

# Constrained Magnetic Resonance Spectroscopic Imaging by Learning Nonlinear Low-Dimensional Models

Fan Lam<sup>1</sup>, Member, IEEE, Yahang Li<sup>2</sup>, Student Member, IEEE, and Xi Peng, Member, IEEE

**Abstract**—Magnetic resonance spectroscopic imaging (MRSI) is a powerful molecular imaging modality but has very limited speed, resolution, and SNR tradeoffs. Construction of a low-dimensional model to effectively reduce the dimensionality of the imaging problem has recently shown great promise in improving these tradeoffs. This paper presents a new approach to model and reconstruct the spectroscopic signals by learning a nonlinear low-dimensional representation of the general MR spectra. Specifically, we trained a deep neural network to capture the low-dimensional manifold, where the high-dimensional spectroscopic signals reside. A regularization formulation is proposed to effectively integrate the learned model and physics-based data acquisition model for MRSI reconstruction with the capability to incorporate additional spatio-spectral constraints. An efficient numerical algorithm was developed to solve the associated optimization problem involving back-propagating the trained network. Simulation and experimental results were obtained to demonstrate the representation power of the learned model and the ability of the proposed formulation in producing SNR-enhancing reconstruction from the practical MRSI data.

**Index Terms**—MR spectroscopic imaging, spectroscopy, low-dimensional models, neural network, manifold learning, spatio-spectral constraint.

## I. INTRODUCTION

MAGNETIC resonance spectroscopic imaging (MRSI) is a potentially powerful noninvasive, label-free molecular imaging modality. It allows for simultaneous detection and quantification of the spatiotemporal variations of many endogenous molecules in the human body [1]–[3]. However, due to the high dimensionality of the underlying imaging problem to produce spatially-resolved spectra and the inherently low

SNR as the molecules of interest typically have very low concentrations, in vivo applications of MRSI have been limited to single-voxel spectroscopy or low-resolution, time-consuming acquisitions [2], [3].

In the past several years, a number of constrained MRSI methods have been proposed to improve the speed, resolution and SNR tradeoffs. The key methodologies in these works typically rely on constructing a reduced-complexity signal model of the high-dimensional MRSI data, by exploiting either sparsity in the spatial and/or spectral domains [4]–[10] or low-rankness of the desired spatio-spectral function [11]–[17]. Additional spatial constraints have also been introduced to further take advantages of other anatomical prior information readily available in an MR experiment [18]–[21]. Recently, a subspace MRSI method called SPICE has been developed to successfully achieve fast, high-resolution MRSI by jointly designing both data acquisition and processing in a subspace imaging framework [22]–[24]. The key feature of this approach is modeling individual voxel spectra as a linear combination of a small number of basis functions, which can be predetermined using specially designed navigator data, thus significantly reducing the degrees-of-freedom and allowing better tradeoffs for speed, resolution and SNR. However, to capture more complicated spatio-spectral variations, the dimension of the linear subspace can increase substantially, reducing its efficiency, thus motivating the need for a more general nonlinear low-dimensional model.

Learning a general nonlinear model for high-dimensional functions is a challenging problem. In the context of MRSI, locally linear embedding (LLE) [25] and Laplacian eigenmaps (LE) [26] have been applied to classify spectra from normal and diseased tissues by estimating the low-dimensional manifolds where each class of spectra was assumed to reside. But incorporating such classification models into the imaging process is difficult for which more accurate representation is required. Meanwhile, the recent success of deep neural network based methods in learning complex functional mapping and extracting nonlinear features from high-dimensional data presents new opportunities to address the model learning problem [27]–[29]. A number of works have been proposed in MRSI, mainly focusing on spectral quantification [30]–[32] or spectral artifact removal [33]. A common approach among these methods is to directly learn the entire inverse function that maps the noisy and artifact-containing signals to the desired artifact-free ones or the spectral parameters (e.g.,

Manuscript received April 15, 2019; revised July 12, 2019 and July 18, 2019; accepted July 18, 2019. Date of publication July 23, 2019; date of current version February 28, 2020. (Corresponding author: Fan Lam.)

F. Lam is with the Department of Bioengineering, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: fanlam1@illinois.edu).

Y. Li is with the Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

X. Peng is with the Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

This article has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author.

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

Digital Object Identifier 10.1109/TMI.2019.2930586

molecule concentrations or lineshape parameters), by training a deep neural network (DNN) [30], [33], [34]. This approach requires the learned function to simultaneously capture all the nuances in the noise, artifacts, as well as underlying true signals at all the possible SNRs, thus dramatically increasing the complexity of this function and the learning problem. Moreover, it is also sensitive to SNR levels and data acquisition designs.

We proposed in this work a different approach to learn a nonlinear low-dimensional model that captures the inherent degrees-of-freedom in spectroscopic signals and to incorporate such a learned representation for constrained MRSI reconstruction. Specifically, we recognized that a general NMR spectrum can be characterized by a small number of parameters based on the well-defined spectral quantification model, thus should reside in a low-dimensional manifold embedded in the original high-dimensional space. Accordingly, we designed a deep autoencoder network (DAE) to capture this manifold. The effectiveness of DAE in learning nonlinear low-dimensional representation for high-dimensional data has been well-documented [27]–[29] but requires a large amount of training data. To this end, we acquired a small amount of training data to estimate the distributions of molecule-dependent parameters and then used the spectral fitting model to generate the data needed to train the DAE. A regularization formulation was devised to integrate the learned representation and the spatio-spectral encoding model for constrained MRSI reconstruction. An efficient algorithm was developed to solve the associated optimization problem. The proposed model was evaluated against linear dimensionality reduction and demonstrated a more efficient representation of MRSI data. Simulation and experimental results were obtained to demonstrate the capability of the proposed method in producing improved spatio-spectral reconstructions over existing methods.

The rest of the paper is organized as follows. Section II provides background on MR spectroscopy modeling and the need for nonlinear dimension reduction, as well as a brief review of autoencoder based neural networks. Section III presents details on the proposed DAE-based learning method, reconstruction formulation, and optimization algorithm. Section IV summarizes the experimental results followed by discussion and conclusion in Sections V and VI.

## II. BACKGROUND

### A. Spectroscopy Signal Modeling

The imaging problem in MRSI can be defined as recovering a high-dimensional spatiotemporal function  $\rho(\mathbf{r}, t)$  from a set of  $(\mathbf{k}, t)$ -space measurements  $d_p(\mathbf{k}, t)$  that can be expressed as

$$d_p(\mathbf{k}, t) = \int_V s_p(\mathbf{r}) \rho(\mathbf{r}, t) e^{i2\pi \delta f(\mathbf{r})t} e^{-i2\pi \mathbf{k}\mathbf{r}} d\mathbf{r} + n_p(\mathbf{k}, t), \quad (1)$$

where  $V$  is the imaging volume of interest,  $s_p(\mathbf{r})$  is the sensitivity map for the  $p$ th coil used in data acquisition,  $\delta f(\mathbf{r})$  denotes the  $B_0$  field inhomogeneity,  $\mathbf{k}$  the coordinates for the Fourier encoding space and  $n_p(\mathbf{k}, t)$  the measurement noise. The  $\mathbf{k}$ -space sampling can be designed based on resolution and time requirements. While  $\rho(\mathbf{r}, t)$  is very high-dimensional and challenging to recover at a high resolution

from the typically noisy  $d_p(\mathbf{k}, t)$ , there is abundant prior information that can be exploited for improved reconstruction. Specifically, the FID signals at individual voxels can be generally modeled as

$$\rho(\mathbf{r}, t) = \sum_{m=1}^M c_m(\mathbf{r}) \phi_m(t) e_m(t; \boldsymbol{\theta}_m(\mathbf{r})), \quad (2)$$

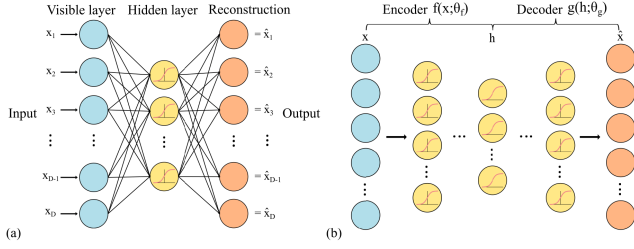
where  $c_m$  and  $\phi_m(t)$  represent the concentration and time-domain basis function of the  $m$ th molecule, respectively, and  $e_m(t; \boldsymbol{\theta}_m)$  captures molecule-dependent time-domain modulation functions that can be described by a few experiment and physiology relevant parameters in  $\boldsymbol{\theta}_m$ . The Fourier transform of  $\phi_m(t)$  is a resonance structure containing the relative frequencies, amplitudes and phases of different spectral peaks for the  $m$ th molecules, which is data independent and can be predicted by spin physics. A most common specification of  $e_m(t; \boldsymbol{\theta}_m)$  is the exponential lineshape model below [35]:

$$\rho(\mathbf{r}, t) = \sum_{m=1}^M c_m(\mathbf{r}) \phi_m(t) e^{-t/T_{2,m}^*(\mathbf{r}) + i2\pi \delta f_m(\mathbf{r})t} e^{-\beta(\mathbf{r})t^2}, \quad (3)$$

