Published in final edited form as:

IEEE Trans Biomed Eng. 2017 May ; 64(5): 1178–1186. doi:10.1109/TBME.2016.2594583.

Spectral Quantification for High-Resolution MR Spectroscopic Imaging with Spatiospectral Constraints

Qiang Ning [Student Member, IEEE],

Department of Electrical and Computer Engineering and the Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Chao Ma [Member, IEEE],

Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, and is now with the Gordon Center for Medical Imaging, Massachusetts General Hospital and Department of Radiology, Harvard Medical School, Charlestown, MA 02129

Fan Lam [Member, IEEE], and

Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Zhi-Pei Liang [Fellow, IEEE]

Department of Electrical and Computer Engineering and the Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Abstract

Objective—To obtain reliable spectral estimation from magnetic resonance spectroscopic imaging (MRSI) data.

Methods—The proposed method takes advantage of prior knowledge: a) along the spectral dimension in the form of spectral bases, and b) along the spatial dimensions in the form of spatial regularizations (e.g., smoothness or transform sparsity) and jointly estimates parameters from all the voxels.

Results—Simulation and in vivo studies have been performed to demonstrate the performance of the proposed method. A Cramér-Rao bound based analysis is also provided.

Conclusion—Incorporation of both spatial and spectral constraints can significantly improve spectral quantification of MRSI data.

Significance—The proposed method is expected to be useful for various quantitative MRSI studies.

Index Terms

MRSI; spectral quantification; sparsity constraint; Cramér-Rao bound

I. Introduction

MRSI is a unique tool for noninvasive, label-free molecular imaging, with a wide range of potential applications from early detection and diagnosis of various diseases [1], [2], to physiological studies [3]. Reliable spectral quantification (i.e., obtaining accurate estimates of spectral parameters) is a key problem for quantitative studies using MRSI. One challenge is due to the inherently low sensitivity of NMR techniques and the nonlinearity associated with the estimation problem. Significant efforts have been made in the past decades to address this challenge. For example, HSVD [4], HLSVD [5], LPSVD [6], and HTLS [7], solve the spectral estimation problem by exploiting the linear prediction structure of the MRSI signals. However, these linear prediction methods have limited capability for incorporating prior knowledge (e.g., resonance structures) [8], [9], thereby limiting their robustness and accuracy in practice.

State-of-the-art methods (e.g., VARPRO [10], LCMoel [11], AMARES [12], QUEST [13], AQSES [14], Young et al. [15], etc.) impose stronger spectral constraints in the form of spectral bases (obtained through quantum mechanics simulations [16] or from in vitro experiments under the same experimental setup), and achieve much more robust quantification than linear prediction methods. Although these methods have been successful in processing experimental MRSI data, they perform quantification voxel-by-voxel separately without exploiting the spatial characteristics of the imaging object. The resulting spectral estimates may still have large uncertainties, especially when the SNR is low (e.g., due to acquisition time constraints and/or high spatial resolution).

To deal with the large uncertainties in estimated parameter maps, several quantification methods exploring spatial prior knowledge have been proposed. Early efforts (e.g., Soher et al. [17] and LCMoel [11]) make use of parameters estimated from neighboring voxels to obtain a better initial guess or a stronger constraint to mitigate the local minimum problem. Recently, methods incorporating spatial prior knowledge through regularization have also been reported, promoting smoothness of the parameter maps (e.g., AQSES-MRSI [18] and Kelm et al. [19]), which have led to significant improvements.

In this work, we describe a new method to jointly estimate spectral parameters for all the voxels of interest. This method enables us to take advantage of both spectral and spatial characteristics of the molecules of interest. More specifically, built on state-of-the-art methods, a regularization framework is introduced to enforce spatial constraints across the entire image object (i.e., spatial smoothness or transform sparsity), and the resulting regularized nonlinear least-squares problem is solved efficiently using a variable-projection based strategy. The proposed method leads to significantly improved spectral estimation, which is validated using both simulation and experimental MRSI data. Theoretical analyses using the constrained Cramér-Rao bound (CRB) [20] further quantify the benefits of using spatial constraints.

The rest of this paper is organized as follows. Sections II and III describe the proposed formulation and algorithm, followed by a theoretical characterization of its performance in Section IV. Section V shows the effectiveness of the proposed method using simulation and

in vivo results, with comparison to a state-of-the-art time-domain quantification method, and Section VI contains the concluding remarks of the paper. A preliminary version of this work was reported in [21].¹

II. Proposed Method

A. Signal Model with Spectral Constraints

Given the spatio-spectral/spatiotemporal reconstruction of an image object, the MRSI signal at voxel location $\mathbf{x}_p$ and time instant t_q can be expressed as

$$d(\mathbf{x}_p, t_q) = \sum_{m=1}^M a_m(\mathbf{x}_p) e^{-t_q/T_{2,m}^*(\mathbf{x}_p)} e^{-i2\pi \Delta f(\mathbf{x}_p) t_q} \varphi_m(t_q) + n_{pq}, \quad (1)$$

For $p = 1, \dots, P$, $q = 1, \dots, Q$, where n_{pq} represents additive noise (often assumed to be complex Gaussian without spatial correlation), a_m , $T_{2,m}^*$, and φ_m represent the concentration, T_2^* relaxation time, and basis function for the m -th spectral component, respectively, and f is the frequency shift caused by local B_0 field inhomogeneity. Specifically, $\{\varphi_m\}_{m=1}^M$ represents the physics-based spectral structures of metabolites; direct incorporation of $\{\varphi_m\}_{m=1}^M$ can be viewed as imposing equality constraints and is usually stronger than linear predictability, spectral sparsity, etc. For notation convenience, (1) is rewritten in a more compact form as

$$\mathbf{d}_p = \mathbf{K}(\boldsymbol{\theta}_p) \cdot \mathbf{a}_p + \mathbf{n}_p, \quad p=1, \dots, P, \quad (2)$$

where

$$\mathbf{d}_p = \begin{bmatrix} d(\mathbf{x}_p, t_1) \\ \vdots \\ d(\mathbf{x}_p, t_Q) \end{bmatrix}, \quad \mathbf{a}_p = \begin{bmatrix} a_1(\mathbf{x}_p) \\ \vdots \\ a_M(\mathbf{x}_p) \end{bmatrix}, \quad \mathbf{n}_p = \begin{bmatrix} n_{p,1} \\ \vdots \\ n_{p,Q} \end{bmatrix},$$

and $\mathbf{K}(\boldsymbol{\theta}_p) \in \mathbb{C}^{Q \times M}$, with $\boldsymbol{\theta}_p$ absorbing all the nonlinear unknown parameters at voxel $\mathbf{x}_p$ (e.g., $f(\mathbf{x}_p)$ and $\{T_{2,m}^*(\mathbf{x}_p)\}_{m=1}^M$) such that

$$k_{q,m}(\boldsymbol{\theta}_p) = e^{-t_q/T_{2,m}^*(\mathbf{x}_p)} e^{-i2\pi \Delta f(\mathbf{x}_p) t_q} \varphi_m(t_q),$$

