Denoising MR Spectroscopic Imaging Data with Low-Rank Approximations

Hien M. Nguyen, *Member, IEEE*, Xi Peng, Minh N. Do, *Senior Member, IEEE*,
Zhi-Pei Liang, *Fellow, IEEE*

***Abstract*—This paper addresses the denoising problem associated with magnetic resonance spectroscopic imaging (MRSI), where signal-to-noise ratio (SNR) has been a critical problem. A new scheme is proposed, which exploits two low-rank structures that exist in MRSI data, one due to partial-separability and the other due to linear predictability. Denoising is performed by arranging the measured data in appropriate matrix forms (i.e., Casorati and Hankel) and applying low-rank approximations by singular value decomposition (SVD). The proposed method has been validated using simulated and experimental data, producing encouraging results. Specifically, the method can effectively denoise MRSI data in a wide range of SNR values while preserving spatial-spectral features. The method could prove useful for denoising MRSI data and other spatial-spectral and spatial-temporal imaging data as well.**

***Index Terms*—MR spectroscopy, MR spectroscopic imaging, denoising, low-rank approximation, partially-separable functions, linear prediction**

I. INTRODUCTION

THE acquired magnetic resonance spectroscopic (MRS) signal in (k,t)-space can be expressed as

$$s(\boldsymbol{k},t)=\iint \rho(\boldsymbol{r},f)e^{-i2\pi\boldsymbol{k}\cdot\boldsymbol{r}}e^{-i2\pi ft}d\boldsymbol{r}df+\xi(\boldsymbol{k},t), \quad (1)$$

where $\boldsymbol{\rho}(\boldsymbol{r},f)$ denotes the desired spatial-spectral function and $\xi(\boldsymbol{k},t)$ is the measurement noise often modeled as a complex white Gaussian process. The function $\boldsymbol{\rho}(\boldsymbol{r},f)$ provides valuable information on the spatial-spectral distribution of metabolites, and is useful for noninvasive metabolite imaging of living systems. For example, ^{13}C MRSI can be used to study glucose metabolism [1]; ^{31}P MRSI is capable of detecting metabolites participating in tissue energy metabolism [1]; ^{1}H MRSI can map out the spatial distributions of N-Acetylaspartate (NAA), creatine, choline, and lactate metabolites that are useful for the investigation of neurological disorders [2]. However, considerable practical challenges remain in obtaining $\boldsymbol{\rho}(\boldsymbol{r},f)$ in both high spatial-spectral resolution and high SNR. These difficulties are due to acquisition time limitations and low concentrations of metabolites (typically thousands-fold below that of tissue water [3]). This paper addresses the low SNR problem.

A straightforward way to improve the SNR of MRSI data is to acquire multiple sets of measurements for signal averaging, but at the expense of lengthening the already long data acquisition time. Another approach is to apply a linear shift-invariant filter such as a Gaussian smoothing filter. However, such an approach often has poor tradeoff between spatial-spectral resolution and SNR. For a better tradeoff between SNR and resolution, many advanced denoising methods proposed for general signal processing applications can be used. Most notable in this class are transform-based methods (e.g., wavelet shrinkage [4], SVD truncation [5], [6], etc.) and PDE-based methods [7]. These methods utilize known signal properties such as piecewise smoothness for feature-preserving denoising. They are effective when the SNR is beyond some threshold, but can not well separate signal from noise in the presence of severe noise contamination, as it is often the case with practical MRSI data.

There are also several denoising methods specifically designed for MRSI signals. In [8], for instance, MRS signals are denoised by consecutive projection onto different domains, represented by a set of linear time-frequency transforms. Explicit parametric models have also been used for denoising MRSI data either prior to or during metabolite quantitation [9]–[13]. These methods are very effective when the models are correct. However, generic parametric models that fully account for all spectral features are not yet available and incorrect models often create significant bias which can be very problematic in practical applications. For example, a popular denoising method described in [9] and originally proposed in [14] makes a strict use of the Lorentzian lineshape of the spectral peaks and is very sensitive to noise. Less-constrained approaches were proposed in [15], [16].

In this work, we propose a new scheme for MRSI denoising, coined LORA (LOw Rank Approximations), by utilizing the low-rank structures of the spatial-spectral data. Specifically, we incorporate low-rank approximation in $(\boldsymbol{k},t)$ domain by assuming that spatial variations are separable from temporal variations to some low order [17]. In addition, we exploit the low-rank structure of the temporal signal due to its linear

H. M. Nguyen is with the Lucas Center for MR Spectroscopy and Imaging, Department of Radiology, Stanford University, Palo Alto CA 94304 (email: hnguyen6@stanford.edu).

M. N. Do is with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana IL 61801 (email: minhdo@illinois.edu).

X. Peng is with the Paul C. Lauterbur Research Center for Biomedical Imaging, Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China. During this work, X. Peng was a visiting scholar in Beckman Institute, University of Illinois at Urbana-Champaign, Urbana IL 61801 (email: xi.peng@siat.ac.cn).

Z.-P. Liang is with the Department of Electrical and Computer Engineering and the Beckman Institute, University of Illinois at Urbana-Champaign, Urbana IL 61801, (email: z-liang@illinois.edu).

This work was supported in part by the following research grants: NIH-P41-EB-001977, NIH-P41-EB015904, NIH-R01-EB013695, NSF-CBET-07-30623, NSF-CCF-06-35234, NSF-CCF-09-16953, and VEF fellowship.

predictability. The use of these two low-rank properties provides efficient spatial-spectral filtering. Our preliminary results were reported in [18]. This paper presents a more detailed description and a more thorough evaluation of the proposed method.

The rest of the paper is organized as follows. Section II presents the proposed method, including a description of the low-rank properties and the associated denoising algorithm. Section III discusses algorithm considerations. Section IV shows simulation and experimental results illustrating denoising performance of the proposed algorithm, followed by the conclusions of the paper in Section V. For easy reference, the following is a list of the key symbols used in this paper.

$\rho(\boldsymbol{r}, f)$:	Spatial-spectral function of a spin system
$\rho(f)$:	MRSI spectrum at an arbitrary spatial location
$s_0(\boldsymbol{k}, t)$:	Ideal noiseless MRSI signals
$s(\boldsymbol{k}, t)$:	Measured MRSI signals
$\xi(\boldsymbol{k}, t)$:	Noise in the measured MRSI signals
$s(t)$:	MRSI signal at an arbitrary spatial location
L_1:	Order of the partially separable (PS) model
L_2:	Order of the linear predictive (LP) model
$c_l(\boldsymbol{r})$:	Spatial basis of the PS model
$\psi_l(f)$:	Spectral basis of the PS model
$\hat{c}_l(\boldsymbol{k})$:	Fourier transform of $c_l(\boldsymbol{r})$
$\hat{\psi}_l(t)$:	Inverse Fourier transform of $\psi_l(f)$
α_l:	Coefficients of the mixture of Lorentzians model
φ_l:	Lorentzian spectral function
$\boldsymbol{C}$:	Casorati matrix
$\boldsymbol{H}$:	Hankel matrix
$\boldsymbol{A}$:	A general matrix
λ:	Singular values of a matrix

II. PROPOSED METHOD

The proposed method is based on low-rank approximations of MRSI data. Specifically, it exploits two low-rank properties (one due to partial separability and the other due to linear predictability) for noise removal.

A. Low-rankness due to spatiotemporal partial separability

The spatial-spectral distribution function $\rho(\boldsymbol{r}, f)$ can be expressed as

$$\rho(\boldsymbol{r}, f) = \sum_{l=1}^{L_1} c_l(\boldsymbol{r})\psi_l(f), \tag{2}$$

where $c_l(\boldsymbol{r})$ can be viewed as the spatial basis and $\psi_l(f)$ as the spectral basis for $\rho(\boldsymbol{r}, f)$. This model is called L_1th-order partially separable (between space and frequency). Equivalently, the noiseless MRSI data $s_0(\boldsymbol{k}, t)$ can be expressed as partially separable between $\boldsymbol{k}$ and t to the L_1th order:

$$s_0(\boldsymbol{k}, t) = \sum_{l=1}^{L_1} \hat{c}_l(\boldsymbol{k})\hat{\psi}_l(t), \tag{3}$$

where $\hat{c}_l(\boldsymbol{k})$ and $\hat{\psi}_l(t)$ are related to $c_l(\boldsymbol{r})$ and $\psi_l(f)$ in (2) by the Fourier transform, respectively. In (3) we assume that field inhomogeneity effects have been previously removed. This task can be achieved using a range of field-correction techniques, such as those in [19]–[21].

Validity of the PS model (2) can be justified as follows. Noting that $\rho(\boldsymbol{r}, f)$ is an $\mathcal{L}_2$-function, model (2) is always valid for $L_1 = \infty$. In practice, the PS model is valid for a finite (small) L_1 because there is a finite number of resonances (spectral components) in any practical MRS experiment. In this case, function $\psi_l(f)$ can be viewed as the spectral function of the l-th resonance component and $c_l(\boldsymbol{r})$ is its corresponding spatial distribution.

Model (2) also arises when $\rho(\boldsymbol{r}, f)$ can be decomposed into a summation of compartmental spectral functions. This simplified form of the PS model was previously used in [22] for spectroscopic imaging. In this case L_1 is the number of compartments, $c_l(\boldsymbol{r}) = 1$, and

$$\psi_l(f) = \begin{cases} \frac{1}{V_l}\int \rho(\boldsymbol{r}, f)d\boldsymbol{r}, & \boldsymbol{r} \in \mathcal{D}_l \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

where $\mathcal{D}_l$ represents the l-th compartment and V_l is its corresponding volume. While compartmental spatial-spectral functions are partially-separable, the PS model in (2) represents a much broader class of functions and does not impose compartmental homogeneity as in [22], which is often problematic for practical MRSI data [23].

