Improved Model-Based Magnetic Resonance Spectroscopic Imaging

Mathews Jacob*, Xiaoping Zhu, Andreas Ebel, Norbert Schuff, and Zhi-Pei Liang

Abstract—Model-based techniques have the potential to reduce the artifacts and improve resolution in magnetic resonance spectroscopic imaging, without sacrificing the signal-to-noise ratio. However, the current approaches have a few drawbacks that limit their performance in practical applications. Specifically, the classical schemes use less flexible image models that lead to model misfit, thus resulting in artifacts. Moreover, the performance of the current approaches is negatively affected by the magnetic field inhomogeneity and spatial mismatch between the anatomical references and spectroscopic imaging data. In this paper, we propose efficient solutions to overcome these problems. We introduce a more flexible image model that represents the signal as a linear combination of compartmental and local basis functions. The former set represents the signal variations within the compartments, while the latter captures the local perturbations resulting from lesions or segmentation errors. Since the combined set is redundant, we obtain the reconstructions using sparsity penalized optimization. To compensate for the artifacts resulting from field inhomogeneity, we estimate the field map using alternate scans and use it in the reconstruction. We model the spatial mismatch as an affine transformation, whose parameters are estimated from the spectroscopy data.

Index Terms—Constrained reconstruction, inhomogeneity compensation, prior information, spectroscopic imaging.

I. INTRODUCTION

MAGNETIC resonance spectroscopic imaging (MRSI) is emerging as a very useful technique for the diagnosis and staging of various diseases such as cancer [1], [2], sclerosis [3], and epilepsy [4]. Unfortunately, the practical utility of this technique has been hampered by its poor sensitivity because of the low *in vivo* concentrations of the metabolites; the use of standard acquisition-reconstruction schemes leads to clinically unacceptably long scan times and noisy reconstructions.

The classical MRSI acquisition scheme is chemical shift imaging (CSI) [5]–[7], where one excitation is required to acquire each spatial phase encode. This limits the number of acquisitions possible in a clinically acceptable scan time. In this context, the standard IFFT-based reconstruction technique provides poor spatial resolution, leading to significant truncation artifacts. Specifically, the slow decay of the sinc point-spread

function results in cross talk between spectra from adjacent spatial regions. This artifact, known as spectral leakage, makes the intrapretation of the reconstructed data difficult. Moreover, magnetic susceptibility differences between air and tissue introduce strong local variations in the magnetic field, leading to spectral line-shape distortions and signal loss. All the above adverted to errors propagate to the subsequent quantification step, thus affecting the reliability and reproducibility of the results.

Recently, several methods that use readout gradients to provide extended k-space coverage in the same scan-time were proposed [8]–[12]. The smaller voxel size offered by these schemes provides a reduction in spectral leakage and inhomogeneity induced signal losses. However, these techniques suffer from a proportional decrease in signal-to-noise ratio (SNR). Moreover, since images have a decaying power spectrum, the higher k-space points often have significantly lower SNRs as compared to the lower k-space samples. Thus, the performance gain in using these techniques is limited.

An alternative approach is to exploit the information extracted from anatomical scans to provide the spatial localization [13]–[17]. These schemes use side information to generate a compartmental model. Since spatial localization is ensured by the model, the scan time can be devoted to acquire the lower k-space encodes with higher SNR. Thus, the model-based schemes offer an attractive framework to ameliorate spatial resolution (and thus eliminate the associated artifacts) without compromising the SNR. In spite of their advantages, the use of the model-based schemes in spectroscopic imaging applications is limited, mainly because of some of the technical problems that limit their performance. The main goal of this paper is to identify some of these problems and propose efficient solutions to them, thereby developing a new scheme that is useful for *in vivo* applications.

The current model-based schemes [13]–[15] limit the number of basis functions to obtain well-posed reconstructions. This restricts the flexibility of the model, thus leading to model misfit and consequently artifacts. Since these approaches use basis functions that are different from the measurement functions,[1] the model-based schemes are more vulnerable to the model-misfit than the Fourier schemes. The signal components that are not captured by the model alias back to the reconstructions as artifacts. We propose a more flexible model with both compartmentalized and shift-invariant basis functions; the compartmental basis functions represent the smooth variations inside the compartments, while the shift-invariant functions capture the localized perturbations because of lesions or segmentation errors.

*M. Jacob is with the Biomedical Engineering Department, University of Rochester, Rochester, NY 14622 USA.

X. Zhu, A. Ebel, and N. Schuff are with the Veterans Association Medical Center, University of California, San Francisco, CA 94121 USA.

Z.-P. Liang is with the Beckman Institute, University of Illinois, Urbana–Champaign, IL 61820 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

---

[1]Since we are sampling on the Fourier space, the measurement functions are Fourier exponentials; the measurements are the inner products of the signal with the exponentials.

We then use a sparse reconstruction scheme to ensure robust and unique reconstruction. Since this scheme adaptively selects the basis vectors from a large class to represent the signal, we obtain fewer artifacts than standard schemes.

The inhomogeneity of the magnetic field often leads to local spatial variations that the model is not able to capture, thus degrading the performance of the model-based schemes. In standard MRSI schemes (e.g., SLIM [13], GSLIM [14], and Fourier inversion [18]), it is a general practice to ignore the field map in order to have a simple reconstruction algorithm [18], [19]. As discussed previously, this may lead to line-shape distortions and signal losses. Following our early work [20], we propose to estimate the field inhomogeneity using additional scans and use it in the reconstruction procedure. This approach leads to fewer artifacts and more accurate results, even when used with simple image models. This scheme can be viewed as the generalization of the inhomogeneity compensated SLIM approach introduced in [21] and [22].

Spatial mismatch between the anatomical scans and MRSI data can lead to significant distortions in the model-based schemes. One reason for the mismatch is the incoherence between the sequences used to acquire the anatomical and MRSI scans. Even when the discrepancy is small, it can lead to severe contamination of the brain spectra by extra-cranial lipid signals. We propose to model this deformation as an affine transformation; the affine parameters are iteratively estimated from the MRSI data during the reconstructions.

The rest of the paper is organized as follows. In the next section, we briefly review the image formation and the standard model based reconstruction procedure. In Section III, we introduce the improvements over the standard model based approach. In Section IV, we pose the problem as an optimization algorithm; we use an iterative reconstruction procedure to solve for the coefficients and the affine parameters simultaneously. In Section V, we perform an analysis of the reconstruction error in a general model-based reconstruction scheme that provides more insight on model based reconstruction schemes. Finally, in Section VI, we validate the algorithm using both experimental and simulation data.

II. PRELIMINARIES

The image formation in MRSI is mathematically modeled as

$$\theta_{\text{meas}}(\mathbf{k}; k_f)$$
$$= \int_{\mathbf{r} \in \mathbb{R}^3} \int_{f \in \mathbb{R}} \rho(\mathbf{r}, f) e^{-j(\mathbf{k}; k_f)^T (\mathbf{r}; f)} e^{-j\gamma \Delta B_0(\mathbf{r}) k_f} d\mathbf{r} df. \quad (1)$$

Here, $\mathbf{k} = (k_x, k_y)^T$ indicates the phase encoding location and k_f denotes the sampling location along the readout axis (time). k_x, k_y, and k_f are the conjugate axes of x, y, and f, respectively. We denote the set of acquired phase encodes as $\mathbf{K} = \{\mathbf{k}_m; m = 0 \ldots M-1\}$. Similarly, the sampling locations along the readout axis are indicated by $\{k_{f_n}; n = 0 \ldots N-1\}$. $\Delta B_0(\mathbf{r})$ specifies the magnetic field inhomogeneity map and γ is the gyromagnetic ratio. Note that if the magnetic field is homogeneous ($\Delta B_0(\mathbf{r}) = 0$), (1) reduces to a simple Fourier transform.

A. Model-Based Reconstruction

The reconstruction of the continuous domain ρ from the finite set of measurements (1) is an ill-posed problem. The standard approach is to use a parametric image model, with finite degrees of freedom, to constrain the reconstructions. All the current MRSI image models are linear

$$s(\mathbf{r}, f) = \sum_{p=0}^{P-1} c_p(f) \varphi_p(\mathbf{r}) = \mathbf{c}(f)^H \boldsymbol{\varphi}(\mathbf{r}) \quad (2)$$

where $\varphi_p; p = 0, \ldots P - 1$ are arbitrary basis functions. Here, $\mathbf{c}(f) = [c_0(f) \ldots c_{P-1}(f)]^H$ and $\boldsymbol{\varphi}(\mathbf{r}) = [\varphi_0(\mathbf{r}) \ldots \varphi_{P-1}(\mathbf{r})]^T$. $\mathbf{c}(f)$ are the unknown coefficients that are estimated from the data. The widely used Fourier model corresponds to $\boldsymbol{\varphi} = \mathbf{e}$, where $\mathbf{e} = [e^{j\mathbf{k}_0^T \mathbf{r}}, \ldots, e^{j\mathbf{k}_{M-1}^T \mathbf{r}}]$ is the vector of Fourier exponentials. $\mathbf{c}^H$ is the Hermitian (conjugate transpose) of $\mathbf{c}$. The SLIM model corresponds to $\varphi_p(\mathbf{r}) = \chi_p(\mathbf{r})$, where

$$\chi_p(\mathbf{r}) = \begin{cases} 1, & \text{if; } \mathbf{r} \text{ is in the } p \text{ tissue class} \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

Plugging (2) into (1) and setting $\Delta B_0(\mathbf{r}) = 0$, we obtain

$$\theta_{\text{sim}}(\mathbf{k}, k_f) = \hat{\mathbf{c}}(k_f)^H \hat{\varphi}(\mathbf{k}) \quad (4)$$

where $\hat{\varphi}$ denotes the 3-D Fourier transform of $\boldsymbol{\varphi}$ and $\hat{\mathbf{c}}(k_f)$ is the Fourier transform of $\mathbf{c}(f)$. The reconstruction problem is now posed as a least square minimization problem; we minimize the discrepancy between the simulated measurements (4) and the experimental measurements

$$\mathcal{C} = \sum_n \underbrace{\left(\sum_{m=0}^{M-1} |\theta_{\text{sim}}(\mathbf{k}_m, k_{f_n}) - \theta_{\text{meas}}(\mathbf{k}_m, k_{f_n})|^2 \right)}_{\mathcal{C}_{k_{f_n}}}. \quad (5)$$

Here, k_{f_n} indicates the sampling locations along time of k_f axis.

B. Point-Spread Functions

When the spatial basis functions are independent of k_f, the estimation of $\hat{\mathbf{c}}(k_f)$, for a specified k_f, can be considered independently of the coefficient vector at other time points; it is obtained as the minimization of $\mathcal{C}_{k_f}$ in (5)

$$\hat{\mathbf{c}}_{\text{est}}(k_f) = \mathbf{Q}\boldsymbol{\theta}_{\text{meas}}(k_f) \quad (6)$$

where $\mathbf{Q}$ is an appropriate reconstruction matrix and $\hat{\rho}(k_f)$ indicate the vector of measurements corresponding to the specified k_f. Computing the inverse Fourier transform of (6) along k_f, and since $\boldsymbol{\theta}_{\text{meas}}(f) = \langle \mathbf{e}, \rho(f) \rangle$, we get

$$\mathbf{c}_{\text{est}}(f) = \mathbf{Q}\hat{\boldsymbol{\theta}}_{\text{meas}}(f)$$
$$= \mathbf{Q} \int_{\mathbb{R}^2} \mathbf{e}^*(\mathbf{r}) \hat{\rho}(\mathbf{r}, f) d\mathbf{r}$$
$$= \int_{\mathbb{R}^2} \underbrace{\mathbf{Q}\mathbf{e}^*(\mathbf{r})}_{\tilde{\boldsymbol{\varphi}}^*} \rho(\mathbf{r}, f) d\mathbf{r} \quad (7)$$
$$= \langle \rho(\mathbf{r}, f), \tilde{\boldsymbol{\varphi}}(\mathbf{r}) \rangle_{L_2}. \quad (8)$$

Here, $\mathbf{e} = [e^{j\mathbf{k}_0^T \mathbf{r}}, \dots, e^{j\mathbf{k}_{M-1}^T \mathbf{r}}]$ is the vector of the exponentials and $\mathbf{e}^*$ indicates its conjugate. In the above equations, $\boldsymbol{\theta}_{\text{meas}}(f)$ is the inverse Fourier transform of $\hat{\boldsymbol{\rho}}_{\text{meas}}(k_f)$. Thus, we obtain the coefficients as the inner-product between ρ and the corresponding $\tilde{\boldsymbol{\varphi}}$ (termed as point-spread functions [15]).

