SNR Enhancement for Multi-TE MRSI Using Joint Low-Dimensional Model and Spatial Constraints

Yahang Li, *Student Member, IEEE*, Zepeng Wang, *Student Member, IEEE*, and Fan Lam, *Senior Member, IEEE*

***Abstract*— We present a novel method to enhance the SNR for multi-TE MR spectroscopic imaging (MRSI) data by integrating learned nonlinear low-dimensional model and spatial constraints. A deep complex convolutional autoencoder (DCCAE) was developed to learn a nonlinear low-dimensional representation of the high-dimensional multi-TE ^{1}H spectroscopy signals. The learned model significantly reduces the data dimension thus serving as an effective constraint for noise reduction. A reconstruction formulation was proposed to integrate the spatiospectral encoding model, the learned model, and a spatial constraint for an SNR-enhancing reconstruction from multi-TE data. The proposed method has been evaluated using both numerical simulations and in vivo brain MRSI experiments. The superior denoising performance of the proposed over alternative methods was demonstrated, both qualitatively and quantitatively. In vivo multi-TE data was used to assess the improved metabolite quantification reproducibility and accuracy achieved by the proposed method. We expect the proposed SNR-enhancing reconstruction to enable faster and/or higher-resolution multi-TE ^{1}H-MRSI of the brain, potentially useful for various clinical applications.**

***Index Terms*— multi-TE ^{1}H-MRSI, denoising, regularized reconstruction, deep learning, complex convolutional neural network, low-dimensional modeling**

I. Introduction

MULTI-TE MRSI (acquiring spatially resolved MR spectroscopy data at multiple echo-times) offers stronger molecular imaging capabilities than standard single-TE MRSI, such as the ability to better resolve molecules with strong spectral overlaps and map metabolite relaxation parameters [1]–[5]. However, the need to acquire spectroscopy data at different TEs inevitably increases the encoding dimensions, thus leading to more limited trade-offs between speed, resolution, and SNR. The SNR challenge is even more serious for data acquired at longer TEs (to encode J-coupling evolution or T_2 decay), which can negatively affect subsequent quantitative analysis. As an example, a typical multi-TE MRSI experiment usually takes around 20 minutes to achieve a good SNR for a centimeter-level resolution [6]. Therefore, the ability to enhance the SNR for multi-TE MRSI will be particularly useful for enabling faster and/or higher-resolution acquisitions and making this technology more clinically relevant.

Denoising is becoming a common practice to address SNR issues for MRSI experiments. The simple and efficient spatial and/or spectral invariant filters, e.g., spatial hamming, temporal exponential, or Gaussian filters, can enhance SNR but at the expense of spatiospectral resolution. More advanced signal processing methods that exploit the unique spectroscopic signal characteristics such as wavelet domain sparsity [7] and linear predictability have been proposed [8]–[10]. These methods can reduce noise levels while better preserving molecular features of interest, but offer limited performance when the SNR is low. Recognizing the readily available anatomical information in any MRSI experiments, a number of methods that incorporate spatial constraints derived from anatomical images [11]–[13] or combine spatial and spectral constraints [14], [15] for improved reconstruction of noisy MRSI data have also been described. While these methods utilized the spatial and spectral domain constraints in an independent fashion, more recent approaches motivated by the fact that high-dimensional MRSI data can be well approximated by low-dimensional partial separability (PS) models due to the inherent strong spatiospectral correlation have produced impressive SNR and/or resolution improvements [9], [16], [17]. The subspace or low-rank models derived from PS have also been integrated with spatial constraints to offer further enhanced performance [18]–[22]. All the methods mentioned above have been extensively investigated for single-TE MRSI.

One straightforward way to adapt the methods mentioned above to multi-TE MRSI is applying them to individual TEs separately. However, such an approach does not effectively utilize spatio-spectral-TE correlations embedded in the high-dimensional multi-TE MRSI data for optimized SNR enhancement. To this end, several methods were developed to exploit the correlations among multiple dimensions for accelerating

This work is supported in part by NSF-CBET 1944249 and NIH-NIBIB 1R21EB029076A. This work was conducted in part at the Biomedical Imaging Center of the Beckman Institute for Advanced Science and Technology at the University of Illinois Urbana-Champaign (UIUC-BI-BIC). Copyright (c) 2021 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

Yahang Li and Zepeng Wang are with the Department of Bioengineering, University of Illinois Urbana-Champaign, Urbana, IL, 61801 USA

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

multi-TE/J-resolved MRSI without a strong SNR penalty, making use of either sparsity in the two-dimensional spectra [23], [24] or low-rank tensor constraints derived from a higher-order PS model [25], [26]. As these constraints typically require many encodings along the TE/J-resolved dimension to be effective, limiting the flexibility in experimental design [27], augmented subspace models that can work with a smaller number of TEs have also been proposed recently [28], [29].

While the subspace model is a powerful tool to achieve better SNRs for MRSI, they still rely on linear dimensionality reduction of high-dimensional data, which can be less efficient (requiring higher model orders) when there are more complicated spectral variations due to either physiological or practical experimental conditions [30], [31]. Nonlinear low-dimensional models have been shown to offer better dimensionality reduction thus SNR improvement [30], [32]–[34]. With the recent advancements in deep representation learning, it is also possible to pre-learn such a nonlinear low-dimensional model from training data and use it for constrained MRSI reconstruction [35]–[37].

We present here a novel method to improve the SNR of multi-TE MRSI by learning a nonlinear low-dimensional model for multi-TE ^{1}H spectroscopy data and using this learned model for spatiospectral reconstruction. Specifically, we proposed to use a deep complex convolutional autoencoder (DCCAE) [38], [39] to learn a nonlinear representation of multi-TE spectra with a stronger dimensionality reduction capability than existing linear subspace models, offering higher potential for noise reduction. The model learning process can be readily adapted to any type of signal excitation strategy. A regularized reconstruction formulation was constructed to integrate the acquisition model, learned model, and a spatial constraint for SNR-enhancing reconstruction from noisy multi-TE ^{1}H-MRSI data. The effectiveness of the proposed method has been evaluated using simulated and experimental data, demonstrating superior denoising performance over existing methods and better metabolite quantification. The reconstruction formulation, modeling learning process, numerical algorithm, and experimental studies are described below in detail.

II. Theory

The data acquisition process for a multi-TE MRSI experiment can be described as

$$d(\mathbf{k},t,TE) = \int \rho(\mathbf{r},t,TE)e^{-i2\pi\delta f(\mathbf{r})t}e^{-i2\pi\mathbf{kr}}d\mathbf{r} + n(\mathbf{k},t,TE), \tag{1}$$

where $\rho(\mathbf{r},t,TE)$ denotes the TE-dependent spatiotemporal function of interest that contains various molecular signal components of interest, t denotes the free induction decay (FID) time, TE the echo time (TE) at which the data is acquired, $d(\mathbf{k},t,TE)$ represents the measured (k,t)-space data, $\delta f(\mathbf{r})$ corresponds to the B_0 field inhomogeneity distribution, and $n(\mathbf{k},t,TE)$ is the measurement noise (commonly assumed to be complex white Gaussian thermal noise [1]). The goal of an SNR-enhancing or denoising reconstruction is to recover the object function of interest ρ from the noisy measurement d with as high an SNR as possible while protecting the biologically or clinically relevant features.

After a proper discretization, the image function ρ can be represented by a tensor structure $\boldsymbol{\rho} \in \mathbb{C}^{N\times D\times M}$, with N being the number of spatial voxels, D the number of FID sample points, and M the number of TEs. A denoising reconstruction can then be generally formulated as the following optimization problem

$$\hat{\boldsymbol{\rho}} = \arg\min_{\boldsymbol{\rho}} \|\mathbf{d} - \mathcal{F}_\Omega\{\mathbf{B}\odot\boldsymbol{\rho}\}\|_2^2 + \lambda\mathcal{R}(\boldsymbol{\rho}), \tag{2}$$

where the first term imposes data consistency and $\mathcal{R}(\cdot)$ imposes various types of model constraints based on prior information about $\boldsymbol{\rho}$ (e.g., sparsity, low-rankness, and smoothness, etc). $\mathbf{B}$ captures the effects of B_0 field inhomogeneity (voxel dependent linear phases along t), $\odot$ denotes a point-wise multiplication, and $\mathcal{F}$ is a multi-TE spatiotemporal encoding operator with a (k,t,TE)-space sampling pattern Ω. λ is the regularization parameter. A key difference for various denoising methods lies in the specific choice of $\mathcal{R}(\boldsymbol{\rho})$ and algorithm to solve Eq. (2).

