LeaRning nonlineAr representatIon and projectIon for faSt constrained MRSI rEconstruction (RAIISE)

Yahang Li [1], Loreen Ruhm [2], Zepeng Wang [2], Ruiyang Zhao [2], Aaron Anderson [2], Paul Arnold [2], Graham Huesmann [2], Anke Henning [2], and Fan Lam [2]

[1]Department of Bioengineering
[2]Affiliation not available

October 31, 2023

Abstract

Learning and utilizing low-dimensional models for high-dimensional spatiospectral imaging problems is an active research area. We present here a novel method for computationally efficient reconstruction from noisy high-dimensional MR spectroscopic imaging (MRSI) data. The proposed method features (a) a novel strategy that jointly learns a nonlinear low-dimensional representation of high-dimensional spectroscopic signals and a neural-network-based projector to recover the low-dimensional embeddings from noisy/limited data; (b) a joint formulation that integrates the forward spatiospectral encoding model, a constraint exploiting the learned representation, and a complementary spatial constraint; and (c) a highly efficient algorithm enabled by a learned projector within an alternating direction method of multipliers (ADMM) framework, circumventing the computationally expensive network inversion subproblem. The proposed method has been evaluated using simulations and in vivo 31P-MRSI and 1H-MRSI data, demonstrating improved performance over state-of-the-art methods. Computational complexity and algorithm convergence analysis have been performed to offer further insights into the effectiveness of the proposed method.

LeaRning nonlineAr representatIon and projectIon for faSt constrained MRSI rEconstruction (RAIISE)

Yahang Li, Loreen Ruhm, Zepeng Wang, Ruiyang Zhao, Aaron Anderson, Paul Arnold, Graham Huesmann, Anke Henning, Fan Lam

***Abstract*— Learning and utilizing low-dimensional models for high-dimensional spatiospectral imaging problems is an active research area. We present here a novel method for computationally efficient reconstruction from noisy high-dimensional MR spectroscopic imaging (MRSI) data. The proposed method features (a) a novel strategy that jointly learns a nonlinear low-dimensional representation of high-dimensional spectroscopic signals and a neural-network-based projector to recover the low-dimensional embeddings from noisy/limited data; (b) a joint formulation that integrates the forward spatiospectral encoding model, a constraint exploiting the learned representation, and a complementary spatial constraint; and (c) a highly efficient algorithm enabled by a learned projector within an alternating direction method of multipliers (ADMM) framework, circumventing the computationally expensive network inversion subproblem. The proposed method has been evaluated using simulations and in vivo ^{31}P-MRSI and ^{1}H-MRSI data, demonstrating improved performance over state-of-the-art methods. Computational complexity and algorithm convergence analysis have been performed to offer further insights into the effectiveness of the proposed method.**

***Index Terms*— MR spectroscopic imaging, representation learning, learning low-dimensional projection, neural network, denoising, regularized reconstruction.**

I. Introduction

MR spectroscopic imaging (MRSI) is a powerful molecular imaging modality that can simultaneously map the spatiotemporal variations of a range of physiologically important endogenous metabolites noninvasively, with many potential applications from basic sciences to clinical translations [1]–[3]. Nevertheless, due to the inherent low abundance of the metabolites of interest, MRSI signals are typically contaminated by a strong level of noise, limiting achievable resolution. Furthermore, the high dimensionality of the imaging problem due to the need to encode and decode both spatial and spectral dimensions prefers faster acquisitions which may further reduce the signal-to-noise ratio (SNR). In the meantime, the multidimensional nature of the MRSI data presents unique opportunities to develop processing methods to enable better SNR, resolution, and speed tradeoffs, by exploiting low-dimensional structures of the underlying signals of interest.

In particular, many constrained reconstruction methods have been proposed to construct a low-complexity model by utilizing MRSI signals' unique characteristics for improved reconstruction and quantification, e.g., spatial and/or spectral domain sparsity [4]–[8] and low rank properties [9]–[15]. To further take advantage of the spatial prior information from anatomical scans readily available in most MRSI experiments, various spatial constraints have also been introduced for MR spatiospectral reconstruction [16]–[20]. Subspace imaging exploited the fact that high-dimensional spectroscopic signals can be well approximated by a low-dimensional subspace which can be prelearned from special training data [21]–[23], which significantly reduced the degrees-of-freedom of the imaging problem to enable fast, high-resolution MRSI, and has been demonstrated with state-of-the-art performance across various field strength and nuclei [23]–[29].

With the significant developments in deep-learning-based MRI reconstruction, efforts have been made to develop learning-based reconstruction and processing methods for MRSI and demonstrate improved performance. One approach is to train end-to-end networks that directly map noisy, artifact-corrupted data to higher-SNR, artifact-reduced data via supervised learning [30]–[34]. As high-SNR, high-resolution MRSI data are rather difficult to acquire, simulated single-voxel spectroscopy or MRSI data have been used in training such networks [33]. But simulated data often do not fully capture the variations observed in experimental data, especially the spatial patterns of different metabolites, which can lead to significant bias for complete data-driven learning. Moreover, end-to-end mapping provides less flexibility for changes in ac-

Correspondence to F. Lam, Ph.D., Department of Bioengineering, Department of Electrical and Computer Engineering, Beckman Institute for Advanced Science and Technology, and Cancer Center at Illinois, University of Illinois Urbana-Champaign, Urbana, IL, 61801 USA e-mail: fanlam1@illinois.edu.

Y. Li and Z. Wang are with the Department of Bioengineering and the Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign. R. Zhao is with the Department of Electrical and Computer Engineering and the Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign. A. Anderson is with the Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign.

P. Arnold and G. Huesmann are with the Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, and Carle Foundation Hospital.

L. Ruhm and A. Henning are with the Advanced Imaging Research Center, University of Texas Southwestern Medical Center, and Max Planck Institute for Biological Cybernetics.

This work is supported in part by NSF-CBET 1944249, NIH-1R21EB029076A, and NIH-1R35GM142969.

quisition parameters during inference. An alternative approach is to learn low-dimensional representations of the desired spectroscopic signals using deep networks, which generalize the linear subspace models to provide better model accuracy. The learned representations in the form of a trained network can then be integrated with the physical forward encoding model in an overall optimization formulation [27], [35]–[37]. While this approach simplified the learning problem and allowed for flexible adaptation for different acquisition parameters and SNRs, it resulted in a high-dimensional, nonconvex problem that required inverting the trained network in each iteration, which is computationally demanding.

Here, we propose a new method that leverages deep learning techniques for efficient MRSI reconstruction exploiting learned low-dimensional models, called RAIISE (joint leaRning of nonlineAr representatIon and projectIon for faSt constrained MRSI rEconstruction). RAIISE is characterized by three key features: (a) a novel strategy to jointly learn low-dimensional representations of desired high-dimensional spectroscopic signals and a network-based projector that can recover the low-dimensional embeddings (or latent variables) from noisy data; (b) a new constrained optimization formulation that jointly solves for the unknown spatiotemporal function of interest and the latent variables, integrating the forward encoding model, a constraint enforced through the learned representation, and a complementary spatial constraint; and (c) a highly efficient algorithm enabled by the learned projector to circumvent network backpropagation during one of the subproblems in an ADMM framework. We have evaluated RAIISE using simulations as well as in vivo ^{31}P-MRSI and ^{1}H-MRSI data. RAIISE achieved impressive SNR-enhancing reconstruction and produced similar or slightly better results than a previously published state-of-the-art network-constrained MRSI reconstruction method but with two orders of magnitude faster processing speed.

The remaining of the paper is organized as follows: Section II provides background on the MRSI reconstruction problem. Section III presents details of the proposed method, including the projector learning strategy, reconstruction formulation, and optimization algorithm. Section IV describes the simulation and experimental settings for evaluations, and Section V presents representative results followed by a discussion and conclusion in Sections VI and VII.