III. PROPOSED APPROACH

In this section, we describe the new method with several improvements over the standard model-based approaches to improve the reconstructions.

A. New Image Model

The proposed model represents the MRSI signal as a linear combination of two sets of functions: 1) compartmental basis functions and 2) local basis functions. These sets have complementary properties as discussed later; by considering the linear combination, the proposed scheme benefits from the advantages of both.

1) Compartmental Basis Functions: Our goal is to introduce a compartmental model that can represent smooth variations in the compartments while being flexible to capture the abrupt transitions at the compartment boundaries. While a single band-limited model can capture the smooth signal variations reasonably well, it is often not sufficient to represent the discontinuities (unless the bandwidth is sufficiently large). We propose to represent the signal as a weighed sum of N_c band-limited functions

$$s(\mathbf{r}, f) = \sum_{i=0}^{N_c - 1} \chi_i(\mathbf{r}) \xi_i(\mathbf{r}, f) \qquad (9)$$

where N_c is the number of compartments, χ_i is the characteristic function of the ith compartment, and $\xi_i(\mathbf{r}, f); i = 0, \dots, N_c - 1$ (for a specified frequency f) are band-limited functions. Here, we represent the signal at a specified frequency f as the weighted sum of N_c band-limited functions; the binary weights are specified by the characteristic functions. While $\rho_i; i = 0 \dots N_c - 1$ capture the smooth variations within the ith compartment, weighting them by χ_i enables us to represent the discontinuities at the compartment boundaries, if any. Replacing the band-limited functions by a linear combination of exponentials, we get

$$s(\mathbf{r}, f) = \sum_{i=0}^{N_c - 1} \chi_i(\mathbf{r}) \underbrace{\sum_{l=0}^{L_i - 1} c_{i,l}(f) e^{j\boldsymbol{\omega}_l^T \mathbf{r}}}_{\xi_i(\mathbf{r}, f)}$$

$$= \sum_{i=0}^{N_c - 1} \sum_{l=0}^{L_i - 1} c_{i,l}(f) \underbrace{e^{j\boldsymbol{\omega}_l^T \mathbf{r}} \chi_i(\mathbf{r})}_{\varphi_{i,l}(\mathbf{r})}. \qquad (10)$$

In the above equation, $e^{j\boldsymbol{\omega}_l^T \mathbf{r}}; p = 0 \dots L_i - 1$ are the exponentials of the band-limited model for the ith compartment. Note that (10) is still a linear model with the basis set as

$$\Phi = \{\varphi_{i,l}; i = 0, \dots, N_c - 1, l = 0, \dots, L_i - 1\}. \qquad (11)$$

In general, the number of exponentials in each compartment can be chosen differently. However, in our applications, we choose $L_i = L; \forall i$. Note that if we select $L = 1$ and set $\boldsymbol{\omega}_0 = 0; \forall i$, we obtain the standard SLIM model.

The assumption of voxels belonging to any one of the tissue class is valid only when we have very high-resolution anatomical references. In 2-D CSI, it is common to choose a thick slab (e.g., thickness of 15 mm) to obtain a reasonable SNR. In this case, a voxel in the reconstructed image would comprise of contributions from multiple tissue classes. Hence, we replace the characteristic function $\chi_k(\mathbf{r})$ in (10) by the partial volume function of the kth compartment: $0 \leq \text{PV}_k(\mathbf{r}) \leq 1$; it indicates the relative concentration of a tissue class at the spatial location $\mathbf{r}$. These functions are obtained by the classification of the anatomical references using standard segmentation algorithms. When we image a thick slab, a few nonoverlapping anatomical slices of the same region are acquired. These slices are segmented and averaged along the z dimension to get the partial volume functions.

2) Local Basis Functions: It is clear that the piecewise exponential model (10) captures the smooth variations within the compartments and the discontinuities at the region boundaries. However, when there are blob-like local perturbations inside the compartments because of lesions or segmentation errors, the piecewise band-limited model is not very efficient to represent them; we require a high value of L to approximate them reasonably well.

However, shift-invariant representations such as the *B*-spline scheme [23] are capable of capturing local perturbations efficiently. Intuitively, a localized perturbation can be represented by a few *B*-spline functions around that point. If we use an exponential model, such a perturbation would lead to significant values for all the Fourier coefficients. The cubic *B*-spline representation of a general 2-D signal is

$$s(\mathbf{r}, f) = \sum_{k=0}^{(K-1)^2} d_k(f) \gamma(\mathbf{r} - \mathbf{r}_k) \qquad (12)$$

where $\mathbf{r}_k \in [0, \dots, (K-1)^2]$ indicate the points on grid on which we reconstruct the signal, and γ is the cubic *B*-spline function on the specified grid. We denote the corresponding basis set as

$$\Gamma = \left\{\gamma(\mathbf{r} - \mathbf{r}_k); k = 0 \dots (K-1)^2\right\}. \qquad (13)$$

3) Optimality Criterion: As mentioned before, we represent the signal as a linear combination of basis vectors in the set

$$\Theta = \{\Phi, \Gamma\}. \qquad (14)$$

The individual sets Φ and Γ have complementary advantages and disadvantages. Φ is efficient to capture the smooth variations in the compartments and sharp transitions at the region boundaries. However, it is not very efficient to capture local perturbations; we would require a large order model to represent them. At the same time, while Γ can represent the local perturbations using few coefficients, we would require all the basis functions of Γ to capture the entire signal. By considering the joint set, we profit from the advantages of both these

schemes. However, since this set is redundant, the standard criterion (5) for solving the coefficients will give nonunique reconstructions. Hence, we pose the reconstruction as a sparse optimization problem. The objective is to choose the smallest subset of Θ that is optimal to represent the measured data. Clearly, the smallest subset is linearly independent and hence leads to a robust reconstruction. Note that Φ is subject-adaptive and Γ is efficient to capture local perturbations. Hence, we expect a good fit to the measured data using a few basis functions.[2] We formulate the reconstruction as the minimization of the criterion

$$\mathcal{C}_0 = \|\boldsymbol{\theta}_{\text{sim}} - \boldsymbol{\theta}_{\text{meas}}\|_{\ell_2}^2 + \lambda \left(\|\mathbf{d}\|_{\ell_0} + \|\mathbf{c}\|_{\ell_0} \right) \quad (15)$$

where $\|\mathbf{c}\|_{\ell_0}$ indicates the number of nonzero coefficients in $\mathbf{c}$. Here, λ is a regularization parameter that controls the number of nonzero coefficients. In (15), we obtain θ_{sim} by using (1), where $s(\mathbf{r}, f)$ is given by

$$s(\mathbf{r}, f) = \sum_{i=0}^{N_c-1} \sum_{l=0}^{L-1} c_{i,l}(f) \varphi_{i,l}(\mathbf{r}) + \sum_{k=0}^{(K-1)^2} d_k(f) \gamma(\mathbf{r} - \mathbf{r}_k)$$
$$= \mathbf{c}(f)^H \boldsymbol{\varphi}(\mathbf{r}) + \mathbf{d}(f)^H \boldsymbol{\gamma}(\mathbf{r}). \quad (16)$$

Note that by minimizing (15), we choose a subset from the over-complete basis (14) that is efficient to represent the signal. This subset is linearly independent; for every linearly dependent subset, there exists a linearly independent set that offers a sparser representation.

Equation (15) is a nonconvex optimization problem; its minimization would involve a combinatorial search, whose computational complexity is prohibitive for practical applications. To make the reconstruction procedure computationally tractable, it is common [24], [25] to approximate (15) by

$$\mathcal{C}_1 = \underbrace{\|\boldsymbol{\theta}_{\text{sim}} - \boldsymbol{\theta}_{\text{meas}}\|_{\ell_2}^2}_{\text{data consistency}} + \lambda \underbrace{\left(\|\mathbf{d}\|_{\ell_1} + \|\mathbf{c}\|_{\ell_1} \right)}_{\text{regularization}} \quad (17)$$

where

$$\|\mathbf{d}\|_{\ell_1} = \int \sum_k \left(|d_k(f)| \right) df. \quad (18)$$

The equivalence of the solutions of (17) and (15) is a research area that is extensively researched [24], [25]. In this paper, we do not dwell on such issues. The ℓ_1 penalty is a widely used tool in image processing in the context of total-variational denoising, and it has been demonstrated to give robust reconstructions. While being motivated by sparsity, obtaining the sparsest representation is not crucial for our scheme; our main objective is to get a robust solution that the ℓ_1 approach can guarantee.

The only parameter to be adjusted in the proposed algorithm is λ; it provides the tradeoff between sparsity (model order and thus the robustness) and approximation error. We derive the optimal value of λ using the L-curve[3] method [26].

[2]We expect that the subset would contain mostly the piecewise exponential basis vectors and a few local basis vectors at the locations of the perturbations.

[3]This approach performs the reconstructions for different values of λ and makes a plot of the data consistency error versus the regularization term. The plot resembles the shape of an L; the lambda corresponding to the vertex of the L provides a good compromise between the two terms.
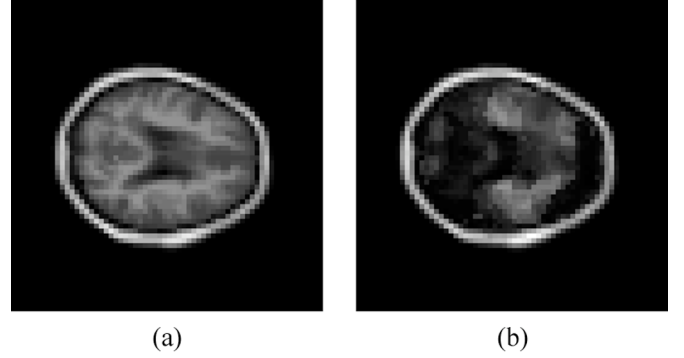

(a)   (b)

Fig. 1. Illustration of effect of magnetic field inhomogeneity. We consider a synthetic $\rho(\mathbf{r}, f)$ that conforms to the SLIM model. [This signal was estimated from experimental data obtained from a human subject (see Fig. 5(d)]. We then generated shifted signal using magnetic field inhomogeneity map estimated from reference water scans obtained from subject. Nonshifted and shifted signals are shown in (a) and (b), respectively. Field inhomogeneity often creates shifts of a few tens of Hertz at 1.5 T. Combined with narrow spectral resonances of metabolites, this shift could lead to substantial spatial variations within compartments. Experiment indicates that even if spectral distributions conform to image model, the observed signal could have drastic signal variations because of magnetic field inhomogeneity. (a) $\rho(\mathbf{r}, f)$. (b) $\rho(\mathbf{r}, f - f_0(\mathbf{r}))$.