where  $\{T_{2,m}^*\}$  and  $\{\delta f_m\}$  denote the tissue and physiological condition dependent lineshape and frequency shifts for each molecule, respectively, and  $e^{-\beta(\mathbf{r})t^2}$  captures additional lineshape distortion due to system imperfection such as intravoxel  $B_0$  inhomogeneity and eddy currents. Accordingly, all spectral signals specified by Eq. (3) with varying  $\{c_m\}$ ,  $\{T_{2,m}^*\}$ ,  $\{\delta f_m\}$  and  $\beta$  can be treated as points residing on a low-dimensional manifold embedded in the high-dimensional vector space [36], [37]. This manifold can be well-approximated by a linear subspace if the parameters are in a relatively narrow range. This property has been exploited in the recent subspace-based reconstruction and spectral quantification methods to successfully achieve fast, high-resolution MRSI (by modeling spectroscopic signals as a linear combination of a number of basis functions) [22]–[24], [38]. However, as the ranges of the parameters increase, e.g., due to physiological variations, increased system nonideality or the presence of a certain pathology, the dimension of the subspace can increase substantially to ensure an accurate approximation, reducing its efficiency. Therefore, a general nonlinear low-dimensional model that can accurately capture all the physiologically relevant spectral variations is desired. A toy example is shown in Fig. S1 (Supplementary Materials) to illustrate this point.

### B. Deep Autoencoders

A number of methods have been developed to learn nonlinear features from MRSI data for classification purpose, e.g., [25], [26], [39]. The incorporation of learned features into constrained reconstruction requires representations with much higher accuracy and flexibility. Motivated by the recent success of deep neural network in representation learning [27]–[29], we propose in this work to use a deep autoencoder to learn an accurate and efficient low-dimensional model for MR spectroscopy signals. To provide a context for the proposed method, we present in this section a brief introduction to autoencoders and deep autoencoders.



**Fig. 1.** Illustration of AEs: (a) A basic three-layer AE with fully connected layers and a hidden layer to extract nonlinear features of the data; (b) A DAE constructed by stacking the basic AEs shown in (a) with multiple nonlinear hidden layers, which can be trained to extract hierarchical features from high-dimensional data and for nonlinear dimensionality reduction.

An autoencoder (AE) is a special type of artificial neural network that is typically composed of fully connected layers and has been developed to learn the underlying representation of data of various types [27]. Figure 1a illustrates a commonly used basic autoencoder with the first layer being linear units to represent an input vector  $\mathbf{x} \in \mathcal{R}^D$ , the middle layer being a set of nonlinear units to capture the “encoded features”  $\mathbf{h} \in \mathcal{R}^{D'}$  of  $\mathbf{x}$ , and the last layer being also linear units to represent an output vector  $\hat{\mathbf{x}}$  with the same dimension as  $\mathbf{x}$ . The mapping from  $\mathbf{x}$  to  $\mathbf{h}$  is referred to as the encoder, denoted as  $f(\mathbf{x})$ , while the mapping from  $\mathbf{h}$  to  $\hat{\mathbf{x}}$  is referred to as the decoder, denoted as  $g(\mathbf{h})$ . For the simple AE in Fig. 1a, the relationships between  $\mathbf{x}$ ,  $\mathbf{h}$  and  $\hat{\mathbf{x}}$  can be mathematically expressed as:

$$\mathbf{h} = f(\mathbf{x}) = \mathbf{a}(\mathbf{W}_1 \mathbf{x}; \mathbf{b}), \quad (4)$$

$$\hat{\mathbf{x}} = g(\mathbf{h}) = \mathbf{W}_2 \mathbf{h} \quad (5)$$

where  $\mathbf{W}_1$  and  $\mathbf{W}_2$  denote the weights from the input layer to the hidden layer and hidden layer to the output layer, respectively, and  $\mathbf{a}(\cdot)$  contains element-wise nonlinear activation functions with bias vector  $\mathbf{b}$ . This model allows one to extract nonlinear low-dimensional features of data for which linear dimension reduction techniques do not offer satisfactory performance. Theoretical connections to nonlinear manifold modeling have also been made [40], [41].

Deep autoencoders (DAEs) are deep neural networks that are constructed by stacking multiple nonlinear encoding and decoding layers from the standard AEs (Fig. 1b). The multiple nonlinear layers introduce more complexity, thus offering stronger representation and feature extraction powers than the standard AEs. Mathematically, the features extracted from the multiple nonlinear layers can be denoted as  $\mathbf{h} = \mathbf{a}_L(\mathbf{W}_L \cdots \mathbf{a}_2(\mathbf{W}_2 \mathbf{a}_1(\mathbf{W}_1 \mathbf{x}; \mathbf{b}_1); \mathbf{b}_2) \cdots; \mathbf{b}_L)$  ( $\{\mathbf{a}_i\}_{i=1}^L$  being nonlinear functions at different layers); and the multi-layer encoders and decoders can be expressed generally as  $f(\mathbf{x}; \boldsymbol{\theta}_f)$  and  $g(\mathbf{h}; \boldsymbol{\theta}_g)$ , respectively. The vectors  $\boldsymbol{\theta}_f$  and  $\boldsymbol{\theta}_g$  contain all the weights and bias in the encoding and decoding layers. DAEs have been trained to learn compact low-dimensional representations of complicated, high-dimensional data (e.g., audios and images) [27], [29]. Specifically, the training can be done by solving the following problem that minimizes an error metric  $\varepsilon$  with respect to  $\boldsymbol{\theta}_f$  and  $\boldsymbol{\theta}_g$  such that the output of the network  $\hat{\mathbf{x}}_s = g(f(\mathbf{x}_s; \boldsymbol{\theta}_f); \boldsymbol{\theta}_g)$  can accurately approximate

the corresponding input  $\mathbf{x}_s$  ( $s$  being the sample index), i.e.,

$$\{\hat{\boldsymbol{\theta}}_f, \hat{\boldsymbol{\theta}}_g\} = \arg \min_{\boldsymbol{\theta}_f, \boldsymbol{\theta}_g} \frac{1}{N} \sum_{s=1}^N \varepsilon(\mathbf{x}_s, g(f(\mathbf{x}_s; \boldsymbol{\theta}_f); \boldsymbol{\theta}_g)). \quad (6)$$

Regularization can be imposed on  $\boldsymbol{\theta}_f$ ,  $\boldsymbol{\theta}_g$ ,  $f(\cdot; \boldsymbol{\theta}_f)$  and/or  $g(\cdot; \boldsymbol{\theta}_g)$  such that the encoded features are not required to be lower-dimensional than the input and still being able to avoid overfitting or simply learning an identity transform [28], [42], [43]. A common choice for  $\varepsilon$  is the  $L_2$  error, i.e.,  $\varepsilon = \|\mathbf{x}_s - \hat{\mathbf{x}}_s\|_2^2$ . Applications of DAEs to MR image reconstruction and MRSI spectral quantification have been recently explored [32], [44], where the networks were trained to generate images or spectral parameters directly from the measured data by extracting low-dimensional features.

### III. PROPOSED METHOD

#### A. Low-Dimensional Spectral Model Learning Using DAE

Recognizing that an inherent low-dimensional representation should exist for the MR spectra specified by Eq. (3), we propose here a strategy to learn this general nonlinear model for MR spectroscopic signals using a DAE. Two major issues arise for this learning task. First, deep neural networks require a large number of high-quality training data which is a generally luxury for many MRSI applications. Secondly, proper designs of architecture and training procedure are needed to learn a representation of the signals that can be useful in the imaging process (e.g., reconstruction from noisy measurements).

We address the first issue by combining the physical model in Eq. (3), quantum mechanical (QM) simulations, and experimentally acquired training data. More specifically, the molecule basis functions,  $\phi_m(t)$ , can be obtained from QM simulations [45], [46]. These resonance structures are molecular structure specific and can be assumed to be invariant with respect to different subjects and experiments. Meanwhile, we acquired high-SNR, low-resolution CSI data from healthy volunteers. These data were subject to spectral fitting from which empirical distributions of the spectral parameters, i.e.,  $c_m$ ,  $T_{2,m}^*$ , and  $\delta f_m$ , were obtained. The empirical distributions were fitted to parametric Gaussian distributions to allow the generation of more randomly distributed parameters (Fig. 2). Finally, a large collection of FIDs can then be synthesized by combining  $\phi_m(t)$  and randomly generated  $c_m$ ,  $T_{2,m}^*$ , and  $\delta f_m$  for training a DAE. Randomly distributed  $\beta$  corresponding to Gaussian linewidths with 1 Hz mean and 0.5 Hz standard deviation were generated and incorporated into the training data. Additional metabolite-dependent phases can also be incorporated to better capture realistic signal variations which were assumed to follow a uniform distribution for simplicity. To effectively extract the low-dimensional features of these complex-valued signals, a special multi-layer DAE was designed. More specifically, we rearranged all the data by concatenating the real and imaginary parts as inputs. The first hidden layer of the proposed DAE was designed to have more units than the input layer such that the data can be lifted to a higher-dimensional space to i) increase the model’s capacity and ii) enable the extraction of complex nonlinear features.



