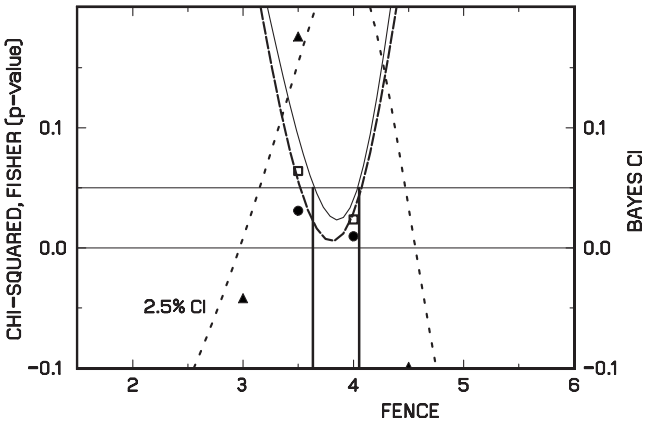


FIG. 3. Statistical tests for association between $TS_{Sum(SD)}$ fence values and pain in the hands. The Yates' chi-squared test p-value is shown by the continuous line, the Fisher's exact test p-value by the long-dashed line, and the Bayes CI (2.5%) by the short-dashed line. The subject's data are plotted as open squares, filled circles, and filled triangles, respectively ($n=15$).

that these fence values will yield sensitivities ranging from 100% to 85%, and specificities ranging from 63% to 75%, respectively.

As both sensitivity and specificity will be related to the fence values by continuous functions in the limit when $n \rightarrow \infty$, it is again instructive to summarize this information graphically. A clearer picture is obtained from the ROC curve, in which the sensitivity is plotted as a function of (1—specificity). This is shown in Fig. 5, where the ROC curve is plotted as a dashed line. The data from the subjects are shown by the filled circles. The likelihood ratio, the gradient of this curve, reaches a maximum for sensitivities from approximately 45% to 85%, revealing that the most useful test will be obtained with a fence value close to $t \sim 3.8$, where the association is also statistically significant. For these data, the likelihood ratio is 3.43.

The corresponding ROC curves for reports of numbness at night (continuous line) and pain in the hands (dashed line) are plotted in Fig. 6. From the statistical tests for association, only fence values in the range $3.6 \leqslant TS_{Sum(SD)} < 4.4$ for the

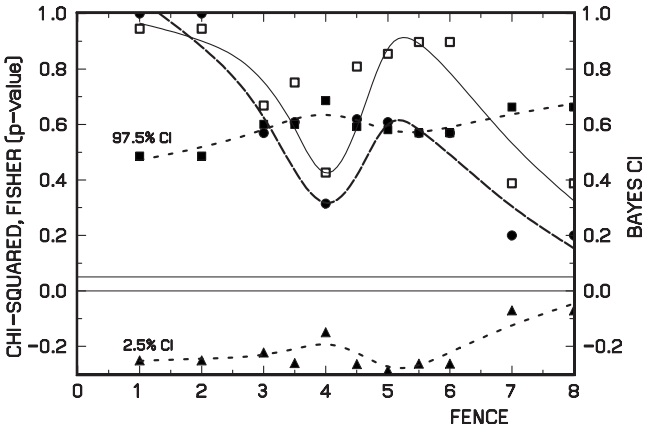


FIG. 4. Statistical tests for association between $TS_{Sum(SD)}$ fence values and neck pain. The Yates' chi-squared test p-value is shown by the continuous line, the Fisher's exact test p-value by the long-dashed line, and the Bayes CI (97.5% and 2.5%) by the short-dashed lines. The subject's data are plotted as open squares, filled circles, and filled squares and triangles, respectively ($n=15$).

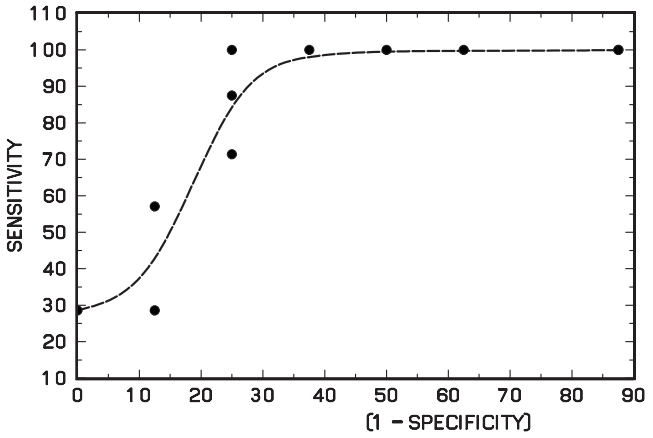


FIG. 5. ROC for subject-reported finger/hand numbness predicted from $TS_{Sum(SD)}$. The subject's data are plotted as filled circles ($n=15$).

former and $3.6 \leqslant TS_{Sum(SD)} < 4.1$, for the latter may be reliably used for a clinical test. The fence values for numbness at night correspond to sensitivities ranging from 100% to 80%, and specificities ranging from 67% to 78%, respectively, while those for pain correspond to sensitivities ranging from 100% to 83%, and specificities ranging from 63% to 75%. For these data, the likelihood ratios are 4.5 for numbness at night and 4.0 for pain in the region of statistical association between $TS_{Sum(SD)}$ and the relevant symptom. The most significant fence values are $t \sim 4.0$ for numbness at night and for pain in the hands.

IV. DISCUSSION

For peripheral neuropathies, present clinical practice tends to rely on electrically stimulated nerve conduction and so does not include the performance of the nerve endings, which are uniquely specialized to serve as transducers for different sensory modalities. In addition, by shock excitation of a nerve trunk, only the gross conduction features of the nerve in response to this artificial stimulus can be established, leaving unanswered questions concerning the normal biochemical function of nerve fibers.

A primary requirement for an effective QST is that it should be associated with the symptom it is intended to con-

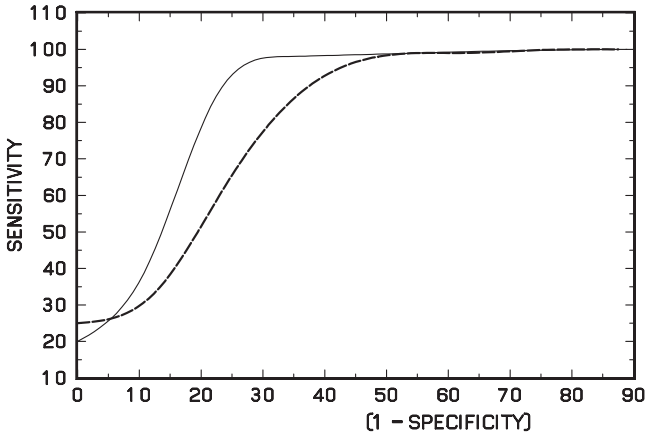


FIG. 6. ROC for subject-reported numbness at night (continuous line) and pain (dashed line) predicted from $TS_{Sum(SD)}$.

firm. The summed normalized threshold shift, constructed as described here, has clearly been shown to be related to reports by these subjects of numbness and pain in the hands using the statistical tests most commonly employed for evaluating the significance of associations in 2×2 contingency tables (Altman, 1991). It is plausible that the strength of the associations would have been reduced if the subjects' reports had not been confirmed, but our data do not permit us to assess this possibility. Even with this caveat, the associations are substantial and have been more rigorously demonstrated than in previous work. Indeed, there is only one study to our knowledge in which associations between mechanoreceptor-specific vibrotactile thresholds and equivalent symptom reports have been observed (Coutu-Wakulczyk *et al.*, 1997). A lack of association between vibrotactile thresholds and symptoms has been an equally common outcome of studies in which the QST has been employed (Shy *et al.*, 2003), as has already been discussed.

The crucial difference in our approach is, we believe, the differentiation between vibrotactile thresholds mediated by mechanoreceptor populations with known roles in the sense of touch (Merkel disks and Meissner corpuscles) (Bolanowski *et al.*, 1988; Johannson *et al.*, 1984; Johnson, 2001), as opposed to measurement procedures in which responses from mixed, or unknown, receptor populations are obtained. In the latter circumstances, responses from Pacinian corpuscles, which are acutely sensitive to vibration, usually dominate (Gescheider *et al.*, 2002). The thresholds from these receptors are strongly influenced by stimulus and stimulation characteristics (e.g., stimulus waveform, stimulating probe diameter, presence of a static surround, probe and surround contact force, and skin indentation by stimulator), as well as the subject's age (Verrillo, 1985). The Merkel disks and Meissner corpuscles that are believed to mediate the responses studied here, while more difficult to stimulate and hence establish their thresholds, are inherently five to six times *less* sensitive to a subject's age than the Pacinian corpuscles (see Brammer *et al.*, 2007a). Moreover, formation of the threshold *shift* from the mean threshold recorded from the hands of healthy persons at the same stimulation frequency further mitigates the effects of aging, as the mean thresholds can be chosen for persons of the same age as the subject.

The small number of subjects in this study (15 persons for whom both vibrotactile thresholds and symptoms were obtained) inevitably restricts the direct calculation of parameters for assessing the performance of the QST. This is evident from Table VI, where the sensitivity and specificity for numbness status have been calculated for different boundaries between the absence and presence of this symptom. The increment in fence values listed in the table coincides with one subject changing categories and therefore includes all meaningful values derivable from this data set. As has already been discussed, all the statistical tests consistently identify essentially the same fence values for which statistically significant associations exist, though the tests identify somewhat different ranges of fence values. Interpolation between the fence values listed in the table is possible by plotting the results graphically, which is equivalent to assuming the results were obtained from a much larger number of subjects. The results have been presented in this form, showing both the data points from our subjects and the interpolated curve fits to the data. Reference to Figs. 1–3 reveals that the Bayes 2.5% CI (short dashes) provides the largest range of significant fence values for our data (i.e., the range for which the CI is greater than zero). The p-values produced by Fisher's exact test provide a range of statistically significant fences values that are always slightly greater than those produced by Yates' chi-squared test, but the difference is not believed to be of consequence. In each case, there exists a common range in the interpolations in which the fence values are judged statistically significant by all the tests (shown by the vertical lines).

TABLE VII. Statistically significant fence values (expressed in units of SDs).

Symptom	Fence, t	
	Lower	Upper
Numbness status	3.3	3.9
Numbness at night	3.6	4.4
Pain in the hands	3.6	4.1

The calculated range of fence values so identified is summarized in Table VII for the symptoms for which statistically significant associations exist with $TS_{Sum(SD)}$, and can be seen to be somewhat less for numbness status than for reports of numbness at night or pain in the hands. An interpretation that is compatible with this result is that numbness at night represents a symptom requiring more loss of tactile sensitivity. The similarity between fence values for numbness at night and pain in the hands could imply that the two symptoms are related. An alternative interpretation would be that the test is responding similarly to the performance of the pain and tactile end organs. As the sensory test is specific to the tactile end organs, this interpretation would appear unlikely. A more plausible interpretation would be that the test is responding to changes in the nerve fibers innervating the sensory end organs. It is well known that the mechanoreceptors are innervated by type Aβ mylineated nerve fibers, while the nociceptors are innervated by smaller diameter mylineated fibers (Aδ) and unmylineated fibers (Lawson, 2005). Thus it is curious, and perhaps coincidental, that the sensory test developed here may reflect changes in several types of nerve endings and nerve fibers.

Finally, it is interesting to note by comparing Tables V and VII that the most significant fence values when using the magnitude of $TS_{Sum(SD)}$ as a QST correspond almost exactly to the boundary between normal and abnormal thresholds, that is thresholds for which the probability of occurrence in the hands of healthy persons is $p=0.05$. This boundary was employed arbitrarily, and perhaps fortuitously, in a previous analysis (Brammer *et al.*, 2007b). The significance, if any, to attach to this observation must await the results of further studies. Also as yet unresolved is whether neurological symptoms will be found to be associated more generally with threshold shifts in both SAI and FAI receptor populations, as in the present group of subjects. It is possible that in some

neuropathies the symptoms will be found to reflect changes in acuity of just one receptor population. The large correlation between the SAI and FAI threshold shifts recorded in a fingertip (Table V) preclude such a distinction to be made here.

V. CONCLUSIONS

The summed normalized threshold shift, constructed as described here, has been shown to be related to reports of numbness and pain in the hands using the statistical tests most commonly employed for evaluating the significance of associations in 2×2 contingency tables. The strengths of the associations are statistically significant ($p<0.02$ for numbness, and $p<0.025$ for pain) and have been more rigorously demonstrated than in previous work.

The small number of subjects in this study restricts the direct calculation of parameters for assessing the performance of a QST based on fence value for $TS_{Sum(SD)}$. Interpolation between the fence values has been performed by plotting the results graphically to overcome this limitation. For each of the statistical tests employed, a common range of fence values can be identified that are judged statistically significantly by all the tests. The range of fence values so identified is somewhat less for numbness status than for reports of numbness at night or pain in the hands. The similarity between fence values for numbness at night and pain in the hands could imply that the test is responding to changes in the nerve fibers innervating the sensory end organs, rather than to the status of the end organs themselves. The most significant fence values for a QST correspond to the boundary between normal and abnormal thresholds, that is, thresholds for which the probability of occurrence in the hands of healthy persons is $p=0.05$.

ACKNOWLEDGMENTS

The authors wish to acknowledge the provision of financial assistance by the Finnish National Board of Forestry and the Finnish Forestry fund.

APPENDIX

The chi-squared measure of association for 2×2 contingency tables was derived by Yule as, with nomenclature given in Table I (see, for example, Kendall and Stuart, 1961):

$$V=\frac{ad-bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}. \tag{A1}$$

If the proposed test, based on a given fence value t, is correct, then true positives and true negatives will be determined correctly. Hence a and d will take high values compared to b and c. In consequence, Yule's measure of association [Eq. (A1)] will be strongly positive. One limitation of the classical formula arises if any of the row or column totals is 0. The following algorithm avoids this limitation by calculating Bayes estimate of the CI (Gellman *et al.*, 2004).

A suitable likelihood model for the 2×2 contingency table is a multinomial model Multinomial(n,p_a,p_b,p_c,p_d), where n denotes the sample size, p_a denotes the probability of true positive, p_b denotes the probability of false positive, p_c denotes the probability of false negative, and p_d denotes the probability of true negative. Now Eq. (A1) can be rewritten in terms of probabilities as

$$V_p=\frac{p_ap_d-p_bp_c}{\sqrt{(p_a+p_b)(p_c+p_d)(p_a+p_c)(p_b+p_d)}}. \tag{A2}$$

For modeling purposes, the non-informative Dirichlet distribution is chosen as the prior distribution. This has a three-fold purpose. First, the choice of the Dirichlet distribution for the multinomial model is convenient in that the posterior distribution is again the Dirichlet distribution. Second, generating a sample from the Dirichlet distribution is straightforward. Finally, by choosing the non-informative Dirichlet prior distribution, the measure of association in computed in a conservative way, in the sense that if the measure is positive then it is definitely positive with a very high confidence. The posterior distribution is described as follows:

$$\text{posterior} \propto \text{likelihood} \times \text{prior},$$

where posterior is Dirichlet$(\alpha_a+a,\alpha_b+b,\alpha_c,+c,\alpha_d+d)$, the likelihood is Multinomial(n,p_a,p_b,p_c,p_d), and the prior distribution is Dirichlet$(\alpha_a,\alpha_b,\alpha_c,\alpha_d)$. As the choice of the prior distribution is non-informative, the values of $\alpha_a,\alpha_b,\alpha_c,\alpha_d$ are arbitrarily assigned to $\alpha_a=\alpha_b=\alpha_c=\alpha_d=0.25$.

The algorithm then has the following steps:

(a) Step 0: Define an array of length of the simulation size, say, N.
(b) Step 1: Generate a sample $\{p_a,p_b,p_c,p_d\}$ from Dirichlet$(\alpha_a+a,\alpha_b+b,\alpha_c+c,\alpha_d+d)$.
(c) Step 2: Compute V_p using Eq. (A2) and record that value in a cell of the array.
(d) Step 3: Go to Step 1 and repeat the process N times (where $N=20\,000$, i.e., 20 000 samples from the posterior distribution are simulated).
(e) Step 4: Finally, once the entire $N(=20\,000)$ estimates of the association have been computed, the average becomes the Bayes estimate of the association from which the 95% CI is derived.

Altman, D. G. (**1991**). *Practical Statistics for Medical Research* (Chapman & Hall, London/CRC, Boca Raton, FL).

Bolanowski, S. J., Gescheider, G. A., Verrillo, R. T., and Checkosky, C. M. (**1988**). "Four channels mediate the mechanical aspects of touch," J. Acoust. Soc. Am. **84**, 1680–1694.

Brammer, A. J., and Piercy, J. E. (**1991**). "Measuring vibrotactile perception thresholds at the fingertips of power-tool operators," in *Proceedings of the United Kingdom Informal Group on Human Response to Vibration*, Buxton, United Kingdom, pp. 1–7.

Brammer, A. J., Peterson, D. R., Cherniack, M. G., and Diva, U. A. (**2006**). "Temporary changes in mechanoreceptor-specific vibrotactile perception to stimuli simulating impact power tools," in *Proceedings Inter-Noise 2006* (Noise Control Foundation, New York).

Brammer, A. J., Piercy, J. E., Nohara, S., Nakamura, H., and Auger, P. L. (**1993**). "Age-related changes in mechanoreceptor-specific vibrotactile thresholds for normal hands," J. Acoust. Soc. Am. **93**, 2361.

Brammer, A. J., Piercy, J. E., Pyykkö, I., Toppila, E., and Starck, J. (**2007a**). "Method for detecting small changes in vibrotactile perception threshold related to tactile acuity," J. Acoust. Soc. Am. **121**, 1238–1247.