B. Magnetic-Field Inhomogeneity

The standard schemes assume $\Delta B_0(\mathbf{r}) = 0$ to obtain simple reconstruction algorithms. We now show that this approach could lead to substantial distortions. Making a change of variable in (1), we obtain

$$\theta_{\text{sim}}(\mathbf{k}; k_f) = \int_{\mathbb{R}^3} s\left(\mathbf{r}, f - f_0(\mathbf{r})\right) e^{-j(\mathbf{k}; k_f)^T(\mathbf{r}; f)} d\mathbf{r} df \quad (19)$$

where $f_0(\mathbf{r}) = \gamma \Delta B_0(\mathbf{r})$ and γ is the gyromagnetic ratio. The above expression implies that the measurements correspond to the Fourier transform of the shifted function $\rho_{\text{shifted}} = \rho(\mathbf{r}, f - f_0(\mathbf{r}))$ rather than $\rho(\mathbf{r}, f)$ itself. Note from Fig. 1 that ρ_{shifted} may vary spatially (in $\mathbf{r}$) quite significantly, even when ρ conforms to the piece-wise constant assumptions; since the spectral resonance lines are quite narrow, even small field inhomogeneities could give rise to substantial spatial variations. If left unaccounted for, this could lead to a model misfit, thus resulting in significant distortions. The other option is to use a more flexible model to accommodate for the inhomogeneity induced spatial variations. However, a higher model order could be less robust to measurement noise, as seen later in the paper.

We propose to estimate $\Delta B_0(\mathbf{r})$ from alternate scans and use it in the reconstruction procedure. In the experiments described in this paper, we estimate the field map from water reference scans. These scans are routinely acquired during MRSI scans to quantify the spectral information [18], [27]. Alternatively, high-resolution field maps estimated using three-point Dixon's scans [28] can be used to obtain a finer and accurate correction. Since we use a more elaborate model than SLIM, the noniterative reconstruction procedure described in [21] and [22] is not a very attractive option in our case.[4]

[4]It would involve the inversion of large matrices repeatedly, while the spatial mismatch is corrected. Moreover, the kernels have to be recomputed within the iteration loop.

C. Spatial Mismatch

The anatomical reference images and the spectroscopy data are acquired using two different sequences. This often results in imperfect spatial correspondence between the two data sets. Subject motion could be another source of mismatch. Even small misalignments between these data sets could lead to significant spectral leakage artifacts, especially from the extra-cranial lipid regions.

The anatomical and the CSI data-sets are drastically different in contrast and resolution. This makes it difficult to align the two using standard registration algorithms (prior to the reconstruction). We propose to include the registration step into the reconstruction algorithm; the new scheme jointly estimates the optimal deformation map and the coefficients from the CSI data set. We model the deformation as an affine transformation

$$\mathbf{r}' = \mathbf{A}\mathbf{r} + \mathbf{b} \qquad (20)$$

where $\mathbf{A}$ and $\mathbf{b}$ are the parameters of the affine transformation.

IV. OPTIMIZATION ALGORITHM

Combining the field-inhomogeneity correction in (19) and the affine transformation given by (20), we obtain θ_{sim} as

$$\theta_{\text{sim}}(\mathbf{k}; k_f) = \int_{\mathbb{R}^3} s\left(\mathbf{A}\mathbf{r} + \mathbf{b}, f - f_0(\mathbf{r})\right) e^{-j(\mathbf{k}; k_f)^T(\mathbf{r}; f)} d\mathbf{r} df. \qquad (21)$$

Note that in (21), we assume the field map to be spatially aligned with the MRSI data. This is because it is obtained by a water reference scan that uses the same pulse sequence that is used to obtain the MRSI data. Furthermore, we assume s to be specified by (16). The reconstruction problem boils down to the minimization of (17). The unknowns are: a) the coefficient vector $\{\mathbf{c}, \mathbf{d}\}$ and b) the affine parameters $\mathbf{A}, \mathbf{b}$. We propose the following two-step alternating optimization algorithm to solve for the unknowns.

Initialize $\mathbf{A} = \mathbf{I}$ and $\mathbf{b} = \mathbf{0}$. Alternate between the following steps until convergence.
  a) Estimate the optimal $\mathbf{c}(f)$ and $\mathbf{d}(f)$ assuming $\mathbf{A}$ and $\mathbf{b}$ from the previous iteration

$$\{\mathbf{c}(f), \mathbf{d}(f)\}_{\text{opt}} = \arg \min_{\mathbf{c}(f), \mathbf{d}(f)} \mathcal{C}_1\left(\mathbf{c}(f), \mathbf{d}(f), \mathbf{A}, \mathbf{b}\right). \qquad (22)$$

  b) Update $\mathbf{A}$ and $\mathbf{b}$ assuming $\{\mathbf{c}(f), \mathbf{d}(f)\}$ from the previous step

$$\{\mathbf{A}, \mathbf{b}\}_{\text{opt}} = \arg \min_{\mathbf{A}, \mathbf{b}} \mathcal{C}_1\left(\mathbf{c}(f), \mathbf{d}(f), \mathbf{A}, \mathbf{b}\right). \qquad (23)$$

A. Estimation of Optimal Coefficient Set: $\{\mathbf{c}, \mathbf{d}\}$

Since the affine parameters are assumed to be constants during this step, we omit them from the equations in this section. Since the coefficients $\mathbf{c}$ and $\mathbf{d}$ are complex variables, it is difficult to use the standard linear programming techniques to solve the ℓ_1 minimization problem as in [24] and [25]. We instead rely on a half quadratic regularization method introduced in [29], after a slight modification of the definition of the ℓ_1 penalty

$$\|\mathbf{d}\|_{\ell_1} \approx \int \left(\sum_k \sqrt{\alpha + |d_k(f)|^2}\right) df \qquad (24)$$

where $\alpha > 0$ is a small constant. Note that this metric is differentiable and converges to (18), when $\alpha \to 0$. In practice, we choose an α that is high enough for the algorithm to be numerically stable. The half-quadratic approach enables us to reformulate the reconstruction problem as a sequence of quadratic optimization problems. We alternate between the minimization of

$$\mathcal{C}_{\text{weighted}} = \|\boldsymbol{\theta}_{\text{sim}} - \boldsymbol{\theta}_{\text{meas}}\|_{\ell_2}^2 + \lambda\left(\|\mathbf{d}\|_{\mathbf{w}_d} + \|\mathbf{c}\|_{\mathbf{w}_c}\right) \qquad (25)$$

where $\|\mathbf{c}\|_{\mathbf{w}_c}$ indicate the weighted norm

$$\|\mathbf{c}\|_{\mathbf{w}_c} = \int \left(\sum w_{c_k(f)} |c_k(f)|^2\right) df$$

and the update of the weights

$$w_{c_k(f)} = \frac{1}{\sqrt{\alpha + |c_k(f)|^2}} \qquad (26)$$

to minimize (22). We start with $\mathbf{w}_c = \mathbf{w}_d = \mathbf{1}$.

Since (25) is a quadratic criterion and the forward model is linear (see from [21]), we use the conjugate gradients optimization algorithm to minimize (25). We derive the expression for the gradient of the criterion with respect to the coefficients in Appendix C. The half-quadratic optimization algorithm is guaranteed to converge to the minimum of (17) [29].

B. Estimation of Affine Parameters $\{\mathbf{A}, \mathbf{b}\}$

We use a steepest descent algorithm to derive the optimal affine parameters. Since this minimization problem is non-linear, the algorithm is not globally convergent. Equation (16) has many parameters, and hence the cost function may have a lot of local minima. To make the criterion well behaved, we run the two-step algorithm with a high value of λ for a few iterations to initialize the affine parameters. Since high values of λ imply sparser models,[5] we expect the cost function to have fewer local minima. This procedure ensures the convergence to the global minimum, analogous to the multiresolution strategy used in nonlinear registration algorithms [30].

Note that the two-step approach is only needed if there is a spatial mismatch between the anatomical and MRSI data sets. If the pulse sequences are matched so that there is perfect correspondence, the second step of the two-step alternating algorithm (solving for the affine parameters) can be omitted.

---

[5]Models with more coefficients will lead to a higher value of criterion; models with fewer parameters are preferred when a high value of λ is chosen.

V. ERROR ANALYSIS OF GENERAL MODEL-BASED SCHEMES

The reconstruction problem in MR spectroscopic imaging is somewhat different from the standard MR imaging setup because of the limited number of spatial Fourier encodes. In standard MRI, the dominant term in the reconstruction error is the variance of the estimate resulting from measurement noise. In contrast, the main part of the reconstruction error in MRSI is the bias (systematic errors due to finite number of measurements, which are also termed as truncation artifacts). In this context, it is highly desirable that the errors resulting from the different terms in a general model-based reconstruction schemes are analyzed.

We restrict our attention to schemes that use linearly independent basis sets for simplicity; we derive a general expression for the error for such schemes. Even though the superset $\boldsymbol{\Theta}$ is linearly dependent, the minimization of (15) will lead to signal reconstruction using a linearly independent subset of $\boldsymbol{\Theta}$. Thus, the error expression derived in this section is applicable to our algorithm by considering this subset as the set of reconstruction basis functions $\boldsymbol{\varphi}$. In the strict sense, the applicability of the analysis to the proposed algorithm depends on the equivalence of (15) and (17), since our algorithm is based on (17) rather than (15).

Since the k_f (readout) dimension is acquired at a sufficiently high sampling rate in CSI, the truncation effects in this dimension can be safely ignored. Hence, we restrict the analysis to the reconstruction of a 2-D spatial function $\rho(x, f)$ (for a specified f) from its Fourier samples. In the current form, the error expression is general enough to be applicable to a wide range of Fourier inversion problems.

Note from (7) that estimating the coefficients from the measurements (in the absence of noise) is equivalent to evaluating an inner product between the original function ρ and the point-spread functions (denoted by $\tilde{\varphi}$). In MRSI, the measurements are constrained to be Fourier samples of ρ. Hence, the point-spread functions (denoted by $\tilde{\varphi}$), in general, are not the same as the corresponding reconstruction basis functions (denoted by $\boldsymbol{\varphi}$). $\boldsymbol{\varphi} = \tilde{\varphi}$ only when we have a band-limited model. Thus, in the general case, the reconstruction procedure is a biorthogonal projection and hence is conceptually similar to the sampling framework considered in [31], [32]. We now quantify the approximation error in reconstructing an arbitrary signal in the model based framework using similar tools.

