

Fan Lam¹, Xi Peng, and Zhi-Pei Liang²

High-Dimensional MR Spatospectral Imaging by Integrating Physics-Based Modeling and Data-Driven Machine Learning

Current progress and future directions



©SHUTTERSTOCK.COM/PAPAPIG

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 the high dimensionality and low signal-to-noise ratio (SNR). Recent technological developments integrating physics-based modeling and data-driven machine learning that exploit the unique physical and mathematical properties of MRSI signals have demonstrated impressive performance in addressing these challenges for rapid high-resolution quantitative MRSI. This article provides a systematic review of recent progress in the context of MRSI physics and offers perspectives on promising future directions.

Introduction

MR spatospectral imaging techniques allow for the noninvasive visualization and quantification of molecule-specific physiological processes in living animals and humans that are inaccessible by conventional anatomical and functional imaging methods. MRSI, in particular, integrates the concepts of spatial encoding used in MRI and spectral encoding used in MR spectroscopy (MRS) to produce spatially resolved 1D or multidimensional spectra. These spectra allow the 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 physiological processes) and clinical applications (including diagnosis, prognosis, and treatment monitoring for cancer, stroke, epilepsy, and many other neurological and psychiatric diseases) [3], [4], [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

Digital Object Identifier 10.1109/MSP.2022.3203867
Date of current version: 17 February 2023

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 3–4 orders of magnitude lower concentrations than the water molecules imaged by MRI
- 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 SNR, resolution, speed, and robustness desired or required by many practical applications.

Recent advances in ultrahigh-field systems, high-sensitivity radio-frequency (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 toward 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 spatio-spectral functions from limited and/or noisy data [7], [8], [9], [10], [11] and quantifying molecular parameters (e.g., concentrations and relaxation times) from the spatio-spectral function to derive quantitative biomarkers for physiological functions and diseases [12], [13], [14].

We present, in this article, a systematic review of recent progress 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 spatio-spectral imaging modalities.

This review is organized as follows. The “MRSI Physics” section introduces basic physical principles underlying MRSI and defines the key challenges associated with in vivo high-dimensional MRSI. The “Physics-Based Spatio-spectral Priors” section reviews physics-based modeling of spectroscopic signals and data-driven low-dimensional model learning approaches inspired by the physics models. The “Spatio-spectral Reconstruction Using Learned Models” section discusses spatio-spectral image reconstruction methods integrating physics-based and learned models. The “Dynamic MRSI Using Learned Models” section extends the discussion to dynamic MRSI, followed by the “Spectral Quantification

Using Learned Models” section, where integrated physics-based and machine learning strategies for metabolite quantification are reviewed. Finally, the “Future Directions” section presents our perspectives on future research directions.

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 (^1H). MRSI, on the other hand, is designed to measure MR signals from ^1H and/or other nuclei (e.g., ^{31}P) in a range of molecules [e.g., N-acetylaspartate (NAA), choline (Cho), creatine (Cr), 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 spatio-spectral encoding strategies used in MRSI. Then, we will introduce the resulting high-dimensional spatio-spectral imaging problem and the unique challenges associated with MRSI.

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 (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., ^1H '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 Figure 1(a), makes different nuclei in the same or different molecules resonate at a range of frequencies specified by

$$f = \gamma B_0(1 - \sigma) \quad (1)$$

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 the nuclei and 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 NAA (one of the most abundant metabolites in the brain) exhibit several frequencies, producing a unique resonance structure [Figure 1(b)]. Different molecules have their unique resonance structures.

Physiologically meaningful high-dimensional FIDs/spectra we acquired should reside in or close to a low-dimensional subspace.

Considering the chemical shift distribution in each imaging voxel, the image function of interest becomes a spatio-spectral 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 an 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 are in (k, t) -space and related to $\tilde{\rho}(\mathbf{r}, t)$ by

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

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 ($\Delta k_x, \Delta k_y, \Delta k_z$ and Δt are the corresponding sampling intervals); and $\epsilon(\cdot)$ represents the measurement noise (modeled as complex white Gaussian). Based on (2), the MRSI problem is to recover $\tilde{\rho}(\mathbf{r}, t)$ from a set of (k, t) -space measurements $\{s(\mathbf{k}_p, t_q)\}_{p=1, q=1}^{P, Q}$, also known as the *spatiospectral encodings*. Coil sensitivity can be included in (2) such that $s(\cdot)$ becomes $s_c(\cdot)$. Without the loss of generality, we ignore the coil index in the following discussion. An illustration of the spatio-spectral imaging problem is shown in Figure 2.