¹A more recent paper on this topic was published online (Laruelo et al., NMR Biomed. 2016) after this paper was accepted for publication.

for $q = 1, \dots, Q$, and $m = 1, \dots, M$. Here, spectral constraints are applied in the form of basis functions, $\{\varphi_m(\cdot)\}_{m=1}^M$, which are prior knowledge of the NMR spectral structure of different molecules and can be obtained from either quantum simulation (e.g., GAVA [16] and NMR-SCOPE [22]) or in vitro experiments.

Given the signal model in (2), a common practice to obtain the estimated parameters is to solve the following nonlinear least squares problem (or maximum likelihood estimation problem for the white Gaussian noise),

$$(\hat{\mathbf{a}}_p, \hat{\boldsymbol{\theta}}_p) = \arg \min_{\mathbf{a}_p, \boldsymbol{\theta}_p} \|\mathbf{d}_p - \mathbf{K}(\boldsymbol{\theta}_p) \mathbf{a}_p\|_2^2, \quad p=1, 2, \dots, P. \quad (3)$$

Since the problem is disjoint for each individual voxel $\mathbf{x}_p$, the spectral parameters in (3) can be solved independently.

However, this approach does not take advantage of the spatial characteristics of the spectral parameters (especially for MRSI data). For example, the line shape of a spectrum is usually similar to those for the surrounding voxels because of the smoothness of field inhomogeneity and tissue properties (two primary factors affecting line shapes) in image objects. Moreover, metabolite concentrations are expected to be smooth within a same tissue. Therefore, we propose a joint estimation formulation to effectively exploit the spatial prior knowledge in $\boldsymbol{\theta}$ and $\mathbf{a}$ in a Bayesian framework, which is described in detail next.

B. Joint Estimation with Spatial Constraints

We rewrite (3) as a joint estimation problem to determine $\mathbf{a}$ and $\boldsymbol{\theta}$ from all the voxels simultaneously: let

$$\mathbf{d} = [\mathbf{d}_1^T, \mathbf{d}_2^T, \dots, \mathbf{d}_P^T]^T, \mathbf{a} = [\mathbf{a}_1^T, \mathbf{a}_2^T, \dots, \mathbf{a}_P^T]^T, \boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \dots, \boldsymbol{\theta}_P^T]^T, \text{ and}$$

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

the spectral estimation problem can be reformulated as

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

where $\mathbf{R}(\mathbf{a}, \boldsymbol{\theta})$ is a regularization term imposing constraints on $\mathbf{a}$ and $\boldsymbol{\theta}$ (note (4) would be equivalent to the subproblems in (3) if $\mathbf{R}(\mathbf{a}, \boldsymbol{\theta}) = 0$).

While many different choices can be used for $\mathbf{R}(\mathbf{a}, \boldsymbol{\theta})$, our current study focuses on imposing piece-wise smoothness regularization, which is motivated by the fact that in most biological

samples, those spectral parameters (e.g., metabolite concentrations and T_2^* relaxation times) have rather smooth spatial distributions within a same tissue type. In addition, we propose to impose regularization for each metabolite individually, i.e.,

$$R(\mathbf{a}, \boldsymbol{\theta}) = \sum_{m=1}^M R^{(m)}(\mathbf{a}, \boldsymbol{\theta}), \quad (5)$$

where the superscript “ (m) ” represents the m -th metabolite. More specifically, we use

$$R^{(m)}(\mathbf{a}, \boldsymbol{\theta}) = \lambda_m \|\mathbf{W}_{\boldsymbol{\theta}} \boldsymbol{\theta}^{(m)}\|_2^2 + \eta_m \|\mathcal{W}_{\mathbf{a}}\{\mathbf{a}^{(m)}\}\|_1, \quad (6)$$

Where λ_m and η_m are regularization parameters. The weighted ℓ_2 penalty term on $\boldsymbol{\theta}$ is applied to promote smoothness of $\boldsymbol{\theta}$, and the weighting matrix $\mathbf{W}_{\boldsymbol{\theta}}$ is designed to preserve edges [23]:

$$\|\mathbf{W}_{\boldsymbol{\theta}} \mathbf{x}\|_2^2 = \sum_{p_1=1}^P \sum_{\substack{p_2 \neq p_1 \\ p_2 \in \Omega_{p_1}}} w_{p_1, p_2} |x_{p_1} - x_{p_2}|^2, \quad (7)$$

Where w_{p_1, p_2} is a weighting coefficient, Ω_{p_1} is the set of voxels that are in the adjacent neighbor to voxel p_1 , and x_p is the p -th element of $\mathbf{x}$, $p = 1, \dots, P$. Conceptually, the weighting coefficients w_{p_1, p_2} should be large in smooth areas (to promote smoothness), and be small across edges (to preserve edges). One practical way to derive them from reference anatomical images was proposed by Haldar et al. [23].

The second penalty term in (6) is to impose spatial prior knowledge (i.e., piece-wise smoothness) in $\mathbf{a}$. A natural choice is to enforce sparsity in a transformed domain such as the wavelet transform and total variation (TV) transform [24]. In this work, we choose $\mathcal{W}_{\mathbf{a}}$ to be the second order total generalized variation (TGV) transform [25], [26]:

$$\|\mathcal{W}_{\mathbf{a}}\{\mathbf{u}\}\|_1 = \min_{\mathbf{p}_1, \mathbf{p}_2} \alpha_0 \|\nabla \mathbf{u} - \mathbf{p}\|_1 + \alpha_1 \sum_{j=1}^4 \|(\mathcal{E}(\mathbf{p}))_j\|_1, \quad (8)$$

where $\alpha_0, \alpha_1 > 0$ are two regularization parameters, and $(\mathcal{E}(\mathbf{p}))_j$ is the j -th entry of the vectorization of $\mathcal{E}(\mathbf{p})$. The operators ∇ and $\mathcal{E}$ are given by