II. Background

The MRSI data acquisition process can be described as:

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

where $d(\mathbf{k},t)$ denotes the corresponding (k,t)-space measurements, $\rho(\mathbf{r},t)$ represents the unknown high-dimensional spatiotemporal function (Fourier counterpart of the desired spatiospectral function), t indexes the free induction decay (FID) time dimension, $\delta f(\mathbf{r})$ denotes the B_0 field inhomogeneity, $\mathbf{k}$ the k-space coordinates, and $n(\mathbf{k},t)$ the measurement noise (commonly assumed to be complex white Gaussian). The overall imaging problem here is defined as recovering $\rho(\mathbf{r},t)$ from a set of extremely noisy measurements $d(\mathbf{k},t)$, for which effective constraints are critical.

A generic constrained MRSI reconstruction formulation can be written as (after proper discretization):

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

The first term imposes data consistency where $\mathbf{B}$ captures the B_0 field inhomogeneity, $\mathcal{F}_\Omega$ represents the encoding operator with a (k,t)-space sampling pattern Ω, and $\odot$ is an element-wise multiplication. The second term $\mathcal{R}(.)$ imposes the constraints based on specific prior information, e.g., the unknown $\boldsymbol{\rho}$ resides on a learned low-dimensional manifold $\mathcal{R}(\boldsymbol{\rho};\mathcal{N}_\theta(\boldsymbol{\rho}))$ with $\mathcal{N}_\theta$ being the trained network parameterized by θ, and λ is the regularization parameter. This formulation offers great flexibility with different acquisition designs, but solving the associated optimization problem can be computationally demanding. For example, with the neural network term, a very time-consuming Jacobian matrix calculation (i.e., $\frac{\partial\mathcal{R}(\boldsymbol{\rho};\mathcal{N}_\theta(\boldsymbol{\rho}))}{\partial\boldsymbol{\rho}}$) is needed for each voxel FID in each iteration. Therefore in this work, we proposed a novel method for computationally efficient reconstruction of MRSI data leveraging the learned low-dimensional representation prior via a learned-projection-based ADMM algorithm, and a novel strategy to jointly learn the nonlinear low-dimensional representation and its associated projector. Details can be found below.

III. Proposed Method

A. Jointly Learning Nonlinear Representation and Projection for MRSI Data

As demonstrated in [35], [36], learned nonlinear representations can serve as effective constraints for high-SNR MRSI reconstruction. And the expressiveness and flexibility of neural networks make them a powerful tool for learning such representations, e.g., using a deep autoencoder (as illustrated in Fig. 1 top branch) [35]. Specifically, let $E(.;\theta_e)$ and $D(.;\theta_d)$ denote the learned encoder and decoder, and θ_e and θ_d containing the corresponding network parameters, low-dimensional modeling of the spectroscopic signals can be enforced by a penalty term $\|D(E(\mathbf{x};\theta_e);\theta_d)-\mathbf{x}\|_2^2$. However, solving an optimization problem with this term can be computationally challenging due to the need of inverting the network. With RAIISE, we sought to reformulate this problem using a projection network. Specifically, we first adapted the learning strategy in [35], [36] for pretraining a representation model ($E(.;\theta_e)$ and $D(.;\theta_d)$). Then, a set of noisy training samples $\widetilde{\mathbf{x}}_m$ were generated by combining the same training data used for representation learning with Gaussian noise at various SNRs (Fig. 1, bottom branch). A projector was trained to recover the low-dimensional embeddings (captured by the pre-trained encoder) from the noisy data as follows:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_p=\arg\min_{\boldsymbol{\theta}_p}\frac{1}{M}\sum_{m=1}^{M}&\epsilon_1(E(\mathbf{x}_m;\boldsymbol{\theta}_e),P(\widetilde{\mathbf{x}}_m;\boldsymbol{\theta}_p))\\&+\gamma\epsilon_2(\mathbf{x}_m,D(P(\widetilde{\mathbf{x}}_m;\boldsymbol{\theta}_p);\boldsymbol{\theta}_d)),\end{aligned} \qquad (3)$$

where $P(.;\boldsymbol{\theta}_p)$ is the projector with parameters $\boldsymbol{\theta}_p$, $\widetilde{\mathbf{x}}_m$ are the noisy input, ϵ_1 and ϵ_2 compute the "projection" errors

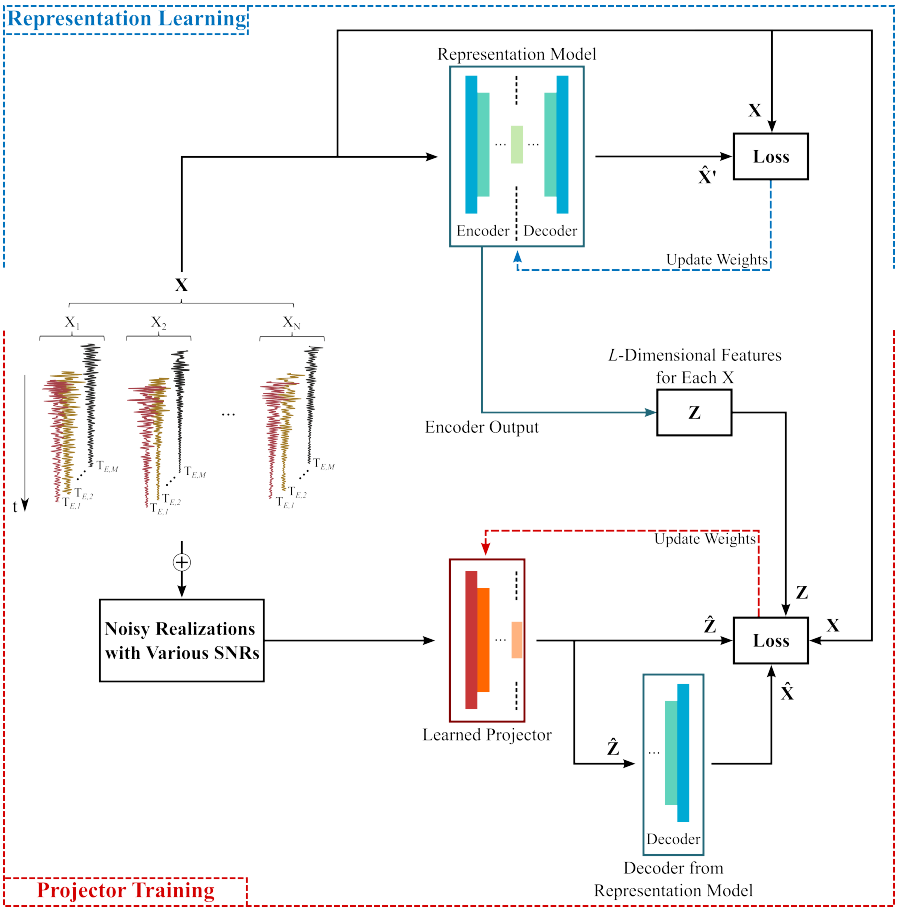

Fig. 1. The proposed strategy to jointly learn a nonlinear low-dimensional representation of spectroscopic signals (top branch) and a projector for recovering the low-dimensional features/embeddings from the corresponding noisy measurements (bottom branch). The representation model was a (complex) deep autoencoder with either fully-connected or convolutional layers (for single- or multi-TE data, respectively) [35], [36]. The projector was trained using the same data (with noise at different SNRs added) and the prelearned decoder from the representation network, aiming to recover low-dimensional embeddings of the spectroscopy signals captured by the representation encoder from their noisy counterparts and to reconstruct the decoded signals accurately (see the long arrow on the right from the top to bottom branch).

for the low-dimensional features and the full signals, respectively, and γ balances the two losses. The encoder $E(.;\theta_e)$ and decoder $D(.;\theta_d)$ are from the representation network (Fig. 1 top branch). The first term is used to minimize the difference between the hidden representation of the noiseless signal captured by the encoder and its counterparts recovered by the learned projector. The second term minimizes the difference between the reconstructed signal by the decoder from the projected embedding and the true noiseless signal. The projector can be realized using the same architecture as the representation encoder with different layer designs adapted to the data characteristics (single- or multi-TE FIDs), and ϵ_1 and ϵ_2 can be chosen separately. In this study, mean squared error (MSE) was used for both ϵ_1 and ϵ_2.