We propose here a new $\mathcal{R}(\cdot)$. Specifically, we set out to learn a nonlinear low-dimensional model that can accurately represent multi-TE ^{1}H spectroscopic signals and enable their effective separation from noise (high-dimensional). To this end, we designed a deep complex convolution autoencoder (DCCAE) and trained the network to discover low-dimensional features that can faithfully reconstruct the high-dimensional multi-TE FIDs. With this learned model, denoted as $\mathcal{N}_\theta(\mathbf{X})$ where $\mathbf{X}$ denotes a tensor representation of the multi-TE spatiotemporal function of interest and θ contains the network parameters, we propose to formulate the reconstruction problem as follows:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{d} - \mathcal{F}_\Omega\{\mathbf{B}\odot\mathbf{X}\}\|_2^2 + \lambda_1\sum_{n=1}^{N}\|\mathcal{N}_\theta(\mathbf{X}_n) - \mathbf{X}_n\|_F^2 + \lambda_2\|\mathbf{D}_w\mathbf{X}\|_F^2. \tag{3}$$

The dimension for $\mathbf{X}$ is $N\times D\times M$, with N denoting the number of voxels, D the number of FID points for each voxel, and M the number of TEs acquired. $\mathbf{X}_n$ is the multi-TE FIDs for the nth voxel. The first regularization term, with parameter λ_1, imposes the constraint that multi-TE FIDs reside on or close to a low-dimensional manifold by penalizing the difference between the signal and its "projection" onto the manifold through $\mathcal{N}_\theta(\cdot)$. The second regularization term, with parameter λ_2, is an edge-preserving spatial constraint term with $\mathbf{D}_w$ a weighted finite-difference operator [12]. $\|\cdot\|_F$ represents the Frobenius norm. The methods for training $\mathcal{N}_\theta(\cdot)$, numerically solving Eq. (3) and experimentally evaluating the proposed method are described in the sections below.

III. Methods

A. Low-Dimensional Representation Learning for Multi-TE MRSI Data

While efforts have been made on learning efficient low-dimensional representation for high-dimensional FIDs or spectra, e.g., [30], [31], the main challenges for multi-TE MRSI

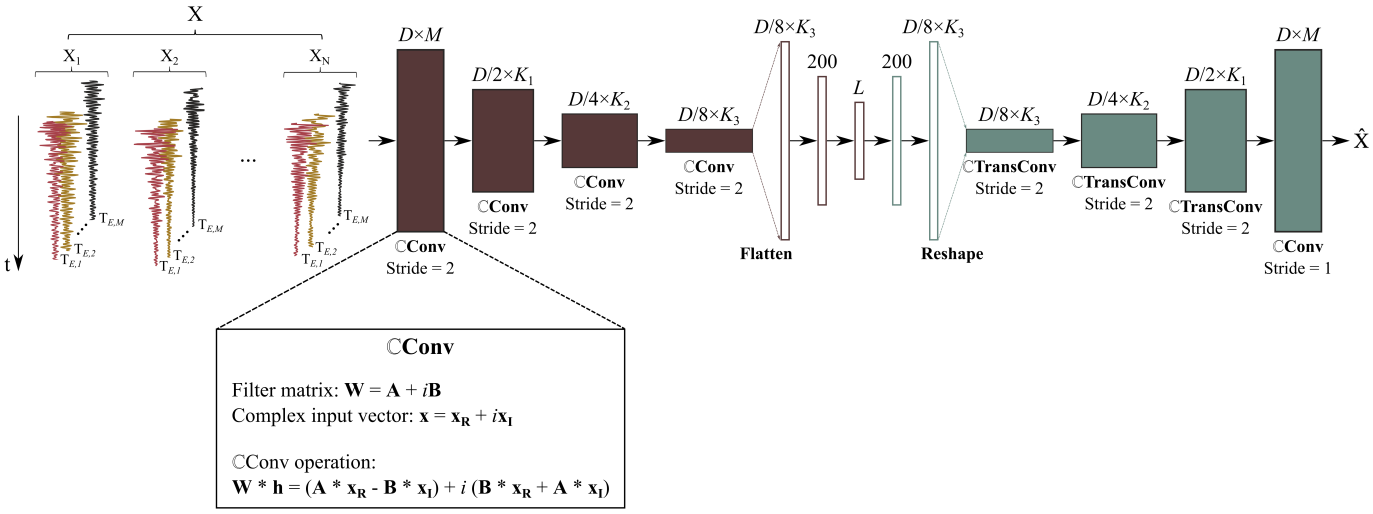

Fig. 1. Illustration of the training strategy of the proposed DCCAE. Here $\mathbf{X}$ represents a collection of multi-TE FID training data, with $\boldsymbol{D}$ denotes the data length and $\boldsymbol{M}$ is the number of TEs. Complex convolutional layers were adapted, where the complex multi-TE MRSI training data was directly used as input with different TEs treated as different channels, followed by complex fully-connected layers, to encode data $\mathbf{X}$ into $\mathbf{L}$-dimensional features. The data dimensions were reduced by half in each complex convolutional block, while the channel dimension $\mathbf{K}$ increased by a small amount (details in texts). The mathematical formulation of the training strategy is provided in Eq. (4).

include the higher dimensions and the need to effectively exploit the relationships across TEs. Methods that treat the TE dimension as separate input channels [40] have been used for image reconstruction recently. Following a similar spirit, we propose here a multi-convolution-channel DCCAE design. Figure 1 illustrates the network structure and training strategy. Specifically, this network has several new features compared to prior works: (1) Automatic feature extraction and correlation exploitation across TEs achieved by treating data acquired at different TEs as separate input channels to the network and using (t, TE) convolutional layers; (2) Combination of convolutional layers and fully connected layers for further low-dimensional feature extraction; and (3) Use of complex-valued units and activation functions to handle the multi-TE FIDs directly (complex operations defined in Fig. 1) [38], [39], instead of either using only the real parts or treating real and imaginary parts separately as in most existing spectroscopy deep learning methods. The training was done by solving:

$$\{\hat{\boldsymbol{\theta}}\} = \arg\min_{\boldsymbol{\theta}} \frac{1}{J}\sum_{j=1}^{J} \epsilon(\mathbf{x}^j, \mathcal{N}(\mathbf{x}^j;\boldsymbol{\theta})) + \lambda \frac{1}{K}\sum_{k=1}^{K} \epsilon(\mathbf{x}^k, \mathcal{N}(\mathbf{x}^k + n^k;\boldsymbol{\theta})), \tag{4}$$

where $\mathbf{x}^j$ and $\mathbf{x}^k$ are training samples, with each $\mathbf{x}^j$, $\mathbf{x}^k \in \mathbb{C}^{D\times M}$ being a multi-TE FID. $\mathcal{N}(\cdot;\boldsymbol{\theta})$ represents the network parameterized by $\boldsymbol{\theta}$, n^k denotes white Gaussian noise vectors with varying SNRs (defined w.r.t. the highest spectra peak from a representative testing spectra) and ϵ is the loss function. Specifically, the first term enforces the NN to learn a low-dimensional representation that can accurately recover the high-dimensional data ($\mathbf{x}^j$), while the second term is a denoising regularization with regularization parameter λ to increase the generalization of the learned model [36]. We found that this specific design with convolutional layers followed by fully-connected layers with "bottle-neck" structures and the denoising regularization can successfully encode the high-dimensional multi-TE data into a set of L-dimensional features, from which the original data can be recovered accurately. We will refer to L as the model order below.

B. Network Training

All the training and testing data for the representation network in Fig. 1 were synthesized using the following multi-TE ^{1}H spectroscopic signal model [31], [41], [42]:

$$x(t,TE) = \sum_{p=1}^{P_{Met}} c_p e^{i(\alpha_{0,TE}+\alpha_p)}\phi_p(t,TE)e^{-TE/T_{2,p}}e^{-t/T^*_{2,p}+i2\pi\delta f_p t}h(t) + \sum_{q=1}^{Q_{MM}} b_q e^{i(\alpha_{0,TE}+\beta_q)}e^{-TE/T_{2,q}}e^{-t^2\frac{\pi^2 W_q^2}{4\ln(2)}+i2\pi\delta f_q t}, \tag{5}$$

where the first summation captures the metabolite components and the second summation captures the MMs. The variables c_p, $T_{2,p}$, $T^*_{2,p}$, and δf_p are concentrations, relaxation parameters, and additional frequency shifts for individual molecules, respectively, $\phi_p(t, TE)$ denotes the TE-dependent metabolites basis (simulated for the sequences used in NMRScopeB [43]), α_p and $\alpha_{0,TE}$ accounts for molecule-dependent and TE-dependent phases to make the mathematical representation more general, and $h(t)$ an additional Gaussian lineshape function. For the MM component, b_q, $T_{2,q}$, W_q, and δf_q represents the coefficients, relaxation parameters, Gaussian linewidths, and frequency shifts for individual MM groups. This parametric MM model has been demonstrated to be mathematically accurate (based on the residuals of fitting in vivo MM data) and very commonly used [31], [41], [44]. Although not offering strong physical interpretation, its mathematical sufficiency ensures that we can generate realistic spectra for

learning an accurate low-dimensional model for dimensionality reduction/denoising. Note that this model means that high-dimensional multi-TE FIDs are characterized by a small set of parameters, thus should reside in a low-dimensional manifold.

