HHS Public Access

Author manuscript

IEEE Signal Process Mag. Author manuscript; available in PMC 2023 August 03.

Published in final edited form as:

IEEE Signal Process Mag. 2023 March ; 40(2): 101–115. doi:10.1109/msp.2022.3203867.

High-Dimensional MR Spatiospectral Imaging by Integrating Physics-Based Modeling and Data-Driven Machine Learning: Current progress and future directions

Fan Lam [Senior Member, IEEE],
Department of Bioengineering, University of Illinois Urbana-Champaign, Urbana, IL, 61801 USA

Beckman Institute for Advanced Science and Technology, Department of Electrical and Computer Engineering and Cancer Center at Illinois, University of Illinois Urbana-Champaign.

Xi Peng [Member, IEEE],
Department of Radiology, Mayo Clinic, Rochester, MN, 55905, USA

Zhi-Pei Liang [Fellow, IEEE]
Beckman Institute for Advanced Science and Technology, Department of Electrical and Computer Engineering and Cancer Center at Illinois, University of Illinois Urbana-Champaign.

Abstract

Magnetic resonance spectroscopic imaging (MRSI) offers a unique molecular window into the physiological and pathological processes in the human body. However, the applications of MRSI have been limited by a number of long-standing technical challenges due to high dimensionality and low signal-to-noise ratio (SNR). Recent technological developments integrating physics-based modeling and data-driven machine learning that exploit unique physical and mathematical properties of MRSI signals have demonstrated impressive performance in addressing these challenges for rapid, high-resolution, quantitative MRSI. This paper provides a systematic review of these progresses in the context of MRSI physics and offers perspectives on promising future directions.

I. Introduction

Magnetic resonance (MR) spatiospectral imaging techniques allow for noninvasive visualization and quantification of molecule-specific physiological processes in living animals and humans that are inaccessible by conventional anatomical and functional imaging methods. MR spectroscopic imaging (MRSI), in particular, integrates the concepts of spatial encoding used in MRI and spectral encoding used in MR spectroscopy (MRS) to produce spatially-resolved one or multi-dimensional spectra. These spectra allow simultaneous detection, quantification, and mapping of numerous endogenous molecules in the human body, providing important insights into the biochemical processes in vivo. This molecular imaging capability, since its inception in 1975 [1], [2], promised to significantly impact many basic science studies (including understanding fundamental

fanlam1@illinois.edu .

physiological processes) and clinical applications (including diagnosis, prognosis and treatment monitoring for cancer, stroke, epilepsy, and many other neurological and psychiatric diseases) [3]–[5]. However, several long-standing technical challenges have hindered the progress and applications of in vivo MRSI, including low sensitivity, poor resolution, low imaging speed, contamination from nuisance signals and sensitivity to system imperfections. The fundamental reasons underlying these challenges are: (1) the inherently low abundance of the molecules of interest in MRSI, e.g., metabolites and neurotransmitters having three to four orders of magnitude lower concentrations than the water molecules imaged by MRI, and (2) the high dimensionality of the imaging problem due to the need to encode and decode both spatial and spectral dimensions with high resolutions. While significant efforts have been devoted to tackle these challenges, the performance of in vivo MRSI still falls short of the signal-to-noise ratio (SNR), resolution, speed, and robustness desired or required by many practical applications.

Recent advances in ultrahigh-field systems, high-sensitivity radiofrequency (RF) receiver arrays, and computational imaging methods have presented new opportunities to address the technological challenges associated with in vivo MRSI, reinvigorating this “old” but relatively unexploited field. Recent years have observed a substantial growth in the efforts pushing towards rapid and high-resolution MRSI (millimeter resolution on humans instead of the conventional centimeter resolution) [6]. Particularly, computational methods that exploit the rich prior information for spectroscopy signals derived from MR physics have been developed to address several key issues on recovering high-dimensional spatiospectral functions from limited and/or noisy data [7]–[11], and quantifying molecular parameters (e.g., concentrations, T2s) from the spatiospectral function to derive quantitative biomarkers for physiological functions and diseases [12]–[14]. We present, in this paper, a systematic review of these recent progresses in the context of MRSI physics and related imaging problems, and provide our perspectives on directions and opportunities for future pursuit. We expect that further technology developments that effectively integrate spin physics and machine learning have the potential to transform MRSI from a slow, low-SNR and poor-resolution modality into a practical, high-resolution, in vivo molecular and metabolic imaging tool for many scientific and clinical applications. These advancements may also benefit and inspire innovations in other high-dimensional spatiospectral imaging modalities.

This review is organized as follows. Section II introduces basic physical principles underlying MRSI and defines the key challenges associated with in vivo high-dimensional MRSI. Section III reviews physics-based modeling of spectroscopy signals and data-driven low-dimensional model learning approaches inspired by the physics models. Section IV discusses spatiospectral image reconstruction methods integrating physics-based and learned models. Section V extends the discussion to dynamic MRSI followed by Section VI where integrated physics-based and machine learning strategies for metabolite quantification are reviewed. Finally, Section VII presents our perspectives on future research directions.

II. MRSI Physics

Atomic nuclei with an odd number of protons and/or an odd number of neutrons possess a physical property, known as nuclear spin, which gives rise to a nonzero magnetic moment

and thus detectable MR signals after an RF excitation. The nucleus of the hydrogen atom is the simplest in nature, consisting of just one proton and no neutron. Because hydrogen is the most common element found in the human body, proton MRI has been widely used for anatomical and functional imaging using mainly the MR signals from water protons (^{1}H). MRSI, on the other hand, is designed to measure MR signals from ^{1}H and/or other nuclei (e.g., ^{31}P) in a range of molecules (e.g., N-Acetyle Aspartate, Choline, Creatine, etc). These signals are properly encoded in data acquisition and decoded in data processing to achieve multiplexed molecular imaging. In this section, we will first review the physical mechanisms underlying the generation of different resonance frequencies and the spatiospectral encoding strategies used in MRSI. Then we will introduce the resulting high-dimensional spatiospectral imaging problem and the unique challenges associated with MRSI.

A. Chemical Shift and MRSI

In a typical MRI experiment, the object of interest is commonly represented as a spatial function $\rho(\mathbf{r})$, whose values at a spatial location $\mathbf{r}$ depend on the abundance of water molecules (or often referred to as spin density) and a few other biophysical parameters of tissue water, e.g., T_1, T_2, and χ (susceptibility). A detailed discussion on this can be found in [15]. However, there can be many different molecules other than water present in the object. Furthermore, one molecule can have the same nucleus (e.g., ^{1}H's) in different functional groups, which are surrounded by varying numbers of orbiting electrons. These orbiting electrons "locally" perturb the magnetic field experienced by the nuclei. This effect, called electron shielding, illustrated in Fig. 1(a), makes different nuclei in the same or different molecules resonate at a range of frequencies specified by

$$f = \gamma B_0(1-\sigma), \tag{1}$$

where B_0 is the strength of the main magnetic field that the object experiences, γ is the gyromagnetic ratio (only nucleus dependent) and σ is the electron shielding constant (dependent on nuclei, molecules as well as the position of a nucleus in a molecule). This frequency dispersion gives rise to the chemical shift phenomenon fundamental to MR spectroscopy. As an example, different protons in N-acetylaspartate (NAA, one of the most abundant metabolites in the brain) exhibit several frequencies, producing a unique resonance structure (Fig. 1(b)). Different molecules have their unique resonance structures.

Considering the chemical shift distribution in each imaging voxel, the image function of interest becomes a spatiospectral function $\rho(\mathbf{r}, f)$. Mapping this function allows one to obtain spatially-resolved spectra, from which we can measure various molecules and quantify their relative abundance, providing useful insight into the physiological conditions of the tissues/organs of interest.

In practical MRSI acquisition with pulse excitations, we use a radiofrequency (RF) coil to pick up time-domain signals called free induction decays (FIDs). Therefore, we often use the Fourier counterpart of $\rho(\mathbf{r}, f)$, i.e., $\tilde{\rho}(\mathbf{r}, t)$, to represent the spatiotemporal function of interest

(where t denotes the FID dimension, sometimes referred to as the spin clock). Thus, the imaging data collected is in (k,t)-space, and related to $\tilde{\rho}(\mathbf{r},t)$ by

$$s(\mathbf{k}_p,t_q)=\int_V \tilde{\rho}(\mathbf{r},t)e^{-i2\pi\gamma\Delta f(\mathbf{r})t_q}e^{-i2\pi\mathbf{k}_p\mathbf{r}}d\mathbf{r}+\epsilon(\mathbf{k}_p,t_q), \tag{2}$$

where V denotes the imaging volume of interest, $\Delta f(\mathbf{r})$ denotes the macroscopic B_0 field inhomogeneity distribution (due to system imperfection and object induced magnetic field perturbation), $\mathbf{k}_p=(p_x\Delta k_x, p_y\Delta k_y, p_z\Delta k_z)$ and $t_q=q\Delta t$ denote the sampling location in the high-dimensional (k,t)-space respectively (Δk_x, Δk_y, Δk_z and Δt are the corresponding sampling intervals), and $\epsilon(\cdot)$ represent the measurement noise (modeled as complex white Gaussian). Based on Eq. (2), the MR spectroscopic imaging problem is to recover $\tilde{\rho}(\mathbf{r},t)$ from a set of (k,t)-space measurement $\{s(\mathbf{k}_p,t_q)\}_{p=1,q=1}^{P,Q}$, also known as the spatiospectral encodings. Coil sensitivity can be included in Eq. (2) such that $s(\cdot)$ becomes $s_c(\cdot)$. Without loss of generality, we ignore the coil index in the following discussion. An illustration of the spatiospectral imaging problem is shown in Fig. 2.

B. Challenges for MRSI

As shown in Eq. (2) and Fig. 2, MRSI is a high-dimensional imaging problem due to the need to encode and decode both spatial and spectral information. Because Δt needs to be small enough to satisfy the Nyquist sampling requirement along the spectral dimension (i.e., 1 to 2 kHz bandwidth for 3T and even higher for ultrahigh-field systems), it is not realistic to cover extended k-space during each readout. Therefore, a standard approach is to acquire all the time points needed for a single or a few k-space locations after each excitation and repeat the process many times. Furthermore, many FID (t) points (on the order of hundreds) need to be acquired to achieve sufficient spectral resolutions for accurately differentiating signals from different molecules. As a result, the imaging time for an MRSI experiment is long and grows exponentially as the desired spatiospectral resolution increases in the conventional Fourier imaging paradigm.