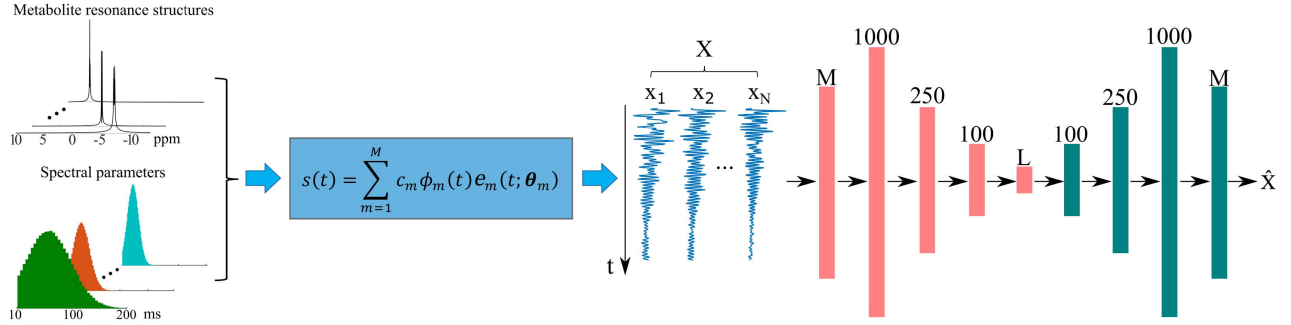


Fig. 2. The proposed strategy to learn a nonlinear low-dimensional representation of spectroscopic signals: (Most left column) Metabolite resonance structures are obtained by quantum mechanical simulations and spectral parameters are generated from distributions estimated from empirical data. This information is fed into the commonly used spectral fitting model (blue box) to generate a large collection of FID data for training, denoted as  $\mathbf{X}$ . (Most right column) A DAE network with fully-connected layers was trained to encode  $\mathbf{X}$  into  $L$ -dimensional features which can accurately reconstruct the original data. The mean-squared error was minimized for training. See the texts for detailed descriptions of the network design.

This layer was then followed by a “bottleneck” structure that gradually reduces the dimensionality with the middle layer being linear to ensure a sufficient dynamic range for the final encoded low-dimensional features. The overall data generation and DAE-based learning strategy are illustrated in Fig. 2. More comparison on the network designs can be found in the supplementary materials. The entire trained network is denoted by a function mapping  $\mathcal{C}(\cdot)$  with fixed parameters, which minimizes the error between input and output, i.e.,  $\|\mathcal{C}(\mathbf{X}) - \mathbf{X}\|_2^2$ .

We will use  $^{31}\text{P}$  spectroscopy data in this work to demonstrate the capability of the proposed approach considering their clearly defined spectral features, absence of nuisance signals and macromolecule baseline (see Discussion for more details on extensions to MRSI of other nuclei). To this end, the QM simulated resonance structures include the commonly observed  $^{31}\text{P}$ -containing metabolites, as described in [46]. The CSI data to estimate the empirical parameter distributions were acquired on a 7T scanner (more details in the Results section).

### B. Constrained MRSI Using the Learned Model

With the trained DAE  $\mathcal{C}(\cdot)$  to extract the low-dimensional features of spectroscopy data, the key issue is how to incorporate this learned model into the imaging process. To this end, we proposed a constrained reconstruction formulation that allows an effective integration of the imaging acquisition model and the learned model through a regularization functional. Specifically, the reconstruction is formulated as follows:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{d} - \Omega\{\mathbf{F}\mathbf{B} \odot \mathbf{X}\}\|_2^2 + \lambda_2 \|\mathbf{D}_w \mathbf{X}\|_F^2 + \lambda_1 \sum_{n=1}^N \|\mathcal{C}(\mathbf{X}_n) - \mathbf{X}_n\|_2^2 \quad (7)$$

where  $\mathbf{X}$  is a matrix representation of the spatiotemporal function of interest with a size of  $N \times T$  and each row being a length- $T$  voxel FID,  $\mathbf{B}$  models the effects induced by  $B_0$  inhomogeneity,  $\odot$  denotes a point-wise multiplication,  $\mathbf{F}$  the Fourier transform,  $\Omega$  the sampling operator, and  $\mathbf{d}$  is a vector containing the  $(k, t)$ -space data. The first term enforces the imaging acquisition model and data consistency, the second term imposes spatial edge-preserving regularization with  $\mathbf{D}_w$

being a weighted finite-difference operator [20], [22], and the third term incorporates the prior that each reconstructed FID ( $\mathbf{X}_n$ ) should reside in the low-dimensional manifold captured by  $\mathcal{C}(\cdot)$ .  $\|\cdot\|_F$  denotes the Frobenius norm. Compared to the common approach of directly learning the highly complex inverse mapping (through either a single neural network or cascaded neural networks), the proposed method represents a different but rigorous way of integrating physics-based forward modeling and data-driven priors within a unified formulation. Solving the optimization problem in Eq. (7) is challenging, as it requires nonlinear programming to handle the regularization term containing the learned neural network, which however is not efficient for the linear least-squares terms.

### C. Optimization Algorithm

We describe here an efficient algorithm to solve Eq. (7). Specifically, to decouple the linear least-squares and the nonlinear regularization functional, we introduce an auxiliary variable  $\mathbf{S} = \mathbf{B} \odot \mathbf{X}$  and rewrite the problem as

$$\begin{aligned} \hat{\mathbf{X}} = \arg \min_{\mathbf{X}} & \|\mathbf{d} - \Omega\{\mathbf{F}\mathbf{S}\}\|_2^2 + \lambda_2 \|\mathbf{D}_w \bar{\mathbf{B}} \odot \mathbf{S}\|_F^2 \\ & + \lambda_1 \sum_{n=1}^N \|\mathcal{C}(\mathbf{X}_n) - \mathbf{X}_n\|_2^2 \\ \text{s.t. } & \mathbf{B} \odot \mathbf{X} = \mathbf{S}. \end{aligned} \quad (8)$$

$\bar{\mathbf{B}}$  represents the conjugate of  $\mathbf{B}$  (element-wise). With this form, the alternating direction method of multipliers (ADMM) can then be applied to solve this equivalent problem [47]. The ADMM algorithm decomposes the original problem into simpler subproblems which can be solved efficiently. More specifically, the following three steps were performed iteratively in our proposed algorithm:

(I) Update  $\mathbf{X}$  with fixed  $\mathbf{S}^{(i)}$  and  $\mathbf{Y}^{(i)}$  with  $i$  denoting the iteration number and  $\mathbf{Y}$  the Lagrangian multiplier

$$\begin{aligned} \mathbf{X}^{(i+1)} = \arg \min_{\mathbf{X}} & \lambda_1 \sum_{n=1}^N \|\mathcal{C}(\mathbf{X}_n) - \mathbf{X}_n\|_2^2 \\ & + \frac{\mu}{2} \left\| \mathbf{B} \odot \mathbf{X} - \mathbf{S}^{(i)} + \frac{\mathbf{Y}^{(i)}}{\mu} \right\|_F^2. \end{aligned} \quad (9)$$

(II) Update  $\mathbf{S}$  with fixed  $\mathbf{X}^{(i+1)}$  and  $\mathbf{Y}^{(i)}$

$$\mathbf{S}^{(i+1)} = \arg \min_{\mathbf{S}} \|\mathbf{d} - \Omega\{\mathbf{FS}\}\|_2^2 + \lambda_2 \|\mathbf{D}_w \bar{\mathbf{B}} \odot \mathbf{S}\|_F^2 + \frac{\mu}{2} \left\| \mathbf{B} \odot \mathbf{X}^{(i+1)} - \mathbf{S} + \frac{\mathbf{Y}^{(i)}}{\mu} \right\|_F^2. \quad (10)$$

(III) Update  $\mathbf{Y}$

$$\mathbf{Y}^{(i+1)} = \mathbf{Y}^{(i)} + \mu \left( \mathbf{B} \odot \mathbf{X}^{(i+1)} - \mathbf{S}^{(i+1)} \right). \quad (11)$$