Proposition 1: Let $\boldsymbol{\varphi}$ denote a P-dimensional vector of linearly independent basis functions. Then, the expected error in reconstructing an arbitrary signal $\rho(\mathbf{r}) \in L_2(\mathbf{R}^2)$ from its noisy Fourier samples (variance σ^2) in $V_{\boldsymbol{\varphi}} = \mathrm{span}(\boldsymbol{\varphi})$ using (6) and (2) is given by

$$
\begin{aligned}
\epsilon &= \mathbf{E}\left(\|\rho^* - \rho\|_{L_2}^2\right) \\
&= \underbrace{\|\rho\|_{L_2}^2 - \|\langle\rho, \boldsymbol{\varphi}_o\rangle\|_{\ell_2}^2}_{\epsilon_{\min}} + \underbrace{\left\|\mathbf{R}^{1/2}\langle\rho, \tilde{\varphi} - \boldsymbol{\varphi}_d\rangle\right\|_{\ell_2}^2}_{\epsilon_{\mathrm{res}}} \\
&\quad + \underbrace{\sigma^2 \mathrm{trace}(\mathbf{Q}^H \mathbf{R} \mathbf{Q})}_{\epsilon_{\mathrm{noise}}}.
\end{aligned}
\tag{27}
$$

The positive-definite correlation matrix[6] $\mathbf{R}$ is given by $\mathbf{R} = \int_{\mathbb{R}^3} \boldsymbol{\varphi}(\mathbf{r})\boldsymbol{\varphi}^H(\mathbf{r})d\mathbf{r}$. The basis set $\boldsymbol{\varphi}_o = \mathbf{R}^{-1/2}\boldsymbol{\varphi}$ spans $V_{\boldsymbol{\varphi}}$ and

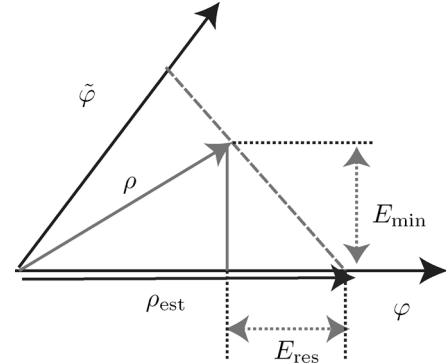[6]If $\boldsymbol{\varphi}$ forms an orthogonal basis set, $\mathbf{R} = \mathbf{I}$.

Fig. 2. Illustration of first two error terms in (27). Consider representation of vector indicated by ρ in terms of basis vector φ. Best possible reconstruction corresponds to orthogonal projection of ρ onto φ. Reconstruction error in this context is indicated by $E_{\min}$ that corresponds to the first term in (27). To achieve this minimum possible error, we have to perform an orthogonal projection of ρ onto φ; this requires evaluation of the inner product between ρ and φ. Now, assume that we could only perform an inner product between ρ and $\tilde{\varphi}$. In this case, the reconstruction is given by ρ_{est}: an oblique projection of ρ, orthogonal to $\tilde{\varphi}$, giving the extra error term $\varepsilon_{\mathrm{res}}$. In this case, we will be making an extra error indicated by E_{res}.

they are orthogonal to each other (i.e., $\int \boldsymbol{\varphi}_o(\mathbf{r})\boldsymbol{\varphi}_o^H(\mathbf{r})d\mathbf{r} = \mathbf{I}$). Similarly, $\boldsymbol{\varphi}_d = \mathbf{R}^{-1}\boldsymbol{\varphi}$ spans $V_{\boldsymbol{\varphi}}$ and they are biorthogonal to $\boldsymbol{\varphi}$ (i.e., $\int \boldsymbol{\varphi}_d(\mathbf{r})\boldsymbol{\varphi}^H(\mathbf{r})d\mathbf{r} = \int \boldsymbol{\varphi}(\mathbf{r})\boldsymbol{\varphi}_d^H(\mathbf{r})d\mathbf{r} = \mathbf{I}$); they are termed the dual basis functions of $\boldsymbol{\varphi}$.

The proof of the proposition is given in Appendix A. We term $\epsilon_{\min}$ the orthogonal projection error. $\|\langle\rho, \boldsymbol{\varphi}_o\rangle\|_{l_2}^2$ in (12) is the l_2 norm of the inner products of ρ with $\boldsymbol{\varphi}_o$. Since $\boldsymbol{\varphi}_o$ is the orthogonal basis set of $V(\boldsymbol{\varphi})$, this term corresponds to the L_2 norm of the orthogonal projection of ρ onto $V(\boldsymbol{\varphi})$. Thus, $\epsilon_{\min}$ depends only on $\boldsymbol{\varphi}$; given the reconstruction basis functions, it is the lower bound of the achievable error.

We term ϵ_{res} the biorthogonal projection residue. Note from (27) that this term would vanish for any ρ if $\tilde{\varphi} = \boldsymbol{\varphi}_d$. However, when the measurements are constrained to be Fourier samples, this is achieved only when $\boldsymbol{\varphi} = \mathbf{R}\mathbf{Q}\mathbf{e}$; i.e., when the reconstruction basis functions can be represented as linear combination of the Fourier exponentials. Thus, only band-limited models give $\epsilon_{\mathrm{res}} = 0$, for any input. However, as the number of Fourier samples tends to infinity, $\mathbf{R}\mathbf{Q}\mathbf{e} \to \boldsymbol{\varphi}_d$ for any model[7]; $\epsilon_{\mathrm{res}} \to 0$ as the number of measurements increase. Thus, ϵ_{res} can be visualized as the truncation error (error due to the number of measurements being limited).

A signal $\rho \in V_{\boldsymbol{\varphi}}$ can be expressed as $\rho = \boldsymbol{\varphi}^H \mathbf{c}$, where $\mathbf{c}$ is an arbitrary set of coefficients. Now, $\langle\boldsymbol{\varphi}_d, \rho\rangle = \langle\mathbf{R}^{-1}\boldsymbol{\varphi}, \boldsymbol{\varphi}^H\rangle\mathbf{c} = \mathbf{c}$. Similarly, we have $\langle\tilde{\varphi}, \rho\rangle = \langle\mathbf{Q}\mathbf{e}, \boldsymbol{\varphi}^H\rangle\mathbf{c} = \mathbf{Q}\mathbf{A}\mathbf{c} = \mathbf{c}$. Hence, $\langle\rho, \boldsymbol{\varphi} - \tilde{\varphi}\rangle = \mathbf{0}$ and consequently $\epsilon_{\mathrm{res}} = 0$ when $\rho \in V_{\boldsymbol{\varphi}}$. This can also be seen geometrically from Fig. 2. Hence, $\epsilon_{\mathrm{res}} = 0$, if the signal ρ is in the span of the reconstruction functions; if there is no model misfit, ϵ_{res} will also vanish, even when $\boldsymbol{\varphi}_d \neq \tilde{\varphi}$.

The first two terms in (27) are the systematic error components; they contribute to bias in the reconstructions. The last

[7]Any $\boldsymbol{\varphi}$ can be represented by a linear combination of exponentials, as the number of exponentials tends to infinity.

term is a measure of the reconstruction error resulting from measurement noise. The noise would be amplified by the reconstruction matrix $\mathbf{Q}$. Note that $\epsilon_{\min}$ is independent of $\mathbf{Q}$, while ϵ_{res} vanish if the model represents the signal well (i.e., $\rho \in V_{\varphi}$). Thus, for well-designed models, the main error term that is dependent on $\mathbf{Q}$ is $\epsilon_{\mathrm{noise}}$. Thus, we derive the optimal reconstruction matrix $\mathbf{Q}$ as the one that minimizes $\epsilon_{\mathrm{noise}}$, subject to the data-consistency criterion $\mathbf{Q}\mathbf{\Psi} = \mathbf{I}$

$$\mathbf{Q} = \left(\mathbf{\Psi}^H \mathbf{R}^{-1}\mathbf{\Psi}\right)^{-1}\mathbf{\Psi}^H \mathbf{R}^{-1}. \tag{28}$$

Here, the $M \times P$ measurement matrix (M is the number of measurements and P is the number of basis functions) $\mathbf{\Psi}$ is defined as

$$[\mathbf{\Psi}]_{i,j} = \hat{\varphi}_j(\mathbf{k}_i); \qquad i=0\ldots,M-1; \quad j=0\ldots P-1. \tag{29}$$

Note that when the basis functions are orthogonal ($\mathbf{R} = \mathbf{I}$), we have $\mathbf{Q} = \left(\mathbf{\Psi}^H\mathbf{\Psi}\right)^{-1}\mathbf{\Psi}^H$ that is the widely used solution. The use of this solution for nonorthogonal φ could lead to increased variance. With (28) as the reconstruction matrix, we get

$$\epsilon_{\mathrm{noise}} = \sigma^2 \mathrm{trace}\left(\left[\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi}\right]^{-1}\right). \tag{30}$$

A. Inferences

Since the signal model in the Fourier scheme is matched to the measurements, it gives reconstructions with $\epsilon_{\mathrm{res}} = 0$ and no noise amplification (since $\mathbf{Q} = \mathbf{I}_{P \times P}$, $\epsilon_{\mathrm{noise}} = \sigma^2 P$, where P is the number of exponentials). These properties, coupled with its computational efficiency, make it an attractive option for high resolution scans. However, the convergence of the Fourier model to piece-wise smooth signals is very slow. This leads to a high value of $\epsilon_{\min}$, when the number of Fourier samples (and consequently the number of exponentials in the model) is small.

However, if designed properly to reduce misfit, the model-based techniques can provide low ϵ_{res} and a much lower $\epsilon_{\min}$ than Fourier schemes. Moreover, these schemes can also significantly reduce the noise contribution. We show that for large number of spatial encodes, we have

$$\epsilon_{\mathrm{noise}} \approx \sigma^2 P \tag{31}$$

where P is the number of reconstruction basis functions (see Appendix B for the proof). When the number of phase encodes is limited, the growth of $\epsilon_{\mathrm{noise}}$ as a function of P is faster than linear since $\mathbf{\Psi}$ becomes ill conditioned. Since model-based schemes give good signal representation using fewer basis functions than the Fourier model, it can significantly reduce $\epsilon_{\mathrm{noise}}$.

In summary, model-misfit could result in biased reconstructions, while noise variance scales with the number of basis functions (free parameters). Thus, a scheme that provides a good representation using fewer basis functions would provide a low value of combined error (bias/systematic error + variance). Conventional model-based schemes such as SLIM use few basis functions to obtain a low noise variance (well-posed reconstructions); in practical applications, this results in misfit

and hence the bias is the dominant error component. The number of basis functions in Fourier schemes scale linearly with the number of encodes. Hence, this approach may lead to large variance, thus degrading the overall performance as the number of encodes increase (in noisy situations). We can also see it from the simulation studies in the next section. The proposed scheme of picking the best (linearly independent) subset from an over-complete dictionary is more flexible than the conventional model-based algorithms. Being an adaptive scheme, it provides a better representation to the signal at hand using fewer basis functions, thus lowering the overall error.

VI. RESULTS

In this section, we present validations of the proposed technique with both simulated and experimental data.

A. Simulations

We used the segmentation of an anatomical data set (considered in Section VI-B) to create a numerical phantom. Using the parameters from [33] and the software package GAMMA [34], we created the line shapes of NAA, Creatine, Choline, and myo-Inositol; these line shapes were broadened to simulate T_2 decay. We also modeled the lipid spectrum with a broad peak around 3.5 ppm. The piece-wise exponential model (10) with 5×5 exponentials per compartment was used to generate the spatial spectral function $\rho(\mathbf{r}, f)$. The coefficients $c_{i,l}(f)$ were obtained as a linear combination of the lines shapes as

$$c_{i,l}(f) = \sum_{s=0}^{S-1} v_{i,l,s}\beta_k(f) \tag{32}$$

where $\beta_k; k = 0, \ldots, S-1$ are the spectral line shapes of the metabolites, $v_{i,p,s}$ are random coefficients, and S is the total number of metabolites in each compartment. Since in practice, we do not get any signal from the cerebral spinal fluid region [white region in the center in Fig. 3(a)], we set $S = 0$ for this compartment. Similarly, the extra-cranial region was assumed to have only lipids. The slice at 2 ppm (corresponding to NAA peak) is shown in Fig. 3(c). Using the inhomogeneity map from the scan considered in Section VI-B, we developed a realistic forward model. White Gaussian distributed noise was added to the simulated measurements. To test the robustness of the algorithms to segmentation errors, we considered the perturbed segmentation map shown in Fig. 3(b); we changed the compartmental structure at the region indicated by the arrow. Although this is a more drastic situation than expected in practical cases, we use this example to illustrate the ability of the algorithm to compensate for segmentation errors.