Another fundamental challenge with MRSI is low SNR. The MR signal from each molecule is proportional to the bulk magnetization $M_{z,m}^0$ given by [15]

$$M_{z,m}^0=\frac{\gamma^2\hbar^2N_m}{4KT_s} \tag{3}$$

where N_m is the number of polarized spins that can be used to generate detectable signals for the mth molecules, $\hbar$ is the Planck constant, K the Boltzmann constant and T_s the absolute temperature. As N_m for the molecules of interest in MRSI, e.g., metabolites such as NAA, creatine (Cr) and choline (Cho) and neurotransmitters such as glutamate (Glu) and GABA, are typically three orders of magnitude smaller than that for water, much weaker signals from these molecules are expected. As a result, a common practice is to prescribe large voxel sizes (poor spatial resolution) in order to maintain sufficient SNRs within a clinically feasible scan time. A related important issue (which is often less discussed) is the large dynamic range for signals from different molecular components due

to their concentration differences. An outstanding example is the strong nuisance water and subcutaneous lipid signals in ^{1}H-MRSI, which can significantly affect our ability to reliably reconstruct and quantify metabolite signals of interest. Therefore, strong prior information is needed to address the dimensionality, SNR and dynamic range challenges to enable fast, high-resolution, high-SNR MRSI for practical applications. The well-established chemical and physical knowledge for the spectroscopic signals and advanced data-driven machine learning tools to leverage this knowledge present new opportunities to tackle these challenges.

III. Physics-Based Spatial-Spectral Priors

Strong physics-based priors can be imposed on $\tilde{\rho}(\mathbf{r}, t)$ for in vivo MRSI based on extensive studies on the signal characteristics of the molecules of interest. Specifically, each time/frequency point needs not be treated as an independent unknown during the imaging process. Given that there are only a finite number of frequencies originating from NMR-sensitive molecules in vivo [4], one straightforward choice is to model the signals as a linear combination of complex exponentials, each capturing a particular frequency

$$\rho(t) = \sum_{k=1}^{K} c_k e^{-\beta_k t} \tag{4}$$

where $\beta_k = a_k + ib_k$ is a complex coefficient modeling both the decaying behaviors of FIDs and the frequency values. While this model has often been used in the analysis of solution MR data in chemistry studies [11], [16], its utility in in vivo MRSI has been limited, because (1) each molecule can have a number of frequency peaks and there are many different molecules that can be simultaneously detected, thus K can be large making this constraint less effective, especially for ^{1}H−MRSI; (2) the frequency components from the same molecule have strong dependence according to quantum mechanics (QM) [4]. Therefore, a better model for in vivo MRSI data at each spatial voxel can be introduced, i.e.,

$$\tilde{\rho}(\mathbf{r}, t) = \sum_{m=1}^{M} c_m(\mathbf{r}) \phi_m(t) e^{-t/T_{2,m}^*(\mathbf{r}) + i2\pi\delta f_m(\mathbf{r})t} \tag{5}$$

where c_m are the molecular concentrations and ϕ_m denote a basis whose Fourier transform is the spectral structure of a specific molecule; ϕ_m can be predicted by QM simulations provided the chemical shifts and J-coupling constants for a particular molecule (from chemical database) and is experiment independent. $T_{2,m}^*$ and δf_m are experiment-dependent and molecule-specific apparent relaxation parameters and additional frequency shifts. As can be seen, the model order reduces from K (the number of frequencies) to M the number of molecules, a significantly smaller degrees of freedom. While one can attempt to directly estimate $c_m(\mathbf{r}), T_{2,m}^*(\mathbf{r})$ and $\delta f_m(\mathbf{r})$ from the noisy and often limited $s(\mathbf{k}, t)$, the strong nonlinearity of the estimation problem requires a high SNR that is usually not available for in vivo MRSI, thus limiting the achievable spatial resolution. However, the physics-based model in Eq. (5) underlies the inherent variations of spectroscopic signals that motivate the development of low-dimensional modeling and machine learning approaches to enable

better tradeoffs in speed, resolution, and SNR for MRSI. We review these developments in the following sections.

A. Subspace Model

Under physiological conditions, the parameters $c_m, T^*_{2,m}(\mathbf{r})$ and $\delta f_m(\mathbf{r})$ have values in a narrow range (e.g., $T^*_{2,m}$~30 to 100ms), and M is usually small, i.e., 10–15 for in vivo ^{1}H-MRSI. Therefore, the physiologically meaningful high-dimensional FIDs/spectra we acquired should reside in or close to a low-dimensional space. One approach to exploiting this low-dimensionality is to approximate the high-dimensional $\tilde{\rho}(\mathbf{r}, t)$ using the partial separability (PS) model [7], [8], [17]:

$$\tilde{\rho}(\mathbf{r}, t) = \sum_{l=1}^{L} u_l(\mathbf{r}) v_l(t) \tag{6}$$

which represent the FIDs/spectra at individual voxels as linear combinations of a small set of basis functions $\{v_l(t)\}$, thus in a low-dimensional subspace. The PS model leads to low-rank structures for the discretized representation of $\tilde{\rho}(\mathbf{r}, t)$, which has been used for denoising [18] or super-resolution reconstruction [19]. Such a low-rank filtering approach requires joint estimation of both the spatial coefficients $\{u_l(\mathbf{r})\}$ and subspace basis $\{v_l(t)\}$ (subspace pursuit). Subspace pursuit often requires relatively high SNR data which is well beyond that of in vivo MRSI data acquired using accelerated acquisition schemes. Alternatively, we can leverage physics-based priors to pre-learn $\{v_l\}$ from high-SNR training data and transform the MRSI problem into the recovery of $\{u_l\}$ which is significantly lower-dimensional than $\tilde{\rho}(\mathbf{r}, t)$, thus enabling more flexible designs of data acquisition and physics-motivated low-rank reconstruction to address the speed, resolution and SNR (SRS) challenges. We will focus on this type of reconstruction methods in this review and more detailed discussions on data acquisition can be found in [20]–[22].

While subspace learning has been well studied in machine learning [23], the unique physical properties of each FID allows for unique training data generation and learning strategies [8]. More specifically, we can draw samples from pre-specified distributions of $c_m, T^*_{2,m}$ and δf_m and synthesize a large number of FIDs from which the subspace structure can be determined, e.g., through PCA, ICA or other dimensionality reduction techniques. The parameter distributions can be specified using literature data with empirical distribution assumptions (e.g.,Gaussian distributions with lower and upper bounds) [24], [25]. In addition, spectral fitting using Eq. (5) or its variants can be applied to experimentally acquired high-SNR training data to extract empirical distributions of the molecular parameters [8], [12] or to supplement the synthetic data. The subspace learned this way incorporates strong physics prior, captures the inherent molecular spectral features of interest, and is less susceptible to experimental artifacts than a direct application of dimensionality reduction to noisy experimental data.

B. Nonlinear Manifold Model

While linear subspace models have been demonstrated highly effective and enables flexible sampling of (k,t)-space for the determination of spatial coefficients. Their effectiveness degrades when the range of spectral parameters increases, which motivates more general models. This is illustrated in Fig. 3. We generated spectra of glutamate (Glu) with varying ranges of c, T_2^* and δf, and compared the representation accuracy between linear subspace and a nonlinear manifold model [26]. As can be seen, as the ranges of T_2^* and δf increase, the approximation errors of subspace increases, as shown by the error spectra in black, while the nonlinear model maintains high accuracy for the same order. Therefore, learning a nonlinear low-dimensional model can offer a more efficient representation of general in vivo MR spectra. As large quantities of training samples can be generated using the physics-based method described above, data-driven learning of nonlinear low-dimensional representations is possible, leveraging recent progresses in deep learning.

Nonlinear dimensionality reduction has been used for classifying single-voxel spectroscopy (SVS) data and demonstrated superior performance than linear dimensionality reduction [27], [28]. But these works did not focus on representation accuracy and did not consider the use of low-dimensional representations for image reconstruction. A deep autoencoder (DAE) based approach was proposed recently to learn an accurate low-dimensional representation for ensembles of spectroscopic signals (FIDs) [26]. Denoting each voxel FID as a vector $\boldsymbol{\rho}$, an encoder $E(\cdot;\theta_e)$ and a decoder $D(\cdot;\theta_d)$ network can be learned such that a low-dimensional embedding encoded from $E(\cdot)$ can accurately reconstruct ρ, i.e., $\rho \approx D(E(\boldsymbol{\rho};\theta_e);\theta_d)$. The idea has been extended to multi-echo-time (multi-TE) spectroscopy data [29], and the improved representation efficiency of learned nonlinear low-dimensional models over linear subspace model has been shown, for not only ^{1}H but also other nuclei data [25], [26], [29]. Figure 4 provides a conceptual illustration of physics-model-based training data generation incorporating QM priors and empirical distributions of spectral parameters, and data-driven learning of low-dimensional representations for high-dimensional spectroscopic signals.

These learned low-dimensional models enabled new ways to formulate the recovery of the high-dimensional image function (a.k.a. spatiospectral reconstruction) from noisy or often sparsely sampled spatiospectral encodings, which are reviewed in the next sections.

IV. Spatiospectral Reconstruction Using Learned Models

A. Subspace-Based Reconstruction

Using the discretized representation, i.e., representing $\tilde{\rho}(\mathbf{r},t)$ at a set of sampled locations $\{\mathbf{r}_n, t_{q'}\}_{n,q'=1}^{N,Q'}$, the subspace model in Eq. (6) can be rewritten in a matrix form $\rho = \mathbf{U}\widehat{\mathbf{V}}$ [17], where $\mathbf{U} \in C^{N\times L}$ and $\widehat{\mathbf{V}} \in C^{L\times Q'}$ are matrix notations for the spatial coefficients and learned subspace. Accordingly, the reconstruction problem can be formulated as estimating the unknown spatial coefficients (with a significantly less number of degrees-of-freedom than) [17]

$$\widehat{\mathbf{U}} = \arg\min_{\mathbf{U}} \| \mathbf{d} - \Omega\left\{\mathbf{FB} \odot \left(\mathbf{U}\widehat{\mathbf{V}}\right)\right\} \|_2^2 + \lambda R\left(\mathbf{U}, \widehat{\mathbf{V}}\right), \tag{7}$$

where $\mathbf{B}$ captures the B_0 field inhomogeneity induced linear phases (i.e., $\mathbf{B}_{nq'} = e^{i2\pi\Delta f(\mathbf{r}_n)t_q'}$), $\mathbf{F}$ is a spatial Fourier transform operator and Ω denotes a (k,t)-space sampling operator. Note that q' can be different from q as truncated FID sampling can be easily implemented. The vector $\mathbf{d}$ contains all the noisy spatiospectral encodings. $R(\cdot)$ is a regularization that can be used to impose spatial constraint on $\mathbf{U}$ or the images at individual FID time point, e.g., an edge preserving penalty [20] or sparsity constraint [30], and λ is the regularization parameter. The final reconstruction $\hat{\rho}$ can be formed as $\hat{\rho} = \widehat{\mathbf{U}}\widehat{\mathbf{V}}$. This joint subspace and spatial constrained reconstruction has achieved impressive improvement in resolution, SNR and speed for both ^{1}H and X-nuclei MRSI in brain and muscle applications over traditional methods [9], [20], [31].