$$\nabla \mathbf{u} = \begin{bmatrix} a_x \mathbf{u} \\ a_y \mathbf{u} \end{bmatrix},$$

$$\mathcal{E}(\mathbf{p}) = \begin{bmatrix} a_x \mathbf{p}_1 & \frac{1}{2}(\partial_y \mathbf{p}_1 + \partial_x \mathbf{p}_2) \\ \frac{1}{2}(\partial_y \mathbf{p}_1 + \partial_x \mathbf{p}_2) & \partial_y \mathbf{p}_2 \end{bmatrix},$$

where ∂_x and ∂_y denote the finite difference along the x and y directions, and $\mathbf{p}_1$ and $\mathbf{p}_2$ are the two segments of $\mathbf{p}$ corresponding to the lengths of $\partial_x \mathbf{u}$ and $\partial_y \mathbf{u}$, respectively. We can see from (8) that in an extreme case that $a_1 \gg a_0$, this second order TGV regularization is the same with TV regularization. As a generalization of TV, TGV seeks an optimized balance between different orders of smoothness, so it is generally better at preserving edges and intracompartment heterogeneity [25], [26]. Meanwhile, we would like to note that TGV is just one choice to demonstrate the impact of spatial constraints in improving MRSI quantification. Other regularization terms can also be included into the proposed approach for further improvement. Combining Eqs. (4)–(6) yields

$$(\hat{\mathbf{a}}, \hat{\boldsymbol{\theta}}) = \arg \min_{\mathbf{a}, \boldsymbol{\theta}} \|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{a}\|_2^2 + \sum_{m=1}^M \lambda_m \|\mathbf{W}_{\boldsymbol{\theta}} \boldsymbol{\theta}^{(m)}\|_2^2 + \sum_{m=1}^M \eta_m \|\mathcal{W}_{\mathbf{a}}\{\mathbf{a}^{(m)}\}\|_1. \quad (9)$$

III. Algorithm

The optimization problem in (9) is a large-scale nonlinear problem and computationally expensive to solve. To address this issue, we propose to solve the problem in two steps:

$$(\mathbf{a}_0, \hat{\boldsymbol{\theta}}) = \arg \min_{\mathbf{a}, \boldsymbol{\theta}} \|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{a}\|_2^2 + \sum_{m=1}^M \lambda_m \|\mathbf{W}_{\boldsymbol{\theta}} \boldsymbol{\theta}^{(m)}\|_2^2, \quad (10)$$

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \|\mathbf{d} - \mathbf{K}(\hat{\boldsymbol{\theta}})\mathbf{a}\|_2^2 + \sum_{m=1}^M \eta_m \|\mathcal{W}_{\mathbf{a}}\{\mathbf{a}^{(m)}\}\|_1, \quad (11)$$

where $\mathbf{a}_0$ in (10) will be used as the initial value for solving (11). These two problems are relatively easy to handle compared to the original problem in (9): (10) can be solved using a variable projection strategy [27], and (11) is an ℓ_1 -norm regularized linear least-squares problem which can be solved by many efficient algorithms.

A. Estimation of $\boldsymbol{\theta}$

The optimization problem in (10) can be reformulated as

$$\begin{cases} \hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})^\dagger \mathbf{d}\|_2^2 + \sum \lambda_m \|\mathbf{W}_{\boldsymbol{\theta}} \boldsymbol{\theta}^{(m)}\|_2^2, \\ \mathbf{a}_0 = \mathbf{K}(\hat{\boldsymbol{\theta}})^\dagger \mathbf{d}. \end{cases} \quad (12)$$

where “ $\dagger$ ” represents the Moore-Penrose pseudo inverse. The replacement of “ $\mathbf{a}$ ” in (10) by “ $\mathbf{K}(\boldsymbol{\theta})^\dagger \mathbf{d}$ ” follows the variable projection strategy [27]. A proof of equivalence to (10) can be found in [28].

One advantage of the variable projection strategy is a significant reduction in the dimensionality of parameter space (by decoupling the nonlinear and linear parameters). Meanwhile, the gradient of the cost function in (12) is still well defined, because

$\|\mathbf{d} - \mathbf{K}\mathbf{K}^\dagger \mathbf{d}\|_2^2$ is the standard variable projection functional as in [27], and $\|\mathbf{W}_{\boldsymbol{\theta}} \boldsymbol{\theta}^{(m)}\|_2^2$ is a quadratic term of $\boldsymbol{\theta}$ (thus differentiable with respect to $\boldsymbol{\theta}$). Therefore, with the initialization of $\boldsymbol{\theta}$ coming from a voxel-by-voxel method (e.g., QUEST), we can use the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm [29], which does not store the Hessian matrix explicitly and consumes less memory during run time (thus suitable for large-scale problems).

B. Estimation of $\mathbf{a}$

After we obtain an estimate of $\boldsymbol{\theta}$ (i.e., $\hat{\boldsymbol{\theta}}$), we solve the ℓ_1 -norm regularized least-squares problem in (11). Here we use an alternating direction method of multipliers (ADMM) [30] which is recently shown to be efficient in solving large-scale convex, non-smooth problems.

We first present the algorithm for the case that $\mathcal{W}_{\mathbf{a}}$ in (11) is a linear operator (e.g., the wavelet and TV transform) and can be represented by a matrix multiplication, i.e., $\mathcal{W}_{\mathbf{a}}\{\mathbf{a}^{(m)}\} = \mathbf{W}_{\mathbf{a}}\mathbf{a}^{(m)}$. As proposed by Guo et al. [26], we need an auxiliary variable $\mathbf{u}$, so that (11) is equivalent to

$$\begin{aligned} \min_{\mathbf{a}, \mathbf{u}} & \|\mathbf{d} - \mathbf{K}(\hat{\boldsymbol{\theta}})\mathbf{a}\|_2^2 + \sum_{m=1}^M \eta_m \|\mathbf{u}^{(m)}\|_1, \\ \text{s. t. } & \mathbf{u}^{(m)} = \mathbf{W}_{\mathbf{a}}\mathbf{a}^{(m)}, \text{ for } m=1, \dots, M. \end{aligned} \quad (13)$$

We can now use ADMM to solve (13). Let

$$L(\mathbf{a}, \mathbf{u}; \tilde{\mathbf{u}}) = \|\mathbf{d} - \mathbf{K}(\hat{\boldsymbol{\theta}})\mathbf{a}\|_2^2 + \sum_{m=1}^M \left[\eta_m \|\mathbf{u}^{(m)}\|_1 + \frac{\mu_m}{2} \|\mathbf{W}_{\mathbf{a}}\mathbf{a}^{(m)} - \mathbf{u}^{(m)} - \tilde{\mathbf{u}}^{(m)}\|_2^2 \right].$$

