Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 308 (2007) 246–267
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

A smoothness index-guided approach to wavelet parameter selection in signal de-noising and fault detection

I. Soltani Bozchalooi, Ming Liang*

Department of Mechanical Engineering, University of Ottawa, 770 King Edward Avenue, Ottawa, Ontario, Canada K1N 6N5

Received 6 December 2006; received in revised form 12 April 2007; accepted 24 July 2007
Available online 4 September 2007

Abstract

Gabor wavelet transform can be used for de-noising impulsive signals measured from faulty bearings. However, it has been a challenging task to select proper wavelet parameters. This paper reports a method to guide the selection process by a smoothness index. The smoothness index is defined as the ratio of the geometric mean to the arithmetic mean of the wavelet coefficient moduli of the vibration signal. For the signal contaminated by Gaussian white noise, we have shown that the modulus of the wavelet coefficients follows Rician distribution. Based on this observation, we then prove that the smoothness index converges to a constant number (0.8455…) in the absence of mechanical faults or for very low signal to noise ratio. This result provides a dimensionless smoothness index upper bound corresponding to the most undesirable case. We have also shown that the smoothness index value decreases in the presence of impulses with properly selected parameters. The proposed method has been successfully used to de-noise both simulated and experimental signals.

1. Introduction

With the advances in digital signal processing methods, there has been an increasingly strong interest in the application of vibration analysis techniques for fault detection and diagnosis of rotating machinery elements. Bearings are among widely used and probably the most critical elements. As such, bearing fault detection and diagnosis has been the subject of extensive research [1–4]. Bearing faults usually appear as cracks or spalls on the surfaces of the bearing especially on the roller, outer race or inner race. When these faults come in contact with mating surfaces during the operation of the bearing, they generate impulses which are the main features to be detected through vibration measurement. The frequency of repetition of these impulses can be related to the rotational frequency of the shaft and the geometry of the bearing. A comprehensive analytical model for the vibration of a roller bearing with a single fault has been described by McFadden and Smith [1]. Comparing the analytical model with the measured vibration may lead to an effective method for fault detection. For this purpose, the fault features should be made evident via proper signal processing means as the measured signals are often contaminated by intensive noise, especially when the faults are at their early stages of development. As a result, a major step in fault

*Corresponding author. Tel.: +1 613 562 5800x6269; fax: +1 613 562 5177.
E-mail addresses: isolt038@uottawa.ca (I.S. Bozchalooi), liang@eng.uottawa.ca (M. Liang).

0022-460X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2007.07.038

Nomenclature

$a(t)$	local signal to noise ratio, which is equal to $\mu(t)/v$
$A_s(G_s)$	arithmetic (geometric) mean of a series
A	amplitude of the simulated impulse
$\mathrm{Ei}(x)$	exponential integral
$f_{X,Y}(x,y)$	joint probability density of random variables X and Y
f_0	modulation frequency of the mother Gabor wavelet
$F_Z(z)$	cumulative probability distribution of random variable Z
${}_1F_1$	confluent hypergeometric function
$F_x(\cdot)$	Fourier transform of a function with respect to variable x
I_0	modified Bessel function of first kind and zeroth order
$r_{G/A}$	ratio of the geometric mean to the arithmetic mean
Re (Im)	real (imaginary) part of a complex value
s	scale
t	time index
u	wavelet translation index
$u(t)$	unit step function
$v(t)$	vibration signature of a faulty bearing in the absence of noise
$V(t)$	measured vibration
$W_f(s,u)$	wavelet transform of the function $f(t)$ associated with scale s and translation index u
$W_f^{s,\sigma}(u)$	Gabor wavelet transform of a signal $f(t)$ associated with scale s and shape factor σ
β	structural damping coefficient
γ	Euler's constant
$\eta(s,\sigma)$	smoothness index as a function of scale s and shape factor σ
$\mu(t)$	modulus of the wavelet coefficients calculated for the noise-free vibration signal
v^2	noise variance
σ	shape factor
$\psi(t)$	mother wavelet
$\psi_{s,u}(t)$	daughter wavelet associated with scale s and translation index u
$\Psi(n)$	Euler's psi function
ω_0	excited resonance frequency

detection is de-noising. One such de-noising method is filter based. With this approach, high signal to noise ratio frequency band of the vibration signal passes through the filter whereas other frequency components are eliminated. Consequently the filtered signal should display the features of a faulty bearing more clearly. Bandpass filtering is also applied as a preprocessing step in the high frequency resonance technique [4]. Besides the filter-based de-noising, wavelet threshold de-noising such as the "soft-thresholding" method proposed by Donoho and Johnstone [5,6] has also been used for bearing vibration de-noising [7,8].

Though the comparison of the two methods seems in favor of the filter-based approach [9,10], its performance largely relies on the proper selection of parameters, i.e., the center frequency and the bandwidth of the filter. Several studies have attempted to address related issues by choosing proper parameters of a daughter Morlet wavelet and by applying the wavelet filter-based de-noising method [9–12]. In this study, we propose a new criterion to specify the center frequency and bandwidth by adjusting the scale and shape factor of the Gabor wavelet. The Gabor wavelet is used because of optimal time and frequency resolution [13]. The proposed criterion is reflected by the smoothness index (SI), which is the ratio of the geometric mean to the arithmetic mean of the wavelet coefficient moduli. We will demonstrate that the envelope of the bandpass filtered signal or the modulus of the wavelet coefficients follows the Rician [14] distribution. This result is later used to prove that the smoothness index converges to a constant number (0.8455…) in the absence of mechanical faults or when the signal to noise ratio of the vibration signal is very low. It is also shown that the smoothness index value decreases in the presence of fault signatures with properly selected parameters. As our purpose is to detect or reveal the fault signature, the selection of the Gabor wavelet parameters can be guided by the process of minimizing smoothness index.

This paper is organized as follows: In Section 2 we provide a brief introduction to wavelet transform, the idea of wavelet filter-based de-noising and the envelope signal to be used in the de-noising process. Section 3 presents the smoothness index and explains its application for the selection of the scale and the shape factor. The experimental results are given in Section 4. Section 5 concludes the paper.

2. Continuous wavelet transform (CWT) and envelope detection

The CWT of $f(t)$ with respect to a wavelet $\psi(t)$is defined as [13,15]

$$W_f(s,u) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\psi^*\left(\frac{t-u}{s}\right)\mathrm{d}t, \tag{1}$$

where s and u are real and asterisk denotes complex conjugate.

Eq. (1) can be written in a more compact form using:

$$\psi_{s,u}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right). \tag{2}$$

On the other hand, the convolution of two signals $h(t)$ and $x(t)$ is defined as

$$h(t) * x(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\,\mathrm{d}\tau. \tag{3}$$

We have [15]

$$f(u) * \psi^*_{s,0}(-u) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\psi^*\left(\frac{t-u}{s}\right)\mathrm{d}t = W_f(s,u) \tag{4}$$

Denoting Fourier transform of a function with respect to the variable x by F_x, we obtain

$$F_u[W_f(s,u)] = F_u[f(u)]F_u[\psi^*_{s,0}(-u)]. \tag{5}$$

According to Eq. (5) for a fixed scale we can consider wavelet transform $W_f(s,u)$ as the output of a filter with impulse response $\psi^*_{s,0}(-u)$ and input $f(u)$. In this paper, we choose Gabor wavelet as the mother wavelet. This wavelet is obtained through frequency modulation of a Gaussian window. Gaussian windows are used because of their optimal time and frequency energy concentration, proved by the Heisenberg Uncertainty theorem [13]. Gabor wavelet is defined as

$$\psi(t) = c\,\mathrm{e}^{-\sigma^2 t^2}\,\mathrm{e}^{\mathrm{i}2\pi f_0 t}. \tag{6}$$

In this paper, we choose constant c as

$$c = \sqrt[4]{\frac{8\sigma^2}{\pi}}, \tag{7a}$$

so that

$$\int_{-\infty}^{\infty} \mathrm{Re}(\psi(t))^2\,\mathrm{d}t \approx \int_{-\infty}^{\infty} \mathrm{Im}(\psi(t))^2\,\mathrm{d}t \approx 1. \tag{7b}$$

To elaborate on this, we write

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{Re}(\Psi(t))^2\,\mathrm{d}t &= \int_{-\infty}^{\infty} [c\,\mathrm{e}^{-\sigma^2 t^2}\cos 2\pi f_0 t]^2\,\mathrm{d}t = c^2\int_{-\infty}^{\infty} \mathrm{e}^{-2\sigma^2 t^2}\frac{1+\cos 4\pi f_0 t}{2}\,\mathrm{d}t \\ &= \frac{c^2}{2}\left[\int_{-\infty}^{\infty}\mathrm{e}^{-2\sigma^2 t^2}\,\mathrm{d}t + \int_{-\infty}^{\infty}\mathrm{e}^{-2\sigma^2 t^2}\cos 4\pi f_0 t\,\mathrm{d}t\right].\end{aligned} \tag{8a}$$

From Ref. [16] the above statement can be written as

$$\frac{c^2}{2}\left[\sqrt{\frac{\pi}{2\sigma^2}} + \mathrm{e}^{-2(\pi f_0)^2/\sigma^2}\sqrt{\frac{\pi}{2\sigma^2}}\right]. \tag{8b}$$

Similarly

$$\int_{-\infty}^{\infty} \mathrm{Im}(\Psi(t))^2\,\mathrm{d}t = \int_{-\infty}^{\infty} \left[c\,\mathrm{e}^{-\sigma^2 t^2}\,\sin 2\pi f_0 t\right]^2 \mathrm{d}t = c^2 \int_{-\infty}^{\infty} \mathrm{e}^{-2\sigma^2 t^2}\,\frac{1-\cos 4\pi f_0 t}{2}\,\mathrm{d}t$$
$$= \frac{c^2}{2}\left[\int_{-\infty}^{\infty} \mathrm{e}^{-2\sigma^2 t^2}\,\mathrm{d}t - \int_{-\infty}^{\infty} \mathrm{e}^{-2\sigma^2 t^2}\,\cos 4\pi f_0 t\,\mathrm{d}t\right] = \frac{c^2}{2}\left[\sqrt{\frac{\pi}{2\sigma^2}} - \mathrm{e}^{-2(\pi f_0)^2/\sigma^2}\sqrt{\frac{\pi}{2\sigma^2}}\right]. \tag{8c}$$

