Journal of Magnetic Resonance **139,** 189–204 (1999)
Article ID jmre.1999.1782, available online at http://www.idealibrary.com on IDEAL®

Accurate Quantification of ^{1}H Spectra: From Finite Impulse Response Filter Design for Solvent Suppression to Parameter Estimation

Tomas Sundin,* Leentje Vanhamme,† Paul Van Hecke,‡ Ioannis Dologlou,§ and Sabine Van Huffel†

**Department of Systems and Control, Uppsala University, P.O. Box 27, SE-75103 Uppsala, Sweden; †Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Kard. Mercierlaan 94, 3001 Leuven, Belgium; ‡Biomedical NMR Unit, Katholieke Universiteit Leuven, Gasthuisberg, 3000 Leuven, Belgium; and §Institute for Language and Speech Processing (ILSP), 6 Epidavrou str., 151 25 Athens Greece*

Received November 24, 1998; revised April 13, 1999

A scheme for accurate quantification of ^{1}H spectra is presented. The method uses maximum-phase finite impulse response (FIR) filters for solvent suppression and an iterative nonlinear least-squares (NLLS) algorithm for parameter estimation. The estimation algorithm takes the filter influence on the metabolites of interest into account and can thereby correctly incorporate a large variety of prior knowledge into the estimation phase. The FIR filter is designed in such a way that no distortion of the important initial samples is introduced. The FIR filter method is compared numerically with the HSVD method for water signal removal in a number of examples. The results show that the FIR method, using an automatic filter design scheme, slightly outperforms the HSVD method in most cases. The good performance and ease of use of the FIR filter method combined with its low computational complexity motivate the use of the proposed method.

Key Words: **magnetic resonance spectroscopy data quantification; finite impulse response filter; solvent suppression; AMARES; nonlinear least squares.**

INTRODUCTION

Accurate and efficient quantification of magnetic resonance spectroscopy (MRS) signals is the essential step prior to the conversion of the estimated signal parameters into biochemical quantities (concentration, pH). MRS signals are characterized by a low signal-to-noise ratio (SNR) and simple signal processing techniques are in general not adequate. Therefore a number of more advanced techniques based on a time domain model function have been developed. The function often used to model the N measured data points is the sum of exponentially damped complex sinusoids,

$$y(n) = \hat{y}(n) + w(n) = \sum_{k=1}^{K} a_k e^{j\phi_k} e^{(-\alpha_k + j2\pi f_k)n\Delta t} + w(n), \quad n = 0, \ldots, N-1, \quad [1]$$

where $j = \sqrt{-1}$, a_k is the amplitude, ϕ_k the phase, α_k the damping, and f_k the frequency of the kth sinusoid ($k = 1, \ldots, K$); Δt is the sampling interval; and w is complex white noise with a circular Gaussian distribution. The caret on y indicates that this quantity represents the model function rather than the actual measurements. The time domain estimation methods can be divided into two classes. On the one hand, there are the so-called black-box methods. Among this class of methods are techniques based on Kumaresan and Tufts' linear prediction (LP) method (*1*) combined with the singular value decomposition (SVD) (*2*). Kung's state-space approach (*3*) combined with SVD and least squares (LS) [called HSVD (*4*)] is a more efficient and more accurate alternative to the LP method as it circumvents the polynomial rooting and root selection. Other variants of the state-space algorithms have been recently proposed in which some forms of prior knowledge can be incorporated (*5–7*), but the limitations to the imposition of prior knowledge about model function parameters are inherent to this type of method. An alternative method which also avoids polynomial rooting is the matrix pencil method as described in (*8*) and references therein. On the other hand, interactive methods exist that are iterative, require user involvement, and allow inclusion of prior knowledge. The algorithms minimize the difference between the nonlinear model function and the data. This approach leads to maximum likelihood (ML) parameter estimates if the underlying assumptions concerning the model function and noise distribution are satisfied. VARPRO (*9*) and the more recent AMARES (*10*) are examples of this type of method.

In this paper we focus our attention on ^{1}H spectra. In the absence of water suppression techniques the signal contribution of the water can have a magnitude that is 10^3 to 10^4 larger than the magnitude of the metabolites of interest, which lie on the broad "tails" of the water resonance. In practice carefully designed measurement sequences are used to suppress the water signal prior to data acquisition. It is, however, impossible to completely eliminate the water signal from *in vivo* signals without affecting the metabolites of interest over a relatively wide frequency range. Therefore a water peak remains present in the spectra that cannot be described by an analytical function, mainly because of magnetic field inhomogeneity and lineshape distortions caused by suppression techniques. Since no model function is available for the water signal, a method,

1090-7807/99 $30.00

e.g., like AMARES, cannot be used without first removing the disturbing peak. Note that the problem is also present if one uses a frequency domain quantification method. In that case care has to be taken of the tilted baseline caused by the presence of the water resonance. As a consequence, irrespective of the method used for the actual quantification of these proton spectra, a preprocessing step is necessary to remove the unwanted water contribution. It is obvious that this preprocessing step should influence the final parameter estimates of the metabolites of interest as little as possible. Furthermore the procedure must be easy to use and have a low computational complexity since in applications such as spectroscopic imaging a large number of spectra have to be processed.

Many preprocessing techniques have already been developed. Some of these methods estimate the water signal part and then subtract it from the original signal. Various methods for obtaining the estimate of the water contribution have been proposed. In (*11*), a low-pass convolution [finite impulse response (FIR) filter] is used to obtain the water peak estimate by suppressing all the high-frequency contents (including the metabolite peaks) of the signal. Variants on the method proposed in (*11*) have been developed to alleviate distortions. An overview and discussion of these types of methods can be found in (*12*). Another example in the same context is the HSVD method described in (*13*) [or the similar TLS-based method used in (*14*)] where HSVD is used to model the water region by a sum of damped complex exponentials. A conceptually different approach is to suppress the water part of the signal by applying filters or convolution techniques directly to the signal. Examples using that approach can be found, e.g., in (*15, 16*). In (*17*) the baseline caused by the presence of the water peak is approximated and removed from the spectrum while the frequency region around the center of the water peak is replaced by random noise. In (*18*) a technique based on SVD is presented.

A general comment on these filter-related papers is the lack of a satisfactory discussion of the design of the proposed filter. Moreover, none of the published papers discuss the influence of the used techniques on the parameter estimates of the metabolites of interest. The methods presented in the literature are mostly evaluated based only on a visual inspection of the filtered signal. In this paper we take a different approach. We first explain how to design a maximum-phase FIR filter which is ideal in the sense that, contrary to the procedures discussed in (*12*) and references therein, the filtered signal does not include any distorted samples and the loss of signal energy is minimized. To automate the procedure and to increase the reproducibility of the method, we propose an automatic filter design procedure based on statistical arguments related to the noise prewhitening interpretation of the filter. In the parameter estimation phase we explicitly take the effect of the filter into account, resulting in parameter estimates that lie close to what theoretically can be expected. We also discuss the influence of the filter design parameters on the estimates of the relevant metabolites and we compare the new method with the often-used HSVD water removal method.

The paper is organized as follows. In the next section the FIR filter theory is presented and desirable properties of the filter are pointed out. The filter influence on the parameters is studied and the choice of the iterative nonlinear least-squares (NLLS) algorithm is motivated. The filter design aspects are covered and an automatic filter design scheme is proposed. In the HSVD section the HSVD method is briefly discussed and the water removal procedure described. The numerical examples section contains an extensive numerical study of the parameter estimation accuracies of the relevant metabolites with respect to the choice of both the filter and HSVD parameters as well as a comparison of the computational complexity of the methods. Finally, in the last section, we formulate the main conclusions.

QUANTIFICATION OF ^{1}H SPECTRA USING MAXIMUM-PHASE FIR FILTERS FOR SOLVENT SUPPRESSION

FIR Filter Theory

A FIR filter is defined by the convolution

$$y_f(n) = \sum_{m=0}^{M-1} h_m y(n-m), \quad n = 0, \ldots, N-1, \qquad [2]$$

where $\{h_m\}_{m=0,\ldots,M-1}$ are the constant (possibly complex) filter coefficients [see, e.g., (*19*)]. A problem with the definition in Eq. [2] is that the samples $y(n)$ for $n < 0$ are not available for filtering. This is normally solved by assuming that the signal is zero outside the time window [i.e., $y(n) = 0$ for $n < 0$] or assuming a cyclic signal [i.e., $y(n) = y(N + n)$ for $n < 0$]. Either of these assumptions will lead to a distortion of the first $M - 1$ samples of the filter output which therefore should be discarded. Let

$$x_k(n) = a_k e^{j\phi_k} e^{(-\alpha_k + j2\pi f_k)n\Delta t}, \quad k = 1, \ldots, K,$$

denote the individual exponentially damped complex sinusoids. Filtering the MRS signal model in Eq. [1] and discarding the first $M - 1$ samples yield

$$\begin{aligned}
\hat{y}_f(n) &= \sum_{m=0}^{M-1} h_m \hat{y}(n-m+M-1) \\
&= \sum_{m=0}^{M-1} h_m (x_1(n-m+M-1) \\
&\qquad + x_2(n-m+M-1) \\
&\qquad + \cdots + x_K(n-m+M-1)) \\
&= \bar{h} b_1 x_1(n) + \bar{h} b_2 x_2(n) + \cdots + \bar{h} b_K x_K(n) \\
&= \sum_{k=1}^{K} \bar{h} b_k x_k(n), \quad n = 0, \ldots, N-M, \qquad [3]
\end{aligned}$$

where

$$\bar{h} \triangleq (h_{M-1} \cdots h_0)$$

and

$$b_k \triangleq (1 e^{(-\alpha_k + j2\pi f_k)\Delta t} \cdots e^{(-\alpha_k + j2\pi f_k)(M-1)\Delta t})^T.$$

From Eq. [3] it is seen that the filtered NMR signal consists of the original damped sinusoids (same frequency and damping) altered by a complex scalar $\bar{h}b_k$. Thus the filter coefficients $\{h_m\}_{m=0,\ldots,M-1}$ can be chosen to suppress certain peaks (i.e., make $|\bar{h}b_k| \simeq 0$) while not suppressing others (i.e., make $|\bar{h}b_k| \simeq 1$). The filter (or convolution) methods are often believed to have difficulties in removing the frequency domain "tail" of peaks without affecting peaks lying on this tail. This misunderstanding comes from the well-established fact that time domain convolution is equal to frequency domain multiplication. It is, however, usually forgotten that this is not exactly true unless the signal is cyclic or identically zero outside the actual time window [see, e.g., (*12, 20*)]. The actual effect of the filtering procedure in the frequency domain can be examined by taking the Fourier transform (FT) of the filtered signal model in Eq. [3],

$$\mathcal{F}\{\hat{y}_f(n)\} \triangleq \hat{Y}_f(f) = \sum_{k=1}^{K} \bar{h}b_k X_k(f) \qquad [4]$$

where $X_k(f)$ is the FT of the individual damped sinusoid (so-called Lorentzian). It is thereby clear that the broad peaks representing the exponentially damped sinusoids including their broad "tails," can be suppressed individually [see also (*12*)]. The same result cannot be obtained by simple operations directly in the frequency domain.