Then (13) can be solved iteratively using the following updating procedure, whose convergence is guaranteed by the ADMM algorithm [31]:

$$\mathbf{u}^{\text{new}} = \arg \min_{\mathbf{u}} L(\mathbf{a}^{\text{old}}, \mathbf{u}; \tilde{\mathbf{u}}^{\text{old}}), \quad (14a)$$

$$\mathbf{a}^{\text{new}} = \arg \min_{\mathbf{a}} L(\mathbf{a}, \mathbf{u}^{\text{new}}; \tilde{\mathbf{u}}^{\text{old}}), \quad (14b)$$

$$\tilde{\mathbf{u}}^{\text{new}} = \tilde{\mathbf{u}}^{\text{old}} + \gamma(\mathbf{u}^{\text{new}} - \mathbf{W}_{\mathbf{a}} \mathbf{a}^{\text{new}}). \quad (14c)$$

Subproblem (14a) can be solved explicitly using shrinkage. Subproblem (14b) is quadratic, and can be readily solved by standard convex optimization tools. Equation (14c) is simply an element-wise update, and γ is the selected step length.

As stated earlier, the iterative algorithm above is designed for linear $\mathcal{W}_{\mathbf{a}}$. Plugging the definition of TGV in (8) into (11), we can see the optimization problem is slightly different:

$$(\hat{\mathbf{a}}, \hat{\mathbf{p}}) = \arg \min_{\mathbf{a}, \mathbf{p}} \|\mathbf{d} - \mathbf{K}(\hat{\boldsymbol{\theta}}) \mathbf{a}\|_2^2 + \alpha_0 \sum_{m=1}^M \eta_m \|\nabla \mathbf{a}^{(m)} - \mathbf{p}^{(m)}\|_1 + \alpha_1 \sum_{m=1}^M \eta_m \sum_{j=1}^4 \|(\mathcal{E}(\mathbf{p}^{(m)}))_j\|_1. \quad (15)$$

However, the algorithm described in (14) can still be extended to this case, by defining

$$\tilde{\mathbf{a}} = \begin{bmatrix} \mathbf{a} \\ \mathbf{p} \end{bmatrix}, \tilde{\mathbf{d}} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}, \tilde{\mathbf{K}}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \mathbf{K}(\hat{\boldsymbol{\theta}}) & \\ & \mathbf{0} \end{bmatrix},$$

$\tilde{\mathbf{V}} = [\alpha_0 \nabla, -\alpha_0 \mathbf{I}]$, and $\tilde{\mathcal{E}}_j(\tilde{\mathbf{a}}^{(m)}) = (\mathcal{E}(\mathbf{p}^{(m)}))_j, j = 1, 2, 3, 4$. Then we can reformulate (15) as

$$\hat{\tilde{\mathbf{a}}} = \arg \min_{\tilde{\mathbf{a}}} \|\tilde{\mathbf{d}} - \tilde{\mathbf{K}}(\hat{\boldsymbol{\theta}}) \tilde{\mathbf{a}}\|_2^2 + \sum_{m=1}^M \eta_m \|\tilde{\mathbf{W}} \tilde{\mathbf{a}}^{(m)}\|_1, \quad (16)$$

where

$$\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\nabla} \\ \alpha_1 \tilde{\mathcal{C}}_1 \\ \vdots \\ \alpha_1 \tilde{\mathcal{C}}_4 \end{bmatrix}.$$

Therefore, when $\mathcal{W}_{\mathbf{a}}$ represents the TGV transform, the iterative algorithm described by (14) can still be used after the reformulation described in (16). For a more detailed summary of the algorithm, please refer to our supplementary file.

IV. Performance Characterization

To gain an insight into the quantitative improvement offered by the proposed method, we used the constrained Cramér-Rao bound (CRB) theory [20], [32] to analyze its performance. It is well known that if the parameter $\mathbf{a}$ in (11) is sparse in the transformed domain, then the problem in (11) can be considered as a relaxed version of the following constrained optimization problem:

$$\begin{aligned} \hat{\mathbf{a}} &= \arg \min_{\mathbf{a}} \|\mathbf{d} - \mathbf{K}(\hat{\boldsymbol{\theta}})\mathbf{a}\|_2^2, \\ \text{s. t. } \|\mathbf{W}_{\mathbf{a}}\mathbf{a}\|_0 &\leq S, \text{ for } m=1, \dots, M, \end{aligned} \quad (17)$$

where S is some integer dependent on $\{\eta_m\}_{m=1}^M$. Therefore, one way to characterize the performance of the proposed method is to characterize (17) instead. Assuming the estimated $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^*$ (we use “*” to denote the true parameter hereafter), the performance of the problem in (17) can be characterized by calculating its constrained CRB, which lower bounds the variances of estimated parameters. Specifically, the total variance of $\hat{\mathbf{a}}$ in (17) is bounded by

$$\sum_i \text{Var}(\hat{a}_i) \geq \text{Tr}(\mathbf{A}[\mathbf{A}^T \mathbf{F} \mathbf{A}]^{-1} \mathbf{A}^T) \triangleq B_c, \quad (18)$$

where $\hat{a}_i$ is the i -th element of $\hat{\mathbf{a}}$, $\mathbf{F}$ is the Fisher information matrix of (17), and $\mathbf{A}$ is selected to be those columns of $\mathbf{W}_{\mathbf{a}}^{-1}$ corresponding to the support of $\mathbf{W}_{\mathbf{a}}\mathbf{a}^{(m)}$ [33].

In contrast, the estimator in (3) has no spatial sparsity constraint, as it performs quantification voxel-by-voxel. Its performance can then be characterized by the unconstrained CRB, which can be obtained by letting $\mathbf{A}$ be the identity matrix:

$$\sum_i \text{Var}(\hat{a}_i) \geq \text{Tr}(\mathbf{F}^{-1}) \triangleq B_u. \quad (19)$$

We further define a quotient Q_F

$$Q_I = \frac{B_u}{B_c}, \quad (20)$$

which characterizes how much improvement the proposed method can possibly achieve, i.e., the amount of variance reduction by introducing spatial regularizations.