B. Reconstruction Using the Learned Projector

With the learned projector, RAIISE reformulates the constrained MRSI reconstruction problem as follows:

$$\hat{\mathbf{X}}, \hat{\mathbf{Z}} = \arg\min_{\mathbf{X},\mathbf{Z}} \|\mathbf{d} - \mathcal{F}_\Omega\{\mathbf{B} \odot \mathbf{X}\}\|_2^2 + \lambda R(\mathbf{X}) \quad s.t. \quad \mathbf{X} = D(\mathbf{Z}), \tag{4}$$

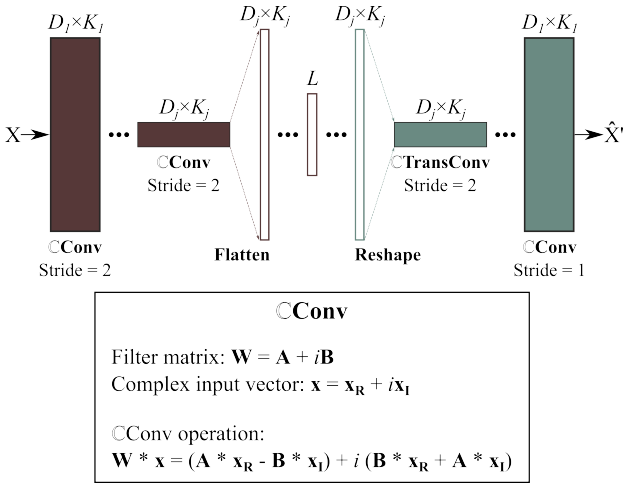

Fig. 2. Network design: $\mathbf{X}$ denotes the collection of training data (single- or multi-TE FIDs) with $\boldsymbol{D_j}$ being the data dimension in each layer. For multi-TE data, convolutional layers were included for exploiting multi-dimensional correlation (TE dimension treated as separate input channels $\boldsymbol{K_j}$), and fully connected middle layers were used for further low-dimensional feature extraction. Complex-valued units and activation functions were used (as defined in the figure) [38], [39].

where $\mathbf{X} = D(\mathbf{Z})$ (D being the decoder) enforces the prior that the underlying spectroscopic signal $\mathbf{X}$ should yield a low-dimensional representation $\mathbf{Z}$ (residing on a low-dimensional manifold). $\mathbf{B}$ models the B_0 inhomogeneity, $\mathcal{F}_\Omega$ is an encoding operator with sampling pattern Ω, $\mathbf{d}$ contains the noisy (k,t)-space measurements, and $R(.)$ imposes a complementary constraint with regularization parameter λ, e.g., a spatial regularization $R(\mathbf{X}) = \|\mathbf{D}_w\mathbf{X}\|_F^2$ in this work, where $\mathbf{D}_w$ is a weighted finite-difference operator with weights derived from a reference image.

An ADMM-based method was used to solve this optimization problem, i.e., with an augmented Lagrangian function:

$$L(\mathbf{X},\mathbf{Z},\mathbf{Y}) = \|\boldsymbol{d} - \mathcal{F}_\Omega\{\mathbf{B}\odot\mathbf{X}\}\|_2^2 + \lambda\|\mathbf{D}_w\mathbf{X}\|_F^2 + \frac{\mu}{2}\|\ \mathbf{X} - D(\mathbf{Z}) + \frac{\mathbf{Y}}{\mu}\|_F^2, \tag{5}$$

where $\mathbf{Y}$ is the Lagrangian multiplier. Then, the following three subproblems were solved in an iterative fashion to solve the overall problem (with i being the iteration index):
Subproblem (I): Update $\mathbf{Z}$ with fixed $\mathbf{X}^{(i)}$ and $\mathbf{Y}^{(i)}$

$$\mathbf{Z}^{(i+1)} = \arg\min_{\mathbf{Z}} \frac{\mu}{2}\|\mathbf{X}^{(i)} - D(\mathbf{Z}) + \frac{\mathbf{Y}^{(i)}}{\mu}\|_F^2. \tag{6}$$

Solving this problem requires backpropagation through the decoder network D (i.e., computing $\frac{\partial D(\mathbf{Z})}{\partial \mathbf{Z}}$) for each voxel FID in each iteration, which is highly time-consuming, especially for 3D data. To this end, we leveraged the learned projector and reformulated the problem into

$$\mathbf{Z}^{(i+1)} = \arg\min_{\mathbf{Z}} \frac{\mu}{2}\|P\left(\mathbf{X}^{(i)} + \frac{\mathbf{Y}^{(i)}}{\mu}; \hat{\boldsymbol{\theta}}_p\right) - \mathbf{Z}\|_F^2, \tag{7}$$

with P being the learned projector and $P(D(\mathbf{Z})) = \mathbf{Z}$, and the solution can be obtained via a one-step forward pass:

$$\mathbf{Z}^{(i+1)} = P\left(\mathbf{X}^{(i)} + \frac{\mathbf{Y}^{(i)}}{\mu}; \hat{\boldsymbol{\theta}}_p\right). \tag{8}$$

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

$$\mathbf{X}^{(i+1)} = \arg\min_{\mathbf{X}} \|\mathbf{d} - \mathcal{F}_\Omega\{\mathbf{X}\}\|_2^2 + \lambda\|\mathbf{D}_w\mathbf{X}\|_F^2 + \frac{\mu}{2}\|\mathbf{X} - D(\mathbf{Z}^{(i+1)}) + \frac{\mathbf{Y}^{(i)}}{\mu}\|_F^2. \tag{9}$$

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

$$\mathbf{Y}^{(i+1)} = \mathbf{Y}^{(i)} + \mu\left(\mathbf{X}^{(i+1)} - D(\mathbf{Z}^{(i+1)})\right). \tag{10}$$

Subproblem (II) is equivalent to a linear least-squares problem with two quadratic regularization terms which is easy to solve. The stopping criteria for the overall ADMM algorithm can be either a specific iteration number (e.g., 15) or the relative changes between $\mathbf{X}^{(i+1)}$ and $\mathbf{X}^{(i)}$ falling below a threshold (e.g., 10^{-3}), whichever reached first.

C. Computational Complexity Analysis

To further analyze the gain in computational efficiency offered by the proposed method, we compared the computational complexity of our algorithm against the NN-constrained reconstruction in [35]. Both methods utilized an ADMM-based algorithm skeleton. For the subproblems with quadratic regularization, Eq. (9) here and Eq. (10) in [35], both were solved by the same conjugate gradient algorithm, thus the computational complexities are almost identical at this step. Therefore, the main differences lie in the other subproblems. Specifically, the NN-constrained reconstruction solved the following overall problem [35]:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{d} - \mathcal{F}_\Omega\{\mathbf{S}\}\|_2^2 + \lambda_2\|\mathbf{D}_w\bar{\mathbf{B}}\odot\mathbf{S}\|_F^2 + \lambda_1\|D(E(\mathbf{X})) - \mathbf{X}\|_F^2 \quad s.t. \quad \mathbf{B}\odot\mathbf{X} = \mathbf{S}, \tag{11}$$

where $\mathbf{S}$ is an auxiliary variable ($\mathbf{S} = \mathbf{B}\odot\mathbf{X}$) to decouple the linear least-squares and the nonlinear regularization terms, $\bar{\mathbf{B}}$ represents the element-wise conjugation of $\mathbf{B}$, and $D(.)$ and $E(.)$ are the corresponding decoder and encoder networks from the learned representation model. λ_1 and λ_2 are the regularization parameters for the NN-based and spatial regularization terms, respectively. A BFGS-based algorithm was used to solve the corresponding network-constrained subproblem:

$$\mathbf{X}^{(i+1)} = \arg\min_{\mathbf{X}} \lambda_1\|D(E(\mathbf{X})) - \mathbf{X}\|_F^2 + \frac{\mu'}{2}\left\|\mathbf{B}\odot\mathbf{X} - \mathbf{S}^{(i)} + \frac{\mathbf{Y}^{(i)'}}{\mu'}\right\|_F^2, \tag{12}$$