A suitable filter can be proposed from analyzing properties of the filter frequency response $H(f)$ defined by

$$H(f) \triangleq \bar{h}\bar{g}(f), \quad f \in [-0.5, 0.5],$$

where f denotes the normalized frequency (sampling frequency = 1), and $\bar{g}(f)$ is the Fourier vector:

$$\bar{g}(f) \triangleq (1 e^{j2\pi f} \cdots e^{j2\pi f(M-1)})^T.$$

The frequency-dependent amplification of the filter is given by the magnitude response $|H(f)|$. Second, the frequency-dependent phase delay is equal to the phase response $\phi(f)$ defined as

$$\phi(f) = \tan^{-1}\left(\frac{\text{imag}(H(f))}{\text{real}(H(f))}\right),$$

where real$(\cdot)$ and imag$(\cdot)$ denote the real and imaginary parts of $(\cdot)$, respectively. Finally, the frequency-dependent time delay (in number of samples) is equivalent to the filter group delay $Gd(f)$:

$$Gd(f) = \frac{d\phi(f)}{df}.$$

A suitable filter type is found by first studying the properties of the so-called linear-phase filter [see, e.g., (*19*)]. This type of filter is characterized by its phase response which is a linear function of the frequency. The linear dependency implies that the filter has a group delay that is equal to a constant (i.e., independent of the frequency). The constant group delay equals a pure time delay of the filtered signal. The time delay is equal to half the filter length [i.e., $(M-1)/2$ samples]. This means that $(M-1)/2$ information-carrying samples are lost when the first $M-1$ samples of the filtered signal are discarded. It is desirable to use a filter with a time delay equal to the filter length (i.e., $M-1$ samples). Such a filter is in general not exactly realizable. Here we propose use of a so-called maximum-phase filter which has the largest possible group delay for a given magnitude response [see, e.g., (*19*)]. This kind of filter is characterized by a phase response that is a nonlinear function of the frequency, leading to a nonconstant group delay. A maximum-phase filter has by definition all its zeroes of the filter coefficient polynomial outside the unit circle. A linear-phase filter can therefore easily be transformed into a maximum-phase filter by reflecting the zeroes of the filter coefficient polynomial outside the unit circle. Note that this operation does not change the magnitude response of the filter.

The properties of the described filter types will be clarified by a numerical example displayed in Fig. 1. A linear-phase filter with filter order $M = 61$ and the corresponding maximum-phase filter are studied. In the top left, the magnitude responses of the filters (both magnitude responses are equal) are displayed. The phase responses of the filters are displayed in the top right. The linear-phase filter has a phase response that varies linearly with frequency while the maximum-phase filter has a phase response with a larger phase delay which is varying nonlinearly with frequency. The bottom left displays the group delay of the two filters. It can be seen that the linear-phase delays the signal exactly 30 samples for all frequencies. For the maximum-phase filter the time delay is close to 60 in the passband and varies slightly as a function of the frequency. Finally in the bottom right the filter coefficients (i.e., the impulse responses) are displayed. The figure clearly shows that the linear-phase filter coefficients are symmetric while the maximum-phase filter has most of its energy in the last coefficients. These are typical properties of the impulse responses of the respective filters [see, e.g., (*19*)].

In Fig. 2 the filters described above are applied to a simulated ^{1}H MRS signal. The exact nature of the signal is de-

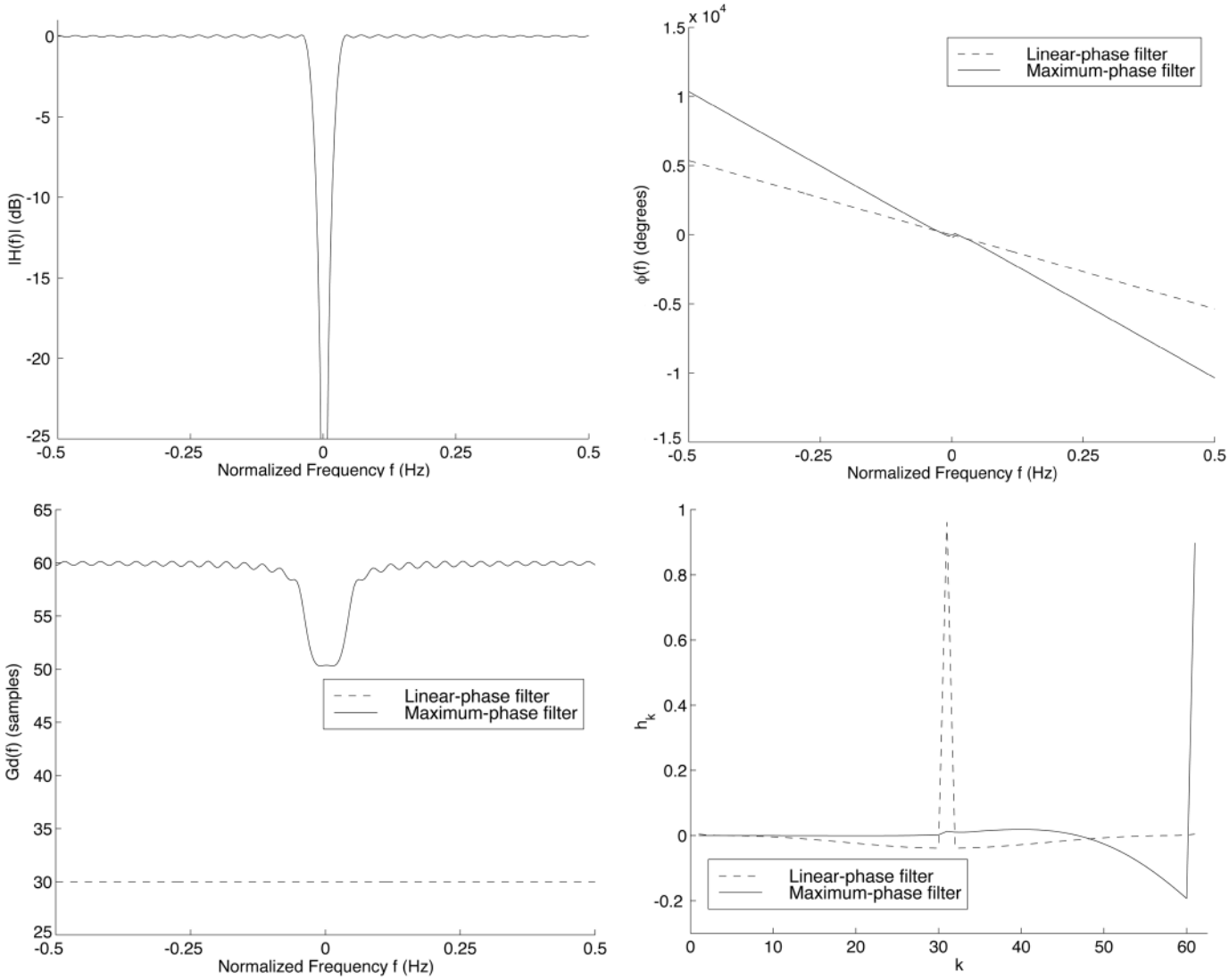

FIG. 1. Filter characteristics of linear-phase and maximum-phase FIR filters. Top left: Magnitude response. Top right: Phase response. Bottom left: Group delay. Bottom right: Filter coefficients (Impulse response).

scribed under Numerical Examples and for our purposes here it is sufficient to recognize the large water peak, located around 0 Hz, that we wish to suppress and the five metabolite peaks of interest at higher frequencies. In Fig. 2a the magnitude spectrum of the signal and a normalized magnitude response of the filter are displayed. Figure 2b displays the magnitude spectrum of the linear-phase filtered signal applied as in Eq. [2] without discarding the initial $M - 1$ distorted samples. This example is interesting since it shows that the spectrum of the filtered signal is nicely explained by the convolution theorem (i.e., the resulting spectrum is exactly given by the multiplication of the signal spectrum and the magnitude response of the filter). In Fig. 2c the magnitude spectrum of the filtered signal is displayed after applying the linear-phase filter and discarding 30 of the distorted samples in the beginning of the signal. The remaining distorted samples give rise to the broad hump in the water peak region. Note that the filtered spectrum can no longer be explained using the convolution theorem. In Fig. 2d the magnitude spectrum of the linear-phase-filtered signal with all distorted samples discarded is displayed. The spectrum shows that there is a perfect suppression of the water peak including the large tails as was predicted by the above discussion and the result in Eq. [3]. In Fig. 2e the result from applying the corresponding maximum-phase filter and discarding all the distorted samples is displayed. Also here the water peak is perfectly removed. The difference between the two last