It is possible to extend Eq. (7) to jointly update both $\mathbf{U}$ and $\mathbf{V}$ [30], which is a special form of low-rank matrix recovery. This type of generic low-rank-model-based methods have been successful in several MRI applications, e.g., dynamic imaging [32], [33], MR relaxometry [34] and hyperpolarized ^{13}C-MRSI [35], where SNR is not a major bottleneck. However, they do not work well when using MRSI data with very low SNRs for subspace pursuit, and are not flexible for incorporatingfuture development physics-motivated subspace learning. Substantial bias may be present in the reconstructed spectra due to errors of subspace estimation from very noisy data, especially for less dominant but important spatiospectral features [20]. A comparison of reconstructions from simulated MRSI data produced by direct low-rank filtering and using a learned subspace is shown in Fig. 5 to demonstrate this effect. Locally low-rank models offer better representation accuracy and capability than global low-rank models [33], [34], [36], [37], but are still limited by the SNR challenge and subspace estimation errors in MRSI applications.

B. Nonlinear Manifold Constrained Reconstruction

As discussed above, generalizing the linear subspace model to nonlinear low-dimensional manifold models can enable more accurate approximation of general spectroscopic signal variations, thus reducing potential modeling errors for diverse physiological and pathological conditions. A key issue is how to integrate such a learned nonlinear model into the reconstruction formulation. One approach is to incorporate a “network representation error” penalty term into a regularized reconstruction formalism:

$$\widehat{\boldsymbol{\rho}} = \arg\min_{\rho} \| \mathbf{d} - \Omega\{\mathbf{FB} \odot \boldsymbol{\rho}\} \|_2^2 + \lambda_1 \sum_{n=1}^{N} \|\mathscr{C}(\boldsymbol{\rho}_n) - \boldsymbol{\rho}_n\|_2^2 + \lambda_2 R(\boldsymbol{\rho}), \tag{8}$$

with $\mathscr{C}(\cdot)$ representing the learned DAE $D\left(E\left(\cdot;\hat{\theta}_e\right);\hat{\theta}_d\right)$. While the first term imposes data consistency same as previous formulations, the second term enforces the prior that the desired FIDs of interest at each voxel, $\boldsymbol{\rho}_n$, should yield small representation errors for the learned low-dimensional model captured by $\mathscr{C}$. If $\mathscr{C}$ reduces to a linear network, this regularization can be viewed as penalizing the error for projection onto a learned subspace,

a softened version of the explicit subspace model in Eqs. (7) – (12). $R(\cdot)$ is an additional regularization term that can be flexibly incorporated and used to impose spatial constraints, either hand-crafted or learned (see the next section for more discussion). Similar to the reconstruction strategies discussed above for the learned subspace, because we kept the forward encoding model and the learned $\mathscr{C}$ is imposing a prior on the underlying true FIDs, once pre-trained, C can be used synergically with different acquisition designs. This is an important difference with another popular unrolling approach, in which an unrolled network is motivated from but does not exactly solve a regularized least-squares problem (the regularization term can not be explicitly defined) and the network typically needs to be retrained for different acquisition schemes. However, the disadvantage of actually solving the optimization problem in Eq. (8) is the need to invert the deep network (involving backpropagation) for each reconstruction, thus higher computational cost compared to the trained end-toend mapping network that only requires forward pass at the inference stage. Figure 6 provides an example to demonstrate the superior denoising reconstruction performance by using the learned nonlinear model (over linear subspace) for an in vivo ^{1}H-MRSI data.

To effectively take advantage of the learned nonlinear model without significantly compromising computational efficiency, projected gradient descent (PGD) based algorithms can be considered. For example, the recently proposed RAIISE method (LeaRning nonlineAr representatIon and projectIon for faSt constrained MRSI rEconstruction) [38] seeks to learn a projection network to extract low-dimensional embedding from high-dimensional, noisy FIDs and use it in an accelerated PGD algorithm. More specifically, the network-constrained reconstruction can be formulated as

$$\widehat{\boldsymbol{\rho}} = \arg \min_{\boldsymbol{\rho} \in D(\mathbf{Z})} \| \mathbf{d} - \Omega\{\mathbf{FB} \odot \boldsymbol{\rho}\} \|_2^2 + \lambda R(\boldsymbol{\rho}), \tag{9}$$

where $\rho \in D(\mathbf{Z}; \boldsymbol{\theta}_d)$ enforces the prior that the underlying signals should yield a low-dimensional representation $\mathbf{Z}$ (residing on a low-dimensional manifold) and $D(\cdot; \boldsymbol{\theta}_d)$ is the decoder from the representation network $\mathscr{C}$. The PGD update step can be realized via a projection operator $\text{Proj}(\cdot) := D(P(\cdot; \boldsymbol{\theta}_p); \boldsymbol{\theta}_d)$, where $P(\cdot)$ is a learned projector that recovers the latent representations from noisy FIDs and trained as follows:

$$\widehat{\boldsymbol{\theta}}_p = \arg \min_{\boldsymbol{\theta}_p} \frac{1}{J} \sum_{j=1}^{J} \epsilon_1(E(\mathbf{x}_j; \boldsymbol{\theta}_e), P(\tilde{\mathbf{x}}_j; \boldsymbol{\theta}_p)) + \lambda \epsilon_2(\mathbf{x}_j, D(P(\tilde{\mathbf{x}}_j; \boldsymbol{\theta}_p); \boldsymbol{\theta}_d)) \tag{10}$$

where $\mathbf{x}_j$ and $\tilde{\mathbf{x}}_j$ are individual training FIDs and their noisy counterparts, respectively. A range of SNRs were considered to generate $\tilde{\mathbf{x}}_j$ for training $P(\cdot; \theta_p)$ parameterized by θ_p, ϵ_1 and ϵ_2 assess the "projection" errors for the low-dimensional features as well as the full signals, respectively, and λ balances the two losses. The encoder $E(\cdot)$ and decoder D(.)are from the representation network discussed above.The projector can be trained with different structures adapted to the data characteristics (single or multi-TE FIDs), and ϵ_1 and ϵ_2 can be chosen separately. It has been demonstrated that leveraging GPU acceleration RAIISE achieved similar or slightly better denoising performance for both single-TE and multi-TE

MRSI with dramatically improved computation time compared to directly minimizing $\|\mathscr{C}(\boldsymbol{\rho}_n) - \boldsymbol{\rho}_n\|_2^2$ in the regularization formulation of Eq. (8) [38].

C. Integrating Learned Spatial and Spectral Priors

While the effectiveness of learned spectral priors incorporating both physics modeling and machine learning have been well documented, the investigations into learned spatial priors have been scarce. One of the major challenges is the difficulty in acquiring high-resolution, high-SNR MRS images as the "gold standard", for either supervised training as commonly done in deep-learning-based MRI reconstruction or self-supervised representation learning. One solution is to simulate spatial distributions of spectra by combining anatomical images and metabolite concentrations and lineshapes from tissue/brain-regionspecific literature values and/or experimental data. For example, a dense U-Net trained completely by synthetic data was used to produce high-resolution metabolite maps from the low-resolution counterpart plus a T1w MRI [10]. The main issue, however, is that the synthetic spatial distributions of metabolites may not capture sufficient variations in real data, especially those related to different pathological conditions.

An alternative approach is to construct generic network-based image representations that do not require high-SNR, high-resolution data for supervised training of image mapping. Recently, efforts have been made to combine subspace constraint and generative network based image representation. The work in [39] used deep image prior (DIP) to model the spatial function $\mathbf{U}$ as $\mathbf{U} = f(\mathbf{Z}; \boldsymbol{\theta})$, where $f(\cdot\,; \boldsymbol{\theta})$ is a network with trainable parameters $\boldsymbol{\theta}$, the reconstruction in Eq. (7) is then changed to

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \| \mathbf{d} - \Omega\left\{\mathbf{FB} \odot \left(f(\mathbf{Z}; \boldsymbol{\theta})\widehat{\mathbf{V}}\right)\right\} \|_2^2 \,. \tag{11}$$

With a chosen network architecture (e.g., U-net in [39]), the parameters $\boldsymbol{\theta}$ are estimated from noisy data with $\mathbf{Z}$ being an anatomical image. This method assumes that the network can account for the contrast difference between the desired spatial coefficient maps and an MR image to leverage high-resolution anatomical prior information. However, DIP can overfit noise by allowing all the network parameters to be updated and anatomical image being the input, the network may introduce undesirable bias into the spatial reconstruction. Thus, choosing a proper early stopping criterion to balance the two factors may be challenging [40]. Another strategy is to leverage a generative network model pretrained using anatomical, multicontrast MRIs with sufficient representation power for images with different contrasts, and solve for the latent variable $\mathbf{Z}$ instead of network parameters $\boldsymbol{\theta}$ during reconstruction. As $\mathbf{Z}$ is usually lower dimensional than the voxel-wise image representation, noise reduction is possible. For example, a GAN-regularized subspace reconstruction can be formulated as [41]

$$\widehat{\mathbf{U}}, \widehat{\mathbf{Z}} = \arg\min_{\mathbf{U}, \mathbf{Z}} \| \mathbf{d} - \Omega\left\{\mathbf{FB} \odot \left(\mathbf{U}\widehat{\mathbf{V}}\right)\right\} \|_2^2 + \lambda \| \mathbf{U} - f(\mathbf{Z}; \theta) \|_F^2 \,. \tag{12}$$

where $f(\mathbf{Z}; \boldsymbol{\theta})$ is a pretrained GAN (e.g., using T1 and T2w images). This problem can be solved using alternating minimization and the step for updating $\mathbf{Z}$ is a GAN inversion

problem, which is an active research topic at the intersection of inverse problem and deep learning. While different GAN designs can be considered for $f(\cdot;\theta)$, StyleGAN combined with intermediate layer optimization (ILO) has been shown to achieve the best inversion performance [42]. A more comprehensive review on this topic can be found in [43]. One important feature of these reconstruction methods is the integration of the physical spatiospectral encoding model and deep learning priors, in contrast to training an end-to-end network for direct mapping from data/noisy images to reconstructions. As a result, the pretrained model can work for any sampling pattern, sequence parameters, and different resolution and SNRs, without the need of retraining for each case. Meanwhile, the computation time required for the network parameter update or GAN inversion is one limitation/consideration.