Without loss of generality, we assume that $\mathbf{X} \in \mathbb{R}^{N\times T}$ with N being the number of voxels and T the number of FID points, and the network has l fully-connected layers with weights $\mathbf{W}_l \in \mathbb{R}^{T\times T}$ for each layer. With some element-wise activation functions, the computational complexity for the forward and backward propagation was then $\mathcal{O}(NlT^2)$. Accordingly, the computational complexity for evaluating the cost function in Eq. (12) was $\mathcal{O}(NlT^2)$, and $\mathcal{O}(N(l+1)T^2)$ for the gradient. Thus, the overall complexity for a K-iteration BFGS algorithm was [40]:

$$\begin{aligned}&\mathcal{O}((l+1)KNT^2)\\ &= \mathcal{O}(KNT^2) + \mathcal{O}(KNlT^2) + \mathcal{O}(KN(l+1)T^2)\end{aligned} \tag{13}$$

Meanwhile, RAIISE solves the NN-constrained subproblem with only one forward pass (Eq. (8)), i.e., $\mathcal{O}(NlT^2)$, which is much lower than the original algorithm in [35].

D. Other Implementation Details

We adapted the network structure and training data generation pipelines described in [35]–[37]. Specifically, fully-connected or convolutional autoencoder with a "bottle-neck" design and complex unit and activation function [38], [39] were used for single-TE and multi-TE data, respectively. The projector was designed to have the same architecture as the encoder in the representation network. Well-defined spectral parametric fitting models were used to synthesize all the training and testing data with model parameters randomly sampled from experimentally estimated distributions. More details can be found in [35]–[37]. A total of 100,000 spectra were simulated for each scenario with 80,000 for training and 20,000 for testing. Each of the simulated FID was added to

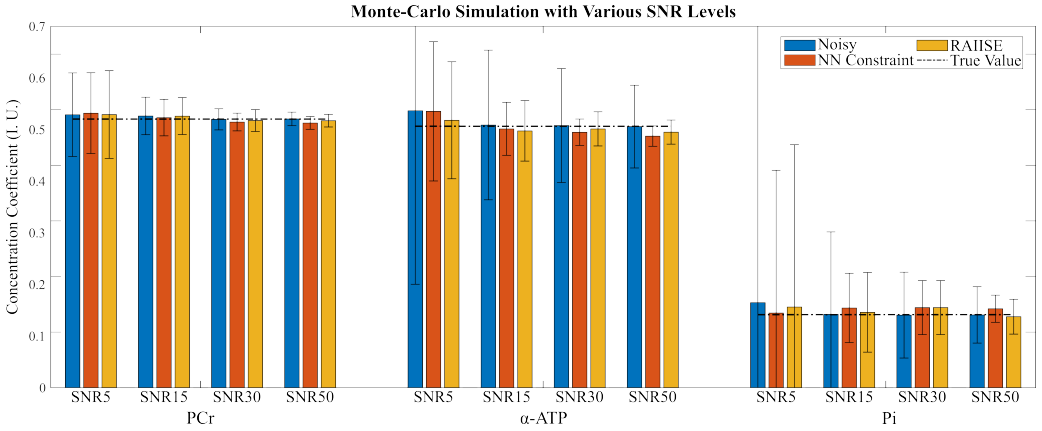

Fig. 3. Monte-Carlo analysis of the denoising effects on metabolite quantification: statistics for the estimated PCr, α-APT, and Pi from the noisy data (Noisy), NN-constrained reconstruction (NN Constraint), and the proposed method (RAIISE) are compared. The true values are shown as dashed lines in the box plots (True Value) and the standard deviations as error bars. Although slight biases were introduced for both NN constraint and RAIISE, significant variance reduction was achieved with denoising, especially for low-concentration metabolites and/or at lower SNRs.

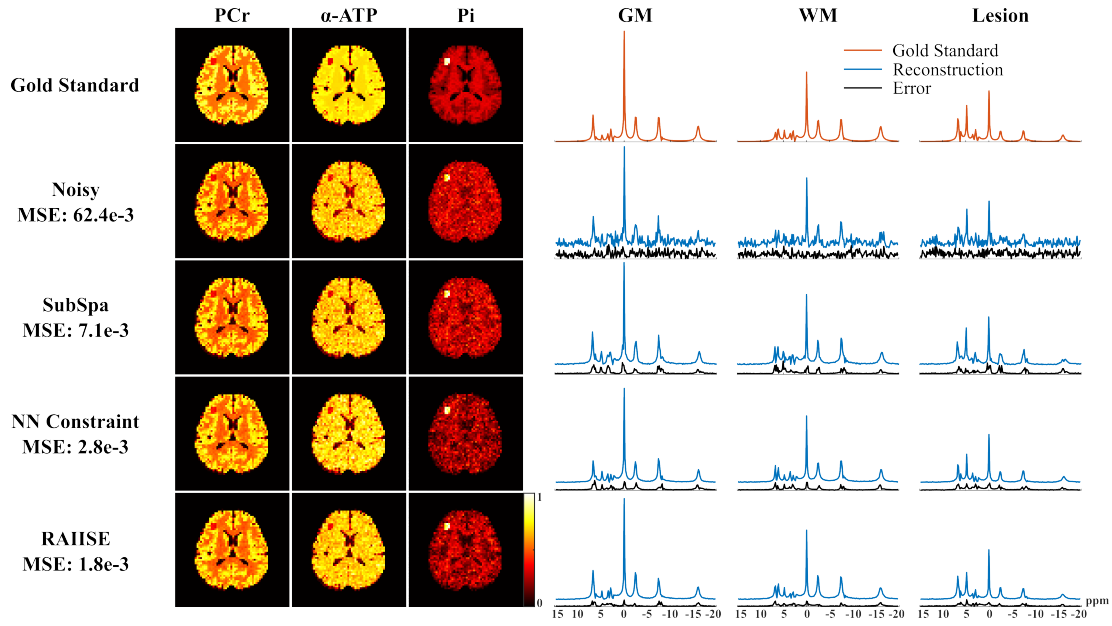

Fig. 4. Spatiospectral reconstruction from the numerical MRSI phantom: The gold standard (Gold Standard), noisy data (Noisy), and results from a subspace-constrained reconstruction (SubSpa) [22], the NN-constrained reconstruction (NN Constraint) [35], and the proposed method (RAIISE) are shown in different rows. MSEs are shown under each method label. The left panel shows metabolite maps of PCr, α-ATP, and Pi, and the right panel compares representative spectra from gray matter (GM), white matter (WM), and the lesion (Lesion) with their corresponding error spectra (in black). All denoising reconstruction methods produced significantly better maps and spectra than noisy data. RAIISE performed slightly better than the NN Constraint method but is dramatically faster, i.e., $\sim$5 mins (RAIISE) vs. $\sim$4.5 hrs of processing time.

different noise realizations with SNR varied between 10 to 100 (SNR defined w.r.t. the highest peak in each spectrum) for the projector training. All network training was performed on a Windows 10 machine with Intel(R) Core(TM) i9-9820X CPU and NVIDIA(R) TITAN RTX(TM) GPU and implemented in PyTorch with Adam optimizer [41]. The batch size was set to 500, and the initial learning rate was 0.001, with 300 epochs, while other parameters remained as default. The same machine was used for the reconstruction.

