Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 175 (2005) 79–91
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Spectral analysis of multichannel MRS data ☆

Niclas Sandgren [a,*], Petre Stoica [a], Frederick J. Frigo [b], Yngve Selén [a]

[a] Systems and Control Division, Department of Information Technology, Uppsala University, P.O. Box 337, SE-751 05 Uppsala, Sweden
[b] GE Healthcare Technologies, Magnetic Resonance Business Group, P.O. Box 414 WTE-893, Milwaukee, WI 53201, USA

Received 13 October 2004; revised 24 March 2005
Available online 27 April 2005

Abstract

The use of phased-array receive coils is a well-known technique to improve the image quality in magnetic resonance imaging studies of, e.g., the human brain. It is common to incorporate proton (^{1}H) magnetic resonance spectroscopy (MRS) experiments in these studies to quantify key metabolites in a region of interest. Detecting metabolites in vivo is often difficult, requiring extensive scans to achieve signal-to-noise ratios (SNR) that provide suitable diagnostic results. Combining the MR absorption spectra obtained from several receive coils is one possible approach to increase the SNR. Previous literature does not give a clear overview of the wide range of possible approaches that can be used to combine MRS data from multiple detector coils. In this paper, we consider the multicoil MRS approach and introduce several signal processing tools to address the problem from different nonparametric, semiparametric, and parametric perspectives, depending on the amount of available prior knowledge about the data. We present a numerical study of these tools using both simulated ^{1}H MRS data and experimental MRS data acquired from a 3T MR scanner.

Keywords: Multichannel magnetic resonance spectroscopy; Phased-array detectors; Damped sinusoidal model; Phantom experimental data

1. Introduction

Obtaining reliable information about the spectral components in MRS signals can be difficult in practical applications, due to strong interference from measurement noise. In magnetic resonance imaging (MRI) applications, low SNR results in poor image quality. The SNR may be improved by combining data obtained from multiple receive coils [1]. As a complement to multichannel (or multicoil) image reconstruction, it can be useful to perform multicoil MRS experiments, to provide accurate quantification of different metabolites in a region of interest. The multichannel approach, which combines MR absorption spectra from several receive coils, has been addressed before in, e.g. [1–6], but so far little emphasis has been put on the signal processing step required to combine data from different receive coils and to obtain a noise reduced spectral estimate of the MRS signal.

Consider the following multichannel magnetic resonance (MR) data equation:

$$y_j(t) = g_j s(t) + e_j(t), \quad t = 1, \ldots, N, \;\; j = 1, \ldots, m, \qquad (1)$$

where $\{y_j(t)\}$ are the m different observations provided by a multicoil MR system, $\{s(t)\}$ is the MR signal of interest, $\{g_j\}$ are unknown gains for each of the m coils, and $\{e_j(t)\}$ are m noise sequences. Note that all quantities in (1) are complex valued.

In this paper, we present the multicoil MRS problem from three different signal processing perspectives:

(a) *Nonparametric approach*. We consider the singular value decomposition (SVD) [7] for multicoil data as a

☆ This work was partly supported by the Swedish Science Council (VR).
* Corresponding author. Fax: +46 18 511925.
E-mail address: niclas.sandgren@it.uu.se (N. Sandgren).

1090-7807/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2005.03.019

method to estimate $\{s(t)\}$ in (1) without making any specific assumptions about the spectral content of the data or the noise sequences $\{e_j(t)\}$. In the following, this technique is referred to as the MC-NP-SVD (multichannel nonparametric singular value decomposition) method. We also present the so-called cross-relation method [6] and show that it is in fact equivalent to MC-NP-SVD.

(b) *Semiparametric approach*. We employ prior knowledge about the data and present a maximum likelihood method [8] which is dependent on the assumption that $\{e_j(t)\}$ are m independent, white, and circularly Gaussian noise sequences with zero mean. We call this new technique MC-SP-MLM (multichannel semiparametric maximum likelihood method). The goal of MC-SP-MLM is (as in the nonparametric case) to compute an accurate estimate of $\{s(t)\}$.

(c) *Parametric approach*. Finally, we present an approach which combines the assumption used in MC-SP-MLM about the noise sequences with the assumption that $\{s(t)\}$ can be well described as a sum of a pre-specified number of damped sinusoids. We derive a modified version of the HSVD method (see, e.g. [9–11]), which we call MC-P-HSVD (multichannel parametric Hankel-(matrix) singular value decomposition), that makes use of the above stated prior knowledge. The output of MC-P-HSVD is an estimated set of signal parameters, which are related to the given number of signal components.

In Section 2, we briefly present the theory of the four signal processing algorithms listed above, which all can be applied to the multicoil data in (1). We also try to point out some advantages, limitations, and drawbacks of each method. In Section 3, we first present a thorough analysis of the performance of these techniques based on simulated ^{1}H MRS data. Then, we investigate the spectral estimation performance of these methods using experimental data from a standard General Electric (GE) spectroscopic phantom containing a solution of metabolites with known concentrations. Finally, in Section 4, we comment on the accuracy and computational complexity of each method and give suggestions on when to use the different approaches.

2. Methods

2.1. Nonparametric approaches

A nonparametric method estimates $\{s(t)\}$ (and possibly also $\{g_j\}$) in (1) without making any specific assumptions on $\{s(t)\}$ nor on $\{e_j(t)\}$. The so-obtained estimate $\{\hat{s}(t)\}$ of $\{s(t)\}$ is then spectrally analyzed via existing methods that could have been applied to $\{s(t)\}$ directly if it were available.

2.1.1. The MC-NP-SVD method

Using the following notations

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix},$$

$$\mathbf{s}^* = [\, s(1) \quad \cdots \quad s(N) \,],$$

where $(\cdot)^*$ denotes the conjugate transpose, we can rewrite (1) as:

$$\mathbf{Y} \triangleq [\, \mathbf{y}(1) \quad \cdots \quad \mathbf{y}(N) \,] = \mathbf{g}\mathbf{s}^* + \boldsymbol{\varepsilon}, \tag{2}$$

where $\boldsymbol{\varepsilon}$ is defined similarly to $\mathbf{Y}$. It follows from (2) that in the high-SNR case the matrix $\mathbf{Y}$ is well approximated by the rank-one matrix $\mathbf{g}\mathbf{s}^*$. Let

$$\mathbf{Y} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^* \tag{3}$$

denote the SVD of $\mathbf{Y}$ where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\boldsymbol{\Sigma}$ is a diagonal matrix with nonnegative main diagonal entries. Also, let σ denote the maximum singular value, and let $\mathbf{u}$ and $\mathbf{v}$ be the corresponding left and right singular vectors, respectively. $\mathbf{g}$ and $\mathbf{s}$ can then be estimated as follows (to within a multiplicative constant):

$$\hat{\mathbf{g}} = \mathbf{u}, \quad \hat{\mathbf{s}} = \sigma\mathbf{v}. \tag{4}$$

The scaling ambiguity in (4) can be eliminated if we know something about the gains $\{g_j\}$, e.g., $g_1 = 1$ or $\|\mathbf{g}\| = 1$.

2.1.2. The cross-relation method

Another possible nonparametric method, which in blind system identification literature (see, e.g., [12,13]) is known as the cross-relation method, is based on the observation (which follows directly from (1)) that, again in the high-SNR case,

$$y_j(t)g_i \approx y_i(t)g_j, \quad i = 1, \ldots, m-1, \quad j = i+1, \ldots, m. \tag{5}$$

Based on the above observation, $\mathbf{g}$ can be estimated by solving the following minimization problem:

$$\hat{\mathbf{g}} = \arg\min_{\mathbf{g}} \sum_{t=1}^{N} \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} |g_i y_j(t) - g_j y_i(t)|^2$$
$$\text{subject to, e.g.,} \quad \|\mathbf{g}\| = 1. \tag{6}$$

Then $\{s(t)\}$ can be estimated from (1) via:

$$\hat{s}(t) = \hat{\mathbf{g}}^* \mathbf{y}(t), \quad t = 1, \ldots, N. \tag{7}$$

This approach has been used for multichannel MRS in [6]. The method based on (5) and (6) seems more intricate than that based on (4). It can be shown, however, that the two approaches are equivalent (see Appendix A for a proof of this claim). For this reason, we consider only the MC-NP-SVD method in Section 3.