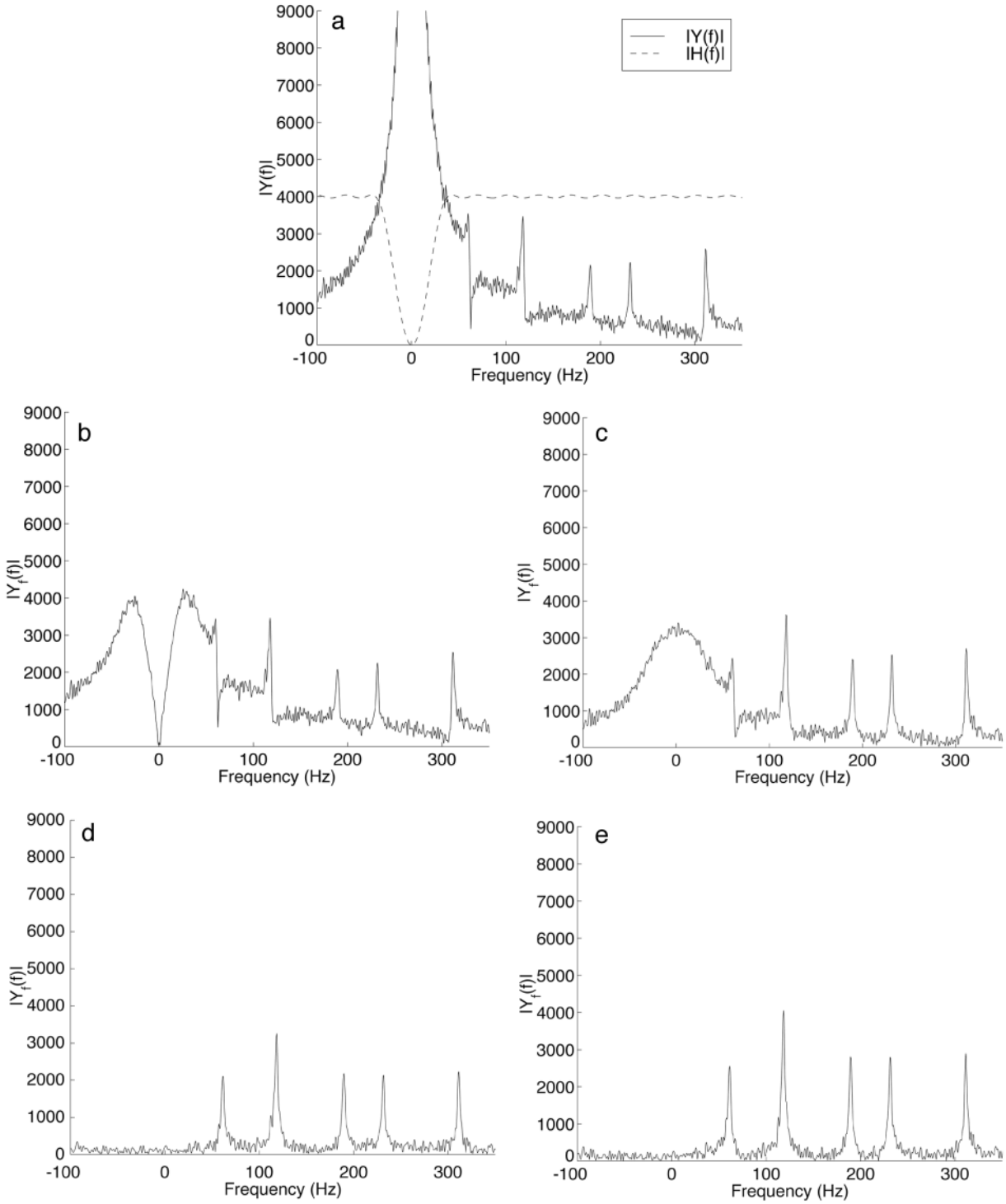

FIG. 2. Illustration of distortions introduced by applying a FIR filter. (a) Magnitude spectrum of ^{1}H MRS signal and normalized magnitude response of 61th-order FIR filter. (b) Magnitude spectrum of linear-phase filtered signal without discarding distorted samples. (c) Magnitude spectrum of linear-phase filtered signal discarding 30 distorted samples. (d) Magnitude spectrum of linear-phase filtered signal with all distorted samples discarded. (e) Magnitude spectrum of maximum-phase filtered signal with all distorted samples discarded.

spectra shows the influence of the different group delays of the filters. The 30-sample group delay of the linear-phase filter leads to a loss of 30 signal samples which contain an important part of the signal. The maximum-phase filter has a group delay close to 60 for frequencies in the passband and therefore there is practically no loss of the initial high-amplitude samples of

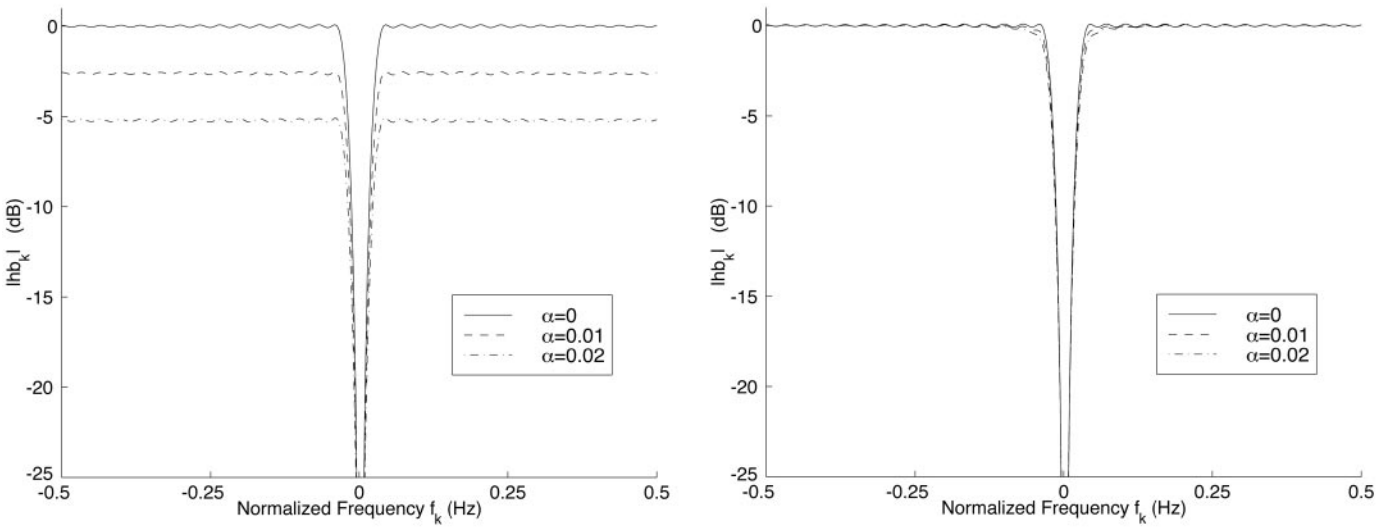

FIG. 3. Magnitude of $\bar{h}b_k$ term as a function of normalized frequency (sampling frequency = 1) for different damping coefficients α. Left: 61-tap, linear-phase FIR filter. Right: 61-tap, maximum-phase FIR filter.

the signal. This can be seen from the higher metabolite peak amplitudes in the maximum-phase-filtered spectrum compared with the linear-phase-filtered spectrum.

The importance of the large group delay of the maximum-phase filter can also be studied using the result derived in Eq. [3]. It is seen that the $\bar{h}b_k$ term gives a measure of the effect of the filter time delay when filtering damped sinusoids and discarding all the distorted samples. In Fig. 3 the loss of SNR as a function of the damping of the peaks is displayed for the linear-phase filter and the maximum-phase filter used above. The magnitude response of the linear-phase FIR filter is displayed on the left-hand side together with the values of the $|\bar{h}b_k|$ term as a function of the frequency for two different damping coefficients. The dampings and frequencies are given in normalized units (sampling frequency = 1). It is seen that for a damping of 0.01 (corresponding to 10 Hz for 1-kHz sampling frequency) the SNR loss is close to 3 dB for the linear-phase filter. The corresponding figures for the maximum-phase filter are found on the right-hand side and it is seen that there is practically no involuntary suppression of the metabolite peaks except for some effects close to the stop band region. The advantage of the use of maximum-phase filters can also be understood in the following way. Note that the b_k vector is equal to the Fourier vector $\bar{g}(f)$ for sinusoids without damping (i.e., $\alpha = 0$). The influence of the damping is that the magnitude response is weighted elementwise by the damped exponentials in b_k: $(1e^{-\alpha_k \Delta t} \cdots e^{-\alpha_k (M-1)\Delta t})^T$. The influence of the weighting will be minimized if the filter has most of its power in the last filter coefficients. This desirable property is as already stated typical of maximum-phase FIR filters.

In summary, the main reasons for using maximum-phase FIR filters to suppress peaks in certain frequency regions are the following. A FIR filter (in general) is powerful since the filter influences the different peaks individually. Thereby it is possible to suppress entire peaks, including the frequency domain "tails," while not influencing other peaks provided all distorted samples are discarded. The resulting signal is undistorted in the sense that the filtered signal is exactly described by the relation in Eq. [3]. Note that this property is not dependent on the type of filter used. The actual choice of filter type becomes important if you study the signal samples lost when the $M - 1$ distorted samples of the filtered signal are discarded. The high group delay of the maximum-phase filter was seen to minimize this loss. The SNR gain compared with using a linear-phase filter can be substantial.

As a final comment we want to point out that the proposed procedure is exactly equivalent to using a minimum-phase filter, filtering backward in time, and discarding the last M samples. This can be understood from the fact that the minimum-phase filter has all its zeros of the filter coefficient polynomial inside the unit circle and can be obtained by reversing the order of the maximum-phase filter coefficients [see, e.g., (*19*)]. The common use of minimum-phase filters and the existing design algorithm for such filters are the reasons for choosing to discuss minimum-phase filters under Filter Design.