The last term inside the bracket in Eqs. (8b) and (8c) is very small for $f_0 \geqslant 1$ and $\sigma \leqslant 1$, which can be neglected. Hence, for $c = \sqrt[4]{8\sigma^2/\pi}$ the above leads to

$$\int_{-\infty}^{\infty} \mathrm{Re}(\psi(t))^2\,\mathrm{d}t \approx \int_{-\infty}^{\infty} \mathrm{Im}(\psi(t))^2\,\mathrm{d}t \approx \frac{c^2}{2}\sqrt{\frac{\pi}{2\sigma^2}} = 1. \tag{8d}$$

According to Eqs. (5) and (6), we have

$$F_u[W(s,u)] = \sqrt{s} F_u[f(u)]\hat{\psi}^*(sf), \tag{9}$$

where $\hat{\psi}(f)$ is the Fourier transform of $\psi(t)$ given as

$$\hat{\psi}(f) = \int_{-\infty}^{\infty} c\,\mathrm{e}^{-\sigma^2 t^2}\,\mathrm{e}^{\mathrm{i}2\pi f_0 t}\,\mathrm{e}^{-\mathrm{i}2\pi f t}\,\mathrm{d}t = c\sqrt{\frac{\pi}{\sigma^2}}\mathrm{e}^{-(\pi^2/\sigma^2)(f-f_0)^2}. \tag{10}$$

Eqs. (9) and (10) show that the wavelet transform at a constant scale acts like a bandpass filtering process with a Gaussian filter. As explained before, for the properly selected bandwidth and center frequency, this process results in a signal with higher signal to noise ratio. The bandwidth of this filter is adjusted by the shape factor σ and the center frequency is adjusted by the scale s as illustrated in Fig. 1.

On the other hand, according to Eq. (10) the filtered signal is analytic for $((\pi f_0)^2/\sigma^2) \gg 1$ [13]. Therefore, the modulus of this analytic result provides the envelope of the bandpass filtered signal [12,13]. Denoting the Gabor wavelet transform of a signal $f(t)$ at scale s and shape factor σ by $W_f^{s,\sigma}(u)$, we have

$$\text{Envelope of the bandpass filtered } f(t) = |W_f^{s,\sigma}(u)| = \sqrt{\mathrm{Re}\,(W_f^{s,\sigma}(u))^2 + \mathrm{Im}\,(W_f^{s,\sigma}(u))^2}. \tag{11}$$

This envelope will later be used in the de-noising process.

3. Ratio of geometric mean to arithmetic mean (G/A ratio) and its implication in parameter selection

3.1. The G/A ratio

The geometric mean of a series is defined as

$$G_s = \left(\prod_{n=1}^{N} S(n)\right)^{1/N}$$

for a positive time series $S(n)$ ($n = 1, 2, \ldots, N$).

Similarly the arithmetic mean of the series is

$$A_s = \frac{1}{N}\sum_{n=1}^{N} S(n).$$

We denote the ratio of the two as $r_{G/A}$, i.e.,

$$r_{G/A} = \frac{G_s}{A_s}. \tag{12}$$

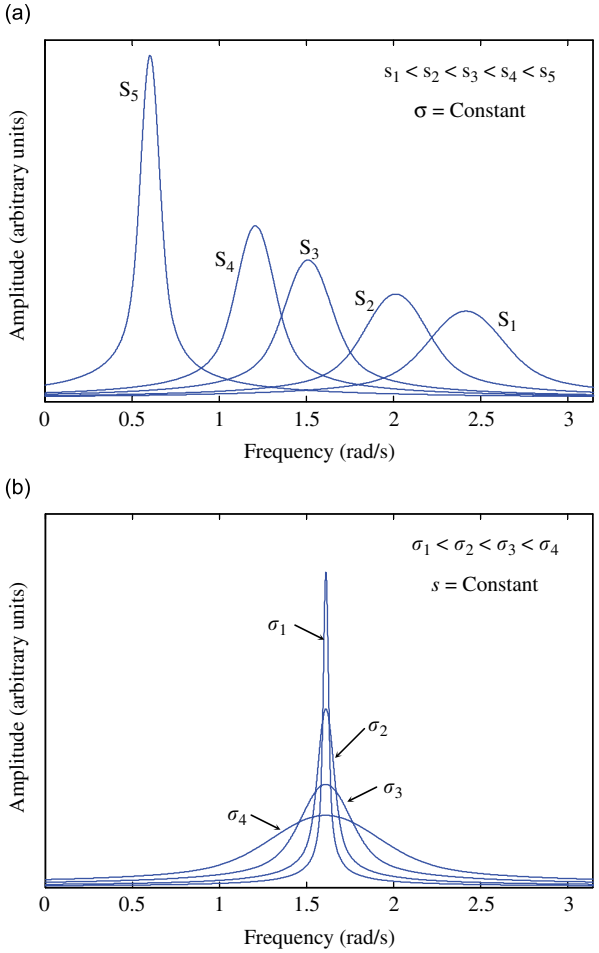

Fig. 1. Gabor wavelet in frequency domain for (a) fixed shape factor and different scales (b) fixed scale and different shape factor.

As $A_s \geqslant G_s$[17] (equality only holds when $S[n] = S[m]$ for all m and n), $r_{G/A}$ is always between zero and one for a positive time series. We rewrite geometric mean as follows:

$$G_s = \exp\left(\frac{1}{N}\sum_{n=1}^{N} \ln\left(S(n)\right)\right). \tag{13}$$

Then for a continuous function of time, we have

$$r_{G/A} = \frac{\exp\left[\frac{1}{T}\int_0^T \ln\left(f(t)\right) \mathrm{d}t\right]}{\frac{1}{T}\int_0^T f(t)\,\mathrm{d}t}, \tag{14}$$

where $f(t)$ is defined over $[0,T]$.

An important property of G/A ratio is that it approaches unity for flat functions and zero for peaky functions. In an extreme case, $r_{G/A} = 1$ if $f(t) =$ constant and $r_{G/A} \approx 0$ when $f(t)$ is highly impulsive in nature. The G/A ratio has been used as a measure of spectral flatness in speech signal processing [18].

3.2. The smoothness index and rationale of using it for parameter selection

As mentioned earlier, wavelet transform is to be used as a bandpass filter for noise reduction. The performance of such a filter depends on the selection of scale and shape factor of the daughter wavelet.

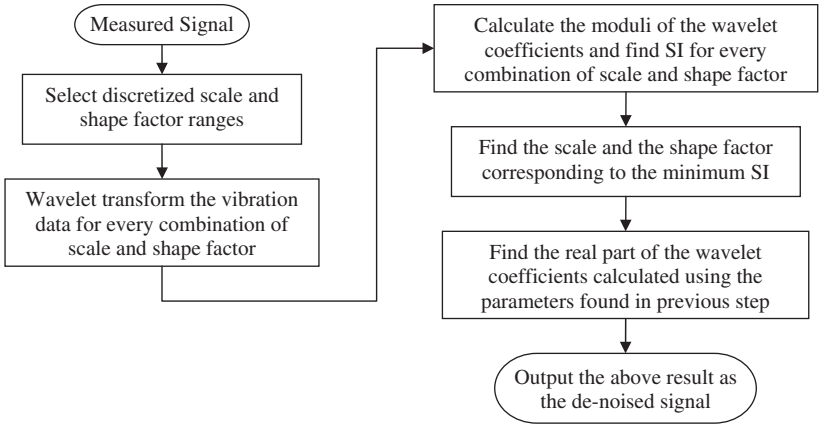

Fig. 2. Flowchart of the proposed de-noising algorithm.

For properly selected scale and shape factor, wavelet transform will result in higher coefficient values in the time interval where the fault generated impulses are located and lower coefficients elsewhere. This way, we can magnify parts of the signal that are of our interest and suppress the other parts. In other words, it is expected that the modulus of the wavelet transform forms a more peaky function of time when proper parameters are used. This is true when the vibration signal contains fault generated impulses. As such, the smoothness index defined below, can be used as a criterion to select the parameters, namely scale s and shape factor σ. Replacing $f(t)$ in Eq. (14) by $\left|W_f^{s,\sigma}(u)\right|$ given in Eq. (11) yields

$$\eta(s,\sigma)=\frac{\exp\left[\frac{1}{T}\int_0^T \ln\left(|W_f^{s,\sigma}(u)|\right)\mathrm{d}u\right]}{\frac{1}{T}\int_0^T |W_f^{s,\sigma}(u)|\,\mathrm{d}u}. \tag{15}$$

More specifically, the best parameters s and σ are achieved if $\eta(s,\sigma)$ is minimized. Fig. 2 summarizes the steps of the proposed method.

3.3. Probability density function (pdf) of wavelet coefficient modulus

To assess the behavior of the smoothness index, the statistical characteristics of the wavelet coefficient modulus are required, which will be detailed in the following. The fault signature resulting from the fault generated impulse may be expressed as

$$S(t)=A\,\mathrm{e}^{-\beta t}\cos(\omega_0 t)u(t), \tag{16}$$

where β is the structural damping coefficient, $u(t)$ is a unit step function and ω_0 is the excited resonance frequency. Due to the resemblance to a theoretical impulse, "fault generated impulse" or "impulse" is used interchangeably to refer to such signatures hereafter. Assuming such an impulse is generated each time when a fault comes in contact with the mating surface, we can model the measured vibration from a faulty bearing as a series of impulses with period T_p which corresponds to the characteristic fault period:

$$V(t)=\sum_{m=-M}^{M} A\,\mathrm{e}^{-\beta(t-mT_p)}\cos\,\omega_0(t-mT_p)u(t-mT_p)+w(t)=v(t)+w(t), \tag{17}$$

where $w(t)$ is continuous white Gaussian noise and $v(t)$ is the impact vibration due to the fault and contains $2M+1$ impulses. Wavelet transforming $V(t)$ using Eqs. (1) and (6), we obtain

$$W_V^{s,\sigma}(t) = \int_{-\infty}^{\infty} V(t')\frac{1}{\sqrt{s}}\psi^*\left(\frac{t'-t}{s}\right)\mathrm{d}t' = \int_{-\infty}^{\infty}\sum_{n=-M}^{M} A\,\mathrm{e}^{-\beta(t'-nT_p)}\cos\,\omega_0(t'-nT_p)u(t'-nT_p) \times\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2-\mathrm{i}(2\pi/s)(t'-t)}\,\mathrm{d}t' + \int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}\mathrm{e}^{-(\sigma/s)^2(t'-t)^2-\mathrm{i}(2\pi/s)(t'-t)}\,\mathrm{d}t'. \tag{18}$$