2.2. The semiparametric approach

A semiparametric technique makes use of some prior information about the data in (1). For the method proposed below we introduce some assumptions on the noise sequences $\{e_j(t)\}$.

2.2.1. The MC-SP-MLM

A reasonable assumption on (1) is that the noise sequences are independent. A more restrictive assumption is that the noise sequences $\{e_j(t)\}, j = 1, \ldots, m$ are independent of each other, white, and circularly Gaussian with means zero and variances $\{\sigma_j^2\}$. Under the previous assumptions, the negative log-likelihood function (see, e.g. [8,14]) of $\{\mathbf{y}(t)\}$ is given by (to within an additive constant):

$$\ln\left(\prod_{j=1}^{m}\sigma_j^2\right)+\frac{1}{N}\sum_{t=1}^{N}[\mathbf{y}(t)-\mathbf{g}s(t)]^*\mathbf{Q}^{-1}[\mathbf{y}(t)-\mathbf{g}s(t)], \quad (8)$$

where

$$\mathbf{Q}=\begin{bmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_m^2\end{bmatrix}=\operatorname{cov}\left(\begin{bmatrix}e_1(t)\\ \vdots\\ e_m(t)\end{bmatrix}\right) \quad (9)$$

and cov denotes the covariance matrix. The function in (8), which is to be minimized with respect to $\mathbf{g}$, $\{s(t)\}$, and $\{\sigma_j^2\}$, is a reasonable fitting criterion even when the $\{e_j(t)\}$ are not Gaussian distributed.

We propose a *cyclic algorithm* for minimizing (8) owing to the computational convenience of such a technique, and its property of monotonically decreasing (1) at each iteration. The two main steps of this algorithm are as follows.

Step 1. For a given $\mathbf{Q}$, the minimizers $\hat{\mathbf{g}}$ and $\hat{s}(t)$ of (8) can be obtained in the following way. Let

$$\tilde{\mathbf{g}}=\mathbf{Q}^{-1/2}\mathbf{g}, \quad \tilde{\mathbf{y}}(t)=\mathbf{Q}^{-1/2}\mathbf{y}(t), \quad (10)$$

where $\mathbf{Q}^{-1/2}$ denotes the square root of $\mathbf{Q}^{-1}$. The minimization of (8) with respect to $s(t)$ gives:

$$\hat{s}(t)=\frac{\mathbf{g}^*\mathbf{Q}^{-1}\mathbf{y}(t)}{\mathbf{g}^*\mathbf{Q}^{-1}\mathbf{g}}=\frac{\tilde{\mathbf{g}}^*\tilde{\mathbf{y}}(t)}{\|\tilde{\mathbf{g}}\|^2}, \quad t=1,\ldots,N. \quad (11)$$

The function left for minimization with respect to $\mathbf{g}$ (or $\tilde{\mathbf{g}}$) becomes:

$$\begin{aligned}\sum_{t=1}^{N}\|\tilde{\mathbf{y}}(t)-\tilde{\mathbf{g}}\hat{s}(t)\|^2&=\sum_{t=1}^{N}\left\|\left[\mathbf{I}-\frac{\tilde{\mathbf{g}}\tilde{\mathbf{g}}^*}{\|\tilde{\mathbf{g}}\|^2}\right]\tilde{\mathbf{y}}(t)\right\|^2\\&=\sum_{t=1}^{N}\tilde{\mathbf{y}}^*(t)\left(\mathbf{I}-\frac{\tilde{\mathbf{g}}\tilde{\mathbf{g}}^*}{\|\tilde{\mathbf{g}}\|^2}\right)\tilde{\mathbf{y}}(t)\\&=\operatorname{tr}\left[\left(\mathbf{I}-\frac{\tilde{\mathbf{g}}\tilde{\mathbf{g}}^*}{\|\tilde{\mathbf{g}}\|^2}\right)(\mathbf{Q}^{-1/2}\hat{\mathbf{R}}\mathbf{Q}^{-1/2})\right],\end{aligned} \quad (12)$$

where

$$\hat{\mathbf{R}}=\sum_{t=1}^{N}\mathbf{y}(t)\mathbf{y}^*(t) \quad (13)$$

and $\mathbf{I}$ is the identity matrix.

It follows from (12) that the problem is reduced to:

$$\max_{\tilde{\mathbf{g}}}\frac{\tilde{\mathbf{g}}^*(\mathbf{Q}^{-1/2}\hat{\mathbf{R}}\mathbf{Q}^{-1/2})\tilde{\mathbf{g}}}{\|\tilde{\mathbf{g}}\|^2} \quad (14)$$

whose solution is well known: $\hat{\tilde{\mathbf{g}}}=$ the maximum eigenvector of $\mathbf{Q}^{-1/2}\hat{\mathbf{R}}\mathbf{Q}^{-1/2}$ (or any scaled version thereof). Consequently:

$$\hat{\mathbf{g}}=\mathbf{Q}^{1/2}\times\max \text{ eigenvector of } (\mathbf{Q}^{-1/2}\hat{\mathbf{R}}\mathbf{Q}^{-1/2}). \quad (15)$$

Step 2. For given $\mathbf{g}$ and $s(t)$, the minimization of (8) with respect to $\{\sigma_j^2\}$ is immediate:

$$\hat{\sigma}_j^2=\frac{1}{N}\sum_{t=1}^{N}|y_j(t)-g_j s(t)|^2, \quad j=1,\ldots,m. \quad (16)$$

The algorithm consists of iterating the above two steps starting, e.g., from an initial estimate of $\mathbf{Q}$. A natural such estimate of $\mathbf{Q}$ when there is no prior information about $\{\sigma_j\}$ is $\mathbf{Q}=\mathbf{I}$. This gives:

$$\hat{\mathbf{g}}=\max \text{ eigenvector of } \hat{\mathbf{R}}, \quad \hat{s}(t)=\hat{\mathbf{g}}^*\mathbf{y}(t) \quad (17)$$

which is nothing but the MC-NP-SVD estimate in (4) (see Appendix A). As a consequence of this observation we may expect that (4) and the present method (MC-SP-MLM) have similar performance when the noise variances in the different coils have similar values, but also that MC-SP-MLM outperforms (4) otherwise.

According to the general discussion on this type of estimation problems in [15] we recommend performing between two and five iterations with the above algorithm, as further iterations will usually not improve the estimates. In Section 3, we perform five iterations of this cyclic minimization algorithm.

2.3. The parametric approach

A parametric method is based on detailed assumptions on both $\{s(t)\}$ and $\{e_j(t)\}$. In what follows we assume that $\{s(t)\}$ consists of a given (known or estimated) number of exponentially damped sinusoids and that the noise sequences $\{e_j(t)\}$ are white. $\{s(t)\}$ can then be modeled as:

$$s(t)=\sum_{k=1}^{n}\beta_k\lambda_k^t, \quad \lambda_k=\mathrm{e}^{-\alpha_k+\mathrm{i}\omega_k}, \quad t=1,\ldots,N, \quad (18)$$

where n denotes the number of signal components, λ_k are the modes of the signal, and $(\beta_k, \alpha_k, \omega_k)$ are the complex amplitude, damping, and angular frequency of the kth component. The sampling period has been absorbed in α_k and ω_k, for notational convenience.

There are several techniques available to estimate the parameters in (18) (see, e.g. [14] for more information). An alternative parametric approach would be to consider the LCModel method [16] in lieu of a method that estimates the parameters in (18). We do not consider the LCModel approach in this paper.

2.3.1. The MC-P-HSVD method

Below we present an extension of the HSVD (see [9–11]) method to multicoil MR data. HSVD is based on the SVD of a certain Hankel data matrix (see below). In this case, we will construct one such data matrix for each coil $j = 1, \ldots, m$. The m Hankel data matrices will be appended to one another, and the signal parameter estimates will be computed considering all coils together (as for the previously presented methods).