V. Dynamic MRSI using learned models

Typical MRSI experiments produce spatially-resolved spectra that reflect a "steady-state" molecular profile, which may not be sufficient to capture the complex in vivo metabolic activities that are time varying in nature. To this end, dynamic MRSI, by introducing an additional time axis (world clock) and producing time-resolved spectra, offers the opportunity to noninvasively visualize metabolism in real time and thus significantly richer biological insights. For example, dynamic ^{31}P-MRSI has been used to study depletion and resynthesis of phosphocreatine (PCr) during exercise-recovery or ischemiareperfusion to assess mitochondrial function in the muscle [44], [45]. Dynamic ^{13}C - and ^{2}H-MRSI, using isotope-enriched endogenous molecules, have become popular to map the imbalance between oxidative phosphorylation and glycolysis pathways in cancer applications [46], [47]. Figure 7 provides an illustration for such a high-dimensional imaging problem. As one would expect, introducing an additional temporal dimension to the already SNR-, speed- and resolution-limited modality further exacerbated these challenges. Therefore, existing studies have been limited by insufficient spatial and/or temporal resolutions, as a result, inaccurate characterization of heterogeneous metabolism. Low-dimensional models with learned priors offer a promising path to addressing these challenges, which we review in this section.

A. Low-Rank Tensor Models with Learned Subspaces

To tackle the increased dimensionality problem, the PS model in Eq. (6) can be generalized to represent new image functions of interest $\tilde{\rho}(\mathbf{r},t,T)$, with T denoting the additional time dimension ($T=1,2,\ldots,N_T$): [9],[49]

$$\tilde{\rho}(\mathbf{r},t,T)=\sum_{m=1}^{M}\sum_{l=1}^{L}\sum_{s=1}^{S}c_{m,l,s}u_m(\mathbf{r})v_l(t)g_s(T), \tag{13}$$

where N_T is the total number of time frame, $\{g_s(T)\}_{s=1}^{S}$ denote the basis functions that span the space of temporal variations and $\{c_{m,l,s}\}_{m,l,s=1}^{M,L,S}$ are the model coefficients. Eq. (13) implies a low-rank tensor structure (generalized from the low-rank matrices) that exploits the spatial-spectral-temporal correlations in dynamic MRSI data with $\{c_{m,l,s}\}$ referred to as the core tensor [49], [50]. M, L and S are the respective model orders (typically much smaller than N and Q). Similarly as the spectral basis $\{v_l(t)\}$ can be learned by

incorporating physics-based model, the temporal basis $\{g_s(T)\}$ can be predetermined from temporal training data [9], leaving only $\{u_m(\mathbf{r})\}$ and $\{c_{m,l,s}\}$ as the unknown in the imaging problem. Eq. (13) can be further adapted into a molecule-dependent tensor model [51]

$$\tilde{\rho}(\mathbf{r},t,T)=\sum_{m=1}^{M}s_m(\mathbf{r})\left(\sum_{l=1}^{L_m}\sum_{s=1}^{S_m}c_{m,l,s}(\mathbf{r})v_{m,l}(t)g_{m,s}(T)\right) \tag{14}$$

where M is the number of molecules of interest, L_m and S_m are the spectral and temporal model orders of the mth molecule, $c_{m,l,s}(\mathbf{r})$ combines the core tensor coefficients and molecule-dependent spatial basis. This model offers up to two orders-of-magnitude reduction in the degrees of freedom (DOF) [51] compared to the canonical truncated Fourier series representation, thus significantly improved resolution and SNR tradeoffs. Furthermore, the model in Eq. (14) provides the flexibility to incorporate molecule-specific physics- and biochemistry-based priors [44]. The spectral basis can be obtained as described in section III.A and the temporal basis can be derived by exploiting existing physiological and metabolic knowledge. More specifically, the temporal variations of the spatially-resolved spectra are often results of metabolic process in living organs, which can be approximated by a set of parametric curves specified by some time constants [51], e.g., the exponential recovery of PCr after ischemia or exercise. The variable $s_m(\mathbf{r})$ can be used to impose molecule-specific spatial support and distribution priors from high-resolution reference images available, further improving the performance (see [51] for more details). Accordingly, the dynamic MRSI reconstruction can be formulated as follows:

$$\widehat{\mathbf{C}}_m=\arg\min_{\{\mathbf{C}\}}\left\|\mathbf{d}-\mathscr{A}\{(\mathbf{s}_m\odot\mathbf{C}_m)\otimes\mathbf{V}_m\otimes\mathbf{G}_m\}\right\|_2^2+\lambda R(\{\mathbf{C}_m\}) \tag{15}$$

where $\mathscr{A}$ describes the forward model as discussed above (capturing field inhomogeneity effects and spatiospectral encoding process), $\mathbf{C}_m$, $\mathbf{V}_m$, $\mathbf{G}_m$ and $\mathbf{s}_m$ are tensor forms of $\{c_{m,l,s}(\mathbf{r})\}$, $\{v_{m,l}(t)\}$, $\{g_{m,s}(T)\}$ and $\{s_m(\mathbf{r})\}$ respectively, where $\mathbf{C}_m\in C^{N\times L_m\times S_m}$, $\mathbf{V}_m\in C^{L_m\times Q'}$, $\mathbf{G}_m\in C^{S_m\times N_T}$ and $\mathbf{s}_m\in C^{N\times 1}$. $\odot$ and $\otimes$ denote column-wise and tensor products. The algorithm in [51] first performed a series of reconstruction of time-dependent spatial coefficients (using per-estimated $\mathbf{V}_m$), i.e., $\{\hat{\mathbf{c}}_{m,l}(T)\}$, and then fit the temporal variations to biochemical models for each molecule to determine $\mathbf{G}_m$. Finally, Eq. (15) was solved using estimated spectral and temporal basis. $R(\{\mathbf{C}_m\})$ imposes regularization to encourage spatial smoothness or sparsity.

Considering the unique properties of the dynamic spectroscopic signals, Eq. (14) can be reformulated by imposing strict separability between the spatial-spectral and spatial-temporal variations [48], i.e.,

$$\tilde{\rho}(\mathbf{r},t,T)=\sum_{m=1}^{M}\left\{\left(\sum_{l=1}^{L_m}a_{m,l}(\mathbf{r})v_{m,l}(t)\right)\left(\sum_{s=1}^{S_m}b_{m,s}(\mathbf{r})g_{m,s}(T)\right)\right\} \tag{16}$$

where $a_{m,l}$ and $b_{m,s}$ are the new spatial coefficients for the spectral and temporal bases, respectively. With this model, the number of unknowns at each voxel is reduced from

$\sum_{m=1}^{M} L_m \times S_m$ to $\sum_{m=1}^{M}(L_m + S_m)$. This does not necessarily increase modeling error for dynamic MRSI specifically, because it is accurate to assume all FID/frequency points corresponding to the same molecule should share the same dynamics at each voxel. Using matrix notations for $a_{m,l}$ and $b_{m,s}$, a similar regularized least-squares problem to Eq. (15) can be formulated.

B. Combining Low-Rank Tensor and Manifold Modeling

As in the "static" MRSI case, it is straightforward to incorporate nonlinear manifold constraints on spectral variations into the reconstruction problems in Eqs. (15) or (16). For example, the learned network can be used to regularize individual FID's at every time point, i.e.,

$$\widehat{\mathbf{C}}_m, \hat{\rho} = \arg\min_{\{\mathbf{C},\rho\}} \| \mathbf{d} - \mathscr{A}\{\rho\} \|_2^2 + \lambda_1 \sum_{T=1}^{N_T} \sum_{n=1}^{N} \| \mathscr{C}(\boldsymbol{\rho}_{n,T}) - \boldsymbol{\rho}_{n,T} \|_2^2 + \lambda_2 R(\{\mathbf{C}_m\}), \quad s.t. \;\; \boldsymbol{\rho} = (\mathbf{s}_m \odot \mathbf{C}_m) \otimes \mathbf{V}_m \otimes \mathbf{G}_m . \tag{17}$$

This formulation assumes that the desired FIDs at each voxel and each time point (i.e., $\boldsymbol{\rho}_{n,T}$) reside on the same learned manifolds and can be solved using alternating minimization. Note that the ensemble temporal variations are still captured by the linear subspace model. To further exploit the prior information available along individual dimensions, the subspace model for temporal variations can be generalized to nonlinear manifolds as well. To this end, the separability representation in Eq. (16) offers additional flexibility to learn and incorporate low-dimensional manifold constraints on molecule-specific temporal dynamics. Specifically, a set of molecule-dependent DAEs, denoted as $\mathscr{G}_m(\cdot)$, were trained (using both synthetic and experimental dynamic data) and integrated into the following low-rank-tensor-based reconstruction problem [48]:

$$\begin{aligned} \widehat{\mathbf{A}}_m, \widehat{\mathbf{B}}_m = \; &= \arg\min_{\{\mathbf{A},\mathbf{B}\}} \|\mathbf{d} - \mathscr{A}\{\boldsymbol{\rho}\}\|_2^2 + \lambda_1 \sum_{T=1}^{N_T} \sum_{n=1}^{N} \|\mathscr{C}(\boldsymbol{\rho}_{n,T}) - \boldsymbol{\rho}_{n,T}\|_2^2 \\ &+ \lambda_2 \sum_{m=1}^{M} \|\mathscr{G}_m(\mathbf{B}_m\mathbf{G}_m) - \mathbf{B}_m\mathbf{G}_m\|_F^2, \quad s.t. \;\; \boldsymbol{\rho} = \sum_{m=1}^{M} (\mathbf{A}_m\mathbf{V}_m) \circ (\mathbf{B}_m\mathbf{G}_m) . \end{aligned} \tag{18}$$

where $\mathbf{A}_m$ and $\mathbf{B}_m$ are matrix forms of $\{a_{m,l}(\mathbf{r})\}$ and $\{b_{m,s}(\mathbf{r})\}$ with $\mathbf{A}_m \in C^{N \times L_m}$ and $\mathbf{B}_m \in C^{N \times S_m}$. $\circ$ denotes the operation defined in Eq. (16), and λ_1 and λ_2 are regularization parameters controlling the trade-off between data consistency and manifold (spectral and temporal) representation errors. Additional regularization terms on the spatial distribution of $\mathbf{a}_{m,l}$ and $\mathbf{b}_{m,s}$ may also be included. The problem in Eq. (18) can be solved via alternating minimization of $\mathbf{A}_m$ and $\mathbf{B}_m$, which also involves backpropagation of the learned networks. Details of the exact algorithm can be found in [48]. While parallelizable, the reconstruction can be computationally demanding due to the need to invert multiple networks for signals at many time points. The formulation used in RAIISE described in Section IV.B may be helpful in addressing this issue. A set of denoising dynamic ^{2}H-MRSI results is shown in Fig. 8 to illustrate the capability of the low-rank-tensor-model-based reconstruction using learned spectral and temporal subspaces. Comparisons of estimated metabolite maps,

spatially-resolved spectra and temporal dynamics from standard Fourier reconstruction, direct low-rank tensor truncation and learned-subspace-based methods were made.