Brammer, A. J., Sutinen, P., Diva, U. A., Pyykkö, I., Toppila, E., and Starck,

J. (**2007b**). "Application of metrics constructed from vibrotactile thresholds to the assessment of tactile sensory changes in the hands," J. Acoust. Soc. Am. **122**, 3732–3742.
Coutu-Wakulczyk, G., Brammer, A. J., and Piercy, J. E. (**1997**). "Association between a quantitative measure of tactile acuity and hand symptoms reported by operators of power tools," J. Hand Surg. [Am] **22**, 873–881.
Dyck, P. J., O'Brien, P. C., Johnson, D. M., Klein, C. J., and Dyck, P. J. B. (**2005**). "Quantification of sensory abnormalities," in *Peripheral Neuropathy*, edited by P. J. Dyck and P. K. Thomas, 4th ed. (W. B. Saunders, Philadelphia), Chap. 42, pp. 1063–1093.
Gelman, A., Carlin, J., Stern, H., and Rubin, D. (**2004**) *Bayesian Data Analysis* (Chapman & Hall, London/CRC, Boca Raton, FL).
Gescheider, G. A., Bolanowski, S. J., Pope, J. V., and Verrillo, R. T. (**2002**). "A four channel analysis of the tactile sensitivity of the fingertip: Frequency selectivity, spatial summation, and temporal summation," Somatosens Mot Res. **19**, 114–124.
ISO 13091-1. (**2001**). Mechanical vibration: Vibrotactile perception thresholds for the assessment of nerve dysfunction—Part 1: Methods of measurement at the fingertips (International Organization for Standardization, Geneva).
ISO 13091-2. (**2003**). Mechanical vibration: Vibrotactile perception thresholds for the assessment of nerve dysfunction—Part 2: Analysis and interpretation of measurements at the fingertips (International Organization for Standardization, Geneva).
Jetzer, T., Haydon, P., and Reynolds, D. (**2003**). "Effective intervention with ergonomics, antivibration gloves, and medical surveillance to minimize hand-arm vibration hazards in the workplace," J. Occup. Environ. Med. **45**, 1312–1317.
Johansson, R. S., and Westling, G. (**1984**). "Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects," Exp. Brain Res. **56**, 550–564.
Johnson, K. O. (**2001**). "The roles and functions of cutaneous mechanoreceptors," Curr. Opin. Neurobiol. **11**, 455–461.
Kendall, M. G., and Stuart, A. (**1961**). *Theory of Advanced Statistics* (Hafner, New York).
Koskimies, K., Pyykkö, I., Starck, J., and Inaba, R. (**1992**). "Vibration syndrome among Finnish forest workers between 1972 and 1990," Int. Arch. Occup. Environ. Health **64**, 251–256.
Lawson, S. N. (**2005**). "The peripheral sensory nervous system: Dorsal root ganglion neurons," in *Peripheral Neuropathy*, edited by P. J. Dyck and P. K. Thomas, 4th ed. (W. B. Saunders, Philadelphia), Chap. 8, pp. 163–202.
Lundborg, G., Dahlin, L. B., Lundström, R. L., Necking, L. E., and Strömberg, T. (**1992**). "Vibrotactile function of the hand in compression and vibration-induced neuropathy: Sensibility index—A new measure," Scand. J. Plast. Reconstr Surg. Hand Surg. **26**, 275–279.
Lundström, R. L., Lundborg, G., and Strömberg, T. (**1992**). "Vibrotactile perception threshold measurements for diagnosis of sensory neuropathy: Description of a reference population," Int. Arch. Occup. Environ. Health **64**, 201–207.
Shy, M. E., Frohman, E. M., So, Y. T., Arezzo, J. C., Cornblath, D. R., Giuliani, M. J., Kincaid, J. C., Ochoa, J. L., Parry, G. J., and Weimer, L. H. (**2003**). "Quantitative sensory testing," Neurology **60**, 898–904.
Sutinen, P., Toppila, E., Starck, J., Brammer, A. J., Zou, J., and Pyykkö, I. (**2006**). "Hand-arm vibration syndrome with use of anti-vibration chain saws: 19-year follow-up of forestry workers," Int. Arch. Occup. Environ. Health **79**, 665–671.
Verrillo, R. T. (**1985**). "Psychophysics of vibrotactile stimulation," J. Acoust. Soc. Am. **77**, 225–232.
Yarnitsky, D. (**1997**). "Quantitative sensory testing," Muscle Nerve **20**, 198–204.

Encapsulated contrast microbubble radial oscillation associated with postexcitation pressure peaks

M. D. Santin
UPMC Univ Paris 06, UMR 7623, LIP, F-75005 Paris, France and CNRS, UMR 7623, Laboratoire d'Imagerie Paramétrique, F-75006 Paris, France

D. A. King
Department of Mechanical Science and Engineering, University of Illinois, 1206 W. Green St., Urbana, Illinois 61801

J. Foiret
UPMC Univ Paris 06, UMR 7623, LIP, F-75005 Paris, France and CNRS, UMR 7623, Laboratoire d'Imagerie Paramétrique, F-75006 Paris, France

A. Haak and W. D. O'Brien, Jr.
Bioacoustics Research Laboratory, Department of Electrical and Computer Engineering, University of Illinois, 405 North Mathews, Urbana, Illinois 61801

S. L. Bridal
UPMC Univ Paris 06, UMR 7623, LIP, F-75005 Paris, France and CNRS, UMR 7623, Laboratoire d'Imagerie Paramétrique, F-75006 Paris, France

(Received 24 July 2009; revised 23 November 2009; accepted 30 November 2009)

This work combines modeling and experiment to assess encapsulated microbubble oscillations associated with broadband pressure peaks detected after microbubble excitation (postexcitation signals). Data were acquired from albumin-shelled and phospholipid-shelled microbubbles using a passive cavitation detector consisting of a confocally aligned 2.8-MHz transmitter and 13-MHz receiver. Microbubbles in weak solutions were insonified with a 5-cycle pulse at a peak rarefactional pressure of 2.0 ± 0.2 MPa. For each microbubble type, at least 100 received signals were identified as individual-microbubble responses. The average broadband noise from signals with postexcitation response was 4.2–7.2 dB higher than from signals without postexcitation. Pressure-time responses for each microbubble type were simulated using the model by Marmottant *et al.* [J. Acoust. Soc. Am. **118**, 3499–3505 (2005)], with insonification conditions matching the experiment. Increased broadband noise predicted for microbubbles with postexcitation response was consistent with that observed experimentally (4.0–8.9 dB). The model predicted that postexcitation signals occur only when the radial oscillation exceeds both the shell break-up threshold and twice the initial radius (free bubble inertial cavitation threshold). © *2010 Acoustical Society of America.*
[DOI: 10.1121/1.3277216]

PACS number(s): 43.80.Vj, 43.35.Ei, 43.35.Yb, 43.60.Bf [CCC] Pages: 1156–1164

I. INTRODUCTION

Ultrasound contrast agents are microbubbles consisting of a gas core encapsulated in a shell that can be composed of lipids, albumin, or polymer. Commercially available agents typically contain polydisperse size distributions with a mean radius near the resonant size for diagnostic imaging ($\sim 2-4$ μm). Useful signal at diagnostic imaging frequencies is generally attributed to microbubbles greater than 1 μm in diameter.[1]

Novel therapeutic techniques have been put into place that take advantage of ultrasound's unique interaction with contrast microbubbles. For example, removal of blood clots[2] and acoustic activation of drug-laden microbubbles[3,4] have been demonstrated using high power destruction of microbubbles. Blood brain barrier disruption has been activated acoustically after injection of contrast agents by applying 10 ms, 260 kHz pulse trains with 0.29–0.57 MPa peak rarefactional pressure amplitude (PRPA).[5] It is widely thought that the inertial cavitation of contrast microbubbles plays a major role in producing therapeutic effects. For example, ultrasound exposure (1–10 cycles at 1 MHz, 0.48–1.32 MPa) with Levovist microbubbles was found to increase sonoporation of DNA into cultured cells, and the effect increased as a function of increasing inertial cavitation dose.[6] It has been shown that using high-intensity focused ultrasound exposures (1.8 MPa at 1.1 MHz), bubble-enhanced heating in biological media can be obtained.[7] Measurements of the inertial cavitation and modeling of the microbubble response attributed the heating to absorption of broadband emissions produced by microbubble inertial cavitation.[7]

Inertial cavitation occurs when a microbubble's collapse is dominated by the surrounding medium's inertia.[8] Passive cavitation detection (PCD) is widely applied in research to detect and obtain an indication of the level of inertial cavitation. Detection of inertial cavitation is sometimes based on

features of the voltage-time signal such as increases in the rms or peak voltages.[9] Many other PCD-based studies have applied criteria based on increased voltage levels or broadband power to characterize inertial cavitation thresholds.[6,10,11] Although, the specific harmonic content of PCD-received signals and their modification when inertial cavitation occurs has remained largely unexplored, it has been demonstrated that subharmonic, second, and third harmonic peaks may also be modified near inertial cavitation thresholds.[5,12]

An important feature of free bubble collapse has been exploited in lithotripsy to evaluate inertial cavitation thresholds. It has been demonstrated[13] that when a shock wave arrives on a cavitation nuclei, a very rapid phase of compression is followed by a phase of much slower gas-body expansion. After reaching some maximum radial expansion, the bubble compresses again violently. This implosion can occur well after the original excitation wave has passed and gives rise to a broadband pressure peak (postexcitation signal).

The postexcitation signal from encapsulated contrast microbubbles has received little attention. Destruction of contrast agent microbubbles has previously been observed *in vitro* using high temporal and spatial resolution microscopy of bubble-wall movement during insonation. For example, using this approach, Chomas *et al.*[14] observed destruction of lipid-shelled microbubbles at 5 MHz (1.5 cycle pulse). Violent fragmentation of bubbles was observed in more than 55% of the bubbles at an incident PRPA of 2.6 MPa. Optical observation showed that microbubble destruction created smaller free bubbles. Such bubbles could be nuclei for postexcitation inertial collapse. In other works, postexcitation signals can be seen in voltage-time responses from ruptured contrast microbubbles acquired with a PCD system, but no comment on these signals is made by the authors.[15] In previous work, we utilized PCD to evaluate isolated Optison microbubbles.[16] Postexcitation signals were observed in the voltage-time response when incident PRPA exceeded a threshold value. A link between postexcitation signals and encapsulated microbubble shell rupture was demonstrated using the model presented by Morgan *et al.*[17] by setting the shell elasticity and viscosity terms to zero, once an Optison microbubble (equilibrium radius of 2 μm) exceeded 3 μm in radius. Once shell parameters were thus removed, the bubble radius underwent large oscillations (radial expansion to more than 5 μm) with repeated cycles of inertial collapse that continued even after the driving pulse ceased to act on the microbubble. Postexcitation signals were thus shown to occur for an encapsulated ultrasound contrast microbubbles when inertial cavitation was associated with microbubble break-up. Modeling in previous work, however, was limited to a unique example at an arbitrarily selected break-up radius and did not examine the dependence of postexcitation emission occurrence on the initial microbubble radius or on the extent of the microbubble oscillation. The broadband noise level specifically associated with the postexcitation response was not evaluated in the previous work, and only albumin-shelled Optison microbubbles were considered.

This work combines modeling and experiment to probe whether postexcitation signals predictably occur when encapsulated contrast agent microbubble shells break up, and to determine how much additional broadband noise is associated with microbubbles exhibiting postexcitation signals. Two lipid-shelled microbubbles are considered in addition to the albumin-shelled agent studied in previous work.

Pressure-time and spectral responses were modeled in this work using the model by Marmottant *et al.*,[18] which specifically describes conditions for shell break-up and resulting surface tension modifications. This model applies even to large amplitude oscillations. Modeling was performed for Definity, Sonovue, and Optison microbubbles with a 2.8 MHz, 5-cycle incident pulse at 2.0 MPa PRPA. The incident pulse used in the model was measured with a hydrophone at the focal position of the transmit transducer to consider nonlinear propagation effects. For each radius-time simulation, radial expansion was related to the microbubble break-up radius, and to occurrence of postexcitation signals on the associated pressure-time simulation.

PCD signals from dilute solutions of Optison, Definity, and Sonovue microbubbles were acquired at a PRPA of 2.0 ± 0.2 MPa (measure mean and uncertainty). At this PRPA, the majority of microbubbles should rupture.[14] This provided PCD data with sufficient signal to noise ratio for detection of both microbubbles with postexcitation signals and weaker responses from microbubbles without postexcitation signals. Comparison of modeled spectra with the experimental spectra allowed interpretation of broadband noise increase with postexcitation signals in terms of the underlying microbubble response.

II. MATERIALS AND METHODS

A. Microbubbles

Three contrast agent microbubbles were used: Optison (Amersham Health, Milwaukee, WI), Sonovue (Bracco Diagnostics, Inc., Geneva, Switzerland), and Definity (Bristol-Myers Squibb Medical Imaging Inc., New York). Both Optison and Definity have an octofluoropropane C_3F_8 gas core, whereas Sonovue has a sulfurhexafluoride SF_6 core. Optison's shell is made of human serum albumin and N-acetyltryptophan caprylic acid. Sonovue and Definity have shells based on different phospholipids.[19]

B. Theoretical modeling of microbubble oscillation

The model proposed by Marmottant *et al.*[18] to describe behavior of lipid coated microbubbles was used. It was derived to describe even large amplitude oscillations and specifically to describe surface tension modifications arising during expansion and at shell break-up. This model considers that the shell of the bubble buckles for bubble radii $R < R_{\text{buckling}}$ and initially breaks up for $R > R_{\text{break-up}}$. This model leads to the dynamic equation

$$\rho_l\left(R\ddot{R}+\frac{3}{2}\dot{R}^2\right)=\left[P_0+\frac{2\sigma(R_0)}{R_0}\right]\left(\frac{R}{R_0}\right)^{-3\kappa}\left(1-\frac{3\kappa}{c}\dot{R}\right)-\frac{2\sigma(R)}{R}-\frac{4\mu\dot{R}}{R}-\frac{4\kappa_S\dot{R}}{R^2}-[P_0+P_{\mathrm{ac}}(t)], \quad (1)$$

where ρ_l is the density of the surrounding liquid, P_0 is the ambient pressure, κ is the polytropic gas exponent, c is the speed of sound in the liquid, μ is the liquid viscosity, and κ_s is the shell surface viscosity. The acoustic pressure as a function of time t, $P_{\mathrm{ac}}(t)$, was measured with a calibrated hydrophone in the configuration used in the PCD experiments described in Sec. II C. This experimentally determined $P_{\mathrm{ac}}(t)$ describes the incident pulse after it has undergone nonlinear propagation distortion in the transducer-to-microbubble water path. The surface tension $\sigma(R)$ is modeled to vary as a function of the value of R as

$$\sigma(R)=\begin{cases}0 & \text{if } R\le R_{\mathrm{buckling}}\\ \chi\left(\dfrac{R^2}{R_{\mathrm{buckling}}^2}-1\right) & \text{if } R_{\mathrm{buckling}}\le R\le R_{\text{break-up}}\\ \sigma_{\mathrm{water}}=0.073\ \mathrm{N/m} & \text{if broken and } R\ge R_{\mathrm{ruptured}},\end{cases} \quad (2)$$

where χ is the shell elastic compression modulus.

For large oscillations, the radius-dependent surface tension initially varies within an elastic regime until a critical break-up tension is reached at $R=R_{\text{break-up}}$. Although initial $R_{\text{break-up}}$ can occur when $\sigma(R)$ is greater than σ_{water} due to shell cohesion, the surface tension relaxes to that of an uncoated microbubble after break-up and remains at this value for all $R>R_{\mathrm{rupture}}$. The surface tension reversibly returns to the elastic regime during the compression phase when $R<R_{\mathrm{rupture}}$ and $R\ge R_{\mathrm{buckling}}$, and returns to the uncoated microbubble surface tension value (σ_{water}) during each subsequent expansion phase when $R\ge R_{\mathrm{rupture}}$. R_{rupture} is generally smaller than $R_{\text{break-up}}$, and is defined as the radius reached during each oscillation following initial break-up when $\sigma(R)$ is equal to σ_{water}.

The pressure radiated by the bubble at a distance r becomes[20]

$$P(t)=\frac{\rho_l}{r}(\ddot{R}R^2+2R\dot{R}^2). \quad (3)$$

The response $P(t)$ from Eq. (3), was convolved with the impulse response of the received transducer, as measured with a calibrated hydrophone, to provide results comparable to experimental data.

It has been shown that albumin-shelled Optison microbubbles buckle, as described in the Marmottant model.[21] Furthermore, it has been analytically determined that surface tension for a thin solid shell layer in the elastic regime varies as $\chi(R^2/R_0^2-1)$ and assumed that R_0 is approximately equal to R_{buckling}.[18] Therefore, Eqs. (1)–(3) were also applied to simulate pressure-time responses for Optison. However, Eqs. (1) and (2), based on behavior of lipid-based shells, assume that shell fragments recombine across the microbubble surface after shell break-up each time the microbubble recontracts to $R<R_{\mathrm{rupture}}$. When Optison's albumin shell breaks, the remains are not likely to behave in the same manner as for a more flexible lipid-shelled agent.[14] Therefore, surface tension conditions described in Eq. (2) were modified to permanently fix $\sigma(R)=\sigma_{\mathrm{water}}$ and to suppress the shell viscosity term ($\kappa_s=0$), once $R=R_{\text{break-up}}$ was reached for the first time, giving expressions for definitive shell break-up as

$$\sigma_{\text{def break-up}}(R)=\begin{cases}0 & \text{if } R\le R_{\mathrm{buckling}} \text{ and shell not broken}\\ \chi\left(\dfrac{R^2}{R_{\mathrm{buckling}}^2}-1\right) & \text{if } R_{\mathrm{buckling}}\le R\le R_{\text{break-up}} \text{ and shell not broken}\\ \sigma_{\mathrm{water}}=0.073\ \mathrm{N/m} & \text{once broken.}\end{cases} \quad (4)$$

This modified equation, used to model the radius-time response for Optison microbubbles, thus considers the microbubble to be permanently uncoated following the initial shell break-up.