Inserting (18) into (1) gives us the parametric data equation:

$$y_j(t) = g_j \sum_{k=1}^{n} \beta_k \lambda_k^t + e_j(t), \quad t = 1, \ldots, N, \;\; j = 1, \ldots, m. \tag{19}$$

Let

$$\mathbf{Z}_j = \begin{bmatrix} y_j(1) & y_j(2) & \cdots & y_j(L) \\ y_j(2) & y_j(3) & \cdots & y_j(L+1) \\ \vdots & \vdots & \ddots & \vdots \\ y_j(M) & y_j(M+1) & \cdots & y_j(L+M-1) \end{bmatrix} \tag{20}$$

be the $(M \times L)$ Hankel data matrix for each coil $j = 1, \ldots, m$, where M is a user parameter usually chosen close or equal to $N/2$ (see, e.g. [9–11,14] for more information on the choice of M) and $L = N - M + 1$.

Now we can define the following $(M \times Lm)$ multichannel data matrix

$$\mathbf{Z} = [\, \mathbf{Z}_1 \quad \cdots \quad \mathbf{Z}_m \,]. \tag{21}$$

In the high-SNR case, the rank of (21) is approximately equal to n. In general n is smaller than $\min(M, L)$, which implies that (20) (and consequently (21)) is rank deficient.

In the following, we outline the procedure for retrieving the modes λ_k in (18). We use the SVD to factor the data matrix in (21) into a product of three matrices:

$$\mathbf{Z} = \tilde{\mathbf{U}} \tilde{\mathbf{\Sigma}} \tilde{\mathbf{V}}^*, \tag{22}$$

where $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ are unitary matrices and $\tilde{\mathbf{\Sigma}}$ is a diagonal matrix with nonnegative main diagonal entries. For the remaining part we will only use the n left singular vectors corresponding to the n largest singular values of $\mathbf{Z}$. Hence we let $\tilde{\mathbf{U}}_n$ denote the n first columns of $\tilde{\mathbf{U}}$.

Next we find the solution $\hat{\mathbf{\Lambda}}$ to the following linear system of equations:

$$\tilde{\mathbf{U}}_u \mathbf{\Lambda} = \tilde{\mathbf{U}}_l, \tag{23}$$

where

$$\tilde{\mathbf{U}}_u = [\mathbf{I}_{M-1} \quad 0] \tilde{\mathbf{U}}_n, \tag{24}$$

$$\tilde{\mathbf{U}}_l = [0 \quad \mathbf{I}_{M-1}] \tilde{\mathbf{U}}_n. \tag{25}$$

The least-squares (LS) solution to (23) is given by (see, e.g. [14])

$$\hat{\mathbf{\Lambda}} = (\tilde{\mathbf{U}}_u^* \tilde{\mathbf{U}}_u)^{-1} \tilde{\mathbf{U}}_u^* \tilde{\mathbf{U}}_l. \tag{26}$$

Now we can obtain estimates $\hat{\lambda}_k$ of the desired modes λ_k, $k = 1, \ldots, n$ in (18) as the eigenvalues of $\hat{\mathbf{\Lambda}}$. See [9–11] for details on the HSVD approach for singlecoil data.

Finally, we address the problem of estimating the complex amplitudes β_k in (18). Use the estimates $\hat{\lambda}_k$ to construct the matrix

$$\mathbf{A} = \begin{bmatrix} \hat{\lambda}_1 & \cdots & \hat{\lambda}_n \\ \vdots & \ddots & \vdots \\ \hat{\lambda}_1^N & \cdots & \hat{\lambda}_n^N \end{bmatrix}. \tag{27}$$

In addition, let

$$\mathbf{c}_j = g_j \boldsymbol{\beta}, \quad \boldsymbol{\beta} = [\, \beta_1 \quad \cdots \quad \beta_n \,]^{\mathrm{T}}, \quad j = 1, \ldots, m. \tag{28}$$

We use LS (see, e.g. [14]) to estimate $\mathbf{c}_j$ for each detector coil $j = 1, \ldots, m$ as

$$\hat{\mathbf{c}}_j = (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \mathbf{y}_j, \tag{29}$$

where

$$\mathbf{y}_j = \begin{bmatrix} y_j(1) & \cdots & y_j(N) \end{bmatrix}^{\mathrm{T}}. \tag{30}$$

Next, construct the matrix

$$\hat{\mathbf{C}} = [\, \hat{\mathbf{c}}_1 \quad \cdots \quad \hat{\mathbf{c}}_m \,]. \tag{31}$$

Let $\bar{\sigma}$ denote the maximum singular value of $\hat{\mathbf{C}}$, and let $\bar{\mathbf{u}}$ and $\bar{\mathbf{v}}$ be the corresponding left and right singular vectors. Note that $\hat{\mathbf{C}}$ is an estimate of $\mathbf{C} = [\, \mathbf{c}_1 \quad \cdots \quad \mathbf{c}_m \,] = \boldsymbol{\beta} [\, g_1 \quad \cdots \quad g_m \,]$. Hence $\mathbf{g}$ and $\boldsymbol{\beta}$ can be estimated as follows (to within a multiplicative constant):

$$\hat{\boldsymbol{\beta}} = \bar{\sigma} \bar{\mathbf{u}}, \quad \hat{\mathbf{g}}^{\mathrm{T}} = \bar{\mathbf{v}}^*. \tag{32}$$

3. Numerical examples

3.1. Simulated data

An artificial biomedical ^{1}H MRS data example is considered in detail to compare the improvement in estimation accuracy of the different methods when the multicoil approach is used in contrast to a singlecoil approach. We consider an MRS signal that includes only the eight largest peaks of a typical in vivo ^{1}H spectrum measured in the human brain. We further assume that the water component has been eliminated and that any baseline distortion has been corrected. The number

of samples N is 512 and the sampling frequency is 500 Hz. An FFT (fast Fourier transform) spectrum of the noise-free (zero-padded) data signal is presented in Fig. 1 where the peaks identified by numbers correspond to the following substances: 1, 3—myoinositol; 2, 5—creatine; 4—choline; 6—NAA (represented as a singlet); 7, 8—lactate. The noise added to the previous signal is zero mean, white, and circularly Gaussian distributed with standard deviation σ_j for each individual coil $j = 1, \ldots, m$. This signal is to be studied using an 8 coil MR system. The channel gains for this system are randomly created with an amplitude chosen from the uniform distribution $[0.8, 1.2]$, and a phase taken from the uniform distribution $[0, 2\pi]$. As a reference, g_1 is always set equal to 1.

Since MC-NP-SVD and MC-SP-MLM are nonparametric methods, which compute an estimate of $\{s(t)\}$, while MC-P-HSVD estimates a list of signal parameters, comparing the performance of these three methods is not a trivial task. For simplicity, we will base most of the graphical presentation in this section on FFT spectra for each of the methods, first considering data sequences from one typical simulated multicoil scan with fixed, but randomly created, noise sequences and channel gains, and second considering the results of Monte Carlo simulations.

In the first simulation, a single scan is considered using the previously introduced 8 coil MR system. The noise standard deviation for each channel is equal to 20 for this example, which corresponds to an SNR of about 2 dB. The number of estimated components n is set to 8 for MC-P-HSVD. In addition, the following parameter settings are used generally throughout this section; for MC-SP-MLM we set $\mathbf{Q}$ equal to $\mathbf{I}$ initially, and we perform five iterations of the cyclic algorithm. For MC-P-HSVD, M is set equal to $\lfloor \frac{N}{2} \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part. The obtained spectra are shown in Fig. 2 for each of the three methods, with $m = 1$ (using only the first coil) and using all available channels ($m = 8$). For $m = 1$, the results of MC-NP-SVD and MC-SP-MLM are equivalent (see Section 2). We see from Fig. 2 that the influence of the relatively high noise level is significantly reduced when the signals from all eight coils are used, as expected. We also note that the results of MC-NP-SVD and MC-SP-MLM are remarkably similar in this scenario even for $m = 8$. As mentioned in Section 2, this is to be expected when the noise standard deviations in the different channels are similar to one another. For MC-P-HSVD, we see that it is difficult to resolve all signal components at this low SNR when $m = 1$ (note that for $m = 1$, MC-P-HSVD reduces to the standard HSVD). Specifically, peaks number 1, 3, and 7 are not visible at all in the estimated HSVD spectrum. However, for $m = 8$ all eight components are well estimated. Note also that by using a parametric approach, such as MC-P-HSVD, it is possible to show the different estimated signal components individually which can simplify substance classification in a scenario where the peaks are overlapping (see Section 3.2).