IV. Simulation and Experimental Settings

A. Simulation

A ^{31}P-MRSI numerical phantom was generated with commonly observed ^{31}P-containing compounds to evaluate the reconstruction performance of the proposed method, where phantom construction details can be found in [35]. In short, brain tissue fractions maps with gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) and spectral parameters were assigned based on experimental data and literature values [42], [43]. Random frequency shifts were simulated to mimic practical conditions. A lesion-like feature was also included. Finally, complex white Gaussian noise was added to the $(\mathbf{r}, t)$-space data to simulate noisy acquisitions.

For the reconstruction comparison, the proposed method, the NN constraint method, as well as a state-of-the-art subspace-constrained reconstruction [22] (with the same spatial constraint) were conducted, where both the nonlinear and linear subspace models were learned with the same

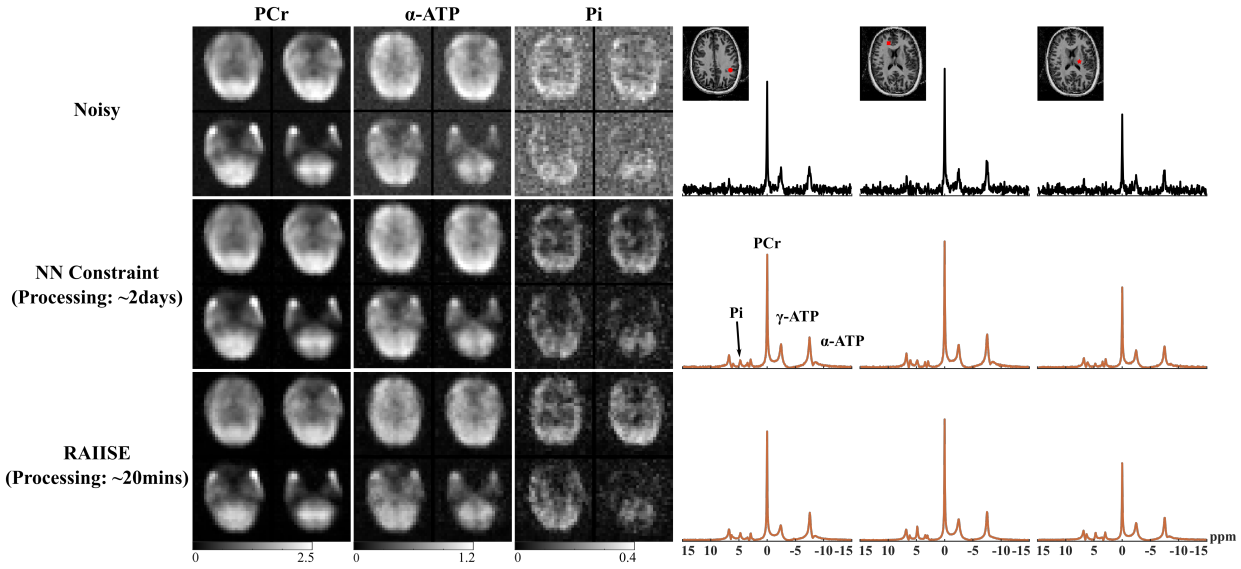

Fig. 5. Spatiospectral reconstructions from in vivo 3D ^{31}P-MRSI data. Results from different methods are shown in different rows. Metabolite maps of PCr, α-ATP, and Pi for selected slices are shown on the first three columns (in institutional unit; left panel), while spatially-resolved spectra from selected voxels (locations marked in the anatomical image inserts) are shown in the subsequent columns (right panel). The SNR improvement is apparent for both the NN Constraint and RAIISE methods, while RAIISE enjoyed a dramatically faster speed, i.e., $\sim$20 mins vs. $\sim$2 days, making it more practically useful for high-resolution volumetric MRSI data.

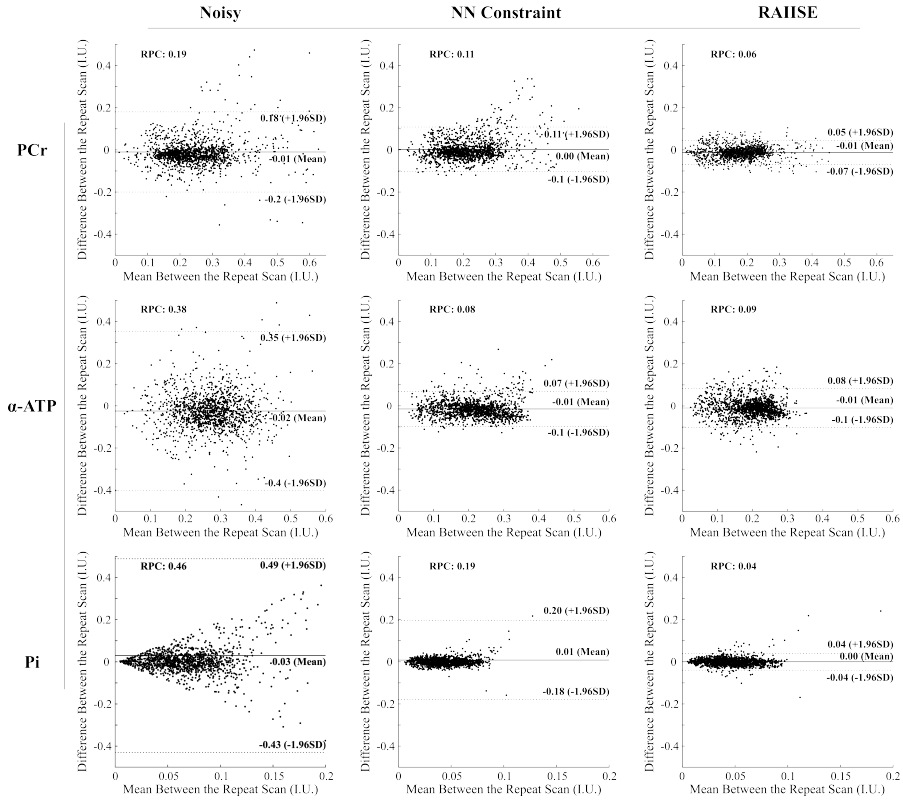

Fig. 6. Bland-Altman analysis of the reconstructions from two repeated experiments. The metabolites were quantified from the reconstruction using a QUEST-based method. Improved consistency between the measurements can be clearly seen for both NN-constrained reconstruction (NN Constraint; second column) and RAIISE (third column), compared to the original noisy data (first column). RAIISE slightly outperformed the NN-constrained reconstruction, particularly for PCr and Pi, as shown by the scatter plots and reproducibility coefficient (RPC in each plot).

training data. The model orders were chosen with the same representation accuracy, 15 for the nonlinear and 32 for the subspace model, respectively (with testing data approximation error around 7%). The regularization parameters for all three methods were fine-tuned to reach the maximum reconstruction accuracy (in terms of MSE) compared to the gold standard.