The first subproblem requires general unconstrained optimization solvers due to the highly nonlinear function  $\mathcal{C}(\cdot)$ . Although the original problem has high dimensionality, it can be solved in a voxel-by-voxel fashion as the Frobenius norm is separable. Furthermore, based on the auto-encoder network design, we can derive the expression of the gradients for individual voxels. More specifically, denote

$$f_n(\mathbf{X}_n) = \lambda_1 \|\mathcal{C}(\mathbf{X}_n) - \mathbf{X}_n\|_2^2 + \frac{\mu}{2} \left\| \left[ \mathbf{B} \odot \mathbf{X}_n - \mathbf{S}^{(i)} + \frac{\mathbf{Y}^{(i)}}{\mu} \right]_n \right\|_2^2 \quad (12)$$

as the cost function for the  $n$ th voxel, the gradient for  $f_n(\cdot)$  can be written as

$$\nabla f_n(\mathbf{X}_n) = 2\lambda_1 (\mathbf{J}_C - \mathbf{I})^T (\mathcal{C}(\mathbf{X}_n) - \mathbf{X}_n) + \mu \mathbf{B}_{(n)}^H \left( \mathbf{B}_{(n)} \mathbf{X}_n - \mathbf{S}_n^{(i)} + \frac{\mathbf{Y}_n^{(i)}}{\mu} \right). \quad (13)$$

where  $\mathbf{J}_C \in \mathcal{R}^{T \times T}$  is the Jacobian for the neural network mapping  $\mathcal{C}(\cdot)$ ,  $\mathbf{I}$  is a  $T \times T$  identity matrix and  $\mathbf{B}_{(n)}$  denotes a diagonal matrix formed by the  $n$ th row of  $\mathbf{B}$ . For a neural network with  $L$  layers and a final linear layer,  $\mathbf{J}_C$  can be derived through back-propagation as

$$\mathbf{J}_C = \mathbf{W}_L^T \times \prod_{l=1}^{L-1} \mathbf{U}_l \mathbf{W}_l^T, \quad (14)$$

where  $\mathbf{W}_l$  contains the linear weights for the  $l$ th layer and  $\mathbf{U}_l$ 's are diagonal matrices containing the derivatives for the nonlinear activation functions at the  $l$ th layer. For instance, if the rectified linear unit (ReLU) is chosen, the diagonal elements in  $\mathbf{U}_l$  are simply unit step functions  $u([\mathbf{X}_{l-1} \mathbf{W}_l + \mathbf{b}_l]_k)$  with  $\mathbf{b}_l$  being the bias vector and  $k$  the output index. These separable problems can then be solved in parallel and efficiently using a quasi-Newton algorithm, i.e., the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method [48], [49].

With the updated  $\mathbf{X}^{(i+1)}$ , the second subproblem is equivalent to solving the following linear equations

$$\mathbf{F}^H \Omega^H \Omega \{\mathbf{FS}\} + \lambda_2 \mathbf{B} \odot \mathbf{D}_w^H \mathbf{D}_w \bar{\mathbf{B}} \odot \mathbf{S} + \frac{\mu}{2} \mathbf{S} = \mathbf{F}^H \Omega^H \{\mathbf{d}\} + \frac{\mu}{2} \left( \mathbf{B} \odot \mathbf{X}^{(i+1)} + \frac{\mathbf{Y}^{(i)}}{\mu} \right), \quad (15)$$

which can be easily solved by exploiting the special structures of  $\mathbf{F}^H \Omega^H \Omega \mathbf{F}$  and  $\mathbf{D}_w^H \mathbf{D}_w$  as suggested in [50]. The iterations were terminated when either  $i$  reaches a certain number or the relative changes in  $\mathbf{X}^{(i+1)}$  and  $\mathbf{X}^{(i)}$  fall below a given threshold (e.g.,  $10^{-3}$ ).

## D. Single-Voxel Spectroscopy Denoising

The proposed method can also be applied to spectroscopy data acquired without spatial encoding. In this case, the formulation in Eq. (7) can be simplified into

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{d} - \mathbf{x}\|_2^2 + \lambda_1 \|\mathcal{C}(\mathbf{x}) - \mathbf{x}\|_2^2 \quad (16)$$

where  $\mathbf{x}$  is a vector representing the desired noiseless FID and  $\mathbf{d}$  is the acquired data. This can be viewed as a special case of the problem in Eq. (12), with the following gradient

$$\nabla f(\mathbf{x}) = 2(\mathbf{x} - \mathbf{d}) + 2\lambda_1 (\mathbf{J}_C - \mathbf{I})^T (\mathcal{C}(\mathbf{x}) - \mathbf{x}). \quad (17)$$

The problem can again be solved efficiently using the BFGS or other nonlinear optimization algorithms. This simplification can be useful for many single-voxel spectroscopy experiments where denoising has mostly been limited to time-domain or wavelet-based filtering using only the noisy data [2], [51]. The proposed method allows an effective and flexible means to incorporate data-driven priors obtained from other high-quality data to improve the results of general single-voxel acquisitions.

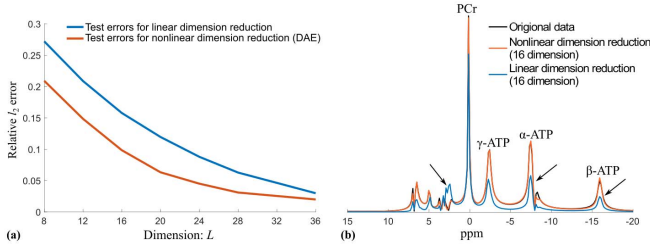
## E. Training and Other Implementation Details

We generated 300,000  $^{31}\text{P}$  MR spectra as described in Section III.A for model learning. The molecules we considered in this work are  $^{31}\text{P}$ -containing compounds commonly observed during in vivo  $^{31}\text{P}$  MRS/MRSI experiments, i.e., PCr,  $\alpha$ -ATP,  $\beta$ -ATP,  $\gamma$ -ATP, Pi, tNAD, PE, PC, GPC, GPE, and MP [46]. The empirical Gaussian distributions for the spectral parameters were estimated from 5 brain CSI data acquired on a 7T scanner (Siemens Magnetom). The  $T_{2,m}^*$  values were lower bounded by 5 ms and upper bounded by 200 ms, and the relative concentrations  $c_m$  were bounded between 0 and 2. For tNAD, GPC, GPE and MP with unreliably estimated spectral parameters (large variances), the literature values were used with a standard deviations at half of the mean values [46], [52]. Metabolite-dependent phases uniformly distributed between  $-\pi/4$  and  $\pi/4$  were included. Among the 300K FIDs, 200K samples were used for training and the remaining 100K for testing. We first implemented the DAE in TensorFlow [53], which was constructed with fully-connected layers in both the encoder and decoder with a dimension structure of  $T - 1000 - 250 - 100 - L - 100 - 250 - 1000 - T$  ( $T = 512$ ). ReLU activation functions were used in the hidden layers except the middle layer. Training was performed using the mean-squared error (MSE) loss and stochastic gradient descent (SGD) with an Adam optimizer [54], with a batch size of 500, learning rate 0.001, 300 epochs and default parameters for the moment estimates. The trained parameters of the DAE were then ported to Matlab and integrated into the proposed formulation and algorithm.

## IV. RESULTS

### A. Numerical Simulations

Simulation studies have been conducted to evaluate the proposed low-dimensional model learning and reconstruction method. We first investigated the approximation accuracy

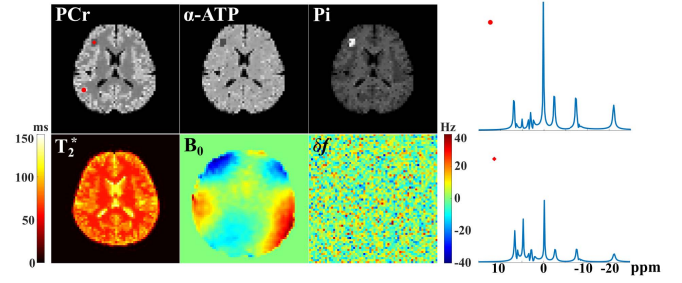


**Fig. 3.** Approximation accuracy of the learned model: (a) Relative  $\ell_2$  errors for the test signals reconstructed from  $L$ -dimensional features encoded by the proposed DAE (red curve), compared to the linear dimension reduction (blue curve); (b) Reconstructions of a single spectrum using 16-dimensional DAE and subspace, respectively. Improved accuracy offered by the DAE is clearly visible. The black arrows in (b) indicate spectral features lost by the linear subspace approximation but recovered by our learned model.

of our learned low-dimensional representation by comparing the dimensionality reduction errors at different model orders (denoted as  $L$ ) with those from a linear dimensionality reduction. The linear dimension reduction was done by projecting the data onto an  $L$ -dimensional subspace determined by PCA of training data ([12], [22]). The results are shown in Fig. 3. As can be seen, the learned DAE-based model yielded higher approximation accuracy (lower relative  $\ell_2$  errors) than the linear subspace approximation across different model orders (Fig. 3a), especially at a smaller  $L$ . As  $L$  increases, the difference becomes smaller. The better representation efficiency offered by the learned model is further demonstrated by comparing the approximations of a fitted experimentally acquired in vivo spectrum (Fig. 3b). A low-dimensional subspace approximation ( $L = 16$ ) leads to noticeable distortion of spectral features which are captured by the learned model.