The computational complexities of the three methods are rather different, at least when comparing MC-P-HSVD to the other two techniques. This is to be expected, since parametric methods are generally more computationally demanding than nonparametric methods. On the other hand, one can argue that the results obtained by a parametric method contain more information than the corresponding results from a nonparametric technique. Table 1 shows the number of flops

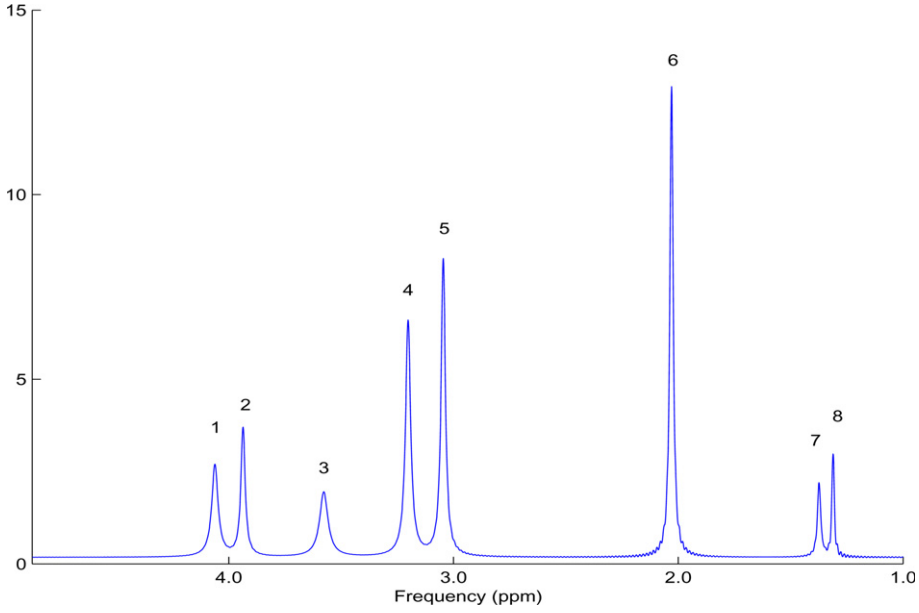

Fig. 1. FFT spectrum of the noise-free simulated 8-peak ^{1}H MRS data.

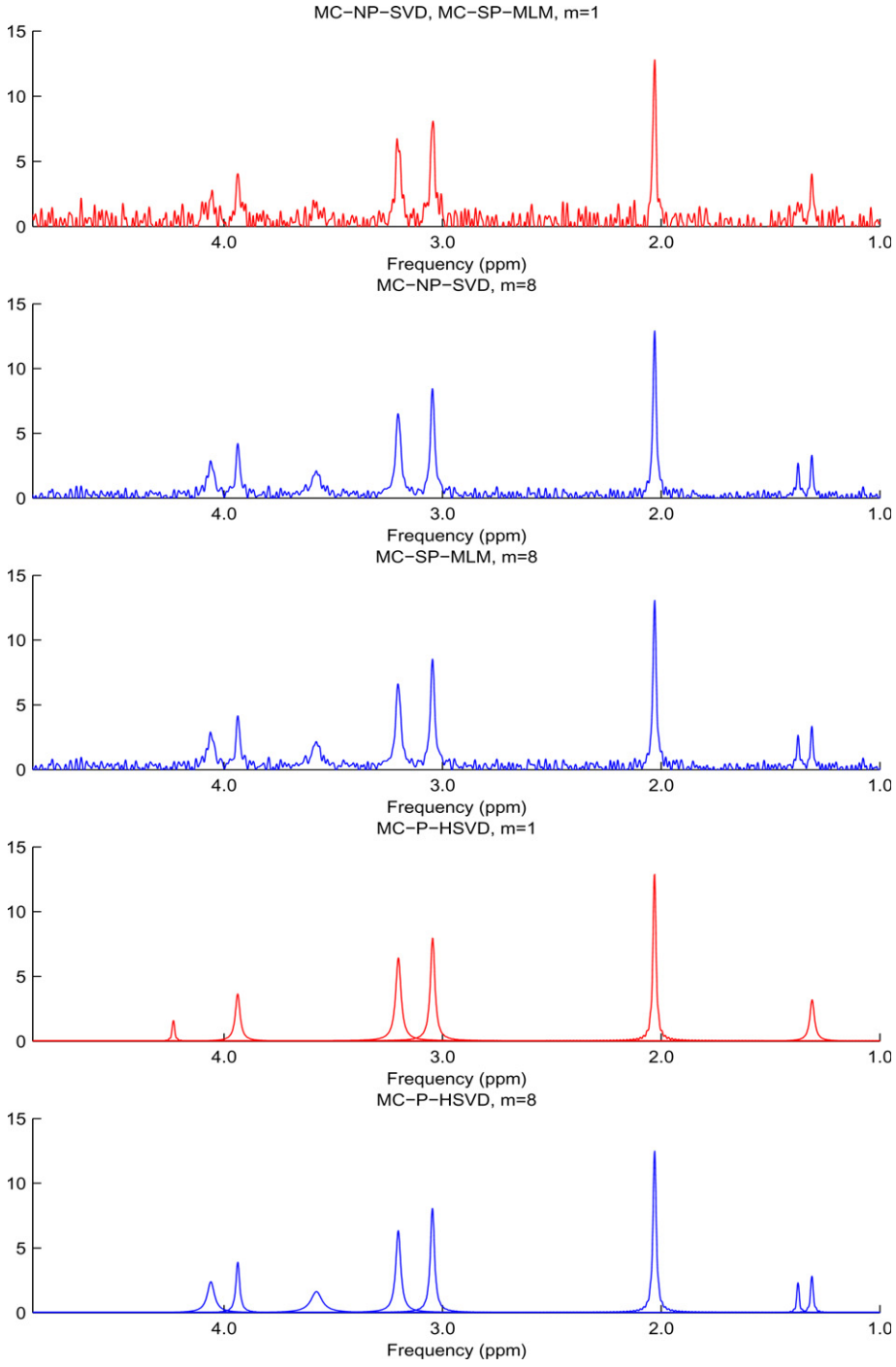

Fig. 2. Eight-peak ^{1}H spectra from simulated data for each of the three methods when $m = 1$ and $m = 8$, respectively. The noise standard deviation (σ) is equal to 20 for all channels.

Table 1
Required number of flops for each of the three methods applied to simulated ^{1}H MRS data when $m = 1$ and $m = 8$

Method	m	
	$m = 1$	$m = 8$
MC-NP-SVD	1.0×10^4	3.4×10^5
MC-SP-MLM	8.4×10^4	2.3×10^6
MC-P-HSVD	7.3×10^8	2.5×10^9

(floating point operations) required to compute each of the spectra in Fig. 2 when $m = 1$ and $m = 8$, respectively, as an indication of the computational speed of the different approaches. MC-NP-SVD and MC-SP-MLM are significantly faster than MC-P-HSVD with MC-SP-SVD having an edge.

In the second simulation, we show Monte Carlo results for the previous example but for noise standard