C. The Motion Issue

The formulations discussed above do not fully consider motions (involuntary or voluntary) during the time window when the dynamic metabolic activities are monitored. This may be an issue when the imaging time is long (e.g., more than just a few minutes) and/or the organ of interest is more susceptible to motion. Prospective and retrospective motion tracking and correction may be used to mitigate its effects [22], [52]. Another possibility is to incorporate motion models into the T dimension, which might lead to an increased model order and degradation in the accuracy of capturing metabolite dynamics. A new dimension can be introduced to represent different motion phases for motion-resolved dynamic MRSI using a higher-order tensor model [50].

VI. Spectral QuAntification using LeARNED Models

With the ability to reconstruct high-SNR, high-resolution spatiotemporal function from noisy data, a key remaining challenge is to extract and quantify individual molecular components from the spatially-resolved FIDs or MR spectra, referred to as the spectral quantification problem. An apparent choice is to use a parametric model similar to Eq. (5) and perform parameter estimation voxel by voxel, e.g., through a nonlinear least-squares (NLLS) fitting [53]:

$$\hat{c}_m, \hat{T}^*_{2,m}, \hat{f}_m, \hat{\beta} = \sum_{q'=1}^{Q'} \left\| \widehat{\rho}_{n,q'} - \sum_{m=1}^{M} c_m \phi_m(t_{q'}) e^{-t_{q'}/T^*_{2,m} + i2\pi\delta f_m t_{q'}} h(t_{q'};\beta) \right\|^2 \tag{19}$$

where $\widehat{\rho}_n$ is the reconstructed FID at the nth voxel and $h(t_{q'};\beta)$ is a molecule-independent modulation function to account for additional spectral distortion, e.g., due to residual intravoxel B_0 inhomogeneity and eddy current effects etc. A Gaussian lineshape function is often used for $h(t)$. Commonly used NLLS solvers such as quasi-Newton methods or variable projection (VARPRO) [53] can be used to solve Eq. (19), but the problem is highly nonlinear, sensitive to noise and model mismatch. It can be time-consuming to solve these voxel-wise nonlinear least-squares for high-resolution acquisitions and computationally expensive to incorporate spatial constraints. Learning-based methods have been recently proposed to address these challenges.

Following a similar motivation from Eq. (5) for constructing the subspace model, a union-of-subspaces (UoSS) approach was described for spectral quantification [12]. Specifically, the FID signals corresponding to individual molecular components can be assumed to reside on their own subspaces spanned by $\{v_{l_m}(t)\}$, i.e.,

$$\tilde{\rho}(\mathbf{r},t) = \sum_{m=1}^{M} \sum_{l_m=1}^{L_m} u_{l_m}(\mathbf{r}) v_{l_m}(t)\,. \tag{20}$$

The basis $\{v_{l_m}(t)\}$ ($\mathbf{V}_m$ in matrix form) can be learned using the strategy described in Section III with ensemble molecule-specific FIDs, either synthesized or obtained from an initial NLLS estimation of the spectral parameters from the reconstruction $\hat{\rho}(r,t)$. Quantification can then be done by solving a UoSS fitting

$$\left\{\widehat{\mathbf{U}}_m\right\} = \arg\min_{\{\mathbf{U}_m\}} \left\| \widehat{\boldsymbol{\rho}} - \sum_{m=1}^{M} \mathbf{U}_m \mathbf{V}_m \right\|_2^2 + \sum_{m=1}^{M} \lambda_m \boldsymbol{R}_m(\mathbf{U}_m), \tag{21}$$

where $\mathbf{U}_m$ are molecule-specific spatial coefficients that can be converted into concentration values (see [12] for more details). This formulation offers stronger robustness against model mismatch owing to the representation power of the subspace model at individual voxels, less sensitivity to noise by transforming the nonlinear fitting into low-dimensional linear problems. It also allows for easier incorporation of spatial priors on individual components through $\boldsymbol{R}_m(\cdot)$. Extension of this UoSS approach to "local" subspaces has also been recently explored [54]. This method was motivated by the linear tangent space concept in smooth manifold recovery which allows a further reduction in the degrees-of-freedom in $\mathbf{U}_m$.

Going beyond subspace models, deep convolutional networks (CNNs) have been used to learn a nonlinear mapping from noisy FIDs/spectra to spectral parameters [13], [55] or clean, lineshape corrected metabolite only spectra [14], [24] for simplified quantification. A self-supervised training strategy was proposed in [13] where a CNN-based regressor $\mathscr{C}(\cdot;\theta)$ (θ the network parameters) was concatenated with a physics-based parametric model $f(t;\alpha)$ to resynthesize spectra from the parameters produced by the CNN. Errors between this resynthesized and original spectra were minimized, which can be formulated as the following optimization problem

$$\hat{\theta} = \arg\min_{\theta} \left\| \widehat{\boldsymbol{\rho}} - f(t; \mathscr{C}(\widehat{\boldsymbol{\rho}}; \theta)) \right\|_2^2, \tag{22}$$

where the spectral parameters α were replaced by the CNN output. This strategy does not require ground truth values for training which are difficult to obtain from practical, in vivo MRS/MRSI data. Processing time reduction was emphasized rather than estimation accuracy compared to standard voxel-wise fitting method. Generalizability may be an issue if there are spectral distortion and contamination not captured by the training data. Since a parametric model is still used to resynthesize the final fits, parameter estimation accuracy still needs to be carefully investigated in the presence of model mismatch. Instead of training a network that directly maps noisy data to spectral parameters, [24] proposed to learn a CNN that is able to extract metabolite-only, lineshape corrected spectra from noisy data with macromolecule (MM) and baseline "contaminations" and lineshape distortion. Estimating of metabolite concentrations from the corrected metabolite-only spectra can then be easily done using a linear least-squares fit to metabolite basis. All these methods are voxel wise thus difficult to incorporate spatial constraints. The learned mapping can be sensitive to changes in acquisition parameters, such as spectral bandwidth, resolution and SNR etc.

VII. Future Directions

Recent advances in computational MRSI using physics-based machine learning have demonstrated good potential for addressing high-dimensional MRSI reconstruction and quantification problems, achieving impressive performance in the combination of speed, resolution, SNR and robustness to system imperfection and practical experimental conditions. These “software” advances are highly synergistic with the on-going “hardware”/ instrumentation developments, e.g., higher field systems, better field inhomogeneity monitoring and shimming devices, and more sensitive coils. We expect to see that synergistic developments in advanced data acquisition and processing will significantly enhance the capability and clinical utility of next generation MRSI technologies.

One area that will see more innovations is MRSI data processing using deep learning networks trained using high-resolution images. Currently, the lack of high-quality MRSI training data makes supervised training, commonly used in learning-based image reconstruction, not directly applicable. One potential direction is to construct an MRSI atlas from highly dedicated experimental data (with many averages followed by artifact correction and registration), to which statistical variations in concentrations, linewidths and other spectral parameters as well as geometry may be introduced for synthetic MRSI data generation as was done in [10], [48] but with more realistic variations. Another potential direction of pursuit is developing learning strategies that do not require companion high-SNR/”clean” data. Unsupervised and self-supervised learning strategies have gained attraction for image denoising and reconstruction applications [56]–[59]. These formulations may offer some inspiration for learning and using models with only noisy MRSI data. As water spectroscopic images are often acquired in an MRSI scan, e.g., through a separate or simultaneous nonwater-suppressed acquisition [21], [22], approaches to leverage these high-dimensional, multicontrast water images beyond just extracting B_0 and B_1 maps to introduce effective spatial constraints should be of great interest. The generative-model-based methods discussed in Section IV.C represent some initial attempts while more work can be done.

While the topics related to spatial-spectral-temporal reconstruction have been the main emphasis in this review, we note that spectral quantification as a critical step towards quantitative MRSI for clinical and basic science applications remains under-investigated. A number of sophisticated parametric models have been presented, mostly for single-voxel spectroscopy, but they often yield significant voxel-to-voxel estimation variations and are computationally demanding for high-resolution data. The notion of voxelby-voxel fitting should continue to be challenged, by new ways to integrate spatial constraints with deep-learning-based solutions. In dynamic or multidimensional MRSI scenarios, quantification is usually done for individual points along time or other additional dimensions (e.g., different TEs). Novel methods for joint quantification, e.g., models and algorithms that simultaneously fit the time-resolved spectra in dynamic MRSI (with reduced number of parameters), should be developed with potential significant practical impact.

We have focused this review on MRSI, but the methodologies discussed here are also adaptable to other high-dimensional spatiospectral imaging problems, where physics-based models and/or simulations can be used to describe the signal changes from which

low-dimensional representations of the intrinsic variations may be learned. Authors in [60] proposed to use a fully-connected feedforward network trained using simulated data to perform voxel-wise quantification of PCr concentration, exchange rate and other system-related parameters from Z-spectra data generated by chemical exchange saturation transfer (CEST) imaging. The learning-based method demonstrated advantages in both performance and computational cost. Applications of physics-based machine learning to other quantitative MRI problems have also been explored (see other reviews in this issue). Similar strategies may also be extended to other spatiospectral imaging modalities besides MR, e.g., mass spectrometry imaging [61]. Beyond addressing the reconstruction and quantitative analysis problems within a single modality, we also expect to see more progresses in connecting and fusing imaging information from different modalities across spatiotemporal scales by integrating physics-based modeling and data-driven learning.

Acknowledgments

This work was supported in part by NSF, grant number: NSF-CBET 1944249 and NIH, grant number 1R35GM142969-01. The authors thank Yahang Li and Yudu Li for their help on the figures.

References

[1]. Lauterbur PC, Kramer DM, House WV, and Chen C-N, “Zeugmatographic high resolution nuclear magnetic resonance spectroscopy: Images of chemical inhomogeneity within macroscopic objects,” J. Amer. Chem. Soc, vol. 97, pp. 6866–6868, 1975.

[2]. Brown TR, Kincaid BM, and Ugurbil K, “NMR chemical shift imaging in three dimensions,” Proc. Natl. Acad. Sci, vol. 79, pp. 3523–3526, 1982. [PubMed: 6954498]

[3]. Ross B and Michaelis T, “Clinical applications of magnetic resonance spectroscopy.” Magn. Reson. Q, vol. 10, no. 4, pp. 191–247, 1994. [PubMed: 7873353]

[4]. de Graaf RA, In Vivo NMR Spectroscopy: Principles and Techniques. Hoboken, NJ: John Wiley & Sons, 2018.