Challenges for MRSI

As shown in (2) and Figure 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., 1–2-kHz bandwidth for 3 T 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 location 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 spatio-spectral 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]

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

where N_m is the number of polarized spins that can be used to generate detectable MR signals for the m th molecule; $\hbar$ is the Planck constant; K is the Boltzmann constant; and T_s is the absolute temperature. As N_m values for the molecules of interest in MRSI, e.g., metabolites such as NAA, Cr, and Cho and neurotransmitters such as glutamate (Glu) and γ -aminobutyric acid (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) 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 ^1H -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

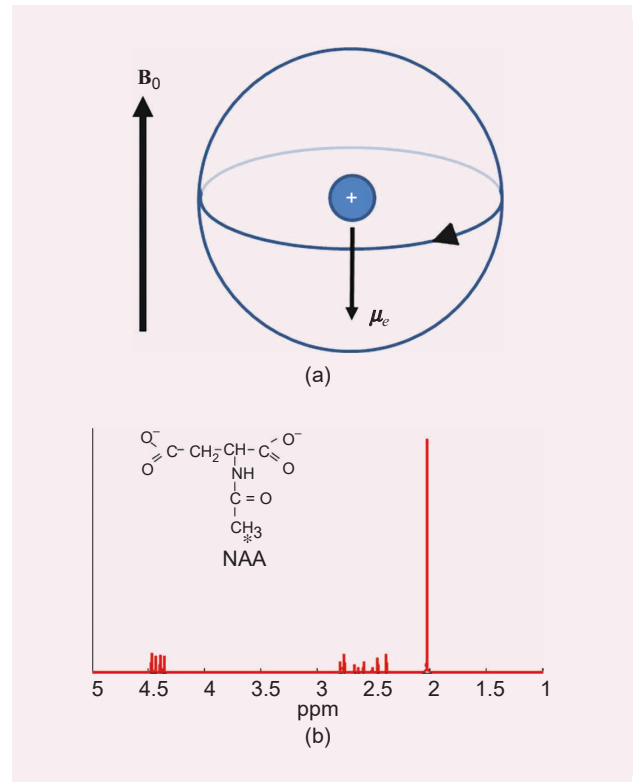


FIGURE 1. An illustration of chemical shift and the resulting frequency distribution. (a) The electron orbiting around the nucleus (denoted by the arrow on the circle) can be viewed as a small current, which generates a magnetic moment μ_e that opposes the main magnetic field B_0 , thus perturbing the magnetic field experienced “locally” by the nucleus and generating a small frequency shift. (b) The resonance structure of the molecule NAA (from quantum mechanical simulation) as a result of the chemical shift phenomenon.

tools to leverage this knowledge present new opportunities to tackle these challenges.

Physics-based spatio-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 does not need to be treated as an independent unknown during the imaging process. Given that there are only a finite number of frequencies originating from nuclear magnetic resonance-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

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

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 of the following:

- 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 ^1H -MRSI.
- 2) The frequency components from the same molecule have a 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.,

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

where c_m is the molecular concentrations and ϕ_m denotes a basis whose Fourier transform is the spectral structure of a specific molecule. ϕ_m can be predicted by QM simulations provided with the chemical shifts and J-coupling constants for a particular molecule (from a chemical database), and it 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 (DOF).

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 (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 (SRS) for MRSI. We review these developments in the following sections.