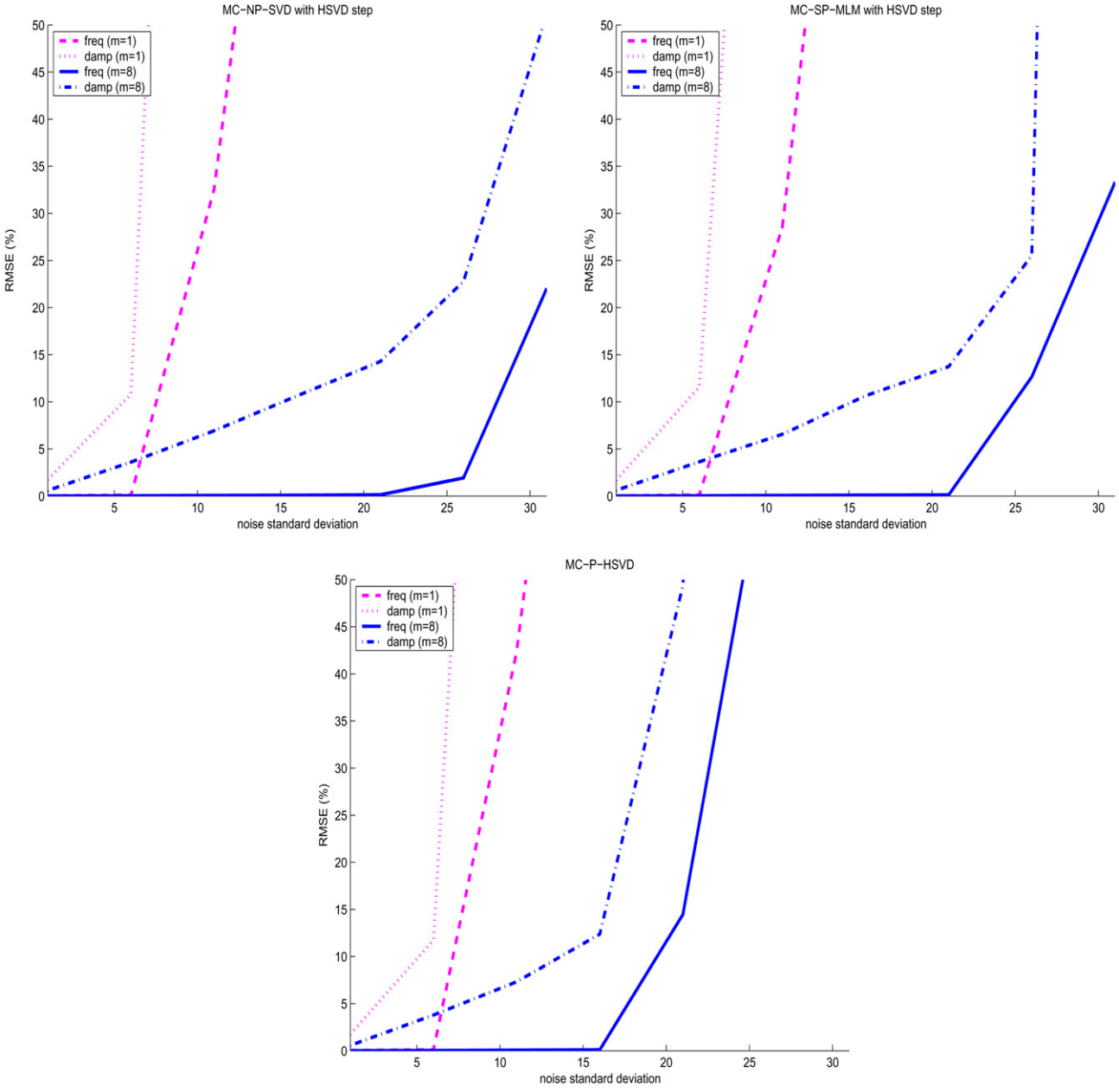

Fig. 3. Frequency (freq) and damping (damp) RMSEs versus noise standard deviation for the eight components in the simulated ^{1}H data when $m = 1$ and $m = 8$ for each of the three methods.

deviations increasing from 1 to 31, in steps of 5. For this purpose, we need to estimate the signal parameters for MC-NP-SVD and MC-SP-MLM (for MC-P-HSVD we get the signal parameters directly). We use the standard HSVD method [9–11] for this task and set the expected number of components n to 8. The quality of the different parameter estimates is measured as the average relative root mean square error (RMSE) [in percent]:

$$\mathrm{RMSE} \triangleq 100\sqrt{\frac{1}{n\Upsilon}\sum_{\upsilon=1}^{\Upsilon}\sum_{k=1}^{n}\frac{(\xi_k-\hat{\xi}_k^{\upsilon})^2}{\xi_k^2}}, \tag{33}$$

where Υ is the number of Monte Carlo runs (we use 1000), ξ_k denotes the relevant parameter, and $\hat{\xi}_k^{\upsilon}$ is its estimate obtained in the υth run. In Fig. 3 the RMSEs

Table 2
Required number of flops for the MC-NP-SVD and MC-SP-MLM methods, including the HSVD step, applied to simulated ^{1}H MRS data when $m = 1$ and $m = 8$

Method	m	
	$m = 1$	$m = 8$
MC-NP-SVD	7.2537×10^8	7.2570×10^8
MC-SP-MLM	7.2545×10^8	7.2768×10^8

of the frequencies (freq) and dampings (damp) are presented for each method for both $m = 1$ and $m = 8$. Fig. 3 shows a significantly reduced sensitiveness to noise interference for all presented methods using the multicoil approach ($m = 8$) in lieu of $m = 1$. We see that there is a significant increase in the RMSEs at a certain noise level for all methods in Fig. 3. The reason for this is the occurrence of a noise component, somewhere in the spectrum, with higher amplitude than one of the eight signal components of interest which leads to a false peak estimate using HSVD with $n = 8$.

Adding the HSVD step to MC-NP-SVD and MC-SP-MLM changes the computational complexity of these two techniques significantly, as expected. Table 2 shows the number of flops required to perform one run of MC-NP-SVD and MC-SP-MLM including the HSVD step

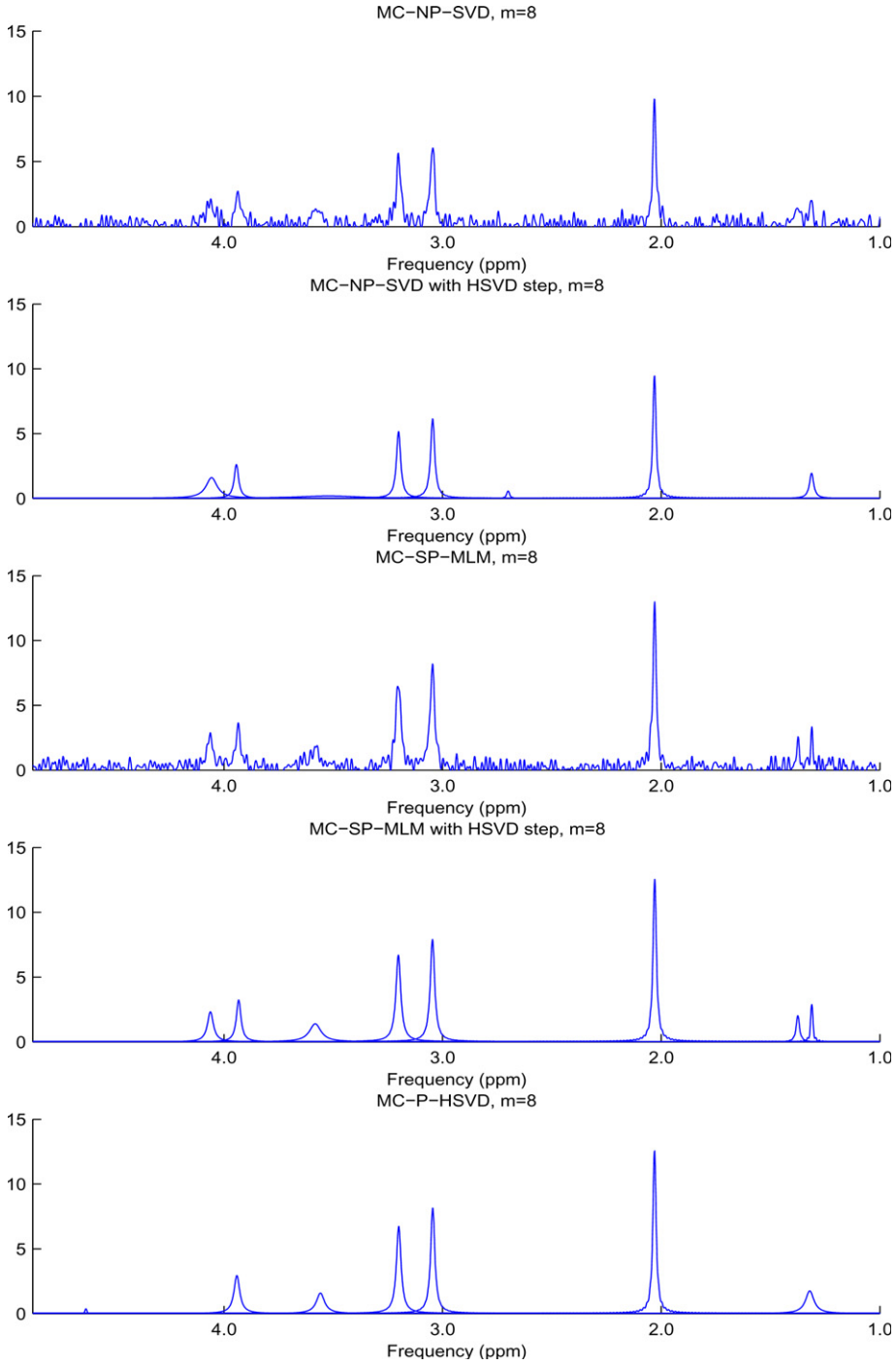

Fig. 4. Eight-peak ^{1}H spectra from simulated data for each of the three methods when $m = 8$ and the noise standard deviations for the different channels are randomly chosen from the uniform distribution $[10, 60]$.