When generating the training data, Gaussian distributions were assumed for the spectral parameters (i.e., c_p, b_q, $T_{2,p/q}$, $T^*_{2,p}$, $\delta f_{p/q}$ and W_q) with means and standard deviations obtained from literature values and fitting results of our own high-SNR MRSI data previously acquired from healthy volunteers [22], [41], [45]–[47]. Randomly sampled values were drawn from these distributions to synthesize multi-TE FIDs using Eq. (5). The concentration values c_p were bounded between 0 to 2, with the mean NAA concentration set to 1. $T_{2,p}$ and $T^*_{2,p}$ were bounded between 5 and 400 ms, $T_{2,q}$ between 5 and 60 ms, the linewidth W_q values between 5 to 70 Hz, and concentration coefficient b_q were lower bounded by 0. Parameters generated with values outside these ranges were excluded. For $\alpha_{0,TE}$, we first generated the global zero-th order phase from the first TE α_{0,TE_1} with Gaussian distribution of zero mean and standard deviation of 20 degrees. Additional phase differences for the subsequent TEs were included with a mean of 3 and a standard deviation of 1 degree (based on the observation from in vivo data). The metabolite/MM-dependent phases α_p and β_q were generated from Gaussian distributions with zero mean and standard deviation of 5 degrees. The metabolites considered here are N-acetylaspartate (NAA), creatine (Cr), choline (Cho), glutamate (Glu), glutamine (Gln), Myo-inositol (mI), gamma-Aminobutyric acid (GABA), taurine (Tau), and lactate (Lac). For MMs, the commonly reported MM resonances with mean δf_q's at 0.9, 1.21, 1.38, 1.63, 2.01, 2.09, 2.25, 2.61, 2.96, 3.11, 3.67, 3.8, and 3.96 ppm were included. Note that here for the in vivo data, the first three MM resonances (0.9, 1.21, and 1.38 ppm) were omitted to accommodate the nuisance removal scheme, where a voxel-by-voxel HSVD-based removal step was used to remove any residual lipids after a union-of-subspace nuisance removal [48].

A total of 100,000 metabolite and 100,000 MM multi-TE FIDs were generated. MM data were scaled with a global scaling factor to mimic experimentally observed metabolite-to-MM ratios, and combined with metabolite spectra to generate 100,000 training samples. 80,000 were used for training and 20,000 for testing. Among those 80,000, 60,000 (J) were used as noiseless data, and noise was added to the remaining 20,000 (K) with SNR (defined w.r.t. the NAA peak in the mean spectra of all training data) uniformly distributed from 5 to 50 (chosen based typical SNRs observed for in vivo MRS/MRSI) for the denoising regularization term in Eq. (4).

C. Optimization Algorithm for Reconstruction Using the Learned Model

Directly solving Eq. (3) with the learned neural network based model constraint is challenging. Here, we adapted the algorithm design from [30], [31] to decouple the linear and nonlinear problems. Specifically, we introduced an auxiliary variable $\mathbf{S} = \mathbf{B} \odot \mathbf{X}$ and reformulated the problem as:

$$\begin{aligned}\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} &\|\mathbf{d} - \mathcal{F}_\Omega\{\mathbf{S}\}\|_2^2 \\ &+ \lambda_1 \sum_{n=1}^{N} \|\mathcal{N}(\mathbf{X}_n) - \mathbf{X}_n\|_F^2 + \lambda_2 \|\mathbf{D}_w \bar{\mathbf{B}} \odot \mathbf{S}\|_F^2 \\ s.t. \quad &\mathbf{B} \odot \mathbf{X} = \mathbf{S},\end{aligned} \tag{6}$$

where $\bar{\mathbf{B}}$ represents the element-wise conjugate of $\mathbf{B}$. The alternating direction method of multipliers (ADMM) was used to solve this equivalent problem [49]. More specifically, Eq. (6) can be solved by iteratively solving the following three subproblems (where i is the iteration index):

Subproblem (I): Update $\mathbf{X}$ with fixed $\mathbf{S}^{(i)}$ and $\mathbf{Z}^{(i)}$, where $\mathbf{Z}$ is the Lagrangian multiplier

$$\begin{aligned}\mathbf{X}^{(i+1)} = \arg\min_{\mathbf{X}} &\lambda_1 \sum_{n=1}^{N} \|\mathcal{N}(\mathbf{X}_n) - \mathbf{X}_n\|_F^2 \\ &+ \frac{\mu}{2} \left\| \mathbf{B} \odot \mathbf{X} - \mathbf{S}^{(i)} + \frac{\mathbf{Z}^{(i)}}{\mu} \right\|_F^2;\end{aligned} \tag{7}$$

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

$$\begin{aligned}\mathbf{S}^{(i+1)} = \arg\min_{\mathbf{S}} &\|\mathbf{d} - \mathcal{F}_\Omega\{\mathbf{S}\}\|_2^2 + \lambda_2 \|\mathbf{D}_w \bar{\mathbf{B}} \odot \mathbf{S}\|_F^2 \\ &+ \frac{\mu}{2} \left\| \mathbf{B} \odot \mathbf{X}^{(i+1)} - \mathbf{S} + \frac{\mathbf{Z}^{(i)}}{\mu} \right\|_F^2;\end{aligned} \tag{8}$$

Subproblem (III): Update $\mathbf{Z}$

$$\mathbf{Z}^{(i+1)} = \mathbf{Z}^{(i)} + \mu \left(\mathbf{B} \odot \mathbf{X}^{(i+1)} - \mathbf{S}^{(i+1)} \right). \tag{9}$$

Subproblem (I) captures the nonlinear network $\mathcal{N}(\cdot)$, and can be solved with general nonlinear optimizers (details for deriving the gradient for the complex-valued network can be found in the Appendix). Note that although Subproblem (I) is very high-dimensional, it can be decoupled into much smaller problems for individual voxels ($\mathbf{X}_n$) due to the separability of the Frobenius norm. We solved these individual-voxel problems in parallel using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm in this work (while more advanced solvers could be used), Subproblem (II) is equivalent to solving a set of linear equations with a quadratic regularization, which we solved using linear conjugate gradient [50]. The overall algorithm is terminated when either a specific iteration number is reached (e.g. 10) or the relative changes between $\mathbf{X}^{(i+1)}$ and $\mathbf{X}^{(i)}$ fall below a given threshold (e.g. 10^{-3}).

D. Other Implementation Details

A fixed spectra bandwidth (BW) of 1250 Hz was used. The convolution layers in Fig. 1 were set up with kernel size 24×24 and increased channel dimension (M-6-12-14) for consecutive layers with stride = 2. The "bottle-neck" structure contains layers with dimension O-200-L-200-O, with O being the flattened output data dimension from the previous complex convolutional block ($O = \mathrm{D}\backslash 8 \times \mathrm{K}_3$ as described in Fig. 1). Complex ReLu activation function ($\mathbb{C}$ReLU) was used in the nonlinear hidden layers, except for the middle

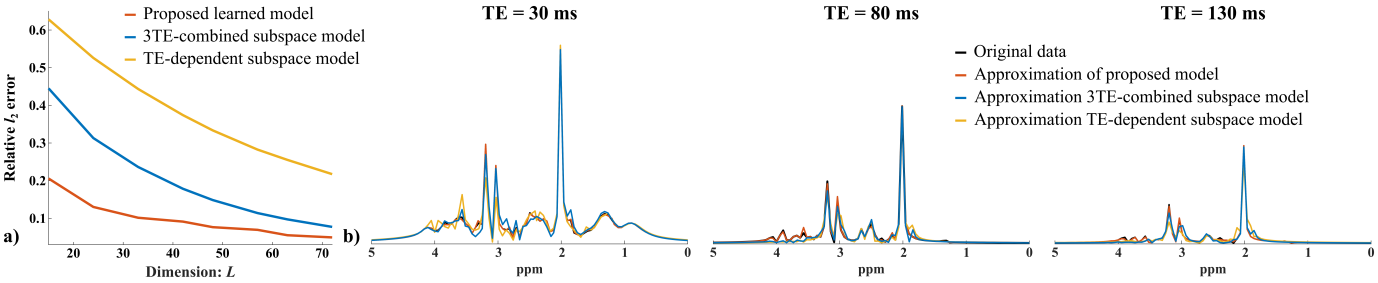

Fig. 2. Representation capability of the proposed method: a) approximation errors (relative ℓ_2) of the trained model (red curve) compared with two linear subspace models: TE-combined subspace (blue curve) and TE-dependent subspace (yellow curve), for a 3-TE case with different model orders (L); b) A representative 3-TE testing spectrum (black) and its approximation by the learned model (red), TE-combined subspace (blue), and TE-dependent subspace (yellow) with $L = 42$. Improved representation accuracy offered by the proposed model can be seen.

layer [38], [39]. The network training was performed on a Windows 10 machine with Intel(R) Core(TM) i9-9820X CPU and NVIDIA(R) TITAN RTX(TM) graphics processing unit and implemented in PyTorch using the Adam optimizer [51]. The batch size was 500, the initial learning rate was 0.001 with 300 epochs while other parameters remained as default. Same machine was used for the constrained reconstruction.