V. Experimental Setup

A. Simulation Study

A 2D MRSI simulation dataset was generated for evaluating and analyzing the performance of the proposed method. The synthetic data were generated using the following model

$$d(x, y, t) = \sum_{m=1}^6 a_m(x, y) e^{-t/T_{2,m}^*(x,y)} \varphi_m(t) + n(x, y, t),$$

The readout bandwidth was 2000 Hz, and the matrix size was 80×80. Six common NMR detectable metabolites in the human brain, n-acetylaspartate (NAA), creatine (Cre), choline (Cho), glutamate (Glu), myo-inositol (mIn), and glutamine (Gln), were used. The metabolite concentration distribution ($\{a_m\}_{m=1}^6$) were designed such that they are smooth within each type of tissue (gray matter, white matter, and cerebrospinal fluid, etc.), but different across different types of tissues. The concentration ratios among metabolites and T_2 values were adjusted based on literature values from [3] (we have chosen the concentration of total choline for Cho). The corresponding spectral basis functions, $\varphi_m(t)$, were obtained by quantum mechanics simulation (GAVA [16]; shown in Fig. 1a). The spectral ℓ_2 integration of the simulated dataset is shown in Fig. 1b. White Gaussian noise was added to the dataset (Fig. 1c). Three typical spectra at different SNR levels are shown in Fig. 1d.

B. In Vivo Study

In vivo experiments were carried out to demonstrate the effectiveness of the proposed method in practice. The MRSI data were acquired from a healthy volunteer on a Siemens 3 Tesla MRI scanner with a two-dimensional bipolar echo-planar spectroscopic imaging (EPSI) sequence, with outer-volume saturation (OVS) [34] and water suppression (CHESS [35]). The residual water and subcutaneous lipid signals were further removed using the union-of-subspaces method proposed by Ma et al. [36]. The readout bandwidth was 100 kHz, the echo time was 30 ms, the echo spacing was 1.42 ms, the scan time was about 31 mins, and the nominal resolution was 4.6 mm (field-of-view 220 mm×220 mm, matrix size 48×48). The reference image to derive $\mathbf{W}_\theta$ is from a co-registered MP RAGE [37] image.

VI. Results

A. Simulation Study

We have evaluated the performance of the proposed method and compared it with QUEST [13], a standard method used in practice. As is done in other state-of-the-art frequency-domain methods (LCModel [11], etc.) and time-domain methods (AQSES [14], etc.), QUEST performs spectral quantification voxel-by-voxel using a set of pre-determined spectral bases without incorporating spatial regularizations. Fig. 2 compares the quantification results from the simulation MRSI dataset for NAA, Cre, Cho, Glx (Glutamate + Glutamine) and mIn using QUEST and the proposed method, respectively. Fig. 3 shows several spectra synthesized using the estimated parameters and their corresponding errors with respect to (w.r.t.) the ground truths.

From Fig. 2, it can be seen that the estimated concentration maps by QUEST (row b) were significantly noisier (larger spatial variations) than those by the proposed method (row c), especially for metabolites with lower concentrations (e.g., choline). This improvement is further illustrated by the error maps and relative errors w.r.t. the ground truth (rows d and e in Fig. 2). The relative errors of each metabolite (shown as red numbers) are defined as

$$\delta = \frac{\|\mathbf{a} - \hat{\mathbf{a}}\|_2}{\|\mathbf{a}\|_2},$$

where $\mathbf{a}$ and $\hat{\mathbf{a}}$ are the true and the estimated concentration maps, respectively. Fig. 3 confirms this improvement: the spectra synthesized using parameters estimated by QUEST showed noticeable errors w.r.t. the ground truths, which were significantly reduced by the proposed method.

B. In Vivo Study

Spectral fittings from some representative voxels were evaluated as shown in Fig. 4 (the fitting was done jointly for NAA, Cre, Cho, Glu, Gln, and mIn). Both methods had residuals close to the noise floor, indicating comparable performance in terms of spectral fitting. The quantified concentration maps are shown in Fig. 5. It can be seen that the QUEST results showed significant spatial variations (indicating large estimation variance), which were reduced by the proposed method. This observation is consistent with the simulation results.

C. Performance Analysis

Simulation data were used to perform the CRB analyses as described in Section IV. To simplify the calculation, the sparsifying transform $\mathcal{W}_{\mathbf{a}}$ was chosen to be the wavelet transform. We assumed that the nonlinear parameter $\boldsymbol{\theta}$ was known a priori, then the CRB calculations for the proposed method and QUEST simply followed (18) and (19), respectively. Fig. 6a shows the calculated CRB for the proposed method (blue) and QUEST (red) at different sparsity levels. We can see that the proposed method has a much lower CRB compared to QUEST, indicating that the estimation variance of the proposed method (i.e., incorporating spatial sparsity constraints) can be smaller. The curve also indicates that

the incorporation of spatial sparsity constraints on $\mathbf{W}_a \mathbf{a}$ can be more effective when $\mathbf{W}_a \mathbf{a}$ is sparser, as expected. The difference in the theoretical lower bounds is consistent with the estimation variances obtained in Monte-Carlo simulations shown in Fig. 6b. The solid line in Fig. 6c plots the quotient Q_I defined in (20), which can be calculated by the ratio between the two lines in (a). Similar to Q_I , we also define Q'_I as the ratio between the two lines in (b), which is plotted as the dashed line in Fig. 6c. We can see the significant improvement resulting from incorporating sparsity regularizations from the lines in Fig. 6c. Note that when there are fewer non-zero entries in $\mathbf{W}_a \mathbf{a}$, the discrepancy between Q_I and Q'_I becomes larger, which is expected in practice because the optimization algorithm does not produce the oracle estimator [33].