A numerical phantom was constructed to validate the proposed MRSI reconstruction method using the learned nonlinear model. Specifically, segmented brain tissue compartments, i.e., gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF), were obtained from an experimentally acquired  $T_1$ -weighted structural image. Regional concentrations and linewidth parameters for the 11 molecules mentioned above were assigned based on literature values [46], [52].<sup>1</sup> These compartmental images (constant concentrations and linewidths in each region) were then combined using the tissue fraction maps (obtained from segmentation) as weightings to generate continuously varying concentration and linewidth maps, which along with the metabolite basis  $\{\phi_m\}$  were fed into Eq. (3) to synthesize spatially varying  $^{31}\text{P}$  spectra. To mimic a more practical scenario, we simulated voxel-dependent random frequency shifts for different molecules (mean zero and a standard deviation of 10 Hz) and residual  $B_0$  field inhomogeneity induced spectral shifts (mean zero and a standard deviation of 10 Hz). A lesion-like feature was also included to evaluate the proposed method's capability in discerning novel spatio-spectral features. The concentrations of

<sup>1</sup>The linewidths for the white matter region was randomly increased by 15%-20% for different molecules to account for the  $T_2$  differences. The concentrations for all molecules in the CSF were assigned to be 0.1% of the concentrations in gray matter.



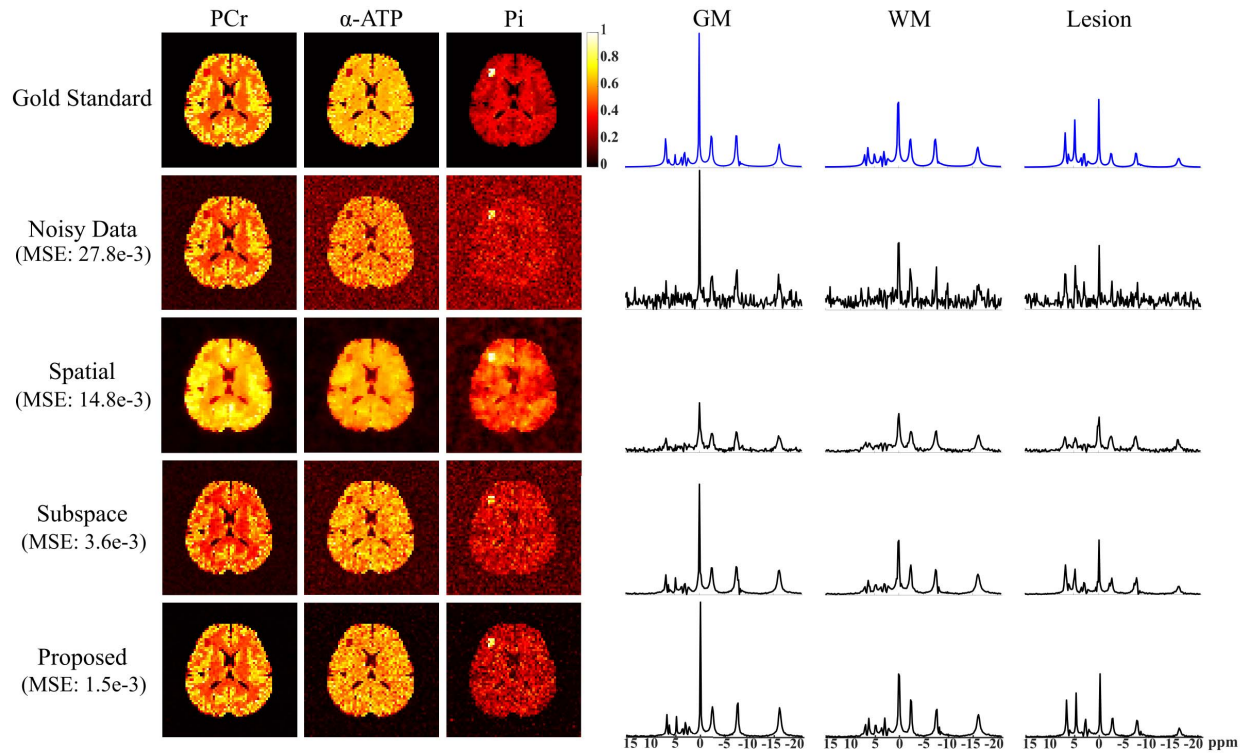
**Fig. 4.** The computational phantom for evaluating different reconstruction methods. The top row shows maps of PCr,  $\alpha$ -ATP and Pi, illustrating distinct contrasts for different molecules. The bottom row shows maps of  $T_2^*$  (PCr),  $B_0$  inhomogeneity and random frequency shifts used in simulation. Representative voxel spectra from the GM and simulated lesion are shown on the most right (locations marked by red symbols).

PCr and ATP were reduced by a factor of 2 in the lesion area compared to the GM, while the concentrations of Pi, PE, and PC were increased by a factor of 3. Figure 4 shows the simulated maps of PCr,  $\alpha$ -ATP, and Pi as well as representative voxel spectra from the GM and lesion, respectively. The  $T_2^*$ ,  $B_0$  inhomogeneity and random frequency shift maps used for simulation are also illustrated. Noisy data were simulated by adding complex white Gaussian noise corresponding to different SNRs for the PCr peak. The SNR is defined as the ratio between the maximum PCr peak amplitude and noise standard deviation ( $\sigma$ ), i.e.,

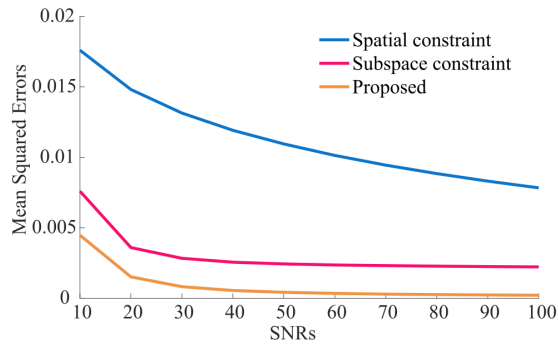
$$\text{SNR} = \frac{\max |\rho(\mathbf{r}, f_{\text{PCr}})|}{\sigma}. \quad (18)$$

A set of denoising results obtained by the proposed method from simulation data with an SNR of 20 is shown in Fig. 5, with comparison to the noisy data (obtained by the standard Fourier reconstruction) as well as results produced by a spatially regularized reconstruction [20] and a subspace constrained reconstruction (SPICE) also with the spatial regularization term [12], [22]. The spatially regularized reconstruction is equivalent to solving Eq. (7) with  $\lambda_1 = 0$ . For the SPICE reconstruction, we used a 24-dimension subspace estimated from the same data used for training the DAE. Normalized mean-squared errors (MSEs) were calculated for each method. The regularization parameters were manually selected to minimize the MSE for the spatially constrained and SPICE reconstructions. As shown by the results (Fig. 5), the spatially constrained reconstruction improved the SNR but tends to oversmooth the images and leads to additional spectral lineshape distortions (third row). The subspace method provided further improved reconstruction (fourth row), while the proposed method produced the best result (last row), both qualitatively and quantitatively, as demonstrated by the reconstructed metabolite maps (obtained by peak integration), voxel spectra and the normalized MSEs. Figure 6 further compares the errors for different methods under various SNRs. The proposed method consistently achieved the lowest errors in all cases. It is worth noting that the spatial regularization serves an auxiliary role to provide additional improvement in the proposed method thus  $\lambda_2$  was not optimized and set to a small value, which helps avoid oversmoothing. To further investigate

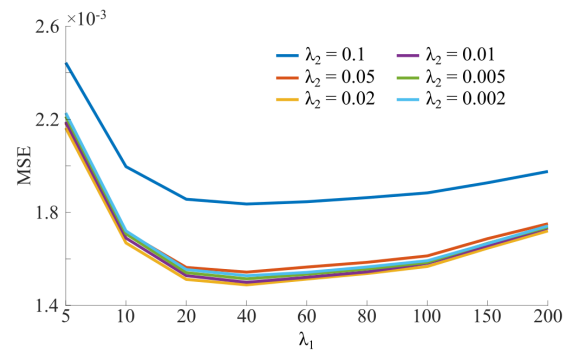




**Fig. 5.** Simulation results showing spatospectral reconstructions from the ground truth (Gold Standard), noisy data (Noisy Data), anatomically constrained reconstruction (Spatial), SPICE reconstruction (Subspace) and the proposed method (Proposed), respectively. The normalized MSEs are shown for each case under the method labels. The first three columns compare the reconstructed maps of PCr,  $\alpha$ -ATP and Pi, and the last three columns show selected voxel spectra from the GM, WM and lesion areas. As can be seen, the proposed method produces significant SNR enhancement and the lowest error while preserving spatospectral features.



**Fig. 6.** Reconstruction errors for different methods (identified by different colored curves) under various SNR levels (SNR defined in texts). As can be seen, the proposed method consistently yields the lowest errors.