For Definity, a shell elastic compression modulus χ of 0.38 N/m and a shell surface dilatational viscosity κ_s of 2.40×10^{-9} kg/s were used.[22] For Sonovue, values of 0.55 N/m for χ and 7.20×10^{-9} kg/s for κ_s were chosen according to Gorce *et al.*[23] For Optison, values of 0.93 N/m for χ and 7.65×10^{-8} kg/s for κ_s were chosen, assuming a shell thickness of 15 nm, according to Chatterjee and Sarkar.[24] The break-up radius $R_{\text{break-up}}$ is related to the maximum surface tension limit of the microbubble $R_{\text{break-up}}=R_{\mathrm{buckling}}\sqrt{1+\sigma_{\text{break-up}}/\chi}$. The value of $R_{\text{break-up}}$ used herein ($1.2R_0$) was selected to be within the range of values obtained by Marmottant *et al.*[18] by fitting the shell parameters in their model to radius-time curves measured for BR14 (a lipid-shelled microbubble) response. Radius-time and pressure-time $P_f(t)$ responses were simulated for microbubbles at a driving PRPA of 2.0 MPa (2.8 MHz, 5 cycles). Microbubble size-ranges were selected within accepted ranges[1] that included radii without postexcitation signals on the modeled pressure-time responses: Definity 0.4–4.6 μm, Sonovue 0.6–5 μm, and Optison 1.8–5 μm. To estimate the spectral response for each microbubble radius, the Fourier transform of each pressure-time response

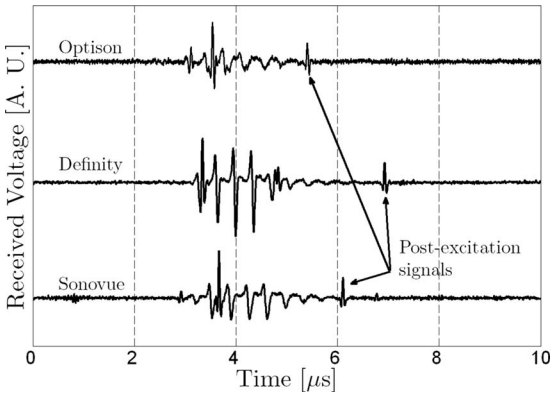


FIG. 1. Typical passive cavitation detection signals presenting postexcitation events for Optison, Definity, and Sonovue microbubbles (incident PRPA = 2.0 ± 0.2 MPa).

$P_f(t)$ was calculated. Average spectra were calculated from responses for different radii using weighting-factors describing a Gaussian size distribution within the selected ranges (mean ± standard deviations of 2.5 ± 0.8, 2.8 ± 0.9, and 3.4 ± 0.6 μm for Definity, Sonovue, and Optison, respectively) with a size-integration technique for simulation of contrast agent spectral density, as demonstrated by Zheng *et al.*[25]

C. Passive cavitation detector

Dilute solutions of ultrasound contrast agents were studied using a PCD with a 2.8-MHz transmitter and 13-MHz receiver (−6 dB bandwidth from 9.3 to 17.6 MHz). Use of the higher frequency receive transducer allows sensitive detection of broadband emissions. Details describing this system have been provided previously.[16] Incident PRPA was characterized in a separate experiment using a calibrated PVDF bilaminar shielded membrane hydrophone (0.5-mm diameter, 699/1/00001/100, GEC Marconi Ltd., Great Baddow, UK) placed at the focal point. The dilution of the microbubbles was such that each signal received by the 13-MHz receiver should, on average, arise from a single microbubble within the confocal volume determined by the intersecting volumes of the two −6 dB beam widths. A total of n, independent, microbubble responses ($n \geq 100$) were acquired with an incident PRPA of 2.0 ± 0.2 MPa. A 5-cycle transmit pulse was used throughout the study.

D. Data analysis

Postexcitation events were observed in some of the voltage-time signals acquired for each type of microbubble (Fig. 1). Microbubble responses were classified as (1) responses with a clear principal response but no postexcitation signal, (2) responses with a clear principal response followed

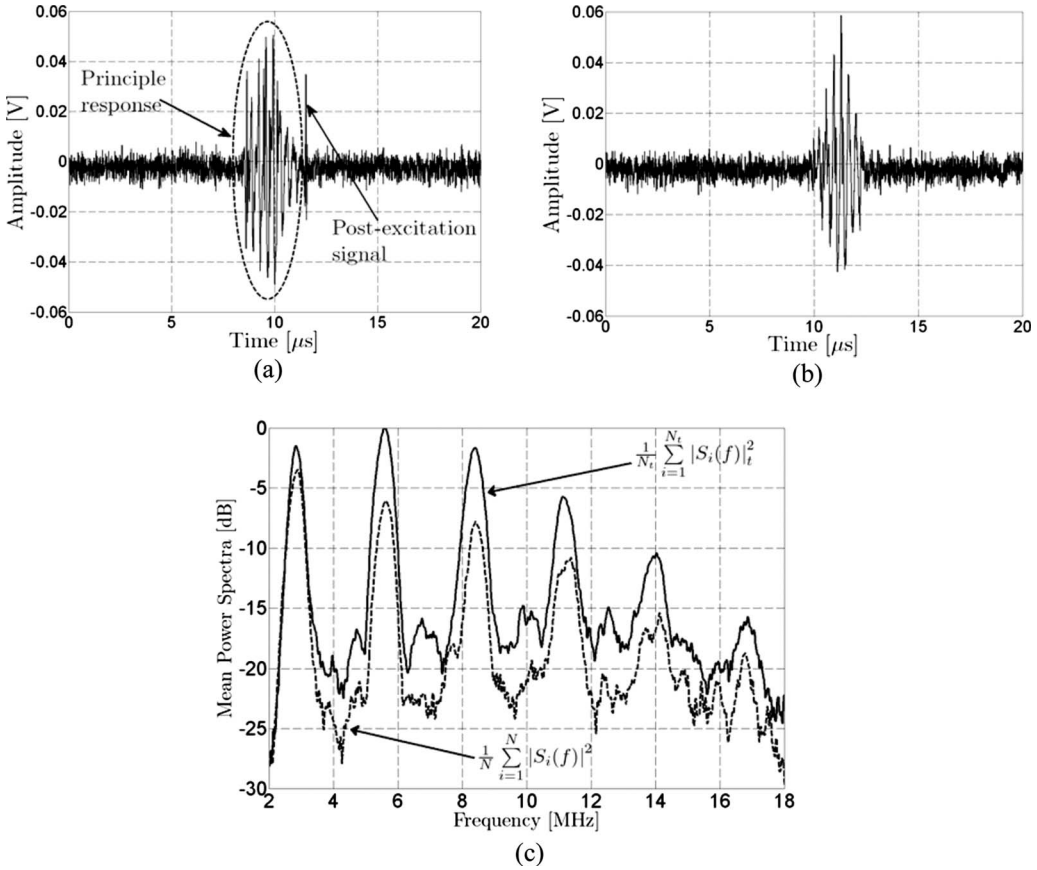


FIG. 2. Illustration of group selection criteria and average spectra. The signals shown were acquired with the passive cavitation detector from Definity insonified at an incident PRPA of 2.0 ± 0.2 MPa (5 cycles, 2.8 MHz). (a) PCD signals presenting a postexcitation signal after the principle response. (b) PCD signal showing only the principle response; no postexcitation signal. (c) Average power spectral density for signals with postexcitation response (N_t = 18) and signals without postexcitation response (N = 17).

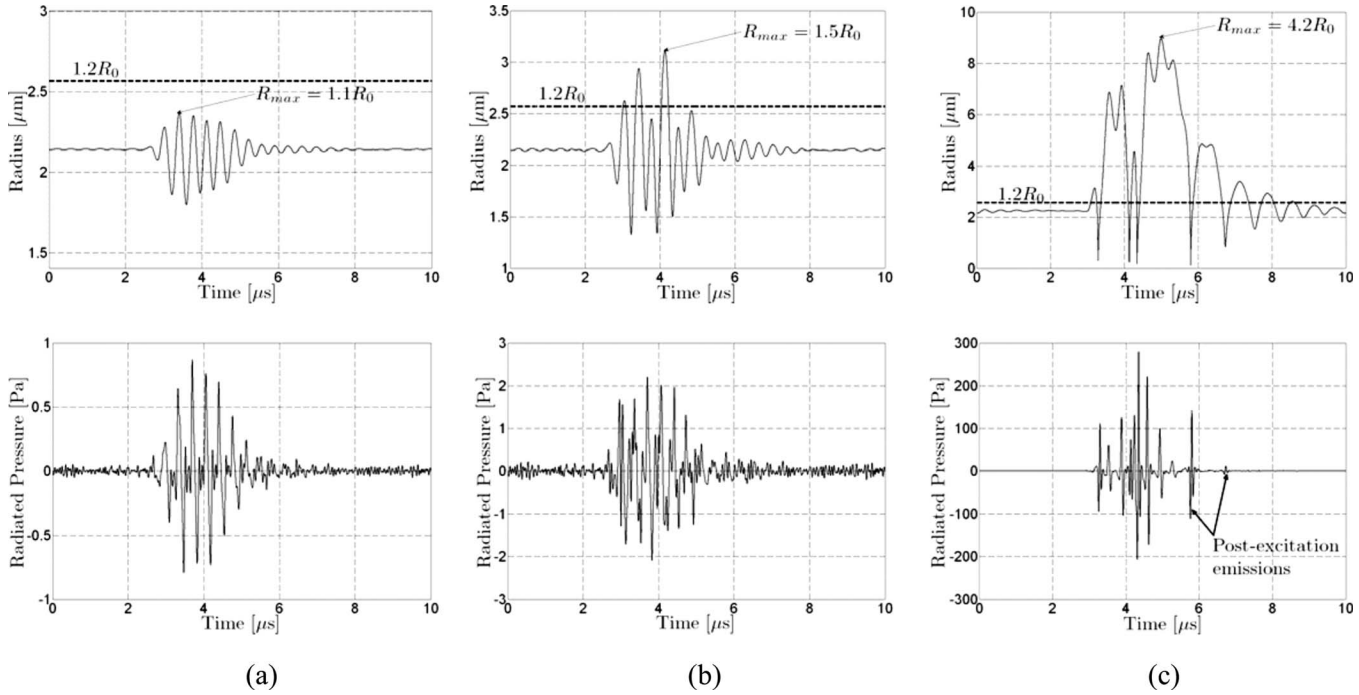


FIG. 3. Simulated radius-time and radiated pressure for a Definity microbubble at resonance radius (R_0=2.14 μm) with a break-up radius $R_{\text{break-up}}=1.2R_0$. Incident pulse at 2.8 MHz with 5 cycles. (a) Incident PRPA of 100 kPa. Radial expansion does not exceed the break-up radius $R_{\max}=1.1R_0$. Radiated pressure does not exhibit a postexcitation signal. (b) Incident PRPA of 400 kPa. Radial expansion exceeds the break-up radius. Radiated pressure does not exhibit a postexcitation signal. (c) Incident PRPA of 2 MPa. Radial expansion exceeds the break-up radius $R_{\max}=4.2R_0$. Postexcitation signals are observed.

by postexcitation signal, or (3) responses removed from further analysis that could not be classified in either groups (e.g., because there was no bubble response with a signal to noise ratio greater than 6 dB, there was unclear distinction of postexcitation signals, or there were possible multiple bubbles). The average spectral density was calculated for each group of N_t signals with postexcitation emissions. The average spectral density was also calculated for the corresponding group of N signals with no postexcitation emission. The group selection criteria and steps toward calculation of average parameters are illustrated schematically in Fig. 2. Broadband noise levels were compared between experimental and modeled spectra within the −6 dB receiving transducer's bandwidth (9.3–17.6 MHz).

III. RESULTS

A. Postexcitation signals in modeled microbubble responses

The microbubble responses simulated using Eqs. (1)–(3) for Definity at a fixed initial microbubble radius and break-up threshold, but with different incident pulse pressures, are presented in Fig. 3. This illustrates the response predicted by the model for the same microbubble under different conditions of radial oscillation. Shell parameters used are described in Sec. II B. As described in Sec. II B, incident pulses (2.8 MHz, 5 cycles) used in the simulation were those measured with a hydrophone, and the pressure-time response was filtered by the response of the receiver used in the PCD system. In Fig. 3(a), the pressure-time curve does not show a postexcitation signal when the shell is not broken (incident PRPA 100 kPa, $R_{\max}=1.1R_0$). In Fig. 3(c), the pressure-time curve has a postexcitation signal when the microbubble responds very strongly after shell break-up (incident PRPA 2 MPa, $R_{\max}=4.2R_0$). However, the pressure-time curve does not exhibit a postexcitation signal for the more modest oscillation shown in Fig. 3(b), even though shell break-up occurs (incident PRPA 400 kPa, $R_{\max}=1.5R_0$).

The postexcitation responses predicted with the model were considered as a function of the maximum microbubble expansion relative to the initial radius for a range of microbubble sizes and for all three types of microbubbles considered in this work ($R_{\text{break-up}}=1.2R_0$). Results are summarized in Fig. 4. For all microbubble types, a few of the simulated pressure-time curves from larger microbubbles (4–5 μm in radius) with $R_{\max}/R_0$ just above the shell break-up threshold do not present postexcitation signals. When $R_{\max}/R_0>2$, however, postexcitation signals appear. Thus, the postexcitation signal appeared only when $R_{\max}>R_{\text{break-up}}$ and $R_{\max}>2R_0$. Some simulated radiated pressure signals for microbubbles <1 μm in radius did not yield detectable postexcitation signals, even though maximum expansion was much greater than $2R_0$.

B. Comparison between modeled and experimental spectral responses

Simulated pressure-time responses were Fourier transformed to model the spectral response, as demonstrated in Fig. 5, using the pressure-time responses presented in Fig. 3 for Definity. The average broadband noise in the −6 dB bandwidth of the PCD receiver (9.3–17.6 MHz) is plotted as a function of incident pressure in Fig. 5(d). The broadband

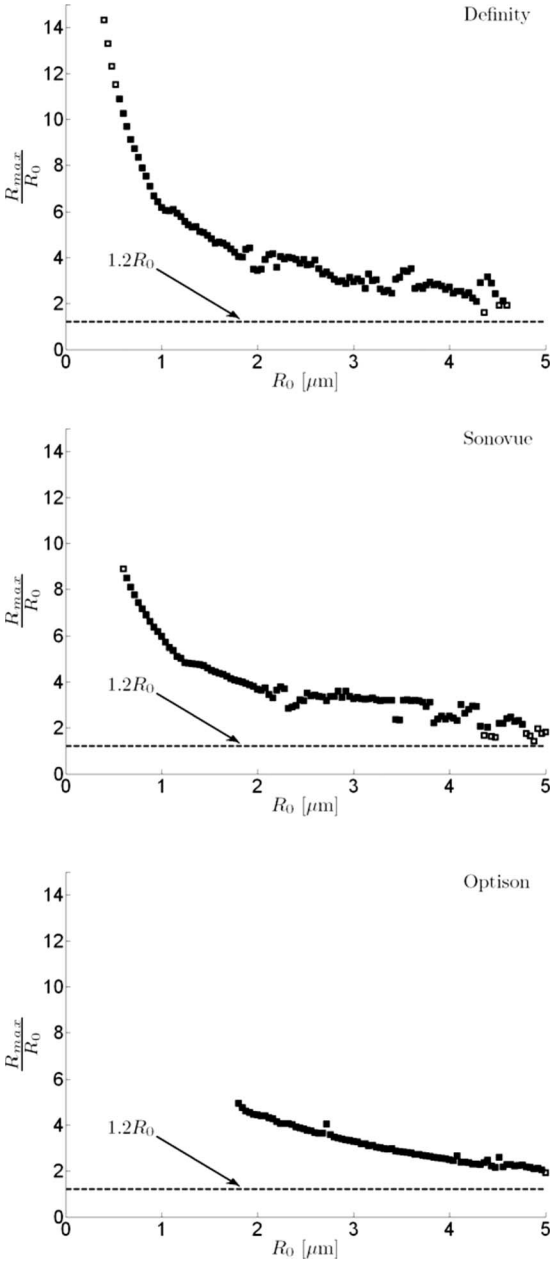


FIG. 4. Postexcitation responses predicted with the model relative to maximum radial expansion R_{max} and break-up radius ($R_{break\text{-}up}=1.2R_0$). Radius-time and pressure-time responses were simulated for each microbubble type as a function of the initial microbubble radius R_0 ($\Delta R_0=20$ nm) for a 2.8 MHz, 5-cycle incident pulse with PRPA of 2 MPa. The value of R_{max} on each radius-time curve divided by the initial radius (R_{max}/R_0) is plotted as a function of R_0 for each microbubble type. Cases for which the associated pressure-time response exhibited a postexcitation signal are plotted as filled symbols, and cases with no postexcitation signal are plotted as open symbols. For display, every other data point is plotted (steps of 40 nm). Simulated responses for large microbubbles presenting R_{max}/R_0 just above the shell break-up threshold, but below a value of two do not present postexcitation signals.

noise increases with incident pressure and presents the highest level when postexcitation signals are present, as anticipated.

Simulated and experimental power spectra are compared between groups of signals presenting postexcitation response and groups without postexcitation response for Sonovue, Definity, and Optison in Fig. 6. Comparisons between simulated and experimental spectra are presented within the −6 dB bandwidth of the PCD's receiving transducer. Modeled spectra for Sonovue and Optison without postexcitation response have sharp features. This is probably due to the small number microbubbles in the size-range selected for simulations that responded in this category. Average broadband power increases for the spectra with postexcitation signals, as compared to those from signals without postexcitation signals, are summarized in Table I for both simulated and experimental data. Relative differences in spectral amplitude in the simulated spectra correspond well to those observed experimentally (highest difference for Definity, and lowest for Optison). The increase in broadband noise between the average spectrum from microbubbles with postexcitation signals and that from same microbubbles without postexcitation signals is on the same order for predicted and experimentally measured spectra.

IV. DISCUSSION

This work further explores the link between postexcitation signals and microbubble break-up. Working at relatively high PRPA (2.0 MPa) for which the majority of microbubbles can be anticipated to break up, we were able to experimentally detect signals with and without postexcitation signals. This was shown experimentally, not only for albumin-shelled Optison microbubbles, but also for phospholipid-shelled Definity and Sonovue microbubbles.