In the analysis above we have assumed $\hat{\boldsymbol{\theta}}$ to be accurate. In practice, however, $\hat{\boldsymbol{\theta}}$ is a random vector and may not be exactly the same as the true $\boldsymbol{\theta}$. In this case, the CRB analyses for QUEST still follows (19) by calculating the Fisher information matrix treating both $\mathbf{a}$ and $\boldsymbol{\theta}$ as unknown parameters. As for the proposed method, we can absorb the model discrepancy caused by $\boldsymbol{\theta} = \boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}$ into an additional noise term. The resulting total noise thus has a larger variance, so that we can adjust the “SNR” accordingly when calculating the constrained CRB. Specifically, let $\mathbf{f}(\boldsymbol{\theta}) = \mathbf{K}(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})^\dagger \mathbf{d}$, and $\mathbf{J}_f$ be the Jacobian matrix of $\mathbf{f}$ at $\boldsymbol{\theta} = \boldsymbol{\theta}^*$, then we can linearize $\mathbf{f}$ by

$$\mathbf{f}(\boldsymbol{\theta}) \approx \mathbf{f}(\boldsymbol{\theta}^*) + \mathbf{J}_f(\boldsymbol{\theta}^* - \boldsymbol{\theta}).$$

Therefore, we can approximate (12) by

$$\begin{aligned} \hat{\boldsymbol{\theta}} &\approx \arg \min_{\boldsymbol{\theta}} \|\mathbf{b} - \mathbf{J}_f \boldsymbol{\theta}\|_2^2 + \|\mathbf{W} \boldsymbol{\theta}\|_2^2, \\ &= (\mathbf{J}_f^H \mathbf{J}_f + \mathbf{W}^H \mathbf{W})^{-1} \mathbf{J}_f^H \mathbf{b}, \end{aligned} \quad (21)$$

where $\mathbf{b} = \mathbf{d} - \mathbf{f}(\boldsymbol{\theta}^*) + \mathbf{J}_f \boldsymbol{\theta}^* = \mathbf{n} + \mathbf{J}_f \boldsymbol{\theta}^*$, and $\mathbf{W}$ is obvious. Let $\mathbf{h}(\boldsymbol{\theta}) = \mathbf{K}(\boldsymbol{\theta}) \mathbf{a}^*$, and $\mathbf{J}_h$ be the Jacobian matrix of $\mathbf{h}$ at $\boldsymbol{\theta} = \boldsymbol{\theta}^*$, then the model discrepancy in solving $\mathbf{a}$ caused by using $\hat{\boldsymbol{\theta}}$ instead of $\boldsymbol{\theta}^*$ can be modeled by an additional noise term $\mathbf{n}$ as follows.

$$\Delta \mathbf{n} = \mathbf{h}(\hat{\boldsymbol{\theta}}) - \mathbf{h}(\boldsymbol{\theta}^*) \approx \mathbf{J}_h \Delta \boldsymbol{\theta}. \quad (22)$$

Therefore, to perform CRB analysis for the case where $\boldsymbol{\theta}$ is not known a priori, we can first approximate $\hat{\boldsymbol{\theta}}$ by (21), then adjust the CRB according to a new “total noise variance” (the sum of the variances of $\mathbf{n}$ and the additional noise $\mathbf{n}$ given by (22)).

The adjusted CRB at different sparsity levels and corresponding Monte Carlo simulation results are shown in Fig. 7. By comparing with the corresponding results in Fig. 6, we can see that when $\boldsymbol{\theta}$ is not known a priori, the performance of the proposed method becomes

slightly worse, which is expected because of the increase in model order and nonlinearity introduced by θ . It can also be seen from Fig. 7 that when the ground truth is sparser, either the adjusted CRB or Monte Carlo simulations will be smaller, indicating again that the incorporation of spatial prior knowledge will be more powerful. The small gap between the two lines at the right end implies that the adjustment method is appropriate at least under this experimental setup.

Fig. 8 further demonstrates the reduction in estimation errors by incorporating spatial sparsity regularizations. We can observe that the trends of the two curves are consistent with those in Fig. 6c. The gap between Q_I and Q'_I implies potential improvement beyond the proposed joint estimation method.

VII. Discussion

We have demonstrated the effectiveness of jointly imposing spatio-spectral constraints in improving MRSI spectral quantification. Quantification methods incorporating spatial prior have been previously reported but in a rather restrictive form. For example, in [17], initial guess of the nonlinear optimization problem is modified based on estimates from neighboring voxels from a previous run. In LCMoel, a “center-out” sweeping strategy is used when dealing with multi-voxel spectroscopic signals, in order to get better starting estimates and “soft constraints” from preceding central voxels for outer voxels [38]. These early methods explore the smoothness of parameter maps to reduce the chance of being trapped in local minima. However, since the problem formulation is the same, the best performance achievable (i.e., CRB) remains unchanged.

Recent methods incorporate spatial prior knowledge as a regularization term, penalizing the parameter variation within neighboring voxels [18], [19]. The proposed method differs from these methods in the following key aspects. Firstly, the proposed method solves the regularized optimization problem jointly for all the voxels of interest, instead of using a specific “visiting schedule” (e.g., row-wise [18] or checkerboard [19]). Secondly, a more general piece-wise smoothness regularization (through using the edge-preserving TGV penalty) is used to better exploit the prior information in the spatial domain, jointly with the spatial regularization on nonlinear parameters (as done in [18], [19]). Thirdly, the spatial regularization terms applied to the nonlinear parameters are also edge-preserving, which is appealing because the T_2 relaxation times are well known to vary among different tissue types. Although it is discussed in [18] that a weighting scalar derived from MRI segmentation or classification can be applied to reflect tissue types, the proposed method constructs the edge-preserving weighting matrix $\mathbf{W}_\theta$ automatically without fine-tuning many parameters. Finally, theoretical performance analysis of the proposed method is provided based on constrained CRB, which has not been adequately done in related works. Some comparison between AQSES-MRSI [18] and the proposed method can be found in our supplementary file.

There are two issues that need to be carefully accounted for when applying the proposed method. The first one is that it requires manual selection of many hyperparameters, i.e., μ_m 's and γ in ADMM, α_0 and α_1 in the definition of TGV, and λ_m 's and η_m 's in (6). It has been

shown that ADMM converges if $\mu_m > 0$ and $0 < \gamma < (\sqrt{5}+1)/2$, but its converging rate is generally unknown [26]. As a result, the specific selection of μ_m and γ for fast convergence requires experience in the ADMM algorithm. α_0 and α_1 control the trade-off between penalizing first- and second-order finite differences of the concentration maps. In the results shown in this paper, we have consistently chosen $\alpha_1 = 10\alpha_0$, which have proven to be a reasonable choice for concentration maps. The other parameters, λ_m 's and η_m 's, are related to the noise level of MRSI signals and the distribution of underlying images (i.e., nonlinear parameter maps and concentration maps). While the noise variance can be readily computed from the tail of the MRSI signals, the distribution of underlying images can be estimated from a previous run of voxel-by-voxel approach. Based on this information, we can first select λ_m 's or η_m 's by making the data consistency term and spatial regularization terms in (10) and (11) comparable, and then tune them by visual inspection. Since η_m 's are the only hyperparameters that directly regularize spatial variation of concentration maps, the quantification results are fairly robust to the selection of other parameters according to our experience. Finally, we note that for a fixed acquisition protocol, these parameters typically only need to be optimized for one data set and used for others.