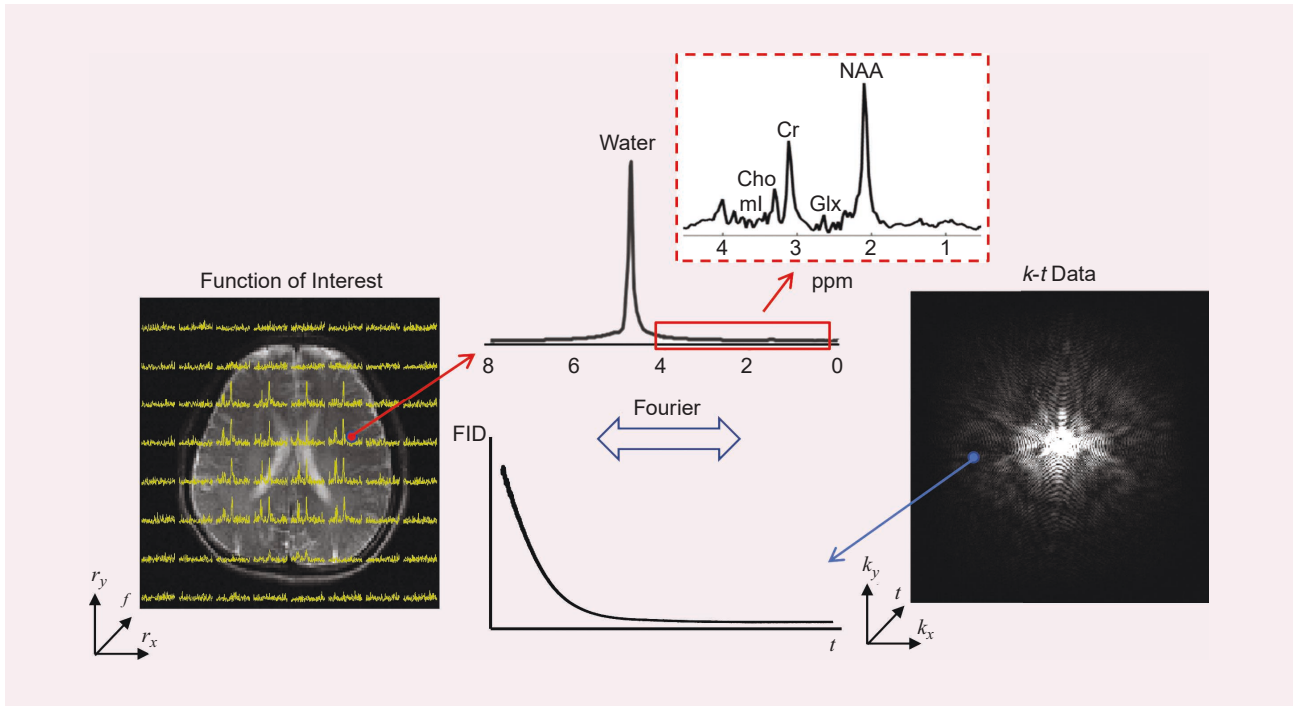


FIGURE 2. An illustration of the imaging problem with MRSI. Left: The desired high-dimensional spatio-spectral function is in the $(\mathbf{r}, f)$ domain where we aim to recover high-resolution spectra at each voxel. Right: Data are acquired in the $(\mathbf{k}, t)$ -space where FIDs are sampled for individual $\mathbf{k}$ -space locations. This imaging problem is inherently higher dimensional than conventional MRI, where only $\mathbf{k}$ -space is sampled. Furthermore, the signals from molecules of interest (e.g., NAA, Cr, Cho, etc.) are three orders of magnitude weaker than water (plots on the top of the middle column), making the problem more challenging. Glx: glutamate + glutamine; ml: myoinositol.

Subspace models

Under physiological conditions, the parameters c_m , $T_{2,m}^*$ ($\mathbf{r}$), and δf_m ($\mathbf{r}$) have values in a narrow range (e.g., $T_{2,m}^* \sim 30\text{--}100$ ms), and M is usually small, i.e., 10–15 for in vivo ^1H -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]

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

which represents the FIDs/spectra at individual voxels as linear combinations of a small set of basis functions $\{v_l(t)\}$, and 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 superresolution reconstruction [19]. Such a low-rank filtering approach requires the joint estimation of both the spatial coefficients $\{u_l(\mathbf{r})\}$ and subspace bases $\{v_l(t)\}$ (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 SRS challenges. We will focus on this type of reconstruction method in this review, and more detailed discussions on data acquisition can be found in [20], [21], [22].

Learning a nonlinear low-dimensional model can offer a more efficient representation of general in vivo MR spectra.

While subspace learning has been well studied in machine learning [23], the unique physical properties of each FID allow for unique training data generation and learning strategies [8]. More specifically, we can draw samples from prespecified 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 principal component analysis, independent component analysis, 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 (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.

Nonlinear manifold models

Linear subspace models have been demonstrated to be highly effective and enable 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 Figure 3. We generated spectra of 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 increase, as shown by the error spectra in black, while the nonlinear model maintains high accuracy for the same order. Therefore, learning a nonlinear