E. Simulation and Experimental Settings

1) Simulations: We first performed numerical simulations to evaluate the representation and generalization capability of the learned model with comparison to two alternative linear subspace-based representations, one using TE-dependent subspaces and another one with a TE-combined augmented subspace (3 TEs as an example here), all obtained from the same training data. For the TE-combined subspace, the Casorati matrix was constructed with the additional TE dimension stacked alone the temporal dimension for the training samples. The relative ℓ_2 approximation errors for a set of test data from the proposed learned model and the subspace projection errors were evaluated at different L's (For the subspaces, L denotes the dimensionality, i.e., the number of basis used for projection).

To evaluate the denoising performance of the proposed method, a 3-TE numerical phantom, with TE = 30, 80, and 130 ms, was constructed with similar procedures well-described in [30], [31]. Specifically, we first acquired the brain tissue segmentation maps from a T_1-weighted structural image to generate tissue fraction maps with different tissue types including gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Next, spectral parameter maps were created by combining the tissue fraction maps with regional metabolite and MM parameters from the literature [41], [42], [45], [47], and then used to synthesize spatially localized FIDs with the multi-TE signal model in Eq. (5). A lesion-like feature was included, where the concentrations of all the metabolites were decreased by a factor of 3 except for Cho (increased by a factor of 3) w.r.t their mean values in GM. Complex Gaussian noise was added to the synthesized (x, t, TE)-space data to simulate noisy acquisitions.

2) In Vivo Experiments: Data were acquired from healthy volunteers to evaluate the performance of the proposed method, with approval from the local Institutional Review Board (Institution: University of Illinois Urbana Champaign; Protocol number: 20132; Date: 2019/09/10). Experimental multi-TE ^{1}H-MRSI data were acquired on a 3T Siemens Prisma scanner using a 20-channel head coil and a multi-TE EPSI sequence [52] that acquires multiple TEs during a single scan. Both 2D and 3D scans with multiple averages were acquired. The parameters for the 2D scans were: TR/TE = 1200/(65, 120, 200) ms, field-of-view (FOV) = 210 mm $\times$ 210 mm, slice thickness = 15 mm (our sequence is a spin-echo sequence with only slice selective excitation thus excitation volume is the same as the FOV), matrix size = 32 $\times$ 32, spectral bandwidth (BW) = 1250 Hz and 256 gradient echo pairs in each EPSI readout. Total acquisition time was about 22 minutes with 11 acquisitions/averages. Such a 2D scan with many averages allowed for some quantitative evaluations of the denoised data (see Results for more details). The parameters for the 3D scans were: TR/TE = 1200/(65, 120, 200) ms, volume-of-interest (VOI) = 220 mm $\times$ 220 mm $\times$ 50 mm, FOV = 220 mm $\times$ 220 mm $\times$ 64 mm, matrix size = 32 $\times$ 32 $\times$ 8, same spectral BW and 320 gradient echoes. Total acquisition time was about 12 minutes. A 60 Hz weak water suppression and outer volume suppression (OVS) bands were used for all scans. Before the denoising reconstruction, an SVD-based coil combination was performed and nuisance water/lipid signals were removed using the method in [48]. Both noisy and denoised multi-TE data were quantified TE-by-TE using LCModel [53] as well as jointly using a ProFit based method [54].

IV. Results

A. Simulation Results

The proposed learned model consistently outperformed the two linear subspace models in terms of dimensionality reduction error (relative ℓ_2) at different model orders, as shown in Fig. 2a. The TE-combined subspace model had a higher representation efficiency than the single-TE subspaces, as expected since it exploited correlation across TEs. To further demonstrate the representation accuracy of the learned model, approximations of representative testing 3-TE spectra from these three models with the same $L = 42$ are compared in Fig. 2b. As can be seen, the proposed learned nonlinear model produced visually apparent better approximations than the linear subspaces with less spectral distortion.

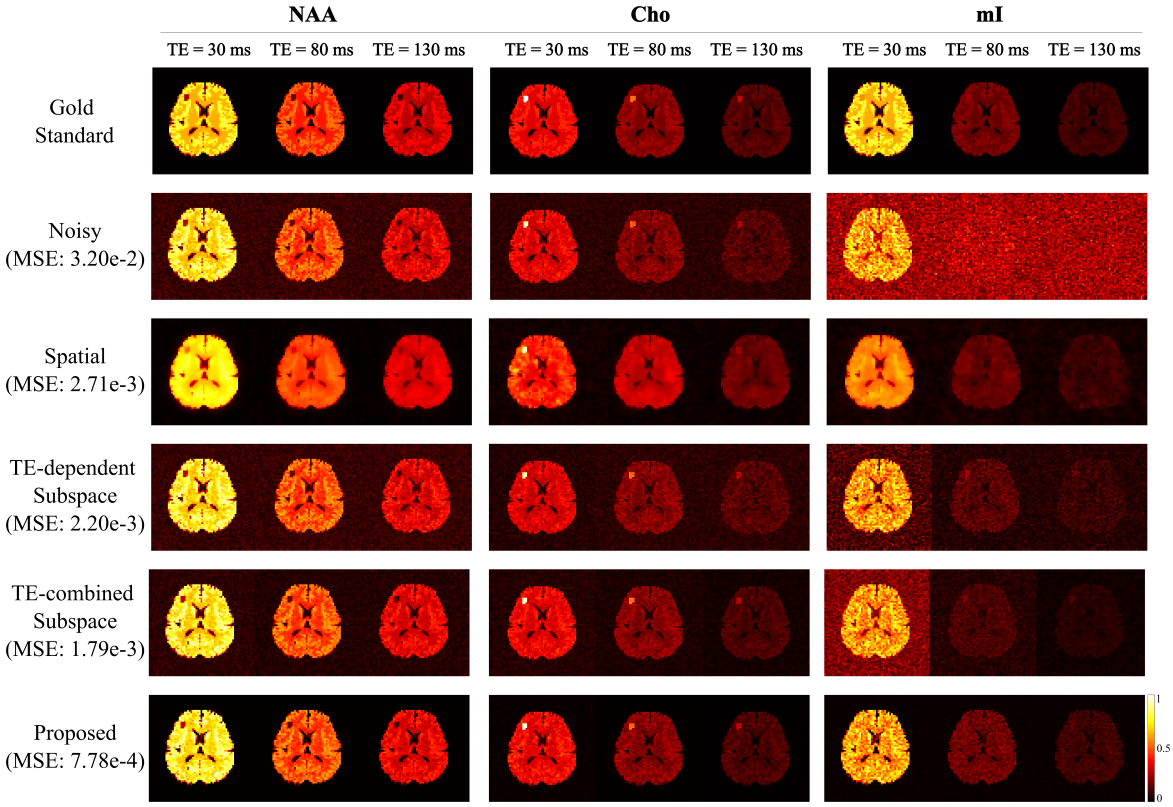

Fig. 3. Simulation results evaluating the denoising performance from the ground truth (Gold Standard), noisy data (Noisy), anatomically constrained denoising (Spatial), TE-dependent subspace denoising (TE-dependent subspace), and TE-combined subspace (TE-combined subspace) with their normalized mean-squared-errors (MSEs), respectively. Metabolite maps of NAA, Cho, and mI for all 3 TEs were shown. As can be seen, the proposed method achieved significant SNR improvement and the lowest MSE while better preserving spatiospectral features.

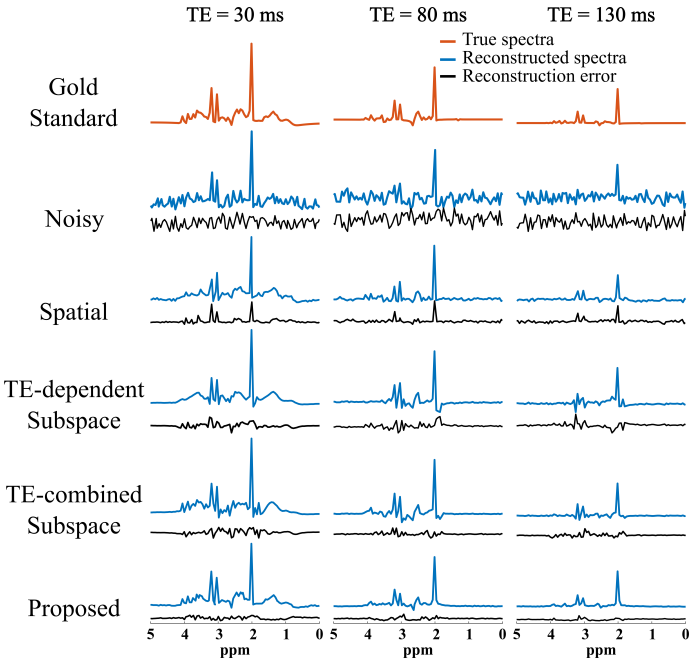

Fig. 4. Additional simulation results of a set of voxel spectra selected from GM. The true spectra, reconstructed spectra from different denoising schemes, and reconstruction errors were shown in red, blue, and black lines, respectively. Here SNR improvements with superior spectral feature preservation achieved by the proposed over alternative methods can be clearly observed.