An important property of model (3) is that the Casorati matrix formed from $s_0(\boldsymbol{k}, t)$ samples is low-rank. Specifically, let

$$\boldsymbol{C}_0 = \begin{bmatrix} s_0(\boldsymbol{k}_1, t_1) & s_0(\boldsymbol{k}_1, t_2) & \dots & s_0(\boldsymbol{k}_1, t_M) \\ s_0(\boldsymbol{k}_2, t_1) & s_0(\boldsymbol{k}_2, t_2) & \dots & s_0(\boldsymbol{k}_2, t_M) \\ & \cdots & & \\ s_0(\boldsymbol{k}_N, t_1) & s_0(\boldsymbol{k}_N, t_2) & \dots & s_0(\boldsymbol{k}_N, t_M) \end{bmatrix} \tag{5}$$

for any point set $\{s_0(\boldsymbol{k}_n, t_m)\}_{n=1,m=1}^{N,M}$. Then, model (3) implies that $\boldsymbol{C}_0$ has at most rank L_1 [17].

In practice, L_1 is much smaller than $\min(N, M)$ due to the small number of spectral components (about $5 \sim 10$ commonly MR-observable metabolites, 4 resonance components from macromolecules, and 5 components from lipids in the human brain [24], [25]). This enables the use of the low-rank property for effective denoising.

B. Low-rankness due to linear predictability

The time-domain signal of a spin system with L_2 spectral components resonating at frequency f_l with damping factor γ_l can be expressed as [3]

$$s_0(t) = \sum_{l=1}^{L_2} \alpha_l e^{-(\gamma_l + j2\pi f_l)t}. \tag{6}$$

Equivalently in the frequency domain, the spectrum consists of L_2 Lorentzian resonance lines $\varphi_l(f)$:

$$\rho(f) = \sum_{l=1}^{L_2} \alpha_l \varphi_l(f),$$

where

$$\varphi_l(f) = \frac{1/\gamma_l}{1 + 4\pi^2(f + f_l)^2/\gamma_l^2} - j\frac{2\pi(f + f_l)/\gamma_l^2}{1 + 4\pi^2(f + f_l)^2/\gamma_l^2}. \tag{7}$$

In practice, resonance lines $\varphi_l(f)$ may deviate from the Lorentzian lineshape due to magnetic field inhomogeneity

and limited spatial resolution, etc. So, an associated practical question is whether the observed resonance lines can be well represented using (7). This question was previously raised in [3], [26] and it was found that any observed lines can be fitted using a linear combination of Lorentzian basis functions. For example, the broad resonance from residual water was found to be well-represented using $3 \sim 10$ Lorentzians [3]. Therefore, (7) is valid (in a mathematical sense), although in this case L_2 does not physically represent the number of spectral peaks. In practice, L_2 is larger than the number of spectral peaks but much smaller than the number of available samples M.

An important property of model (6) is that the discrete time-domain signal is linearly predictable, i.e.,

$$s_0[m] = \sum_{l=1}^{L_2} \beta_l s_0[m-l], \tag{8}$$

where $s_0[m] = s_0(m\Delta t)$ with Δt being the sampling interval. A Hankel matrix formed from $s_0[m]$ is low-rank. Specifically, $\{s_0[m]\}_{m=1}^{M}$ is L_2th-order linearly predictable as defined in (8) if and only if the following Hankel matrix

$$\boldsymbol{H}_0 = \begin{bmatrix} s_0[1] & \dots & s_0[K] \\ s_0[2] & \dots & s_0[K+1] \\ & \cdots & \\ s_0[M-K+1] & \dots & s_0[M] \end{bmatrix} \tag{9}$$

has rank L_2 [27].

C. Proposed algorithm

For the measured noisy data $s(\boldsymbol{k},t) = s_0(\boldsymbol{k},t) + \xi(\boldsymbol{k},t)$, the corresponding Casorati and Hankel matrices in (5) and (9) can be expressed respectively as $\boldsymbol{C} = \boldsymbol{C}_0 + \boldsymbol{E}_C$ and $\boldsymbol{H} = \boldsymbol{H}_0 + \boldsymbol{E}_H$, where $\boldsymbol{E}_C$ and $\boldsymbol{E}_H$ denote noise matrices. The proposed algorithm reduces $\xi(\boldsymbol{k},t)$ by performing low-rank approximation on $\boldsymbol{C}$ and $\boldsymbol{H}$, respectively.

1) Low-rank matrix approximation

Assume that an arbitrary rank-L matrix $\boldsymbol{A}_0$ is perturbed by noise (denoted by a random matrix $\boldsymbol{E}$) as

$$\boldsymbol{A} = \boldsymbol{A}_0 + \boldsymbol{E}. \tag{10}$$

In the case when rank L is known *a priori*, the low-rank approximation of $\boldsymbol{A}$ is given by

$$\bar{\boldsymbol{A}} = \arg \min_{\text{rank}(\hat{\boldsymbol{A}})=L} ||\boldsymbol{A} - \hat{\boldsymbol{A}}||, \tag{11}$$

where the norm $||\cdot||$ denotes either spectral or Frobenius norm. It is well known that $\bar{\boldsymbol{A}}$ can be obtained using singular value decomposition (SVD) as

$$\bar{\boldsymbol{A}} = \sum_{l=1}^{L} \lambda_l(\boldsymbol{A}) \boldsymbol{u}_l \boldsymbol{v}_l^H, \tag{12}$$

where λ_l, $\boldsymbol{u}_l$, and $\boldsymbol{v}_l$ are the singular values, left singular vectors, and right singular vectors of $\boldsymbol{A}$, respectively.

2) Rank determination

Case 1: $\boldsymbol{A}_0$ is a general $N \times M$ low-rank matrix and $\boldsymbol{E}$ is an $N \times M$ random Gaussian matrix. In this case, L can be determined using the results from random matrix theory. Specifically, we first estimate noise standard deviation σ_0 from the measured data $s(\boldsymbol{k}_n, t_m)$ and use σ_0 to estimate $||\boldsymbol{E}||_2$ based on the Marchenko-Pastur distribution of the eigenvalues of $\boldsymbol{E}^H\boldsymbol{E}$ [28]. We then choose rank $\hat{L}$ so that

$$\lambda_{\hat{L}+1}(\boldsymbol{A}) \le ||\boldsymbol{E}||_2 \le \lambda_{\hat{L}}(\boldsymbol{A}). \tag{13}$$

Case 2: $\boldsymbol{A}_0$ is a low-rank Hankel matrix and $\boldsymbol{E}$ is a Gaussian Hankel matrix. The rank selection problem in this case has been well studied and is related to the model order estimation problem for autoregressive models. The rank $\hat{L}$ can be determined using various criteria such as the Akaike Information Criterion (AIC) [29]. Specifically, $\hat{L}$ is chosen to yield a minimum change in $AIC(\hat{L}) - AIC(\hat{L}+1)$, where

$$AIC(\hat{L}) = M \log e(\hat{L}) + 2\hat{L} \tag{14}$$

and $e(\hat{L})$ is the error of a least squares fit of model (6) to the noisy data, given a particular candidate value $\hat{L}$.

3) Denoising by LOw-Rank Approximations (LORA)

The proposed denoising algorithm consists of two key steps: PS-based low-rank filtering and LP-based low-rank filtering, which are summarized below:

1) Given noisy data $s(\boldsymbol{k}_n, t_m)$, construct matrix $\boldsymbol{C}$ according to (5) and solve the following optimization problem:
$$\bar{\boldsymbol{C}} = \arg \min_{\text{rank}(\hat{\boldsymbol{C}})=L_1} ||\boldsymbol{C} - \hat{\boldsymbol{C}}|| \tag{15}$$
by performing rank-L_1 approximation on $\boldsymbol{C}$ using SVD.
2) Take the 2-D discrete Fourier transform (DFT) of each column of matrix $\bar{\boldsymbol{C}}$ to obtain $\tilde{s}(\boldsymbol{r}_n, t_m)$.
3) For each voxel $\boldsymbol{r}_n$, construct matrix $\boldsymbol{H}$ from $\tilde{s}(\boldsymbol{r}_n, t_m)$ according to (9) and solve the following optimization problem:
$$\bar{\boldsymbol{H}} = \arg \min_{\text{rank}(\hat{\boldsymbol{H}})=L_2} ||\boldsymbol{H} - \hat{\boldsymbol{H}}|| \tag{16}$$
by performing rank-L_2 approximation using SVD.
4) Denoised data at each voxel $\boldsymbol{r}_n$ is obtained by extracting the elements from the first row and last column of $\bar{\boldsymbol{H}}$.

III. ALGORITHM CONSIDERATIONS

A. Improved LP-based low-rank filtering

Our empirical results indicate that PS-based low-rank filtering can be applied in a wide range of SNRs, while the performance of LP-based low-rank filtering is rather sensitive to the noise level. One reason for the high sensitivity of LP-based low-rank filtering is that the signal subspace in the Hankel matrix $\boldsymbol{H}$ is usually not well-conditioned and is not well separated from the noise subspace. This is a well-known limitation with LP-based analysis of MRSI data [9], [14]. We have observed that for the low SNR values in typical in-vivo MRSI data, the LP-based low-rank filtering step alone often resulted in spectral artifacts such as loss of signal components and introduction of spurious peaks. This is especially the case if both low-rank and Hankel structures are imposed on $\hat{\boldsymbol{H}}$ as it is done in the Cadzow signal enhancement algorithm [14], which has been used for denoising [4], [9], [10]. By not imposing the Hankel structure on $\hat{\boldsymbol{H}}$ in step 3 of the proposed algorithm, the "rank" constraint is actually weakly

imposed. This "softening" of the LP-based low-rank filtering seems to enable the algorithm to accommodate non-Lorentzian lines and avoid spectral artifacts in low-SNR cases. In addition, it is important that the PS-based low-rank filtering is applied first, which is followed by the LP-based low-rank filtering. In this way, the data SNR is improved before LP-based low-rank filtering is applied. Another step that can further help to reduce the noise sensitivity of LP-based low-rank filtering is to perform "segmented" low-rank filtering. Specifically, instead of forming a single Hankel matrix using the entire signal, we can divide a given signal $\{s[m]\}_{m=0}^{N-1}$ into two (or multiple) segments, say $\{s[m]\}_{m=0}^{N/2-1}$ and $\{s[m]\}_{m=N/2}^{N-1}$, and then form two (or multiple) Hankel matrices, each of which will be processed using the proposed LP-based low-rank filtering.