To predict microbubble response under the experimental conditions considered in this work, we used the model by Marmottant *et al.*[18] that specifically describes variations of microbubble surface tension for lipid-shelled microbubbles as the shell expands and contracts. For application to albumin-shelled Optison, the model was modified to permanently suppress the shell viscosity and to set the surface tension term equal to that of water, once the break-up radius was reached. This suppression of shell elasticity and viscosity at a critical radius is similar to the approach previously applied to model shell break-up of a 2 μm Optison microbubble using the modified Herring equation.[16] Modeling of the pressure-time response for Optison, Sonovue, and Definity microbubbles across typical ranges of radii demonstrated that postexcitation signals occurred only when radial oscillation exceeded both the break-up threshold and $R_{max}>2R_0$. Modeled responses for microbubbles with oscillations not exceeding the break-up threshold or exceeding the break-up threshold without expanding larger than $2R_0$ did not yield postexcitation signals. $R_{break\text{-}up}$ was chosen to have a value consistent with those previously found by fitting the model to experimentally recorded radius-time responses measured for a lipid-shelled agent. Varying $R_{break\text{-}up}$ from 1.2 to 1.5 for Definity microbubbles consistently led to postexcitation signal appearance for $R_{max}>2R_0$. Flynn and co-worker[8,26] established a radial oscillation threshold of $R>2R_0$ for inertial cavitation of free bubbles. Thus, the model predicts that postexcitation signals occur for microbubbles in a range of sizes that respond, following break-up with radial oscillation exceeding the free bubble inertial cavitation threshold. However, a range of responses without postexci-

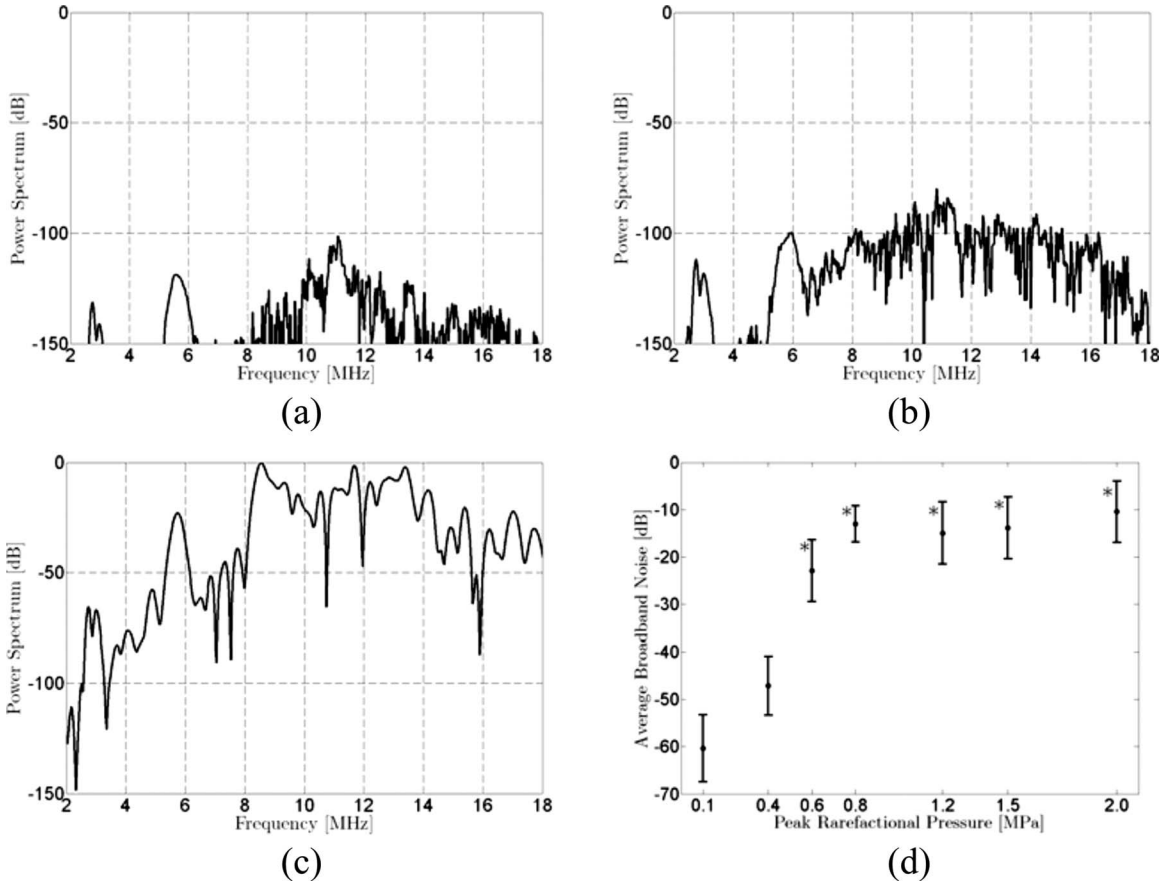


FIG. 5. Simulated spectral responses for a Definity microbubble at resonance radius (R_0=2.14 μm) with a break-up radius $R_{\text{break-up}}=1.2R_0$. Incident pulse at 2.8 MHz with 5 cycles. All spectra are presented with spectral amplitude relative to the peak value of the spectrum calculated for the response to the incident 2 MPa PRPA pulse. Incident PRPA: (a) 100 kPa, (b) 400 kPa, and (c) 2 MPa. (d) The average broadband noise (9.3–17.6 MHz) as a function of peak rarefactional pressure in the incident pulse. Error bars represent the standard deviation in the bandwidth. Stars indicate at which PRPA levels postexcitation emissions were identified on the simulated pressure-time curves.

tation signals from microbubbles exceeding the break-up radius but remaining less than the free bubble inertial cavitation threshold was predicted for each microbubble type considered. The simulated radiated pressure signals become very weak for the smallest microbubbles considered in simulations. Although some of these small microbubbles responded with maximum expansion greater than $2R_0$, no postexcitation signals were detected. Postexcitation signals might be hidden by surrounding noise, but it could also be due to limitations of the model.

Pressure-time response calculated using the model by Marmottant *et al.*[18] provides a good description of experimentally observed postexcitation signals when the shell breaks up and when radial expansion exceeds $2R_0$. This is the first time that this model feature has been demonstrated. A free bubble model should also produce postexcitation signals when radial oscillation is sufficient, but will lack terms considering the shell effects prior to initial break-up and (for lipid-shelled bubbles) during subsequent periods of recompression when $R<R_{\text{ruptured}}$. Postexcitation signals can also be obtained with the Marmottant model for non-lipid shelled agents, but only after modifying the shell parameters to eliminate shell viscosity and surface tension for all time after $R_{\text{break-up}}$ is initially reached.

Although the PRPA used in this study is within the range used in diagnostic imaging (mechanical index on the order of 1.2), such high pressures are not typically used in ultrasound contrast imaging, except when clearing microbubbles from the image plane for perfusion imaging. This high pressure was selected because it provided sufficient signal to noise ratio, not only for experimental detection of PCD responses with postexcitation response, but also for detection of weaker responses without postexcitation signals.

The response of the microbubbles with postexcitation signals relative to that of microbubbles without postexcitation signals demonstrated broadband noise increases of approximately 4–7 dB (experiment) and 4–9 dB (modeled spectra). Data were analyzed to investigate how much of this noise can be attributed to the response of the postexcitation signal. By excluding the postexcitation signals from the experimentally acquired voltage-time trace used for the Fourier transform, we estimated the spectral content, minus the contribution due to postexcitation signals, and compared this to the spectral content of the entire voltage-time traces. The postexcitation signals contributed averages of 0.2 ± 0.1, 0.8 ± 0.2, and 0.5 ± 0.2 dB of the broadband noise level observed for Definity, Sonovue, and Optison, respectively. This leads to the conclusion that the increase in broadband noise

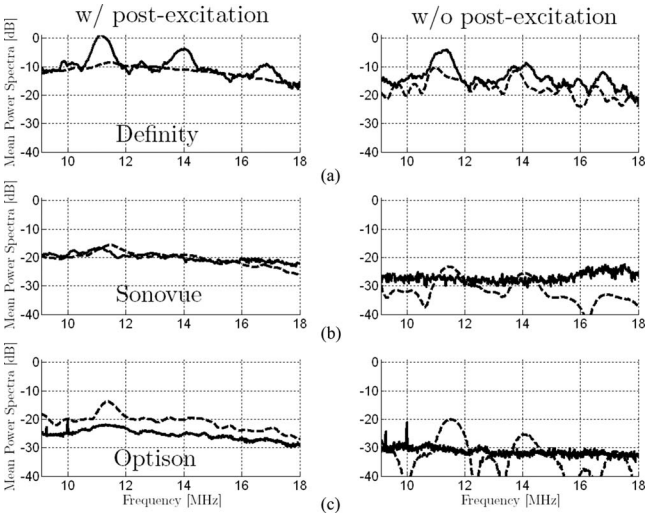


FIG. 6. Comparison of average spectra from signals with and without postexcitation response for simulated (dashed lines) and experimental (solid lines) results for (a) Definity, (b) Sonovue, and (c) Optison (incident peak rarefactional pressure amplitude of 2.0 MPa). All experimentally estimated spectra are presented on the same scale relative (re: 0 dB) to the peak value for Definity with postexcitation signals. The scaling factor that provided best alignment between simulated and experimental curves for Definity with postexcitation signals was calculated. This uniform off-set was applied to all simulated spectra. Highest broadband noise levels are observed for Definity and lowest levels for Optison. For each microbubble type, the average spectrum is higher for the case with postexcitation response than for the case without postexcitation response. Overall, the amplitudes of the simulated spectra vary consistently with the amplitudes of the experimentally assessed spectra (similar differences between microbubble types and between cases with and without postexcitation signals).

observed when postexcitation signals are detected is primarily due to increased pressure radiated during the principal response. These weak postexcitation signals on the pressure-time response are likely to be difficult to detect *in vivo* and can only be observed in the time domain when a single microbubble response is isolated. However, the spectral content associated with the inertial cavitation response may offer insight that is useful for *in vivo* studies. To estimate the onset of inertial cavitation in PCD measurements, increased broadband noise from insonified contrast microbubbles is generally compared to baseline levels observed in the absence of contrast agent. For example, Sassaroli and Hynynen[10] used an increase of at least 20 dB above background in the wideband noise to characterize Optison inertial cavitation. Tu *et al.*[11] found relative increases in normalized inertial cavitation doses from 10 to 20 dB between solutions with and without Optison. The increases in broadband noise associated with inertial cavitation, as reported in the current work, are lower (on the order of 4–7 dB) because comparison is not being made between the response from a group of microbubbles and noise baseline, but is instead being made between two groups of microbubbles under the same insonification conditions, but responding differently. The model predicts that the group with postexcitation signals oscillates above the inertial cavitation threshold and that the group without postexcitation signals predominantly oscillates below this threshold. For a well characterized microbubble population, it could be possible to apply the model to predict which microbubbles in the population will respond with inertial cavitation, and how modification of the insonification conditions can be used to adjust the inertial cavitation response.

TABLE I. Average difference between the mean power spectrum from signals with postexcitation response and the mean power spectrum from signals without postexcitation response are summarized for spectra from experimental data and for spectra calculated from simulated pressure-time responses. Spectral differences are calculated within a bandwidth from 9.3 to 17.6 MHz.

	Average broadband power increase (9.3–17.6 MHz) associated with postexcitation signals (dB)	
	Experimental data	Simulated data
Definity	4.2	4.0
Sonovue	7.2	6.7
Optison	6.1	8.9

An important limitation of this study is that the PCD configuration used does not provide information about the precise initial bubble size for a given voltage-time trace. The size distribution of microbubbles in dilutions may be modified compared to the pre-dilution distribution.[27] Furthermore, the size distribution in solution is time-dependent. Postexcitation signal prediction with the Marmottant model was explored across approximate size distributions to investigate the link with radial oscillation and shell break-up. The ranges considered were selected to be coherent with reported ranges and to include microbubble sizes near the shell break-up threshold at which simulated pressure-time curves did not present postexcitation signals.

However, in spite of the approximate nature of the microbubble size distributions used for modeling, the relative amplitude of the broadband noise predicted by the model between different types of microbubbles was consistent with experimental observations. Relative differences predicted between groups of microbubbles with and without postexcitation responses were also of the same order for the model and the experiment. Better comparisons between experimental and simulated spectra could be obtained if the microbubble size in the experimentally sampled population were known.

Results using the model by Marmottant *et al.*[18] have been presented to explore the relationship between postexcitation signal, microbubble size, break-up, inertial cavitation, and broadband noise levels. Modeled signals only presented postexcitation signals when both the break-up radius and the inertial cavitation threshold were exceeded. Comparable levels of added broadband noise in experimental and simulated data when postexcitation signals were present suggest that the Marmottant model can provide insight concerning spectral response for cavitating microbubbles.

ACKNOWLEDGMENTS

This work was supported by the cooperative project for biomedical engineering between the University of Illinois at Urbana-Champaign and the Centre National de la Recherche Scientifique, France. The authors acknowledge financial support from Agence Nationale de la Recherche under Grant No.

ANR-05-BLAN-0236 of project ACUVA and from the U.S. National Institutes of Health under Grant No. R37EB02641.

[1]A. Bouakaz and N. de Jong, "WFUMB safety symposium on echo-contrast agents: Nature and types of ultrasound contrast agents," Ultrasound Med. Biol. **33**, 187–196 (2007).
[2]H. Takeuchi, K. Ogunyankin, N. G. Pandian, T. P. McCreery, R. H. Sweitzer, V. E. Caldwell, E. C. Unger, E. Avelar, M. Sheahan, and R. Connolly, "Enhanced visualization of intravascular and left atrial appendage thrombus with the use of a thrombus-targeting ultrasonographic contrast agent (MRX-408A1): In vivo experimental echocardiographic studies," J. Am. Soc. Echocardiogr. **12**, 1015–1021 (1999).
[3]A. L. Klibanov, "Microbubble contrast agents—Targeted ultrasound imaging and ultrasound-assisted drug-delivery applications," Invest. Radiol. **41**, 354–362 (2006).
[4]A. Kheirolomoom, P. A. Dayton, A. F. H. Lum, E. Little, E. E. Paoli, H. R. Zheng, and K. W. Ferrara, "Acoustically-active microbubbles conjugated to liposomes: Characterization of a proposed drug delivery vehicle," J. Controlled Release **118**, 275–284 (2007).
[5]N. McDannold, N. Vykhodtseva, and K. Hynynen, "Targeted disruption of the blood-brain barrier with focused ultrasound: Association with cavitation activity," Phys. Med. Biol. **51**, 793–807 (2006).
[6]C.-Y. Lai C.-H. Wu, C.-C. Chen, and P.-C. Li, "Quantitative relations of acoustic inertial cavitation with sonoporation and cell viability," Ultrasound Med. Biol. **32**, 1931–1941 (2006).
[7]C. C. Coussios, C. H. Farny, G. Ter Haar, and R. A. Roy, "Role of acoustic cavitation in the delivery and monitoring of cancer treatment by high-intensity focused ultrasound (HIFU)," Int. J. Hyperthermia **23**, 105–120 (2007).
[8]H. G. Flynn, "Cavitation dynamics. II. Free pulsations and models for cavitation bubbles," J. Acoust. Soc. Am. **58**, 1160–1170 (1975).
[9]C. Rota, C. H. Raeman, S. Z. Child, and D. Dalecki, "Detection of acoustic cavitation in the heart with microbubble contrast agents in vivo: A mechanism for ultrasound-induced arrhythmias," J. Acoust. Soc. Am. **120**, 2958–2964 (2006).
[10]E. Sassaroli and K. Hynynen, "Cavitation threshold of microbubbles in gel tubes by focused ultrasound," Ultrasound Med. Biol. **33**, 1651–1660 (2007).
[11]J. Tu, J. H. Hwang, T. J. Matula, A. A. Brayman, and L. A. Crum, "Intravascular inertial cavitation activity detection and quantification in vivo with Optison," Ultrasound Med. Biol. **32**, 1601–1609 (2006).
[12]E. Biagi, L. Breschi, E. Vannacci, and L. Masotti, "Stable and transient subharmonic emissions from isolated contrast agent microbubbles," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **54**, 480–497 (2007).
[13]C. C. Church, "A theoretical-study of cavitation generated by an extracorporeal shock-wave lithotripter," J. Acoust. Soc. Am. **86**, 215–227 (1989).
[14]J. E. Chomas, P. Dayton, K. Morgan, J. Allen, and K. W. Ferrara, "Mechanisms of contrast agent destruction," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **48**, 232–248 (2001).
[15]W. S. Chen, T. J. Matula, and L. A. Crum, "Ultrasound contrast agent behaviour near the fragmentation threshold," Proceedings of the IEEE Ultrasonic Symposium (2000), p. 1935.
[16]A. Ammi, R. O. Cleveland, J. Mamou, G. Wang, S. L. Bridal, and W. D. O'Brien, Jr., "Ultrasonic contrast agent shell rupture detected by inertial cavitation and rebound signals," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **53**, 126–136 (2006).
[17]K. E. Morgan, J. S. Allen, P. A. Dayton, J. E. Chomas, A. L. Klibanov, and K. W. Ferrara, "Experimental and theoretical evaluation of microbubble behaviour: Effect of transmitted phase and bubble size," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 1494–1509 (2000).
[18]P. Marmottant, S. Van der Meer, M. Emmer, M. Versluis, N. de Jong, S. Hilgenfeldt, and D. Lohse, "A model for large amplitude of coated bubbles accounting for buckling and rupture," J. Acoust. Soc. Am. **118**, 3499–3505 (2005).
[19]A. P. Miller and N. C. Nanda, "Contrast echocardiography: New agents," Ultrasound Med. Biol. **30**, 425–434 (2004).
[20]T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1994).
[21]P. A. Dayton, K. E. Morgan, A. L. Klibanov, G. H. Brandenburger, and K. W. Ferrara, "Optical and acoustical observations of the effects of ultrasound on contrast agents," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**, 220–232 (1999).
[22]E. Kimmel, B. Krasovitski, A. Hoogi, D. Razansky, and D. Adam, "Subharmonic response of encapsulated microbubbles: Conditions for existence and amplification," Ultrasound Med. Biol. **33**, 1767–1776 (2007).
[23]J. M. Gorce, M. Arditi, and M. Schneider, "Influence of bubble size distribution on the echogenicity of ultrasound contrast agents: A study of Sonovue," Invest. Radiol. **35**, 661–671 (2000).
[24]D. Chatterjee and K. Sarkar, "A Newtonian rheological model for the interface of microbubble contrast agents," Ultrasound Med. Biol. **29**, 1749–1757 (2003).
[25]H. R. Zheng, A. Barker, and R. Shandas, "Predicting backscatter characteristics from micron- and submicron-scale ultrasound contrast agents using a size-integration technique," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **53**, 639–644 (2006).
[26]H. G. Flynn and C. C. Church, "Transient pulsations of small gas bubbles in water," J. Acoust. Soc. Am. **84**, 985–998 (1988).
[27]D. Goertz, N. de Jong, and A. F. W. Van der Steen, "Attenuation and size distribution measurements of Definity and manipulated Definity populations," Ultrasound Med. Biol. **33**, 1376–1388 (2007).