Figure 3 shows a set of denoising results from the simulation phantom (at SNR = 20 defined w.r.t. the maximum NAA peak), produced by four alternative methods, i.e., the proposed, spatially constrained reconstruction (Eq. (3) with $\lambda_1 = 0$), subspace constrained denoising using TE-dependent subspaces and TE-combined subspace (a mild spatial constraint was used for both subspace methods). The learned model and linear subspace model were learned/estimated from the same training data with a model order $L = 52$ such that approximation error for the testing data was around 7% for the learned model. Note that the network did not see any data from the simulated phantom. The regularization parameter λ_1 for the proposed method was selected based on a single voxel denoising performance, while λ_2 was chosen based on discrepancy principle and then fine-tuned by minimizing the relative ℓ_2 error of the final spatiospectral reconstructions. As can be seen, the spatially constrained reconstruction (third row) improved the SNR but oversmoothed the images. The two subspace methods (fourth and fifth row) further improved the reconstruction (with TE-combined subspace slightly better than TE-dependent subspaces). The proposed method produced the best results (last row), as demonstrated qualitatively by the metabolite maps and quantitatively by normalized MSEs calculated for the entire spatiospectral image.

Figure 4 shows an additional comparison of the different denoising methods using selected voxel spectra from GM. Both the reconstructed (blue) and the error spectra (black) are

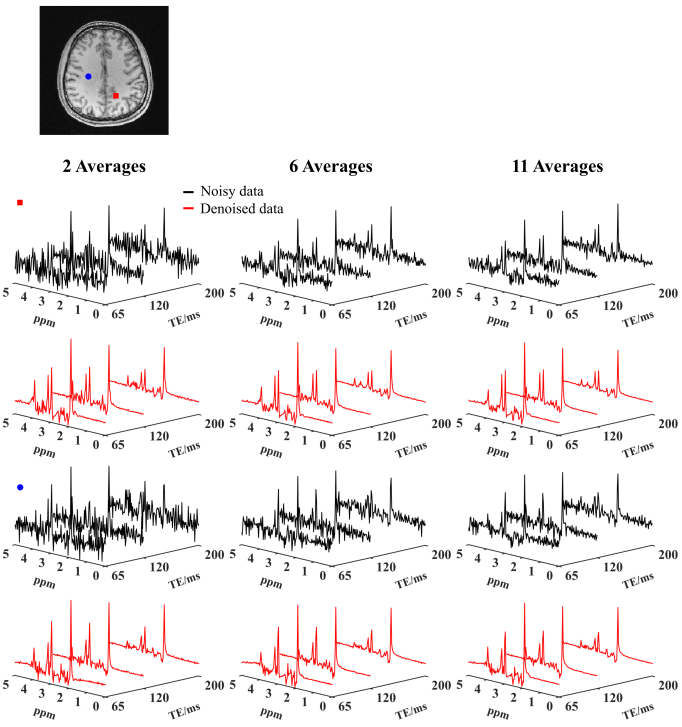

Fig. 5. Noisy (black) and denoised multi-TE spectra (red) for two different voxels from a 2D brain data set (voxels marked by the blue and red marks in the anatomical image on top). Results for different numbers of averages are shown in different columns. A significant SNR enhancement can be seen for the denoised data, and the proposed method produced more consistent reconstruction across different cases. A more quantitative comparison can be found in Fig. S1, which confirmed improved consistency due to denoising.

shown. As can be seen, the spatially constrained reconstruction produced noticeable spectra distortion (third row), while both subspace methods (fourth and fifth row) offered improved performance, with the TE-combined subspace exhibiting lower errors for longer TEs compared to TE-dependent subspaces. The proposed method achieved significant SNR improvement and best preserved spatiospectral features with the lowest reconstruction errors. The spectral components concealed by noise, e.g., the weak J-coupled metabolite and MM peaks, were better recovered by the proposed method.

B. In Vivo Results

We performed denoising to the in vivo 2D data with different numbers of averages (using the same regularization parameters from simulation), and compared the results. Convincing SNR enhancement and improved consistency across different numbers of averages offered by the proposed method are demonstrated by the localized multi-TE spectra shown in Fig. 5. The denoised spectra (red curves) better revealed spectral features and exhibit more realistic T_2 decays compared to their noisy counterparts. Quantification results (from LCModel) are shown in Fig. 6, where the metabolite maps (first three rows) along with estimated standard deviation (SD) maps (last three rows) from noisy measurement, subspace denoising method, and the proposed method are compared, for 2 averages, 6 averages, and 11 averages, respectively. For the noisy data, the SNR increment and estimation variance reduction as the number of averages increases can be clearly observed, as excepted. The proposed method led to the most improved quantification (especially for less averaged data), as shown by the decreased SD maps across different averages and TEs (Fig. 6, bottom panel). Particularly, the denoised data by the proposed method from 2 averages achieved significantly more consistent results to the data with 11 averages than the noisy data (i.e., more averages are needed for the noisy data to achieve similar estimation performance to the denoised data), and the results from 6 averages achieved almost the same quantification as 11 averages. This indicates that the proposed denoising should effectively reduce the data acquisition time. Representative spectra are shown in Fig. S2 to illustrate the LCModel fitting quality.

LCModel only fits spectra TE-by-TE, thus not taking full advantage of the multi-TE data for quantifying the J-coupled molecules. To this end, we adopted the ProFit strategy to jointly fit the data from all TEs [54]. Figure 7 shows the quantification results including molecules mI and Glx, for the noisy and denoised data. As can be seen, the SNR improvement offered by the proposed method does lead to improved quantification, revealing better tissue contrast. A comparison of the regression of Glx levels w.r.t. GM fraction for the noisy and denoised data can be found in Fig. S3. To further illustrate the general applicability of the proposed method, a denoising reconstruction for a 12-TE data set was performed and results are shown in Fig. S4. Improved metabolite spectra and maps from the proposed method can also be observed.

The performance of the proposed method was also evaluated quantitatively using a test-retest study leveraging the multiple averages. Specifically, the reconstruction from the first and last 5 averages are compared. The reproducibility of metabolite quantification was evaluated using a linear regression analysis of the metabolite estimates (from multi-TE fitting) of these two reconstructions in Fig. 8. As can be seen, the consistency of metabolite estimates between the two repeats is substantially improved by the proposed denoising, confirmed by the higher correlation coefficients for NAA, Cr, Cho, and Glx (the same analysis for other metabolites are shown in Fig. S5).

Spatially localized spectra from noisy and denoised in vivo 3D multi-TE data are shown in Fig. 9. As can be seen, the proposed method produced significantly improved SNR and well-preserved spectral features, especially for the data acquired at longer TEs. To demonstrate that those metabolite peaks revealed by denoising are not due to overfitting from the learned network, we also show a set of spectra from a background voxel (noise-only region, green mark). Minimum signals were recovered for this voxel by the proposed method with clearly observed noise variance reduction. Quantified metabolite maps (joint multi-TE fitting) are shown in Fig. 10. Higher quality metabolite maps with significantly less spatial uncertainty were produced from the denoised data with visually better spatial distributions, e.g., better-preserved ventricle features and clearer tissue contrasts. T_2 maps (for NAA, Cr, and Cho) were also estimated and compared in Fig. S6, where the denoised data yielded better estimations which are more consistent with the literature [55]. The