Figure 1 shows a set of simulation results to illustrate the sensitivity (to noise and model order) of LP-based low-rank filtering. Figure 1(a) shows ground truth (red) and noisy (black) spectra at two SNR levels. The denoised spectra in Figs. 1(b)-(e) were obtained using LP-based low-rank filtering with the Hankel constraint (as is done in Cadzow [14]) and without the Hankel constraint (step 2 of LORA), respectively, for two model orders ($L_2 = 20$ and 8). Note that in the high-SNR case, the Hankel constraint helped to produce almost "perfect" denoised result with a correct model order (top row, (b)) but lost some spectral components with an under-estimated model order (top row, (d)). In the low SNR case, the Hankel constraint introduced spurious spectral features (bottom row, (b)) even with the correct model order. The problem can be alleviated with a small model order but at the expense of missing spectral components (such as the loss of the smallest peak at 6.05 ppm and over-smoothing of the non-Lorentzian, closely-spaced spectral peaks in the regions from 2.2 ppm to 2.8 ppm and from 3.3 ppm to 3.8 ppm, see bottom row in (d)). Both problems (spurious peaks and missing peaks) were alleviated by not imposing the Hankel structure (see (e)).

B. Rank selection

For PS-based low-rank filtering, in the case of high/moderate SNR (ideally as long as $||\boldsymbol{E}||_2 \leq \lambda_L(\boldsymbol{A_0})$ or in the range above 18 dB for the simulated ^1H MRSI dataset described in Section IV) we choose rank L_1 of $\boldsymbol{C}_0$ according to (13). To evaluate the performance of the described rank selection method, we performed rank selection on the simulated dataset with true rank $L_1 = 8$. Specifically, we performed a Monte-Carlo study with 9 different noise levels, 32 noise realizations per noise level and estimated $||\boldsymbol{E}||_2$ using the Marchenko-Pastur distribution ([28]). The effective rank $\hat{L}_1$ was selected according to (13). Figure 7(a) shows some representative noisy spectra at the highest, medium, and lowest SNRs used in the experiment. For this experiment and the rest of the paper, we define SNR in terms of signal energy and signal amplitude as

$$\begin{aligned} \text{SNR}_\text{e} &= 10\log_{10}\frac{||s_0||_2}{||s-s_0||_2}, \\ \text{SNR}_\text{p} &= \frac{\text{amplitude of the largest peak}}{\text{standard deviation of the noise}}, \end{aligned} \qquad (17)$$

where s and s_0 are the noisy and noiseless signals, respectively.

Table I shows the mean effective rank $\bar{L}_1$, averaged over all the noise realizations. Notice that $||\boldsymbol{E}||_2$ is closely estimated using the Marchenko-Pastur distribution at every SNR level, as can be seen from the reported relative error $\frac{||\boldsymbol{E}||_2 - ||\hat{\boldsymbol{E}}||_2}{||\boldsymbol{E}||_2}$. The results also show that as the SNR decreases, the discussed approach under-estimates L_1. In practice, the choice of L_1 can also be guided by the known number of spectral components present in the spectrum [30].

For LP-based low-rank filtering, the choice of rank L_2 of matrix $\boldsymbol{H}_0$ is based on the AIC criteria as specified in (14). We performed a Monte-Carlo study on rank L_2 selection with the simulated dataset at 3 different noise levels, for each noise level 256 noise realizations were considered. The corresponding representative noisy spectra are shown in Fig. 8(a). Figure 2 shows the computed mean AIC values (averaged over the noise realizations) as a function of candidate rank values. The effective rank $\hat{L}_2$ is typically chosen to minimize AIC. However, this usually leads to an over-estimated rank [29]. To achieve a closer estimate, we suggest choosing the rank to be the threshold point at which there is no significant reduction in the AIC value when the rank is further increased. In practice, the choice of L_2 can also be guided by the known number of spectral peaks present in the spectrum.

Inaccurate rank estimation is more problematic for LP-based low-rank filtering than for PS-based low-rank filtering. Specifically, an over-estimated rank L_2 ($\hat{L}_2 > L_2$) can result in spurious peaks. Using an under-estimated rank (without imposing the Hankel structure) can alleviate the problem. However, a highly under-estimated rank $\hat{L}_2$ can lead to a loss of useful spectral information, which is undesirable. Typically, peaks with small amplitudes are distorted or lost first [31]. Possible peak loss due to rank under-estimation can be alleviated by not imposing the Hankel constraint as discussed in Section III-A. For PS-based low-rank filtering, severe rank under-estimation can lead to spatial blurring and spectral line broadening. An over-estimated rank L_1 in this case does not directly introduce bias and spurious peaks, but would reduce its filtering effectiveness.

IV. RESULTS AND DISCUSSIONS

In this section, results from simulated and experimental data are presented to demonstrate the denoising performance of the proposed technique.

A. Simulations

We simulate a spatial-spectral distribution function based on literature values of ^1H metabolites and experimental data. We consider 5 commonly MR-observable metabolites in the human brain [24], [30]: *N-acetylaspartate (NAA), creatine (Cr), choline (Cho), glutamate/glutamine (Glx),* and *myo-inositol (m-Ins).* For each metabolite, a spectral profile was obtained from quantum mechanical simulations of a spin-echo MR experiment [32]. The spatial distribution of each metabolite was created based on commonly reported literature values from 3 segmented regions of cerebrospinal fluid, grey matter

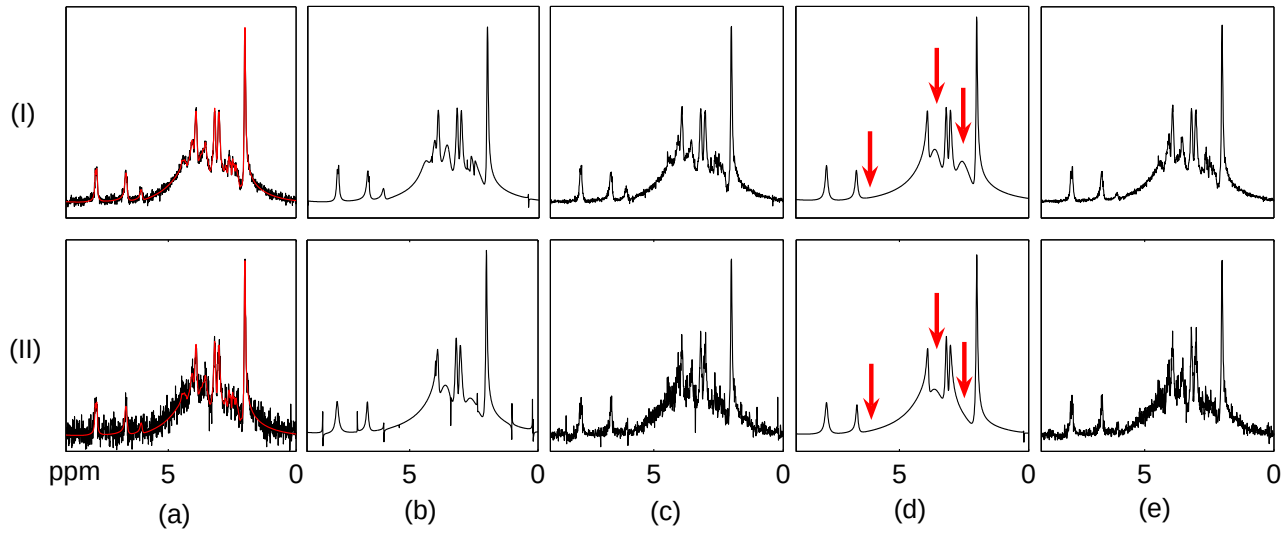


Fig. 1. Simulation results of LP-based low-rank filtering with and without Hankel constraint: (a) original (red) and noisy (black) spectra at low (I) and high (II) noise levels; (b) and (d) denoised spectra with Hankel constraint (Cadzow) for $L_2 = 20$ and 8, respectively; (c) and (e) denoised spectra without Hankel constraint (step 2 of LORA) for $L_2 = 20$ and 8, respectively. Note that the Hankel constraint helped reduce noise but at the expense of spectral artifacts.

TABLE I

MONTE-CARLO EXPERIMENT ON CHOOSING EFFECTIVE RANK $\bar{L}_1$ AT 9 DIFFERENT SNR LEVELS. TRUE RANK $L_1 = 8$. VOXEL 1 IS IN THE REGION OF LOW-SNR, WHILE VOXEL 2 IS IN THE REGION OF HIGH-SNR. ALL SNR VALUES ARE REPORTED IN DECIBELS.

$\text{SNR}_\text{e}(s(\boldsymbol{k},t))$	23.72	22.23	20.94	19.83	18.85	17.97	17.17	16.45	15.78
$\text{SNR}_\text{e}(s(f))$ at voxel 1	6.42	5.17	4.45	3.83	3.26	2.83	2.27	1.99	1.86
$\text{SNR}_\text{e}(s(f))$ at voxel 2	32.46	31.31	29.96	28.52	27.82	27.01	26.22	25.41	24.83
$\frac{\|\|\boldsymbol{E}\|\|_2 - \|\|\hat{\boldsymbol{E}}\|\|_2}{\|\|\boldsymbol{E}\|\|_2}$	0.0057	0.0051	0.0053	0.0058	0.0046	0.0050	0.0054	0.0050	0.0058
$\bar{L}_1$	8	8	7.03	6.63	6.06	6.00	5.25	5.03	4.97

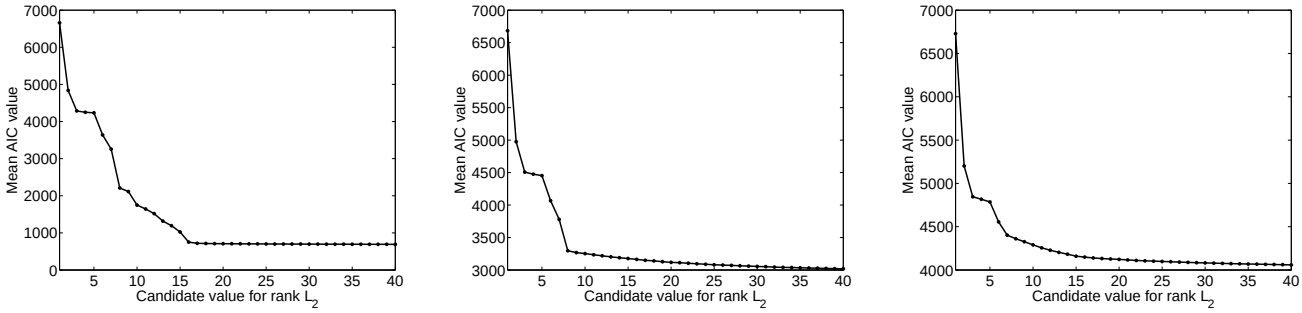


Fig. 2. Monte-Carlo study of rank L_2 selection based on AIC: top row shows mean AIC value (averaged over 256 noise realizations) as a function of $\hat{L}_2$ at (a) low noise level with SNR_e=26.61 dB, SNR_p=94.4; (b) medium noise level with SNR_e=17.03 dB, SNR_p=30.4; (c) high noise level with SNR_e=12.76 dB, SNR_p=17.9. The true rank L_2 in this experiment was 20.