Substituting

$$W_v^{s,\sigma}(t) = \int_{-\infty}^{\infty}\sum_{m=-M}^{M} A\,\mathrm{e}^{-\beta(t'-mT_p)}\,\cos\,\omega_0(t'-mT_p)u(t'-mT_p)\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2-\mathrm{i}(2\pi/s)(t'-t)}\,\mathrm{d}t', \tag{19}$$

into Eq. (18) yields

$$W_V^{s,\sigma}(t) = W_v^{s,\sigma}(t) + \frac{c}{\sqrt{s}}\int_{-\infty}^{\infty} w(t')\mathrm{e}^{-(\sigma/s)^2(t'-t)^2-i(2\pi/s)(t'-t)}\,\mathrm{d}t', \tag{20}$$

where $W_v^{s,\sigma}(t)$ is the wavelet transform in the absence of noise. Decomposing the noise-related component into real and imaginary parts, we obtain

$$\begin{aligned} W_w^{s,\sigma}(t) &= \frac{c}{\sqrt{s}}\int_{-\infty}^{\infty} w(t')\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2-\mathrm{i}(2\pi/s)(t'-t)}\,\mathrm{d}t' \\ &= \int_{-\infty}^{\infty} w(t')\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\frac{c}{\sqrt{s}}\cos\frac{2\pi}{s}(t'-t)\mathrm{d}t' \\ &\quad -\mathrm{i}\int_{-\infty}^{\infty} w(t')\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\frac{c}{\sqrt{s}}\sin\frac{2\pi}{s}(t'-t)\mathrm{d}t'. \end{aligned} \tag{21}$$

As both real and imaginary parts in the above equation are linear transforms of Gaussian random variables, they are Gaussian random variables with zero mean and variance given by

$$\begin{aligned} v^2 &= \mathrm{var}(\mathrm{Re}(W_w^{s,\sigma}(t))) \\ &= \mathrm{var}\left(\int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\cos\frac{2\pi}{s}(t'-t)\,\mathrm{d}t'\right) \\ &= E\left\{\left[\int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\cos\frac{2\pi}{s}(t'-t)\,\mathrm{d}t'\right]^2\right\} \\ &= E\left\{\left[\int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\cos\frac{2\pi}{s}(t'-t)\,\mathrm{d}t'\right]\left[\int_{-\infty}^{\infty} w(t'')\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t''-t)^2}\,\cos\frac{2\pi}{s}(t''-t)\,\mathrm{d}t''\right]\right\} \\ &= E\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(t')w(t'')\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\cos\frac{2\pi}{s}(t'-t)\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t''-t)^2}\,\cos\frac{2\pi}{s}(t''-t)\,\mathrm{d}t'\,\mathrm{d}t''\right\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E[w(t')w(t'')]\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\cos\frac{2\pi}{s}(t'-t)\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t''-t)^2}\,\cos\frac{2\pi}{s}(t''-t)\,\mathrm{d}t'\,\mathrm{d}t'', \end{aligned}$$

where E is the statistical expectation operator. In the above equation, if $t'\neq t''$, $E[(w(t')w(t'']=0$ ($w(t')$ and $w(t'')$ are independent). For $t'=t''$, v^2 is given as follows:

$$\begin{aligned} v^2 = \mathrm{var}(\mathrm{Re}(W_w^{s,\sigma}(t))) &= \int_{-\infty}^{\infty} E\{w(t)'^2\}\left[\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\cos\frac{2\pi}{s}\,(t'-t)\right]^2\mathrm{d}t' \\ &\approx \mathrm{var}(w(t))\int_{-\infty}^{\infty}\left[\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\cos\frac{2\pi}{s}(t'-t)\right]^2\mathrm{d}t'. \end{aligned} \tag{22a}$$

On the other hand, according to Eq. (7b),

$$\int_{-\infty}^{\infty}\left[\frac{c}{\sqrt{s}}\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)\right]^2\mathrm{d}t' \approx \int_{-\infty}^{\infty}\left[\frac{c}{\sqrt{s}}\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\sin\frac{2\pi}{s}(t'-t)\right]^2\mathrm{d}t' \approx 1.$$

So we can write

$$\begin{aligned} v^2 &= \mathrm{var}\left(\mathrm{Re}\left(W_w^{s,\sigma}(t)\right)\right) = \mathrm{var}\left(w(t)\right)\int_{-\infty}^{\infty}\left[\frac{c}{\sqrt{s}}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)\right]^2\mathrm{d}t' \\ &\approx \mathrm{var}(\mathrm{Im}(W_w^{s,\sigma}(t))) = \mathrm{var}(w(t))\int_{-\infty}^{\infty}\left[\frac{c}{\sqrt{s}}\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\sin\frac{2\pi}{s}(t'-t)\right]^2\mathrm{d}t' \approx \mathrm{var}(w(t)). \end{aligned} \tag{22b}$$

According to Eqs. (20) and (22b), $W_V^{s,\sigma}(t)$ is a complex Gaussian random variable, and the real and imaginary parts of this random variable have means $\mathrm{Re}(W_v^{s,\sigma}(t))$ and $\mathrm{Im}(W_v^{s,\sigma}(t))$, respectively, and variance equal to the background noise variance.

In order to form the envelope of the wavelet transformed signal, we denote $X(t) = \mathrm{Re}(W_V^{s,\sigma}(t))$ and $Y(t) = \mathrm{Im}(W_V^{s,\sigma}(t))$ and obtain

$$Z(t) = |W_V^{s,\sigma}(t)| = \sqrt{X(t)^2 + Y(t)^2}. \tag{23}$$

Denoting the joint pdf of $X(t)$ and $Y(t)$ by $f_{x,y}(x,y)$, the cumulative probability distribution of $Z(t)$, $F_Z(z)$, can be written as [19]

$$F_{Z(t)}(z) = P(Z \leqslant z) = \int_{y=-z}^{z}\int_{x=-\sqrt{z^2-y^2}}^{\sqrt{z^2-y^2}} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

The Leibnitz differentiation rule is used to find the pdf of the wavelet coefficient modulus as follows:

$$\frac{\partial}{\partial z}\int_{a(z)}^{b(z)} f(x,z)\,\mathrm{d}x = \int_{a(z)}^{b(z)}\frac{\partial f}{\partial z}\,\mathrm{d}x + f(b(z),z)\frac{\partial b}{\partial z} - f(a(z),z)\frac{\partial a}{\partial z}.$$

Letting $g(y,z) = \int_{x=-\sqrt{z^2-y^2}}^{\sqrt{z^2-y^2}} f_{X,Y}(x,y)\mathrm{d}x$, we have

$$\begin{aligned} f_Z(z) &= \frac{\partial}{\partial z}F_Z(z) = \frac{\partial}{\partial z}\int_{y=-z}^{z} g(y,z)\,\mathrm{d}y \\ &= \int_{-z}^{z}\frac{\partial}{\partial z}g(y,z)\mathrm{d}y + g(z,z) - g(-z,z) \\ &= \int_{-z}^{z}\frac{\partial}{\partial z}g(y,z)\,\mathrm{d}y \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial z}g(y,z) &= \frac{\partial}{\partial z}\int_{x=-\sqrt{z^2-y^2}}^{\sqrt{z^2-y^2}} f_{X,Y}(x,y)\,\mathrm{d}x \\ &= \frac{z}{\sqrt{z^2-y^2}}\left\{f_{X,Y}\left(\sqrt{z^2-y^2},y\right) + f_{X,Y}\left(-\sqrt{z^2-y^2},y\right)\right\}, \end{aligned}$$

so that

$$f_Z(z) = \int_{-z}^{z}\frac{z}{\sqrt{z^2-y^2}}\left\{f_{X,Y}\left(\sqrt{z^2-y^2},y\right) + f_{X,Y}\left(-\sqrt{z^2-y^2},y\right)\right\}\mathrm{d}y. \tag{24}$$

According to Eqs. (20) and (21), $X(t)$ and $Y(t)$ are affine transformations of independent Gaussian random variables. As a result, linear combinations of the two or $aX(t)+bY(t)$ are Gaussian random variables for all a and b. Consequently we can consider joint normality for $X(t)$ and $Y(t)$ [19]. For this purpose, we define $C_{X,Y}$ as the covariance of $X(t)$ and $Y(t)$, i.e.,

$$C_{X,Y}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}E\{w(t')w(t'')\}\,\mathrm{e}^{-(\sigma/s)^2(t'-t)^2}\,\mathrm{e}^{-(\sigma/s)^2(t''-t)^2}\,\cos\frac{2\pi}{s}(t'-t)\,\sin\frac{2\pi}{s}(t''-t)\,\mathrm{d}t'\,\mathrm{d}t''=0. \tag{25}$$

Therefore

$$f_{X,Y}(x,y)=\frac{1}{2\pi v^2}\,\mathrm{e}^{-[(x-\mathrm{Re}\,[W_v^{s,\sigma}(t)])^2+(y-\mathrm{Im}\,[W_v^{s,\sigma}(t)])^2]/2v^2}. \tag{26}$$