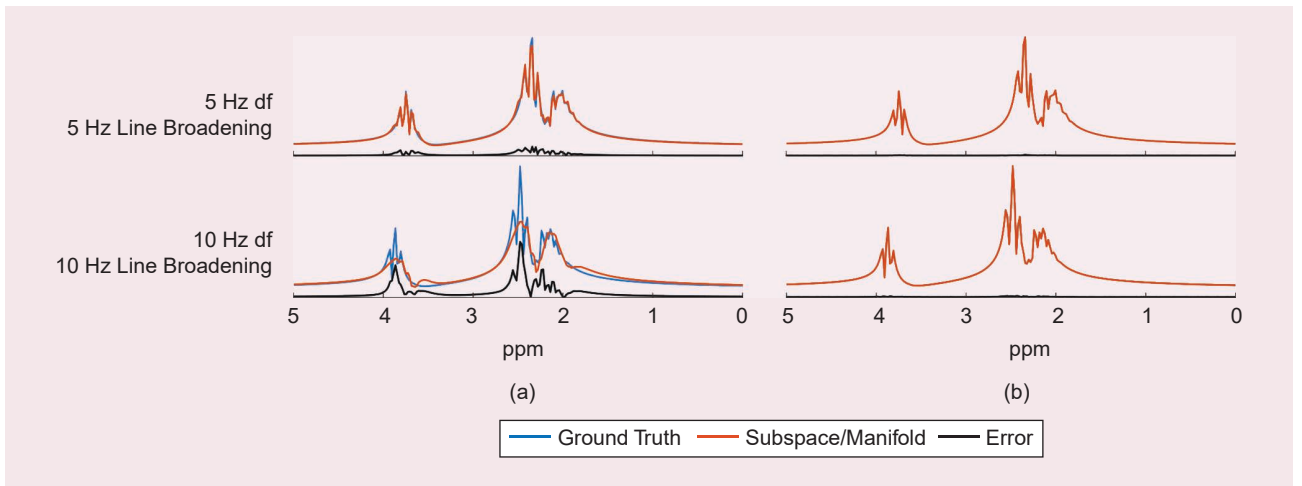


FIGURE 3. Linear versus nonlinear low-dimensional models of spectroscopic signals. (a) An approximation of the Glu spectra using subspaces learned from an ensemble of simulated data. The accuracy reduces as the ranges for the spectral parameters used to generate the training and testing data increase (from the top to the bottom rows). (b) An approximation by the nonlinear model learned from the same data, with the same model order of three. The color coding is specified in the figure legend.

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 previously, data-driven learning of nonlinear low-dimensional representations is possible, leveraging recent progress 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 ρ , 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., $\rho \approx D(E(\rho; \theta_e); \theta_d)$.

The idea has been extended to multi-echo time (multi-TE) spectroscopic data [29], and the improved representation efficiency of learned nonlinear low-dimensional models over linear subspace models has been shown for not only ^1H 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 (as well as spatio-spectral reconstruction) from noisy or

often sparsely sampled spatio-spectral encodings, which are reviewed in the next sections.

Spatiospectral reconstruction using learned models

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 (6) can be rewritten in a matrix form $\rho = \mathbf{U}\hat{\mathbf{V}}$ [17], where $\mathbf{U} \in \mathbb{C}^{N \times L}$ and $\hat{\mathbf{V}} \in \mathbb{C}^{L \times Q'}$ 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 lower number of DOF than ρ) [17]

$$\hat{\mathbf{U}} = \underset{\mathbf{U}}{\text{argmin}} \|\mathbf{d} - \Omega\{\mathbf{F}\mathbf{B} \odot (\mathbf{U}\hat{\mathbf{V}})\}\|_2^2 + \lambda R(\mathbf{U}, \hat{\mathbf{V}}) \quad (7)$$

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 spatio-spectral encodings. $R(\cdot)$ is a regularization that can be used to impose spatial constraint on $\mathbf{U}$ or the images at an 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 $\hat{\rho} = \hat{\mathbf{U}}\hat{\mathbf{V}}$. This joint subspace and spatial constrained reconstruction has achieved impressive improvement in resolution, SNR, and speed for both ^1H and X-nuclei MRSI in brain and muscle applications over traditional methods [9], [20], [31].

