# Constrained Spectral Estimation for Magnetic Resonance Spectroscopic Imaging

Qiang Ning

***Abstract*—Spectral estimation is an important problem in MR Spectroscopic Imaging (MRSI). Although existing methods have effectively taken advantage of the linear prediction structure of MRSI signals, which has indeed improved spectral estimation over nonparametric methods, very limited attention has been paid to exploiting both the spectral and spatial characteristics of the spectral parameters. As a result, spectral estimation remains a challenging problem now, especially for low signal-to-noise ratio (SNR) scenarios. This work addresses the spectral estimation problem in MRSI by jointly estimating the spectra over all the voxels of interest, incorporating both spectral constraints (in the form of basis functions) and spatial regularization. Both simulated and experimental MRSI data have been used to demonstrate the proposed spectral estimation method.**

***Index Terms*—MRSI, constrained spectral estimation**

## I. Introduction

MRSI is a unique, non-invasive tool to acquire *in vivo* biochemical information without using molecular probes or radionuclide tracers. In contrast to MRI, which acquires anatomical information, MRSI integrates the capability of MRI and MR Spectroscopy to obtain a spectrum at each spatial location (Fig. 1). These spectra have proved to be very valuable biochemical information about the tissue, which can be used for the detection, diagnosis, and treatment of diseases [1], [2], and for fundamental sciences like metabolomics [3]. However, the inherent low sensitivity of MR scans and the resulting low SNR make spectral estimation for MRSI a challenging problem in practice.

A simple and still widely used spectral estimation method is to directly Fourier transform $d(\mathbf{x}, t)$, but its practical use is often limited to MRSI experiments with high SNR. Since MRSI signals follow a damped complex sinusoidal model, linear prediction–based methods can also be used for spectral estimation. However, very limited attention has been paid to exploiting both spectral and spatial prior knowledge of the MRSI signals (e.g., smoothness of the metabolite distribution). This work hence addresses the spectral estimation problem by incorporating both spectral and spatial prior information acquired beforehand, which significantly reduces the degree of freedom of the parameter space and leads to better estimation results. The improvement of the proposed method is validated using both a simulation dataset and an *in vivo* experiment.

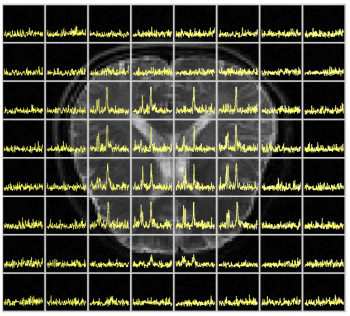

Fig. 1. Example MRSI data acquired from a healthy human subject.

## II. Background

Mathematically, MRSI can be viewed as a spectral estimation problem formulated as

$$d(\mathbf{x}, t) = \int \rho(\mathbf{x}, f) e^{j2\pi f t} df + n(\mathbf{x}, t), \qquad (1)$$

where $d(\mathbf{x}, t)$ is the measured data in the time-domain, $\rho(\mathbf{x}, f)$ is the desired spectrum, and $n(\mathbf{x}, t)$ represents additive Gaussian noise. The main challenge for MRSI lies in the extremely low SNR.

Over the past few decades, researchers have made significant efforts to address this spectral estimation problem. There are two primary categories of existing spectral estimation methods for MRSI: nonparametric methods and parametric methods. An early nonparametric method was to perform the Fourier transform, that is,

$$\hat{\rho}(\mathbf{x}, f) = \int d(\mathbf{x}, t) e^{-j2\pi f t} dt.$$

The corresponding metabolite at spatial location $\mathbf{x}$ would then be simply calculated as the area beneath

the corresponding peak in $\hat{\rho}(\mathbf{x}, f)$. In high-SNR scenarios, the Fourier transform method is useful and efficient; however, the SNR is often so low in practice that the peak locations cannot even be identified within the resulting spectrum, much less accurate peak values.