We also performed a Monte Carlo (MC) simulation to assess the proposed method quantitatively. As performing the

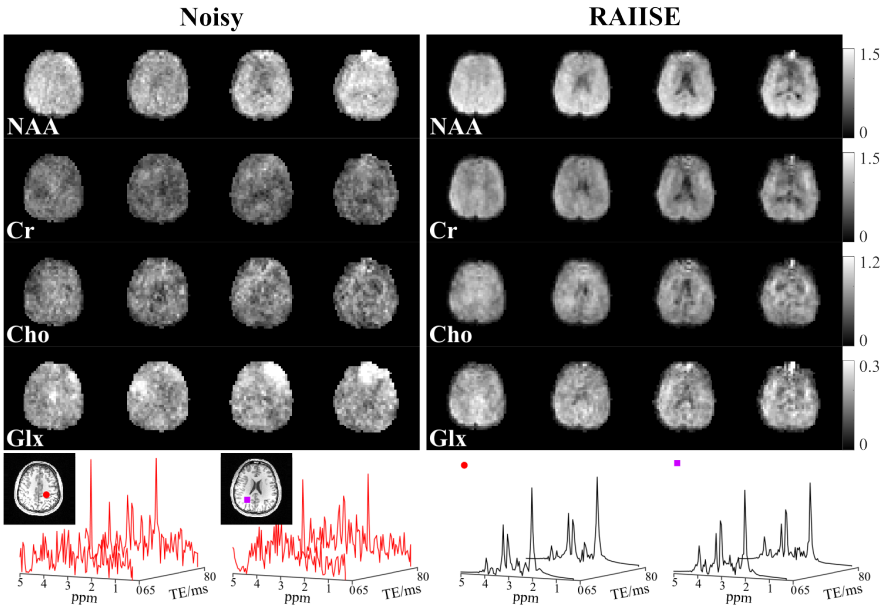

Fig. 7. In vivo ^{1}H-MRSI results from a healthy volunteer: The top panel compares metabolite maps of NAA, Cr, Cho, and Glx from noisy data (left section) and RAIISE (right section). The bottom panel compares spatially-resolved noisy (red curves; left) and reconstructed (black curves; right) spectra, with voxel locations marked by the colored symbols shown in the $\boldsymbol{T_1}$-weighted anatomical image inserts. Significant SNR enhancement and better tissue contrasts can be observed in our reconstruction.

MC study by reconstructing the whole MRSI phantom using the NN-constrained reconstruction in [35] takes too long, we randomly selected a noiseless FID (gold standard) from the numerical phantom and generated 100 noisy realizations at peak SNRs of 5, 15, 30, and 50. Using this single-voxel data, we compared the effects of denoising from different methods on metabolite quantification. A QUEST-based method was used for the quantification here [44].

B. In Vivo Experiments

All in vivo data were acquired with approval from local Institutional Review Boards. ^{31}P-MRSI data were acquired from a healthy volunteer on a Siemens Magnetom 9.4T whole-body imaging system using a double-tuned phased array coil with 8 transceivers/2Rx channels for phosphorus. The acquisition parameters were: TR/TE = 250/1.3 ms, FOV = 180×200×180 mm^3, matrix size = 28×30×13, 16 averages with Hanning weighted k-space averaging, spectral bandwidth = 5000 Hz and 512 FID points [45]. The total acquisition time was about 74 min. Here both the simulation and in vivo ^{31}P-MRSI data shared the same learned nonlinear model (as well as the corresponding projector). For the reconstruction parameters selection, the penalty parameter for the ADMM algorithm μ was kept the same with the simulation, and a very mild spatial regularization parameter was used ($\lambda = 0.03$).

^{1}H-MRSI data were acquired from both healthy volunteers and post-traumatic epilepsy (PTE) patients on a Siemens Prisma 3T system using a multi-TE 3D-EPSI sequence [46] with a 20-channel head coil: TR/TE = 1000/(65, 80) ms, FOV = 220×220×64 mm^3, matrix size = 42×42×8, spectral BW = 1087 Hz and 200 FID points. Total acquisition time was about 14.4 mins. A 60 Hz weak water suppression and outer volume suppression were used for all scans. Before the denoising reconstruction, nuisance water/lipid signals were

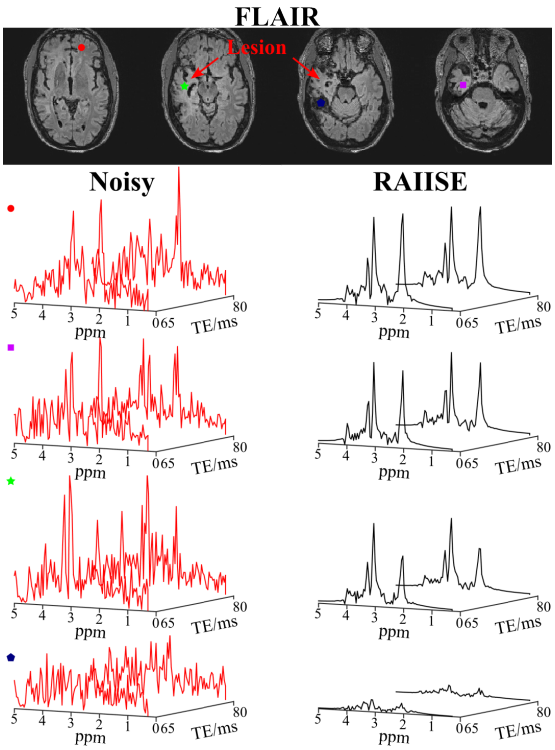

Fig. 8. ^{1}H-MRSI results from a PTE patient: The top panel shows the anatomical images (FLAIR) for different slices across the imaging volume, with a hyperintensity region indicated by the red arrows. The spatially-resolved spectra from RAIISE (black plots; right) exhibit significant SNR enhancements and better differentiate normal-appearing tissues from the lesions than the noisy data (red plots; left). Lesion-specific metabolic profile abnormality buried under noise was better revealed by the proposed reconstruction. The voxel locations were marked by colored symbols.

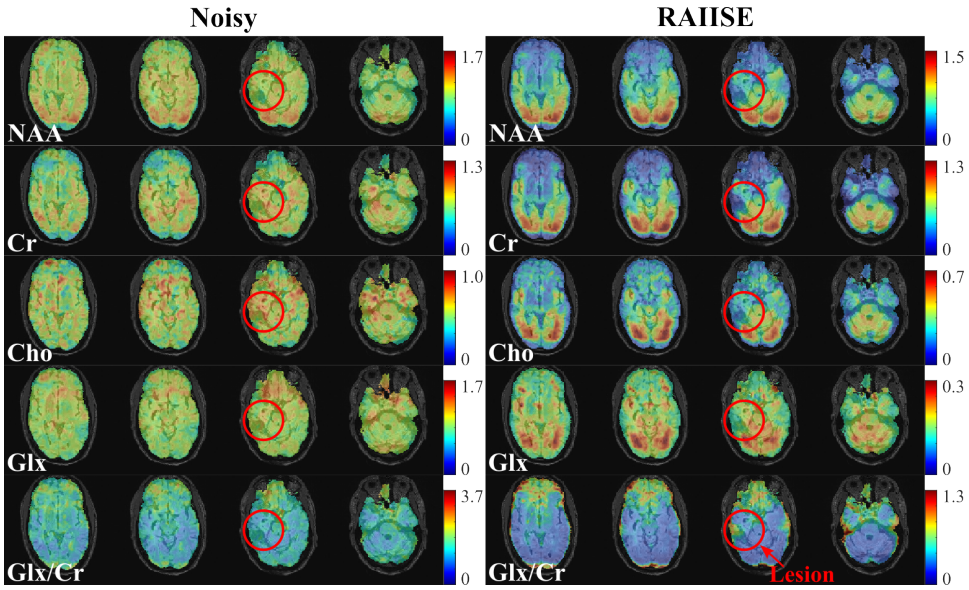

Fig. 9. Metabolite maps, i.e., NAA, Cr, Cho, and Glx produced from the same data in Fig. 8: Results from the noisy data and the proposed method are shown on the left and right panels, respectively. The proposed method obtained Higher-quality metabolite maps better delineating the lesion (red circles), i.e., reduced NAA, reduced Cr, and increased Glx to Cr ratio.

removed using the method in [47]. The same representation model and projector were used here for both healthy and PTE patients. A very mild spatial prior was also introduced ($\lambda = 0.03$) for the reconstruction. All the noisy and denoised data were quantified with a QUEST- and a ProFit-based [48] method for single- and multi-TE data, respectively.