Erratum: Evaluating signal-to-noise ratios, loudness, and related measures as indicators of airborne sound insulation [J. Acoust. Soc. Am. 126, 1219–1230 (2009)]

H. K. Park
Chonnam National University, Gwangju 500-757, Korea

J. S. Bradley
National Research Council, Montreal Road, Ottawa, Ontario K1A 0R6, Canada

(Received 23 November 2009; accepted 23 November 2009)

[DOI: 10.1121/1.3277151]

PACS number(s): 43.60.Vx, 43.55.Hy, 43.10.Vx [NX]

The affiliation and address of the first author should have been as above. The Acknowledgment should have read as follows.

ACKNOWLEDGMENTS

H.K.P.'s contribution to this work was supported by a grant from the Korean Ministry of Education, Science and Technology (The Regional Core Research Program/Biohousing Research Institute) and a Korean Research Foundation Grant (MOEHRD) under Grant No. KRF-2006-352-D00200. The authors would also like to thank Dr. Brad Gover for helpful discussions during this project.

ACOUSTICAL NEWS

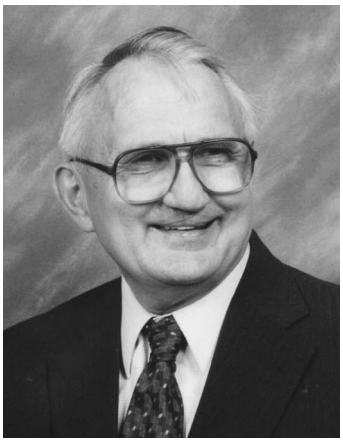
Elaine Moran
Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502

Editor's Note: *Readers of this journal are encouraged to submit news items on awards, appointments, and other activities about themselves or their colleagues. Deadline dates for news and notices are 2 months prior to publication.*

New Fellows of the Acoustical Society of America

Egbert de Boer—For contributions to cochlear mechanics and psychoacoustics

Richard Freyman—For contributions to spatial hearing and speech perception



Paul M. Gammell—For contributions to ultrasonic signal generation and processing

David Knobles—For contributions to inversion methods in underwater acoustics

Allard Jongman—For contributions to the nature of phonetic representations

Enrique Lopez-Poveda—For contributions to cochlear non-linearity

Michael I. Proctor Selected Recipient of the 2009 Stetson Award

ASA member Michael I. Proctor was selected the recipient of the 2009 Raymond H. Stetson Scholarship in Phonetics and Speech Production. Mr. Proctor was a graduate student in the Linguistics Department at Yale University, New Haven, Connecticut.

Michael Proctor received B.A. and B.E. degrees from the University of New South Wales, an M.A. in Linguistics from the University of Queensland, and M.A. and M.Phil. from Yale University. The subject of his Ph.D. dissertation is "The Phonetic Basis of Phonological Classhood: The Case of Liquids."

The Stetson Scholarship, which was established in 1998, honors the memory of Professor Raymond H. Stetson, a pioneer investigator in phonetics and speech science. Its purpose is to facilitate the research efforts of

 0001-4966/2010/127(2)/1166/2/$25.00

Michael I. Proctor

promising graduate students and postgraduates. The Scholarship includes a $4000 stipend for one academic year.

Applications for the award are due in March each year. For further information about the award, please contact the Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502, Tel: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; Web: http://asa.aip.org. Application information can also be found on the ASA Home Page at http://asa.aip.org/fellowships.html.

Special Notice Regarding the Distinguished Service Citation and Honorary Fellowship in the Acoustical Society of America

The Distinguished Service Citation is awarded to any present or former Member or Fellow of the Society in recognition of outstanding service to the Society.

An Honorary Fellowship is awarded from time to time to a rare individual for eminence in, or outstanding service to, acoustics; candidates in general should not be members of the Society.

Nominations may be made in writing by any Member or Fellow. They shall be submitted to the Committee on Medals and Awards for review and forwarded to the Executive Council for approval.

Calendar of Meetings and Congresses

Compiled by the Information Service of the International Commission for Acoustics

2010

8–10 March	Chofu, Japan. Spring Meeting of the Acoustical Society of Japan. http://www.asj.gr.jp/eng/futuremeetings.html
15–18 March	Berlin, Germany. Meeting of the German Association for Acoustics DAGA 2010. http://www.daga-tagung.de/2010
15–19 March	Dallas, TX, USA. International Conference on Acoustics, Speech, and Signal Processing. http://icassp2010.org
7–9 April	Cambridge, UK. David Weston Sonar Performance Assessment Symposium. http://www.ioa.org.uk/events
19–23 April	Baltimore, MD, USA. Joint meeting: 159th Meeting of the Acoustical Society of America and Noise-Con 2010. http://asa.aip.org/meetings.html
27–30 April	Ghent, Belgium. Institute of Acoustics/Belgian Acoustical Association Joint Meeting. http://www.ioa.org.uk/viewupcoming.asp
6–7 May	Paris, France. 2nd International Symposium on Ambisonics and Spherical Acoustics. http://ambisonics10.ircam.fr
10–12 May	Bergen, Norway. Baltic-Nordic Acoustics Meeting2010. E-mail: sanordby@norsonic.com
9–11 June	Aalborg, Denmark. 14th Conference on Low Frequency Noise and Vibration. http://lowfrequency2010.org
13–16 June	Lisbon, Portugal. INTERNOISE2010. http://www.internoise2010.org
5–9 July	Istanbul, Turkey. 10th European Conference on Underwater Acoustics. http://ecua-2010-istanbul.org
18–22 July	Cairo, Egypt. 17th International Congress on Sound and Vibration (ICSV17). http://www.icsv17.org
23–27 August	Sydney, Australia. International Congress on Acoustics 2010. http://www.ica2010sydney.org
23–27 August	Seattle, USA. 11tth International Conference on Music Perception and Cognition. http://depts.washington.edu/icmpc11
29–31 August	Melbourne, Australia. International Symposium on Room Acoustics (ISRA2010). http://www.isra2010.org
14–18 September	Kyoto, Japan. 5th Animal Sonar Symposium. http://cse.fra.affrc.go.jp/akamatsu/AnimalSonar.html
15–18 September	Ljubljania, Slovenia. Alps-Adria-Acoustics Association Meeting joint with EAA. E-mail: mirko.cudina@fs.uni-lj.si
20–22 September	Leuven, Belgium, International Conference on Noise and Vibration Engineering (ISMA2010), http://www.isma-isaac.be
26–30 September	Makuhari, Japan. Interspeech 2010–ICSLP. http://www.interspeech2010.org
11–14 October	San Diego, CA USA. IEEE 2010 Ultrasonics Symposium. E-mail: b.potter@vectron.com
14–16 October	Niagara-on-the Lake, Ont., Canada. Acoustics Week in Canada. http://caa-aca.ca/E/index.html
13–15 October	Leon, Spain. 41st Spanish Congress of Acoustics and 6th Iberian Acoustics Congress. http:// www.sea-acustica.es
18–22 October	Nagahama, Japan. 10th International Workshop on Railway Noise (IWRN10). http://www.rtri.or.jp/IWRN10/first.announcement.html
15–19 November	Cancún, Mexico. 2nd Pan-American Meeting on Acoustics. http://asa.aip.org/meetings.html

2011

22–27 May	Prague, Czech Republic. International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2011). http://www.icassp2011.com
23–27 May	Seattle, WA, USA. 160th meeting of the Acoustical Society of America. http://asa.aip.org/meetings.html
27 June–1 July	Aalborg, Denmark. Forum Acusticum 2011. http://www.fa2011.org
24–28 July	Tokyo, Japan. 19th International Symposium on Nonlinear Acoustics (ISNA 19). http://www.isna19.com
27–31 August	Florence, Italy. Interspeech 2011. http://www.interspeech2011.org
4–7 September	Osaka, Japan. Internoise 2011. http://www.internoise2011.com
5–8 September	Gdansk, Poland. International Congress on Ultrasonics. http://icu2011.ug.edu.pl/index.html

2012

20–25 March	Kyoto, Japan. IEEE International Conference on Acoustics, Speech, and Signal Processing. http:// www.icssp2012.com

2013

26–31 March	Vancouver, Canada. 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). http://www.icassp2013.com
2–7 June	Montréal, Canada. 21st International Congress on Acoustics (ICA 2013). http://www.ica2013montreal.org

BOOK REVIEW

P. L. Marston
Physics Department, Washington State University, Pullman, Washington 99164

These reviews of books and other forms of information express the opinions of the individual reviewers and are not necessarily endorsed by the Editorial Board of this Journal.

Building Acoustics and Vibration—Theory and Practice

Osama A. B. Hassan

World Scientific Publishing Co. Pte. Ltd., Singapore, 2009. 972 pp. Price: $109.00 (hardcover), ISBN: 13 978-283-833-9 ISBN: 10 981-283-833-3.

This book is an impressive collection of information and prediction techniques needed by consultants working in the building acoustics area. The large number of worked examples and practical guidance tips is very useful. The prediction techniques include both theoretical and empirical methods. Five of the eleven chapters cover the important area of sound insulation: single number ratings and test methods, airborne sound insulation, impact sound isolation and insulation of floors, flanking sound transmission, and sound insulation of doors and windows.

The chapter on sound insulation provides some theory and a great deal of practical advice. The airborne sound insulation of single leaf walls is divided into three frequency regions: stiffness controlled, mass controlled, and damping controlled. A resonance region overlaps the top of the stiffness controlled and the bottom of the mass controlled regions, while a coincidence region overlaps the top of the mass controlled and the bottom of the damping controlled regions. The reviewer was surprised to see a theoretical treatment of normal incidence sound insulation in the stiffness controlled region before the more important mass and damping controlled regions were covered. The important field incidence mass law is given, followed by approximations to Cremer's 1942 equation for sound insulation above the critical frequency. The terms multiplying the logarithm of the frequency are 33.22, 33, and 30 in three different approximations to Cremer's equation. Presumably, the purpose of 33 rather than 30 is to counteract the approximation of f/f_c-1 by f/f_c. Sewell's 1970 low frequency correction terms for finite plate size and the approach to the critical frequency from below would be better placed in the mass controlled region section rather than the damping controlled section. The treatment of single leaf walls concludes with two approximate methods due to Watters 1959 and Ingemansson and Elvhammer 1977.

The normal incidence mass-air-mass resonant frequency for double leaf walls is derived, and Sharp's 1973 theory for double leaf walls without mechanical connections between the leaves is presented. Because Sharp assumed that the wall cavity contains sound absorbing materials, a formula from Vér and Holmer 1971 for the case of a double leaf wall without cavity absorption, and a formula from Brekke 1980 for the case of a double glazed window whose non-glass internal cavity surfaces are lined with sound absorbing materials are included. Sharp's theory for the structural transmission by rigid point, line, and edge connections is presented. This is followed by an approach for non-rigid studs due to Kristensen and Rindel 1989. This chapter on airborne sound insulation continues on with practical advice and consideration of composite, laminated, orthotropic, sandwich and finite size plates, and holes and slits.

The chapter on floors derives formulas for the normalized impact sound pressure level of a heavyweight homogeneous floor. Then, formulas for the improvement in impact sound isolation due to floor coverings are derived. The author points out the important fact that the improvement in impact isolation due to floor coverings is less on lightweight floors than on heavyweight floors. Impact sound isolation in practice is covered before a consideration of floating floors. A very short section points out that particle type materials such as sand can be used to improve impact sound insulation due to the fact that they add both damping and mass to the floor. The airborne sound insulation of floor/ceiling combinations is covered, including the sound transmission between adjacent rooms via the plenum space above a suspended ceiling. This chapter concludes with the well known relationship between impact sound isolation and airborne sound insulation for heavyweight homogeneous floors.

The first two sections of the chapter on flanking sound transmission cover airborne and impact sound insulation following ISO 15712:2005 (EN12354:2000) parts 1 and 2. The third section considers the very important prediction of the vibration attenuation across wall junctions. Then the laboratory and *in-situ* measurement of flanking transmission and its prediction by statistical energy analysis (SEA) is considered. It is pointed out that SEA can be used to extend the methods of ISO 15712:2005 (EN12354:2000). The chapter then discusses practical issues of flanking transmission calculations, gives an approximate model of flanking transmission and returns to the issue of the measurement of flanking transmission before giving common practical measures to mitigate the effect of sound flanking.

There is a separate chapter on the sound insulation of doors and windows because, as the author points out, "doors and windows are the acoustically weak elements of a partition because of their low surface density and the gaps around their edges." This chapter is mainly practical advice and includes information on the variation of sound insulation with angle of incidence and with sealing. Laminated, double, and triple glazings are also discussed together with facades and balconies. The chapter's appendix contains octave band sound reduction indices and the weighted sound reduction index for a number of windows and doors.

The chapter on single number ratings and test methods compares ASTM and ISO standards. It also discusses British and Swedish sound insulation requirements. The chapters on sound insulation are followed by a chapter on outdoor sound propagation. Clearly, it is important to be able to predict the exterior noise exposure level of a building. The chapter covers the effects of atmospheric conditions and barriers on the propagation of the sound. A chapter on room acoustics is included so that the internal building sound fields can be calculated. Its appendix contains octave band sound absorption data for a large number of surfaces and objects. The chapter on air distribution system noise covers fan noise, attenuation in straight lined and unlined ducts, attenuation at bends, end reflections, duct break-out and break-in noise, grille noise, and silencers. The last chapter in the book is on vibrations in buildings. It mainly covers vibration isolation and the important topic of the low frequency vibration of floors.

The first two chapters of the book are entitled *An Overview of Building Acoustics* and *Vibration and Fundamental Building Acoustics*. They provide the extensive background that is needed to understand the rest of the book. Appendix 1 provides a summary of building acoustic measurements, while Appendix 2 gives R_w, $L_{n,w}$, *STC*, and/or *IIC* for a large range of walls and floors. Appendix 3 provides octave band sound reduction indices and R_w for some common wall and floor constructions. The book concludes with a bibliography and index.

One of the strengths of this book is that, because the author is based in Sweden, he has been able to provide a good coverage of both North American and European standards and developments. Thus, both ASTM and ISO standards are well covered. This book is a treasure trove of practical advice and calculation methods. It should be of great assistance to acoustical consultants. However, for research purposes, I would prefer Hopkins 2007 *Sound Insulation*, Fahy and Gardonio 2007 *Sound and Structural Vibration: Radiation, Transmission and Response*, and Cremer, Heckl, and Petersson 2005 *Structure-Borne Sound: Structural Vibrations and Sound Radiation at Audio Frequencies*. However, having said this, the wide range of references,

 0001-4966/2010/127(2)/1168/2/$25.00

from which this book draws material, also makes it very useful to the building acoustics researcher. It would also be useful as a textbook. However, its length means that a judicious selection of material from it would have to be made by the course coordinator.

JOHN LAURENCE DAVY
CSIRO Materials Science and Engineering,
P.O. Box 56, Highett, Victoria 3190, Australia

REVIEWS OF ACOUSTICAL PATENTS

Sean A. Fulop
Dept. of Linguistics, PB92
California State University Fresno
5245 N. Backer Ave., Fresno, California 93740

Lloyd Rice
11222 Flatiron Drive, Lafayette, Colorado 80026

The purpose of these acoustical patent reviews is to provide enough information for a Journal reader to decide whether to seek more information from the patent itself. Any opinions expressed here are those of reviewers as individuals and are not legal opinions. Printed copies of United States Patents may be ordered at $3.00 each from the Commissioner of Patents and Trademarks, Washington, DC 20231. Patents are available via the internet at http://www.uspto.gov.