Using the above-mentioned setup, we compare the new approach, SLIM with inhomogeneity compensation, and the Fourier scheme. We reconstruct the signal from 16×16-phase encoded CSI data using these schemes and the results are shown in Fig. 3(d)–(f). Note that the new scheme recovers the smooth signal variations in the compartments and the signal variations in the regions corresponding to wrong segmentations reasonably well. It captures the signal in the wrongly segmented regions using a few local basis functions, thus preventing it from aliasing back to the other regions. However, the SLIM
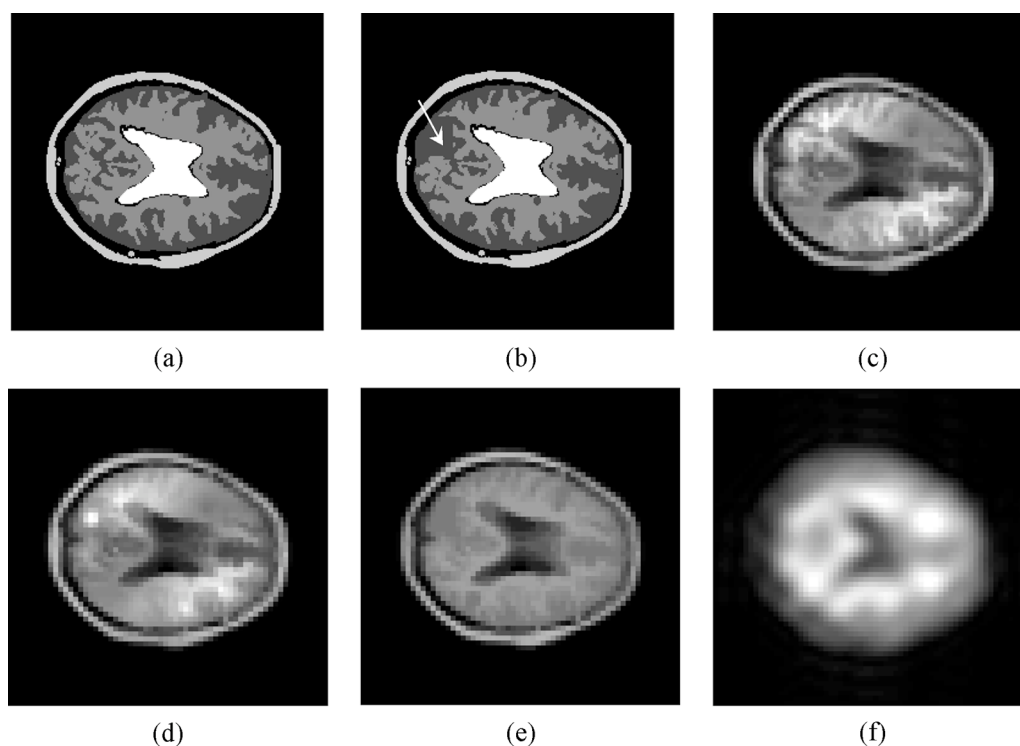
Fig. 3. Comparison of algorithms on simulated data. We generated numerical spectroscopy phantom, whose peak at 2 ppm is shown in (c), using template obtained by segmentation of anatomical scans, discussed in previous section. Slice of template is shown in (a). To simulate effect of segmentation errors, we use template in (b) (by changing the region indicated by the arrows) to perform the reconstruction. New, SLIM, and Fourier reconstructions from 16×16 encoded data with SNR of 10 dB are shown in (d), (e), and (f), respectively. Note that the new reconstructions gave good reconstructions, successfully capturing the signal variations within the compartments as well as the regions with the segmentation errors. This indicates that model can accommodate for a wide range of region inhomogeneities and segmentation errors. (a) Original template. (b) Reference template. (c) Ground truth. (d) New 16×16; $\mathrm{S/E} = 22.47$ dB. (e) SLIM 16×16; $\mathrm{S/E} = 6.62$ dB. (f) Fourier 16×16; $\mathrm{S/E} = 10.07$ dB.
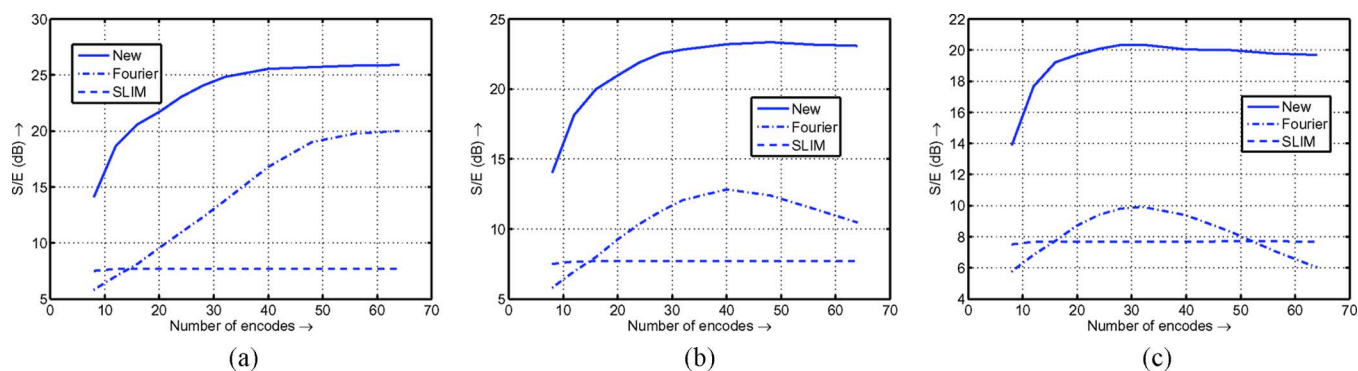

Fig. 4. Comparison of reconstruction schemes (New approach, Fourier reconstruction, and inhomogeneity compensated SLIM) for different number of phase encodes and different noise levels. Note that signal to error ratio of the SLIM is fairly constant indicating that it is very robust to measurement noise. However, the main source of error in this case is approximation error/bias resulting in poor results. New method gave the best overall performance, with its signal to error ratio saturating after 25×25-phase encodes. For higher noise levels, performance of Fourier scheme decreased as number of phase encodes increase. See text for details. (a) Input SNR = 20 dB. (b) Input SNR = 10.4 dB. (c) Input SNR = 6.02 dB.

reconstructions suffer from artifacts, mainly because of the violated spatial homogeneity assumptions and segmentation errors. Similarly, the Fourier reconstructions are extremely blurred and thus do not provide much useful information. Also note the lower SNR figures with SLIM and Fourier reconstructions.

In Fig. 4, we studied the effect of varying the number of phase encodes and the noise level on the algorithm. It is seen that the performance of the SLIM scheme is unaffected by the noise level, mainly resulting from fewer model parameters, thus giving a low value of $\epsilon_{\mathrm{noise}}$. Unfortunately, the dominant error

term is the bias or systematic error components ($\epsilon_{\min}$ and ϵ_{res}); since there is a significant misfit of the model to the original signal, it gives poor reconstructions. The performance of the Fourier scheme improves with the number of encodes at low noise level, while it declines at high noise levels. The model order (number of exponentials) of the Fourier reconstruction scheme is equal to the number of phase encodes. Since the energy of typical images are concentrated at the low end of the Fourier spectrum, the SNR of the high-frequency samples is much lower than that of the low-frequency samples. Thus, as
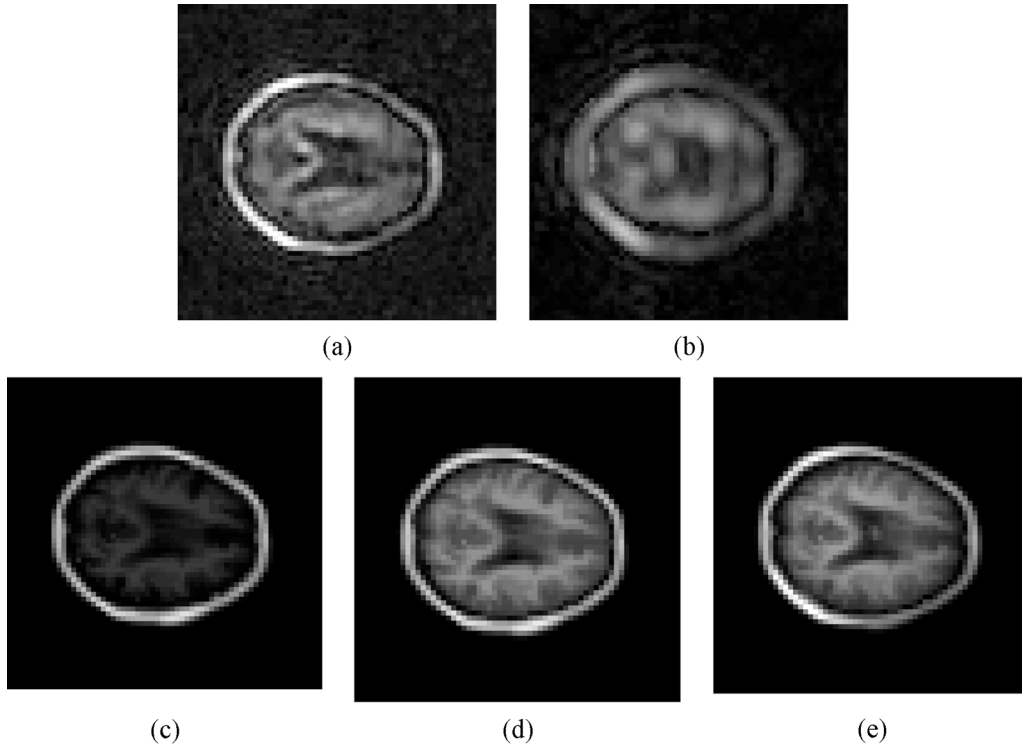
Fig. 5. Comparison of algorithms on experimental data. Fourier reconstruction of image corresponding to NAA peak (2 ppm) from 34×34 encoded CSI data is shown in (a). Considering this data set as a reference, we now compare performance of reconstruction schemes from 16×16 encodes. Reconstructions using Fourier scheme, SLIM without and with inhomogeneity compensation, and new approach are shown in (b)–(e), respectively. (a) Fourier 34×34. (b) Fourier 16×16; $S/E = 3.38$ dB. (c) SLIM without inhomogeneity compensation 16×16; $S/E = 2.93$ dB. (d) SLIM with inhomogeneity compensation 16×16; $S/E = 5.43$ dB. (e) New 16×16; $S/E = 9.54$ dB.

we increase the number of Fourier samples beyond a point, the model order increases (thus considerably increasing $\epsilon_{\mathrm{noise}}$). On the other hand, the newly added samples provide limited information (do not reduce $\epsilon_{\min}$ significantly). Thus, in a noisy scenario, the performance of Fourier reconstructions will degrade as we increase the number of acquisitions beyond a particular number, as seen in Fig. 4.

The performance of the new approach ameliorates with the number of encodes, saturating around 25–30 encodes at high noise levels. The model order of the new algorithm varies with the number of Fourier samples, but it is not linearly proportional to the number of encodes. The order of the model is determined by the choice of λ that is derived using the L-curve method [26]. Thus, the proposed reconstruction scheme gives a good compromise between approximation error and robustness. It adaptively chooses the model to obtain a good compromise. Note that the performance of the new approach saturates after a while; this indicates that increasing the number of measurements (so the scan-time) beyond the saturation point is not beneficial.