V. Results

A. Simulation Results

Figure. 3 shows MC simulation results, i.e., comparing the means and variances of the estimated metabolite coefficients (PCr, α-ATP, and Pi) from the noisy data, NN-constrained reconstruction using the same representation model, and the proposed method. As can be seen, high fitting accuracy can be achieved with low variance for the noisy measurements when SNR is high, while both NN-constrained reconstruction and the proposed method introduced a small amount of bias. Note that the reconstruction altered the inherent statistics of the noise, making the least-squares estimation biased. This indicates that denoising is not necessarily needed when the SNR is sufficiently high. However, in the low-SNR cases and for lower-concentration metabolites (e.g., ATP and Pi), the two denoising methods achieved significant variance reduction with a small bias compared to the noisy Fourier reconstruction. Similar levels of variance reduction were observed for both NN-constrained reconstruction and the proposed method, while the proposed method attained a dramatically higher reconstruction speed. Figure. 4 shows a set of reconstructions from the ^{31}P-MRSI phantom at SNR = 20 (defined w.r.t. the PCr peak). We compared RAIISE with the reconstruction from the NN constraint method as well as the subspace-constrained reconstruction. As can be seen, significant SNR improvement and higher reconstruction accuracy were achieved by the subspace-based method and both nonlinear-model-based methods compared to noisy data (bottom three rows). But the two nonlinear-model-based methods produced better results compared to the subspace reconstruction, as shown by the MSE, metabolite maps (left panel), and spectral error plots (right panel, fourth and fifth rows). RAIISE obtained slightly better accuracy (fifth row). More importantly, RAIISE is dramatically faster than the previous NN-constrained method on the same machine, i.e., $\sim$5 mins vs. $\sim$4.5 hrs.

B. In Vivo Results

Reconstructions from an in vivo *3D* ^{31}P-MRSI dataset are shown in Fig. 5. It is evident that both the NN-constrained reconstruction (second row) and RAIISE (third row) exhibited substantial SNR improvements compared to the noisy data (first row). This enhancement is particularly pronounced for low-abundance metabolites, such as Pi. Notably, the proposed method achieved significantly faster reconstruction, taking approximately 20 mins compared to approximately 2 days for the original NN-constrained method (which required network inversion), an acceleration of about 100$\times$.

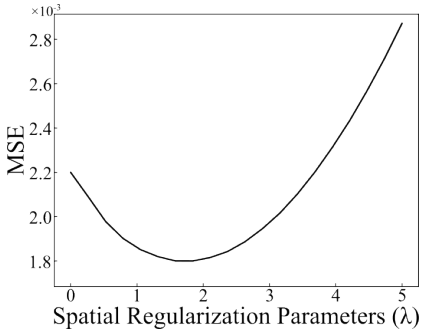

Fig. 10. Reconstruction performance's dependence (in terms of MSE) on the spatial regularization parameter ($\boldsymbol{\lambda}$), evaluated using the simulated ^{31}P-MRSI phantom shown in Fig. 4.

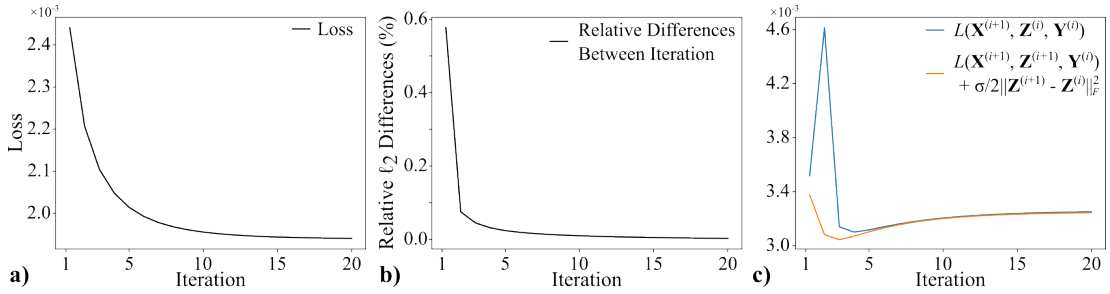

Fig. 11. Convergence analysis: a) Loss function values versus iterations for the MRSI phantom; b) Relative ℓ_2 differences in $\mathbf{X}$ between iterations (i.e., $\|\mathbf{X}^{(i+1)} - \mathbf{X}^{(i)}\|_F / \|\mathbf{X}^{(i)}\|_F$); c) Empirical evidence of the σ-strong convexity ($\sigma = 0.1$) for the first 20 iterations.

The quantitative performance of the proposed method was further evaluated via a test-retest experiment, and the repeatability of the metabolite quantification (PCr, α-APT, and Pi) from two repeated ^{31}P-MRSI scans was demonstrated in Fig. 6. The results demonstrate a significant improvement in the consistency of metabolite estimations from two repeated measurements for both the NN constraint and proposed methods. Particularly, the proposed method yielded slightly better consistency, as indicated by the reduced reproducibility coefficient (RPC) values shown in each subplot.

Spatiospectral reconstruction results from the multi-TE ^{1}H MRSI experiments were shown in Figs. 7 (healthy volunteer), 8, and 9 (PTE patients). Figure 7 compares the metabolite maps for different slices across the imaging volume (top panel) and spatially localized spectra (bottom panel), for the noisy Fourier (left section) and denoised RAIISE reconstructions (right section). As can be seen, the proposed method produced significantly improved SNR with well-preserved spatiospectral features and high-quality metabolite maps, e.g., better GM/WM/CSF contrast. Figures 8 and 9 compare the reconstructions from a PTE patient. Specifically, Figure 8 shows spatially localized spectra of one voxel from anterior gray matter (red circle), two from lesion regions (green star and purple square; selected from the hyper-intensity regions shown in the FLAIR image), and the fourth from a CSF-filled no-tissue region (blue pentagon). Quantified metabolite maps are shown in Fig. 9. Again, the proposed method produced significant SNR improvement with less spectra distortion and overfitting, revealing tissue-specific biochemical features. It also enhanced the quality of metabolite maps with better lesion delineation (see red circles in Fig. 9). Note that this metabolic information can have great potential as additional biomarkers for lesion subtyping, e.g., epileptogenic or not.

VI. Discussion

Our proposed method inherits similar advantages from the previously reported NN-constrained reconstruction method. First, unlike the direct end-to-end network approaches, the learned model aims to capture the inherent variations of the true MRSI signals, thus one trained network can be flexibly used for data acquired at different settings, e.g., resolutions and (k, t)-space sampling patterns (accounted for by the forward model) without re-training. Second, our formulation has the capability to incorporate additional constraints like the spatial regularization introduced in Eq. (4) or even other NN-based priors, e.g., deep image prior [49], which can be an interesting direction to explore in future work. Finally, the complex convolutional autoencoder network design enables a scalable adaption for multi-TE/multi-dimensional MRSI signals than standard fully-connected networks [35].

For the selection of the regularization parameter, Fig. 10 shows the reconstruction performance w.r.t. the value of λ, evaluated using MSE on the numerical ^{31}P-MRSI phantom. The results indicate that the denoising performance remains robust across a relatively large range of parameter values. However, it is worth exploring alternative strategies for parameter selection, including those optimized for metrics beyond MSE and learning-based approaches. Additionally, we used an ℓ_2 regularization for a consistent comparison of algorithms and their complexities. But other forms of regularization, e.g., ℓ_1 or other non-quadratic penalties, can be exploited. We noticed similar performances between our weighted-ℓ_2 and ℓ_1 regularization if the parameter was well chosen.