(GM), and white matter (WM) [30], [33], [34]. To evaluate the ability of the proposed method to handle non-ideal conditions, Gaussian (instead of pure Lorentzian) lineshape was used with an additional baseline signal. Lineshape parameters were chosen based on commonly reported values in the literature [35], [36]. The baseline signal was extracted from single-voxel CSI PRESS experimental data of the brain with TE=30 ms. MRSI data were simulated at the magnetic field strength of 3 T with a spectral bandwidth of $1,200$ Hz and complex Gaussian noise was added.

Figures 3 and 4 show the spatial distribution of the spatial-spectral function at a high-SNR frequency of 3.94 ppm and a low-SNR frequency of 6.66 ppm, both indicated in Fig. 5(f) as "f_1" and "f_2", respectively. We compare the proposed approach with 3D wavelet soft shrinkage (Daubechies 4-tap kernel filter, 4 levels), sparse 3D transform-domain collaborative filtering [37] (code is available at http://www.cs.tut.fi/~foi/GCF-BM3D/), and conventional Gaussian apodization. The wavelet shrinkage threshold and apodization constant were chosen to yield the same level of the residual noise variance as that for the proposed approach (threshold $T = 2.28\sigma_0$). It can be clearly seen that the corresponding wavelet-denoised

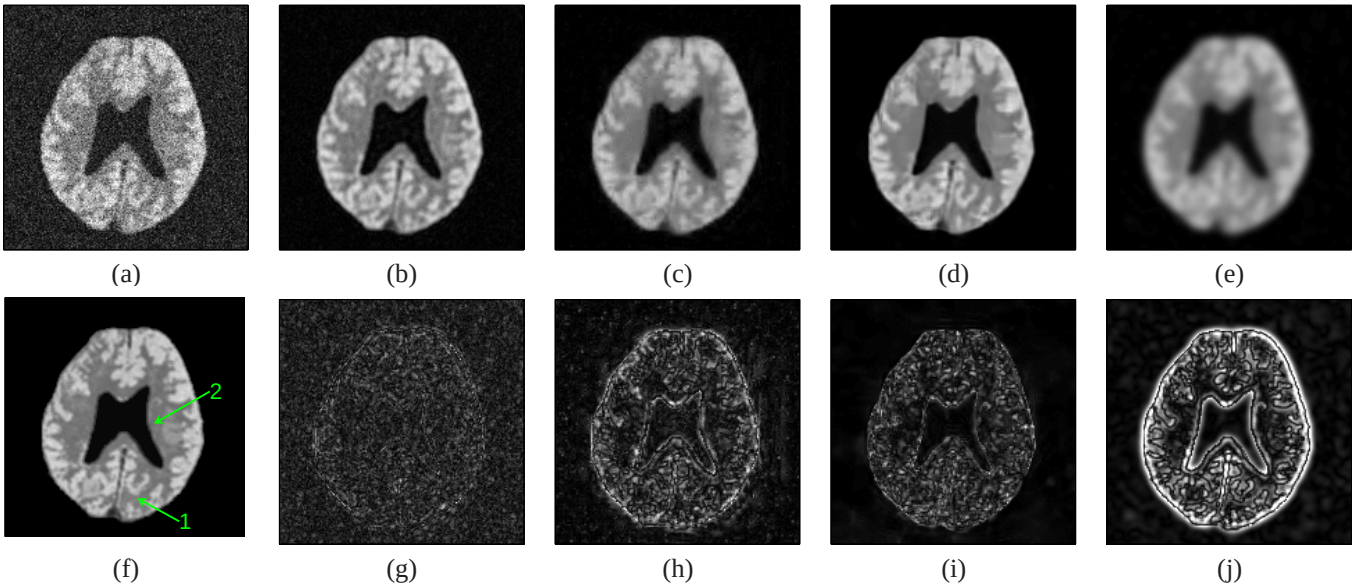


Fig. 3. Denoising results - spatial distributions $\rho(\boldsymbol{r}, f_1)$ at a high-SNR frequency (3.94 ppm, see Fig. 5(f)) obtained from (a) noisy data; (b) LORA denoising; (c) wavelet denoising; (d) sparse 3D transform-domain collaborative filtering; (e) Gaussian apodization; and (f) noiseless data. Corresponding errors of (b)-(e) are shown in (g)-(j), respectively.

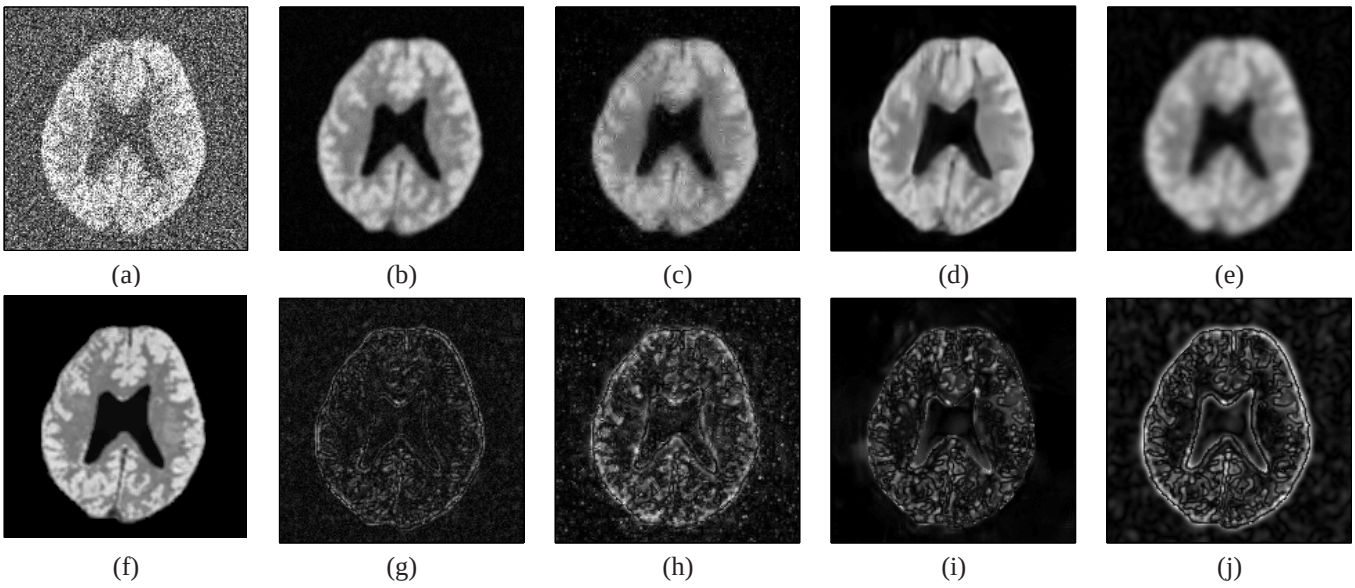


Fig. 4. Denoising results - spatial distributions $\rho(\boldsymbol{r}, f_2)$ at a low-SNR frequency (6.66 ppm, see Fig. 5(f)), obtained from (a) noisy data; (b) LORA denoising; (c) wavelet denoising; (d) sparse 3D transform-domain collaborative filtering; (e) Gaussian apodization; and (f) noiseless data. Corresponding errors of (b)-(e) are shown in (g)-(j), respectively.

spatial-spectral function in Fig. 3(c) and Fig. 4(c) has an improved SNR at the expense of blurring the spatial features significantly, while the proposed denoising preserves the spatial features better. Collaborative filtering performed well at a high-SNR frequency such as the one shown in Fig. 3, but at a lower-SNR freqency shown in Fig. 4 the spatial features such as edges in the GM region were smoothed out noticeably by the collaborative filtering while the residual noise variance in the denoised spectrum was larger than that of obtained from LORA (see Fig. 5(d)). Gaussian apodization shown in Fig. 3(e) and Fig. 4(e) has the poorest trade-off between SNR and resolution among all methods, as expected.

Figure 5 shows representative spectra from a particular voxel marked as "1" in Fig. 3(f). Both Gaussian apodization and wavelet denoising improve SNR at an expense of line broadening and loosing metabolite amplitudes, with Gaussian apodization performing the worst. This amplitude loss can be seen from the corresponding residual spectra, which are not centered at zero, as shown in Figs. 5(h),(j).