B. Experimental Results

Experimental MRI and MRSI data were obtained from an Alzheimer's disease patient, who was scanned using a 1.5-T Siemens Vision system (Siemens Medical System, Germany). The study was approved by the Committee of Human Research at UCSF and written informed consent was obtained from the volunteer before the study began. We used the pulse sequence introduced in [18] to extract the 34×34 encoded MRSI data; it uses a spin echo 34×34 CSI acquisition with parameters TE-25, $\mathrm{TR} = 1.8$ s, bandwidth $= 1538.5$, We then used a subset (16×16 encodes) of these spatial phase encodes to verify the algorithm. The MRI intensities were segmented into tissue classes of gray matter, white matter, CSF, and non-brain tissue classes using a segmentation program developed in house [35]. The magnetic field inhomogeneity map is generated by a reference water scan [18].

The Fourier reconstruction scheme involved zero-padding the phase encoded data to 64×64 k-space samples, 3-D Fourier reconstruction, and post correction of inhomogeneity artifacts using the estimated inhomogeneity map. We considered the SLIM approach with and without the inhomogeneity compensation. Finally, we used the new method with $\mathrm{L} = 81$ (81 exponentials per compartment). Note that this number decides the size of the over-complete basis set Φ and not the final model order. The order selection is part of the reconstruction, thanks to the minimization of (17). As discussed before, we choose the optimal λ using the L-curve approach. For illustration purposes, we display a slice of the reconstructions corresponding to a chemical shift of 2 ppm (location of the NAA peak[8]) in the $x - f$ space in Fig. 5.

---

[8]Note that this image is not the NAA image after metabolite quantification. The lipid peak being very broad can still be seen in this image since the amplitude of the lipid signal is about ten times stronger than the signal from NAA.
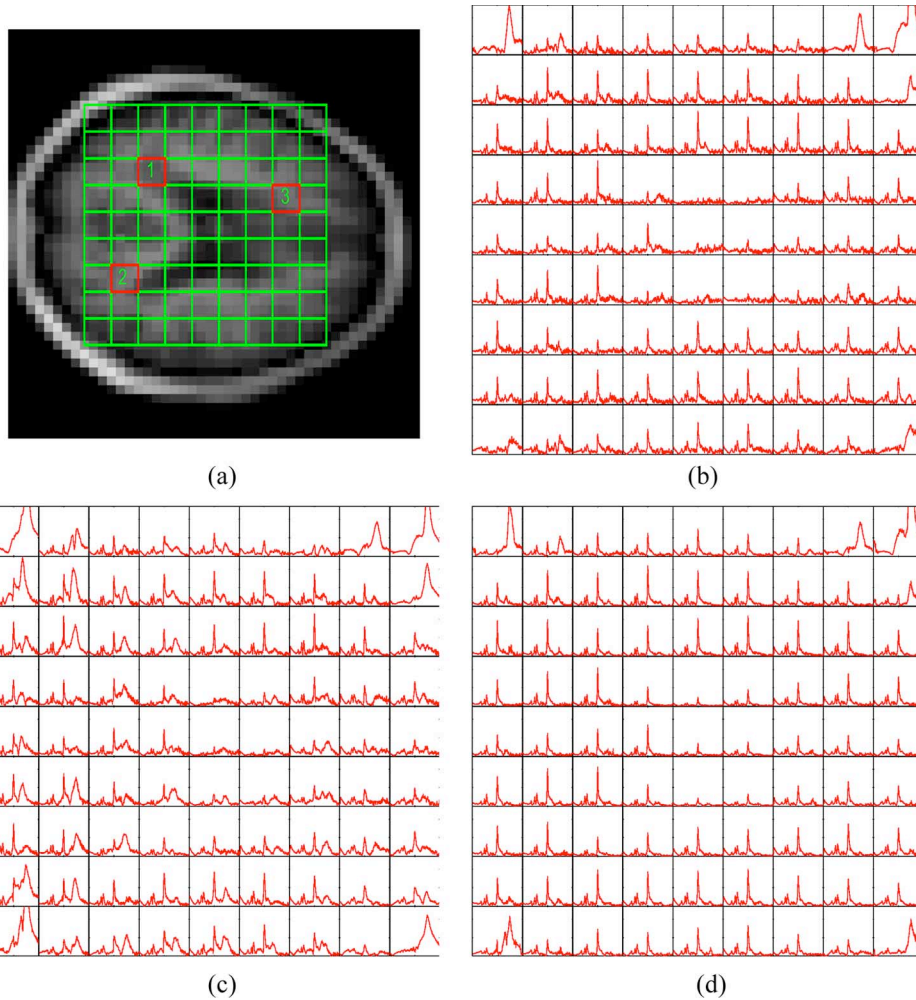
Fig. 6. Averaged spectra from voxels indicated by squares in (a). Note that FOV was 280×280 and slice included lipid regions as well. We just chose this subset for illustration purposes. (b) Corresponds to Fourier reconstructions from 34×34, while (c) is Fourier reconstruction from 16×16 encodes. (d) Spectra obtained from 16×16 encodes using proposed approach. Note that Fourier reconstructions both from 34×34 and 16×16 encodes suffer from a lot of leakage artifacts at lipid resonant frequency. However, new scheme gives a reasonable suppression of lipid peaks (normal subjects will not have lipid signals from the voxels inside the brain). Also, note that the line shapes of the peaks are lower for new reconstructions. Line shape variations seen in new reconstructions result from uncorrected intravoxel field variations; we are assuming a low-resolution field map (derived using water reference scans).

Since the ground truth is not available, we compare the Fourier coefficients of the reconstructed signal from 16×16 encodes with the available 34×34 measurements. The metric we used for quantitative comparison is

$$S/E = \frac{\sum_{n \in \mathbf{N}} \left(\sum_{\mathbf{k} \in \mathbf{K}} |\theta_{\text{recon}}(\mathbf{k}, k_{f_n})|^2 \right)}{\sum_{n \in \mathbf{N}} \left(\sum_{\mathbf{k} \in \mathbf{K}} |\theta_{\text{recon}}(\mathbf{k}, k_{f_n}) - \theta_{\text{meas}}(\mathbf{k}, k_{f_n})|^2 \right)} \quad (33)$$

where $\mathbf{K}; \mathbf{N}$ indicates the set of available measurements (34×34 in this case). θ_{recon} is obtained by injecting the reconstruction into the measurement (21).

It is seen that the Fourier reconstruction from 34×34 encodes, shown in Fig. 5(a), still suffers from significant leakage and ringing artifacts (note the oscillatory bands around the lipid region and the contamination of the spectral shapes in Fig. 6). The reconstructions using the SLIM model without and with inhomogeneity compensation from 16×16 encodes is shown in Fig. 5(c) and (d) respectively. The improvements in Fig. 5(d)

over Fig. 5(c) clearly illustrate the improvement in using inhomogeneity compensation. The reconstruction using the proposed method from the same number of phase encodes is shown in Fig. 5(e). Note that the new approach gives significantly better results, thus illustrating the utility of this scheme in practical applications.

The average spectra from the squares (corresponding to 3×3 pixels) are displayed in Fig. 6. Note that the acquisition was performed on the entire slice, including the lipid regions. A zoom of three of these spectral lines is shown in Fig. 7. Note that the spectra reconstructed using the new algorithm displays much lower leakage artifacts and better line shapes as compared to the Fourier reconstructions (both from 34×34 and 16×16) encodes. It is seen that the Fourier reconstructions exhibit significant spectral leakage, even when the number of Fourier encodings is large (34×34).

The magnetic field is inhomogeneous in the region around box 3 in Fig. 6(a), because of its proximity to the sinus regions. This explains the line shape variations and reduced amplitude of the metabolite peaks in these regions in both the Fourier
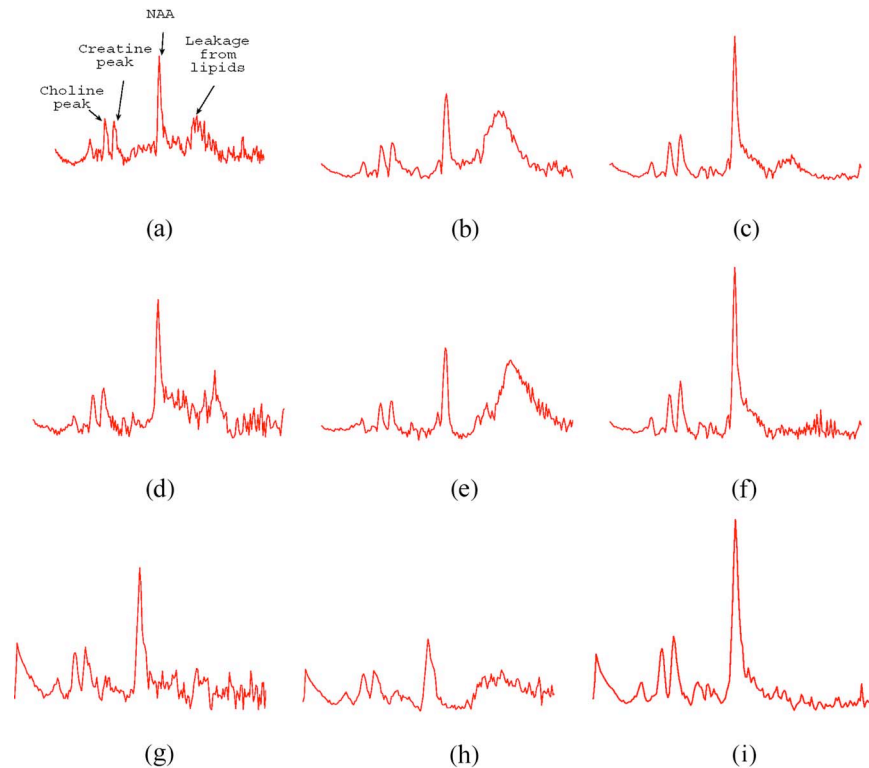
Fig. 7. Zoomed versions of averaged spectra in regions indicated by gray boxes in Fig. 6(a). We have labeled prominent peaks in normal proton spectra in (a). First and the second columns consists of Fourier reconstructions from 34×34 and 16×16 encodes, respectively. Last column indicates spectra obtained from 16×16 encodes using proposed approach. Rows correspond to spectra in boxes 1, 2, and 3 in Fig. 6(a), respectively. Note that new algorithm provides better line shapes and reduced noise. Region around box 3 in Fig. 6(a) corresponds to very inhomogeneous magnetic field. This is seen from line shapes and decreased amplitudes (due to intra-voxel dephasing) of both Fourier reconstructed spectra in the last row. Note that new approach corrects most of these artifacts. (a) Fourier 34×34. (b) Fourier 16×16. (c) New 16×16. (d) Fourier 34×34. (e) Fourier 16×16. (f) New 16×16. (g) Fourier 34×34. (h) Fourier 16×16. (i) New 16×16.

reconstructions. This is due to intravoxel field map variations that are not accounted in the reconstruction. However, the proposed technique corrects most of these variations and restored the amplitudes.

VII. DISCUSSION

A general concern with the model-based schemes that use anatomical constraints is that they would bias the reconstructions, thus making the results insensitive to metabolic variations that are not reflected in the anatomy. The proposed scheme uses a richer basis set; unlike standard model-based schemes, it contains local basis functions that are independent of the anatomy, in addition to the compartmental basis functions. The compartment basis functions capture a large part of the signal that corresponds to lipids, residual water, and components that are smoothly varying in space. This enables the local basis functions to capture the rather subtle metabolic variations that may be indicative of the pathology. We have demonstrated the potential of this scheme to account for such variations in simulation studies as well as phantom studies. However, the application of this scheme to real problems is the ultimate test.