For theoretical analysis of our algorithm, global convergence of a general ADMM algorithm with nonlinear equality constraints was recently established [50]. Specifically, with some basic assumptions (i.e., [50]; Section 2), the authors laid out three technical conditions for convergence, i.e., Eqs. (6.7) and (6.8) in [50], and an additional σ-strong convexity assumption. In our case, these three conditions are:

$$L(\mathbf{X}^{(i+1)}, \mathbf{Z}^{(i)}, \mathbf{Y}^{(i)}) \leq L(\mathbf{X}^{(i)}, \mathbf{Z}^{(i)}, \mathbf{Y}^{(i)}), \tag{14}$$

$$\nabla_{\mathbf{Z}} L(\mathbf{X}^{(i+1)}, \mathbf{Z}^{(i+1)}, \mathbf{Y}^{(i)}) = 0, \tag{15}$$

and the σ-strong convexity assumption being expressed as:

$$\begin{aligned} L(\mathbf{X}^{(i+1)}, \mathbf{Z}^{(i+1)}, \mathbf{Y}^{(i)}) + \frac{\sigma}{2}\|\mathbf{Z}^{(i+1)} - \mathbf{Z}^{(i)}\|_F^2 \\ \leq L(\mathbf{X}^{(i+1)}, \mathbf{Z}^{(i)}, \mathbf{Y}^{(i)}). \end{aligned} \tag{16}$$

Eq. (14) is automatically satisfied as the $\mathbf{X}$-subproblem is solved with a conjugate gradient method. For the $\mathbf{Z}$-subproblem, assuming that the neural network we used has a sufficient representation power so that we can learn a projector $\mathbf{P}$ such that $\mathbf{P}(\mathbf{X}^{(i)} + \frac{\mathbf{Y}^{(i)}}{\mu}) \in \arg\min_{\mathbf{Z}} \frac{\mu}{2}\|\mathbf{X}^{(i)} + \frac{\mathbf{Y}^{(i)}}{\mu} - D(\mathbf{Z})\|_F^2$, Eq. (15) can be readily satisfied. Proving the σ-strong convexity condition analytically with the deep network is difficult, thus we showed some empirical evidence for Eq. (16) in Fig. 11c using our MRSI phantom with $\sigma = 0.1$. Additionally, we showed empirical convergence using the loss function curve (Fig. 11a) as well as the relative changes in $\mathbf{X}$ (Fig. 11b) throughout the iterations.

VII. Conclusion

We proposed a novel MRSI reconstruction method that features a strategy to jointly learn a nonlinear low-dimensional representation of high-dimensional spectroscopic signals and a projector to recover the low-dimensional embeddings, a formulation combining the forward physical model and the spectral constraint enforced via the learned representation, a complementary spatial constraint, and an efficient ADMM-based algorithm facilitated by the learned projector. SNR and quantification improvements with highly efficient computation were demonstrated by both simulations and in vivo MRSI data acquired at different experimental settings. We expect the proposed method to provide a widely applicable processing tool to enhance the quality of MRSI or other high-dimensional spatiospectral imaging data.

References

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

[2] S. Posse *et al.*, "MR spectroscopic imaging: principles and recent advances," *J Magn Reson Imaging*, vol. 37, pp. 1301–1325, 2013.

[3] A. A. Maudsley *et al.*, "Advanced magnetic resonance spectroscopic neuroimaging: Experts' consensus recommendations," *NMR Biomed*, vol. 34, p. e4309, 2021.

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

[5] S. Hu *et al.*, "Compressed sensing for resolution enhancement of hyperpolarized 13C flyback 3D-MRSI," *J Magn Reson*, vol. 192, pp. 258–264, 2008.

[6] B. L. Burns *et al.*, "Group sparse reconstruction of multi-dimensional spectroscopic imaging in human brain in vivo," *Algorithms*, vol. 7, pp. 276–294, 2014.

[7] 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, pp. 1366–1373, 2015.

[8] P. Cao and E. X. Wu, "Accelerating phase-encoded proton MR spectroscopic imaging by compressed sensing," *J Magn Reson Imaging*, vol. 41, pp. 487–495, 2015.

[9] 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.

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

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

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

[13] Y. Zhang *et al.*, "Magnetic resonance spectroscopy with linear algebraic modeling (SLAM) for higher speed and sensitivity," *J Magn Reson*, vol. 218, pp. 66–76, 2012.

[14] X. Qu *et al.*, "Accelerated NMR spectroscopy with low-rank reconstruction," *Angew Chem Int Ed*, vol. 54, pp. 852–854, 2015.

[15] 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," *MRM*, vol. 81, pp. 2841–2857, 2019.

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

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

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

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

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

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

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

[23] R. Guo *et al.*, "Simultaneous QSM and metabolic imaging of the brain using SPICE: Further improvements in data acquisition and processing," *MRM*, vol. 85, pp. 970–977, 2021.

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

[25] X. Peng *et al.*, "Simultaneous QSM and metabolic imaging of the brain using SPICE," *MRM*, vol. 79, pp. 13–21, 2018.

[26] H. Lee *et al.*, "High resolution hyperpolarized 13C MRSI using SPICE at 9.4 T," *MRM*, vol. 80, pp. 703–710, 2018.

[27] Y. Li *et al.*, "Machine learning-enabled high-resolution dynamic deuterium MR spectroscopic imaging," *IEEE TMI*, vol. 40, pp. 3879–3890, 2021.

[28] Y. Zhao *et al.*, "Fast high-resolution 19F-MRSI of perfluorocarbon nanoemulsions for MRI cell tracking using SPICE with learned subspaces," in *Proc of ISMRM*, 2021.

[29] R. Guo *et al.*, "Fast high-resolution 1H-MRSI of the human brain at 7T," *Proc of ISMRM*, vol. 3713, 2021.

[30] S. P. Kyathanahally *et al.*, "Deep learning approaches for detection and removal of ghosting artifacts in MR spectroscopy," *MRM*, vol. 80, pp. 851–863, 2018.

[31] H. H. Lee and H. Kim, "Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain," *MRM*, vol. 82, pp. 33–48, 2019.

[32] S. Dong *et al.*, "High-resolution magnetic resonance spectroscopic imaging using a multi-encoder attention U-Net with structural and adversarial loss," in *Proc of IEEE EMBC*, 2021, pp. 2891–2895.

[33] Z. Iqbal *et al.*, "Super-resolution 1H magnetic resonance spectroscopic imaging utilizing deep learning," *Front Oncol*, vol. 9, p. 1010, 2019.

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

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

[36] Y. Li *et al.*, "SNR enhancement for multi-TE MRSI using joint low-dimensional model and spatial constraints," *IEEE TBME*, 2022.

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

[38] C. Trabelsi *et al.*, "Deep complex networks," *arXiv:1705.09792*, 2017.

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

[40] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999.

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

[42] D. K. Deelchand *et al.*, "Quantification of in vivo 31P NMR brain spectra using LCModel," *NMR Biomed*, vol. 28, pp. 633–641, 2015.

[43] J. Ren *et al.*, "31P-MRS of healthy human brain: ATP synthesis, metabolite concentrations, ph, and T1 relaxation times," *NMR Biomed*, vol. 28, pp. 1455–1462, 2015.

[44] H. Ratiney *et al.*, "Time-domain quantitation of 1H short echo-time signals: background accommodation," *MAGMA*, vol. 16, p. 284, 2004.

[45] L. Ruhm *et al.*, "3D 31P MRSI of the human brain at 9.4 Tesla: Optimization and quantitative analysis of metabolic images," *MRM*, vol. 86, pp. 2368–2383, 2021.

[46] Z. Wang *et al.*, "High-resolution, 3D multi-TE 1H MRSI using fast spatiospectral encoding and subspace imaging," *MRM*, vol. 87, pp. 1103–1118, 2022.

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

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

[49] K. Gong *et al.*, "Arterial spin labeling MR image denoising and reconstruction using unsupervised deep learning," *NMR Biomed*, vol. 35, p. e4224, 2022.

[50] J. Bolte *et al.*, "Nonconvex lagrangian-based optimization: monitoring schemes and global convergence," *Math Oper Res*, vol. 43, pp. 1210–1232, 2018.