Common parametric approaches rely on a damped complex sinusoidal model for MRSI signals [4], [5]:

$$\rho(\mathbf{x}, f) = \int \left( \sum_{l=1}^{L} a_l(\mathbf{x}) e^{-\theta_l(\mathbf{x})t - j2\pi f_l(\mathbf{x})t} \right) e^{-j2\pi f t} dt,$$

or in the time-domain,

$$\rho(\mathbf{x}, t) = \sum_{l=1}^{L} a_l(\mathbf{x}) e^{-\theta_l(\mathbf{x})t - j2\pi f_l(\mathbf{x})t}, \qquad (2)$$

where $a_l(\mathbf{x})$, $\theta_l(\mathbf{x})$, and $f_l(\mathbf{x})$ are the concentration, damping factor, and resonance frequency, respectively, of the $l^{\text{th}}$ peak at location $\mathbf{x}$, and where

$$d(\mathbf{x}, t) = \rho(\mathbf{x}, t) + n(\mathbf{x}, t).$$

This model motivates the application of linear prediction–based methods for spectral estimation, such as LPSVD [4], HSVD [5], and HLSVD [6]. Let $d_{\mathbf{x}}[m]$ denote $d(\mathbf{x}, m\Delta t)$, i.e., the discretized time sequence at location $\mathbf{x}$. Then the LPSVD method, for instance, says that $\forall \mathbf{x}$, $d_{\mathbf{x}}[m], m = 0, 1, \ldots, M-1 (M > L)$, satisfies the following equation

$$\begin{bmatrix} d_{\mathbf{x}}[L-2] & d_{\mathbf{x}}[L-3] & \cdots & d_{\mathbf{x}}[0] \\ d_{\mathbf{x}}[L-1] & d_{\mathbf{x}}[L-2] & \cdots & d_{\mathbf{x}}[1] \\ \vdots & \vdots & \ddots & \vdots \\ d_{\mathbf{x}}[M-2] & d_{\mathbf{x}}[M-3] & \cdots & d_{\mathbf{x}}[M-L] \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_L \end{bmatrix} = \begin{bmatrix} d_{\mathbf{x}}[L-1] \\ d_{\mathbf{x}}[L] \\ \vdots \\ d_{\mathbf{x}}[M-1] \end{bmatrix},$$

from which we can determine the linear prediction coefficients $\beta_1, \ldots, \beta_L$ by solving the linear equation in the least squares sense (or in the total least squares sense). Rooting the polynomial equation

$$z^L + \beta_1 z^{L-1} + \cdots + \beta_{L-1} z + \beta_L = 0,$$

yields estimates of the "poles" in (2). The linear parameters (i.e., $a_l(\mathbf{x})$) can then be determined by projecting $d(\mathbf{x}, t)$ onto the subspace spanned by $\{e^{-\theta_l(\mathbf{x})t - j2\pi f_l(\mathbf{x})t}\}_{l=1}^{L}$. By using a parametric model, linear prediction–based methods can perform better spectral estimation than nonparametric methods. However, an inherent drawback of those methods is their limitation in incorporating prior knowledge (despite the existence of some variants of linear prediction methods that can impose some special forms of prior knowledge, e.g., [7], [8]). This limitation strongly prohibits linear prediction methods from being applied to low-SNR scenarios, which is often the case in practice.

## III. Proposed Method

In contrast to linear prediction–based methods, we address the spectral estimation problem for MRSI by incorporating both spectral and spatial prior knowledge. The prior knowledge can greatly reduce the parameter search space, helping to improve the estimation variance of the parameters.

### A. Model

We propose to impose spectral constraints in the form of spectral basis functions. This is motivated by the fact that within one spectrum $\rho(\mathbf{x}, f)$, different peaks originating from the same metabolite should be related to each other in terms of their concentration ratios, damping factors, and frequency locations (for example, Fig. 2 shows the spectra of several typical basis functions). Without loss of generality, $d(\mathbf{x}, t)$ can be modeled using basis functions

$$d(\mathbf{x}, t) = \sum_{n=1}^{N} a_n(\mathbf{x}) e^{-\theta_n(\mathbf{x})t} \varphi_n(t) + n(\mathbf{x}, t), \qquad (3)$$

where $\varphi_n(t)$ is the so-called basis function for the $n^{\text{th}}$ metabolite, which can be accurately obtained beforehand using quantum simulation (e.g., GAVA [9]). Obviously, the introduction of basis functions $\varphi_n(t)$ greatly reduces the dimension of the parameter space.