Reviewers for this issue:

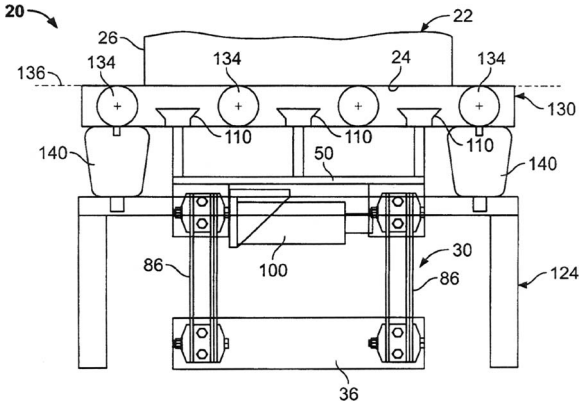
GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
SEAN A. FULOP, *California State University, Fresno, 5245 N. Backer Avenue M/S PB92, Fresno, California 93740-8001*
MARK KAHRS, *Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261*
DAVID PREVES, *Starkey Laboratories, 6600 Washington Ave. S., Eden Prairie, Minnesota 55344*
CARL J. ROSENBERG, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
NEIL A. SHAW, *Menlo Scientific Acoustics, Inc., Post Office Box 1610, Topanga, California 90290*
ROBERT C. WAAG, *Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14627*

7,556,421

43.20.Tb VIBRATORY APPARATUS AND METHOD FOR SETTLING THE CONTENTS OF A CONTAINER

Michael J. Lindbeck and John J. Valuet, assignors to Martin Engineering Company
7 July 2009 (Class 366/114); filed 9 October 2007

"Contents may settle during shipment." Device 20 vibrates containers 22 horizontally, as opposed to vertically in prior art, on conveying plate 136. Containers are restrained by suction cups 110. A more efficient means of settling the contents is said to result, despite the inevitability of the opening statement.—NAS

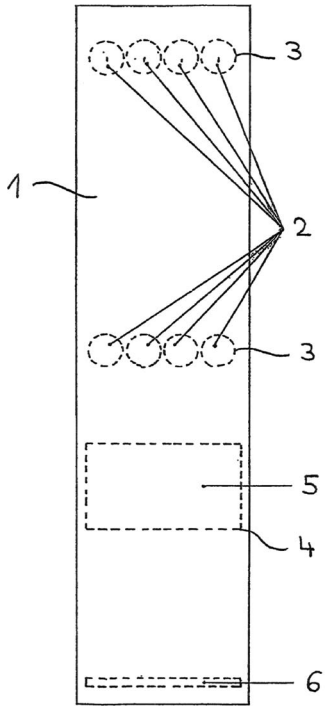


7,548,630

43.38.Ja INTEGRATED LOUDSPEAKER SYSTEM

Berthold Schlenker *et al.*, assignors to Fraunhofer-Gesellschaft zur Forderung der angewandten Forshcung e.V.
16 June 2009 (Class 381/385); filed in Germany 11 March 2003

Wall panel 1 is sized so that it can be mounted directly to structural supports (studs) in residential and commercial partitions; in fact, the panel can be used for all of a wall. Openings 3 are fitted with transducers 2, which can be conventional electrodynamic cone types (in which a plurality of circular openings are used) or distributed mode loudspeaker(s) 10 (with drivers 11), as required by the installation. Amplifier assembly 5 fits in rectangular recess 4, and connections 12–5 are via connector panel 6 (which can be accessed directly or via a surface raceway system). The patent claims that the wall panels can therefore be placed as needed for multi-channel playback or electronic acoustic enhancement systems without the limitations other transducer types may have, such as conventional box enclosures and picture frame speakers.—NAS



7,551,749

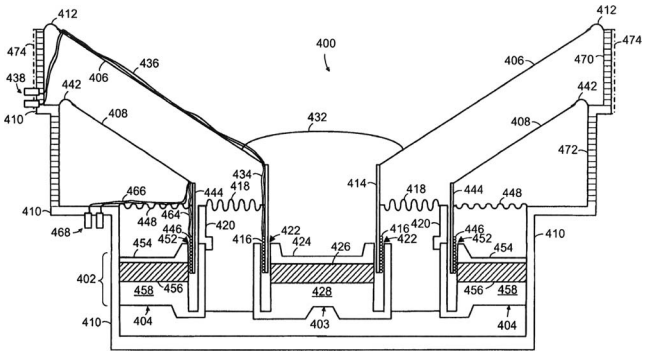
43.38.Ja BAFFLE VIBRATION REDUCING

Michael D. Rosen and Richard Warren Little, assignors to Bose Corporation
23 June 2009 (Class 381/401); filed 30 November 2004

Two moving mechanisms (those for cones 406 and 408) are mounted

 0001-4966/2010/127(2)/1170/11/$25.00

in a single frame 410. The drive to voice coils 416 and 446 are inverted with respect to one another. The vibration of frame 410 is said to thereby be reduced and therefore the excitation of a flimsy baffle, such as may be found in an automobile, to which assembly 400 is attached, is reduced, and therefore so is the baffle vibration. Various implementations, including, but of course not limited to, conventional ported and multi-chamber enclosures are also described, which may or may not use separate drivers and/or processing of the signals for the coils.—NAS

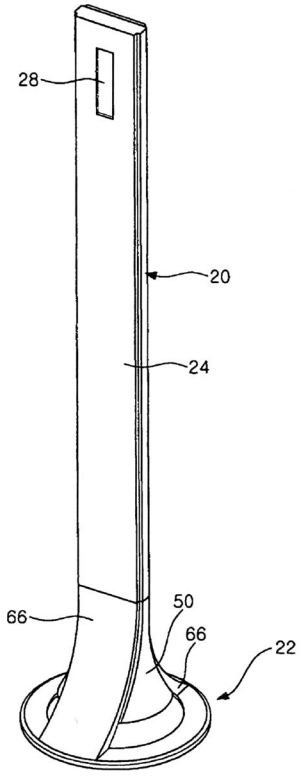


7,575,095

43.38.Ja SPEAKER

Chang Wuk An and Jung Woo Choi, assignors to LG Electronics Incorporated
18 August 2009 (Class 181/199); filed in Republic of Korea 28 December 2005

This patent describes a tower type loudspeaker that might be used in a home theater system. The embodiment shown is noteworthy for its disregard of elementary trade-offs involved in low frequency sound reproduction. A down-firing woofer is located in the base of the tower. One might expect the



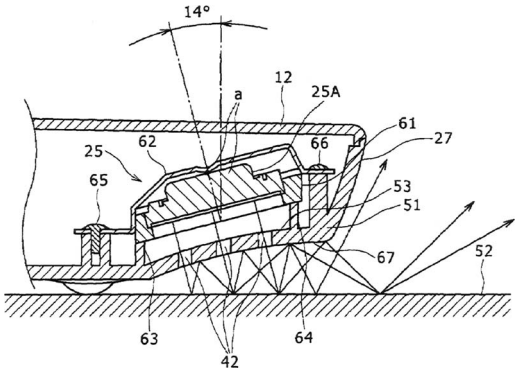
thin rectangular plinth to provide transmission line loading, but no, there is only a small, sealed back chamber barely large enough to house the woofer. The remainder of the cabinet is simply a support for the tweeter. Then there is the issue of prior art. The patent claims set forth a two-way speaker system, "…wherein the first speaker unit is installed to face a front through a front end of the housing, and the second speaker unit is installed to face a floor on which the support assembly is placed." Moreover, there is a "…guide cone…facing the second speaker unit to change a direction of sound outputted from the second speaker unit." All of which accurately describes a drum shaped loudspeaker system marketed by Audio Empire in the 1960s.—GLA

7,576,983

43.38.Ja PORTABLE ELECTRONIC APPARATUS WITH IMPROVED AUDIO QUALITY THROUGH A CURVED AND SLOPED SURFACE OF A HOUSING

Koichi Yamaguchi and Toshiharu Tajiri, assignors to Sony Corporation
18 August 2009 (Class 361/683); filed in Japan 12 September 2006

This is a trading stamp patent. That is, it provides little actual protection and discloses nothing very unusual or interesting. It merely adds to



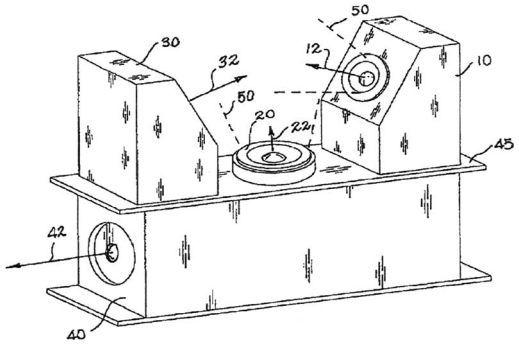
Sony's arsenal of patents that can be bargained in cross-licensing agreements. This is not the first instance of down-firing loudspeakers in laptop computers. The only novel feature seems to be the concave curvature of panel 67.—GLA

7,577,265

43.38.Ja LOUDSPEAKER SYSTEM PROVIDING IMPROVED SOUND PRESENCE AND FREQUENCY RESPONSE IN MID AND HIGH FREQUENCY RANGES

Ira Pazandeh and Yorba Linda, California
18 August 2009 (Class 381/160); filed 15 October 2004

In spite of its appearance, the assembly shown is not a compact stereo system but rather a single-channel reproducer. Mid-range speakers 50 are



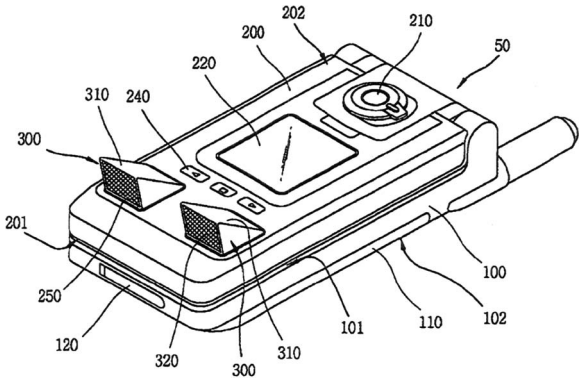
housed in separate enclosures on either side of upward-firing tweeter 20. "Sound from the first and third radiators impinges on the second radiator so as to cause an echo effect, improving sound presence." The theory may be hazy but under the right conditions this arrangement might well produce a pleasing spatial effect.—GLA

7,583,938

43.38.Ja MOBILE COMMUNICATION DEVICE HAVING SPEAKER UNIT

Hee-Jun Kim, assignor to LG Electronics Incorporated
1 September 2009 (Class 455/90.3); filed in Republic of Korea 4 November 2004

This flip-top cellular phone is intended to provide "high quality sound" whether the unit is open or closed. When open, as shown, two teeny-tiny speaker compartments 310 rise out of the case to expose even tinier speakers. When closed, the speakers retract like concealed headlights and sound is conducted through short waveguides to openings on the opposite side of the case. A rotating cylindrical speaker compartment would make better use of available space—has this already been patented?—GLA

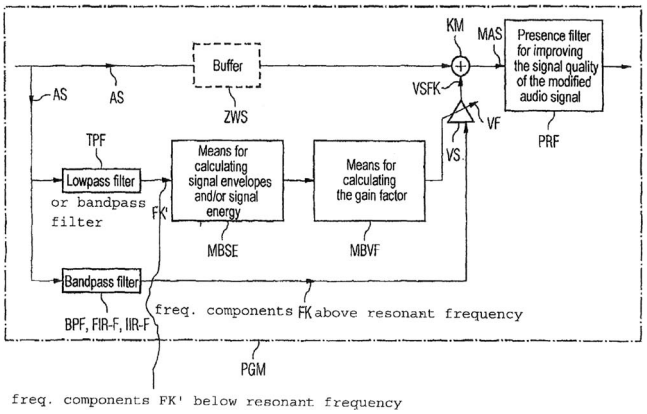


7,574,009

43.38.Lc METHOD AND APPARATUS FOR CONTROLLING THE REPRODUCTION IN AUDIO SIGNALS IN ELECTROACOUSTIC CONVERTERS

Roland Aubauer, 81669 München and Stefano Ambrosius Klinke, 50169 Kerpen, both of Germany
11 August 2009 (Class 381/98); filed 21 September 2001

The human hearing mechanism can be tricked into perceiving low frequency tones that do not exist. The current popularity of small, portable audio devices has prompted a renewed interest in circuits that can create



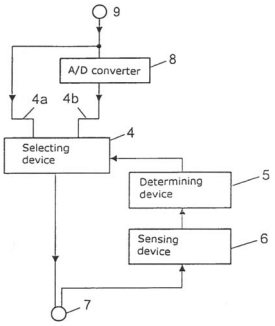
such "synthetic bass." (Historically, organ builders have practiced the art for more than 100 years.) Most electronic devices generate additional harmonics that lie within the passband of a bass-deficient loudspeaker, but this patent takes a slightly different approach. Signal content within a limited range is boosted, and the amount of boost is controlled by the level of energy below the loudspeaker's resonance frequency. For example, a 40 Hz fundamental tone would be blocked but existing harmonics in the range from 80 to 150 Hz would be boosted proportionately.—GLA

7,574,011

43.38.Lc DETECTION DEVICE

Rolf Meyer and Klaus-Michael Polten, assignors to Sennheiser electronic GmbH & Company KG
11 August 2009 (Class 381/113); filed in Germany 7 May 2003

The patent describes a "smart" XLR connector that can determine whether a received signal is analog or digital. "The selecting device then selects either an analog or a digital connection, in dependence on the determination signal." Another embodiment covers the opposite situation in which a smart microphone senses whether it is connected to a digital or analog input and then supplies the appropriate signal. The presence of phantom power, the phantom power voltage, source impedance, and signal characteristics may all be used to make the proper determination.—GLA

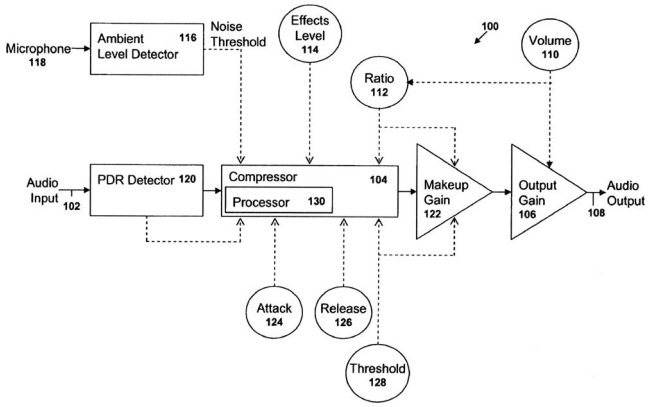


7,580,531

43.38.Lc DYNAMIC RANGE REDUCING VOLUME CONTROL

Charles Anderson, assignor to Cirrus Logic, Incorporated
25 August 2009 (Class 381/106); filed 6 February 2004

This patent informs us that audio program material generally includes both loud passages and quiet passages. 50 years ago that may have been true, but today's recordings have a dynamic range of about 6 dB, which is further squashed during radio transmission. Nevertheless, there may be level differences between different program sources, between speech and music, or indeed between loud and soft passages in high quality recordings. In a noisy environment, reproduction at loud listening levels (say, 80–90 dB) may be acceptable whereas softer passages may become inaudible when the volume is turned down. To counteract this effect, compression ratio 112 is linked to volume setting 110. It can also be linked to a background noise threshold detector 116. This basic scheme is expounded at much greater length in the 43 patent claims.—GLA



7,548,851

43.38.Md DIGITAL MULTIMEDIA JUKEBOX

Jack Lau *et al.*, Hong Kong, China
16 June 2009 (Class 704/201); filed 11 October 2000

In all of this verbiage, the only questionably original claim is a two step bit allocation scheme for a MP3 coder.—MK

7,548,875

43.38.Md MEDIA DELIVERY PLATFORM

John Mikkelsen, Minneapolis, Minnesota and Robert Freidson, St. Petersburg, Russian Federation
16 June 2009 (Class 705/26); filed 26 June 2002

Whenever schematics are presented, one has a better impression of the reality of the invention. Unfortunately, using a portable telephone as a music player is not original.—MK

7,552,389

43.38.Md COMPUTER PROGRAM AND METHODS FOR AUTOMATICALLY INITIALIZING AN AUDIO CONTROLLER

Thomas M. Drewes *et al.*, assignors to Polycom, Incorporated
23 June 2009 (Class 715/727); filed 20 August 2003

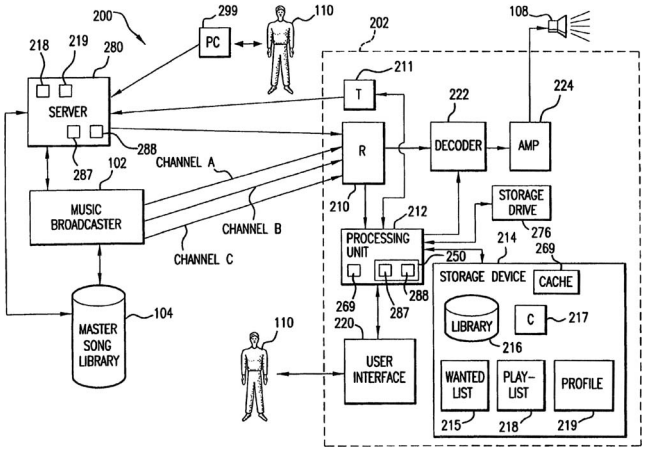
It is a sad commentary on the United States Patent Office when patents are issued for a description of a software driver for a speakerphone.—MK

7,555,539

43.38.Md PERSONALIZED AUDIO SYSTEM AND METHOD

Jeremy C. Rosenberg and Ronald M. Yurman, assignors to Music Choice
30 June 2009 (Class 709/219); filed 10 March 2006

Suppose that a music server has a large collection of musical tracks of different genres. Each portable listener wants a different selection. How can this be achieved with a broadcast mechanism? One strategy is to broadcast metadata before the tracks. If the metadata match the filter, then play the track. In addition, with a large storage medium, tracks can be stored for delayed playing (time shifting).—MK



7,558,637

43.38.Md DIGITAL MIXER CAPABLE OF PROGRAMMING MIXER CONFIGURATION, MIXER CONFIGURATION EDITING APPARATUS, AND CONTROL APPLICATION PROGRAM TO CONTROL DIGITAL MIXER

Satoshi Takemura and Yoshinori Kawase, assignors to Yamaha Corporation
7 July 2009 (Class 700/94); filed in Japan 13 July 2004

This patent is very specific to the Yamaha DME digital mixer interconnect. The question is how to compile a graphical representation of a mix into the microcode needed to run the mixing digital signal processor. The writing discusses the software on the host in graphic detail but leaves the mechanism of compiling each step into the microcode unsaid.—MK

7,563,975

43.38.Md MUSIC PRODUCTION SYSTEM

Daniel Leahy *et al.*, assignors to Mattel, Incorporated
21 July 2009 (Class 84/619); filed 13 September 2006