Now letting [19]

$$\mu(t)=\sqrt{\mathrm{Re}\,(W_v^{s,\sigma}(t))^2+\mathrm{Im}\,(W_v^{s,\sigma}(t))^2}, \tag{27a}$$

$$y(t)=z(t)\,\sin\,\theta, \tag{27b}$$

$$\mathrm{Re}(W_v^{s,\sigma}(t))=\mu(t)\,\cos\,\phi, \tag{27c}$$

$$\mathrm{Im}(W_v^{s,\sigma}(t))=\mu(t)\,\sin\,\phi \tag{27d}$$

and substituting Eqs. (26) and (27) into Eq. (24), we obtain

$$\begin{aligned}
f_Z(z)&=\frac{1}{2\pi v^2}\int_{-\pi/2}^{\pi/2}\frac{z}{z\,\cos\,\theta}\Big\{\mathrm{e}^{-[(z\,\cos\,\theta-\mu(t)\cos\,\phi)^2+(z\,\sin\,\theta-\mu(t)\sin\,\phi]^2/2v^2}\\
&\quad+\mathrm{e}^{-[(-z\,\cos\,\theta-\mu(t)\cos\,\phi)^2+(z\,\sin\,\theta-\mu(t)\sin\,\phi]^2/2v^2}\Big\}z\,\cos\,\theta\,\mathrm{d}\theta\\
&=\frac{z\,\mathrm{e}^{-\left(z^2+\mu(t)^2\right)/2v^2}}{2\pi v^2}\int_{-\pi/2}^{\pi/2}(\mathrm{e}^{z\mu(t)\cos\,(\theta-\phi)/v^2}+\mathrm{e}^{-z\mu(t)\cos\,(\theta+\phi)/v^2})\,\mathrm{d}\theta\\
&=\frac{z\,\mathrm{e}^{-\left(z^2+\mu(t)^2\right)/2v^2}}{2\pi v^2}\left(\int_{-\pi/2}^{\pi/2}\mathrm{e}^{z\mu(t)\cos\,(\theta-\phi)/v^2}\,\mathrm{d}\theta+\int_{\pi/2}^{3\pi/2}\mathrm{e}^{z\mu(t)\cos\,(\theta-\phi)/v^2}\,\mathrm{d}\theta\right)\\
&=\frac{z\,\mathrm{e}^{-\left(z^2+\mu(t)^2\right)/2v^2}}{2\pi v^2}\int_0^{2\pi}\mathrm{e}^{z\mu(t)\cos\,(\theta-\phi)/v^2}\,\mathrm{d}\theta=\frac{z\,\mathrm{e}^{-(z^2+\mu(t)^2)/2v^2}}{\pi v^2}\int_0^{\pi}\mathrm{e}^{z\mu(t)\cos\,(\theta)/v^2}\,\mathrm{d}\theta,
\end{aligned}$$

i.e.,

$$f_Z(z)=\frac{z\,\mathrm{e}^{-(z^2+\mu(t)^2)/2v^2}}{v^2}I_0\left(\frac{z\mu(t)}{v^2}\right), \tag{28}$$

where I_0 is the modified Bessel function of the first kind and zeroth order [17] defined as

$$I_0(\kappa)=\frac{1}{\pi}\int_0^{\pi}\mathrm{e}^{\kappa\,\cos\,(\theta)}\,\mathrm{d}\theta.$$

This is the expression for a Rician distributed random variable. Fig. 3 shows this distribution for different values of v^2 and $\mu(t)$. Letting $H(t)=Z(t)/v$ and $a(t)=\mu(t)/v$ and denoting the pdf of the random variable $H(t)$ by $f_H(h)$, Eq. (28) becomes [14]

$$f_H(h)=h\mathrm{e}^{-\left(h^2+a(t)^2\right)/2}I_0(a(t)h). \tag{29}$$

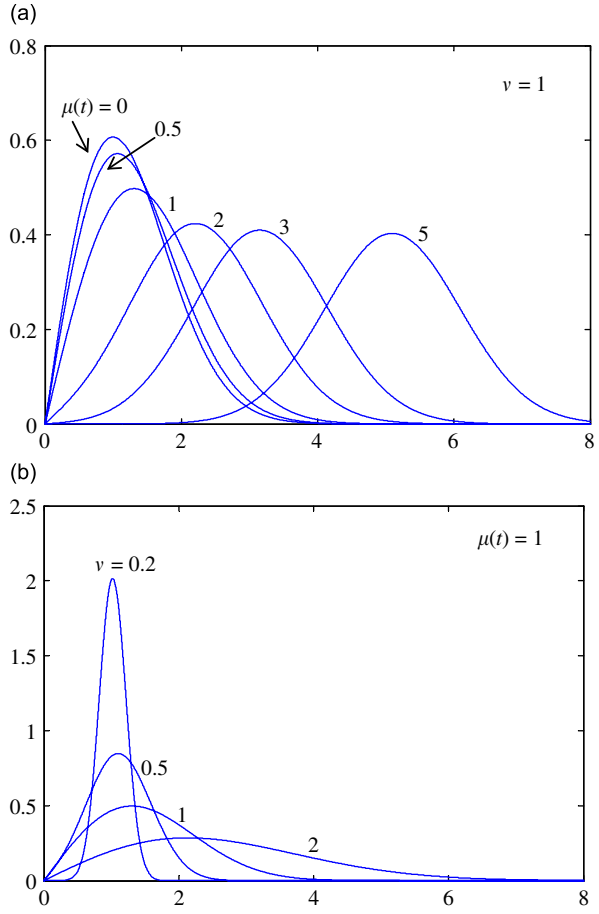

Fig. 3. Rician distribution for (a) $v = 1$ and different $\mu(t)$ and (b) $\mu(t) = 1$ and different v.

3.4. Behavior of smoothness index for different values of local signal to noise ratio ($a(t)$)

As shown in Eq. (19), $W_v^{s,\sigma}(t)$ is a periodic function of time with period T_p. Hence, for a large M, the smoothness index can be written as

$$\eta(s,\sigma) = \frac{\exp\left[\int_{-MT_p}^{(M+1)T_p} \frac{1}{(2M+1)T_p} \ln(|W_V^{s,\sigma}(t)|)\,\mathrm{d}t\right]}{\frac{1}{(2M+1)T_p}\int_{-MT_p}^{(M+1)T_p} |W_V^{s,\sigma}(t)|\,\mathrm{d}t} \approx \frac{\exp\left[\int_0^{T_p} \frac{1}{T_p} E\{\ln(Z(t))\}\,\mathrm{d}t\right]}{\frac{1}{T_p}\int_0^{T_p} E\{Z(t)\}\,\mathrm{d}t}. \tag{30}$$

Replacing $Z(t)$ with $H(t)v$ in Eq. (30) leads to

$$\eta(s,\sigma) \approx \frac{\exp\left[\ln v + \int_0^{T_p} \frac{1}{T_p} E\{\ln(H(t))\}\,\mathrm{d}t\right]}{\frac{v}{T_p}\int_0^{T_p} E\{H(t)\}\,\mathrm{d}t} = \frac{\exp\left[\int_0^{T_p} \frac{1}{T_p} E\{\ln(H(t))\}\,\mathrm{d}t\right]}{\frac{1}{T_p}\int_0^{T_p} E\{H(t)\}\,\mathrm{d}t}. \tag{31}$$

The above equation provides a general expression of smoothness index. As mentioned earlier, the smoothness index approaches zero for signals of impulsive nature. However, the smoothness index behavior should be analyzed when the signal is corrupted with background noise. More detailed analysis is given in the following.

3.4.1. Smoothness index behavior under low local signal to noise ratio

For very low signal to noise ratio, i.e., $a(t) \ll 1$, the following holds [17]:

$$\lim_{a(t)h\to 0} I_0(a(t)h) = \frac{1}{\Gamma(1)} = 1. \tag{32a}$$

From the above result, when $a(t)$ approaches zero, Eq. (29) can be written as

$$\lim_{a(t)\to 0} f_H(h) = h\,\mathrm{e}^{-h^2/2}. \tag{32b}$$

Eq. (32b) is the expression of Rayleigh distribution. Considering Eq. (32b), (31) can be written as

$$\eta(s,\sigma) = \frac{\exp[E\{\ln(H)\}]}{E\{H\}}, \tag{33}$$

where H is a Rayleigh distributed random variable. As $E\{\ln(H)\} = \int_0^\infty \ln(h) h\,\mathrm{e}^{-h^2/2}\,\mathrm{d}h$, we obtain the following with a change of variable $r = h^2$:

$$\begin{aligned} E\{\ln(H)\} &= \int_0^\infty \ln(\sqrt{r})\sqrt{r}\,\mathrm{e}^{-r/2}\frac{1}{2\sqrt{r}}\,\mathrm{d}r \\ &= \frac{1}{4}\int_0^\infty \ln(r)\,\mathrm{e}^{-r/2}\,\mathrm{d}r = \frac{1}{4}\int_0^\infty (\ln\frac{r}{2} + \ln 2)\,\mathrm{e}^{-r/2}\,\mathrm{d}r \\ &= \frac{1}{4}\left[\int_0^\infty \ln\frac{r}{2}\,\mathrm{e}^{-r/2}\,\mathrm{d}r + \ln 2\int_0^\infty \mathrm{e}^{-r/2}\,\mathrm{d}r\right] \\ &= \frac{1}{4}\left[2\int_0^\infty \ln\frac{r}{2}\,\mathrm{e}^{-r/2}\,\mathrm{d}\frac{r}{2} + 2\ln 2\int_0^\infty \mathrm{e}^{-r/2}\,\mathrm{d}\frac{r}{2}\right] \\ &= \frac{1}{4}[-2\gamma + 2\ln 2] = -\frac{1}{2}\left(\gamma + \ln\frac{1}{2}\right), \end{aligned} \tag{34a}$$

where $\gamma = \lim_{s\to\infty}\left(\sum_{m=1}^{s}(1/m) - \ln(s)\right) = 0.5772156649\ldots$ is the Euler's constant.

In addition, the denominator of Eq. (33) can be written as [16]

$$E\{H\} = \int_0^\infty h f_H(h) = \int_0^\infty h^2\,\mathrm{e}^{-h^2/2}\,\mathrm{d}h = \sqrt{\frac{\pi}{2}}. \tag{34b}$$

Substituting Eqs. (34a) and (b) into Eq. (33) yields

$$\lim_{a(t)\to 0}\eta(s,\sigma) = \frac{\exp\left[-\frac{1}{2}\left(\gamma + \ln\frac{1}{2}\right)\right]}{\sqrt{\frac{\pi}{2}}} = 0.845501287\ldots. \tag{35}$$

According to Eq. (35), for very low signal to noise ratio or in the absence of any faults where the vibrations behave randomly, the smoothness index approaches a constant number.

3.4.2. Smoothness index behavior under higher local signal to noise ratio

In this section, we analyze the expected value and the expected logarithm of a Rician distributed random variable for a larger $a(t)$. The mean of such a random variable with distribution shown in Eq. (28) is [14]

$$E\{Z(t)\} = v\sqrt{2}\Gamma\left(\frac{3}{2}\right){}_1F_1\left(-\frac{1}{2}, 1, -\frac{\mu(t)^2}{2v^2}\right), \tag{36}$$

where ${}_1F_1$ is a confluent hypergeometric function [17]. According to Ref. [17, Eq. (13.1.5)], ${}_1F_1$ can be expressed as

$${}_1F_1\left(-\frac{1}{2}, 1, -\frac{\mu(t)^2}{2v^2}\right) = \frac{\Gamma(1)}{\Gamma(3/2)}\left(\frac{\mu(t)^2}{2v^2}\right)^{1/2}\left[1 + O\left(\left(\frac{\mu(t)^2}{2v^2}\right)^{-1}\right)\right].$$

Substituting the above equation into Eq. (36), we obtain

$$E\{Z(t)\} = v\sqrt{2}\Gamma\left(\frac{3}{2}\right)\frac{\Gamma(1)}{\Gamma(3/2)}\left(\frac{\mu(t)^2}{2v^2}\right)^{1/2}\left[1+O\left(\left(\frac{\mu(t)^2}{2v^2}\right)^{-1}\right)\right]$$
$$= \mu(t)\left(1+O\left(\left(\frac{\mu(t)^2}{2v^2}\right)^{-1}\right)\right), \tag{37}$$

where $O(x)$ is the order of the argument x. In the above equation we used the relation $\Gamma(1)=1$.