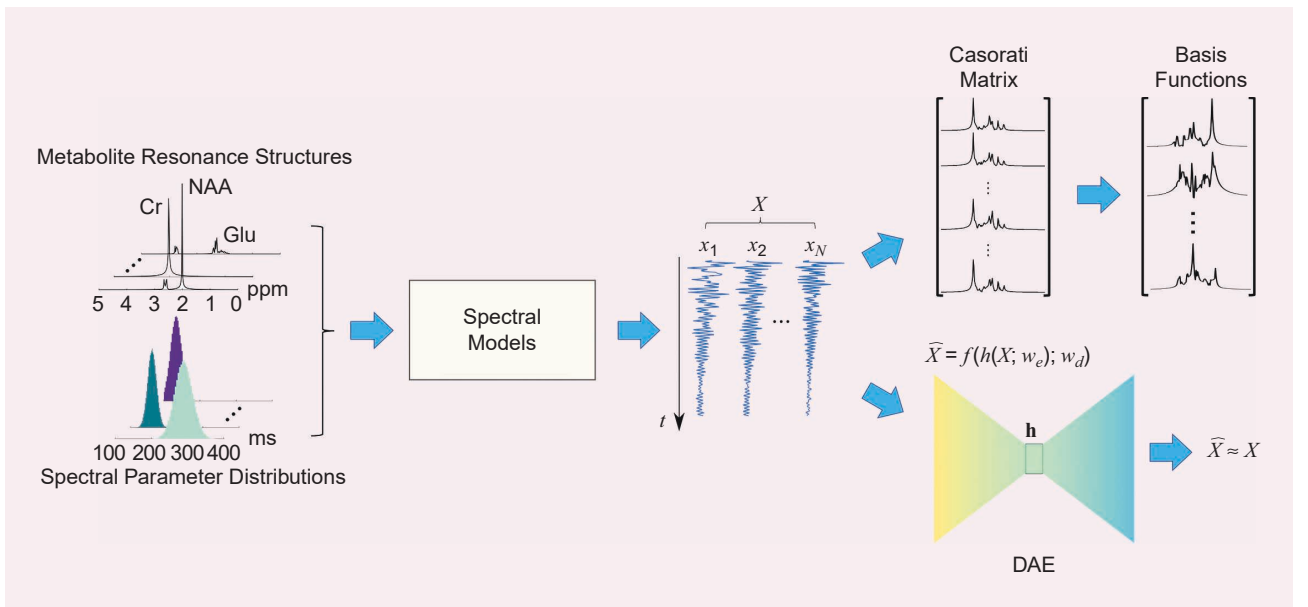


FIGURE 4. Physical-model-driven data generation and low-dimensional representation learning for MRSI data. Both QM-simulated resonance structures of individual molecules (metabolite basis, top of the left column) and spectral parameters sampled from empirical distributions (from literature values or experimental data, bottom of the left column) are fed into a spectral fitting model to generate a large quantity of training data ($\mathbf{X}$). 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.

It is possible to extend (7) to jointly update both $\mathbf{U}$ and $\mathbf{V}$ [30], which is a special form of low-rank matrix recovery. These types 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 incorporating future development of 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 spatio-spectral features [20]. A comparison of reconstructions from simulated MRSI data produced by direct low-rank filtering and using a learned subspace is shown in Figure 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.

Nonlinear manifold constrained reconstruction

As discussed previously, 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

$$\hat{\rho} = \underset{\rho}{\operatorname{argmin}} \|\mathbf{d} - \Omega\{\mathbf{FB} \odot \rho\}\|_2^2 + \lambda_1 \sum_{n=1}^N \|C(\rho_n) - \rho_n\|_2^2 + \lambda_2 R(\rho) \quad (8)$$

with $C(\cdot)$ representing the learned DAE $D(E(\cdot; \hat{\theta}_e); \hat{\theta}_d)$. While the first term imposes data consistency the same as previous formulations, the second term enforces the prior that the desired FIDs of interest at each voxel, ρ_n , should yield small representation errors for the learned low-dimensional model