Parameter Estimation

From the filtered signal, $y_f(n)$, parameter estimates can be obtained using any of the standard estimation methods. This is possible since the filtered damped sinusoids are still damped sinusoids with altered amplitude and phase as can be seen by Eq. [3]. Applying an estimation method to the filtered signal $y_f(n)$ results in estimates (called $\tilde{a}_k$, $\tilde{\alpha}_k$, $\tilde{\omega}_k$, and $\tilde{\phi}_k$, respectively) that are related to the parameters of interest in the following way:

$$a_k = \frac{\tilde{a}_k}{|\bar{h}b_k|},$$

$$\alpha_k = \tilde{\alpha}_k,$$

$$f_k = \tilde{f}_k,$$

$$\phi_k = \tilde{\phi}_k - \tan^{-1}\left(\frac{\text{imag}(\bar{h}b_k)}{\text{real}(\bar{h}b_k)}\right). \quad [5]$$

These corrections are easily made after the estimation phase by using the estimated values $\tilde{\alpha}_k$ and $\tilde{f}_k$ to find the b_k vectors. However, the estimation cannot be performed with standard methods if prior knowledge concerning the amplitudes or phases is to be taken into account. The filter influence on these parameters has to be considered directly in the estimation procedure to yield correct estimates. Here we propose use of a NLLS fit where the filtered signal $y_f(n)$ is fitted to the filtered model function derived in Eq. [3]:

$$\min_{a_k,\phi_k,\alpha_k,f_k} \sum_{n=0}^{N-M} |y_f(n) - \sum_{k=1}^{K} \bar{h}b_k x_k(n)|^2. \quad [6]$$

It is straightforward to correctly incorporate any kind of prior knowledge into the above criterion. The minimization in Eq. [6] has been numerically implemented for evaluation by modifying the AMARES algorithm. The new algorithm is referred to as AMARES$_f$ in the following and briefly described under Quantification of ^{1}H Spectra.

The use of the NLLS fit in Eq. [6] can be motivated by studying the ML estimator for the model in Eq. [1] based on a more general noise assumption. Let

$$R = E[\bar{w}^*\bar{w}] \quad [7]$$

denote the covariance matrix of the circular Gaussian distributed noise, where $\bar{w} = (w(0) \cdots w(N-1))$, E is the expected value operator, and $*$ denotes the Hermitian conjugate. The ML estimator is then given by the weighted NLLS (WNLLS) fit

$$\min_{a_k,\phi_k,\alpha_k,f_k} (\bar{y} - \sum_{k=1}^{K} \bar{x}_k)^* R^{-1} (\bar{y} - \sum_{k=1}^{K} \bar{x}_k), \quad [8]$$

where

$$\bar{y} = (y(0) \cdots y(N-1))^T$$

and

$$\bar{x}_k = (x_k(0) \cdots x_k(N-1))^T.$$

It is easy to check that Eq. [8] simplifies to

$$\min_{a_k,\phi_k,\alpha_k,f_k} \sum_{n=0}^{N-1} |y(n) - \sum_{k=1}^{K} x_k(n)|^2 \quad [9]$$

for white noise, with covariance matrix $R = \sigma^2 I$, where I denotes the identity matrix and σ^2 is the noise variance. The NLLS fit in Eq. [9] is the one used in algorithms such as VARPRO and AMARES.

The relation between the WNLLS fit in Eq. [8] and the proposed NLLS fit of the filtered signal in Eq. [6] is easily seen if the latter is slightly reformulated. Let

$$H = \begin{pmatrix} h_{M-1} & h_{M-2} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & h_{M-1} & h_{M-2} & \cdots & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_{M-1} & h_{M-2} & \cdots & h_0 \end{pmatrix} \quad [10]$$

denote the $(N - M + 1) \times N$ FIR filter matrix. Using Eq. [10], the following matrix multiplication replaces the convolution sum in Eq. [3]:

$$\bar{y}_f = H\bar{y}. \quad [11]$$

The NLLS fit of the filtered signal in Eq. [6] can then be written as

$$\min_{a_k,\phi_k,\alpha_k,f_k} (\bar{y} - \sum_{k=1}^{K} \bar{x}_k)^* H^* H (\bar{y} - \sum_{k=1}^{K} \bar{x}_k). \quad [12]$$

When we compare the NLLS fit in Eq. [12] with Eq. [8] we see that the modified NLLS fit is equal to the ML estimator if the FIR filter is an ideal prewhitening filter:

$$H^*H = R^{-1}. \quad [13]$$

The basic idea of the FIR filter technique is to regard the water signal as a part of the noise term and use the filter to whiten the total noise term prior to the estimation phase. It is, however, important to note that the ML interpretation of the filtered NLLS fit in Eq. [6] is valid only if the noise term is Gaussian distributed. This is not the case when a non-Gaussian signal (such as the water signal) is included in the noise term. Furthermore, it is impossible to completely decorrelate the noise term and the metabolite signals since Eq. [13] cannot be solved in general. The use of an approximate prewhitening FIR filter and the NLLS fit in Eq. [6] for quantification of the ^{1}H spectra is still expected to perform well and its good performance is illustrated under Numerical Examples.

Filter Design

The objective of this section is to give guidelines on how to find a well-matched prewhitening filter that minimizes the filter influence on the final parameter estimates. An automatic scheme for finding the appropriate values of the filter design parameters is proposed. This scheme reduces the required user interaction and ensures the reproducibility of the proposed method.

The design of maximum-phase filters is, as explained above, closely related to the design of minimum-phase filters. The design of the latter filter type is often based on the approach proposed in (*21*) where primarily a linear-phase FIR filter is designed and transformed to a minimum-phase filter by spectral factorization (i.e., reflecting all zeroes of the filter coefficient polynomial inside the unit circle). The maximum-phase filter is then found by reversing the order of the filter coefficients. The design of the initial linear-phase FIR filter is in our case done with a constrained least-squares fit [using the algorithms proposed (and provided) by (*22*)] in which the filter order M, normalized cutoff frequency f_c, stopband suppression sup, and passband ripple r have to be specified [see, e.g., (*19*)]. The linear-phase filter in Fig. 1 was designed using the described method and the following parameter values were used: $M = 60$, $f_c = 0.05$, sup $= 10^{-70/20}$ (corresponds -70 dB), and $r = 0.01$. The maximum-phase filter in the same figure is obtained by a spectral factorization of the linear-phase filter and reordering of the filter coefficients. The spectral factorization is done by rooting the filter coefficient polynomial and reflecting the zeroes inside the unit circle and thereafter reconstructing the impulse response of the filter. Using the above design algorithm transforms the problem of finding a suitable prewhitening filter into choosing appropriate values of the design parameters M, f_c, sup, and r. This choice can be automated using the following scheme based on some estimates obtained from the original signal. The filter design scheme is based on the assumption that the water peak consists mainly of one exponentially damped sinusoid,

$$y_w(n) = a_w e^{j\phi_w} e^{(-\alpha_w + j2\pi f_w)n\Delta t},$$

and that we easily can find estimates of a_w, α_w, f_w, and the noise standard deviation σ. It is important to note that this assumption is merely used to estimate the energy content and the frequency localization of the water peak and does not impose restrictions on the actual shape of the water signal.

To obtain the estimates, the frequency domain magnitude of the water peak, s_0, is calculated as

$$s_0 = \max|Y(\omega_n)|,$$

where $Y(\omega_n)$ denotes the discrete Fourier transform (DFT) of $y(n)$. The estimate of the water peak frequency is given by the corresponding frequency:

$$\tilde{f}_w = \frac{1}{2\pi} \arg\max_{\omega_n} |Y(\omega_n)|.$$

The width f_{wp} of the water peak (in Hz) at half-height $s_0/2$ gives the following estimate of the water peak damping:

$$\tilde{\alpha}_w = \pi f_{wp}.$$

The estimate of the water peak amplitude is given by

$$\tilde{a}_w = s_0 \hat{\alpha}_w.$$

Finally an estimate of the noise standard deviation can be found from the last samples of the original data sequence if the sinusoidal components have been sufficiently damped out:

$$\tilde{\sigma} = \sqrt{\frac{1}{P-1} \sum_{p=1}^{P} (y(N-p) - m_y)^* (y(N-p) - m_y)},$$

where m_y is the estimated mean,

$$m_y = \frac{1}{P} \sum_{p=1}^{P} y(N-p),$$

and P is chosen equal to a (small) number of samples containing mainly noise. Based on these estimates the following automatic procedure to design the FIR filter is proposed.

Filter Design Scheme

1. Calculate the estimates of the noise standard deviation σ, the damping α_w, frequency f_w, and amplitude a_w of the water peak as described above.
2. To suppress the water signal below the noise level, choose the suppression sup_0 to be equal to

$$\text{sup}_0 = \frac{\tilde{\sigma}}{2\tilde{a}_w}$$

and determine starting values for the filter order (e.g., $M = 50$) and passband ripple (e.g., $r = 5\%$).
3. Correct the suppression as a function of the damping of the water signal:

$$\text{sup} = \text{sup}_0 \left[\frac{M}{\sum_{m=0}^{M-1} e^{-\tilde{\alpha}_w m}} \right] = \text{sup}_0 \left[\frac{M(1 - e^{-\tilde{\alpha}_w M})}{1 - e^{-\tilde{\alpha}_w}} \right].$$

The reason for this correction is that the filter suppression is given for the magnitude response of the filter. The damped sinusoids in the water signal will be less suppressed due to the

weighting of the filter coefficients by the damping term (as explained above). The $M(1 - e^{-\hat{\alpha}_w M})/(1 - e^{-\hat{\alpha}_w})$ term approximately compensates for this loss.

4. Design a linear-phase filter (using the constrained LS algorithm above) with the lowest possible cutoff frequency f_c that fulfills the following constraints:

$$|\overline{hf}(0)| \leq \text{sup},$$

$$|\overline{hf}(\tilde{\alpha}_w)| \leq \text{sup}.$$

5. If this is not possible for

$$f_c < f_1,$$

where f_1 is the frequency for the metabolite of interest that lies closest to the water peak, increase the filter order M by 10 and restart from 3.

6. Shift the filter to be centered around the water peak [if the water peak is not located at zero frequency (i.e., $\tilde{f}_w \neq 0$)]:

$$\bar{h} = \bar{h} \odot (1e^{-j2\pi\tilde{f}_w} \cdots e^{-j2\pi\tilde{f}_w(M-1)}),$$

where $\odot$ denotes elementwise multiplication.