[5]. Li Y, Wang T, Zhang T, Lin Z, Li Y, Guo R, Zhao Y, Meng Z, Liu J, Yu X, Liang Z-P, and Nachev P, “Fast high-resolution metabolic imaging of acute stroke with 3D magnetic resonance spectroscopy,” Brain, vol. 143, pp. 3225–3233, 2020. [PubMed: 33141145]

[6]. Bogner W, Otazo R, and Henning A, “Accelerated MR spectroscopic imaging—a review of current and emerging techniques,” NMR Biomed, vol. 34, no. 5, p. e4314, 2021. [PubMed: 32399974]

[7]. Liang Z-P, “Spatiotemporal imaging with partially separable functions,” in Proc IEEE Int. Symp. Biomed. Imag, Arlington, VA, USA, 2007, pp. 988–991.

[8]. Lam F, Li Y, Guo R, Clifford B, and Liang Z-P, “Ultrafast magnetic resonance spectroscopic imaging using SPICE with learned subspaces,” Magn Reson Med, vol. 83, pp. 377–390, 2020. [PubMed: 31483526]

[9]. Ma C, Clifford B, Liu Y, Gu Y, Lam F, Yu X, and Liang Z-P, “High-resolution dynamic 31P-MRSI using a low-rank tensor model,” Mag. Reson. Med, vol. 78, pp. 419–428, 2017.

[10]. Iqbal Z, Nguyen D, Hangel G, Motyka S, Bogner W, and Jiang S, “Super-resolution 1H magnetic resonance spectroscopic imaging utilizing deep learning,” Frontiers in Oncology, vol. 9, 2019.

[11]. Qu X, Huang Y, Lu H, Qiu T, Guo D, Agback T, Orekhov V, and Chen Z, “Accelerated nuclear magnetic resonance spectroscopy with deep learning,” Angew. Chem. Int. Ed, vol. 59, no. 26, pp. 10297–10300, 2020.

[12]. Li Y, Lam F, Clifford B, and Liang Z, “A subspace approach to spectral auantification for MR spectroscopic imaging,” IEEE Trans Biomed Eng, vol. 64, pp. 2486–2489, 2017. [PubMed: 28829303]

[13]. Gurbani SS, Sheriff S, Maudsley AA, Shim H, and Cooper LA, “Incorporation of a spectral model in a convolutional neural network for accelerated spectral fitting,” Magn Reson Med, vol. 81, no. 5, pp. 3346–3357, 2019. [PubMed: 30666698]

[14]. Lee H and Kim H, “Bayesian deep learning-based 1h-mrs of the brain: Metabolite quantification with uncertainty estimation using monte carlo dropout,” in Proc Intl Soc Mag Reson Med, 2021, p. 2014.

[15]. Liang Z-P and Lauterbur PC, Principles of Magnetic Resonance Imaging: A Signal Processing Perspective. New York: IEEE Press, 2000.

[16]. Barkhuysen H, de Beer R, and van Ormondt D, “Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals,” J. Magn. Reson, vol. 73, pp. 553–557, 1987.

[17]. Lam F and Liang Z-P, “A subspace approach to high-resolution spectroscopic imaging,” Magn Reson Med, vol. 71, pp. 1349–1357,2014. [PubMed: 24496655]

[18]. Nguyen HM, Peng X, Do MN, and Liang Z-P, “Denoising MR spectroscopic imaging data with low-rank approximations,” IEEE Trans. Biomed. Eng, vol. 60, pp. 78–89, 2013. [PubMed: 23070291]

[19]. Kasten J, Klauser A, Lazeyras F, and Van De Ville D, “Magnetic resonance spectroscopic imaging at superresolution: Overview and perspectives,” J. Magn. Reson, vol. 263, pp. 193–208, 2016. [PubMed: 26766215]

[20]. Lam F, Ma C, Clifford B, Johnson CL, and Liang Z-P, “High-resolution 1H-MRSI of the brain using SPICE: Data acquisition and image reconstruction,” Magn Reson Med, vol. 76, pp. 1059–1070, 2016. [PubMed: 26509928]

[21]. Peng X, Lam F, Li Y, Clifford B, and Liang Z-P, “Simultaneous QSM and metabolic imaging of the brain using SPICE,” Magn Reson Med, vol. 79, pp. 13–21, 2018. [PubMed: 29067730]

[22]. Guo R, Zhao Y, Li Y, Wang T, Li Y, Sutton B, and Liang Z-P, “Simultaneous QSM and metabolic imaging of the brain using SPICE: Further improvements in data acquisition and processing,” Magn. Reson. Med, vol. 85, no. 2, pp. 970–977, 2021. [PubMed: 32810319]

[23]. Parsons L, Haque E, and Liu H, “Subspace clustering for high dimensional data: A review,” SIGKDD Explor. Newsl, vol. 6, no. 1, p. 90–105, 2004.

[24]. Lee HH and Kim H, “Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain,” Magn Reson Med, vol. 82, no. 1, pp. 33–48, 2019. [PubMed: 30860291]

[25]. Li Y, Wang Z, Sun R, and Lam F, “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. [PubMed: 33395390]

[26]. Lam F, Li Y, and Peng X, “Constrained magnetic resonance spectroscopic imaging by learning nonlinear low-dimensional models,” IEEE Trans Med Imaging, vol. 39, pp. 545–555, 2020. [PubMed: 31352337]

[27]. Tiwari P, Rosen M, and Madabhushi A, “Consensus-locally linear embedding (C-LLE): Application to prostate cancer detection on magnetic resonance spectroscopy,” in MICCAI 2008. Springer Berlin Heidelberg, 2008, pp. 330–338.

[28]. Yang G, Raschke F, Barrick TR, and Howe FA, “Manifold learning in MR spectroscopy using nonlinear dimensionality reduction and unsupervised clustering,” Magn. Reson. Med, vol. 74, no. 3, pp. 868–878, 2015. [PubMed: 25199640]

[29]. Li Y, Wang Z, and Lam F, “SNR enhancement for multi-TE MRSI using joint low-dimensional model and spatial constraints,” IEEE Trans. Biomed. Eng, 2022, in Press.

[30]. Klauser A, Courvoisier S, Kasten J, Kocher M, Guerquin-Kern M, Van De Ville D, and Lazeyras F, “Fast high-resolution brain metabolite mapping on a clinical 3T MRI by accelerated 1H-FID-MRSI and low-rank constrained reconstruction,” Magn. Reson. Med, 2018, 10.1002/mrm.27623.

[31]. Chen Y, Li Y, and Xu Z, “Improved low-rank filtering of mr spectroscopic imaging data with pre-learnt subspace and spatial constraints,” IEEE Transactions on Biomedical Engineering, vol. 67, pp. 2381–2388, 2020. [PubMed: 31870975]

[32]. Lingala SG, Hu Y, DiBella E, and Jacob M, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR," IEEE Trans. Med. Imaging, vol. 30, no. 5, pp. 1042–1054, 2011. [PubMed: 21292593]

[33]. Trzasko JD, "Exploiting local low-rank structure in higher-dimensional MRI applications," in Wavelets and Sparsity XV, Ville DVD, Goyal VK, and Papadakis M, Eds., vol. 8858, International Society for Optics and Photonics. SPIE, 2013, pp. 551–558.

[34]. Zhang T, Pauly JM, and Levesque IR, "Accelerating parameter mapping with a locally low rank constraint," Magn. Reson. Med, vol. 73, no. 2, pp. 655–661, 2015. [PubMed: 24500817]

[35]. "Using a local low rank plus sparse reconstruction to accelerate dynamic hyperpolarized 13C imaging using the bSSFP sequence," J. Magn. Reson, vol. 290, pp. 46–59, 2018. [PubMed: 29567434]

[36]. Haldar JP, "Low-rank modeling of local *k*-space neighborhoods (LORAKS) for constrained MRI," IEEE Trans. Med. Imaging, vol. 33, no. 3, pp. 668–681, 2014. [PubMed: 24595341]

[37]. Vizioli L, Moeller S, Dowdle L, Akccakaya M, De Martino F, Yacoub E, and Uugurbil K, "Lowering the thermal noise barrier in functional brain mapping with magnetic resonance imaging," Nat. Commun, vol. 12, no. 1, p. 5181, 2021. [PubMed: 34462435]

[38]. Li Y, Ruhm L, Henning A, and Lam F, "LeaRning nonlineAr representatIon and projectIon for faSt constrained MRSI rEconstruction (RAIISE)," in Proc Intl Soc Mag Reson Med, 2022, p. TBD.

[39]. Gong K, Han PK, Marin T, Fakhri GE, Li Q, and Ma C, "High resolution MR spectroscopic imaging using deep image prior constrained subspace modeling," in Proc Intl Soc Mag Reson Med, 2020, p. 388.

[40]. Ulyanov D, Vedaldi A, and Lempitsky VS, "Deep image prior," 2018 IEEE/CVF CVPR, pp. 9446–9454, 2018.

[41]. Zhao R, Wang Z, and Lam F, "Generative image prior constrained subspace reconstruction for high-resolution mrsi," in Proc Intl Soc Mag Reson Med, 2022, p. TBD.

[42]. Daras G, Dean J, Jalal A, and Dimakis A, "Intermediate layer optimization for inverse problems using deep generative models," in Proceedings of the 38th International Conference on Machine Learning, vol. 139, 2021, pp. 2421–2432.

[43]. Xia W, Zhang Y, Yang Y, Xue J-H, Zhou B, and Yang M-H, "GAN inversion: A survey," arXiv, 2021,2101.05278..

[44]. Liu Y, Gu Y, and Yu X, "Assessing tissue metabolism by phosphorous-31 magnetic resonance spectroscopy and imaging: a methodology review," Quant Imaging Med Surg, vol. 7, pp. 707–726, 2017. [PubMed: 29312876]

[45]. Santos-Diaz A and Noseworthy MD, "Phosphorus magnetic resonance spectroscopy and imaging(31p-mrs/mrsi) as a window to brain and muscle metabolism: A review of the methods," Biomedical Signal Processing and Control, vol. 60, 2020.

[46]. Brender JR, Kishimoto S1, Merkle H, Reed G, Hurd RE, Chen AP, Ardenkjaer-Larsen JH, Munasinghe J, Saito K, Seki T, Oshima N, Yamamoto K, Choyke PL, Mitchell J, and Krishna MC, "Dynamic imaging of glucose and lactate metabolism by 13c-mrs without hyperpolarization," Scientific Reports, vol. 9, 2019.

[47]. Feyter HMD, Behar KL, Corbin ZA, Fulbright RK, Brown PB, McIntyre S, Nixon TW, Rothman DL, and de Graaf RA, "Deuterium metabolic imaging (dmi) for mri-based 3d mapping of metabolism in vivo," Scientific Reports, vol. 4, 2018.