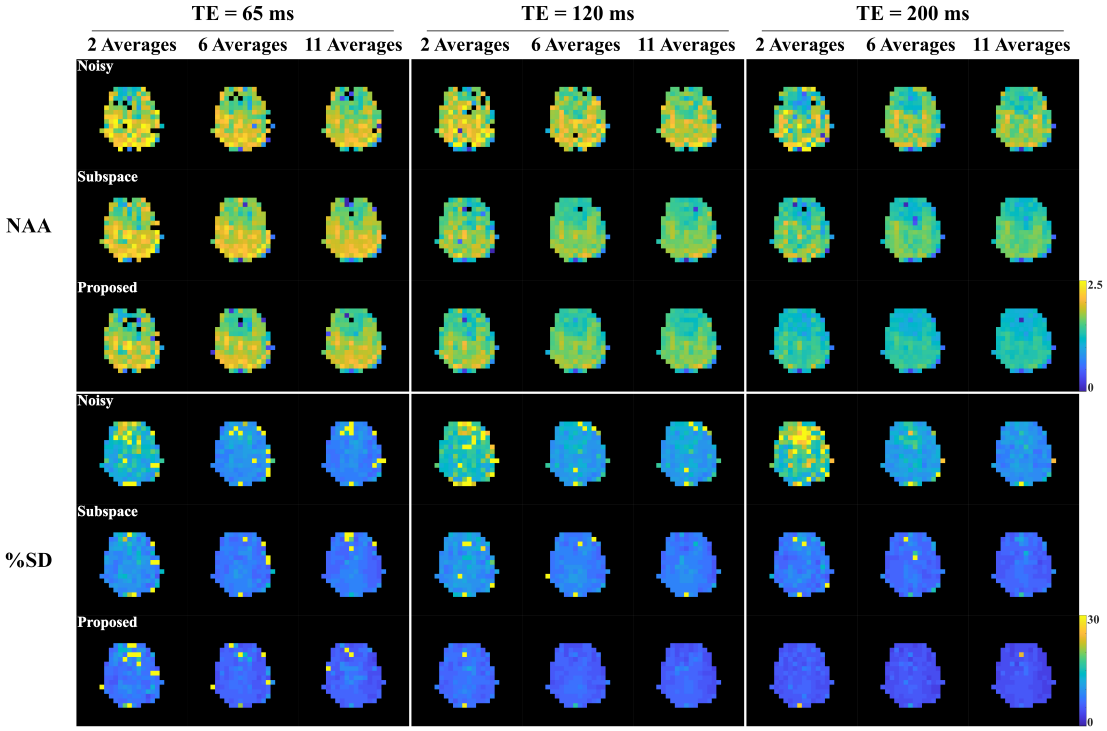

Fig. 6. Metabolite quantification by LCModel for the 2D in vivo data with different numbers of averages, from the Fourier reconstruction (Noisy), subspace denoising (Subspace), and the proposed method (Proposed). The top three rows are TE-dependent NAA maps, while the bottom three rows are the estimated standard deviation maps expressed in percentage of the estimated concentrations (%SD). Results for different TEs are in different panels. Each panel shows results for noisy data, subspace denoising, and the proposed method in different rows, and different columns display the results for different numbers of averages, respectively. It can be observed that for the noisy data, the estimation variance is reduced with more averages. Both denoising methods offer reduced SD. The proposed method yielded more improvement (especially for less averaged data) and consistent estimation across different cases. The values are shown in an institutional unit (I.U.).

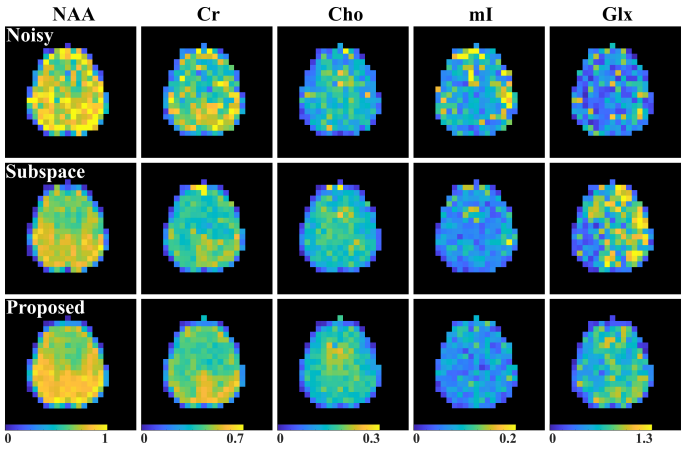

Fig. 7. Joint multi-TE metabolite fitting results (coefficients c_p in Eq. (5) in institutional units) for NAA, Cr, Cho, mI, and Glx from 2D in vivo data with 3 TEs and 6 averages. Results from the noisy data (Noisy), and denoised data produced by the subspace method (Subspace), and the proposed method (Proposed) are shown in the first, second, and third rows, respectively. As can be seen, the proposed method offers significantly improved metabolite maps with higher image quality, including for the J-coupled molecules.

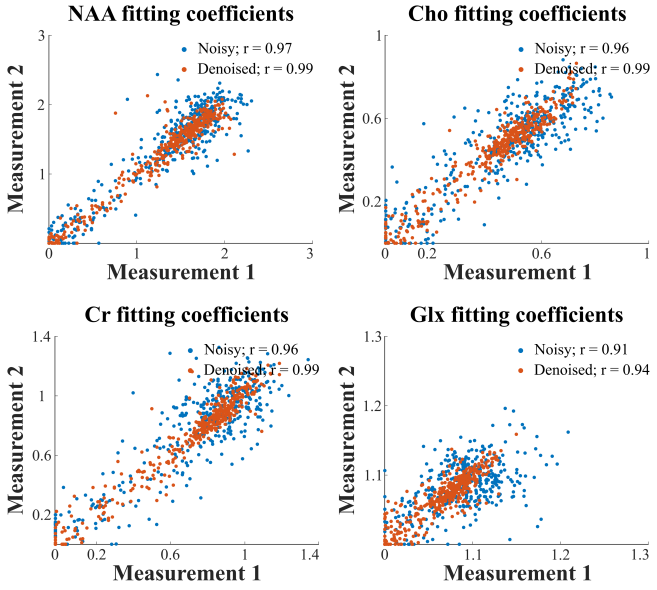

Fig. 8. Regression analysis of metabolite estimates from the noisy (blue dots) and denoised data (red dots). Results for four metabolites from the first and last 5 averages of the 2D multi-average scan are compared (each dot representing a voxel). A noticeably better consistency with higher correlation coefficients between the two measurements can be observed for the proposed denoising reconstruction, indicating a stronger reproducibility and less uncertainty.

processing time for this entire 3D data set was about one week (see more detailed discussion regarding computation time considerations in the next section).

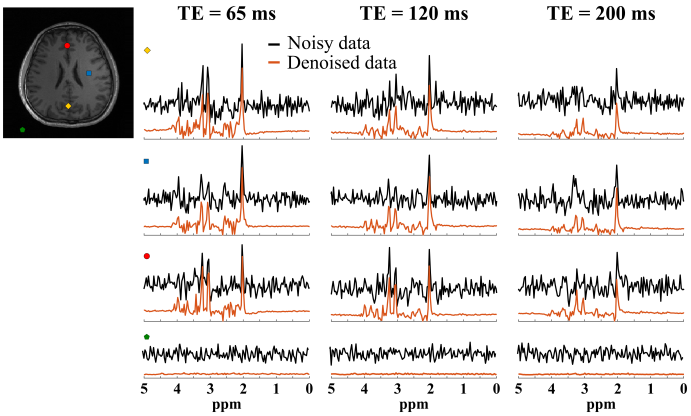

Fig. 9. Spatially resolved spectra from a 3D data set, before (black) and after denoising (orange). One voxel from posterior gray matter (yellow diamond), one from white matter region (blue square), one from anterior gray matter (red circle), and the fourth from the background region (green pentagon) are shown. Apparent SNR enhancement can be seen for the denoised data. Minimum signal is reconstructed from the background, indicating a small bias with a large variance reduction.

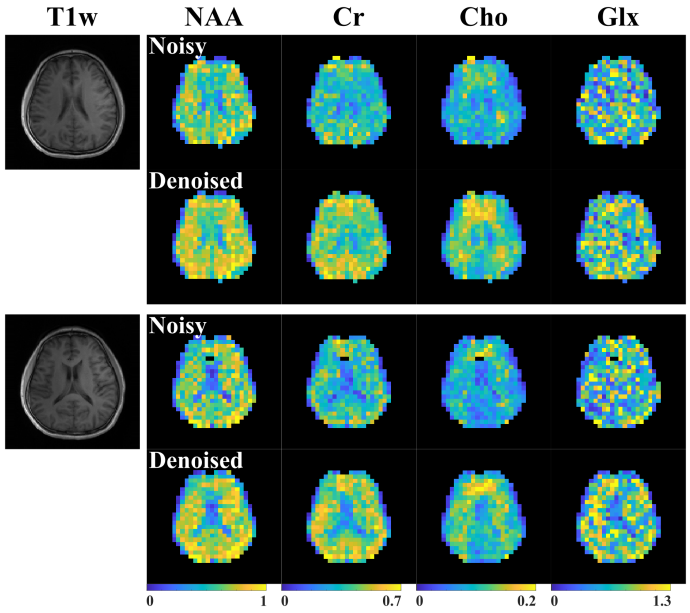

Fig. 10. Metabolite maps from the in vivo 3D data. The first column shows different slices of the anatomical images (T1w). The subsequent columns display maps of NAA, Cr, Cho, and Glx for the corresponding slice. Maps from the noisy data and the proposed method (Denoised) are shown in different rows. Significantly better metabolite distributions can be visualized for the denoised data.