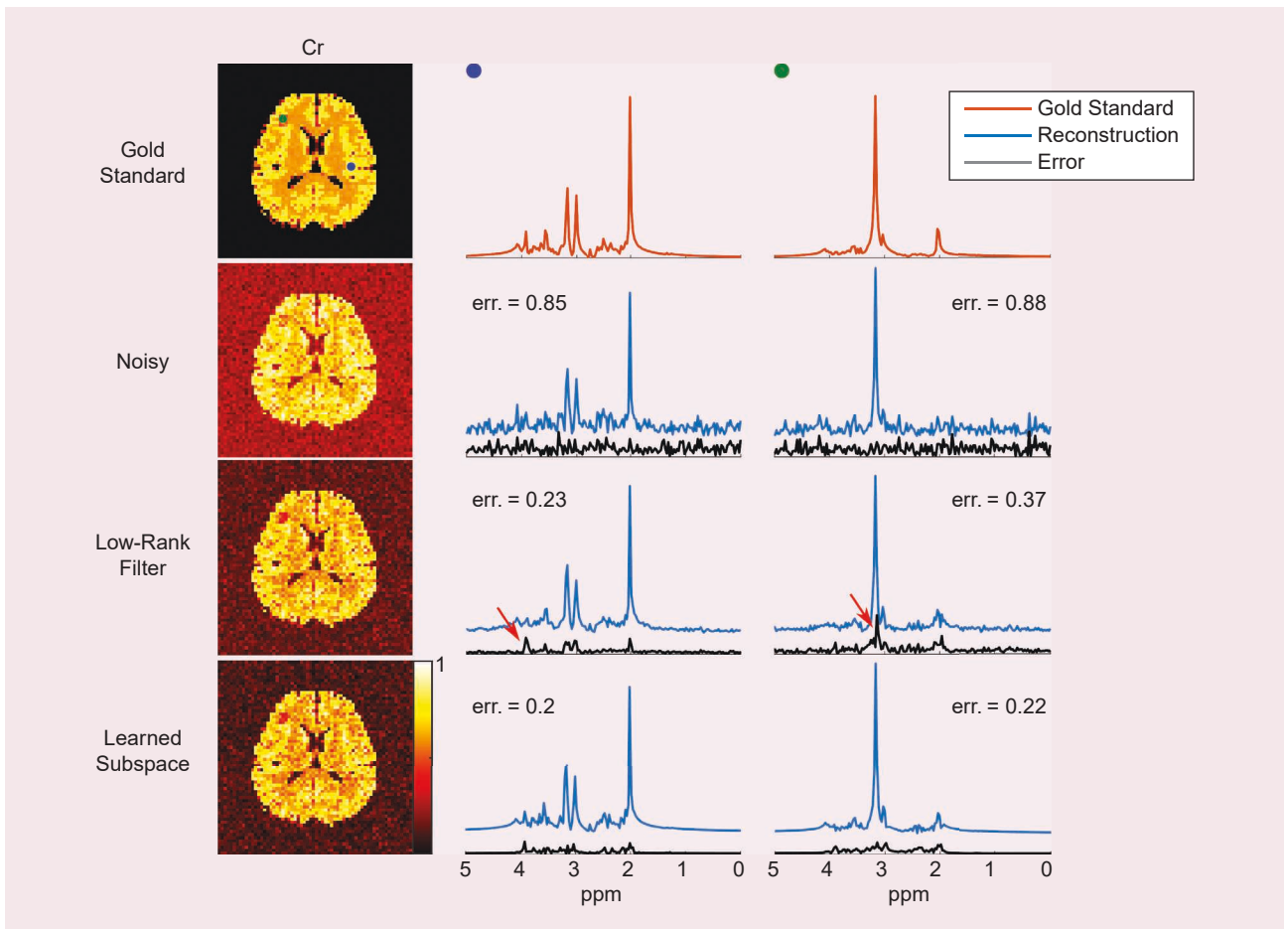


FIGURE 5. Reconstructions (simulated data) with and without using a prelearned subspace. The first column shows metabolite maps (Cr) from different cases [rows 1–4: 1) gold standard, 2) Fourier reconstruction, 3) low-rank filtering, and 4) reconstruction using a learned subspace], and the subsequent 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.

captured by $\mathcal{C}$. If $\mathcal{C}$ 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 (7)–(12). $R(\cdot)$ is an additional regularization term that can be flexibly incorporated and used to impose spatial constraints, either handcrafted or learned (see the next section for more discussion).

Similar to the reconstruction strategies discussed previously for the learned subspace, because we kept the forward encoding model and the learned $\mathcal{C}$ is imposing a prior on the underlying true FIDs, once pretrained, $\mathcal{C}$ 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 cannot 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 (8) is the need to invert the deep network (involving backpropagation) for each reconstruction, and thus, higher computational cost compared to the trained end-to-end mapping network that requires only 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 ^1H -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 (Learning Nonlinear Representation and Projection for Fast Constrained MRSI Reconstruction) method [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

$$\hat{\rho} = \underset{\rho \in D(\mathbf{Z})}{\operatorname{argmin}} \|\mathbf{d} - \Omega\{\mathbf{F}\mathbf{B} \odot \rho\}\|_2^2 + \lambda R(\rho) \quad (9)$$

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

$$\hat{\theta}_p = \underset{\theta_p}{\operatorname{argmin}} \frac{1}{J} \sum_{j=1}^J \epsilon_1 (E(\mathbf{x}_j; \theta_e), P(\tilde{\mathbf{x}}_j; \theta_p)) + \lambda \epsilon_2 (\mathbf{x}_j, D(P(\tilde{\mathbf{x}}_j; \theta_p); \theta_d)) \quad (10)$$