As a result, as $a(t)=\mu(t)/v$ increases, the mean of $Z(t)$ approaches $\mu(t)$ or the modulus of the wavelet coefficients of the noise-free vibration signal at a rate in the order of $(\mu(t)/2v^2)^{-1}$.

Moreover, considering $a'(t)=a(t)/\sqrt{2}$ and $H'(t)=H(t)/\sqrt{2}$, Eq. (29) can be written as

$$f_{H'}(h') = 2h'\,\mathrm{e}^{-(h'^2+a'(t)^2)}\,I_0(2a'(t)h'). \tag{38}$$

From the above expression we find the expected logarithm of $H'(t)$ as follows:

$$E\{\ln(H'(t))\} = \int_0^\infty 2h'\,\ln\,h'\,\mathrm{e}^{-(h'^2+a'(t)^2)}I_0(2a'(t)h')\,\mathrm{d}h'$$
$$= 2\mathrm{e}^{-a'(t)^2}\int_0^\infty h'\,\ln\,h'\,\mathrm{e}^{-h'^2}I_0(2a'(t)h')\,\mathrm{d}h'. \tag{39}$$

On the other hand $I_0(2a'(t)h')$ can be expressed as [17]

$$I_0(2a'(t)h') = \sum\nolimits_{k=0}^{\infty}\frac{(a'(t)h')^{2k}}{k!\Gamma(k+1)}. \tag{40}$$

Substituting Eq. (40) into Eq. (39), we obtain

$$E\{\ln(H'(t))\} = 2\mathrm{e}^{-a'(t)^2}\int_0^\infty h'\,\ln\,h'\,\mathrm{e}^{-h'^2}\sum_{k=0}^{\infty}\frac{(a'(t)h')^{2k}}{k!\Gamma(k+1)}\,\mathrm{d}h'$$
$$= 2\mathrm{e}^{-a'(t)^2}\sum_{k=0}^{\infty}\frac{a'(t)^{2k}}{k!\Gamma(k+1)}\int_0^\infty \ln\,h'\,\mathrm{e}^{-h'^2}h'^{2k+1}\,\mathrm{d}h'. \tag{41}$$

Letting $r=h'^2$ yields

$$\int_0^\infty \ln\,h'\,\mathrm{e}^{-h'^2}h'^{2k+1}\,\mathrm{d}h' = \int_0^\infty \frac{1}{4}\ln\,r\,\mathrm{e}^{-r}r^k\sqrt{r}\frac{1}{\sqrt{r}}\,\mathrm{d}r = \frac{1}{4}\int_0^\infty \ln\,r\,\mathrm{e}^{-r}r^k\,\mathrm{d}r.$$

From Eq. (4.352.1) in Ref. [16], we get

$$\frac{1}{4}\int_0^\infty \ln\,r\,\mathrm{e}^{-r}r^k\,\mathrm{d}r = \frac{1}{4}\Gamma(k+1)\Psi(k+1), \tag{42}$$

where $\Psi(n)$ is the Euler's psi function defined as $\Psi(n)=-\gamma+\sum_{k=1}^{n-1}k^{-1}$ and γ as mentioned before is the Euler's constant. Replacing the integral of Eq. (41) with Eq. (42) leads to

$$E\{\ln(H'(t))\} = 2\,\mathrm{e}^{-a'(t)^2}\sum\nolimits_{k=0}^{\infty}\frac{a'(t)^{2k}}{k!\Gamma(k+1)}\frac{1}{4}\,\Gamma(k+1)\Psi(k+1) = \frac{1}{2}\,\mathrm{e}^{-a'(t)^2}\sum\nolimits_{k=0}^{\infty}\frac{a'(t)^{2k}\Psi(k+1)}{k!}. \tag{43}$$

It is proven in Ref. [20] that

$$\mathrm{e}^{-a'(t)^2}\sum\nolimits_{k=0}^{\infty}\frac{a'(t)^{2k}\Psi(k+1)}{k!} = \ln(a'(t)^2) - \mathrm{Ei}(-a'(t)^2), \tag{44}$$

where $\mathrm{Ei}(x)$ is the exponential integral, defined as

$$\mathrm{Ei}(x) = -\int_{-x}^{\infty}\frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t.$$

Substituting Eq. (44) in Eq. (43) results in

$$E\{\ln(H'(t))\} = \frac{1}{2}\ln(a'(t)^2) - \frac{1}{2}\mathrm{Ei}(-a'(t)^2). \tag{45}$$

As $a'(t)^2$ increases, the exponential integral in Eq. (45) approaches zero. Considering the relations $a'(t) = \mu(t)/v\sqrt{2}$ and $H'(t) = Z(t)/v\sqrt{2}$, we obtain

$$E\left\{\ln\left(\frac{Z(t)}{v\sqrt{2}}\right)\right\} = \ln\left(\frac{\mu(t)}{v\sqrt{2}}\right) - \frac{1}{2}\mathrm{Ei}\left(-\left(\frac{\mu(t)}{v\sqrt{2}}\right)^2\right)$$

$$\Rightarrow E\{\ln(Z(t))\} = \ln(\mu(t)) - \frac{1}{2}\mathrm{Ei}\left(-\left(\frac{\mu(t)}{v\sqrt{2}}\right)^2\right). \tag{46}$$

Consequently, as the local signal to noise ratio or $a(t)$ increases, the expected logarithm of $Z(t)$ approaches $\ln(\mu(t))$ at a rate specified by the second term of the above equation. Substituting Eqs. (37) and (46) in Eq. (30) yields:

$$\eta(s,\sigma) = \frac{\exp\left[\frac{1}{T_p}\int_0^{T_p}\ln(\mu(t)) - \frac{1}{2}\mathrm{Ei}\left(-\left(\frac{\mu(t)}{v\sqrt{2}}\right)^2\right)\mathrm{d}t\right]}{\frac{1}{T_p}\int_0^{T_p}\mu(t)\left(1+O\left(\left(\frac{\mu(t)}{v\sqrt{2}}\right)^{-2}\right)\right)\mathrm{d}t}. \tag{47}$$

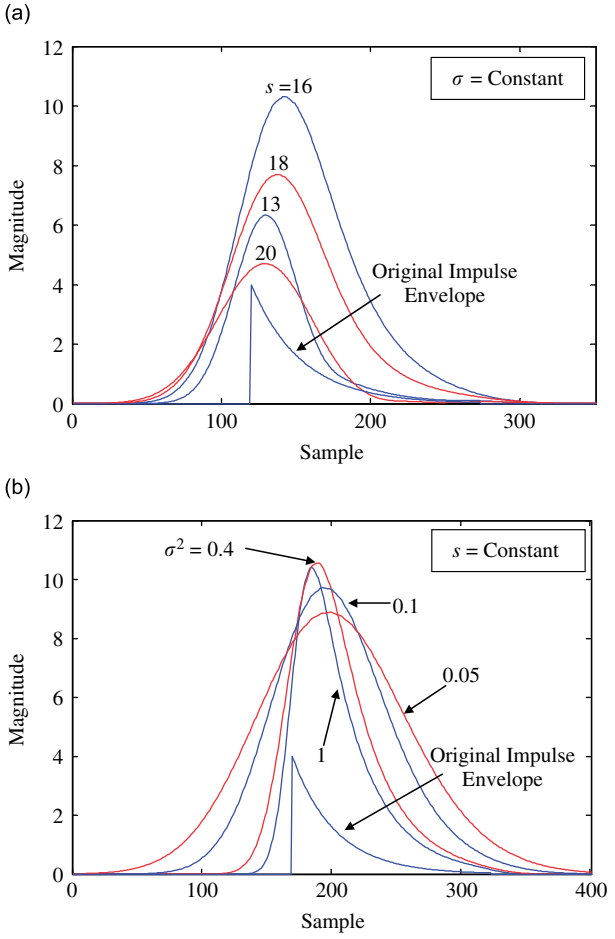

Fig. 4. (a) Wavelet coefficient modulus at different scales with fixed shape factor σ and (b) wavelet coefficient modulus for different shape factors at the scale associated with the resonance frequency.

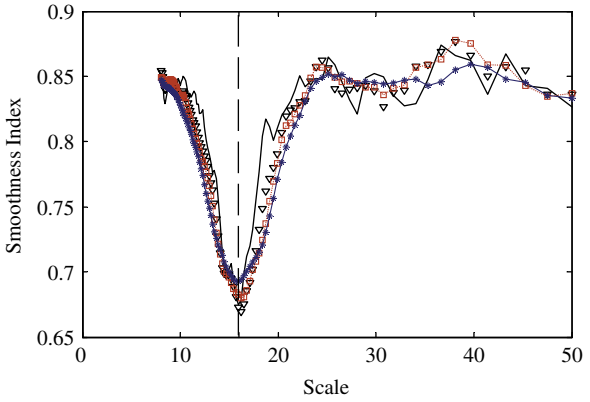

Fig. 5. Smoothness index versus scale for different shape factors calculated based on the simulated signal (Fig. 8(a)) with noise added (scale corresponding to the resonance frequency is shown by the vertical dashed line) ((—) $\sigma^2 = 0.07$, ($\triangledown$) $\sigma^2 = 0.19$, ($\square$) $\sigma^2 = 0.31$, ($*$) $\sigma^2 = 0.61$).

Eq. (47) reveals the degree of vigilance of the smoothness index to the background noise intensity. According to this equation, as $a(t) = \mu(t)/v$ increases, smoothness index approaches its noise-free environment (pure impulsive vibration) level which is closer to zero due to the impulsive nature of the faulty bearing vibration.