7. Transform the linear-phase FIR filter to a minimum-phase filter by spectral factorization and reorder the filter coefficients to obtain the final maximum-phase filter.

The only parameter the user has to specify using this scheme is the approximate frequency, f_1, of the peak of interest that lies closest to the water peak.

Quantification of ^{1}H Spectra

The quantification of ^{1}H spectra is performed by the AMARES$_f$ algorithm. The main difference between AMARES and AMARES$_f$ is the cost function which is minimized. Whereas AMARES solves Eq. [9], AMARES$_f$ minimizes

$$\min_{a_k,\phi_k,\alpha_k,f_k} \sum_{n=0}^{N-M} |\sum_{m=0}^{M-1} h_m y(n-m+M-1) - \sum_{k=1}^{K} \bar{h} b_k x_k(n)|^2$$

using the same NLLS solver as AMARES. The input to the AMARES$_f$ algorithm thus consists of the unfiltered signal $y(n)$ and the filter coefficients $\{h_m\}_{m=0,\ldots,M-1}$. The calculation of the Jacobian in AMARES$_f$ has been adapted to take into account the changed cost function. AMARES$_f$ allows the imposition of the same prior knowledge and parameter constraints as AMARES. The final quantification scheme can be summarized as follows.

Quantification Scheme

1. The noise standard deviation, water peak amplitude, frequency, and damping are estimated from the original signal as described above.
2. A frequency less or equal to the frequency of the metabolite peak closest to the water peak is defined by the user.
3. A suitable FIR filter is designed as described by the scheme above.
4. The signal is quantified (and filtered) using AMARES$_f$.

The computational complexity of the above quantification scheme is an important issue. The parameter estimation in the first step above introduces only a modest computational load. The filter design using the automatic scheme is performed by a constrained LS fit which has a computational complexity of $\mathcal{O}(M^3)$ floating-point opperations [where $\mathcal{O}(\cdot)$ denotes the order of magnitude] per iteration. The number of iterations is dependent on the filter specifications and cannot be known beforehand. However, the algorithm is known to be efficient in the sense that the required number of iterations is low. The spectral factorization can be solved by finding the roots of the filter polynomial of length M. The standard solution to solve this problem involves finding the eigenvalues of an $M \times M$ matrix leading to a computational complexity of $\mathcal{O}(M^3)$ floating-point operations. The filtering is included in the AMARES$_f$ algorithm and thereby the computational burden is increased by $\mathcal{O}(NM)$ floating-point operations compared with the standard AMARES algorithm. In summary the most computationally intensive parts in the quantification scheme are the filter design step and the spectral factorization. However, note that for applications such as spectroscopic imaging where a large number of spectra have to be processed and where the water contributions between spectra are similar it is sufficient to design the filter only once. This will decrease the computational burden significantly.

We want to point out that the newly developed method AMARES$_f$ and the automated filter design scheme will be incorporated into a future release of the MRUI software package (*23, 24*).

HSVD FOR SOLVENT SUPPRESSION

HSVD is a subspace-based parameter estimation method in which the noisy signal space is subdivided in a "signal" subspace and a "noise" subspace using a SVD of a Hankel data matrix. The "signal" subspace is found by truncating the SVD of this matrix to rank M, the number of exponentials that models the underlying signal. In general, HSVD provides a mathematical fit of the data by a sum of exponentially damped complex-valued sinusoids. HSVD can therefore be used to approximate the complicated features of the water resonance, including its large tails. The fitted water region is subsequently subtracted from the original signal.

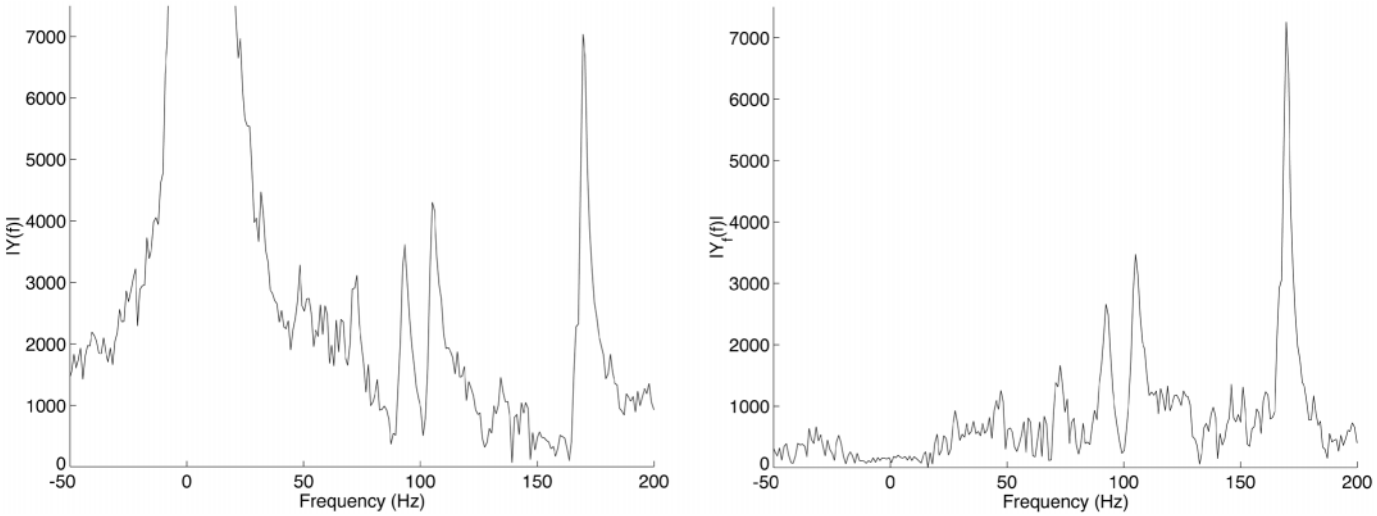

FIG. 4. Water-suppressed proton spectrum from a $2 \times 2 \times 2$-cm volume in the white matter of the brain of a healthy volunteer, acquired with the STEAM sequence (TR/TE/TM = 2000/20/30 ms) at 1.5 T. Note that no line broadening was applied to display the signals. Left: Magnitude spectrum of eddy current corrected signal. Right: Magnitude spectrum of the same signal after FIR filtering.

The algorithm is computationally intensive since it requires the computation of the SVD of the $(N/2 \times N/2)$ matrix demanding $\mathcal{O}(N^3)$ floating-point operations. HLSVD (*25*) and HLR (*26*) are fast versions of the HSVD algorithm in which the computation of a full SVD is circumvented. The gain in efficiency of these fast methods, however, decreases when the number of data points decreases and/or the model order increases (*26*). HLSVD and HLR offer approximately the same computational savings but HLR does not suffer from the problems associated with the Lanczos procedure (*26*). Since the filter design scheme and the fast versions of HSVD are all iterative methods, the differences in computation time are signal dependent and no exact statements can be made about the actual differences. However, the number of floating-point operations associated with the solvent suppression schemes using HSVD and HLR are evaluated for a number of examples under Numerical Examples.

In this paper we use the following scheme to process proton spectra.

1. The user specifies the model order M and a cutoff frequency f_r, which defines a so-called water region $[-f_r, f_r]$.
2. HSVD is used to model the original signal by a sum of M exponentially damped complex-valued sinusoids.
3. The peaks with frequencies belonging to this user-defined water region are used to reconstruct the water peak, after which the reconstructed water signal is subtracted from the original signal.
4. The residual signal is quantified with AMARES.

Note that the user is responsible for choosing the model order. As illustrated under Numerical Examples, this choice is more important than previously believed.

NUMERICAL EXAMPLES

In this section the proposed FIR filter-based suppression technique is evaluated to determine the sensitivity of the final parameter estimates with respect to the choice of filter design parameters. The method is compared with the often-used HSVD method described above. We investigate the water suppression abilities of both methods as a function of the noise level and varying frequency distances between the metabolite peaks and the water peak. A comparison of the computational complexity of the methods is also included.

First we visually illustrate the water suppression abilities of the FIR filter-based method by applying the method to an *in vivo* and an *in vitro* proton MRS signal. The *in vivo* signal is taken from a $2 \times 2 \times 2$-cm volume in the white matter of the brain of a healthy volunteer. The signal was acquired with the STEAM sequence (TR/TE/TM = 2000/20/30 ms) at 1.5 T (Vision, Siemens) and eddy current corrected using the method described in (*27*), which is based on earlier work described in (*28*). The *in vitro* proton MRS signal is acquired from a water solution of 100 mM creatine (CH_2 singlet, CH_3 singlet), 100 mM acetate (CH_3 singlet), 50 mM t-butyl alcohol ($3 \times CH_3$ singlet), and 10 mM TSP ($3 \times CH_3$ singlet). A single-voxel signal from a spherical phantom was acquired at 1.5 T (Vision, Siemens) using the STEAM sequence (TR/TE/TM = 20000/20/30 ms). The results are displayed in Figs. 4 and 5, respectively. To the left the magnitude spectra of the eddy current corrected signals are shown. To the right the magnitude spectra of the FIR-filtered signals are displayed. In the figures it can be seen that the water signals, including the tilted baselines distorting the nearby peaks, are perfectly removed by the filters obtained by the above filter design scheme.