[48]. Li Y, Zhao Y, Guo R, Wang T, Zhang Y, Chrostek M, Low WC, Zhu X-H, Liang Z-P, and Chen W, "Machine learning-enabled high-resolution dynamic deuterium MR spectroscopic imaging," IEEE Trans. Med. Imaging, vol. 40, pp. 3879–3890, 2021. [PubMed: 34319872]

[49]. He J, Liu Q, Christodoulou AG, Ma C, Lam F, and Liang Z, "Accelerated high-dimensional MR imaging with sparse sampling using low-rank tensors," IEEE Trans Med Imaging, vol. 35, pp. 2119–2129, 2016. [PubMed: 27093543]

[50]. Christodoulou AG, Shaw JL, Nguyen C, Yang Q, Xie Y, Wang N, and Li D, "Magnetic resonance multitasking for motion-resolved quantitative cardiovascular imaging," Nat. Biomed. Eng, vol. 2, pp. 215–226, 2018. [PubMed: 30237910]

[51]. Clifford B, Gu Y, Liu Y, Kim K, Huang S, Li Y, Lam F, Liang Z-P, and Yu X, “High-resolution dynamic 31P-MR spectroscopic imaging for mapping mitochondrial function,” IEEE Transactions on Biomedical Engineering, vol. 67, pp. 2745–2753, 2020. [PubMed: 32011244]

[52]. Bogner W, Hess AT, Gagoski B, Tisdall MD, van der Kouwe AJ, Trattnig S, Rosen B, and Andronesi OC, “Real-time motion- and B0-correction for LASER-localized spiral-accelerated 3D-MRSI of the brain at 3T,” NeuroImage, vol. 88, pp. 22–31, 2014. [PubMed: 24201013]

[53]. Vanhamme L, van den Boogaart A, and Van Huffel S, “Improved method for accurate and efficient quantification of MRS data with use of prior knowledge,” J. Magn. Reson, vol. 129, no. 1, pp. 35–43, 1997. [PubMed: 9405214]

[54]. Ma C and Fakhri GE, “MRSI spectral quantification using linear tangent space alignment (LTSA)-based manifold learning,” in Proc Intl Soc Mag Reson Med, 2022, p. 243.

[55]. Turco F, Zubak I, and Slotboom J, “Deep learning based mrs metabolite quantification: Cnn and resnet versus non linear least square fitting,” in Proc Intl Soc Mag Reson Med, 2021, p. 2018.

[56]. “Noise2noise: Learning image restoration without clean data,” in 35th International Conference on Machine Learning, ICML 2018, ser. Proceedings of Machine Learning Research, vol. 7, 2018, pp. 4620–4631.

[57]. Yaman B, Hosseini SAH, Moeller S, Ellermann J, Uugurbil K, and Akccakaya M, “Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data,” Magnetic Resonance in Medicine, vol. 84, no. 6, pp. 3172–3191,2020. [PubMed: 32614100]

[58]. Liu J, Sun Y, Eldeniz C, Gan W, An H, and Kamilov US, “RARE: Image reconstruction using deep priors learned without groundtruth,” IEEE J. Sel. Top. Signal Process, vol. 14, no. 6, pp. 1088–1099, 2020.

[59]. Akcakaya M, Yaman B, Chung H, and Ye JC, “Unsupervised deep learning methods for biological image reconstruction and enhancement: An overview from a signal processing perspective,” IEEE Signal Process. Mag, vol. 39, no. 2, pp. 28–44, 2022. [PubMed: 36186087]

[60]. Chen L, Schar M, Chan KWY, Huang J, Wei Z, Lu H, Qin Q, Weiss RG, van Zijl PCM, and Xu J, “In vivo imaging of phosphocreatine with artificial neural networks,” Nat. Comm, vol. 11, p. 1072, 2020.

[61]. Spengler B, “Mass spectrometry imaging of biomolecular information,” Anal. Chem, vol. 87, pp. 64–82, 2015. [PubMed: 25490190]

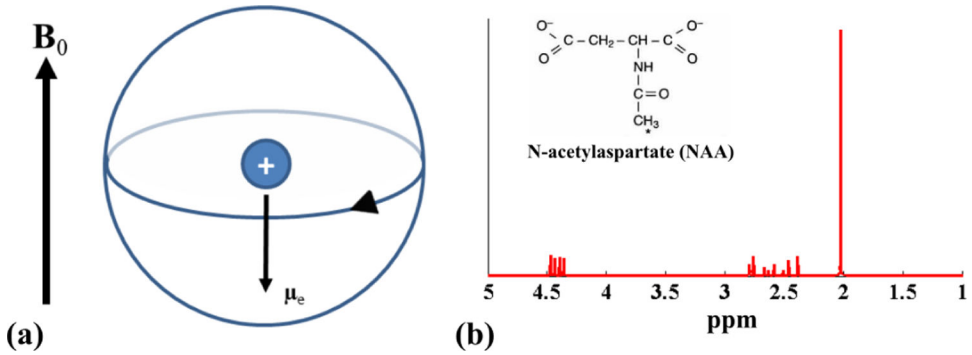

Fig. 1.

Illustration of chemical shift and the resulting frequency distribution: (a) the electron orbiting around the nucleus (denoted by the arrow on the circle) can be viewed as a small current, which generates a magnetic moment μ_e that opposes the main magnetic field B_0, thus perturbing the magnetic field felt "locally" by the nucleus and generating a small frequency shift; (b) resonance structure of the molecule NAA (from quantum mechanical simulation) as a result of the chemical shift phenomenon.

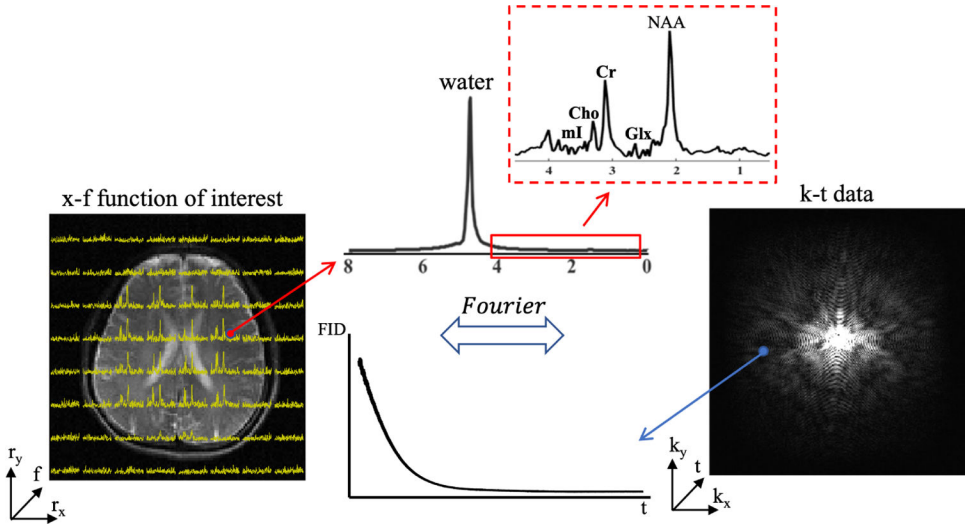

Fig. 2.

Illustration of the imaging problem with MRSI: The desired high-dimensional spatiospectral function is in (x, f) domain where we aim to recover high-resolution spectra at each voxel (left). Data are acquired in (k, t)-space where FIDs are sampled for individual k-space locations (right). This imaging problem is inherently higher dimensional than conventional MRI where only k-space is sampled. Furthermore, the signals from molecules of interest (e.g., NAA, Cr, and Cho etc) are three orders of magnitude weaker than water (plots on top of the middle column), making the problem more challenging.

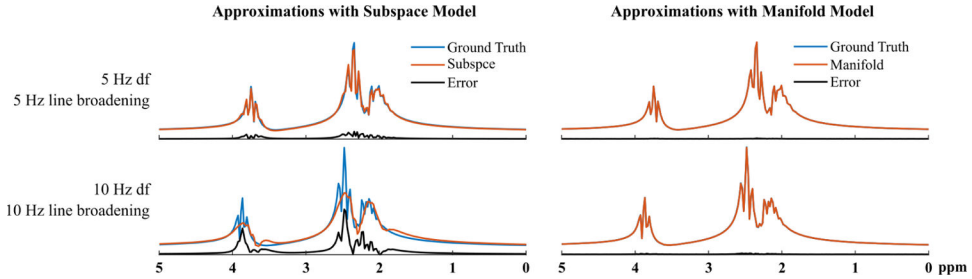

Fig. 3.

Linear v.s. nonlinear low-dimensional models of spectroscopic signals: (Left) Approximation of glutamate spectra using subspaces learned from an ensemble of simulated data. The accuracy reduces as the ranges for the spectral parameters used to generate the training and testing data increase (from top to bottom rows); (Right) Approximation by the nonlinear model learned from the same data, with the same model order of 3. The color coding is specified in the figure legend.

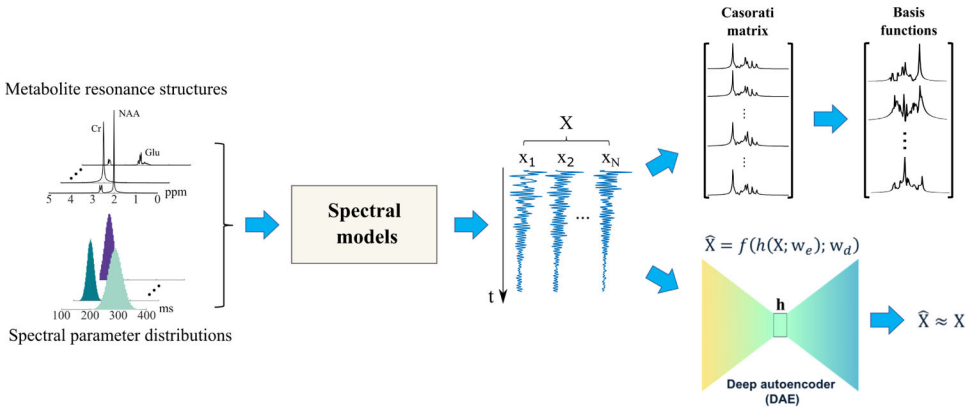

Fig. 4.

Physical-model-driven data generation and low-dimensional representation learning for MRSI data. Both QM-simulated resonance structures of individual molecules (metabolite basis, top of left column) and spectral parameters sampled from empirical distributions (from literature values or experimental data, bottom of left column) are fed into a spectral fitting model to generate a large quantity of training data (X). These data can be either formed into a Casorati matrix from which a set of basis can be estimated (upper branch, linear subspace model) or used to train a DAE to capture a nonlinear low-dimensional manifold where high-dimensional spectroscopic signals reside.