3.5. *Effect of wavelet parameters on smoothness index*

Fig. 4a shows the wavelet coefficient moduli associated with different scales for a constant shape factor. These moduli are calculated for a single noise-free simulated impulse. The scale corresponding to resonance frequency is 16. As one can see, the magnitude of the wavelet coefficient modulus at scale 16 is the highest of all the scales for all samples. Since noise variance v^2 remains constant as shown in Eq. (22b), the local signal to noise ratio, $a(t) = \mu(t)/v$, increases and according to Eq. (47) the smoothness index approaches its noise-free environment level. This was also expected intuitively—it is well known that the resonance frequency band corresponds to the high signal to noise ratio frequency region.

As shown in Fig. 5, the lowest smoothness index appears at a scale very close to the scale associated with resonance frequency which is in agreement with the above explanations. The effect of shape factor σ can also be explained in a similar manner. Fig. 4b shows the modulus of the wavelet coefficients at the scale associated with resonance frequency for different shape factor values. As one can see, by decreasing the shape factor, the magnitude of the wavelet coefficient modulus increases in most of sampling points whereas again according to Eq. (22b) the noise variance remains constant. As explained earlier, this in fact results in a higher local signal to noise ratio. To illustrate, we consider Fig. 6. Figs. 6a and b show the power spectral density (PSD) of the simulated signal of Fig. 8a with different noise intensity levels. Fig. 6 shows that increasing noise leads to a narrower high signal to noise ratio band in the frequency domain. This means that a smaller shape factor should be chosen to make the bandwidth of the filter narrower. However, wavelet transform is the weighted averaging and lowering the shape factor σ increases this averaging effect which results in a flatter signal with a higher signal to noise ratio. This phenomenon can also be seen in Fig. 4b. Besides, this smoothing effect causes the neighboring impulses to alias. It should also be noted that the decrease in σ has decreasing or advert influence on the wavelet coefficients through the constant c given in Eq. (7a). This effect is stronger for smaller shape factors.

Considering an extreme case where the bandwidth of the filter is extremely narrow (very small shape factor) and only the center frequency could pass, the corresponding envelope signal would be a flat function. On the other hand, for very large bandwidth (large shape factor) all the frequency components can pass through the filter and no noise reduction could take place. Hence the proper parameter is obtained when a desirable

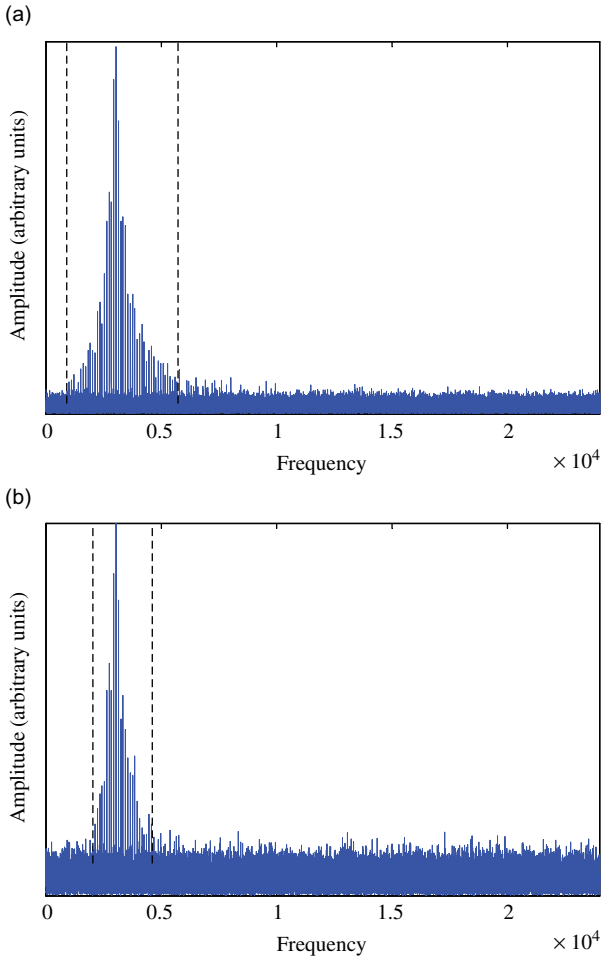

Fig. 6. PSD of the simulated signal from Fig. 8a: (a) signal to noise ratio of −7 dB and (b) signal to noise ratio of −13 dB. The frequencies between the vertical dashed lines indicate the approximate high signal to noise ratio band.

balance between the increase of signal to noise ratio and flatter time signal is achieved so that the fault generated impulses can be better identified from the result.

3.6. Comparison of the smoothness index minimization method with the kurtosis maximization criterion

As detailed earlier, the proposed smoothness index can quantify the impulsiveness of the bandpass filtered signal and hence be used as a criterion to guide the search for better wavelet parameters. A similar criterion is kurtosis which is defined as follows:

$$\mathrm{Kurt}(h) = \frac{E\{(H - E\{H\})^4\}}{E\{(H - E\{H\})^2\}^2} - 3, \tag{48}$$

where H is a random variable. This index is a measure of non-Gaussianity [21] of a dataset and has higher values for signals of impulsive nature. Accordingly, kurtosis has been widely employed in fault diagnosis research. As the bandpass filtering is usually a preprocessing step for envelope spectrum analysis, it has been

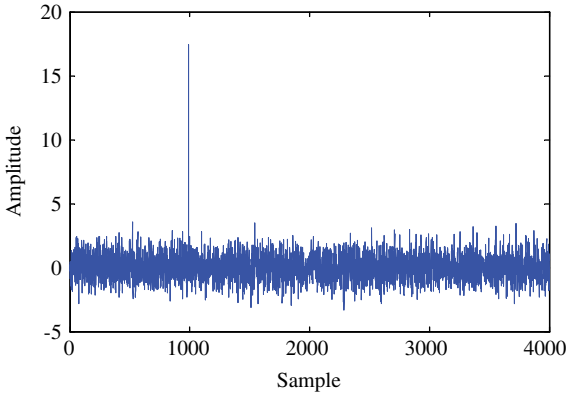

Fig. 7. Simulated white Gaussian noise with one outlier included in the signal.

proposed [3] that the proper bandpass filter be found by maximizing the kurtosis of the envelope of the filtered signal. Kurtosis of the filtered signal has also been maximized for bandpass filter adjustment [11].

Though the kurtosis-based method has its merits, it is very sensitive to the outliers in the data. This may lead the parameter selection algorithm to the wrong frequency bands of the measured vibration and consequently result in poor de-noising performance. Such a drawback may be better illustrated using an example. Fig. 7 shows simulated white Gaussian noise consisting of 4000 sampling points. This dataset also contains one outlier, purposely added to the data. Clearly, a robust indicator of impulsiveness should not be susceptible to a small number of such outliers. Hence, to reflect the white Gaussian nature of the dataset, the smoothness index should be very close to 0.8455 and, as proven later, the kurtosis value for the envelope of the noise should be close to 0.2451 (the kurtosis of the noise itself should be roughly zero), if the two criteria were both insensitive to the outlier (out of 4000 points). The proof of the latter statement is similar to that of Eq. (35) and is presented in the following.

As proven earlier, for a signal with no impulsive features or with extremely low signal to noise ratio, H follows Rayleigh distribution. Substituting Eq. (34b) into Eq. (48), we have

$$\mathrm{Kurt}(h)=\frac{E\left\{H^4-4\sqrt{\frac{\pi}{2}}H^3+3\pi H^2-2\pi\sqrt{\frac{\pi}{2}}H+\frac{\pi^2}{4}\right\}}{E\left\{H^2-2\sqrt{\frac{\pi}{2}}H+\frac{\pi}{2}\right\}^2}-3. \tag{49}$$

According to Eq. (34b), $E\{H^2\}=2$, $E\{H^3\}=3\sqrt{\pi/2}$ and $E\{H^4\}=8$. Substituting these results into Eq. (49) yields:

$$\mathrm{Kurt}(h)=\frac{8-\frac{3\pi^2}{4}}{\left(2-\frac{\pi}{2}\right)^2}-3=0.2451.$$

The smoothness index calculated for the envelope signal (simulated noise plus a single outlier) was 0.8409, very close to the expected constant 0.8455. However, this was not the case for kurtosis which was found to be 20.49 for the simulated noise plus an outlier and 77.93 for the associated envelope, substantially higher than the anticipated levels, i.e., zero and 0.2451, respectively. We then recalculated the two indices for simulated noise with no outliers added. The associated smoothness index of the envelope was 0.8450 but kurtosis value reduced to -0.04 and 0.287, respectively, for the original noise and its envelope. This result clearly shows the robustness of the proposed smoothness index and the peculiar overreaction of kurtosis to merely a single outlier. It should be pointed out that similar high kurtosis value could be obtained from a highly impulsive dataset without any outliers. Therefore, the kurtosis value does not necessarily represent the impulsiveness of a signal since the high kurtosis value could mean either high impulsiveness or the existence of outliers. This will be illustrated later in the experimental results. In addition, a few outliers generally do not reflect the true