### B. Formulation

Note that in the discrete case, (3) can be written in a more compact, vector-matrix form

$$\mathbf{d}_{\mathbf{x}} = \mathbf{K}(\boldsymbol{\theta}_{\mathbf{x}}) \mathbf{a}_{\mathbf{x}} + \mathbf{n}_{\mathbf{x}},$$

where $\mathbf{d}_{\mathbf{x}}$ and $\mathbf{n}_{\mathbf{x}}$ contain all the data and noise at location $\mathbf{x}$, where $\mathbf{a}_{\mathbf{x}}$ and $\boldsymbol{\theta}_{\mathbf{x}}$ are the concentration vector and damping factor vector at $\mathbf{x}$, where $\mathbf{K}(\cdot)$ represents the model matrix with basis functions,

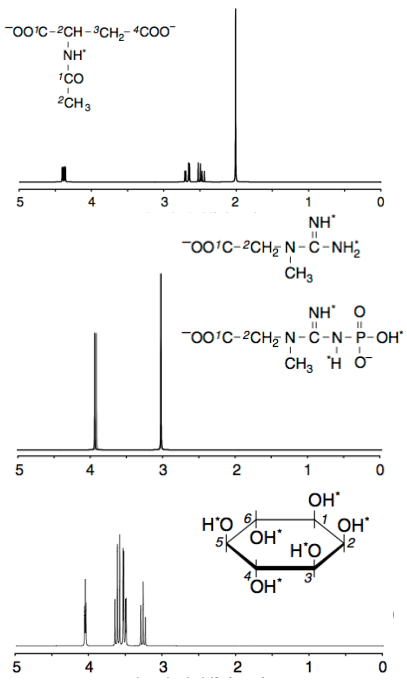

Fig. 2. Example spectra of basis functions and molecular structures of N-acetyl aspartate (NAA), creatine, and myo-inositol, three common metabolites in human brains [3].

and where $\mathbf{x} = \mathbf{x}_1, \ldots, \mathbf{x}_P$ (for $P$ total locations of interest). Instead of independently performing spectral estimation point-by-point (as existing methods do), we propose to perform joint spectral estimation incorporating prior spatial information. Specifically, let $\mathbf{d} = [\mathbf{d}_{\mathbf{x}_1}^T,\ \mathbf{d}_{\mathbf{x}_2}^T,\ \cdots,\ \mathbf{d}_{\mathbf{x}_P}^T]^T$, $\mathbf{a} = [\mathbf{a}_{\mathbf{x}_1}^T,\ \mathbf{a}_{\mathbf{x}_2}^T,\ \cdots,\ \mathbf{a}_{\mathbf{x}_P}^T]^T$, $\boldsymbol{\theta} = [\boldsymbol{\theta}_{\mathbf{x}_1}^T,\ \boldsymbol{\theta}_{\mathbf{x}_2}^T,\ \cdots,\ \boldsymbol{\theta}_{\mathbf{x}_P}^T]^T$, and

$$\mathbf{K}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{K}(\boldsymbol{\theta}_{\mathbf{x}_1}) & & \\ & \ddots & \\ & & \mathbf{K}(\boldsymbol{\theta}_{\mathbf{x}_P}) \end{bmatrix}.$$

We formulate the spectral estimation problem as

$$(\hat{\mathbf{a}}, \hat{\boldsymbol{\theta}}) = \arg\min_{\mathbf{a},\boldsymbol{\theta}} \|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{a}\|_2^2 + \mathrm{R}(\mathbf{a}, \boldsymbol{\theta}), \quad (4)$$

where $\mathrm{R}(\mathbf{a}, \boldsymbol{\theta})$ is a regularization functional imposing spatial constraints on $\mathbf{a}$ and $\boldsymbol{\theta}$. This incorporation of spatial prior knowledge can help further reduce the parameter search space and improve the spectral estimation results.

In this work, we limit our focus to imposing smoothness constraints on the concentration vector $\mathbf{a}$, although other spatial constraints are possible. The use of smoothness constraints is motivated by the fact that in most biological samples, there are only a few tissue types, inside which metabolites should spatially have rather smooth concentrations. Using this motivation, we can use an $\ell_1$-regularizer for $\mathrm{R}(\mathbf{a}, \boldsymbol{\theta})$. More precisely, we reformulate the problem in (4) as:

$$(\hat{\mathbf{a}}, \hat{\boldsymbol{\theta}}) = \arg\min_{\mathbf{a},\boldsymbol{\theta}} \frac{1}{2}\|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{a}\|_2^2 + \eta\|\mathcal{W}\{\mathbf{a}\}\|_1, \quad (5)$$

where $\eta$ is a regularization parameter and $\mathcal{W}$ is an operator representing the wavelet transform, total variation (TV) transform, or total generalized variation (TGV) transform [10], etc. We found TGV to be suitable for metabolite concentration maps, promoting piecewise smoothness, although it could be easily replaced by another sparsifying transform.