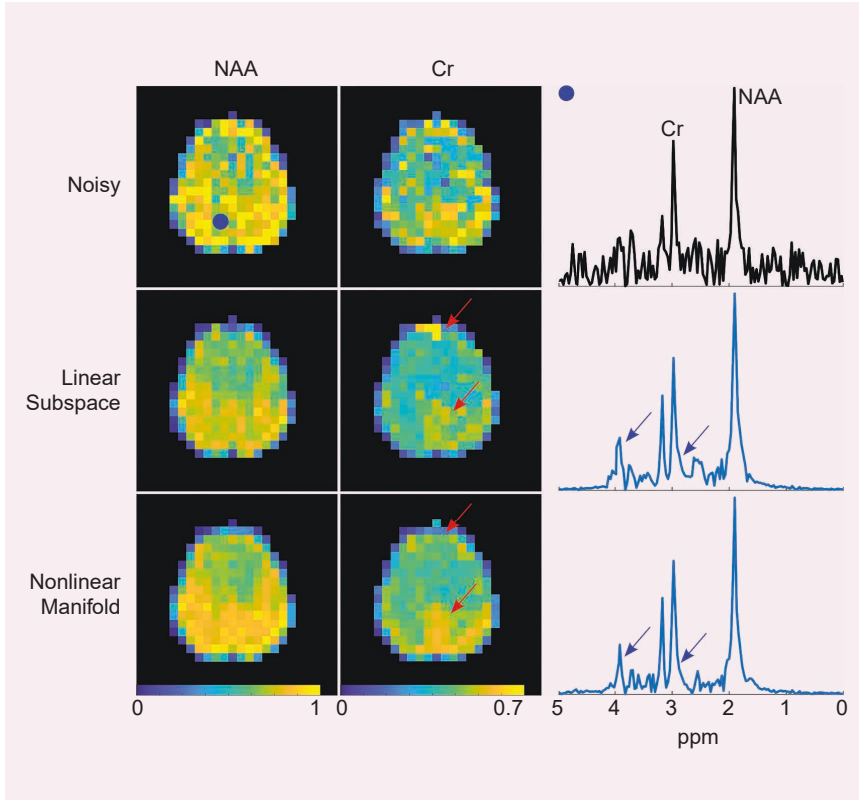


FIGURE 6. A set of in vivo results from reconstructions using learned subspace (linear subspace) and nonlinear models (nonlinear manifold). The latter produced metabolite maps with less noise contamination and better recovered tissue-dependent features (as indicated by the red arrows) as well as spectra with sharper line shapes (third column, blue arrows). A more thorough quantitative analysis without a gold standard can be found in [29].

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 $D(\cdot)$ are from the representation network discussed previously. 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 by 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 $\|\mathcal{C}(\rho_n) - \rho_n\|_2^2$ in the regularization formulation of (8) [38].

Integrating learned spatial and spectral priors

While the effectiveness of learned spectral priors incorporating both physics modeling and machine learning has 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 line shapes from tissue/brain-region-specific 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 T1-weighted MR image [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 the 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 $\boldsymbol{\theta}$, and the reconstruction in (7) is then changed to

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \|\mathbf{d} - \Omega\{\mathbf{FB} \odot (f(\mathbf{Z}; \boldsymbol{\theta}) \hat{\mathbf{V}})\}\|_2^2. \quad (11)$$

With a chosen network architecture (e.g., U-net in [39]), the parameter $\boldsymbol{\theta}$ is 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 with an 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 generative adversarial network (GAN)-regularized subspace reconstruction can be formulated as [41]

$$\hat{\mathbf{U}}, \hat{\mathbf{Z}} = \arg\min_{\mathbf{U}, \mathbf{Z}} \|\mathbf{d} - \Omega\{\mathbf{FB} \odot (\mathbf{U} \hat{\mathbf{V}})\}\|_2^2 + \lambda \|\mathbf{U} - f(\mathbf{Z}; \boldsymbol{\theta})\|_F^2 \quad (12)$$

where $f(\mathbf{Z}; \boldsymbol{\theta})$ is a pretrained GAN (e.g., using T1 and T2-weighted 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 $f(\cdot; \boldsymbol{\theta})$, StyleGAN combined with intermediate layer optimization 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 spatio-spectral encoding model and deep learning priors, in contrast to training an end-to-end network for direct mapping from noisy data/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 for retraining for each case. Meanwhile, the computation time required for the network parameter update or GAN inversion is one limitation/consideration.

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 ischemia–reperfusion to assess mitochondrial function in the muscle [44], [45]. Dynamic ^{13}C - and ^2H -MRSI, using isotope-enriched endogenous molecules, has 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, resulting in 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.

Low-rank tensor models with learned subspaces

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

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

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. Equation (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_T and Q). Similarly as the spectral basis $\{v_l(t)\}$ can be learned by incorporating a 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 unknowns in the imaging problem. Equation (13) can be further adapted into a molecule-dependent tensor model [51]

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

where M is the number of molecules of interest; L_m and S_m are the spectral and temporal model orders of the m th molecule; $c_{m,l,s}(\mathbf{r})$ combines the core tensor coefficients and molecule-dependent spatial basis and $s_m(\mathbf{r})$ incorporates spatial priors. This model offers up to two orders-of-magnitude reduction in the DOF [51] compared to the canonical truncated Fourier series representation, and thus, significantly improved resolution and SNR tradeoffs. Furthermore, the model in (14) provides

the flexibility to incorporate molecule-specific physics- and biochemistry-based priors [44].

The spectral basis can be obtained as described in the “Subspace Models” section, 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 the results of metabolic processes 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 available high-resolution reference images, further improving the performance (see [51] for more details). Accordingly, the dynamic MRSI reconstruction can be formulated as follows:

$$\hat{\mathbf{C}}_m = \underset{\{\mathbf{C}_m\}}{\operatorname{argmin}} \|\mathbf{d} - \mathcal{A}\{(\mathbf{s}_m \odot \mathbf{C}_m) \otimes \mathbf{V}_m \otimes \mathbf{G}_m\}\|_2^2 + \lambda R(\{\mathbf{C}_m\}) \quad (15)$$

where $\mathcal{A}$ describes the forward model as discussed previously (capturing the field inhomogeneity effects and spatio-spectral