V. Discussion

Our proposed method effectively integrated a learned nonlinear low-dimensional model of high-dimensional multi-TE spectroscopic signals and a spatial constraint for optimized SNR-enhancing MRSI reconstruction. Our complex-valued convolutional and fully-connected neural network blocks efficiently exploited the correlations among TEs and allowed for the extraction of accurate low-dimensional embedding from the high-dimensional multi-TE data. There are several unique advantages to our approach. First, unlike other methods with end-to-end mapping networks [56], [57], it provides strong flexibility for different acquisitions. Specifically, the learned model captures the inherent variations of the true signals and can be readily applicable to data acquired at different SNRs, resolutions, and sampling patterns. Our formulation can work with an arbitrary (k, t)-space sampling design (through the forward spatiospectral encoding operator) with the same learned model, e.g., the 3-TE in vivo 2D and 3D data were acquired with different acquisition parameters but processed using the same representation model. While retraining the network with a denoising regularizer at an SNR better matched with individual data set may further improve the result, it is not required for the proposed method, thus making it more flexible/generalizable. The model needs to be retrained for different excitation schemes (with the same network structure) as in this case the QM basis for various metabolites are different.

Another advantage of the proposed method is the model scalability, which is important when processing such high-dimensional data. Instead of concatenating real, imaginary parts, and signals from different TEs as in previous representation model learning approaches [30], [31], which will result in a huge fully-connected network, a more scalable combination of convolutional and fully-connected blocks with complex-valued units were used to better accommodate the complex nature of the multi-TE signals. This allows us to learn models for many more TEs while maintaining a manageable network size. Here we choose a 3-TE acquisition in this work as an example to demonstrate the SNR enhancement utility of the proposed method. But as mentioned before, the proposed network and algorithm designs can handle an arbitrary number of TEs (by changing the number of input channels to the network), including many-TE (e.g., the 12-TE case shown in Fig. S4) as well as single-TE MRSI data. Also, note that the choices of TE values here can be optimized to minimize the estimation variances for specific sets of molecules, e.g., through a Cramer-Rao bound analysis, which is beyond the scope of this work.

The regularization parameter selection (for λ_1 and λ_2 here) is one important issue. While results obtained with different combinations of λ_1 and λ_2 indicate that the denoising performance remains robust to a large range of parameter values (with smaller values for λ_2 achieving better balance between additional SNR enhancement and resolution loss; figure not shown due to space consideration), better parameter selection strategies, including those optimized for metrics beyond MSE and machine learning based approaches may be considered [58].

In addition, although the learned representation currently serves as a voxel-wise constraint, it can introduce time-dependent bias into the final reconstruction thus frequency-dependent spatial resolution. While the traditional metrics (like FWHM of point spread functions) are not sufficient to analyze such resolution effects with the nonlinearity of neural networks, a more careful spatial resolution analysis of the denoised reconstruction should be conducted in future research (e.g., using new analysis tools such as [59]).

Processing speed remains a challenge for the current im-

plementation. While the proposed algorithm is efficient (additional convergence analysis can be found in [31]) and allows for parallel processing of individual voxel data in one of the subproblems, the overall computation time is still demanding, especially for 3D data with high spatial resolutions and a large number of TEs (approximately one week to process the entire 3D data set). The most computationally expensive step is Subproblem (I) in Eq. (7) where a backpropagation is needed for the gradient update. More efficient sparse matrix treatments for the convolution and faster algorithms are highly desirable. The current Matlab implementation of the denoising algorithm uses both CPU and GPU, creating communication overheads. Implementing the entire algorithm on GPU should provide further acceleration. Increased computational efficiency will allow more flexible incorporation of advanced regularization terms, e.g., ℓ_1 or other non-quadratic penalties. These are currently being pursued. For the parametric model used in the data generation process, a Lorentzian/Gaussian lineshape function was adopted for metabolites/MMs. But more sophisticated lineshape distortion functions can be considered to improve model generalization. All the training data for the representation network were synthetic. Combining synthetic and high-quality experimental training data may further improve the model accuracy for practical data.

We used LCModel for the metabolite quantification, which is a well-established tool in the MR spectroscopy community but only fits data TE by TE, thus not taking full advantage of the TE-dependent signal evolution in the multi-TE data. While there are limitations associated with LCModel analysis for our multi-TE data (e.g., some poorly fitted voxels), it serves to provide a consistent comparison between the noisy and denoised data as well as demonstrate the denoising effects for individual TEs. More sophisticated quantification methods for J-resolved spectroscopy data may be considered to improve the estimate of the challenging molecules. These are directions worth investigating in future research.

VI. Conclusion

The proposed method effectively integrated a learned nonlinear low-dimensional model and spatial constraint to enhance the SNR for multi-TE MRSI. The deep complex-valued convolutional autoencoder learned an efficient nonlinear low-dimensional representation of high-dimensional multi-TE spectra and offered stronger dimensionality reduction capability for noise reduction than existing subspace models. Simulated and experimental data demonstrated improved multi-TE ^{1}H-MRSI spatiospectral reconstruction and subsequent quantitative metabolite analysis.

VII. Appendix

A. Gradient Calculation for Subproblem (I)

The gradients for individual voxels can be efficiently calculated through backpropagation based on the DCCAE design. More specifically, denote:

$$f_n(\mathbf{X}_n) = \lambda_1 \|\mathcal{N}(\mathbf{X}_n) - \mathbf{X}_n\|_F^2 + \frac{\mu}{2}\left\| \left[\mathbf{B} \odot \mathbf{X} - \mathbf{S}^{(i)} + \frac{\mathbf{Z}^{(i)}}{\mu} \right]_n \right\|_F^2 \tag{10}$$

as the loss function for the n-th voxel, then the gradient can be written as:

$$\begin{aligned} &\nabla f_n[\mathbf{vec}(\mathbf{X}_n)] \\ &= 2\lambda_1 \left(\mathbf{J}_{\mathcal{N}} - \mathbf{I}\right)^T \mathbf{vec}\left[(\mathcal{N}(\mathbf{X}_n)) - \mathbf{X}_n\right] \\ &+ \mu \mathbf{vec}\left[\mathbf{B}_n^H \odot \left(\mathbf{B}_n \odot \mathbf{X}_n - \mathbf{S}_n^{(i)} + \frac{\mathbf{Z}_n^{(i)}}{\mu} \right) \right]. \end{aligned} \tag{11}$$

where $J_{\mathcal{N}} \in \mathbb{C}^{(D\times M)\times(D\times M)}$ is the Jacobian of the NN mapping. $\mathbf{I}$ is an identity matrix with the same dimension as $J_{\mathcal{N}}$. B_n denotes the B_0 field inhomogeneity matrix in n-th voxel. For a neural network with S layers and a linear final layer, the Jacobian matrix $J_{\mathcal{N}}$ can be derived through backpropagation [30]:

$$\mathbf{J}_{\mathcal{N}} = \mathbf{W}_S^T \times \prod_{s=1}^{S-1} \mathbf{U}_s \mathbf{W}_s^T, \tag{12}$$

where $\mathbf{W}_s$ represents the complex weights for the s-th layer (note that here the weights for convolutional and fully-connected layer can both be denoted as $\mathbf{W}_s$). $\mathbf{U}_s$ represents the diagonal matrices with the s-th layer's nonlinear activation function derivatives. This matrix multiplications can be done with similar complex multiplication rules described in Fig 1.

References

[1] R. A. de Graaf, *In Vivo NMR Spectroscopy: Principles and Techniques*. Hoboken, NJ: John Wiley and Sons, 2007.

[2] M. A. Thomas *et al.*, "Localized 2D J-resolved 1H MR spectroscopy of human brain tumors in vivo," *J Magn Reson Imag*, vol. 6, pp. 453–459, 1996.

[3] W. Bogner *et al.*, "1D-spectral editing and 2D multispectral in vivo 1H-MRS and 1H-MRSI - Methods and applications," *Anal Biochem*, vol. 529, pp. 48–64, 2017.

[4] R. F. Schulte *et al.*, "Improved two-dimensional J-resolved spectroscopy," *NMR Biomed*, vol. 19, pp. 264–270, 2006.

[5] Y. Li *et al.*, "Comparison of T1 and T2 metabolite relaxation times in glioma and normal brain at 3T," *J Magn Reson Imag*, vol. 28, pp. 342–350, 2008.

[6] Y. Li *et al.*, "Three-dimensional J-resolved H-1 magnetic resonance spectroscopic imaging of volunteers and patients with brain tumors at 3T," *Magn Reson Med*, vol. 58, pp. 886–892, 2007.

[7] H. de Greiff *et al.*, "Signal de-noising in magnetic resonance spectroscopy using wavelet transforms," *Concept Magn Reson*, vol. 14, pp. 388–401, 2002.

[8] A. Diop *et al.*, "Improvements of quantitation by using the Cadzow enhancement procedure prior to any linear-prediction methods," *J Magn Reson*, vol. 105, pp. 17–24, 1994.

[9] H. M. Nguyen *et al.*, "Denoising MR spectroscopic imaging data with low-rank approximations," *IEEE Trans Biomed Eng*, vol. 60, pp. 78–89, 2012.

[10] J. Ying *et al.*, "Hankel matrix nuclear norm regularized tensor completion for N-dimensional exponential signals," *IEEE Trans Signal Process*, vol. 65, pp. 3702–3717, 2017.