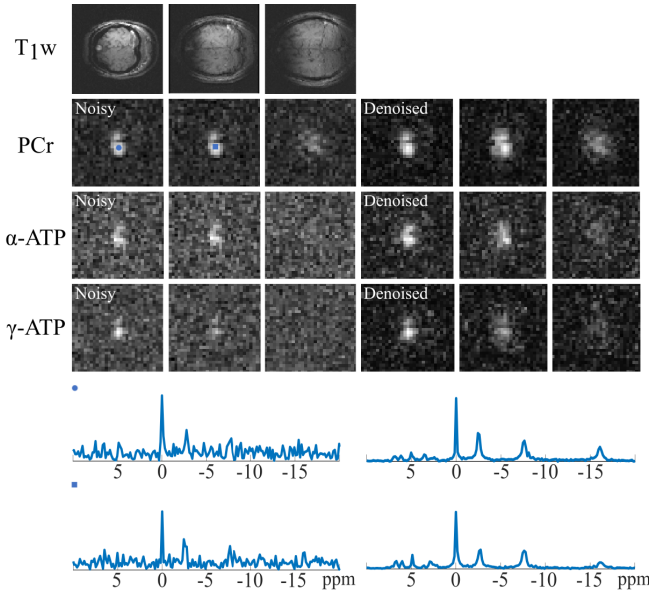
**Fig. 7.** Effects of  $\lambda_1$  and  $\lambda_2$  on the proposed reconstruction. Each curve plots the MSEs w.r.t. different values of  $\lambda_1$  and a fixed  $\lambda_2$ , while the results with different  $\lambda_2$ 's are encoded in different colors.

the effects of regularization parameters, we performed reconstructions with different values of  $\lambda_1$  and  $\lambda_2$  at an SNR of 20. The errors are compared in Fig. 7 and Fig. S2. As expected, a too small regularization parameter has minimal denoising effects (larger errors) while a too large value can also increase the error. But the reconstruction performance remains robust to a large range of parameter values.

### B. Experimental Studies

The performance of the proposed method under practical experimental conditions has been evaluated using brain  $^{31}\text{P}$ -MRSI data acquired from healthy volunteers on a

7T system (Siemens Magnetom) using a double-tuned  $^{31}\text{P}$ - $^1\text{H}$  surface coil. Data were acquired using a CSI sequence [55] with the following parameters: TR/TE = 170/2.3 ms, field-of-view (FOV) =  $200 \times 200 \times 100 \text{ mm}^3$ , matrix size =  $32 \times 32 \times 10$  (corresponding to a nominal voxel size of approximately  $6 \times 6 \times 10 \text{ mm}^3$ ), and 512 FID samples at 5 kHz spectral bandwidth (BW). The total acquisition time was 26 min with elliptical sampling and two signal averages. Gaussian distributed zeroth-order phases were included in generating the data for training the DAEs used for in vivo reconstruction (with mean and standard deviation estimated from the experimental data). The acquisition delay was also considered by defining



**Fig. 8.** Reconstructions from the in vivo data acquired using a surface coil. The first row contains the anatomical images within the imaging volume ( $T_1w$ ). The maps of PCr,  $\alpha$ -ATP and  $\gamma$ -ATP for the corresponding slices are shown in the subsequent rows and representative voxel spectra in the last two rows. Results from the Fourier reconstruction were on the left while those from the proposed method on the right. Apparent SNR enhancement in both the metabolite maps and spectra can be observed for the proposed method. The visual differences in the FOVs between the anatomical and  $^{31}P$  images were due to the different sensitivity profiles for the  $^1H$  and  $^{31}P$  channels.

proper sampling time vectors during data generation and reconstruction. Spatiospectral reconstructions obtained by the conventional Fourier reconstruction and the proposed method are shown in Fig. 8. The Fourier reconstruction produced very noisy results; the proposed method achieved significantly improved SNR and spectral quality, revealing spatiospectral features previously concealed by noise.

Another data set was acquired using a double-tuned volume  $^{31}P$ - $^1H$  coil on the same 7T system with matched TR, TE and spectral BW. The FOV was modified to  $220 \times 220 \times 150$  mm<sup>3</sup> and matrix size to  $32 \times 32 \times 8$  (total acquisition time of about 25 min). The spatiospectral reconstruction results from this data are shown in Fig. 9 and Fig. S3 (Supplementary Materials). Figure 9 compares the results from the SPICE (subspace) reconstruction and the proposed method. As can be seen, the proposed method produced metabolite maps with higher SNR and qualitatively better spectra. Note that while the volume coil has larger and more uniform coverage, the raw data from this experiment were considerably noisier than the previous case. However, the proposed method was still able to achieve significant SNR improvement. A more comprehensive comparison of the noisy Fourier reconstruction, the SPICE reconstruction and the proposed method is shown in Fig. S3.

### C. Special Case: Single-Voxel Spectroscopy Denoising

As discussed in Section III.D, the proposed method can also be applied to denoise single-voxel spectroscopy (SVS) data by solving the optimization problem in Eq. (16), which allows the incorporation of data-driven priors to improve the SNR of

general single-voxel acquisitions. Figure 10 shows an example of such results to demonstrate this capability. Specifically, a noisy brain  $^{31}P$  spectrum was acquired in vivo and denoised by the proposed method. In addition, a “reference” spectrum was generated by spectral fitting of this data (the residual was inspected to ensure high-quality fitting) and compared to the denoising result. As can be seen, the proposed method achieved effective noise reduction with excellent preservation of spectral features as compared to the fitted spectrum. The peaks of PCr,  $\alpha$ -ATP,  $\gamma$ -ATP, and even the weak signals of  $\beta$ -ATP and Pi were well recovered, although the performance for some of the weaker and overlapping peaks (e.g., GPC and GPE) can be further improved. A more comprehensive and quantitative evaluation of the proposed method for SVS studies will be done in future research.

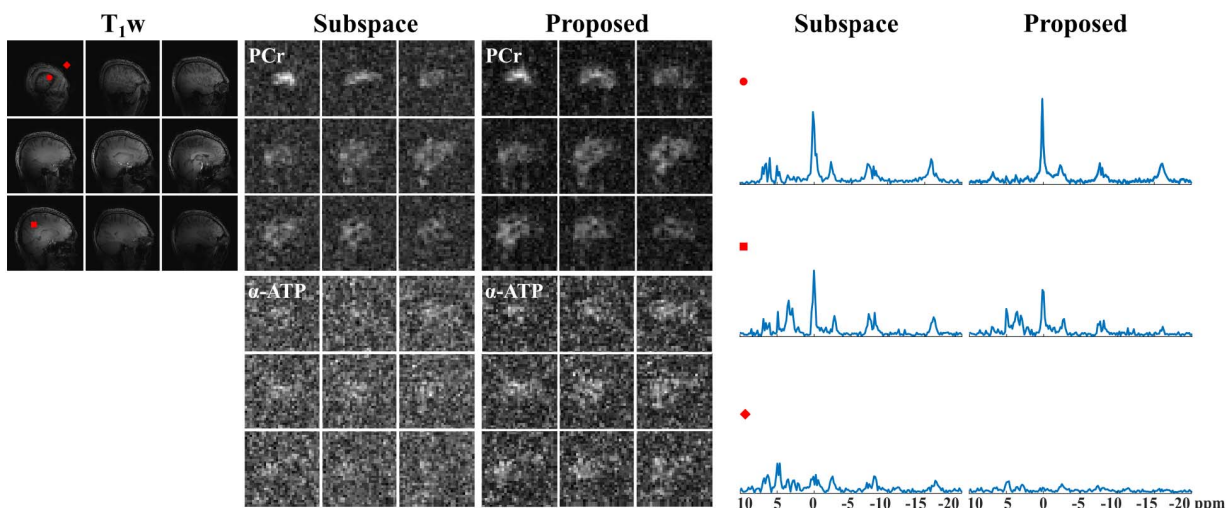
## V. DISCUSSION

The proposed method has several key differences compared to other learning-based denoising/reconstruction strategies. First of all, by not directly learning the inverse transform that maps the noisy, artifact containing data to the desired signals or parameters, we simplified the learning problem and allowed the constructed DAE to focus on learning to extract the physiologically meaningful/molecule-specific low-dimensional features of spectroscopic data (for a particular field strength) instead of the nuisance noise and artifacts. Second, the proposed formulation represents a different way to incorporate deep learning into the imaging process that effectively combines the physics-based data acquisition model and learned signal model. It offers higher flexibility and can work with different noise levels, acquisition parameters and other spatiospectral constraints without having to retrain the network. Third, our algorithm directly solves the resulting optimization problem as opposed to previous works which mapped an iterative process to a cascade of networks that approximately solve a similar regularized reconstruction formulation (e.g., [56], [57]). These “unrolling” based methods can be efficient but still fall into the category of directly learning the entire inverse mapping and can be sensitive to SNR levels and acquisition designs.

An important issue with the proposed method is the selection of regularization parameters. As shown by our comparison of reconstructions obtained with different combinations of  $\lambda_1$  and  $\lambda_2$  (Figs. 7 and S2), the denoising performance remains robust to a large range of parameter values. Furthermore, the selection of  $\lambda_1$  can be done by performing single voxel denoising (due to the nature of this constraint), which according to our simulation and experiment studies serves as a good initial guess for the overall reconstruction problem. The other parameter,  $\lambda_2$ , can be chosen based on discrepancy principle and then fine tuned together with  $\lambda_1$  by visual inspection of spatiospectral reconstruction quality. Meanwhile, a careful choice of  $\lambda_2$  is not required for effective denoising as the learned model is the main contributor, as shown by the result with  $\lambda_2 = 0$  in Fig. S2. More sophisticated parameter selection strategies can also be explored in future research.