The other issue is that the joint estimation framework is computationally expensive to solve. Specifically, a typical nonlinear least squares solution algorithm (e.g., BFGS) has the computational complexity of $\mathcal{O}(n^2)$, where n is the number of parameters to estimate. In the case of quantification of MRSI signals, a voxel-by-voxel approach has the computational complexity of $\mathcal{O}(PN^2)$, where P is the total number of voxels to quantify and N is the number of parameters within each voxel. On the other hand, if the estimation is done jointly, the computational complexity will be $\mathcal{O}(P^2N^2)$, which is tremendously larger than voxel-by-voxel approaches given the huge number of voxels P . We have used several strategies to address the computation issue. Firstly, a previous run of a voxel-by-voxel approach is used for better initialization and convergence of our joint estimation. Secondly, we decoupled the estimation of θ and $\mathbf{a}$ into two steps to simplify the optimization problem suitable for efficient algorithms (i.e., L-BFGS and ADMM). Finally, for the solution of nonlinear parameters, a variable-projection strategy is used to further reduce the number of parameters. The combination of these strategies has reduced the computational time significantly, although more advanced algorithms can be developed to better address the computational problem, especially when extending the proposed method to volumetric MRSI data.

VIII. Conclusion

We have presented a new method for spectral quantification from noisy MRSI data. The proposed method jointly estimates the spectral parameters from all the voxels of interest, enabling the effective use of both spatial and spectral constraints. Both simulation and experimental results demonstrated a significant improvement in estimation accuracy by the proposed method, which is also theoretically confirmed by a constrained Cramér-Rao bound analysis. The proposed method may prove useful for spectral quantification in various MRSI studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work presented in this paper was supported in part by the National Institutes of Health (NIH-1RO1-EB013695 and NIH-R21EB021013-01) and Beckman Postdoctoral Fellowship.

References

1. Gonen O, Oberndorfer TA, Inglese M, Babb JS, Herbert J, Grossman RI. Reproducibility of three whole-brain N-acetylaspartate decline cohorts in relapsing-remitting multiple sclerosis. *American journal of neuroradiology*. 2007; 28(2):267–271. [PubMed: 17296992]
2. Gomes WA, Lado FA, de Lanerolle NC, Takahashi K, Pan C, Hetherington HP. Spectroscopic imaging of the pilocarpine model of human epilepsy suggests that early NAA reduction predicts epilepsy. *Magnetic resonance in medicine*. 2007; 58(2):230–235. [PubMed: 17654595]
3. de Graaf, RA. *In Vivo NMR Spectroscopy: Principles and Techniques*. John Wiley & Sons; 2013.
4. Barkhuijsen H, de Beer R, van Ormondt D. Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals. *Journal of Magnetic Resonance*. 1987;553–557.
5. Pijnappel WWF, van den Boogaart A, de Beer R, van Ormondt D. SVD-based quantification of magnetic resonance signals. *Journal of Magnetic Resonance*. 1992; 97(1):122–134.
6. Barkhuijsen H, de Beer R, Bovée WMMJ, van Ormondt D. Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure. *Journal of Magnetic Resonance*. 1985; 61(3):465–481.
7. Vanhuffel S, Chen H, Decanniere C, van Hecke P. Algorithm for time-domain NMR data fitting based on total least squares. *J Magn Reson Ser A*. 1994; 110(2):228–237.
8. Chen H, van Huffel S, van Ormondt D, de Beer R. 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*. 1996; 119(2):225–234.
9. Stoica P, Selén Y, Sandgren N, van Huffel S. Using prior knowledge in SVD-based parameter estimation for magnetic resonance spectroscopy-the ATP example. *Biomedical Engineering, IEEE Transactions on*. 2004; 51(9):1568–1578.
10. van der Veen JWC, de Beer R, Luyten PR, van Ormondt D. Accurate quantification of in vivo ^{31}P NMR signals using the variable projection method and prior knowledge. *Magnetic Resonance in Medicine*. 1988; 6(1):92–98. [PubMed: 3352510]
11. Provencher SW. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magnetic Resonance in Medicine*. 1993; 30(6):672–679. [PubMed: 8139448]
12. Vanhamme L, van den Boogaart A, van Huffel S. Improved method for accurate and efficient quantification of MRS data with use of prior knowledge. *Journal of Magnetic Resonance*. 1997; 129(1):35–43. [PubMed: 9405214]
13. Ratiney H, Sdika M, Coenradie Y, Cavassila S, van Ormondt D, Graveron-Demilly D. Time-domain semi-parametric estimation based on a metabolite basis set. *NMR in Biomedicine*. 2005; 18:1–13. [PubMed: 15660450]
14. Pouillet JB, Sima DM, Simonetti AW, de Neuter B, Vanhamme L, Lemmerling P, van Huffel S. An automated quantitation of short echo time MRS spectra in an open source software environment: AQSES. *NMR in Biomedicine*. 2007; 20(5):493–504. [PubMed: 17167819]
15. Young K, Govindaraju V, Soher BJ, Maudsley AA. Automated spectral analysis I: Formation of a priori information by spectral simulation. *Magnetic Resonance in Medicine*. 1998; 40(6):812–815. [PubMed: 9840824]
16. Soher BJ, Young K, Bernstein A, Aygula Z, Maudsley AA. GAVA: spectral simulation for in vivo MRS applications. *Journal of Magnetic Resonance*. 2007; 185(2):291–299. [PubMed: 17257868]