B. Denoising performance

To analyze the denoising performance of the proposed low-rank filtering, we rewrite $\boldsymbol{C}$ and $\bar{\boldsymbol{C}}$ as $\boldsymbol{C} = \boldsymbol{C_0} + \boldsymbol{E}$ and $\bar{\boldsymbol{C}} = \boldsymbol{C_0} + \boldsymbol{\Delta}$, where $\boldsymbol{C_0}$ is an $N \times M$ data matrix with rank L_1 as described in (5), $\boldsymbol{E}$ is the noise matrix, $\bar{\boldsymbol{C}}$ is formed from

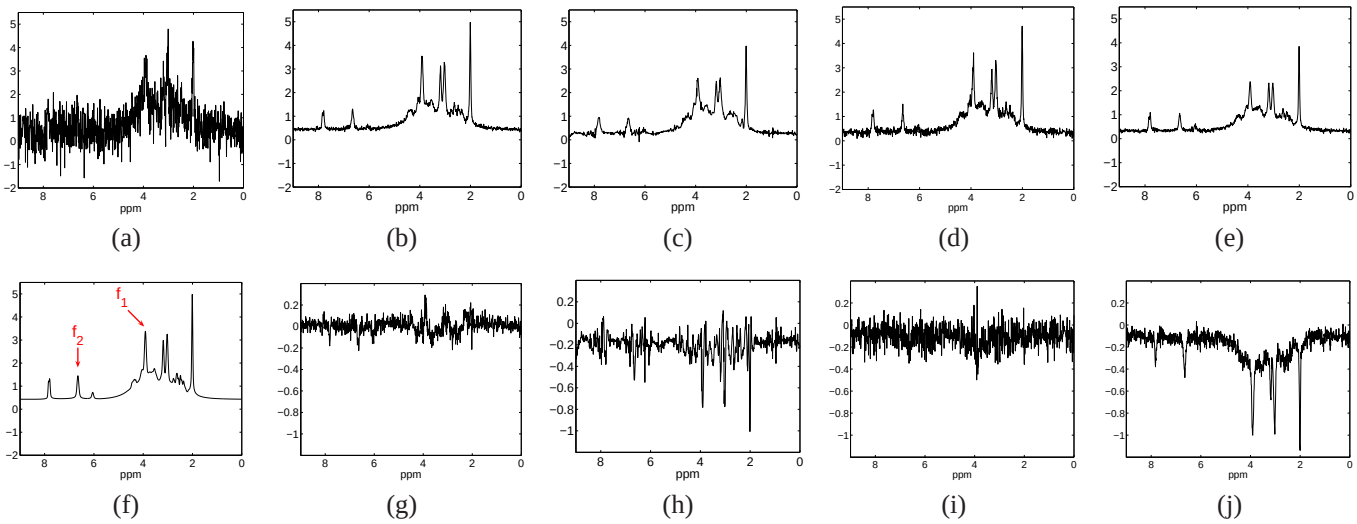


Fig. 5. Denoising results - spectra at a particular voxel, marked as "1" in Fig. 3(f), obtained from (a) noisy data; (b) LORA denoising; (c) wavelet shrinkage; (d) sparse 3D transform-domain collaborative filtering; (e) Gaussian apodization; (f) noiseless data. Corresponding errors of (b)-(e) are shown in (g)-(j), respectively. All spectra are shown in the real-valued mode.

the denoised data as obtained from (15), and $\boldsymbol{\Delta}$ is the residual error, which can contain both residual noise and possible signal loss due to low-rank filtering. Without loss of generality, we assume $M < N$. The denoised matrix $\bar{\boldsymbol{C}}$ is obtained according to (15). We define the noise reduction factor as

$$g = \frac{||\boldsymbol{E}||_F}{||\boldsymbol{\Delta}||_F}. \tag{18}$$

By defining the SNR before and after denoising as $20\log_{10}\frac{\mathrm{P_s}}{||\boldsymbol{E}||_F}$ and $20\log_{10}\frac{\mathrm{P_s}}{||\boldsymbol{\Delta}||_F}$, where $\mathrm{P_s}$ denotes the signal power, the SNR gain can be calculated as $20\log_{10} g$. Ideally, one would want $g = \infty$. In the full-rank case when $L_1 = M$ and no denoising is achieved, g equals to 1.

Computing g as in (18) requires knowing the ground truth in order to compute $||\boldsymbol{\Delta}||_F$. However, notice that we only need to estimate the norm $||\boldsymbol{\Delta}||_F$ rather than knowing $\boldsymbol{\Delta}$ completely. Thus, one would hope that there is a signal-independent expression to predict $||\boldsymbol{\Delta}||_F$ closely. In general, it is hard to accurately characterize singular values of $\boldsymbol{C}_0$, $\boldsymbol{C}$, and $\bar{\boldsymbol{C}}$. However, under assumptions that:

Cond1: the signal and noise (before and after denoising) subspaces are uncorrelated in the sense that $\boldsymbol{C_0}^H\boldsymbol{E} = \boldsymbol{0}$ and $\boldsymbol{C_0}^H\boldsymbol{\Delta} = \boldsymbol{0}$;

Cond2: the noise is complex white, i.e. $\boldsymbol{E}^H\boldsymbol{E}$ and $\boldsymbol{\Delta}^H\boldsymbol{\Delta}$ are scalar multiples of identity matrix [38], we can obtain the following approximate formula based on the theoretical results of SVD of noisy matrices in [38]:

$$g = \sqrt{\frac{\sum_{i=1}^{M}\lambda_i^2(\boldsymbol{E})}{\sum_{i=1}^{\hat{L}_1}\lambda_i^2(\boldsymbol{E})}}, \quad \text{if } \hat{L}_1 \geq L_1. \tag{19}$$

Expression (19) provides a signal-independent prediction of g. It suggests that (i) in the case of accurate rank estimation ($\hat{L}_1 = L_1$) and $\lambda_i(\boldsymbol{E})$ are all equal, the noise reduction factor is on the order of $\sqrt{\frac{M}{\hat{L}_1}}$; (ii) in the case of over-estimated rank ($\hat{L}_1 > L_1$), the residual noise level is larger than the residual noise level with correct rank determination. In practice, the conditions (Cond1) and (Cond2) which are required for (19) to hold are never satisfied exactly. However, we observed that a mild violation of these conditions may still give a rough guess of g, as shown in our simulation study discussed below.

Based on the simulated ^1H MRSI dataset described in Section IV-A, we performed a Monte-Carlo study to compute g according to (18), averaged over 32 noise realizations. The noise reduction factor g is plotted as a function of the estimated rank $\hat{L}_1$ in Fig. 6. We observe that when the rank is properly chosen, g reaches its maximum value (about 7.6 for the simulated dataset). As the rank is over-estimated ($\hat{L}_1 > L_1$) and approaches the full rank, g decreases to 1, as expected. The reduction of g reflects the residual noise variance in the over-estimated rank case and the signal loss in the under-estimated rank case. A nice feature of this plot is that it shows how much of the noise reduction we can achieve even in the case when the rank is incorrectly estimated. For example, if the rank was over-estimated to be 32, while the true rank is 8, then according to this plot, we still achieve a factor of about 3 (or 9.5 dB) of improvement in the SNR. Thus, for the sake of improving SNR, estimating rank precisely is not a critical point, as long as we do not under-estimate the rank causing the signal loss. Note that with the considered dataset in Fig. 6, the relative error $\frac{||\boldsymbol{E}^H\boldsymbol{E}-M\sigma_0^2\boldsymbol{I}||_F}{M\sigma_0} = 0.99$ for condition (Cond1) and $\frac{||\boldsymbol{C}_0^H\boldsymbol{E}||_F}{\sqrt{||\boldsymbol{C}_0^H\boldsymbol{C}_0||_F||\boldsymbol{E}^H\boldsymbol{E}||_F}} = 0.15$ for condition (Cond2). These numbers did not significantly change as the SNR was varied from low to high (see Fig. 7(a) for representative SNR levels). Conditions (Cond1) and (Cond2) for $\boldsymbol{\Delta}$ are satisfied closer as $\hat{L}$ approaches M. Predicted value from (19) is plotted in the same figure as the empirical value of g in Fig. 6. Notice that in the region $\hat{L}_1 > L_1$ the theoretically predicted values of g approach its empirical value as $\hat{L}_1$ tends to the full-rank case.

We have performed a similar analysis on the noise reduction factor g, for the case of low-rank filtering of linear prediction matrix. However, to evaluate the level of SNR improvement

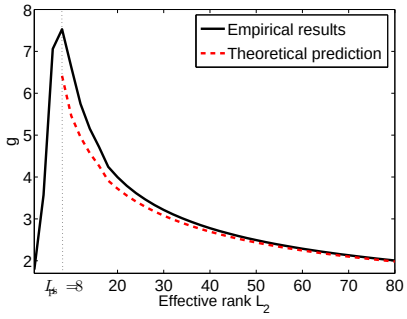
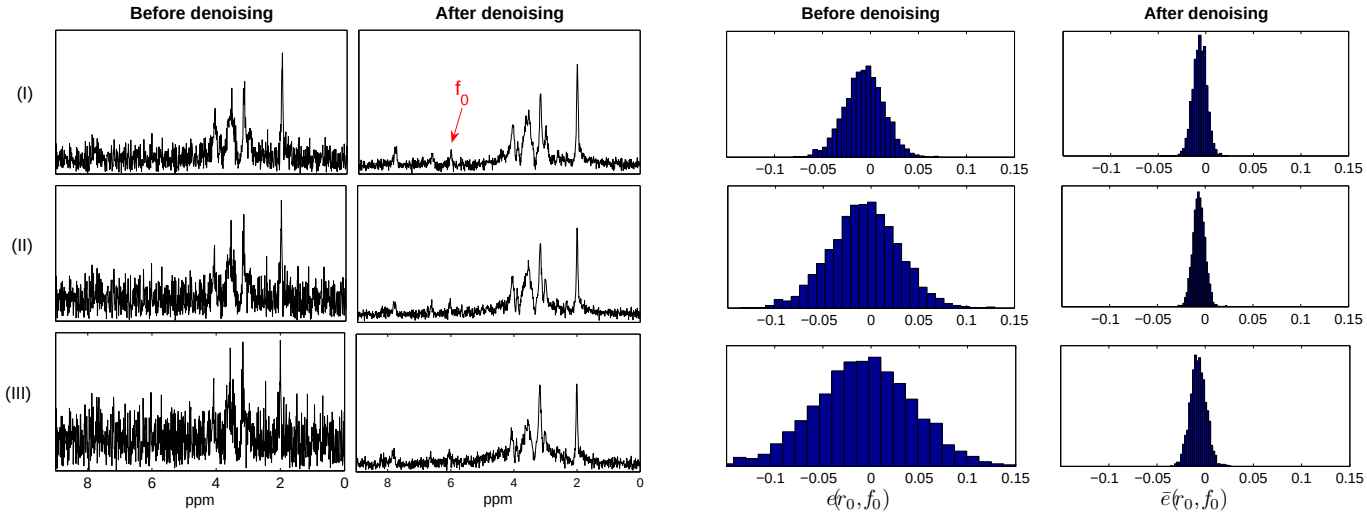


Fig. 6. Noise reduction factor g for low-rank filtering based on partial separability as a function of the estimated rank$\hat{L}_1$. Empirical values of g (black) and corresponding theoretical approximations (red) were computed according to (18) and (19), respectively. The true rank $L_1 = 8$. Notice that in the region $\hat{L}_1 > L_1$ the theoretically predicted values of g tend to its empirical values and both approach 1 in the full-rank case.