Several other issues are worth investigating in future research. In particular, the network for learning the nonlinear

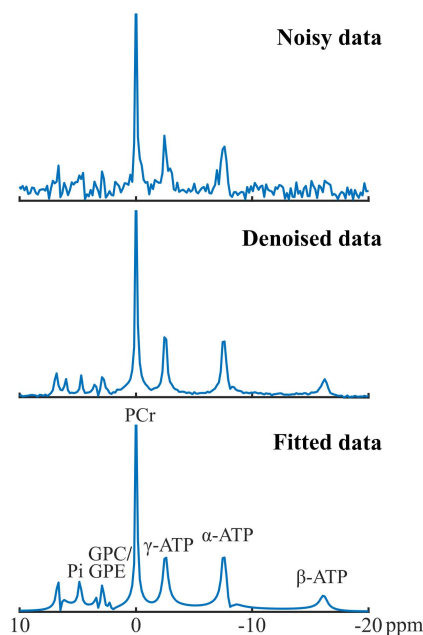




**Fig. 9.** Results from the data acquired using a volume coil. The maps of PCr and  $\alpha$ -ATP are shown for the SPICE (Subspace) and the proposed method (Proposed), along with spatially-resolved spectra from voxels marked by red symbols in the  $T_1$ -weighted anatomical images shown on the most left ( $T_1$ w). As shown by the metabolite maps and spectra, the proposed method yields a qualitatively better reconstruction. The spectra for a voxel from outside the brain (diamond-shaped symbol) exhibit less noise-induced bias for the proposed method than the subspace reconstruction, implying less modeling errors.

low-dimensional representation can be further optimized. For example, structures for better handling complex-valued data, different layer designs (convolutional versus fully-connected), asymmetric encoder and decoder networks, and choices of activation functions may be studied. Characterizations of the connections between network complexity, dimensionality of the nonlinear features, and the complexity of the spectral functions will be topics of both theoretical and practical interest. High-quality in vivo data can be acquired to improve the estimation of spectral parameter distributions or directly augment the neural network training data, which should make the learned model more adaptive to experimental variations not captured by synthesized data alone. One unique advantage of the proposed nonlinear model is that the dimensionality is not affected by the range of the spectral parameters, thus more generalizable to patient populations with larger parameter changes. However, if there exist novel molecular features in certain patient groups that are not present in the training data, patient-specific data need to be collected to adapt the learned model to capture such features. A more general lineshape distortion function  $h(\mathbf{r}, t)$  can be considered, and incorporated into the system encoding operator and separated from the ideal spectral signals. For the proposed algorithm, the most computationally intensive step is solving the subproblem in Eq. (9), which however is highly parallelizable and should greatly benefit from translating the current implementation to parallel computing platforms.

While we demonstrated the capability of the proposed approach using  $^{31}\text{P}$ -MRSI data, application to other nuclei is possible. Extending the proposed method to  $^1\text{H}$ -MRSI will require addressing a number of important issues such as residual nuisance water/lipid signals, presence of macromolecule baseline in short-TE data, more spectral features, and potentially stronger lineshape distortion due to larger intravoxel  $B_0$  inhomogeneity. These issues, although nontrivial, we believe



**Fig. 10.** Application of the proposed method to denoise a single-voxel  $^{31}\text{P}$  spectrum. The noisy data and denoised spectra are shown in the first and second rows, respectively, with comparison to a noiseless spectrum (last row) obtained by performing a spectral fitting of the noisy data.

are addressable. For example, a separate nonlinear low-dimensional model can be learned for the macromolecule signals by exploiting their unique spectral structures and lineshapes [34], [58]. This model can be incorporated into a modified reconstruction formulation that imposes the learned models for metabolites and macromolecules separately.

It is also worth mentioning that as the proposed method integrates the data acquisition model and the learned nonlinear model through a regularization formalism, it can be

readily extended to incorporate other spatio-spectral constraints (e.g., spatio-spectral sparsity using non-quadratic regularizations) and even potentially learned spatial priors. The improved SNR can allow faster speeds (e.g., with less number of averages and/or shorter TRs) and higher resolution acquisitions. Although we only demonstrated denoising reconstruction, the proposed formulation can be extended to other scenarios such as sparse sampling in the  $(k,t)$ -space with modifications of the sampling operator. These directions are currently being pursued and could lead to new opportunities in synergizing physics-based modeling and machine learning to push the resolution and SNR limits of in vivo MRSI.

## VI. CONCLUSION

We have presented a new method to model MRSI data by learning a low-dimensional nonlinear representation and use the learned model for spatio-spectral reconstruction. The model was learned using a deep autoencoder based neural network which can accurately capture the inherent low-dimensional features of high-dimensional spectral variations and enable effective dimensionality reduction. The proposed constrained reconstruction method incorporates the learned model through a regularization formalism which was solved by an efficient ADMM-based algorithm. Significantly improved spatio-spectral reconstruction over conventional methods achieved by the proposed method has been demonstrated using both simulation and experimental data. The proposed method represents a new means to incorporate deep learning into the imaging process and may be extended in various ways including improved network designs, training data generation, choices of error metrics and combinations with other spatio-spectral constraints.

## ACKNOWLEDGMENT

The authors would like to acknowledge Drs. Wei Chen and Xiao-Hong Zhu for providing the  $^{31}\text{P}$ -MRSI data, and Dr. Dinesh Deelchand for sharing the simulated metabolite basis.

## REFERENCES

- [1] T. R. Brown, B. M. Kincaid, and K. Ugurbil, "NMR chemical shift imaging in three dimensions," *Proc. Nat. Acad. Sci. USA*, vol. 79, no. 11, pp. 3523–3526, 1982.
- [2] R. A. De Graaf, *Vivo NMR Spectroscopy: Principles and Techniques*. Hoboken, NJ, USA: Wiley, 2007.
- [3] S. Posse, R. Otazo, S. R. Dager, and J. Alger, "MR spectroscopic imaging: Principles and recent advances," *J. Magn. Reson. Imag.*, vol. 37, no. 3, pp. 1301–1325, 2013.
- [4] S. Hu *et al.*, "Compressed sensing for resolution enhancement of hyperpolarized  $^{13}\text{C}$  flyback 3D-MRSI," *J. Magn. Reson.*, vol. 192, pp. 258–264, Jun. 2008.
- [5] R. Eslami and M. Jacob, "Robust reconstruction of MRSI data using a sparse spectral model and high resolution MRI priors," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1297–1309, Jun. 2010.
- [6] P. Cao and E. X. Wu, "Accelerating phase-encoded proton MR spectroscopic imaging by compressed sensing," *J. Magn. Reson. Imag.*, vol. 41, no. 2, pp. 487–495, 2014.
- [7] R. Otazo, D. K. Sodickson, A. Yoshimoto, and S. Posse, "Accelerated proton echo-planar spectroscopic imaging using parallel imaging and compressed sensing," in *Proc. ISMRM*, Honolulu, HI, USA, 2009, p. 331.
- [8] I. Chatnuntawech, B. Gagoski, B. Bilgic, S. F. Cauley, K. Setsompop, and E. Adalsteinsson, "Accelerated  $^1\text{H}$  MRSI using randomly under-sampled spiral-based  $k$ -space trajectories," *Magn. Reson. Med.*, vol. 74, no. 1, pp. 13–24, 2014.
- [9] B. L. Burns, N. E. Wilson, and M. A. Thomas, "Group sparse reconstruction of multi-dimensional spectroscopic imaging in human brain in vivo," *Algorithms*, vol. 7, no. 3, pp. 276–294, 2014.
- [10] R. Nagarajan *et al.*, "Accelerated echo planar  $J$ -resolved spectroscopic imaging in prostate cancer: A pilot validation of non-linear reconstruction using total variation and maximum entropy," *NMR Biomed.*, vol. 28, no. 11, pp. 1366–1373, 2015.
- [11] J. Kasten, F. Lazeyras, and D. Van De Ville, "Data-driven MRSI spectral localization via low-rank component analysis," *IEEE Trans. Med. Imag.*, vol. 32, no. 10, pp. 1853–1863, Oct. 2013.
- [12] H. M. Nguyen, X. Peng, M. N. Do, and Z.-P. Liang, "Denoising MR spectroscopic imaging data with low-rank approximations," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 78–89, Jan. 2013.
- [13] Y. Liu, C. Ma, B. A. Clifford, F. Lam, C. L. Johnson, and Z.-P. Liang, "Improved low-rank filtering of magnetic resonance spectroscopic imaging data corrupted by noise and  $B_0$  field inhomogeneity," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 841–849, Apr. 2016.
- [14] Y. Zhang, R. E. Gabr, M. Schär, R. G. Weiss, and P. A. Bottomley, "Magnetic resonance spectroscopy with linear algebraic modeling (SLAM) for higher speed and sensitivity," *J. Magn. Reson.*, vol. 218, pp. 66–76, May 2012.
- [15] X. Qu, M. Mayzel, J.-F. Cai, Z. Chen, and V. Orekhov, "Accelerated NMR spectroscopy with low-rank reconstruction," *Angew. Chem. Int. Ed.*, vol. 54, no. 3, pp. 852–854, Jan. 2015.
- [16] I. Bhattacharya and M. Jacob, "Compartmentalized low-rank recovery for high-resolution lipid unsuppressed MRSI," *Magn. Reson. Med.*, vol. 78, no. 4, pp. 1267–1280, 2017.
- [17] A. Klauser *et al.*, "Fast high-resolution brain metabolite mapping on a clinical 3T MRI by accelerated H-FID-MRSI and low-rank constrained reconstruction," *Magn. Reson. Med.*, vol. 81, no. 5, pp. 2841–2857, 2019.
- [18] M. Jacob, X. Zhu, A. Ebel, N. Schuff, and Z. P. Liang, "Improved model-based magnetic resonance spectroscopic imaging," *IEEE Trans. Med. Imag.*, vol. 26, no. 10, pp. 1305–1318, Oct. 2007.
- [19] Y. Bao and A. A. Maudsley, "Improved reconstruction for MR spectroscopic imaging," *IEEE Trans. Med. Imag.*, vol. 26, no. 5, pp. 686–695, May 2007.
- [20] J. P. Haldar, D. Hernando, S.-K. Song, and Z.-P. Liang, "Anatomically constrained reconstruction from noisy data," *Magn. Reson. Med.*, vol. 59, pp. 810–818, Apr. 2008.
- [21] J. Kornak, K. Young, B. J. Soher, and A. A. Maudsley, "Bayesian  $k$ -space-time reconstruction of MR spectroscopic imaging for enhanced resolution," *IEEE Trans. Med. Imag.*, vol. 29, no. 7, pp. 1333–1350, Jul. 2010.
- [22] F. Lam and Z.-P. Liang, "A subspace approach to high-resolution spectroscopic imaging," *Magn. Reson. Med.*, vol. 71, no. 4, pp. 1349–1357, 2014.
- [23] F. Lam, C. Ma, B. Clifford, C. L. Johnson, and Z.-P. Liang, "High-resolution  $^1\text{H}$ -MRSI of the brain using SPICE: Data acquisition and image reconstruction," *Magn. Reson. Med.*, vol. 76, pp. 1059–1070, Oct. 2016.
- [24] C. Ma, F. Lam, Q. Ning, C. L. Johnson, and Z.-P. Liang, "High-resolution  $^1\text{H}$ -MRSI of the brain using short-TE SPICE," *Magn. Reson. Med.*, vol. 77, pp. 467–479, Feb. 2017.
- [25] P. Tiwari, M. Rosen, and A. Madabhushi, "Consensus-locally linear embedding (C-LLE): Application to prostate cancer detection on magnetic resonance spectroscopy," in *Proc. MICCAI*. Berlin, Germany: Springer, 2008, pp. 330–338.
- [26] G. Yang, F. Raschke, T. R. Barrick, and F. A. Howe, "Manifold learning in MR spectroscopy using nonlinear dimensionality reduction and unsupervised clustering," *Magn. Reson. Med.*, vol. 74, pp. 868–878, Sep. 2015.
- [27] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [28] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, no. 12, pp. 3371–3408, Dec. 2010.
- [29] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