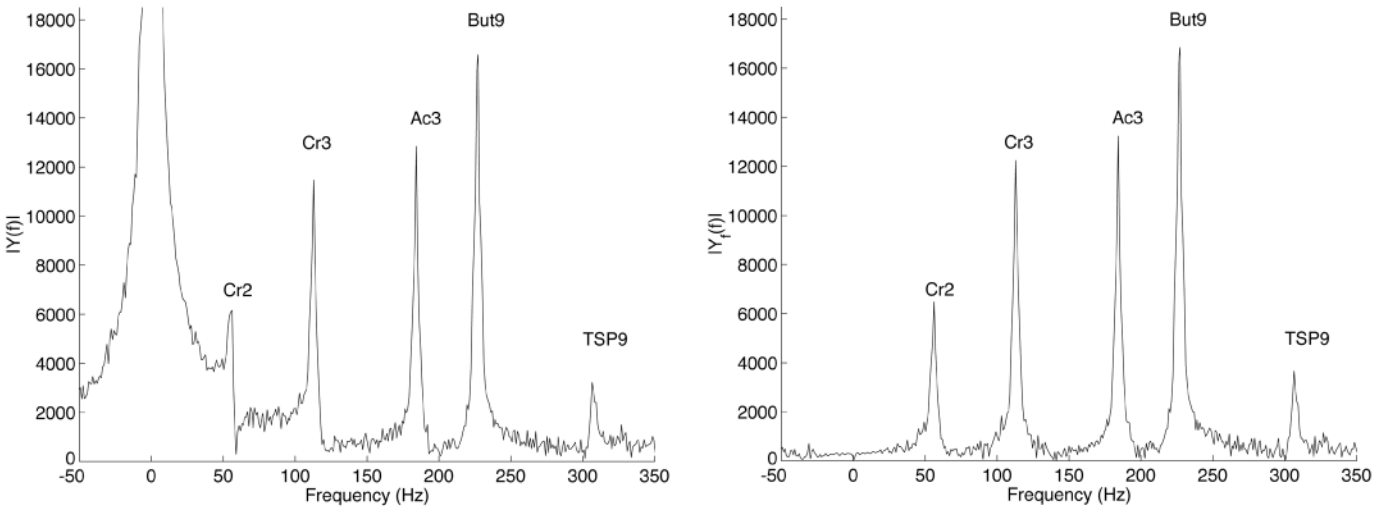

FIG. 5. Water-suppressed proton spectrum from a water solution of 100 mM creatine, 100 mM acetate, 50 mM t-butyl alcohol, and 10 mM TSP. A single-voxel signal from a spherical phantom was acquired at 1.5 T using the STEAM sequence (TR/TE/TM = 20000/20/30 ms). Note that no line broadening was applied to display the signals. Left: Magnitude spectrum of eddy current corrected signal. Right: Magnitude spectrum of the same signal after FIR filtering.

Usually, the phased, real part of the spectrum is displayed in the frequency domain. Here we choose to visualize the magnitude of the spectra. The reason is twofold. First, applying a FIR filter to a signal, results in slight phase and amplitude changes of the signal, which are taken into account in the estimation phase. To display the signal in the usual way, however, a phase correction of the filtered signal is needed. An approximate phase correction can easily be calculated from the filter phase response, but we prefer to display the magnitude spectrum to circumvent the user-dependent phasing of the signal. Second, by displaying the magnitude spectra, all information present in the signal is visualized.

Since experimental signals contain errors introduced by factors such as unknown lineshape, data acquisition errors, and eddy currents, all inevitably present in *in vivo* experiments, we use simulated signals to evaluate the performance of the proposed quantification scheme. The simulation signals are derived from the phantom signal described above in the following way. The acquired phantom signal was quantified with HSVD using a high model order ($M = 100$). The water signal was subsequently reconstructed with all the exponentially damped sinusoids with frequencies between -30 and 30 Hz and with amplitudes above the estimated noise level $\tilde{\sigma} \simeq 7.5$. The parameters of the seven peaks used to reconstruct the water resonance are found in Table 1. Five metabolite peaks were added as exponentially damped sinusoids with frequency, phase, and damping close to what was measured in the phantom experiment. The amplitudes of the peaks were chosen to be approximately equal to the estimated TSP9 amplitude and set equal for all peaks except for the two creatine peaks whose 2:3 ratio was kept. In Table 2 the exact parameters used in the simulation examples are given. The added complex noise is white and circular Gaussian distributed. The noise standard deviation was varied to simulate a number of SNRs. The SNR for each peak is measured in decibels (dB) and defined as

$$\text{SNR peak } k \triangleq 20 \log\left(\frac{a_k}{\sigma}\right).$$

In the following examples the quality of the amplitude estimates is measured as the relative root mean squared error (RRMSE) in percent,

$$\text{RRMSE peak } k \triangleq 100 \sqrt{\frac{1}{L}\sum_{l=1}^{L}\frac{(a_k - \tilde{a}_k^l)^2}{a_k^2}},$$

where L is the number of simulation runs and $\tilde{a}_k^l$ denotes the estimate of a_k obtained in simulation run l. The RRMSE is compared with the relative Cramér–Rao lower bound (CRB). The CRB is calculated from a model consisting of the five metabolite peaks without the water signal. The CRB indicates

TABLE 1
Estimated Water Signal Parameters Used in the Reconstruction of the Water Peak

f_{wk} (Hz)	α_{wk} (Hz)	ϕ_{wk} (deg)	a_{wk} (a.u.)
−8.48	5.10	−90.88	15.01
−5.25	8.28	−45.19	64.74
−2.16	10.51	−2.95	321.25
−0.18	12.45	179.97	1142.30
−0.17	4.24	−170.39	251.92
3.09	6.79	36.77	201.11
6.31	4.00	81.11	12.30

TABLE 2
Metabolite Parameters Used in the Simulated Signals

Peak k	f_k (Hz)	α_k (Hz)	ϕ_k (deg)	a_k (a.u.)
1	61	7	0	20
2	118	7	0	30
3	189	7	0	20
4	231	7	0	20
5	311	7	0	20

the best possible accuracy of an estimate for any unbiased estimator [see, e.g., (*29*)].

Simulation A: Influence of FIR Filter Parameters on Estimation Precision

The simulation signal described above is used here to examine the influence of the filter design procedure on the final parameter estimates as a function of the SNR. In Fig. 6 the RRMSE results obtained from 400 simulation runs are compared with the CRB for the amplitude estimates of peak 1 and peak 4. The estimation results for peaks 2 to 5 are practically equal and therefore only the results of peak 4 are shown. In addition to the automatic filter design procedure (cf. Filter Design) three different filters are chosen to investigate the filter influence on the parameter estimates. The filters have the same order, $M = 50$, and passband ripple, $r = 0.01$, but different values of suppression: 40, 60, and 80 dB with corresponding cutoff frequencies of 20, 30, and 40 Hz, respectively. Note that the influence of the filter order M is very modest. The order has to be chosen high enough to fulfill the requirements for the stop band suppression and cutoff frequency. Furthermore the passband ripple r is not critical either since this effect is taken into account in the estimation phase. As expected, the results for peak 1 are more sensitive to the filter design procedure than the results for peaks further away from the water peak. The results for the different filters show that a low suppression and cutoff frequency are desirable for low SNR, whereas a higher suppression is needed to sufficiently suppress the water peak for high SNR. These results are consistent with the theoretical reasoning that the filter should be designed to decorrelate the noise term (including the water signal) with the signals of interest as much as possible. A suitable level of the suppression must therefore be used not to deteriorate the final estimates. The automatic filter design procedure matches the filter characteristics with the water signal using the simple estimates described above. The resulting filter has a suitable suppression level and cutoff frequency, leading to good results for all SNRs in the example.

Simulation B: Influence of HSVD Parameters on Estimation Precision

For the HSVD method described above two parameters (M and f_r) have to be defined. The simulation signal is intended to examine the sensitivity of the final parameter estimates to the choice of these user parameters. In Fig. 7 the RRMSE results obtained from 400 simulation runs for the HSVD method are displayed. The HSVD method was applied using different model orders ($M = 10$, 12, 20, 30, and 40) and cutoff frequencies ($f_r = 15$, 35, and 55 Hz). Only the results for $f_r = 35$ Hz for each model order are shown since our simulations showed that the choice of cutoff frequency has a minor influence on the final parameter estimates. In our simulation example the water peak is exactly modeled by seven exponentially

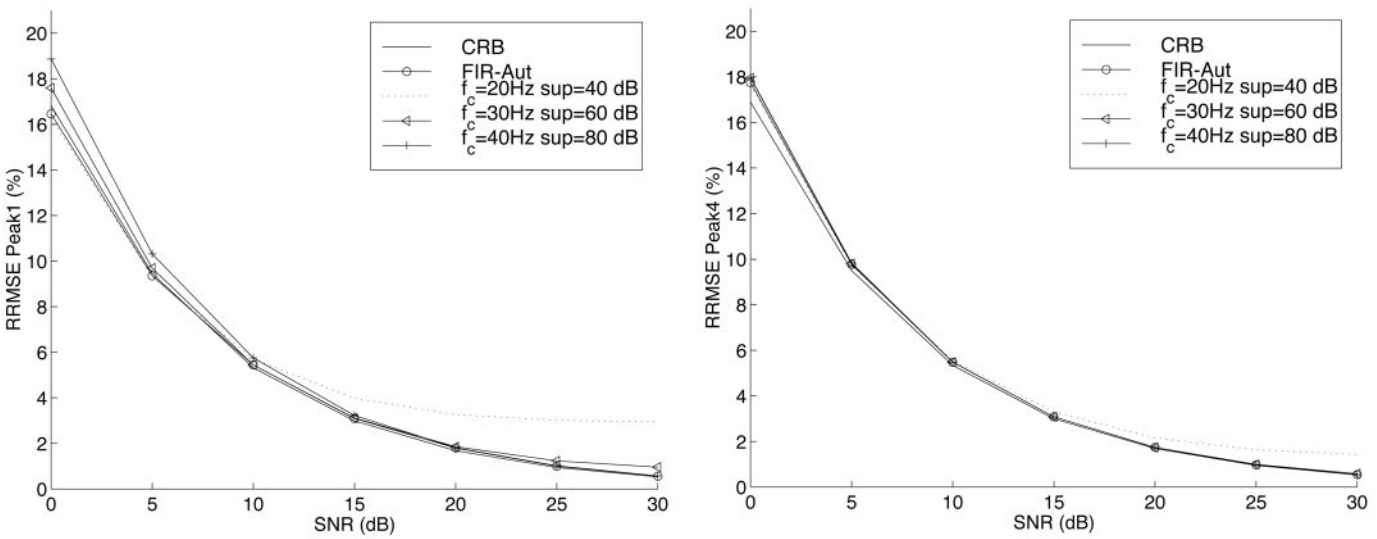

FIG. 6. CRB and RRMSE of amplitude estimates as a function of SNR obtained from 400 simulation runs using different FIR filters. Left: Peak 1. Right: Peak 4.