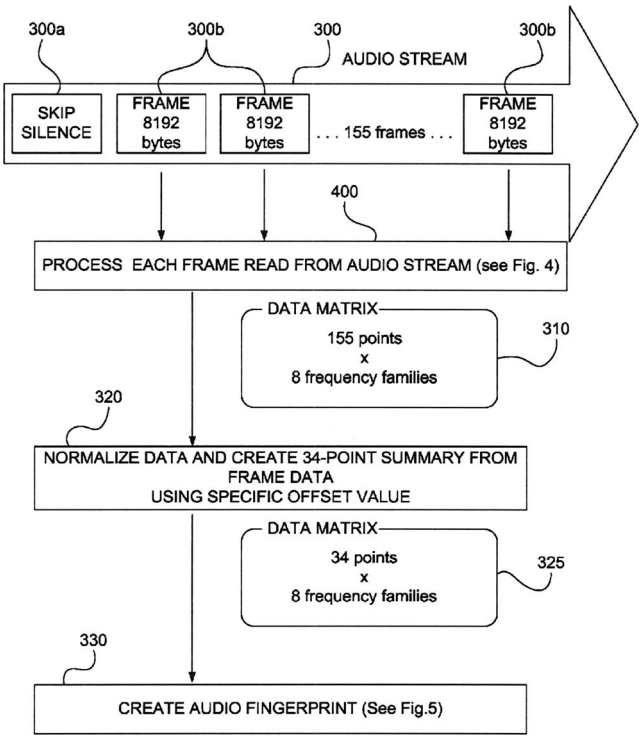
An animated rock band can be controlled by a microphone wielding user by (a) correcting pitch (b) finding notes, and (c) composing an arrangement on the fly. Unfortunately, the steps outlined in the patent are not original or novel.—MK

7,567,899

43.38.Md METHODS AND APPARATUS FOR AUDIO RECOGNITION

Vladimir Askold Bogdanov, assignor to All Media Guide, LLC
28 July 2009 (Class 704/231); filed 30 December 2004

An audio fingerprint is a concise representation of an audio recording. This disclosure uses a time method based on curve fitting to a fixed (and



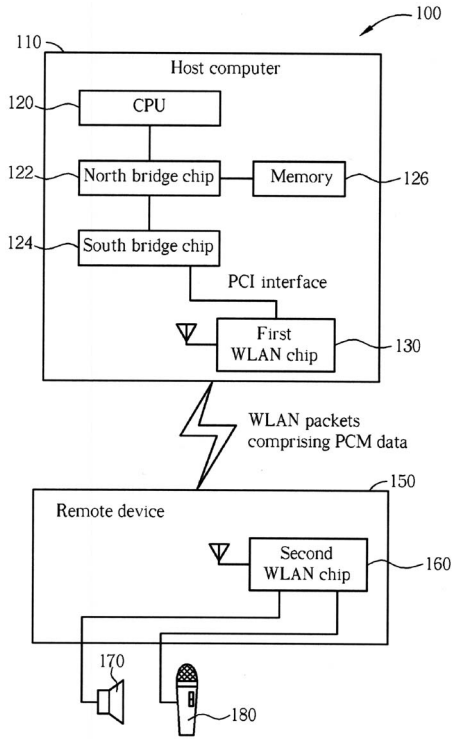
small) set of templates 16. No statistics are given, so how well this works in practice is left to the imagination.—MK

7,577,261

43.38.Si WIRELESS AUDIO SYSTEM USING WIRELESS LOCAL AREA NETWORK

Yen-Ku Liu and Jia-Bin Huang, assignors to Realtek Semiconductor Corporation
18 August 2009 (Class 381/77); filed in Taiwan 13 October 2005

Wireless networks may deliver audio program material to a variety of devices, not just computers. However, existing protocols generally rely on computer sound cards for analog-to-digital and digital-to-analog processing.



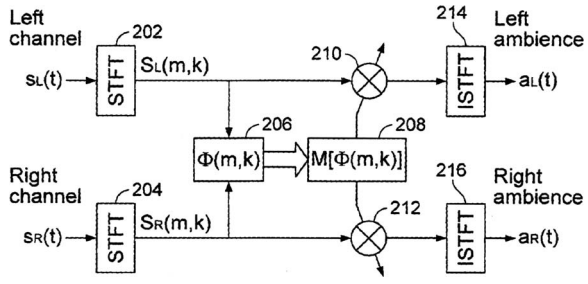
This short patent discloses an alternative method requiring "a first computing unit for performing audio processing," "a second WLAN unit for receiving the audio data," and "an audio reproducing unit coupled to the second WLAN unit…wherein the first device registers the first WLAN unit as an audio processing unit…and second WLAN unit as an audio in/out interface of the first computing unit."—GLA

7,567,845

43.38.Vk AMBIENCE GENERATION FOR STEREO SIGNALS

Carlos M. Avendano and Jean-Marc M. Jot, assignors to Creative Technology Limited
28 July 2009 (Class 700/94); filed 4 June 2002

Given N input channels, they can be "upmixed" to M output channels ($M \geq N$). In addition, each output channel has additional "ambience" adde



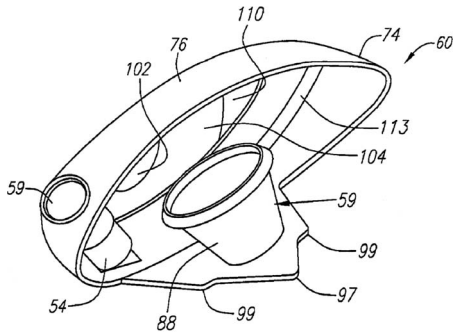
d. As shown in the figure, for stereo ($M=N=2$), a spatial coherence function can be computed in the frequency domain after using the short time Fourier transform (STFT). A nonlinear function M can be applied to create ambience and to shade the coefficients before applying the inverse STFT.—MK

7,556,567

43.40.Tm MULTIPLE MATERIAL GOLF CLUB HEAD

J. Andrew Galloway, assignor to Callaway Golf Company
7 July 2009 (Class 473/224); filed 21 January 2008

Two means of enhancing the sound made when the golf club face hits a golf ball are described. In one, a conical piece 59, which can also be washer-like in shape, is attached to the exterior portion 99 of club head 60. Alternately, an "arcuate piece" of a composite material (something made of more than one substance) can be fitted to the exterior of face component 60.—NAS



7,549,505

43.50.Jh ACOUSTIC NOISE REDUCTION DEVICE FOR ELECTRONIC EQUIPMENT, INCLUDING PERSONAL COMPUTERS

Maher S. Kawar, Mountain View, California
23 June 2009 (Class 181/210); filed 7 December 2007

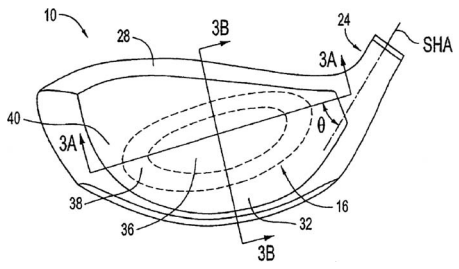
There has been insufficient attention paid to the fan noise of computers. This inventor proposes the use of various shaped barriers that must accommodate fans and heat exhaust in many different computer models.—MK

7,553,242

43.50.Pn COMPOSITE METAL WOOD CLUB

Jeffrey W. Meyer *et al.*, assignors to Acushnet Company
30 June 2009 (Class 473/314); filed 1 October 2008

Hitting face 16 has three areas, 36, 38, and 40, each with a different thickness, stiffness, and relative area. The patent asserts that the hit area on the face of a driver club can be described as elliptical, and thus areas 36 and 38 have this shape. Angle θ is also a parameter that can be varied and th



e coefficient of restitution of the golf club head is aligned with the ball impact pattern. Thus the sweet zone and spin range, as well as other performance aspects of the club, can be tailored (not to be confused with the homophone brand name) to suit the golfer's predilection and preference. In addition, which seems to be a requirement in patents such as this, a means of

varying the sound made when the club strikes the ball is included, in this case by the means with which the club face is attached to the club body.—NAS

7,565,951

43.55.Ev WALL MOUNTABLE ACOUSTIC ASSEMBLY FOR INDOOR ROOMS

Joab Jay Perdue, Amarillo, Texas
28 July 2009 (Class 181/287); filed 4 August 2006

A sound absorbing panel attached to a wall by a hinge at its bottom allows for variability and tuning of acoustics in a listening room. When the panel is allowed to open away from the wall, the absorptive quality of the panel is increased due to the airspace behind the panel.—CJR

7,571,790

43.55.Ev SOUND ABSORPTION BLOCK AND METHOD OF CONSTRUCTING THE SAME

Bae-Young Kim, Jeollanam-Do 526-853, Republic of Korea
11 August 2009 (Class 181/293); filed in Republic of Korea 30 June 2004

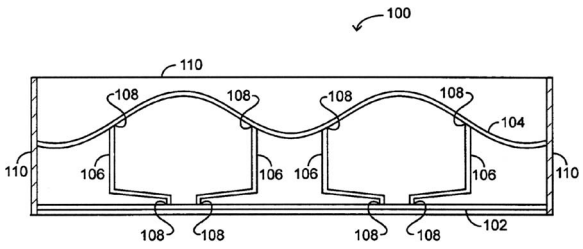
These are multiple plates, perhaps made of metal, that can be stacked together to form a block. Angled side walls of multiple openings funnel sound from one layer to the layers below. In this somewhat convoluted manner, a porous sound absorptive block is created.—CJR

7,565,950

43.55.Rg SOUND SUPPRESSION MATERIAL AND METHOD

Gary F. Hawkins and Michael J. O'Brien, assignors to The Aerospace Corporation
28 July 2009 (Class 181/207); filed 26 January 2005

The two sides of a wall panel are held together by flexible L-shaped brackets instead of by rigid framing. This allows for the outer face of the panel to move in a manner that sets up destructive interference in the radiating sound wave at certain frequencies, and this improves the sound barrier performance of the panel. This light-weight panel is best suited for aircraft or automobiles.—CJR



7,562,743

43.55.Ti ACOUSTICAL WINDOW AND DOOR COVERING

Lisa A. Beeson and Joseph G. Gaddone, assignors to Quietly Making Noise, LLC
21 July 2009 (Class 181/287); filed 2 December 2004

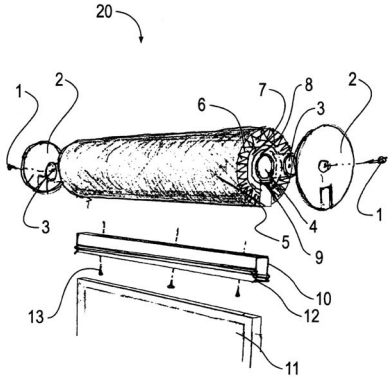
The patent describes an acoustically designed shutter for a window or door that is exposed to exterior noise. The shutter comes in a variety of configurations and includes, as one package, a decorative face, a noise barrier middle element, and a sound absorptive finish that would face inward toward the window or door opening. This makes a nice all-contained package for upgrading the sound reduction from environmental noise, though hardly different from any well-designed sound barrier shutter.—CJR

7,574,833

43.55.Ti ACOUSTIC DEVICE FOR USE ON OFFICE PARTITIONS

Charles H. Stark, Maumee, Ohio
18 August 2009 (Class 52/144); filed 28 February 2007

This design calls for a tubular pillow (a glass fiber cylinder) to be placed on top of a typical open-office partition. An impinging sound wave is either diffracted or absorbed (or both), thus improving privacy between adjacent cubicles.—CJR

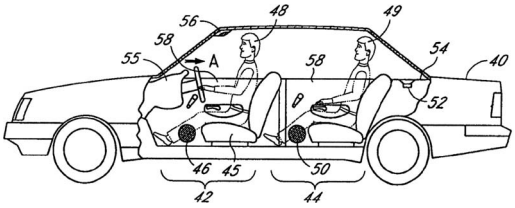


7,555,130

43.60.Dh ACOUSTIC CORRECTION APPARATUS

Arnold I. Klayman and Alan D. Kraemer, assignors to SRS Labs, Incorporated
30 June 2009 (Class 381/1); filed 10 November 2005

A method of processing a stereo audio signal is described wherein the stereo image is corrected to account for the less than optimal placement of transducers 46 50 52 56 in an automobile. For many, the antecedents of this patent are familiar as it is installed in televisions (identified by the SRS logo).—NAS



7,539,083

43.60.Vx REMOTE VOICE DETECTION SYSTEM

Fletcher A. Blackmon and Lynn T. Antonelli, assignors to The United States of America as represented by the Secretary of the Navy
26 May 2009 (Class 367/198); filed 25 June 2007

This patent describes a scheme for using a laser Doppler vibrometer aimed at a person's throat, in order to detect the voice sound. At first, this seemed sensible, until it was made clear that the intent of all of this is to effectively substitute for a microphone. That is, it is not simply the presence

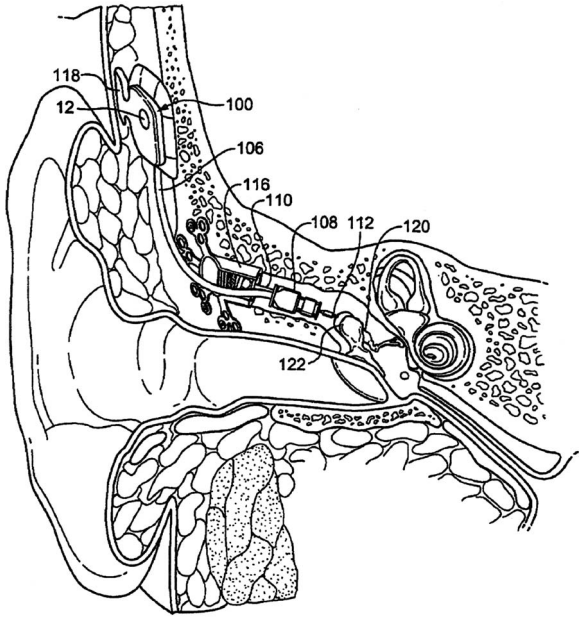
of voicing that is to be detected but the actual sound of the speech. This application seems futile since it is, one might have thought, generally known that the sound of a person's speech is not entirely represented in the vibrations of their throat.—SAF

7,556,597

43.66.Ts ACTIVE VIBRATION ATTENUATION FOR IMPLANTABLE MICROPHONE

Scott Allan Miller III *et al.*, assignor to Otologics, LLC
7 July 2009 (Class 600/25); filed 5 November 2004

A motion transducer is added to microphone/power/processing package 100, of an implantable hearing instrument that provides direct stimulation of the incus 122, to detect and so reduce undesirable motion of the microphone diaphragm.—NAS

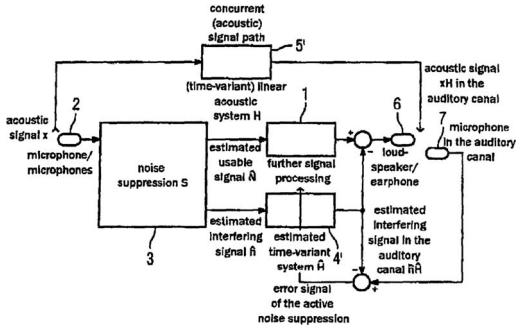


7,574,012

43.66.Ts HEARING AID WITH NOISE SUPPRESSION, AND OPERATING METHOD THEREFOR

Joachim Eggers and Volkmar Hamacher, assignor to Siemens Audiologische Technik GmbH
11 August 2009 (Class 381/317); filed in Germany 10 December 2003

Hearing aid fittings with large vents that leave the ear canal open can produce a comb filtering frequency response caused by an out-of-phase parallel combination of the direct acoustic path signal and the hearing aid-processed signal. To eliminate the comb filtering effect, a usable signal and



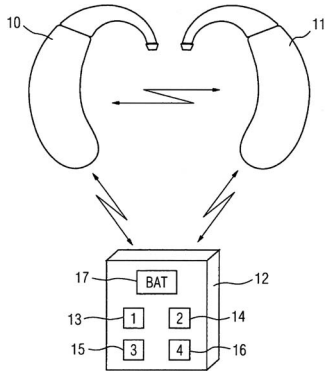
an interference signal are estimated. Interference compensation, which is fed in phase opposition to the ear canal, is generated by applying a simulated acoustic path transfer function to the interference signal. The simulated transfer function may be varied with an ear canal microphone pickup.—DAP

7,577,267

43.66.Ts HEARING AID AND HEARING AID SYSTEM

Roland Barthel *et al.*, assignors to Siemens Audiologische Technik gmbH
18 August 2009 (Class 381/315); filed in Germany 21 May 2004

Status information, such as expected remaining battery life, is determined and presented to the hearing aid wearer as an acoustic output signal comprising a tone rising or falling in frequency when the parameter is above or below a pre-defined threshold, respectively. Also claimed is an input



device for requesting status, e.g., by acoustic command or a mechanical control element, the status being provided via an acoustic output voice signal. Coverage for binaural hearing aid fittings is a further element of the patent.—DAP

7,577,268

43.66.Ts STABILIZATION OF THE SYSTEM CLOCK IN A HEARING AID

Eghart Fischer, assignor to Siemens Audiologische Technik GmbH
18 August 2009 (Class 381/315); filed in Germany 2 August 2004

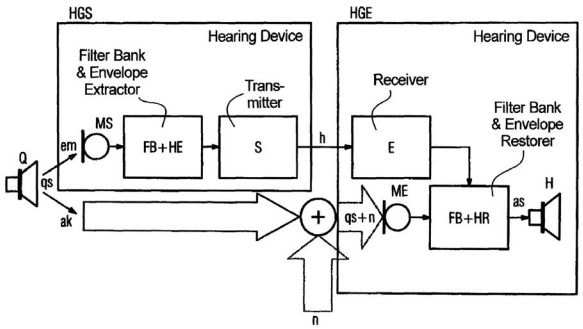
In order for digital hearing aids to achieve correct performance and functionality, their system clocks must remain stable, ideally, by employing a quartz crystal for the oscillator. However, since quartz crystals are too large for many hearing aids, instead, a periodic, externally generated clock is transmitted electromagnetically to the hearing aid to stabilize the system clock without using a quartz crystal. A remote control may be employed for the electromagnetic transmitter.—DAP

7,580,534

43.66.Ts METHOD AND ARRANGEMENT FOR TRANSMITTING SIGNALS TO A HEARING AID

Eghart Fischer, assignor to Siemens Audiologische Technik GmbH
25 August 2009 (Class 381/314); filed in Germany 5 March 2004

The data rate required for ear-to-ear transmissions between two wireless hearing aids in a binaural fitting is reduced by sending, electromagnetically only, the envelopes extracted in multiple frequency bands from the