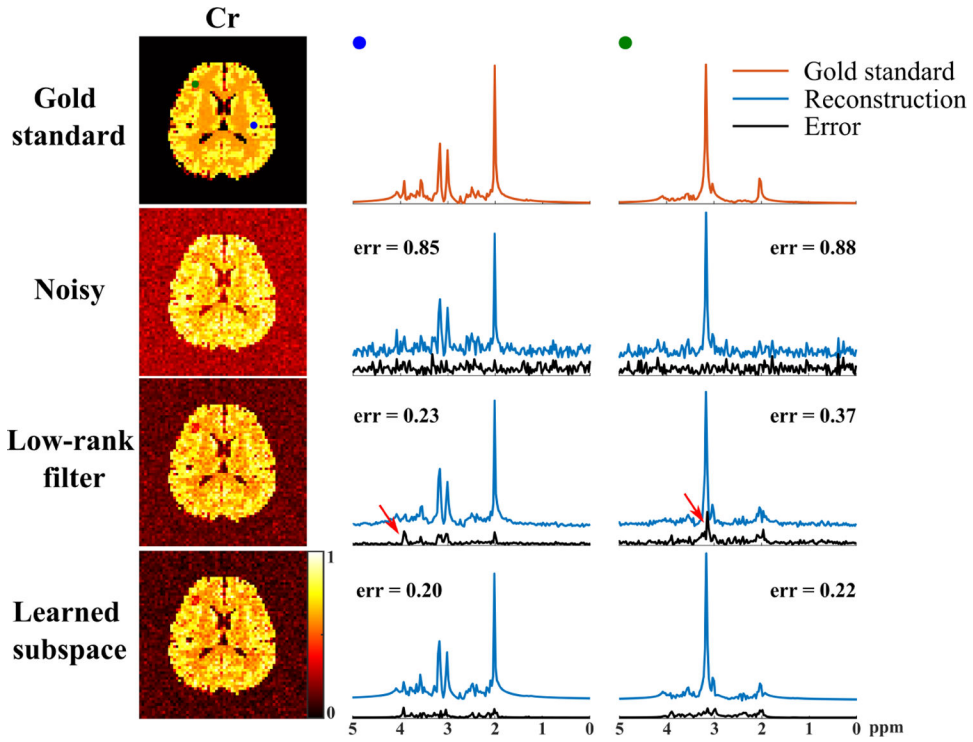

Fig. 5.

Reconstructions (simulated data) with and without using a pre-learned subspace. The first column shows metabolite maps (Cr) from different cases (rows 1–4: gold standard, Fourier reconstruction, low-rank filtering and reconstruction using a learned subspace), and the subsequent columns show localized spectra from two voxels. Direct low-rank filtering (jointly estimating the spatial coefficients and subspace) produced significantly larger errors (black curves) than learned subspace, especially in small features with a distinct spectral pattern (green dot). Relative ℓ_2 errors were included (err) for quantitative comparison.

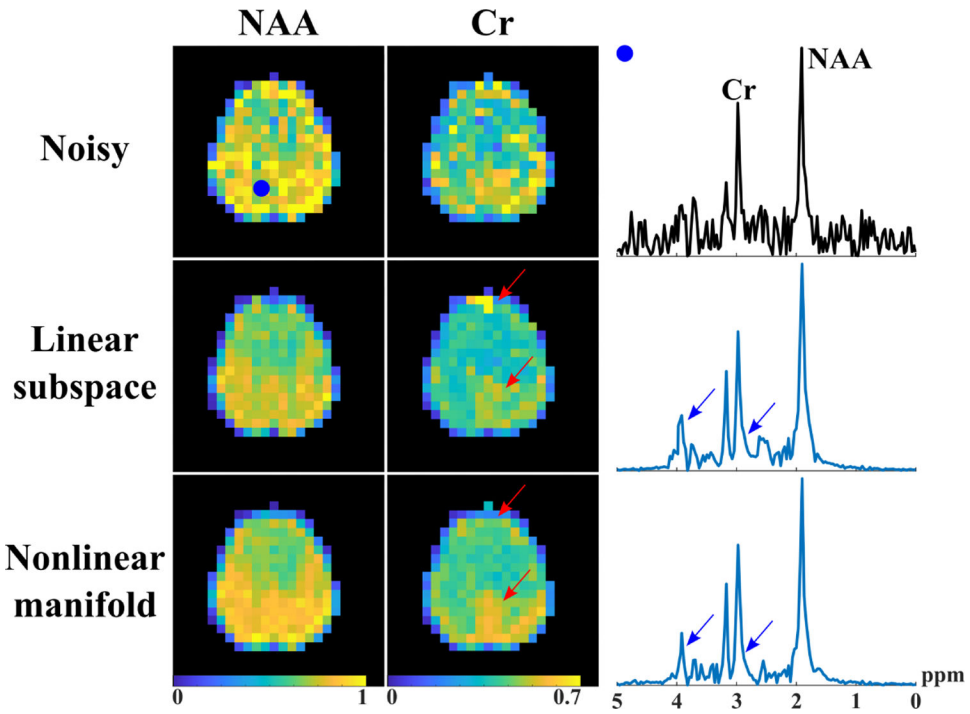

Fig. 6.

A set of in vivo results from reconstructions using learned subspace (Linear subspace) and nonlinear models (Nonlinear manifold). The latter produced metabolite maps with less noise contamination and better recovered tissue-dependent features (as indicated by the red arrows), as well as spectra with sharper lineshapes (third column, blue arrows). A more thorough quantitative analysis without a gold standard can be found in [29].

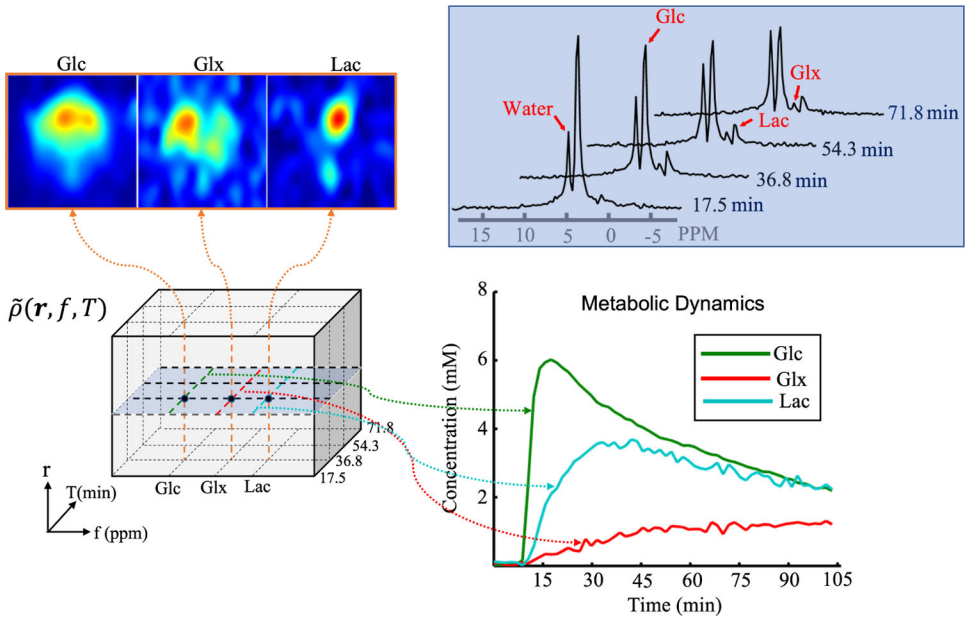

Fig. 7.

The dynamic MRSI problem (dynamic ^{2}H-MRSI of a rat brain in this case): The center cube illustrates the high-dimensional image function of interest. Orange lines taken from different frequencies (f) and a single time point (T) yield various metabolite maps (top left, spatial dimension); the blue plane represents time-resolved spectra (top right, spectral-temporal dimensions) at a single voxel (r); lines from different f 's and a single **r** capture metabolic dynamics (bottom right). The low-rank-tensor-based reconstruction with learned manifold constraints produced an impressive combination of speed, resolution and SNR [48], i.e., 1.8min frame rate at a resolution of 17 × 17 × 5 matrix size over 28 × 28 × 24mm^3 FOV.

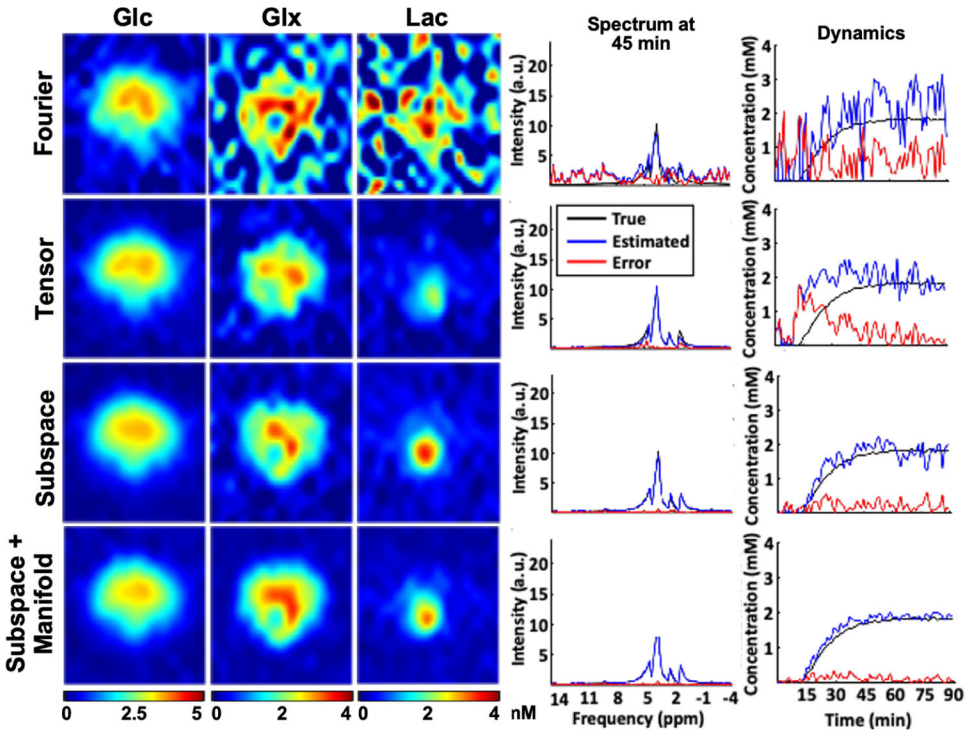

Fig. 8.

A set of dynamic ^{2}H-MRSI results from a rat brain by the standard Fourier reconstruction (Fourier), a direct low-rank tensor filtering of the noisy data (Tensor), a tensor reconstruction with learned subspaces (Subspace), and joint subspace model and manifold regularization (Subspace+Manifold) [48]. For each method, maps of three metabolites (columns 1–3), localized spectra at a particular time point (4th column) and temporal dynamics for the same voxel (5th column) are shown. The learned subspace methods produced improved metabolite maps, and lower errors for the spectra and temporal dynamics. Manifold regularization further improved the temporal fidelity. Both the Tensor and Subspace methods used the same model order ($L_m = 6$ and $S_m = 30$). Colors for different curves are noted in the plot.