17. Soher BJ, Young K, Govindaraju V, Maudsley AA. Automated spectral analysis III: application to in vivo proton MR spectroscopy and spectroscopic imaging. *Magnetic Resonance in Medicine*. 1998; 40(6):822–831. [PubMed: 9840826]
18. Croitor Sava AR, Sima DM, Poulet J-B, Wright AJ, Heerschap A, van Huffel S. Exploiting spatial information to estimate metabolite levels in two-dimensional MRSI of heterogeneous brain lesions. *NMR in Biomedicine*. 2011; 24(7):824–835. [PubMed: 21834006]
19. Kelm BM, Kaster FO, Henning A, Weber MA, Bachert P, Boesiger P, Hamprecht FA, Menze BH. Using spatial prior knowledge in the spectral fitting of MRS images. *NMR in Biomedicine*. 2012; 25(1):1–13. [PubMed: 21538636]
20. Gorman JD, Hero AO. Lower bounds for parametric estimation with constraints. *Information Theory, IEEE Transactions on*. 1990; 36(6):1285–1301.
21. Ning, Q., Ma, C., Liang, Z-P. Spectral estimation for magnetic resonance spectroscopic imaging with spatial sparsity constraints. *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on; IEEE; 2015. p. 1482-1485.*
22. Graveron-Demilly D, Diop A, Briguet A, Fenet B. Product-operator algebra for strongly coupled spin systems. *Journal of Magnetic Resonance, Series A*. 1993; 101(3):233–239.
23. Haldar JP, Hernando D, Song SK, Liang ZP. Anatomically constrained reconstruction from noisy data. *Magnetic Resonance in Medicine*. 2008; 59(4):810–818. [PubMed: 18383297]
24. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*. 2007; 58(6):1182–1195. [PubMed: 17969013]
25. Bredies K, Kunisch K, Pock T. Total generalized variation. *SIAM Journal on Imaging Sciences*. 2010; 3(3):492–526.
26. Guo W, Qin J, Yin W. A new detail-preserving regularization scheme. *SIAM Journal on Imaging Sciences*. 2014; 7(2):1309–1334.
27. Golub GH, Pereyra V. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal of Numerical Analysis*. Apr.1973 10(2)
28. Ning, Q. MS thesis. University of Illinois; Urbana-Champaign: Dec. 2015 Spectral estimation with spatio-spectral constraints for magnetic resonance spectroscopic imaging.
29. Nocedal, J., Wright, S. Numerical optimization, series in operations research and financial engineering. Springer; New York, USA: 2006.
30. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® Machine Learning*. 2011; 3(1):1–122.
31. He B, Liao LZ, Han D, Yang H. A new inexact alternating directions method for monotone variational inequalities. *Mathematical Programming*. 2002; 92(1):103–118.
32. Ben-Haim Z, Eldar YC. The Cramér-Rao bound for estimating a sparse parameter vector. *Signal Processing, IEEE Transactions on*. 2010; 58(6):3384–3389.
33. Lam, F., Ma, C., Liang, Z-P. Performance analysis of denoising with low-rank and sparsity constraints. *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on; IEEE; 2013. p. 1223-1226.*
34. Le Roux P, Gilles RJ, McKinnon GC, Carlier PG. Optimized outer volume suppression for single-shot fast spin-echo cardiac imaging. *Journal of Magnetic Resonance Imaging*. 1998; 8:1022–1032. [PubMed: 9786138]
35. Haase A, Frahm J, Hanicke W, Matthaei D. ¹H NMR chemical shift selective (CHESS) imaging. *Phys Med Biol*. 1985; 30:341–344. [PubMed: 4001160]
36. Ma C, Lam F, Johnson CL, Liang ZP. Removal of nuisance signals from limited and sparse ¹H MRSI data using a union-of-subspaces model. *Magnetic Resonance in Medicine*. 2016; 75(2):488–497. [PubMed: 25762370]
37. Brant-Zawadzki M, Gillan GD, Nitz WR. MP RAGE: a three-dimensional, T1-weighted, gradient-echo sequence—initial experience in the brain. *Radiology*. 1992; 182(3):769–775. [PubMed: 1535892]
38. Provencher, S. [Accessed Jun 3, 2016] LCModel & LCMgui user's manual. 2015. <http://s-provencher.com/pub/LCModel/manual/manual.pdf>

Fig. 1.

(a) Spectral basis functions of the six metabolites used to synthesize the simulation dataset (TE=30 ms); note that the bases of NAA and Cho are both truncated to be a single peak. (b) Whole spectral integral of the synthesized dataset. (c) Spectral integral after adding white Gaussian noise. (d) Representative noisy spectra (top to bottom: gray matter, white matter, CSF), with their locations indicated by the red dots in (c).

Fig. 2.

Simulation results: (a) true metabolite distributions for NAA, Cre, Cho, Glx, and mIn; (b) metabolite distributions estimated by QUEST; (c) metabolite distributions estimated by the proposed joint estimation method; (d) estimation errors for QUEST w.r.t. ground truth; (e) estimation errors for the proposed method, with the relative errors shown as red numbers (the gray scale of rows d and e are 5 times higher than the other rows). Because of the well-known difficulty in separating Glu from Gln in a single-TE data, we are showing Glx (Glu + Gln) instead. Note the significant improvement of the proposed method over QUEST.

Fig. 3.

Simulation results: (a) ground truth spectra; (b) spectra estimated by QUEST (upper part) and errors (lower part) compared to the ground truth in (a); (c) spectra estimated by the proposed method (upper part) and errors (lower part). Voxel locations and order of each column are shown on the left.

Fig. 4.

The spectral fitting results from some representative voxels for the in vivo data. Locations of the voxels are marked in Fig. 5 (left to right: circle, triangle, and square). (a) Spectral fitting using QUEST (black: original spectra, red: fitted spectra); (b) fitting residuals of (a); (c) spectral fitting using the proposed method; (d) fitting residuals of (c). Note the similar levels of residuals for the two methods.

Fig. 5.

The estimated concentration maps by QUEST (top) and the proposed method (bottom) from the in vivo EPSI data, with the corresponding anatomical image (T_1 weighted) and spatial weighting coefficients used in (7) shown on the left.

Fig. 6.

Performance analysis of the proposed method at different sparsity levels (i.e., non-zero entries in $\mathbf{W}_a \mathbf{a}$) assuming $\boldsymbol{\theta}$ is known a priori. (a) CRB for the proposed method (blue) and QUEST (red). Note that the proposed method achieves lower CRB if the parameter is sparser, while the CRB of QUEST is approximately a constant. (b) The total variance of concentration maps obtained from corresponding Monte Carlo simulations using the proposed method (blue) and QUEST (red), respectively. Note the two lines in (b) are consistent with but strictly higher than the lines in (a). (c) Performance improvement of the proposed method over QUEST, obtained by comparing their CRB (solid), or the total variance from Monte Carlo simulations (dashed).

Fig. 7.

Performance analysis of the proposed method at different sparsity levels (i.e., non-zero entries in $W_a \mathbf{a}$) assuming $\boldsymbol{\theta}$ is not known a priori. Solid line: CRB after noise level adjustment (described in the text). Dashed line: the total variance of $\hat{\mathbf{a}}$ from Monte Carlo simulations.

Fig. 8.

Performance improvement factor R/R' after adjusting the noise level. Solid line: the ratio of the CRBs between QUEST and the proposed. Dashed line: the ratio of the total variance from Monte Carlo simulations between QUEST and the proposed.