(a) (b)

Fig. 7. Monte-Carlo study of low-rank filtering based on partial separability: (a) representative noisy and denoised spectra at three SNR levels; (b) error distributions calculated as the histograms of e and $\bar{e}$ according to (21) at a particular location (r_0, f_0). For this illustrative example, r_0 is a representative voxel in the white matter region marked as "2" in Fig. 3(e) and f_0 is the frequency point marked in Fig. 7(a). Notice the reduced noise variances achieved by denoising.

for low-rank filtering of linear prediction matrix, we rather define the following factor:

$$h = \frac{||\boldsymbol{\xi}||_2}{||\boldsymbol{\eta}||_2}, \tag{20}$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are vectors extracted from the first row and last column of $\boldsymbol{E}$ and $\boldsymbol{\Delta}$, respectively. We performed a Monte-Carlo study with 256 noise realizations to compute h at different noise levels shown in Fig. 8(a). The reduction factor h depends on the conditioning of the Hankel matrix $\boldsymbol{H}_0$. For the case considered here, we observed that h achieves the value of around 1.7. However, for better-conditioned cases with well-separated spectral peaks, h can be significantly bigger (in the range of 5 for a spectrum with 3 peaks, separated by 0.57 ppm). This experiment suggests that LP-based low-rank filtering can be less-effective than PS-based low-rank filtering, although both are used in LORA.

We performed a Monte-Carlo study to further evaluate the denoising effectiveness of low-rank filtering with 2048 noise realizations to obtain histograms of the errors before and after denoising as

$$\begin{aligned} e &= s - s_0 \\ \bar{e} &= \bar{s} - s_0, \end{aligned} \tag{21}$$

where $\bar{s}$ denotes the denoised data. Figure 7 shows some representative results for the low-rank filtering based on partial separability. It is easy to see that a significant reduction in the noise variance was achieved at every SNR level. Note that low-rank filtering is a "biased" estimator. However, the resulting bias is negligible once the rank is correctly chosen and SNR is not extremely low. In the case of extremely low SNR, using an over-estimated rank reduces bias but at the expense of reduced filtering effect.

Figure 8 shows some representative results from a Monte-Carlo study of low-rank filtering based on linear predictability. Notice that the denoising performance is frequency-dependent. Histogram at a low-SNR frequency f_2 in Fig. 8(c) shows that a significant noise variance reduction is still achieved, however a larger bias is introduced, compared to the case of a higher-SNR frequency f_1 in Fig. 8(b). For the frequency regions that

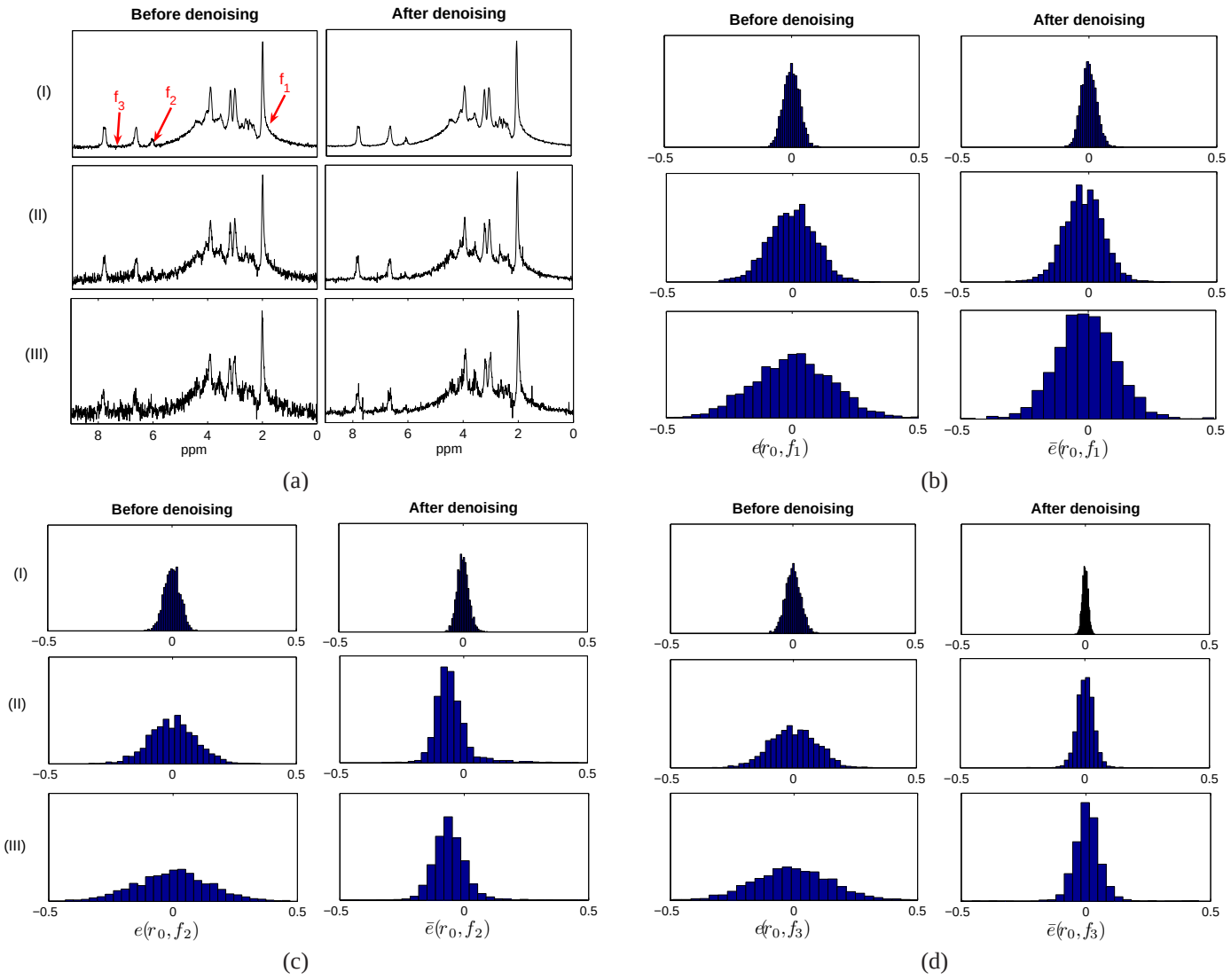


Fig. 8. Monte-Carlo study of low-rank filtering based on linear predictability: (a) representative noisy and denoised spectra at three noise levels; (b)-(d) error distributions calculated as the histograms of $\boldsymbol{e}$ and $\bar{\boldsymbol{e}}$ according to (21) at a particular locations $(\boldsymbol{r}_0, f_1)$, $(\boldsymbol{r}_0, f_2)$, and $(\boldsymbol{r}_0, f_3)$, respectively. For this illustrative example, $\boldsymbol{r}_0$ is a representative voxel in the white matter region marked as "2" in Fig. 3(f) and f_1, f_2, f_3 are the frequency points marked in Fig. 8(a). Notice that the denoising performance is frequency-dependent.

contain mostly noise (such as the frequency f_3 in Fig. 8(d)), we observe significant noise variance reduction with no bias.

C. In-vivo experiments

We have applied LORA to denoise in-vivo MRSI data of the mouse brain, acquired using a Varian INOVA 11.74 T (500 MHz) MRI scanner at Washington University in St. Louis. This dataset was previously described and used in [39]. A craniectomy was performed on the mouse and the Middle Cerebral Artery (MCA) was electrocoagulated. In-vivo MRSI data was acquired using a CSI sequence with CHESS water suppression, TE=270 ms, TR=1500 ms, bandwidth=6000 Hz, 1024 FID data points, and 8 averages. The CSI dataset was pre-processed to compensate for field inhomogeneity and remove water resonance using HSVD [3]. Notice that the residual water and artifacts from field inhomogeneity was not completely removed. For a detailed description of the dataset, see [39]. Before performing LORA-denoising, data was preprocessed with spatial denoising using anatomical reconstruction [40] to protect spatial features from field inhomogeneity artifacts during the low-rank filtering due to spatiotemporal partial separability. NAA, lactate (Lac), and glutamate/glutamine (Glx) spatial maps were obtained by integrating the complex spectrum at 2.02 ppm, 1.33 ppm, and in the region from 2.7 to 2.8 ppm respectively. Figure 9 shows NAA and Lac maps obtained from the measured, spatially-denoised, and LORA-denoised data after each individual low-rank filtering. Figure 10 shows spectra obtained from these reconstructions at voxels outside of electrocogulation area (voxels marked as "1" and "3" in Fig. 9(a)) and in the region of electrocogulation (voxel marked as "2" in Fig. 9(a)). It can be clearly seen that the noise has been significantly suppressed. Electrocoagulation blocked MCA, resulting in a stoppage of blood supply and lack of oxygen. Thus, we expect elevated Lac and reduced NAA concentrations in the