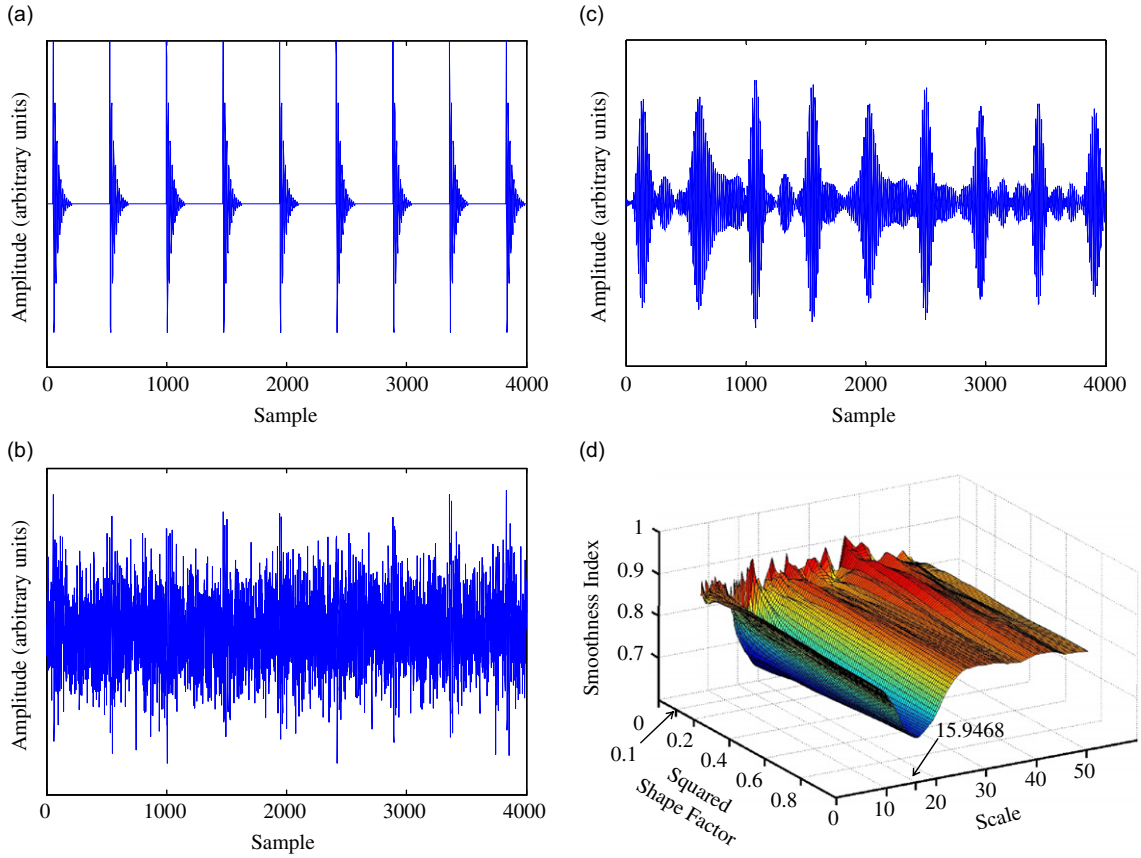

Fig. 8. (a) Simulated faulty bearing signal at 48,000 Hz sampling rate, 1739.4 rev/min rotational speed, fault characteristic frequency = 3.5 × rotational speed (Hz) and resonance frequency of 3000 Hz corresponding to scale 16; (b) simulated signal with noise added, the resulting noisy signal has a signal to noise ratio of −7; (c) de-noised signal using the scale and the shape factor corresponding to the minimum smoothness index; and (d) smoothness index for different s–σ^2 combinations.

machinery condition and are often caused by events of random nature, e.g., certain measuring flaws, or noise of other natures. Furthermore, as mentioned earlier in this section, the filtering algorithm guided by kurtosis maximization tends to settle in a frequency band which leads to a time-domain result containing a few larger data points. These large data points act like outliers and cause ambiguity.

To further compare the two criteria, we calculated the smoothness index and kurtosis values for the simulated faulty bearing vibrations presented in Fig. 8a. They were 0.0591 for the smoothness index, 19.3 and 9.91 for the kurtosis of the simulated signal and its envelope, respectively. Obviously, the small smoothness index value was consistent with the impulsive nature of the data. On the other hand, the kurtosis value for such a highly impulsive signal was even lower than that of the white Gaussian noise with only one outlier shown in Fig. 7. This once again leads us to believe that the proposed smoothness index is a more robust criterion for fault detection and wavelet parameter selection as compared with kurtosis.

We also wish to point out that the above comparisons may not be considered conclusive. More comprehensive studies may be needed to examine the merits of the two approaches.

3.7. Performance assessment of the proposed index using the simulated data

3.7.1. Simulated signal and noise

A considerable amount of noise is added to the simulated signal shown in Fig. 8a. The resulting signal is shown in Fig. 8b. This signal is then wavelet transformed using Gabor wavelet at the scales corresponding to the range $0.04F_{Nyq} \leqslant f \leqslant 0.25F_{Nyq}$. F_{Nyq} is one half of the sampling frequency (48,000 Hz). Hence the search

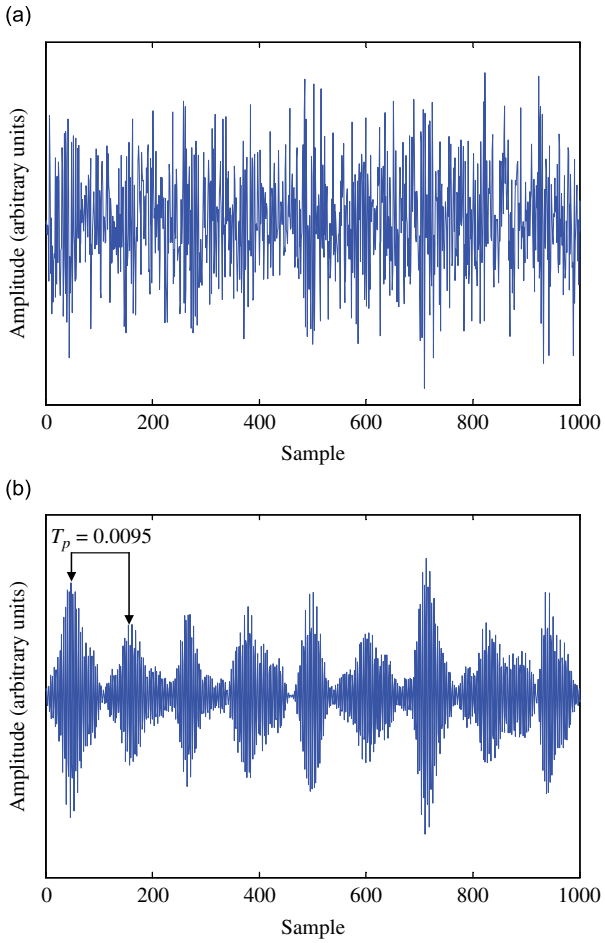

Fig. 9. (a) Measured vibration from a bearing with outer race fault [22] with added noise and (b) de-noised version of the same signal using the parameters corresponding to minimum smoothness index.

range is $(0.04 \times 48{,}000/2,\ 0.25 \times 48{,}000/2) = (960,\ 5960)$ Hz. This range is discretized with the step of 50 Hz, which leads to 100 frequencies, i.e., $f = (960, 1010, 1060, \ldots, 3010, \ldots, 5960\ \text{Hz})$. The associated 100 scales are $s = (50.0000, 47.5248, 45.2830, \ldots, 15.9468, \ldots, 8.0537)$. The squared shape factor is also considered over the range $0.01 \leqslant \sigma^2 \leqslant 1$ with the step of 0.01. The real part of the wavelet coefficients corresponding to the minimum smoothness index achieved for scale $s = 15.9468$ and $\sigma^2 = 0.1$ is presented as the de-noised version of the simulated signal in Fig. 8c. As expected, the best scale corresponds to the frequency of 3010 Hz which is very close to the ringing frequency of the simulated fault impulses (3000 Hz). Fig. 8d illustrates the smoothness index found for different s–σ^2 combinations.

3.7.2. *Real bearing signal with additional simulated noise*

To assess the performance of the proposed method, we took real bearing signals sampled at 12,000 Hz (with faults on the outer and inner races) from Ref. [22] and increased the noise levels. The signal–noise mixtures are shown in Figs. 9a and 10a. Their de-noised results are plotted in Figs. 9b and 10b, respectively. The fault generated impulses can be clearly identified in the de-noised signal and the time interval between two consecutive impulses matches with the characteristic fault frequencies of 107.30 and 162.18 Hz for the outer and inner race faults, respectively.

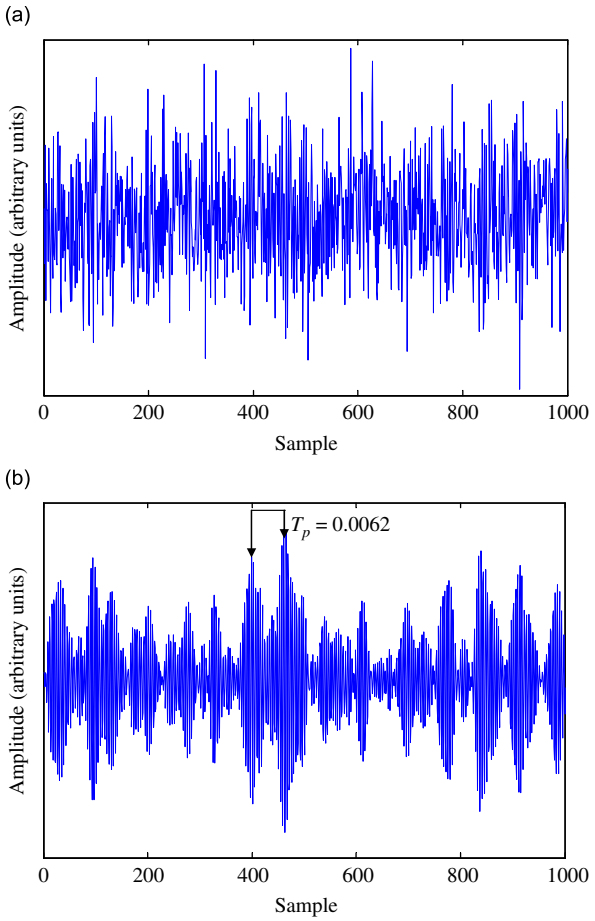

Fig. 10. (a) Measured vibration from a bearing with inner race fault [22] with added noise and (b) de-noised version of the same signal using the parameters corresponding to minimum smoothness index.

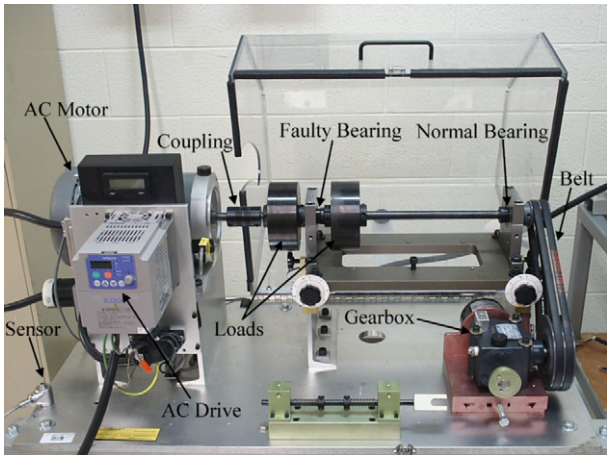

Fig. 11. Experimental setup.