[11] Z.-P. Liang and P. C. Lauterbur, "A generalized series approach to MR spectroscopic imaging," *IEEE Trans Med Imaging*, vol. 10, pp. 132–137, 1991.

[12] J. P. Haldar *et al.*, "Anatomically constrained reconstruction from noisy data," *Magn Reson Med*, vol. 59, pp. 810–818, 2008.

[13] J. Kornak *et al.*, "Bayesian k-space-time reconstruction of MR spectroscopic imaging for enhanced resolution," *IEEE Trans Med Imaging*, vol. 29, pp. 1333–1350, 2010.

[14] H. M. Nguyen *et al.*, "Denoising of MR spectroscopic imaging data with spatial-spectral regularization," in *Proc IEEE Int Symp Biomed Imag*. IEEE, 2010, pp. 720–723.

[15] R. Eslami and M. Jacob, "Robust reconstruction of MRSI data using a sparse spectral model and high resolution MRI priors," *IEEE Trans Med Imaging*, vol. 29, pp. 1297–1309, 2010.

[16] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," in *Proc IEEE Int Symp Biomed Imag*, 2007, pp. 988–991.

[17] Y. Liu *et al.*, "Improved low-rank filtering of magnetic resonance spectroscopic imaging data corrupted by noise and B_0 field inhomogeneity," *IEEE Trans Biomed Eng*, vol. 63, pp. 841–849, 2015.

[18] F. Lam and Z.-P. Liang, "A subspace approach to high-resolution spectroscopic imaging," *Magn Reson Med*, vol. 71, pp. 1349–1357, 2014.

[19] F. Lam *et al.*, "High-resolution 1H-MRSI of the brain using SPICE: Data acquisition and image reconstruction," *Magn Reson Med*, vol. 76, pp. 1059–1070, 2016.

[20] J. Kasten *et al.*, "Data-driven MRSI spectral localization via low-rank component analysis," *IEEE Trans Med Imaging*, vol. 32, pp. 1853–1863, 2013.

[21] Y. Chen *et al.*, "Improved low-rank filtering of MR spectroscopic imaging data with pre-learnt subspace and spatial constraints," *IEEE Trans Biomed Eng*, vol. 67, pp. 2381–2388, 2019.

[22] Y. Li *et al.*, "A subspace approach to spectral quantification for MR spectroscopic imaging," *IEEE Trans Biomed Eng*, vol. 64, pp. 2486–2489, 2017.

[23] M. Sarma *et al.*, "Accelerated echo-planar J-resolved spectroscopic imaging in the human brain using compressed sensing: A pilot validation in obstructive sleep apnea," *AJNR Am J Neuroradiol*, vol. 35, pp. S81–S89, 2014.

[24] N. E. Wilson *et al.*, "Accelerated five-dimensional echo planar J-resolved spectroscopic imaging: Implementation and pilot validation in human brain," *Magn Reson Med*, vol. 75, pp. 42–51, 2016.

[25] C. Ma *et al.*, "Accelerated high-resolution multidimensional 1H-MRSI using low-rank tensors," in *Proc Intl Soc Mag Reson Med*, 2016, p. 379.

[26] J. He *et al.*, "Accelerated high-dimensional MR imaging with sparse sampling using low-rank tensors," *IEEE Trans Med Imaging*, vol. 35, pp. 2119–2129, 2016.

[27] C. S. Bolliger *et al.*, "On the use of Cramer-Rao minimum variance bounds for the design of magnetic resonance spectroscopy experiments," *NeuroImage*, vol. 83, pp. 1031–1040, 2013.

[28] F. Lam *et al.*, "Accelerated J-resolved MRSI using joint subspace and sparsity constraints," in *Proc Intl Soc Mag Reson Med*, 2017, p. 1202.

[29] L. Tang *et al.*, "Accelerated J-resolved 1H-MRSI with limited and sparse sampling of (k, t_1, t_2)-space," *Magn Reson Med*, vol. 85, pp. 30–41, 2021.

[30] F. Lam *et al.*, "Constrained magnetic resonance spectroscopic imaging by learning nonlinear low-dimensional models," *IEEE Trans Med Imaging*, vol. 39, pp. 545–555, 2020.

[31] Y. Li *et al.*, "Separation of metabolites and macromolecules for short-TE 1H-MRSI using learned component-specific representations," *IEEE Trans Med Imaging*, vol. 40, pp. 1157–1167, 2021.

[32] P. Tiwari *et al.*, "Consensus-locally linear embedding (C-LLE): application to prostate cancer detection on magnetic resonance spectroscopy," in *MICCAI 2008*. Springer, 2008, pp. 330–338.

[33] G. Yang *et al.*, "Manifold learning in MR spectroscopy using nonlinear dimensionality reduction and unsupervised clustering," *Magn Reson Med*, vol. 74, pp. 868–878, 2015.

[34] P. Tiwari *et al.*, "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, 2009.

[35] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, pp. 504–507, 2006.

[36] P. Vincent *et al.*, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion." *J Mach Learn Res*, vol. 11, 2010.

[37] Y. Bengio *et al.*, "Representation learning: A review and new perspectives," *IEEE Trans Pattern Anal Mach Intell*, vol. 35, pp. 1798–1828, 2013.

[38] C. Trabelsi *et al.*, "Deep complex networks," *arXiv preprint arXiv:1705.09792*.

[39] E. Cole *et al.*, "Analysis of deep complex-valued convolutional neural networks for MRI reconstruction and phase-focused applications," *Magn Reson Med*, vol. 86, pp. 1093–1109, 2021.

[40] B. Bilgic *et al.*, "Improving parallel imaging by jointly reconstructing multi-contrast data," *Magn Reson Med*, vol. 80, pp. 619–632, 2018.

[41] 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, 2019.

[42] R. Birch *et al.*, "Influence of macromolecule baseline on 1H MR spectroscopic imaging reproducibility," *Magn Reson Med*, vol. 77, pp. 34–43, 2017.

[43] Z. Starcuk Jr *et al.*, "Simulation of coupled-spin systems in the steady-state free-precession acquisition mode for fast magnetic resonance (MR) spectroscopic imaging," *Meas Sci Technol*, vol. 20, p. 104033, 2009.

[44] C. Cudalbu *et al.*, "Contribution of macromolecules to brain 1H MR spectra: Experts' consensus recommendations," *NMR Biomed*, vol. 34, p. e4393, 2021.

[45] K. Landheer *et al.*, "Concentration and effective T2 relaxation times of macromolecules at 3T," *Magn Reson Med*, vol. 84, pp. 2327–2337, 2020.

[46] F. Lam *et al.*, "Ultrafast magnetic resonance spectroscopic imaging using SPICE with learned subspaces," *Magn Reson Med*, vol. 83, pp. 377–390, 2020.

[47] P. O. Wyss *et al.*, "In vivo estimation of transverse relaxation time constant (T2) of 17 human brain metabolites at 3T," *Magn Reson Med*, vol. 80, pp. 452–461, 2018.

[48] C. Ma *et al.*, "Removal of nuisance signals from limited and sparse ^{1}H MRSI data using a union-of-subspaces model," *Magn Reson Med*, vol. 75, pp. 488–497, 2016.

[49] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing," *SIAM J. Sci. Comput.*, vol. 33, pp. 250–278, 2011.

[50] J. Nocedal and S. J. Wright, "Conjugate gradient methods," *Numerical optimization*, pp. 101–134, 2006.

[51] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[52] Z. Wang *et al.*, "High-resolution, 3D multi-TE 1H MRSI using fast spatiospectral encoding and subspace imaging," *Magn Reson Med*, 2021.

[53] S. W. Provencher, "Estimation of metabolite concentrations from localized in vivo proton NMR spectra," *Magn Reson Med*, vol. 30, pp. 672–679, 1993.

[54] R. F. Schulte and P. Boesiger, "Profit: two-dimensional prior-knowledge fitting of J-resolved spectra," *NMR Biomed*, vol. 19, pp. 255–263, 2006.

[55] S.-Y. Tsai *et al.*, "Fast mapping of the T2 relaxation time of cerebral metabolites using proton echo-planar spectroscopic imaging (PEPSI)," *Magn Reson Med*, vol. 57, pp. 859–865, 2007.

[56] X. Qu *et al.*, "Accelerated nuclear magnetic resonance spectroscopy with deep learning," *Angew Chem Int Ed*, vol. 132, pp. 10 383–10 386, 2020.

[57] D. Chen *et al.*, "Review and prospect: deep learning in nuclear magnetic resonance spectroscopy," *Chem Eur J*, vol. 26, pp. 10 391–10 401, 2020.

[58] S. Liu and J. Zhang, "Machine-learning-based prediction of regularization parameters for seismic inverse problems," *Acta Geophys*, pp. 1–12, 2021.

[59] C.-C. Chan and J. P. Haldar, "Local perturbation responses and checkerboard tests: Characterization tools for nonlinear MRI methods," *Magn Reson Med*, vol. 86, pp. 1873–1887, 2021.