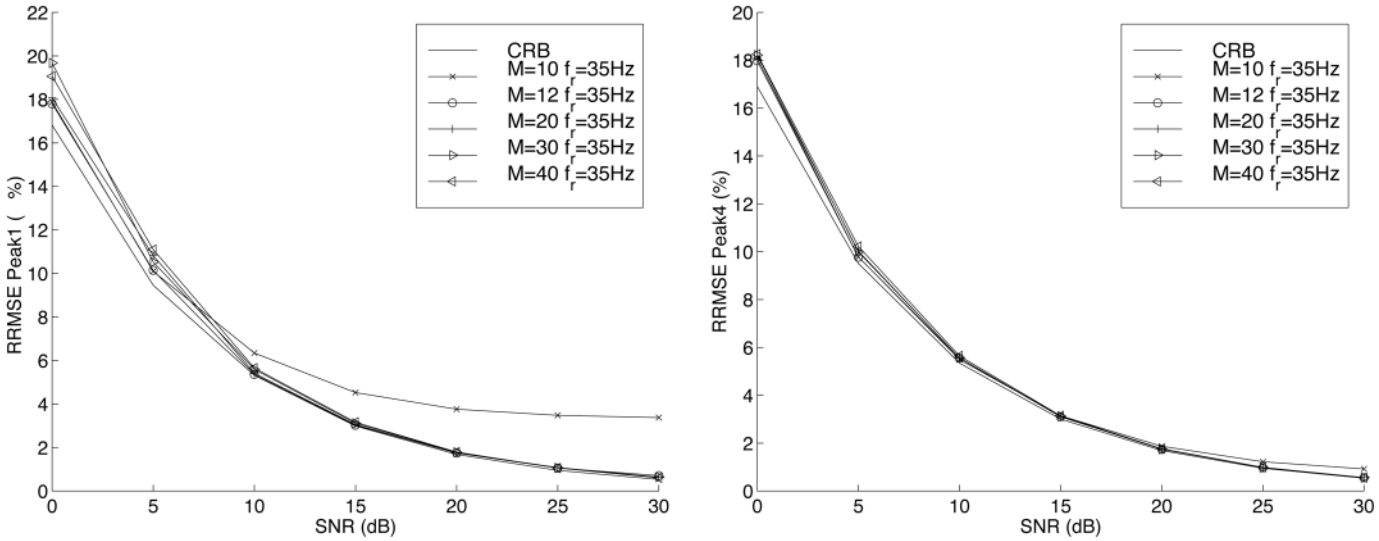

FIG. 7. CRB and RRMSE of amplitude estimates as a function of SNR obtained from 400 simulation runs using HSVD with different model orders. Left: Peak 1. Right: Peak 4.

damped sinusoids. This implies that the correct order of the signal subspace is equal to 12 (7 sinusoids used to model the water signal and 5 to model the metabolite peaks). The results show that the correct choice of the model order gives the best results for all SNRs. It is interesting to note the relatively large variations in the final estimates for peak 1 resulting from under- or overestimation of the order of the signal subspace. A low model order (i.e., $M = 10$) gives good results for low SNR while the results are very poor for high SNR. Overmodeling (i.e., $M > 12$), and on the other hand, gives poor estimates for low SNR. These results can be explained in a similar way as was done for the filter method above. Due to undermodeling the water signal is not completely removed which deteriorates the accuracy of the estimates at high SNR. This corresponds to the results obtained above using filters with too low suppression. On the other hand, overmodeling leads to modeling parts of the noise by damped sinusoids. Subtraction of these sinusoids from the original signal adds new features to the signal. These artificially introduced sinusoids can have a relatively high amplitude when the background noise is strong (low SNR) and influence the parameter estimates significantly. This can be compared with using an unnecessarily high-suppression FIR filter as described above. The simulation example is probably unrealistically simple for high SNR in which case the HSVD method has no problem in finding good estimates of the seven damped sinusoids used in the water signal reconstruction. However, the example shows that the model order selection for the HSVD method does influence the final parameter estimates and should be made with some care. We want to point out that methods exist that make an automatic choice of model order based on different information criteria and the values of the dominant singular values [see, e.g., (*8*) and references therein]. It is beyond the scope of this work to look further into this subject since the main purpose of this paper is to investigate the method based on FIR filters. Since no standard way of obtaining the "optimal" model order has been desribed in the literature, the HSVD method is applied with the correct model order ($M = 12$) in the following.

Simulation C: Comparison of FIR Filter Method and HSVD Method with Respect to Estimation Precision

In this simulation example the accuracy of the FIR filter method using the automatic filter design scheme is compared with the HSVD method using the correct model order ($M = 12$). The influence on the estimation accuracy as a function of the position of the metabolite peaks is examined by modifying the basic experiment described above. The frequency of peak 1 is set to different values (61, 51, 41, 31, and 21 Hz) while the frequencies of the other peaks are left unchanged. Four hundred simulation runs with four different noise levels are used to quantify the estimation errors. In Fig. 8 the RRMSE for peaks 1 and 4 are displayed as a function of the frequency for peak 1. The results for peak 1 show that the estimates are degrading when the metabolite peak is closer to the water peak. The reason is that the distortions introduced in the frequency region of the removed water signal by both methods correlate the noise term with nearby peaks. There is a difference between the distortions introduced by the two methods. The FIR filter uses a linear combination of the data samples to suppress the water peak. The distortions are in this case introduced by the inevitable suppression of the previously white noise in the water signal frequency region. The distortion introduced by the HSVD method is due to the estimation errors of the parameters of each sinusoid used to reconstruct the water signal. The

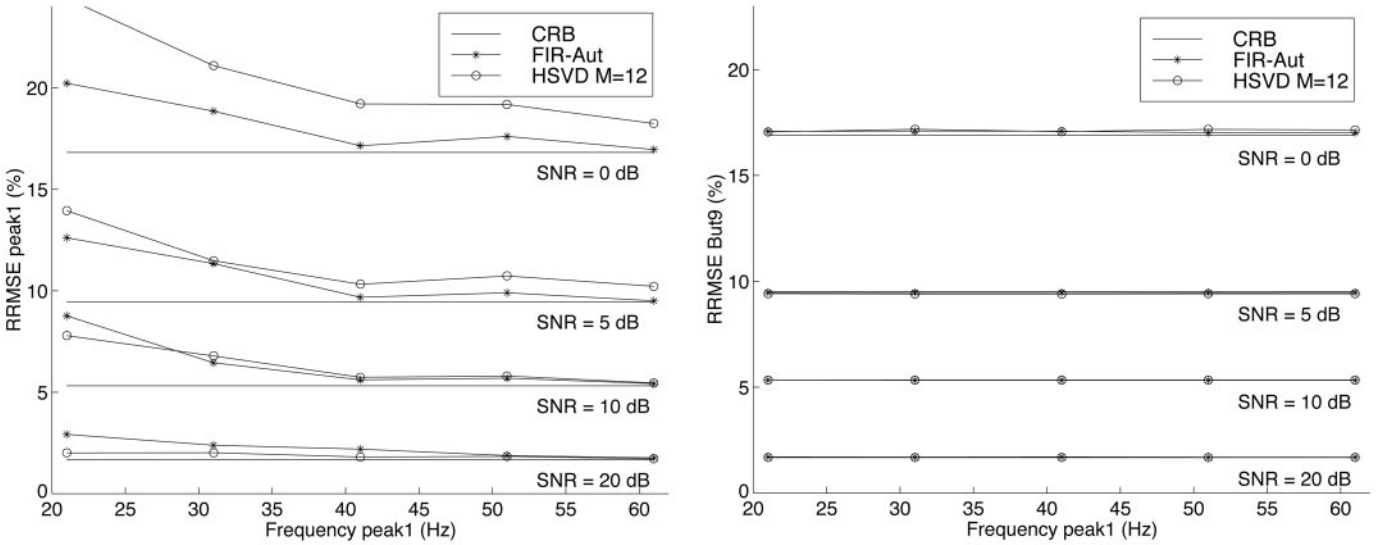

FIG. 8. CRB and RRMSE of amplitude estimates as a function of the frequency of peak 1. The results are obtained from 400 simulations runs for four different SNRs using the FIR filter method and the HSVD method. Left: Peak 1. Right: Peak 4.

subtraction of the water signal reconstruction will not lead to a perfect cancellation of the water signal and thereby introduce new features distorting the spectrum. These distortions are relatively small for high SNR where the HSVD procedure finds an excellent fit to the water signal but for realistic SNR levels the distortions have an effect on the final estimates as we can see in this example. In our example the HSVD method is outperformed by the FIR filter method for lower SNRs which indicates that the distortions introduced by filtering are less important than the distortions introduced by the HSVD method.

Simulation D: Comparison of FIR Filter, HSVD, and HLR Methods for Water Removal with Respect to Computational Complexity

In this simulation example the computational complexity of the FIR filter method is compared with that of the HSVD and HLR methods. The comparison is performed by calculating the number of flops (floating-point operations) used in matlab when applying the water suppression techniques to the simulation signal described above for $N = 512$, 1024, and 2048 data points, respectively.