### C. Algorithm

The joint quantitation problem in (5) is a nonlinear least squares problem regularized by an $\ell_1$ term, so we solve it using a modified alternating direction method of multiplier (ADMM) [11]. For notation convenience, we present the algorithm for the case that $\mathcal{W}\{\mathbf{a}\}$ in (5) is a linear operator of $\mathbf{a}$ (e.g., the wavelet and TV transform) and can thus be represented in a matrix form, i.e., $\mathcal{W}\{\mathbf{a}\} = \mathbf{W}\mathbf{a}$. The proposed algorithm can be easily extended to the case that $\mathcal{W}\{\mathbf{a}\}$ represents the TGV transform as in [10].

As proposed by Guo et al. [11], we introduce an auxiliary variable $\mathbf{u}$, so that (5) is equivalent to

$$\min_{\mathbf{a},\boldsymbol{\theta},\mathbf{u}} \frac{1}{2}\|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{a}\|_2^2 + \eta\|\mathbf{u}\|_1, \ \text{ s.t. } \ \mathbf{u} = \mathbf{W}\mathbf{a}. \quad (6)$$

We then decompose (6) into the following subproblems, whose convergence is guaranteed by the classic ADMM algorithm [12]:

$$\mathbf{u}^{(n+1)} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 + \frac{\mu}{2\eta}\|\mathbf{u} - \mathbf{W}\mathbf{a}^{(n)} - \tilde{\mathbf{u}}^{(n)}\|_2^2, \quad (7)$$

$$(\mathbf{a}^{(n+1)}, \boldsymbol{\theta}^{(n+1)}) = \arg\min_{\mathbf{a},\boldsymbol{\theta}} f(\mathbf{a}, \boldsymbol{\theta}), \quad (8)$$

$$\tilde{\mathbf{u}}^{(n+1)} = \tilde{\mathbf{u}}^{(n)} + \gamma(\mathbf{W}\mathbf{a}^{(n+1)} - \mathbf{u}^{(n+1)}), \quad (9)$$

where

$$f(\mathbf{a}, \boldsymbol{\theta}) = \mu\|\mathbf{x}^{(n+1)} - (\mathbf{W}\mathbf{a} + \tilde{\mathbf{x}}^{(n)})\|_2^2 + \frac{1}{2}\|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{a}\|_2^2. \quad (10)$$

Subproblem (7) can be solved explicitly using shrinkage:

$$\mathbf{x}^{(n+1)} = \text{shrinkage}(\mathbf{W}\mathbf{a}^{(n)} + \tilde{\mathbf{x}}^{(n)}, \eta/\mu), \quad (11)$$

where we define

$$\text{shrinkage}(\mathbf{v}, \eta) = \mathbf{v}. * \max(1 - \eta./|\mathbf{v}|, 0).$$

Subproblem (8) is a challenging nonlinear optimization problem. We use the variable projection strategy [13] to efficiently solve the problem. More specifically, we solve the following two-step problem instead:

$$\boldsymbol{\theta}^{(n+1)} = \arg\min_{\boldsymbol{\theta}} f(\mathbf{a}_{\text{opt}}(\boldsymbol{\theta}), \boldsymbol{\theta}), \quad (12)$$

$$\mathbf{a}^{(n+1)} = \mathbf{a}_{\text{opt}}(\boldsymbol{\theta}^{(n+1)}), \quad (13)$$

where $\mathbf{a}_{\text{opt}}(\boldsymbol{\theta}) \triangleq \arg\min_{\mathbf{a}} f(\mathbf{a}, \boldsymbol{\theta})$ is a quadratic problem and has a closed-form solution:

$$(\beta\mathbf{K}(\boldsymbol{\theta})^H\mathbf{K}(\boldsymbol{\theta}) + \mu\mathbf{W}^H\mathbf{W})\mathbf{a}_{\text{opt}}(\boldsymbol{\theta}) = \beta\mathbf{K}(\boldsymbol{\theta})^H\mathbf{d} + \mu\mathbf{W}^H\mathbf{u}, \quad (14)$$

where $\mathbf{u} = \mathbf{x}^{(n+1)} - \tilde{\mathbf{x}}^{(n)}$. Eq. (12) is a nonlinear least squares problem, which we solve using the classic descent method where only the gradient is required. The derivative of $d\mathbf{a}_{\text{opt}}/d\boldsymbol{\theta}$ can be obtained by taking derivative of both sides of (14),

$$(\beta\mathbf{K}^H\mathbf{K} + \mu\mathbf{W}^H\mathbf{W})\frac{\partial\mathbf{a}_{\text{opt}}}{\partial\boldsymbol{\theta}} = \beta\frac{\partial\mathbf{K}^H}{\partial\boldsymbol{\theta}}(\mathbf{d} - \mathbf{K}\mathbf{a}_{\text{opt}}) - \beta\mathbf{K}^H\frac{\partial\mathbf{K}}{\partial\boldsymbol{\theta}}\mathbf{a}_{\text{opt}}. \quad (15)$$

Finally, the subproblem (9) is only an updating operation.