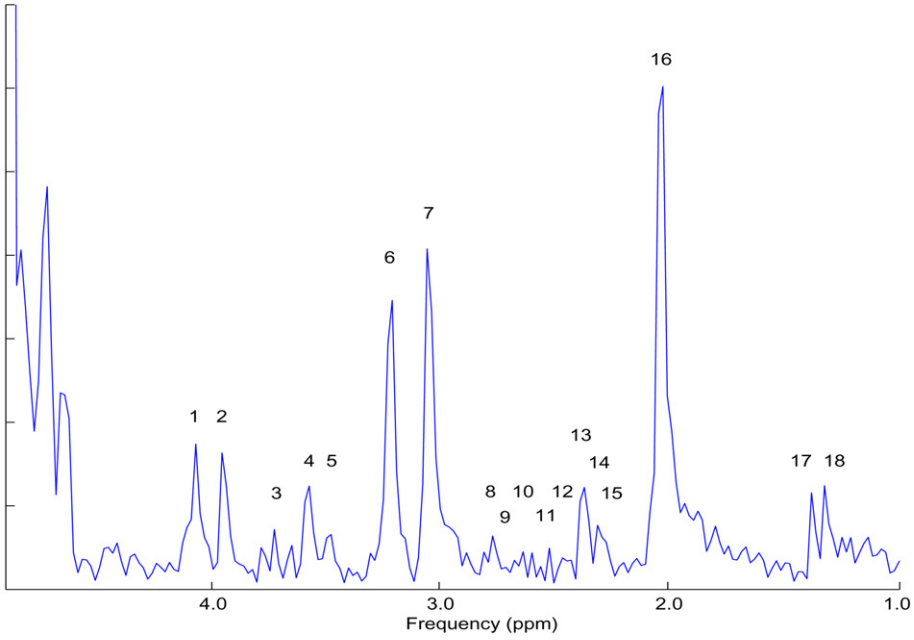

Fig. 5. FFT spectrum of the 18-peak standard GE phantom data acquired using a single element of the 8-channel domed head coil and averaging over 16 frames.

as in Fig. 3 when $m = 1$ and $m = 8$, respectively. These numbers are to be compared with those of the MC-P-HSVD method in Table 1.

To conclude the study on the simulated ^{1}H MRS data, we show a numerical example to support our claim in Section 2 regarding the difference in accuracy especially between MC-NP-SVD and MC-SP-MLM when the noise standard deviations in various channels are relatively different from one another. We let the noise standard deviation for each of the eight channels be randomly chosen from the uniform distribution $[10, 60]$. As in the first simulation, we consider a typical multichannel scan with fixed, but randomly created, noise sequences (with standard deviations as above) and channel gains. The obtained spectra (for all three methods) are presented in Fig. 4. This figure also shows the results for MC-NP-SVD and MC-SP-MLM followed by the HSVD step, which (similar to MC-P-HSVD) provide spectra that show the estimated signal components separately. MC-SP-MLM shows significantly better performance than the other two techniques in this case when the noise standard deviations have largely different values, as expected.

3.2. Phantom data

We consider a GE MRS phantom with known solutions of metabolites whose concentrations are approximately the same as what might be found in a human brain. The phantom contains a solution of 12.5 mM NAA, 10.0 mM creatine (Cr), 3.0 mM choline (Ch), 7.5 mM myoinositol (mI), 12.5 mM L-glutamic acid (Glu), 5.0 mM lactate, and 0.5 mM γ-aminobutyric acid (GABA). In Fig. 5 the FFT spectrum obtained by averaging 16 phase-corrected, water-suppressed frames (see below) for the region of interest is shown where the peaks identified by numbers correspond to the following substances: 1, 5—myoinositol; 2, 7—creatine; 3, 4, 12–15—glutamate; 6—choline; 8–11, 16—NAA; 17, 18—lactate. It can be seen from Fig. 5 that some of the peaks are hard to detect since they are more or less buried in noise.

We will compare spectra obtained from two different coils with similar volume. Data from a single channel quadrature head coil manufactured by GE Healthcare Technologies, Milwaukee, WI, USA, will be used as a singlecoil reference and data from an 8 channel domed head coil manufactured by MRI Devices, Waukesha, WI, USA, will be used for the multicoil analyses. Data were collected using a 3.0 T GE Signa MR scanner equipped with a high-bandwidth (1 MHz) data acquisition subsystem and a TwinSpeed gradient coil capable of 40 mT/cm at a maximum slew rate of 150 T/m/s. A conventional point resolved spectroscopy (PRESS) pulse sequence was used to acquire the data. The region of interest for these MRS experiments was an 8 cm^3 volume located at iso-center. For each single-voxel MRS scan, a set of nonwater-suppressed reference data was collected along with a corresponding set of water-suppressed data for which water was suppressed using a chemical shift selective (CHESS) technique [17]. The (water-suppressed) metabolite data were phase-corrected, with removal of residual water, using high-SNR reference data (see, e.g. [18]). No zero-filling or apodization was applied. Timing parameters were TE = 35 ms and TR = 1500 ms. The spectral bandwidth