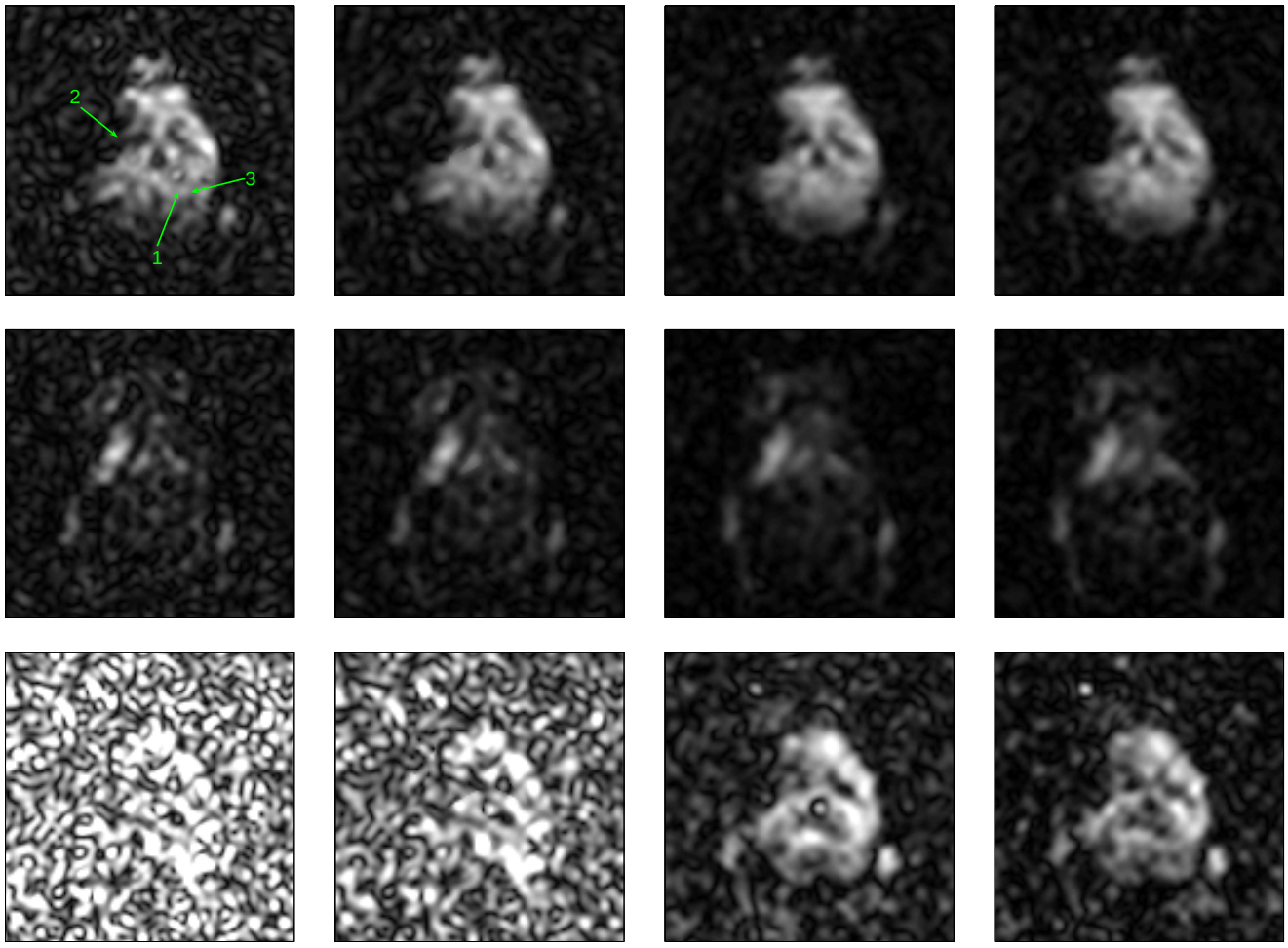


Fig. 9. Denoising results from in-vivo experimental data. Spatial distributions corresponding to NAA in the region from 1.9 to 2.1 ppm (top row), Lac in the region from 1.2-1.37 ppm (middle row), and Glx in the region from 2.7 to 2.8 ppm (bottom row) obtained from: (a) noisy data; (b) spatial denoising; (c) PS-based low-rank filtering; (d) LP-based low-rank filtering of the results in (c) (the final output of LORA).

electrocogualation region, which can be seen from the LORA-denoised reconstruction in Fig. 9(d).

It is also interesting to observe the filtering results of each individual filtering step. While spatial filtering is useful, its output is often still too noisy for practical application. The results were significantly improved by subsequent PS-based and LP-based low-rank filterings. Note that while PS-based low-rank filtering performs spatial-spectral denoising, LP-based low-rank filtering is just a spectral filter. These filtering characteristics can be clearly observed from Figs. 9 and 10. Note also that the denoised spectra in Fig. 10(d) show minimal spectral distortions with impressive SNR improvement as compared to the results in Fig. 10(a).

V. Conclusions

This paper has presented a new method (LORA) for spatial-spectral filtering of MR spectroscopic imaging data. LORA exploits the low-rank properties of MRSI data, one due to partial separability and the other due to linear predictability. The combined use of partial separability and linear predictability provides a new principled way to improve SNR for spatiotemporal imaging. Simulation and experimental results demonstrate that LORA can effectively denoise MRSI data with very low SNRs. It should prove useful for practical MRSI applications.

References

[1] R. A. de Graaf, *In vivo NMR spectroscopy*, 2nd ed. John Wiley & Sons, Ltd, 2007.

[2] A. D. Costanzo, F. Trojsi, M. Tosetti, T. Schirmer, S. M. Lechner, T. Popolizio, and T. Scarabino, "Proton MR spectroscopy of the brain at 3 T: an update," *Eur. Radiol.*, vol. 17, pp. 1651–1662, 2007.

[3] A. van den Boogaart, D. van Ormondt, W. W. F. Pijnappel, R. de Beer, and M. Ala-Korpela, "Removal of the water resonance from 1H magnetic resonance spectra," *in Mathematics in Signal Processing III*, vol. J. G. McWhirter, Ed., Oxford, 1994, pp. 175–195, Clarendon Press.

[4] H. F. C. de Greiff, R. Ramos-Garcia, and J. V. Lorenzo-Ginori, "Signal de-noising in magnetic resonance spectroscopy using wavelet transforms," *Concepts Magn. Reson.*, vol. 14, pp. 388–401, 2002.

[5] H. C. Andrews and C. L. Patterson, "Singular value decompositions and digital image processing," *IEEE Trans. Acoust. Speech Sign. Process.*, vol. 24, pp. 26–53, 1976.

[6] R. Kumaresan and D. W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, pp. 833–840, 1982.

[7] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, pp. 490–530, 2005.

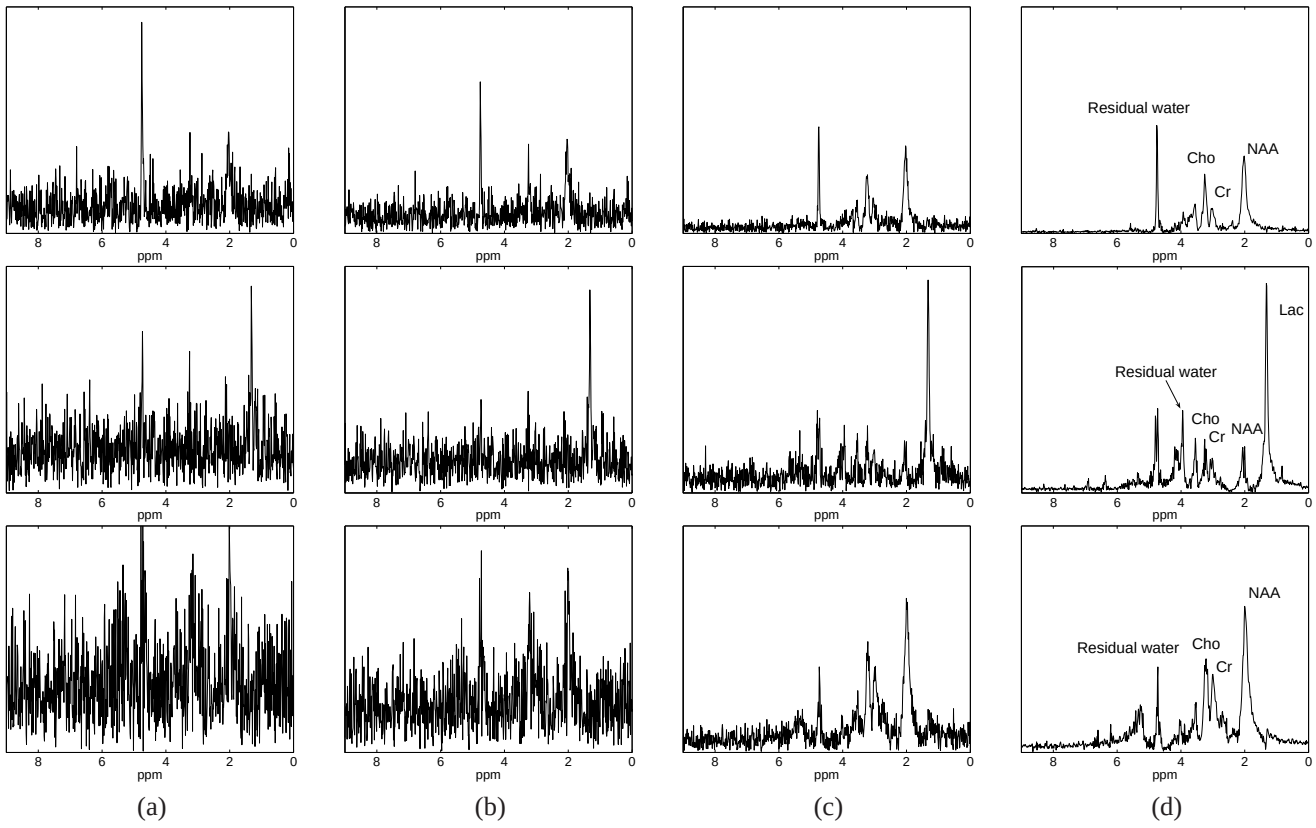


Fig. 10. Denoising results from in-vivo experimental data set: spectra at three voxels in the region outside of electrocogualation (voxel 1, top row and voxel 3, bottom row) and inside of electrocogualation area (voxel 2, middle row), obtained from (a) noisy data; (b) spatial denoising; (c) PS-based low-rank filtering of the spatial filtering results; (d) LP-based low-rank filtering of PS-based low-rank filtering results (which is the final output of LORA). All spectra are shown in absolute mode and on the same scale.