## IV. Results and Discussion

### A. Validation with Simulated MRSI Data

The performance of the proposed joint spectral estimation method has been evaluated on a simulated MRSI dataset. The quantification results are compared with HSVD [5], a standard linear prediction–based method that has been widely used in practice, which performs spectral estimation voxel-by-voxel without incorporating spectral and spatial priors.

The dataset consists of five representative metabolites of interest: N-acetyl aspartate (NAA), creatine (Cr), choline (Cho), glutamate/glutamine (Glx), and myo-inositol (mI). Fig. 3a) shows a typical comparison between the ground truth, HSVD, and the proposed method, based on the spectral estimation results of the concentration map (i.e., vector $\mathbf{a}$) of NAA, Cr, and Glx, and Fig. 3b) demonstrates the typical SNR levels of the simulation, which have been selected to represent realistic MRSI measurements. From Fig. 3, we can see that the estimation results of HSVD still had large spatial variations, while the proposed method significantly reduced the estimation variance. Therefore, the proposed method had a visually more robust estimation to the ground truth. Fig. 4 illustrates that the proposed method produced better fitting compared to existing methods. It shows one typical set of synthetic spectra from the true spectral parameters and from the estimated parameters for one specific voxel. As can be observed from the second row, the synthetic spectrum estimated by HSVD showed noticeable errors, which were significantly reduced by using the proposed method incorporating spatial prior information.

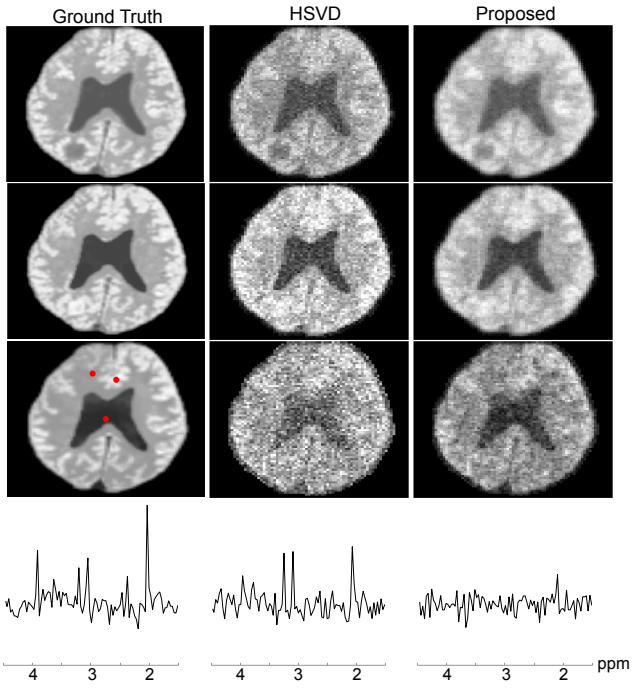


Fig. 3. Simulation results: a) the metabolite concentration maps for NAA, creatine, and Glx (glutamate/glutamine), respectively; and b) three spectra at the three spatial locations marked by red dots in a), demonstrating the typical SNR levels of the simulated data.

### B. Validation with Experimental MRSI Data

To further validate the proposed method, we acquired an *in vivo* dataset from a healthy volunteer on a 3.0 Tesla scanner using an echo-planar spectroscopic imaging (EPSI) sequence with a 30 ms echo time. Fig. 5 shows the NAA and creatine maps obtained using HSVD and the proposed method.

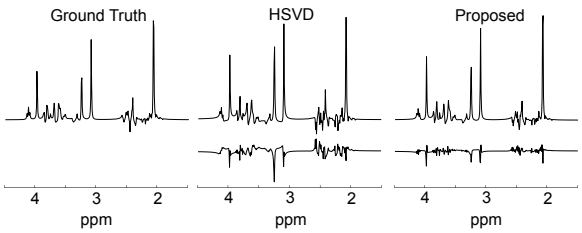


Fig. 4. The first row shows the spectra synthesized using the true spectral parameters from the simulation phantom in Fig. 1, and the estimated parameters from HSVD and the proposed method, respectively. The second row shows the difference between the estimated spectra and the true one.