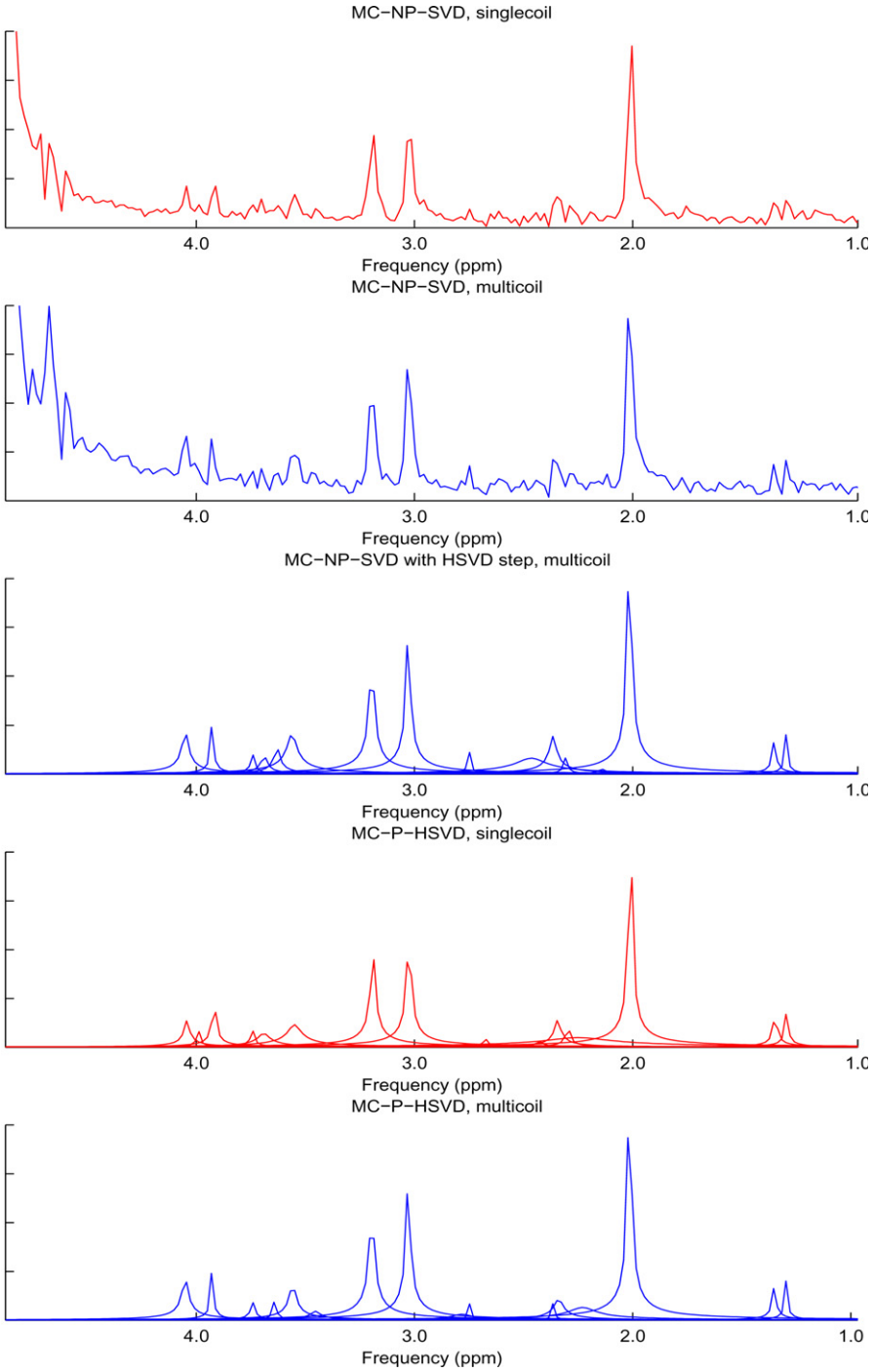

Fig. 6. Spectra from the 18-peak standard GE phantom data example for MC-NP-SVD and MC-P-HSVD for singlecoil and multicoil data, respectively. The results of MC-SP-MLM are similar to those of MC-NP-SVD and have therefore been omitted.

was 2500 Hz, and $N = 2048$ complex valued points were sampled for each readout.

The spectrum in Fig. 5 is acquired by averaging 16 phase-corrected, water-suppressed frames measured by one element of the 8 channel domed head coil. The averaging procedure reduces the noise interference at the cost of an increased scanning time. By using only a single acquisition with data from 8 receive coils processed by one of the proposed methods, satisfactory MRS results may be obtained at a significantly reduced scan time. Alternatively, the performance can be improved by using 8 receive coils *and* data which are an average of 16 frames. In the following, we will consider data from a single acquisition. The different spectra obtained by using MC-NP-SVD and MC-P-HSVD are presented in Fig. 6. The results of MC-SP-MLM are similar to

those of MC-NP-SVD (as in Fig. 2) and have therefore been omitted from Fig. 6. The spectra obtained using the single channel quadrature head coil data are referred to as 'singlecoil' and the spectra obtained using the 8 channel domed head coil data are referred to as 'multicoil.' The number of estimated components n is set to 36 for both MC-P-HSVD and the HSVD step added to MC-NP-SVD. The main reason why the selected value for n is larger than the number of signal components of interest is that several modes are used to model the residual water component (around 4.7 ppm in Fig. 5). Since this component is not of interest here (for the graphical presentation), the corresponding estimates have been omitted from Fig. 6. Note also that the results obtained by using MC-NP-SVD with the HSVD step for singlecoil data are equivalent with using the MC-P-HSVD for singlecoil data directly.

A general difference between the singlecoil and multicoil spectra in Fig. 6 is that several signal components appear to have lower amplitudes (and hence worse SNR) in the singlecoil spectra than in the corresponding multicoil plots. Specifically, the lactate peaks located around 1.3 ppm and the myoinositol and creatine components around 4.0 ppm are harder to separate from the noise in the singlecoil case compared to the multicoil spectra. The concentrations of the choline and creatine components at about 3.2 and 3.0 ppm, respectively, do not appear to be correctly estimated using the singlecoil approach since the creatine component is supposed to be larger than the choline peak. This information is much better represented in the multicoil spectra. Regarding the smallest components, such as the glutamate peaks, it is hard to get any reliable information from any of the spectra in Fig. 6 in this scenario due to the relatively high level of noise.

4. Conclusions

The numerical results in Section 3 show that multichannel MR spectroscopy can improve the SNR and increase the spectral resolution in practical applications where there is a possibility to use multiple receive coils. In addition, a reduction in scan time may be obtained using a multichannel approach for MRS studies in lieu of a single channel technique to acquire a satisfactory MR absorption spectrum. The main goals of this paper are to present the multicoil problem from different signal processing perspectives, invoke prior knowledge when it is reasonable to do so, and present several different solutions to the multichannel MRS problem. We do not aim at including all possible methods here, but rather to cover the range of solutions from nonparametric ones like MC-NP-SVD, which requires no prior knowledge about the data, to the fully parametric MC-P-HSVD, which requires detailed assumptions regarding both the measurement noise in each channel and how to model the MR signal.

If we consider each of the three techniques (MC-NP-SVD, MC-SP-MLM, and MC-P-HSVD) individually we can conclude that MC-NP-SVD performs well from a nonparametric perspective, without requiring any specific assumptions on $\{s(t)\}$ nor on $\{e_j(t)\}$. On the other hand it can be hard to separate the signal components even for a relatively low level of noise. If we expect the noise standard deviations for different coils to be largely different from one another, then MC-SP-MLM is a natural choice in lieu of MC-NP-SVD since the spectral estimation performance can be significantly improved at only a minor increase in the computational complexity. Note that the general advantages in signal estimation performance obtained by using multiple receive coils, as presented in this paper, can be partly lost when the FFT is used for generation of the spectra (as in Figs. 2, 4, and 6) due to the limitations and drawbacks of the FFT (see, e.g. [14]). One alternative solution, which does not impose these limitations on the results, is to consider the parameter estimates of the individual signal components (as in Fig. 3). This approach requires, however, some knowledge about the expected number of components in the data.

A general problem with MC-P-HSVD, compared to MC-NP-SVD and MC-SP-MLM, is that individual signal components can have spurious estimates if the SNR is too low. However, in many applications, when there is enough prior information available about the spectral content of a sample, such spurious estimates can be disregarded. An advantage with the parametric approach is the appealing feature of computing the signal parameters corresponding to each individual component separately. Even if MC-P-HSVD is slower than MC-SP-SVD and MC-NP-MLM, the required computational time can be of minor importance in a scenario where individual presentation of the signal components is a requirement for successful signal quantitation. An alternative approach to obtain spectra which show the estimated signal components separately is to use either MC-SP-SVD or MC-NP-MLM followed by the HSVD step as explained in Section 3.

Appendix A. Proof of equivalence between the MC-NP-SVD method and the cross-correlation method

The first step is to prove by induction that

$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}|g_i y_j(t)-g_j y_i(t)|^2 = \|\mathbf{y}(t)\|^2\|\mathbf{g}\|^2-|\mathbf{g}^*\mathbf{y}(t)|^2 \tag{A.1}$$

for each $t=1,\ldots,N$.

For $m = 2$, a simple calculation shows that:

$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}|g_i y_j(t) - g_j y_i(t)|^2 = |g_1 y_2(t) - g_2 y_1(t)|^2$$
$$= |g_1|^2|y_2(t)|^2 + |g_2|^2|y_1(t)|^2 - 2\mathrm{Re}(g_1^* y_2^*(t) g_2 y_1(t))$$

and

$$\|\mathbf{y}(t)\|^2\|\mathbf{g}\|^2 - |\mathbf{g}^*\mathbf{y}(t)|^2$$
$$= |y_1(t)|^2|g_1|^2 + |y_2(t)|^2|g_2|^2 + |y_1(t)|^2|g_2|^2 + |y_2(t)|^2|g_1|^2$$
$$- |g_1|^2|y_1(t)|^2 - |g_2|^2|y_2(t)|^2 - 2\mathrm{Re}(g_1^* y_1(t) g_2 y_2^*(t))$$
$$= |g_1|^2|y_2(t)|^2 + |g_2|^2|y_1(t)|^2 - 2\mathrm{Re}(g_1^* y_2^*(t) g_2 y_1(t)).$$

Hence (A.1) holds for $m = 2$.