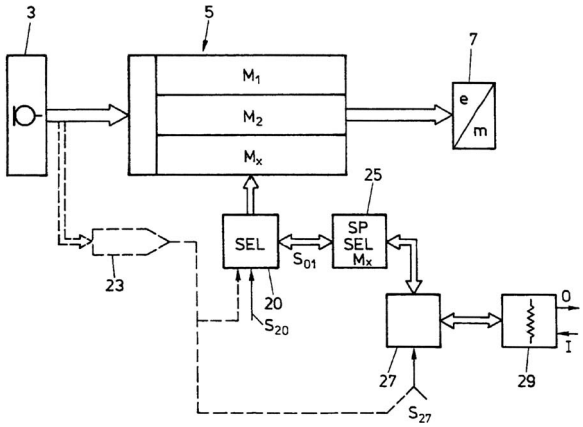
sound input signal picked up by the first microphone. At the receiving end, an output signal for the second hearing aid is generated by modulating the output from the second hearing aid microphone with the envelopes of the transmitted electromagnetic signals.—DAP

7,580,535

43.66.Ts BINAURAL SYNCHRONIZATION

Herbert Baechler, assignor to Phonak AG
25 August 2009 (Class 381/315); filed in the World Intellectual Property Organization 15 October 1999

Modern hearing aids may include several switch-selected modes of operation. If a user is fitted with aids in both ears, the selection of identical modes can become confusing. One prior art solution is a separate wireless



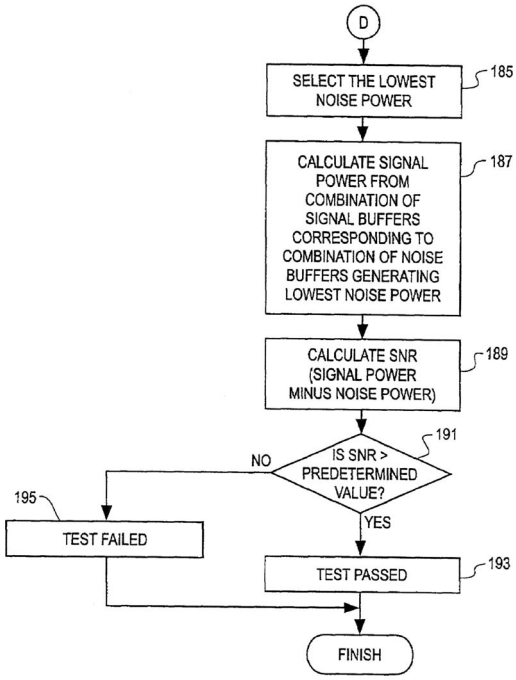
remote control unit, which has obvious disadvantages. This patent suggests an arrangement whereby one (master) unit is switchable and transmits information to a receiver in the second (slave) unit.—GLA

7,574,007

43.66.Yw STIMULUS ANALYSIS SYSTEM AND METHOD HAVING ADAPTIVE ARTIFACT REJECTION

Gregory R. Shaw and Mead C. Killion, assignor to Etymotic Research, Incorporated
11 August 2009 (Class 381/94.7); filed 1 November 2001

A signal averaging method is suggested for determining whether a response signal is acceptable, for example, in distortion product and transient evoked otoacoustic emission (DPOAE and TEOAE) testing. The response signal is repeatedly separated into signal and noise components, from



which noise power is calculated and, if it is low enough, the signal and noise components are stored into buffers. For signal components corresponding to the noise components having the lowest noise power, a signal to noise ratio is calculated and compared to a predetermined value to determine if the response is kept or discarded.—DAP

7,568,963

43.72.Fx INTERACTIVE TOYS

Alon Atsmon *et al.*, assignors to BeepCard Limited
4 August 2009 (Class 446/175); filed in Israel 16 September 1998

Picture a toy mother goose commanding a small squadron of toy goslings. The invention covers the use of acoustic signals between the leader (in this case, mom) and the followers. Other extensions of acoustic signaling are given, e.g., whip sounds and horse galloping.—MK

7,542,905

43.72.Ja METHOD FOR SYNTHESIZING A VOICE WAVEFORM WHICH INCLUDES COMPRESSING VOICE-ELEMENT DATA IN A FIXED LENGTH SCHEME AND EXPANDING COMPRESSED VOICE-ELEMENT DATA OF VOICE DATA SECTIONS

Reishi Kondo, assignor to NEC Corporation
2 June 2009 (Class 704/258); filed in Japan 28 March 2001

A compression scheme for a typically designed concatenative "voice-rule" speech synthesizer is described. The general method appears to involve compressing every other frame of each speech unit that is stored, "by using data of at least one preceding frame and/or at least one succeeding frame," though the compression method itself is intentionally left vague. The next step is to discard the preceding and following frames that were used to derive the compression, but which were not themselves compressed. The remaining compressed frame is then expanded to cover the missing spaces in the signal. It is plain that this would reduce storage requirements, whether it would sound very nice is less plain.—SAF

7,546,241

43.72.Ja SPEECH SYNTHESIS METHOD AND APPARATUS, AND DICTIONARY GENERATION METHOD AND APPARATUS

Masayuki Yamada *et al.*, assignors to Canon Kabushiki Kaisha
9 June 2009 (Class 704/260); filed in Japan 5 June 2002

This patent lays out a garden-variety concatenative speech synthesis scheme, in which the added innovation is a "spectrum correction filter." This filter is described in forms that would apply to speech "microsegments" parametrized using linear prediction or cepstrum coefficients. The filter is then applied to windowed microsegments in order to correct the spectral blurring introduced by the tapering function, thus obviating the disadvantages of concatenating windowed microsegments. It is not immediately clear how this would work, nor is it explained or demonstrated. Seven embodiments are presented which provide slight variations on the method of filter calculation.—SAF

7,574,358

43.72.Ne NATURAL LANGUAGE SYSTEM AND METHOD BASED ON UNISOLATED PERFORMANCE METRIC

Sabine Deligne *et al.*, assignors to International Business Machines Corporation
11 August 2009 (Class 704/243); filed 28 February 2005

The patent describes a business task management system based on the recognition of spoken commands. Separate modules perform speech recognition and task organization and management. Like most such systems, both the speech and the business task recognition are based on statistical analyses of the respective domains. What is arguably new about this system is that the training of such statistics between the language and the business management is inter-related. Joint model training across all domains can be repeated as required until the desired performance level is reached.—DLR

7,574,362

43.72.Ne METHOD FOR AUTOMATED SENTENCE PLANNING IN A TASK CLASSIFICATION SYSTEM

Marilyn A. Walker *et al.*, assignors to AT&T Intellectual Property II, L.P.
11 August 2009 (Class 704/275); filed 11 March 2002

This interactive dialog system may involve either speech or text for both input and output sides of the user interaction. By being designed to handle a specific task domain, in this case, travel arrangements, the system can begin with a specific set of goals, such as the date, departure time and place, and destination place. During the conversation, these goals may be fulfilled in any sequence, using an intelligent interactive protocol to proceed. It may be of some interest to those involved with speech processing that most of the patent deals with language issues. The speech analysis and text-to-speech aspects are all but taken for granted.—DLR

7,584,103

43.72.Ne AUTOMATED EXTRACTION OF SEMANTIC CONTENT AND GENERATION OF A STRUCTURED DOCUMENT FROM SPEECH

Juergen Fritsch *et al.*, assignors to Multimodal Technologies, Incorporated
1 September 2009 (Class 704/257); filed 20 August 2004

The patented system could be called a report generator. Based on a spoken audio input stream, the system would organize the subject matter and produce a structured document describing the transaction. There is no provision for interaction. It is described as a data entry process, although there does seem to be some possibility of an error recovery step. In this case, the subject matter is a medical patient presentation. The speech processor has access to a variety of language models covering various aspects of the patient/physician interaction as described in the speech entry sequence. The system then organizes the collected information and produces a standardized report describing the transaction. Other suggested areas in which the system could be applied include an attorney's interaction with a client.—DLR

7,539,616

43.72.Pf SPEAKER AUTHENTICATION USING ADAPTED BACKGROUND MODELS

Zhengyou Zhang and Ming Liu, assignors to Microsoft Corporation
26 May 2009 (Class 704/233); filed 20 February 2006

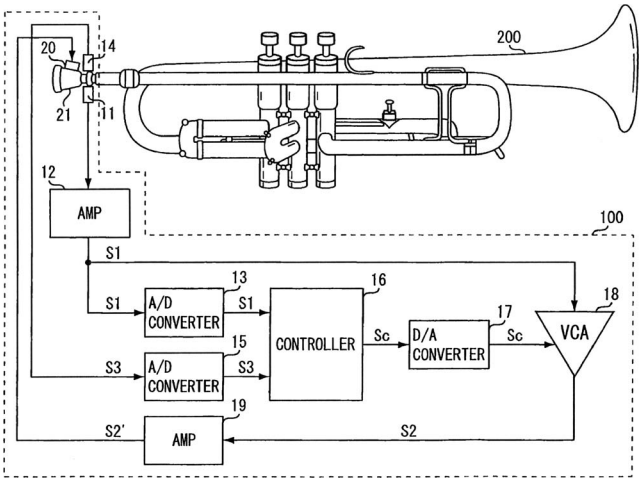
Here is a typical Microsoft patent which is concise, readable, sufficiently detailed, and actually patents something which is worth the effort of filing. Methods for speaker verification are described, which use standard frame-by-frame mel-frequency cepstral coefficient vectors in novel ways to compare users to a speaker pool. In a text-independent embodiment, the key innovations are a universal background model trained from the entire speaker pool, as well as speaker and background model adaptation procedures and similarity tests that have a unique design. A text-dependent embodiment adds a hidden Markov model of the authentication utterance as an element of the procedure. All the above mentioned aspects have equations provided and adequate discussion of how they are incorporated into the overall process.—SAF

7,554,028

43.75.Fg PERFORMANCE ASSIST APPARATUS OF WIND INSTRUMENT

Junji Fujii, assignor to Yamaha Corporation
30 June 2009 (Class 84/741); filed in Japan 27 December 2005

This mixed analog and digital system is designed to "assist" the wind



(brass) player by digitizing both pressure and vibration. A controller generates analog control via a voltage controlled amplifier which in turn controls the actuator.—MK

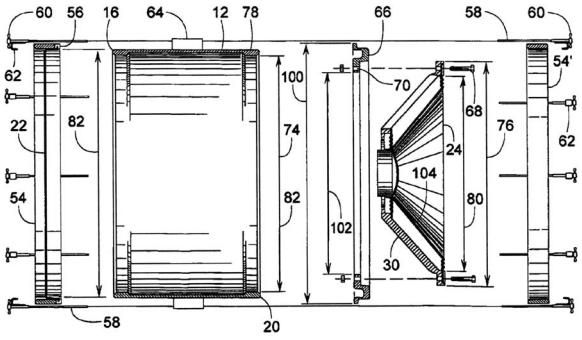
7,560,632

43.75.Hi BASS DRUM WITH COMPLIANT RESONANT HEAD

Kenneth W. Lanzel and Ryan R. Torgerson, both of La Crosse, Wisconsin

14 July 2009 (Class 84/411 R); filed 17 July 2006

Without explanation or excuse, the authors believe that adding a passive radiator to a bass drum (specifically, an emasculated loudspeaker) will add the desired "thump."—MK



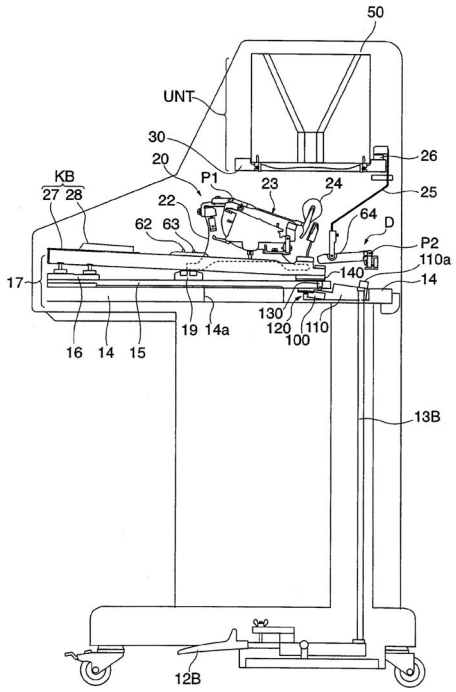
7,560,629

43.75.Kk KEYBOARD-TYPE PERCUSSION INSTRUMENT

Norishige Terada and Tetsuo Hotta, assignors to Yamaha Corporation

14 July 2009 (Class 84/236); filed in Japan 11 January 2007

One of a series of recent celeste patents by Yamaha, this disclosure is concerned specifically with the design of the pedal damping mechanism. Readers interested in celeste design will enjoy the diagrams.—MK



7,560,627

43.75.Mn SYSTEM AND METHOD FOR DRIVING ACTUATORS IN A REPRODUCING PIANO

Wayne Stahnke, Sunnyvale, California

14 July 2009 (Class 84/20); filed 4 April 2005

In a now primitive implementation, the author wants to use a finite state machine (FSM) for hammer actuation in a digital piano. The FSM clock can be varied to slow down the action.—MK

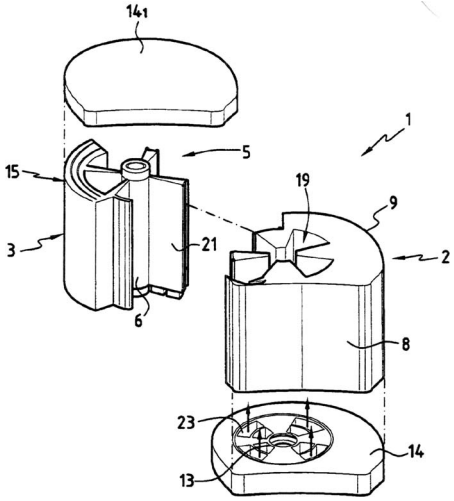
7,550,662

43.75.Pq MUSICAL INSTRUMENT WITH FREE REEDS

Pierre Bonnat, 01290 Grieges, France

23 June 2009 (Class 84/330); filed in France 27 November 2001

The accordion is being redesigned from an inward/outward motion to



a rotatary motion. The bellows are eliminated and replaced with the pistons shown.—MK

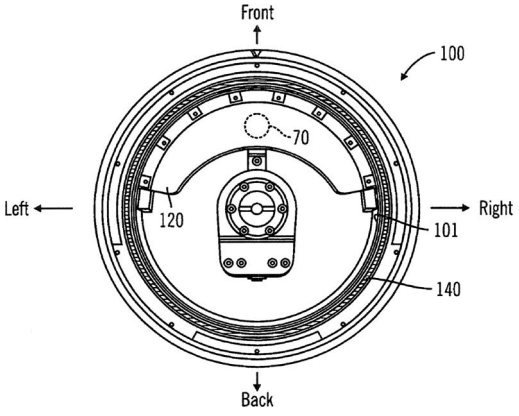
7,560,638

43.75.Tv ELECTRONIC PERCUSSION INSTRUMENT, SYSTEM, AND METHOD WITH VIBRATION

Ryo Susami, assignor to Roland Corporation

14 July 2009 (Class 84/723); filed in Japan 8 January 2004

An electronic high-hat includes both a displacement sensor 60 located



between the two cymbals and a vibration sensor 70 located on the upper-

cymbal. In addition, the conical coil spring separating the two cymbals 615 will press upon a printed resistor 613, so that the compression can be measured as a voltage change.—MK

7,569,758

43.75.Tv ELECTRONIC PERCUSSION SYSTEM AND ELECTRONIC PERCUSSION INSTRUMENT INCORPORATED THEREIN

Minoru Harada and Yuichirou Suenaga, assignors to Yamaha Corporation
4 August 2009 (Class 84/411 M); filed in Japan 7 August 2002

In this shamefully bad Japanese-to-English translation, the inventors propose adding a piezoelectric sensor to sense ambient vibrations in an electronic drum. Various construction details are included, the most noteworthy being the microcontroller algorithm.—MK

7,575,551

43.80.Vj BIOLOGICAL SIGNAL MONITOR DEVICE

Yoshinobu Watanabe *et al.*, assignors to Panasonic Corporation
18 August 2009 (Class 600/443); filed in Japan 1 September 2003

Movement in an ultrasound b-mode image is detected and compensated to show, for example, the contour of a blood vessel wall with less motion-induced degradation.—RCW

7,575,552

43.80.Vj ULTRASONIC PROBE WITH ACOUSTIC MEDIUM

Koetsu Saito, assignor to Panasonic Corporation
18 August 2009 (Class 600/459); filed in Japan 10 June 2004

The ultrasound transducer in this probe rotates in a chamber that contains 1,2-butylene glycol.—RCW

7,578,790

43.80.Vj SYSTEMS AND METHODS FOR DETECTING AND PRESENTING TEXTURAL INFORMATION FROM MEDICAL IMAGES

Shashidhar Sathyanarayana, assignor to Boston Scientific SciMed, Incorporated
25 August 2009 (Class 600/443); filed 20 July 2004

Texture in an uncompressed b-mode image is processed to obtain a color image that is superimposed on a log-compressed version of the original image.—RCW

7,578,791

43.80.Vj AUTOMATED MYOCARDIAL CONTRAST ECHOCARDIOGRAPHY

Patrick Rafter, assignor to Koninklijke Philips Electronics NV
25 August 2009 (Class 600/447); filed 16 April 2005

Ultrasound images of the heart are automatically acquired in two orientations of a plane by using a two-dimensional array system with settings that have been optimized. The acquisition sequence is performed initially with the heart unstressed and repeated with the heart under stress. The initial images and the images acquired during stress are compared to analyze reperfusion of the heart myocardium.—RCW

7,578,792

43.80.Vj AUTOMATIC OPTIMIZATION IN SPECTRAL DOPPLER ULTRASOUND IMAGING

Chi-Yin Lee *et al.*, assignors to Siemens Medical Solutions USA, Incorporated
25 August 2009 (Class 600/453); filed 21 July 2003

Doppler measurement parameters such as time-gate position and length, transmit frequency, filter settings, angle of intersection between the gate and the scan line, and aperture size are optimized by a processor that starts with parameters set by a user.—RCW