As can be seen, the metabolite concentration maps estimated by HSVD showed large spatial variations, including noisy "spikes" at some locations that are biologically unrealistic. As expected, the proposed method significantly reduced the estimation variance compared to HSVD. The performance improvement of the proposed method observed from the experimental data was consistent with the simulation results in Fig. 3.

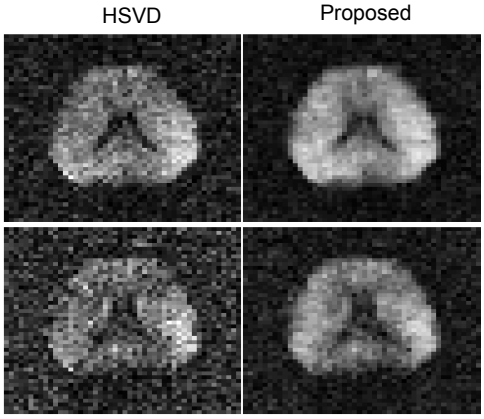


Fig. 5. Results from *in vivo* MRSI data. The concentration maps for two metabolites, NAA (top) and creatine (bottom), are presented. Note that the HSVD results show significant spatial variations (indicating large estimation variance), which are reduced considerably by using the proposed method.

## V. Conclusion

Spectral estimation is a key underlying problem in MRSI, which can have great impact on clinical applications and physiological scientific researches. Existing methods take advantage of the linear prediction structure of MRSI signals and obtain improved results compared with nonparametric methods. However, by solving the spectral estimation problem voxel-by-voxel, they treat the metabolic information at different voxels as independent from each other. Therefore, we have proposed a novel method to address MRSI spectral estimation as a joint estimation problem, so that various prior information on spectral and spatial characteristics can be incorporated. Preliminary results show that the proposed method produced significantly improved spectral estimation results over linear prediction. The proposed spectral estimation method should prove useful in many practical studies.

## References

[1] R. A. Moats, T. Ernst, T. K. Shonk, and B. D. Ross, "Abnormal cerebral metabolite concentrations in patients with probable alzheimer disease," *Magnetic resonance in medicine*, vol. 32, no. 1, pp. 110–115, 1994.

[2] D. P. Soares and M. Law, "Magnetic resonance spectroscopy of the brain: review of metabolites and clinical applications," *Clinical radiology*, vol. 64, no. 1, pp. 12–21, 2009.

[3] R. A. de Graaf, *In vivo NMR spectroscopy: principles and techniques*, John Wiley & Sons, 2013.

[4] H. Barkhuijsen, R. de Beer, W. M. M. J. Bovée, and D. van Ormondt, "Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure," *Journal of Magnetic Resonance*, vol. 61, no. 3, pp. 465–481, 1985.

[5] H. Barkhuijsen, R. de Beer, and D. van Ormondt, "Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals," *Journal of Magneic Resonance*, pp. 553–557, 1987.

[6] W. W. F. Pijnappel, A. van den Boogaart, R. de Beer, and D. van Ormondt, "SVD-based quantification of magnetic resonance signals," *Journal of Magnetic Resonance*, vol. 97, no. 1, pp. 122–134, 1992.

[7] H. Chen, S. van Huffel, D. van Ormondt, and R. de Beer, "Parameter estimation with prior knowledge of known signal poles for the quantification of NMR spectroscopy data in the time domain," *Journal of Magnetic Resonance, Series A*, vol. 119, no. 2, pp. 225–234, 1996.

[8] P. Stoica, Y. Selén, N. Sandgren, and S. van Huffel, "Using prior knowledge in svd-based parameter estimation for magnetic resonance spectroscopy-the atp example," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 9, pp. 1568–1578, 2004.

[9] B. J. Soher, K. Young, A. Bernstein, Z. Aygula, and A. A. Maudsley, "GAVA: spectral simulation for in vivo mrs applications," *Journal of Magnetic Resonance*, vol. 185, no. 2, pp. 291–299, 2007.

[10] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

[11] W. Guo, J. Qin, and W. Yin, "A new detail-preserving regularizaiton scheme," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 1309–1334, 2014.

[12] B. He, L.-Z. Liao, D. Han, and H. Yang, "A new inexact alternating directions method for monotone variational inequalities," *Mathematical Programming*, vol. 92, no. 1, pp. 103–118, 2002.

[13] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal of Numerical Analysis*, vol. 10, no. 2, April 1973.