The only free parameter in the reconstruction procedure is λ. Currently, this parameter is determined using the L-curve method. In practice, the optimal parameter value varies little across data sets that are acquired using same parameter settings. This parameter can be chosen from a training set and can be used in later studies.

Although the proposed algorithm is designed for CSI data, it can also be adapted to fast-scan schemes that use readout gradients [8]–[12]. Since the model order of this scheme does not increase linearly with the number of encodes, as with the Fourier reconstructions, it can efficiently compensate for the lower SNR associated with these schemes. In the current implementation, we only considered a 2-D problem. This scheme can be generalized to the 3-D case, where the scan time reduction in comparison to conventional CSI would be more appreciable.

The effect of the artifacts resulting from the magnetic field inhomogeneity will become more prominent with increasing magnetic field strength. Hence, accounting for these imperfections in the forward model is especially important for MRSI studies at 3 T and higher.

In this paper, we did not concentrate on the choice of the k-space samples. An optimization of the sampling pattern as in [15] and [16], or random sampling as in [24] and [25] is worth investigating. Good sampling patterns could lead to a better conditioned system of equations and consequently better reconstructions.

The current MATLAB implementation takes roughly 10 min to reconstruct a $64 \times 64 \times 256$ spatial spectral volume on an aging 3.06-GHz Intel Pentium processor with 1.5-GB RAM. A more optimized implementation could lower the execution time. Since the use of a high-resolution field map would further increase the reconstruction time and memory demand, a more efficient implementation is desirable to accommodate them.

VIII. CONCLUSION

We have identified some of the technical problems that limited the performance of standard model-based MRSI schemes and proposed efficient solutions to rectify them. Specifically, we introduced a new image model that is more efficient to reduce the model misfit in comparison to the standard schemes. We also accounted for the nonideal assumptions that negatively affected the performance such as magnetic field inhomogeneity and spatial mismatch. Experimental and simulation results clearly demonstrated the performance improvement of the the new scheme. The new approach gave fewer artifacts and better quality spectral information, even from fewer numbers of phase encodes. The results suggest that a properly designed model can efficiently exploit the available spatial priors, thus significantly improving the performance in *in vivo* MRSI applications.

APPENDIX A

We consider the error in reconstructing $\rho(\mathbf{r}) \in L_2(\mathbb{R}^3)$ from the available measurements

$$\boldsymbol{\theta}_{\mathrm{meas}} = \langle \rho, \mathbf{e} \rangle + \boldsymbol{\eta} \tag{34}$$

where $\boldsymbol{\eta}$ is the noise vector of variance σ^2. The optimal coefficients are derived from these measurements as

$$\mathbf{c}^* = \mathbf{Q}\boldsymbol{\theta}_{\mathrm{meas}} = \underbrace{\langle \rho, \underbrace{\mathbf{Qe}}_{\tilde{\boldsymbol{\varphi}}} \rangle + \underbrace{\mathbf{Q}\boldsymbol{\eta}}_{\mathbf{c}_2}}_{\mathbf{c}_1}. \tag{35}$$

In the above expression, the noise component of the measurement is denoted by $\mathbf{c}_2$. Thus, the expectation of the measurement error is given by

$$
\epsilon = \mathbf{E}\left[\int_{\mathbb{R}^3} |\rho(\mathbf{r}) - \mathbf{c}^H \boldsymbol{\varphi}(\mathbf{r})|^2 \, d\mathbf{r} \right]
$$
$$
= \underbrace{\int_{\mathbb{R}^3} |\rho(\mathbf{r}) - \mathbf{c}_1^H \boldsymbol{\varphi}(\mathbf{r})|^2 \, d\mathbf{r}}_{\epsilon_1} + \underbrace{\mathbf{E}\left[\int \boldsymbol{\varphi}(\mathbf{r})^H \mathbf{c}_2 \mathbf{c}_2^H \boldsymbol{\varphi}(\mathbf{r}) d\mathbf{r} \right]}_{\epsilon_2}.
$$
$$\tag{36}$$

We moved the expectation to the second term since the first term does not have any random components. We simplify the second term as

$$
\epsilon_2 = \int \boldsymbol{\varphi}(\mathbf{r})^H \mathbf{Q} \underbrace{E[\boldsymbol{\eta}\boldsymbol{\eta}^H]}_{\sigma^2 \mathbf{I}} \mathbf{Q}^H \boldsymbol{\varphi}(\mathbf{r})^H d\mathbf{r}
$$
$$
= \sigma^2 \int \mathrm{trace}\left(\mathbf{Q}^H \boldsymbol{\varphi}(\mathbf{r}) \boldsymbol{\varphi}(\mathbf{r})^H \mathbf{Q} \right) d\mathbf{r}
$$
$$
= \sigma^2 \mathrm{trace}\left(\mathbf{Q}^H \underbrace{\int \boldsymbol{\varphi}(\mathbf{r}) \boldsymbol{\varphi}^H(\mathbf{r}) d\mathbf{r}}_{\mathbf{R}} \mathbf{Q} \right). \tag{37}
$$

We now simplify the first term of (36) as

$$
\epsilon_1 = \int_{\mathbb{R}^3} |\rho(\mathbf{r})|^2 \, d\mathbf{r} + \mathbf{c}_1^H \underbrace{\int_{\mathbb{R}^3} \boldsymbol{\varphi}(\mathbf{r}) \boldsymbol{\varphi}^H(\mathbf{r}) d\mathbf{r}}_{\mathbf{R}} \mathbf{c}_1
$$
$$
- 2\mathbf{c}_1^H \underbrace{\int_{\mathbb{R}^3} \boldsymbol{\varphi}(\mathbf{r}) \rho^H(\mathbf{r}) d\mathbf{r}}_{\mathbf{d}}. \tag{38}
$$

Since $\boldsymbol{\varphi}$ forms a linearly independent basis, the matrix $\mathbf{R}$ is positive definite and hence can be decomposed as $\mathbf{R} = \mathbf{R}^{1/2}\mathbf{R}^{1/2}$. We denote $\mathbf{c}' = \mathbf{R}^{1/2}\mathbf{c}_1$ and $\mathbf{d}' = \mathbf{R}^{-1/2}\mathbf{d} = \langle \rho, \underbrace{\mathbf{R}^{-1/2}\boldsymbol{\varphi}}_{\boldsymbol{\varphi}_o} \rangle_{L_2}$

$$
\epsilon_1 = \|\rho(\mathbf{r})\|_{L_2}^2 + \|\mathbf{c}'\|_{l_2}^2 - 2\mathbf{c}'^H \mathbf{d}'
$$
$$
= \|\rho(\mathbf{r})\|_{L_2}^2 - \|\mathbf{d}'\|_{l_2}^2 + \|\mathbf{c}' - \mathbf{d}'\|_{l_2}^2. \tag{39}
$$

The basis set $\boldsymbol{\varphi}_o = \mathbf{R}^{-1/2}\boldsymbol{\varphi}$ spans the same space as $\boldsymbol{\varphi}$ and they are orthogonal to each other (i.e., $\int \boldsymbol{\varphi}_o(\mathbf{r}) \boldsymbol{\varphi}_o^H(\mathbf{r}) d\mathbf{r} = \mathbf{I}$). Similarly, $\boldsymbol{\varphi}_d = \mathbf{R}^{-1}\boldsymbol{\varphi}$ spans the same space as $\boldsymbol{\varphi}$ and they are bi-orthogonal to $\boldsymbol{\varphi}$ (i.e., $\int \boldsymbol{\varphi}_o(\mathbf{r}) \boldsymbol{\varphi}^H(\mathbf{r}) d\mathbf{r} = \int \boldsymbol{\varphi}(\mathbf{r}) \boldsymbol{\varphi}_o^H(\mathbf{r}) d\mathbf{r} = \mathbf{I}$); we call them the dual basis functions. Hence, the second term in (39) can be simplified as $\|\langle \rho, \boldsymbol{\varphi}_o \rangle\|_{L_2}^2$; since φ_o indicates the orthogonal basis of V_φ, this term corresponds to the L_2 norm of the orthogonal projection of ρ onto the span of $\boldsymbol{\varphi}$. Thus

$$
\epsilon_1 = \|\rho(\mathbf{r})\|_{L_2}^2 - \|\langle \rho, \boldsymbol{\varphi}_o \rangle\|_{l_2}^2 + \left\| \langle \rho, \mathbf{R}^{1/2}\tilde{\boldsymbol{\varphi}} - \mathbf{R}^{-1/2}\boldsymbol{\varphi} \rangle \right\|_{l_2}^2
$$
$$
= \|\rho(\mathbf{r})\|_{L_2}^2 - \|\langle \rho, \boldsymbol{\varphi}_o \rangle\|_{l_2}^2 + \left\| \mathbf{R}^{1/2} \langle \rho, \tilde{\boldsymbol{\varphi}} - \underbrace{\mathbf{R}^{-1}\boldsymbol{\varphi}}_{\boldsymbol{\varphi}_d} \rangle \right\|_{l_2}^2
$$
$$
= \underbrace{\|\rho(\mathbf{r})\|_{L_2}^2 - \|\langle \rho, \boldsymbol{\varphi}_o \rangle\|_{l_2}^2}_{\epsilon_{\min}} + \underbrace{\left\| \mathbf{R}^{1/2} \langle \rho, \tilde{\boldsymbol{\varphi}} - \boldsymbol{\varphi}_d \rangle \right\|_{\ell_2}^2}_{\epsilon_{\mathrm{res}}}. \tag{40}
$$

Combining (40) and (37), we obtain (27). In the above expression, we denote the weighted norm $\|\mathbf{Rc}\|^2 = \|\mathbf{c}\|_{\mathbf{R}}^2$.

APPENDIX B

Here, we show that the variance of the reconstructions increases, as indicated by (31), as the number of basis functions in the model increase. From (27) and (28), we have

$$
\epsilon_{\mathrm{noise}} = \sigma^2 \mathrm{trace}\left([\boldsymbol{\Psi}^H \mathbf{R}^{-1}\boldsymbol{\Psi}]^{-1} \right) \tag{41}
$$
$$
= \sigma^2 \mathrm{trace}\left(\left[\underbrace{(\mathbf{R}^{-\frac{1}{2}}\boldsymbol{\Psi})^H}_{\boldsymbol{\Psi}_o^H} \underbrace{(\mathbf{R}^{-\frac{1}{2}}\boldsymbol{\Psi})}_{\boldsymbol{\Psi}_o} \right]^{-1} \right). \tag{42}
$$

The jth row of $\boldsymbol{\Psi}_o$ is $\hat{\boldsymbol{\varphi}}_o(\mathbf{k}_j)$, where $\boldsymbol{\varphi}_o$ is the vector of orthonormalized basis functions. Thus

$$
\left[\boldsymbol{\Psi}_o^H \boldsymbol{\Psi}_o \right]_{i,j} = \sum_{m=0}^{M-1} \hat{\varphi}_{o,i}^*(\mathbf{k}_m) \hat{\varphi}_{o,j}^*(\mathbf{k}_m). \tag{43}
$$

Here, M is the number of phase encodes. As $M \rightarrow \infty$, the above sum[9] becomes

$$\left[\mathbf{\Psi}_o^H \mathbf{\Psi}_o\right]_{i,j} = \langle \varphi_{o,i}, \varphi_{o,i} \rangle = \delta_{i,j} \qquad (44)$$

where $\delta_{i,j}$ is the Kroneker delta function.