- [30] N. Hatami, M. Sdika, and H. Ratiney, "Magnetic resonance spectroscopy quantification using deep learning," 2018, *arXiv:1806.07237*. [Online]. Available: <https://arxiv.org/abs/1806.07237>
- [31] D. Das, E. Coello, R. F. Schulte, and B. H. Menze, "Quantification of metabolites in magnetic resonance spectroscopic imaging using machine learning," in *Proc. MICCAI*, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. Collins, and S. Duchesne, Eds. Cham, Switzerland: Springer, 2017, pp. 462–470.
- [32] S. S. Gurbani, S. Sheriff, A. A. Maudsley, H. Shim, and L. A. Cooper, "Incorporation of a spectral model in a convolutional neural network for accelerated spectral fitting," *Magn. Reson. Med.*, vol. 81, no. 2, pp. 3346–3357, 2019.
- [33] S. P. Kyathanahally, A. Döring, and R. Kreis, "Deep learning approaches for detection and removal of ghosting artifacts in MR spectroscopy," *Magn. Reson. Med.*, vol. 80, pp. 851–863, Sep. 2018.
- [34] H. H. Lee and H. Kim, "Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain," *Magn. Reson. Med.*, vol. 82, pp. 33–48, Jul. 2019.
- [35] C. S. Bolliger, C. Boesch, and R. Kreis, "On the use of Cramer–Rao minimum variance bounds for the design of magnetic resonance spectroscopy experiments," *NeuroImage*, vol. 83, pp. 1031–1040, Dec. 2013.
- [36] G. Peyré, "Manifold models for signals and images," *Comput. Vis. Image Understand.*, vol. 113, no. 2, pp. 249–260, 2009.
- [37] U. Nakarmi, Y. Wang, J. Lyu, D. Liang, and L. Ying, "A kernel-based low-rank (KLR) model for low-dimensional manifold recovery in highly accelerated dynamic MRI," *IEEE Trans. Med. Imag.*, vol. 36, no. 11, pp. 2297–2307, Nov. 2017.
- [38] Y. Li, F. Lam, B. Clifford, and Z.-P. Liang, "A subspace approach to spectral quantification for MR spectroscopic imaging," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 10, pp. 2486–2489, Oct. 2017.
- [39] P. Tiwari, M. Rosen, and A. Madabhushi, "A hierarchical spectral clustering and nonlinear dimensionality reduction scheme for detection of prostate cancer from magnetic resonance spectroscopy (MRS)," *Med. Phys.*, vol. 36, pp. 3927–3939, Sep. 2009.
- [40] D. DeMers and G. W. Cottrell, "Non-linear dimensionality reduction," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 5, 1993, pp. 580–587.
- [41] R. Hecht-Nielsen, "Replicator neural networks for universal optimal source coding," *Science*, vol. 269, no. 5232, pp. 1860–1863, 1995.
- [42] M. Ranzato, C. Poultney, S. Chopra, and Y. LeCun, "Efficient learning of sparse representations with an energy-based model," in *Proc. Adv. Neural Inf. Process. Syst.*, B. Schölkopf, J. C. Platt, and T. Hoffman, Eds. Cambridge, MA, USA: MIT Press, 2007, pp. 1137–1144.
- [43] D. Wu, K. Kim, G. El Fakhri, and Q. Li, "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2479–2486, Dec. 2017.
- [44] J. Mehta and A. Majumdar, "RODEO: Robust DE-aliasing autoencoder for real-time medical image reconstruction," *Pattern Recognit.*, vol. 63, pp. 499–510, Mar. 2017.
- [45] D. Stefan *et al.*, "Quantitation of magnetic resonance spectroscopy signals: The jMRUI software package," *Meas. Sci. Technol.*, vol. 20, no. 10, 2009, Art. no. 104035.
- [46] D. K. Deelchand, T.-M. Nguyen, X.-H. Zhu, F. Mochel, and P.-G. Henry, "Quantification of *in vivo*  $^{31}\text{P}$  NMR brain spectra using LCMModel," *NMR Biomed.*, vol. 28, pp. 633–641, Jun. 2015.
- [47] J. Yang and Y. Zhang, "Alternating direction algorithms for  $\ell_1$ -problems in compressive sensing," *SIAM J. Sci. Comput.*, vol. 33, no. 1, pp. 250–278, 2011.
- [48] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Program.*, vol. 45, no. 1, pp. 503–528, 1989.
- [49] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.
- [50] S. Ramani and J. A. Fessler, "Parallel MR image reconstruction using augmented Lagrangian methods," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 694–706, Mar. 2011.
- [51] H. F. Cancino-De-Greiff, R. Ramos-Garcia, and J. V. Lorenzo-Ginori, "Signal de-noising in magnetic resonance spectroscopy using wavelet transforms," *Concepts Magn. Reson.*, vol. 14, no. 6, pp. 388–401, 2002.
- [52] J. Ren, A. D. Sherry, and C. R. Malloy, " $^{31}\text{P}$ -MRS of healthy human brain: ATP synthesis, metabolite concentrations, pH, and  $T_1$  relaxation times," *NMR Biomed.*, vol. 28, pp. 1455–1462, Nov. 2015.
- [53] M. Abadi *et al.* *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Accessed: Mar. 2016. [Online]. Available: <http://tensorflow.org/>
- [54] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [55] H. Lei, X.-H. Zhu, X.-L. Zhang, K. Ugurbil, and W. Chen, "*In vivo*  $^{31}\text{P}$  magnetic resonance spectroscopy of human brain at 7 T: An initial experience," *Magn. Reson. Med.*, vol. 49, pp. 199–205, Feb. 2003.
- [56] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic mr image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 491–503, Feb. 2017.
- [57] H. K. Aggarwal, M. P. Mani, and M. Jacob, "Modl: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019.
- [58] K. L. Behar, D. L. Rothman, D. D. Spencer, and O. A. C. Petroff, "Analysis of macromolecule resonances in  $^1\text{H}$  NMR spectra of human brain," *Magn. Reson. Med.*, vol. 32, pp. 294–302, Sep. 1994.