Next, assume (A.1) holds for $m = k$ and let us prove it also holds for $m = k + 1$. We have, for $m = k + 1$:

$$\sum_{i=1}^{k}\sum_{j=i+1}^{k+1}|g_i y_j(t) - g_j y_i(t)|^2$$
$$= \sum_{i=1}^{k-1}\sum_{j=i+1}^{k+1}|g_i y_j(t) - g_j y_i(t)|^2 + |g_k y_{k+1}(t) - g_{k+1} y_k(t)|^2$$
$$= \sum_{i=1}^{k-1}\sum_{j=i+1}^{k}|g_i y_j(t) - g_j y_i(t)|^2 + |g_k y_{k+1}(t) - g_{k+1} y_k(t)|^2$$
$$+ \sum_{i=1}^{k-1}|g_i y_{k+1}(t) - g_{k+1} y_i(t)|^2$$
$$= \sum_{i=1}^{k-1}\sum_{j=i+1}^{k}|g_i y_j(t) - g_j y_i(t)|^2 + \sum_{i=1}^{k}|g_i y_{k+1}(t) - g_{k+1} y_i(t)|^2. \quad (A.2)$$

Letting

$$\bar{\mathbf{y}}(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_k(t) \end{bmatrix}, \quad \bar{\mathbf{g}} = \begin{bmatrix} g_1 \\ \vdots \\ g_k \end{bmatrix} \quad (A.3)$$

and using the fact that, by assumption (A.2) holds for $m = k$, we get from (A.3):

$$\sum_{i=1}^{k}\sum_{j=i+1}^{k+1}|g_i y_j(t) - g_j y_i(t)|^2$$
$$= \|\bar{\mathbf{y}}(t)\|^2\|\bar{\mathbf{g}}\|^2 - |\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t)|^2 + \sum_{i=1}^{k}|g_i y_{k+1}(t) - g_{k+1} y_i(t)|^2$$
$$= \|\bar{\mathbf{y}}(t)\|^2\|\bar{\mathbf{g}}\|^2 - |\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t)|^2 + |y_{k+1}(t)|^2\|\bar{\mathbf{g}}\|^2 + |g_{k+1}|^2\|\bar{\mathbf{y}}(t)\|^2$$
$$- 2\mathrm{Re}(y_{k+1}^*(t) g_{k+1}\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t)). \quad (A.4)$$

Next, note that

$$\|\mathbf{g}\|^2\|\mathbf{y}(t)\|^2 - |\mathbf{g}^*\mathbf{y}(t)|^2$$
$$= (\|\bar{\mathbf{g}}\|^2 + |g_{k+1}|^2)(\|\bar{\mathbf{y}}(t)\|^2 + |y_{k+1}(t)|^2) - |\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t) + g_{k+1}^* y_{k+1}(t)|^2$$
$$= \|\bar{\mathbf{g}}\|^2\|\bar{\mathbf{y}}(t)\|^2 + \|\bar{\mathbf{g}}\|^2|y_{k+1}(t)|^2 + |g_{k+1}|^2\|\bar{\mathbf{y}}(t)\|^2 + |g_{k+1}|^2|y_{k+1}(t)|^2$$
$$- |\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t)|^2 - |g_{k+1}|^2|y_{k+1}(t)|^2 - 2\mathrm{Re}(y_{k+1}^*(t) g_{k+1}\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t))$$
$$= \|\bar{\mathbf{y}}(t)\|^2\|\bar{\mathbf{g}}\|^2 - |\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t)|^2 + |y_{k+1}(t)|^2\|\bar{\mathbf{g}}\|^2 + |g_{k+1}|^2\|\bar{\mathbf{y}}(t)\|^2$$
$$- 2\mathrm{Re}(y_{k+1}^*(t) g_{k+1}\bar{\mathbf{g}}^*\bar{\mathbf{y}}(t)). \quad (A.5)$$

Comparing (A.4) and (A.5) concludes the proof of (A.1).

It follows that the relevant problem (5) can be equivalently formulated as:

$$\max_{\mathbf{g}} \underbrace{\sum_{t=1}^{N}|\mathbf{g}^*\mathbf{y}(t)|^2}_{\mathbf{g}^*\mathbf{Y}\mathbf{Y}^*\mathbf{g}} \quad \text{subject to, e.g.,} \quad \|\mathbf{g}\| = 1. \quad (A.6)$$

The solution is given by the dominant eigenvector of $\mathbf{Y}\mathbf{Y}^*$, or equivalently by the dominant left singular vector of $\mathbf{Y}$, which hence shows that the estimate of $\mathbf{g}$ in (6) coincides with that in (4).

Regarding the estimates of $\{s(t)\}$, for (7) we have (by the properties of SVD):

$$\hat{\mathbf{s}}^* = [\,\hat{s}(1) \quad \cdots \quad \hat{s}(N)\,] = \hat{\mathbf{g}}^*\mathbf{Y} = \sigma\mathbf{v}^* \quad (A.7)$$

and hence the estimate of $\mathbf{s}$ in (7) coincides with that in (4).

References

[1] P.B. Roemer, W.A. Edelstein, C.E. Hayes, S.P. Souza, O.M. Mueller, The NMR phased array, Magn. Reson. Med. 16 (1990) 192–225.

[2] G. Tan, W. Song, A. Jesmanowics, J.S. Hyde, T. Raidy, S.J. Li, Multi-channel magnetic resonance spectroscopy, in: Proceedings of the SMRM 12th Annual Meeting, 1993, p. 370.

[3] L.L. Wald, S.E. Moyher, M.R. Day, S.J. Nelson, D.B. Vigneron, Proton spectroscopic imaging of the human brain using phased array coils, Magn. Reson. Med. 34 (1995) 440–445.

[4] P.B. Barker, J. Gillen, P.C. van Zijl, X. Golay, Phased-array multi-slice proton MR spectroscopic imaging at 3 Tesla, in: Proceedings of the 11th Int. Soc. Magn. Reson. Med., 2003, p. 1133.

[5] T. Schirmer, S. Kohler, D. Gultekin, T. Raidy, Simple absolute signal scaling for spectroscopic data acquired with phased-array coils at 1.5 T, in: Proceedings of the 12th Int. Soc. Magn. Reson. Med., 2004, p. 2259.

[6] F.J. Frigo, Two-dimensional spectral estimation techniques with applications to magnetic resonance spectroscopy. PhD Dissertation, Marquette University, Milwaukee, WI, April 2004.

[7] Y. Hua, T.K. Sarkar, On SVD for estimating generalized eigenvalues of singular matrix pencil in noise, IEEE Trans. Signal Process. 39 (April) (1991) 892–900.

[8] P. Stoica, K.C. Sharman, Maximum likelihood methods for direction of arrival estimation, IEEE Trans. Acoust. Speech Signal Process. 38 (July) (1990) 1132–1143.

[9] S.Y. Kung, K.S. Arun, B. Rao, State space and singular value decomposition based on approximation methods for harmonic retrieval, J. Opt. Soc. Am. (December) (1983) 1799–1811.

[10] H. Barkhuijsen, R. de Beer, D. van Ormondt, Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals, J. Magn. Reson. 73 (1987) 553–557.

[11] S. Van Huffel, H. Chen, C. Decanniere, P. Van Hecke, Algorithm for time-domain NMR data fitting based on total least squares, J. Magn. Reson. A 110 (October) (1994) 228–237.

[12] W. Qiu, Y. Hua, Performance comparison of three methods for blind channel identification, Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP) 5 (1996) 2423–2426.

[13] H. Liu, G. Xu, L. Tong, A deterministic approach to blind identification of multi-channel FIR systems, Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP) 4 (1994) 581–584.
[14] P. Stoica, R. Moses, Introduction to Spectral Analysis, Prentice Hall, Upper Saddle River, NJ, 1997.
[15] P. Stoica, J. Li, On nonexistence of the maximum likelihood estimate in blind multichannel identification, IEEE Signal Process. Mag. 22 (May) (2005).
[16] S.W. Provencher, Estimation of metabolite concentrations from localized in vivo proton NMR spectra, Magn. Reson. Med. 30 (1993) 672–679.
[17] N. Salibi, M.A. Brown, Clinical MR Spectroscopy First Principles, Wiley, NY, 1998.
[18] B.J. Soher, R.E. Hurd, N. Sailasuta, P.B. Barker, Quantitation of automated single-voxel proton MRS using cerebral water as an internal reference, Magn. Reson. Med. 36 (1996) 335–339.