Thus

$$\epsilon_{\text{noise}} = \sigma^2 \text{trace}(\mathbf{I}_{P \times P}) \qquad (45)$$
$$= \sigma^2 P \qquad (46)$$

where $\mathbf{I}$ denotes the identity matrix.

APPENDIX C

As discussed previously, we use a two-step optimization algorithm to update $\mathbf{c}(f)$ and the affine parameters $\mathbf{A}$ and $\mathbf{b}$. For the ease of notation, we define the operator S_{f_0} that denotes a shift along the frequency axis as $S_{f_0}(\rho) = \rho(\mathbf{r}, f - f_0(\mathbf{r}))$. The adjoint of this operator is $S_{f_0}^*(\rho) = \rho(\mathbf{r}, f + f_0(\mathbf{r}))$. Again, we denote the operator that maps ρ to the Fourier transform evaluated at the sampling locations as $\mathcal{F}(\rho) = \boldsymbol{\rho}_{\text{sim}}$; $\mathcal{F} : \mathbb{R}^3 \rightarrow [0, \ldots, M-1] \times [0, \ldots, N-1]$. The adjoint of this operator is denoted as $\mathcal{F}^* : [0, \ldots, M-1] \times [0, \ldots, N-1] \rightarrow \mathbb{R}^3$. Both $\mathcal{F}$ and $\mathcal{F}^*$ are implemented using fast Fourier transform.

In the first step, we estimate the optimal $\mathbf{c}(f)$, assuming $\mathbf{A}$ and $\mathbf{b}$ from the previous iteration. We start by initializing $\mathbf{f}_0'(\mathbf{r}) = \mathbf{f}_0(\mathbf{A}\mathbf{r} + \mathbf{b})$ and $\boldsymbol{\varphi}'(\mathbf{r}) = \boldsymbol{\varphi}(\mathbf{A}\mathbf{r} + \mathbf{b})/|\mathbf{A}|$, where $|\mathbf{A}|$ is the determinant of $\mathbf{A}$. As mentioned previously, we estimate the coefficients $\mathbf{c}$ using conjugate gradients optimization.

We now compute the gradient of $\mathcal{C}$ that is needed for the conjugate gradients update. Substituting (2) in (1) on the image model $\mathbf{c}(f)^H \varphi(\mathbf{r})$ using the above notations, we get

$$\boldsymbol{\theta}_{\text{sim}}(\mathbf{k}; k_f) = \mathcal{F}S_{f_0'}\left(\mathbf{c}(f)^H \varphi'(\mathbf{r})\right). \qquad (47)$$

We now consider a perturbation in $\mathbf{c}(f) \rightarrow \mathbf{c}(f) + \delta\mathbf{c}(f)$ and derive the corresponding perturbation in the criterion

$$\mathcal{C}(\boldsymbol{c}(f) + \delta\mathbf{c}(f)) - \mathcal{C}(\boldsymbol{\delta}(f))$$

$$\approx 2\mathbb{R}e \int \sum_{m=0}^{M-1}\left[\underbrace{(\theta_{\text{sim}}(\mathbf{k}_m; k_f) - \theta_{\text{meas}}(\mathbf{k}_m; k_f))}_{E(\mathbf{k}_m; k_f)}\right.$$
$$\left. \times \left(\mathcal{F}S_{f_0'}\left(\delta\mathbf{c}(f)^H \varphi'(\mathbf{r})\right)\right)^*\right]dk_f$$

$$= 2\mathbb{R}e\left\langle S_{f_0'}^* \mathcal{F}^* E(\mathbf{k}_m; k_f), \delta\mathbf{c}(f)^H \varphi'(\mathbf{r})\right\rangle_{L_2(\mathbb{R}^3)} \qquad (48)$$

$$= 2\mathbb{R}e\left\langle \delta\mathbf{c}(f)^H, \int_{\mathbb{R}^2}\left(S_{f_0'}^* \mathcal{F}^* E(\mathbf{k}_m; k_f)\right)\varphi'^*(\mathbf{r})d\mathbf{r}\right\rangle_{L_2(\mathbb{R})}. \qquad (49)$$

Note that in (48) and (49), the domains over which the inner product is defined are different. We thus obtain the gradient of the coefficients as

$$\nabla_{\mathbf{c}(f)} = 2\int_{\mathbb{R}^2} S_{f_0}^* \mathcal{F}^*\left(E(\mathbf{k}_m; k_f)\right)\varphi'(\mathbf{r})d\mathbf{r}. \qquad (50)$$

REFERENCES

[1] A. Pirzkall, T. R. McKnight, E. E. Graves, M. P. Carol, P. K. Sneed, W. W. Wara, S. J. Nelson, L. J. Verhey, and D. A. Larson, "MR-spectroscopy guided target delineation for high-grade gliomas," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 50, pp. 915–928, 2001.
[2] S. J. Nelson, "Analysis of volume MRI and MR spectroscopic imaging data for the evaluation of patients with brain tumors," *Magn. Resonance Med.*, vol. 46, pp. 228–239, 2001.
[3] W. D. Rooney, D. E. Goodkin, N. Schuff, D. J. Meyerhoff, D. Norman, and M. W. Weiner, "1H MRSI of normal appearing white matter in multiple sclerosis," *Multiple Sclerosis*, vol. 3, pp. 231–237, 1997.
[4] M. Hajek, M. Dezortova, and V. Komarek, "1H MR spectroscopy in patients with mesial temporal epilepsy," *MAGMA*, vol. 7, pp. 95–114, 1998.
[5] T. R. Brown, M. B. Kincaid, and K. Ugurbil, "NMR chemical shift imaging in three dimensions," *Proc. Nat. Acad. Sci.*, vol. 79, pp. 3523–3526, 1982.
[6] A. A. Maudsley, S. K. Hilal, W. H. Perman, and H. E. Simon, "Spatially resolved high-resolution spectroscopy by four dimensional NMR," *J. Magn. Resonance*, vol. 51, pp. 147–152, 1983.
[7] L. D. Hall, V. Rajanayagam, and S. Sukumar, "Chemical-shift-resolved tomography using four-dimensional FT imaging," *J. Magn. Resonance*, vol. 61, pp. 188–191, 1985.
[8] A. R. Guimaraes, J. R. Baker, B. G. Jenkins, P. L. Lee, R. M. Wesskoff, B. R. Rosen, and R. G. Gonzalez, "Echoplanar chemical shift imaging," *Magn. Resonance Med.*, vol. 41, pp. 877–882, 1999.
[9] E. Adalsteinsson, P. Irarrazabal, S. Topp, C. Meyer, A. Macovski, and D. M. Speilman, "Volumetric spectroscopic imaging with spiral-based k-space trajectories," *Magn. Resonance Med.*, vol. 39, pp. 889–898, 1998.
[10] A. Adalsteinsson and D. Speilman, "Spatially resolved two dimensional spectroscopy," *Magn. Resonance Med.*, vol. 41, no. 8–12, 1999.
[11] P. Irarrazabal and D. Nishimura, "Fast three dimensional magnetic resonance imaging," *Magn. Resonance Med.*, vol. 33, pp. 656–662, 1995.
[12] S. Posse, C. DeCarli, and D. Le-Bihan, "Three-dimensional echo-planar MR spectroscopic imaging at short echo times in human heart," *Radiology*, vol. 192, pp. 733–738, 1994.
[13] X. Hu, D. N. Levin, P. C. Lauterbur, and T. Spraggins, "SLIM: Spectral localization by imaging," *Magn. Resonance Med.*, vol. 8, pp. 314–322, 1988.
[14] Z. P. Liang and P. C. Lauterbur, "A generalized series approach to spectroscopic imaging," *IEEE Trans. Med. Imag.*, vol. 16, no. 1, pp. 132–137, Jan. 1991.
[15] M. Kienlin and R. Mejia, "Spectral localization with optimal point spread function," *J. Magn. Resonance*, vol. 94, pp. 268–287, 1991.
[16] S. K. Plevritis and A. Macovski, "MRS imaging using anatomically based k-space sampling and extrapolation," *Magn. Resonance Med.*, vol. 34, pp. 686–693, 1995.
[17] S. K. Plevritis and A. Macovski, "Spectral extrapolation of spatially bounded images," *IEEE Trans. Med. Imag.*, vol. 14, no. 3, pp. 487–497, Sep. 1995.
[18] B. J. Soher, P. Vermathen, N. Schuff, D. Wiedermann, D. J. Meyerhoff, M. W. Weiner, and A. A. Maudsley, "Short TE *in vivo* H MR spectroscopic imaging at 1.5 T," *Magn. Resonance Imag.*, vol. 18, pp. 1159–1165, 2000.
[19] A. Ebel, B. J. Soher, and A. A. Maudsley, "Assessment of 3-D proton MR echo-planar spectroscopic imaging using automated spectral analysis," *Magn. Resonance Med.*, vol. 46, pp. 1072–1078, 2001.

[9]Note that as $M \rightarrow \infty$, we are considering the entire Fourier series of the basis functions. Hence, we use the Parseval's relation to obtain (44).

[20] M. Jacob, T. Deller, X. Zhu, and Z. P. Liang, "High-resolution spectroscopic imaging using a deformable spatial spectral model," presented at the Int. Soc. Magn. Resonancen Med. Conf., Miami, FL, 2005.

[21] I. Khalidov, D. Van De Ville, M. Jacob, F. Lazeyras, and M. Unser, "BSLIM: Spectral localization by imaging with explicit B0 field inhomogeneity compensation," *IEEE Trans. Med. Imag.*, vol. 26, no. 7, pp. 990–1000, 2007.

[22] A. Bashir and D. A. Yablonskiy, "Natural linewidth chemical shift imaging (NL-CSI)," *Magn. Resonance Med.*, vol. 56, no. 1, pp. 7–18, 2006.

[23] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, 1999.

[24] D. L. Donoho and M. Elad, "Maximal sparsity representation via l-1 minimization," in *Proc. Nat. Acad. Sci.*, 2003, vol. 23, pp. 2197–2202.

[25] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information?," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 489–509, Apr. 2006.

[26] P. C. Hansen, The L-curve and its use in the numerical treatment of inverse problems Dept. Math. Model, Tech. Univ. of Denmark, Lyngby, Denmark, Tech. Rep., 1999.

[27] A. A. Maudsley, G. B. Matson, J. W. Hugg, and M. W. Weiner, "Reduced phase encoding in spectroscopic imaging," *Magn. Resonance Med.*, vol. 31, pp. 645–651, 1994.

[28] T. E. Skinner and C. H. Glover, "An extended two-point Dixon algorithm for calculating separate water, fat, and B0 images," *Magn. Resonance Med.*, vol. 37, pp. 628–630, 1997.

[29] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Image Process.*, vol. 4, no. 7, pp. 932–946, Jul. 1995.

[30] J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1427–1442, Nov. 2003.

[31] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2783–2795, Oct. 1999.

[32] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2915–2925, Nov. 1994.

[33] V. Govindaraju, K. Young, and A. A. Maudsley, "Proton NMR chemical shifts and coupling constants for brain metabolites," *NMR Biomed.*, 2000.

[34] S. A. Smith, T. O. Levante, B. H. Meier, and R. R. Ernst, "Computer simulations in magnetic resonance. An object oriented programming approach," *J. Magn. Resonance*, vol. 106 a.

[35] J. Tanabe, "Tissue segmentation of the brain in Alzheimer's disease," *Amer. J. Neuroradiol.*, vol. 18, pp. 115–123, 1997.