[8] O. A. Ahmed, "New denoising scheme for magnetic resonance spectroscopy signals," *IEEE Trans. Med. Imag.*, vol. 24, pp. 809–816, 2005.

[9] A. Diop, Y. Zaim-Wadghiri, A. Briguet, and D. Graveron-Demilly, "Improvements of quantitation by using the Cadzow enhancement procedure prior to any linear-prediction methods," *J. Magn. Reson.*, vol. 105, pp. 17–24, 1994.

[10] R. D. Beer, "Quantitative in vivo NMR," *University of Technology Delft Lecture Notes.*, 1994. Available online at http://www.intermnet.ua.es/inteRMNet/c59/index.html.

[11] A. C. Sava, D. Sima, J. Poullet, and S. V. Huffel, "Exploiting spatial information for estimating metabolite concentration in MRSI," *17th Proc. Intl. Soc. Mag. Reson. Med.*, p. 4280, 2009.

[12] Y. Bao and A. A. Maudsley, "Improved reconstruction for MR spectroscopic imaging," *IEEE Trans. Med. Imag.*, vol. 26, pp. 686–695, 2007.

[13] J. Kornak, K. Young, N. Schuff, A. A. Maudsley, and M. W. Weiner, "Bayesian reconstruction of low resolution magnetic resonance imaging modalities," *Leeds Annual Statistical Research Workshop - Bioinformatics, Images, and Wavelets*, pp. 89–92, 2004.

[14] J. A. Cadzow, "Signal enhancement - a composite property mapping algorithm," *IEEE Trans. Acoust. Speech Sign.*, vol. 36, pp. 49–62, 1988.

[15] R. Eslami and M. Jacob, "Reduction of distortions in MRSI using a new signal model," *Proc. IEEE Intl. Symposium Biomed. Imag.*, pp. 438–441, 2009.

[16] H. M. Nguyen, J. P. Haldar, M. N. Do, and Z.-P. Liang, "Denoising of MR spectroscopic imaging data with spatial-spectral regularization," *Proc. IEEE Intl. Symposium Biomed. Imag.*, pp. 720–723, 2010.

[17] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," *Proc. IEEE Intl. Symposium Biomed. Imag.*, pp. 988–991, 2007.

[18] H. M. Nguyen, X. Peng, M. N. Do, and Z.-P. Liang, "Spatiotemporal filtering of MR spectroscopic imaging data by low-rank approximations," in *Proc. IEEE Intl. Symposium Biomed. Imag.*, 2011, pp. 857 – 860.

[19] X. Peng, H. Nguyen, J. Haldar, D. Hernando, X.-P. Wang, and Z.-P. Liang, "Correction of field inhomogeneity effects in limited k-space coverage MRSI data with anatomical constraint," *Proc. Intl. Soc. Mag. Reson. Med.*, pp. 883–886, 2010.

[20] A. Bashir and D. A. Yablonskiy, "Natural linewidth chemical shift imaging (NL-CSI)," *Mag. Res. Imag.*, pp. 7–18, 2006.

[21] I. Khalidov, D. V. D. Ville, M. Jacob, F. Lazeyras, and M. Unser, "BSLIM: spectral localization by imaging with explicit B_0 field inhomogeneity compensation," *IEEE Trans. Med. Imag.*, pp. 990–1000, 2007.

[22] X. Hu, D. N. Levin, P. C. Lauterbur, and T. Spraggins, "SLIM: Spectral localization by imaging," *Magn. Reson. Med.*, vol. 8, pp. 314 – 322, 1988.

[23] Z.-P. Liang and P. C. Lauterbur, "A theoretical analysis of the SLIM technique," *J. Magn. Reson.*, vol. 102, pp. 54–60, 1993.

[24] P. B. Barker and D. D. M. Lin, "In vivo proton MR spectroscopy of the human brain," *Prog. NMR Spect.*, vol. 49, pp. 99–128, 2006.

[25] U. Seeger and U. Klose, "Parameterized evaluation of macromolecules and lipids in proton MR spectroscopy of brain diseases," *Magn. Reson. Med.*, vol. 49, pp. 19–28, 2003.

[26] W. W. F. Pijnappel, A. van den Boogaart, R. de Beer, and D. van Ormondt, "SVD-based quantification of magnetic resonance signals," *J. Magn. Reson.*, vol. 97, pp. 122–134, 1992.

[27] I. S. Iohvidov, *Hankel and Toeplitz matrices and forms*. Birkhauser, 1982.

[28] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR - Sbornik*, vol. 1, pp. 457 – 483, 1967.

[29] R. Shibata, "Selection of the order of an autoregressive model by Akaike's information criterion," *Biometrika*, vol. 63, pp. 117–126, 1976.

[30] V. Govindaraju, K. Young, and A. A. Maudsley, "Proton NMR chemical shifts and coupling constants for brain metabolites," *NMR Biomed.*, vol. 13, pp. 129–153, 2000.

[31] F. S. V. Bazan and P. L. Toint, "Conditioning of infinite hankel matrices of finite rank," *Syst Control Lett*, vol. 41, pp. 347–359, 2000.

[32] B. J. Soher, K. Young, A. Bernstein, Z. Aygula, and A. A. Maudsley, "GAVA: Spectral simulation for in vivo MRS applications," *J. Magn. Reson.*, vol. 185, pp. 291–299, 2007.

[33] R. Srinivasan, N. Sailasuta, R. Hurd, S. Nelson, and D. Pelletier,

"Evidence of elevated glutamate in multiple sclerosis using magnetic resonance spectroscopy at 3 T," *Brain*, vol. 128, pp. 1016–1025, 2005.
[34] Y. Wang and S.-J. Li, "Differentiation of metabolic concentrations between gray matter and white matter of human brain by in vivo ^{1}H magnetic resonance spectroscopy," *Magn. Reson. Med.*, vol. 39, pp. 28–33, 1998.
[35] F. Traber, W. Block, R. Lamerichs, J. Gieseke, and H. H. Schild, "^{1}H metabolite relaxation times at 3.0 Tesla: measurements of T1 and T2 values in normal brain and determination of regional differences in transverse relaxation," *J. Magn. Reson. Imag.*, vol. 19, pp. 537–545, 2004.
[36] C. Choi, N. J. Coupland, P. P. Bhardwaj, S. Kalra, C. A. Casault, K. Reid, and P. S. Allen, "T_2 measurement and quantification of glutamate in human brain in vivo," *Magn. Reson. Med.*, vol. 56, pp. 971–977, 2006.
[37] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, pp. 2080–2095, 2007.
[38] B. D. Moor, "The singular value decomposition and long and short spaces of noisy matrices," *IEEE Sign. Process.*, vol. 41, pp. 2826–2838, 1993.
[39] J. P. Haldar, T.-H. Wu, Q. Wang, C.-I. Chen, S.-K. Song, and Z.-P. Liang, "Further development in anatomically constrained MR image reconstruction: application to multimodal imaging of mouse stroke," *Proc. Intl. IEEE Eng. Med. Biol. Soc.*, pp. 422–425, 2008.
[40] J. P. Haldar, D. Hernando, S.-K. Song, and Z.-P. Liang, "Anatomically constrained reconstruction from noisy data," *Magn. Reson. Med.*, vol. 59, pp. 810–818, 2008.

Minh N. Do received his Dr.Sci. degree in communication systems from the Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland, in 2001. He is currently Associate Professor in the Department of Electrical and Computer Engineering. His research interests include image and multidimensional signal processing, wavelets and multiscale geometric analysis, computational imaging, augmented reality, and visual information representation. He received a Silver Medal from the 32nd International Mathematical Olympiad in 1991, a University Medal from the University of Canberra in 1997, a Doctorate Award from the EPFL in 2001, a CAREER Award from the National Science Foundation in 2003, and a Young Author Best Paper Award from IEEE in 2008. He is a member of the IEEE Signal Processing Theory and Methods and Image, Video, and Multidimensional Signal Processing Technical Committees, and an Associate Editor of the IEEE Transactions on Image Processing.

Hien Nguyen received her B.S. degree in Computer Science from Hanoi University of Technology, Vietnam, in 2002 and the M.S./Ph.D. degree in Electrical and Computer Engineering from University of Illinois at Urbana-Champaign in 2007/2011. She is currently a postdoctoral researcher at Stanford University, Department of Radiology, Richard M. Lucas center for Imaging. Her research interests include multi-dimensional signal processing, parameter estimation, and inverse problems, with particular emphasis on biomedical imaging applications.

Zhi-Pei Liang received his Ph.D. degree in Biomedical Engineering from Case Western Reserve University in 1989. He is currently the Franklin W. Woeltge Professor of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign (UIUC). Dr. Liang's research covers image formation theory, algorithms and biomedical applications. He is a recipient of the 1990 Sylvia Sorkin Greenfield Best Paper Award (Medical Physics), an NSF CAREER Award (1995), and the 1999 IEEE-EMBS Early Career Achievement Award, and the 2012 IFMBE Otto Schmitt Award. He was named Henry Magnuski Scholar (1999-2001), and University Scholar (2001-2004) at UIUC. He is a Fellow of IEEE (2006) and ISMRM (2010), and was elected to the American Institute for Medical and Biological Engineering in 2005 and to the International Academy of Medical and Biological Engineering in 2012. Dr. Liang served as President of IEEE Engineering in Medicine and Biology Society (EMBS) from 2011-2012.

Xi Peng received the B.Eng. and M.S. degree in measurement, control and instruments in 2006 and 2008 respectively, and the Ph.D. degree in Electrical Engineering in 2012, all from Wuhan University, China. During the doctoral process, he was a visiting scholar in Beckman Institute, University of Illinois at Urbana-Champaign. He is currently Assistant Professor in the Paul C. Lauterbur Research Center for Biomedical Imaging, Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. His research interests include image reconstruction, parameter estimation, and signal processing for biomedical applications.