The results for the FIR filter method are given in Table 3. The computational load is divided into three parts: the linear-phase filter design (constrained LS), the spectral factorization, and the filtering operation. Spectral factorization is seen to be the most computationally intensive step. The filter design procedure can be made more efficient by using alternative minimum-phase filter design techniques that circumvent the spectral factorization [see, e.g., (*30*) and references therein]. Note also the low computational load associated with the actual filtering procedure. This implies that in applications, such as spectroscopic imaging in which the filter design has to be performed only once, the computational complexity of the FIR filter method can be decreased by an order of magnitude. The results for the HLR and HSVD methods are given in Table 4. Comparing the results of these methods with the total computational load of the FIR filter method shows that HLR requires 12 times more computations ($N = 1024$) than the automatic FIR filter scheme (including the filter design steps), even for a

TABLE 3
Floating-Point Operations ($\times 10^6$) Required by the FIR Filter Method for Water Removal[a]

N	*M*	Const. LS	Spe. Fac.	Filtering	Total
512	30	0.13	0.41	0.13	0.66
	50	0.40	2.05	0.21	2.66
	70	0.76	5.00	0.29	6.06
	Aut	1.01	2.05	0.21	3.27
1024	30	0.13	0.41	0.25	0.79
	50	0.40	2.05	0.42	2.87
	70	0.76	5.00	0.58	6.35
	Aut	1.01	2.02	0.42	3.44
2048	30	0.13	0.41	0.51	1.05
	50	0.40	2.05	0.84	3.29
	70	0.76	5.01	1.16	6.93
	Aut	1.01	2.02	0.84	3.87

[a] The FIR filter computations are divided into the design of the linear-phase filter by constrained LS (Const. LS), spectral factorization (Spe. Fac.) and the filtering operation (Filtering). *M* is the filter length and *N* is the number of data points.

TABLE 4
Floating-Point Operations ($\times 10^6$) Required by the HLR and HSVD Methods for Water Removal[a]

N	M	HLR	HSVD
512	10	14.45	733.0
	12	21.21	734.2
	20	58.77	741.2
	30	137.2	755.6
1024	10	33.24	5697
	12	40.95	5700
	20	120.4	5113
	30	275.6	5740
2048	10	79.16	45086
	12	92.95	45091
	20	244.1	45116
	30	544.2	45167

[a] M is the model order and N is the number of data points.

correct choice of model order ($M = 12$). When the correct model order is unknown a sufficiently high model order should be used since undermodeling could lead to loss of accuracy (see Simulation B). The computational load of HLR is 80 times ($N = 1024$) higher than that of the automatic FIR filter scheme for a model order of $M = 30$. The computational load of HSVD is 1656 ($N = 1024$) times higher than for the FIR filter method, even for a low model order ($M = 12$).

In summary this simulation example shows that the FIR filter method, including the filter design step, is computationally always at least one order of magnitude more efficient than the fastest available algorithms based on the HSVD method for water removal. The difference is even higher if a high model order is required to reconstruct the water signal or when a large number of data points are used.

CONCLUSIONS

In this paper a scheme for accurate quantification of ^{1}H spectra is presented. The method uses maximum-phase FIR filters for solvent suppression and an iterative NLLS algorithm for parameter estimation (an adaption of AMARES). The estimation algorithm takes the filter influence on the metabolites of interest into account and thereby allows us to correctly incorporate a large variety of prior knowledge into the estimation phase. The theory describing the FIR filters is presented and it is shown that using a maximum-phase filter makes it possible to minimize the loss of signal energy when the distorted samples of the filtered MRS signal are discarded, and therefore the accuracy of the final parameter estimates is improved. The design of the FIR filter is thoroughly discussed and an automatic design procedure is proposed. The filtering procedure can be interpreted as a prewhitening operation, thereby giving a statistical motivation for the used estimation procedure. The computational complexity of the proposed scheme is low and at least one order of magnitude lower than for the HSVD-based water removal method even if a fast algorithm such as HLR is used. Note also that the FIR filter method can be used for solvent suppression by itself followed by parameter estimation by any existing estimation method. In this case the corrections of the filter influence have to be done afterward and it is not possible to incorporate prior knowledge of amplitude and/or phase relations.

The FIR filter method is compared numerically with the HSVD method for water peak removal in a number of simulation examples. The performance of the methods is examined as a function of the metabolite frequencies and the results show that the FIR filter method using the automatic design scheme slightly outperforms the HSVD method in most cases. The use of the automatic filter design scheme leads to small estimation errors and ensures the reproducibility of the results. The HSVD method is seen to be sensitive to the choice of the model order and we want to point out that there is a need to investigate automatic model order estimation techniques in this context. The good performance and ease of use of the FIR filter method combined with the low computational complexity motivate the use of the proposed method as an alternative to the often-used HSVD method.

ACKNOWLEDGMENTS

This work is supported by the Belgian Programme on Interuniversity Poles of Attraction (IUAP-4/2 and 24), initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture; by the EU program "Training and Mobility of Researchers" project (Contract ERBFMRXCT970160) entitled "Advanced Signal Processing for Medical Magnetic Resonance Imaging and Spectroscopy"; by a Concerted Research Action (GOA) project of the Flemish Community, entitled "Model-Based Information Processing Systems"; by FWO Grant G.0360.98; and by the Swedish Foundation for Strategic Research. L.V. is a Ph.D. student funded by the IWT (Flemish Institute for Support of Scientific-Technological Research in Industry). S.V.H. is a Research Associate with the FWO (Fund for Scientific Research–Flanders). This work was carried out at ESAT-SISTA/COSIC, Katholieke Universiteit Leuven, Heverlee, Belgium.

REFERENCES

1. R. Kumaresan and D. Tufts, Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise, *IEEE Trans. Accoustics Speech Signal Process.* **ASSP 30,** 833–840 (1982).
2. H. Barkhuysen, R. de Beer, W. Bovée, and D. van Ormondt, Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure, *J. Magn. Reson.* **61,** 465–481 (1985).
3. S. Kung, K. Arun, and D. B. Rao, State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem, *J. Opt. Soc. Am.* **73,** 1799–1811 (1983).
4. H. Barkhuysen, R. de Beer, and D. van Ormondt, Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals, *J. Magn. Reson.* **73,** 553–557 (1987).

5. H. Chen, S. Van Huffel, D. van Ormondt, and R. de Beer, Parameter estimation with prior knowledge of known signal poles for the quantification of NMR spectroscopy data in the time domain, *J. Magn. Reson. A* **119,** 225–234 (1996).
6. H. Chen, S. Van Huffel, A. van den Boom, and P. van den Bosch, Subspace-based parameter estimation of exponentially damped sinusoids using prior knowledge of frequency and phase, *Signal Process.* **59,** 129–136 (1997).
7. S. Van Huffel, Subspace-based exponential data modeling using prior knowledge, *in* "Proceedings of the IEEE Benelux Chapter Signal Processing Symposium (IEEEBSPS), Leuven, Belgium (1998)," pp. 211–214.
8. Yuang-Ya Lin, P. Hodgkinson, M. Ernst, and A. Pines, A novel detection-estimation scheme for noisy NMR signals: Applications to delayed acquisition data, *J. Magn. Reson.* **128,** 30–41 (1997).
9. J. van der Veen, R. de Beer, P. Luyten, and D. van Ormondt, Accurate quantification of *in vivo* ^{31}P NMR signals using the variable projection method and prior knowledge, *Magn. Res. Med.* **6,** 92–98 (1988).
10. L. Vanhamme, A. van den Boogaart, and S. Van Huffel, Improved method for accurate and efficient quantification of MRS data with use of prior knowledge, *J. Magn. Reson.* **129,** 35–43 (1997).
11. D. Marion, M. Ikura, and A. Bax, Improved solvent suppression to one- and two-dimensional NMR spectra by convolution of time-domain data, *J. Magn. Reson.* **84,** 425–430 (1989).
12. C. Craven and J. Waltho, The action of time-domain convolution filters for solvent suppression, *J. Magn. Reson.* **106,** 40–46 (1995).
13. A. van den Boogaart, D. van Ormondt, W. Pijnappel, R. de Beer, and M. Ala-Korpela, Removal of the water resonance from ^{1}H magnetic resonance spectra, *J. Magn. Reson.* **1,** 175–195 (1991).
14. J. Leclerc, Distortion-free suppression of the residual water peak in proton spectra by postprocessing, *J. Magn. Reson. Ser. B* **103,** 64–67 (1994).
15. K. Cross, Improved digital filtering technique for solvent suppression, *J. Magn. Reson.* **101,** 220–224 (1993).
16. G. Sobering, M. Kienlin, C. Moonen, P. van Zijl, and A. Bizzi, Post-acquisition reduction of water signals in proton spectroscopic imaging of the brain, *in* "Proceedings, SMRM," p. 771 (1991).
17. M. Deriche and X. Hu, Elimination of water signal by postprocessing, *J. Magn. Reson.* **101,** 229–232 (1992).
18. G. Zhu, D. Smith, and Y. Hua, Post-acquisition solvent suppression by singular-value decomposition, *J. Magn. Reson.* **124,** 286–289 (1997).
19. A. V. Oppenheim and R. W. Schafer, "Discrete-Time Signal Processing," Prentice–Hall, Englewood Cliffs, NJ (1989).
20. R. Pintelon, J. Schoukens, and G. Vandersteen, Frequency domain system identification using arbitrary signals, *IEEE Trans. Autom. Control* **AC-42,** 1717–1720 (1997).
21. O. Herrmann and H. Schüssler, Design of nonrecursive digital filters with minimum phase, *Electronics Lett.* **6,** 329–330 (1970).
22. I. Selesnick, M. Lang, and C. Burrus, Constrained least square design of FIR filters without specified transition bands, *IEEE Trans. Signal Process.* **44,** 1879–1892 (1996).
23. http://azur.univ-lyon1.fr/TMR/tmr.html.
24. http://www.mrui.uab.es/mrui/mruiHomePage.html.
25. W. Pijnappel, A. van den Boogaart, R. de Beer, and D. van Ormondt, SVD-based quantification of magnetic resonance signals, *J. Magn. Reson.* **97,** 122–134 (1992).
26. L. Vanhamme, R. Fierro, S. Van Huffel, and R. de Beer, Fast removal of residual water in proton spectra, *J. Magn. Reson.* **132,** 197–203 (1998).
27. U. Klose, In vivo proton spectroscopy in the presence of eddy currents, *Magn. Res. Med.* **14,** 26–36 (1990).
28. G. A. Morris, Compensation of instrumental imperfections by deconvolution using an internal reference signal, *J. Magn. Reson.* **80,** 547–552 (1988).
29. T. Söderström and P. Stoica, "System Identification," Prentice–Hall International, London (1989).
30. A. Groth and H. Göckler, Design of equiripple minimum phase FIR-filters, *in* "Proceedings, EUSIPCO 98, Rhodos, Greece (1998)," Vol. III, pp. 1897–1900.