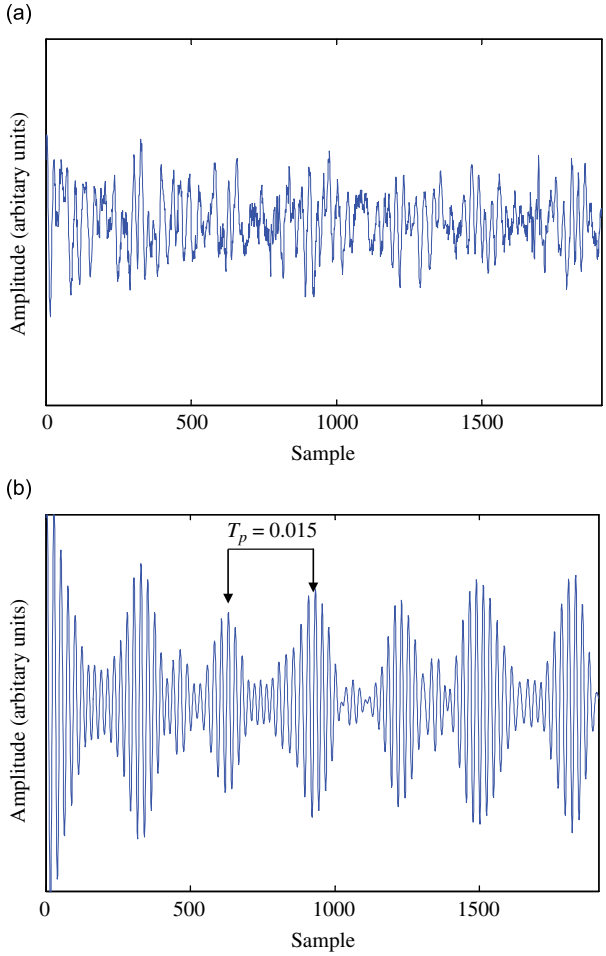

Fig. 12. (a) Measured vibration signal and (b) de-noising result.

4. Experimental evaluation

The proposed method was further evaluated using the vibration data measured in our lab. The experiment was carried out using a SpectraQuest Machinery Fault Simulator (MFK-PK5M) as shown in Fig. 11. Two bearings (type ER10K) were used to support a two-mass rotor test kit. The two-mass rotors were well balanced and are 2″ thick, 4″ in diameter and 11.1 lb each. The rotors were fitted into a 5/8″ steel shaft to provide the radial load. The simulator was driven by a 3-hp AC motor with a Hitachi controller (SJ200-022NFU). The shaft speed was set at 1326 rev/min (22.1 Hz). The left bearing has a pre-seeded single fault on the outer race with a characteristic frequency of 67 Hz ($=3.052f_r$). A Montronix model VS100-100 accelerometer with 100 mV/g sensitivity and 1–12 kHz sensitivity range was used to measure the vibration signal. The signal was fed to an NI AT-MIO-16DE-10 DAQ card and then collected through LabVIEW. The signal processing was done using MATLAB on a Pentium® 4/2.52 GHz PC.

Considering the fact that it may not be realistic to install an accelerometer in the near vicinity of a bearing in an industrial setting, we mounted the accelerometer at a spot on the simulator base that is away from the faulty bearing (Fig. 11). In addition to the rotor load, we also connected a gearbox to the driving shaft using a belt shown on the figure to generate more interference. The vibration data were acquired at 20,000 samples/s. A portion of the measured data is plotted in Fig. 12a and is de-noised using the proposed method. To find the

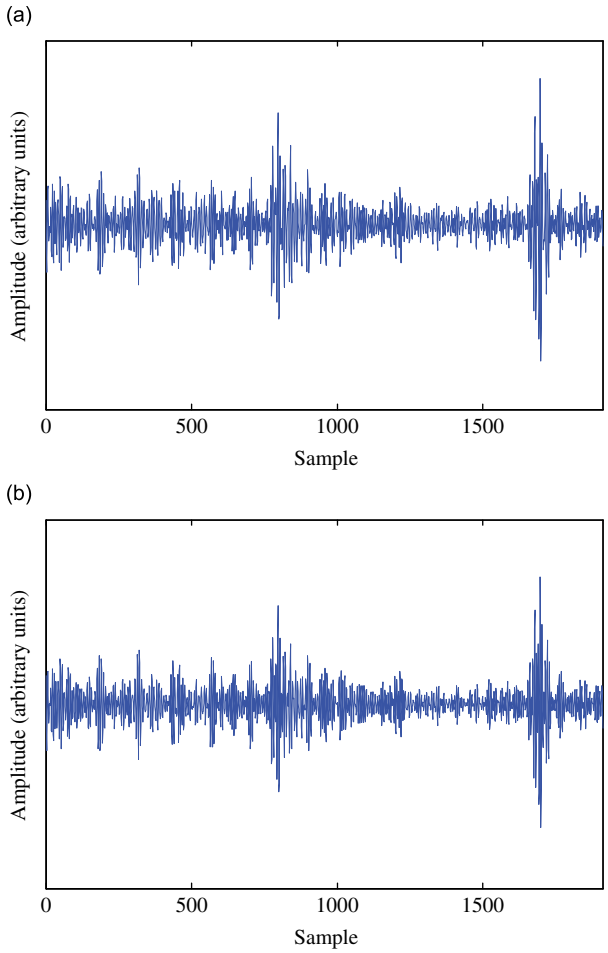

Fig. 13. De-noising result based on maximizing the kurtosis of: (a) bandpass filtered signal and (b) envelope of the bandpass filtered signal.

best $(s-\sigma^2)$ combination that yields the minimum smoothness index, we searched the scale range corresponding to the frequency interval [200, 6000] Hz and the σ^2 range from 0.01 to 1. The search step lengths were 50 Hz for frequency and 0.01 for σ^2. The best $(s-\sigma^2)$ combination was (23.53, 0.24) corresponding to the minimum smoothness index of 0.7467. The de-noising result associated with this $(s-\sigma^2)$ combination is displayed in Fig. 12b. As shown in the figure, the time interval between two consecutive impulses is about 0.015 s which precisely reflects the fault characteristic frequency 67 Hz as mentioned above.

For comparison, the scale and shape factor were also selected by maximizing the kurtosis value calculated for both the filtered signal and its envelope. The $(s-\sigma^2)$ combinations associated with the maximum kurtosis of the filtered signal (6.67) and that of the envelope (8.8151) were (6.55, 0.96) and (6.55, 0.5), respectively. The de-noising results corresponding to the two $(s-\sigma^2)$ combinations are shown in Figs. 13a and b. While some impulsive features can be seen from both results, no periodic component reflecting the fault characteristic frequency of the bearing could be detected. This result shows that the kurtosis maximization criterion has led the selection process to a wrong frequency band of the vibration signal.

5. Conclusion

A smoothness index-guided search approach has been proposed to find the best combination of wavelet scale and shape factor and hence the best Gabor daughter wavelet for de-noising impulsive signals. In this

study, we have proven that the smoothness index value approaches 0.8455 for data without impulse components or with extremely low signal to noise ratio. We have also shown that the smoothness index value decreases if a proper combination of wavelet scale and shape factor is selected. The proposed method has been tested using both simulated and experimental data. All our tests have shown that the proposed method can provide very reliable de-noising results.

Acknowledgments

We wish to thank Professor K.A. Laparo of Case Western Reserve University for his kind permission to use their bearing data. The major part of this study was supported by Natural Science and Engineering Research Council of Canada. This work was also supported recently by Ontario Centers of Excellence in partnership with GasTOPS, Ltd. Their support is greatly appreciated.

References

[1] P.D. McFadden, J.D. Smith, Model for the vibration produced by a single point defect in a rolling element bearing, *Journal of Sound and Vibration* 96 (1) (1984) 69–82.
[2] P.D. McFadden, J.D. Smith, The vibration produced by multiple point defects in a rolling element bearing, *Journal of Sound and Vibration* 98 (2) (1985) 263–273.
[3] J. Antoni, R.B. Randal, The spectral kurtosis: application to the vibratory surveillance and diagnosis of rotating machines, *Mechanical Systems and Signal Processing* 20 (2) (2006) 308–331.
[4] P.D. McFadden, J.D. Smith, Vibration monitoring of rolling element bearing by the high-frequency resonance technique—a review, *Tribology International* 17 (1984) 3–10.
[5] D.L. Donoho, De-noising by soft-thresholding, *IEEE Transactions on Information Theory* 41 (3) (1995) 613–627.
[6] D.L. Donoho, I.M. Johnstone, Ideal spatial adaptation via wavelet shrinkage, *Biometrika* 81 (1994) 425–455.
[7] J. Lin, L. Qu, Feature extraction based on Morlet wavelet and its application in mechanical fault diagnosis, *Journal of Sound and Vibration* 234 (1) (2000) 135–148.
[8] J. Lin, M.J. Zuo, K.R. Fyfe, Mechanical fault detection based on the wavelet de-noising technique, *Journal of Vibration and Acoustics* 126 (1) (2004) 9–16.
[9] H. Qiu, J. Lee, J. Lin, G. Yu, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognosis, *Journal of Sound and Vibration* 289 (4–5) (2006) 1066–1090.
[10] H. Qiu, J. Lee, J. Lin, G. Yu, Robust performance degradation assessment methods for enhanced rolling element bearing prognostics, *Advanced Engineering Informatics* 17 (3–4) (2003) 127–140.
[11] J. Lin, M.J. Zuo, Gearbox fault diagnosis using adaptive wavelet filter, *Mechanical Systems and Signal Processing* 17 (6) (2003) 1259–1269.
[12] N.G. Nikolaou, I.A. Antoniadis, Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets, *Mechanical Systems and Signal Processing* 16 (4) (2002) 677–694.
[13] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1998.
[14] S.O. Rice, Mathematical analysis of random noise, *Bell System Technical Journal* 24 (1945) 46–156.
[15] R.M. Rao, A.S. Bopardikar, *Wavelet Transforms: Introduction to Theory and Application*, Addison-Wesley, Reading, MA, 1998 (Chapter 1).
[16] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series and Products*, sixth ed., Academic Press, San Diego, 2000.
[17] M. Abramowitz, A.I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover Publications, New York, 1972.
[18] S.M. Kay, The effect of noise on the autoregressive spectral estimator, *IEEE Transactions on Acoustics, Speech and Signal Processing* 27 (5) (1979) 478–485.
[19] A. Papoulis, S.U. Pillai, *Probability, Random Variables and Stochastic Processes*, fourth ed., McGraw-Hill, New York, 2002.
[20] A. Lapidoth, S.M. Moser, Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels, *IEEE Transactions on Information Theory* 49 (10) (2003) 2426–2467.
[21] A. Hyvarinen, J. Karhunen, E. Oja, *Independent component analysis*, Wiley, New York, 2001.
[22] Case Western Reserve University bearing data center, ⟨http://www.eecs.cwru.edu/laboratory/bearing⟩, 2006.