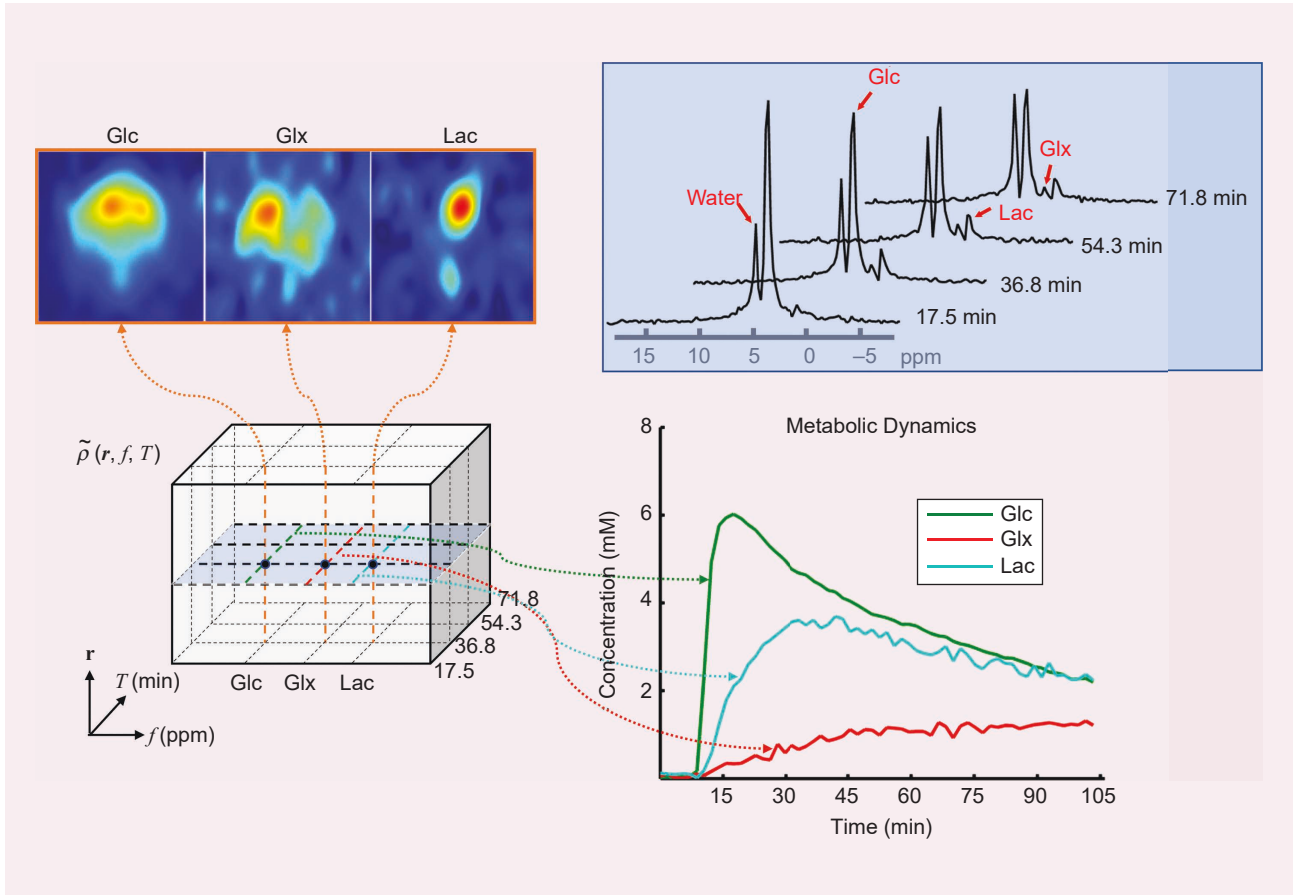


FIGURE 7. The dynamic MRSI problem (dynamic ^2H -MRSI of a rat brain in this case). The center cube illustrates the high-dimensional image function of interest. Orange lines taken from different frequencies (f) and a single time point (T) yield various metabolite maps (top left, spatial dimension); the blue plane represents time-resolved spectra (top right, spectral-temporal dimensions) at a single voxel ($\mathbf{r}$); and lines from different f 's and a single $\mathbf{r}$ capture metabolic dynamics (bottom right). The low-rank tensor-based reconstruction with learned manifold constraints produced an impressive combination of SRS [48], i.e., a 1.8-min frame rate at a resolution of $17 \times 17 \times 5$ matrix size over $28 \times 28 \times 24$ mm³ field of view (FOV). Glc: glucose; Lac: lactate.

encoding process), and $\mathbf{C}_m, \mathbf{V}_m, \mathbf{G}_m$ 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 $\mathbf{C}_m \in \mathbb{C}^{N \times L_m \times S_m}$, $\mathbf{V}_m \in \mathbb{C}^{L_m \times Q'}$, $\mathbf{G}_m \in \mathbb{C}^{S_m \times N_T}$, and $\mathbf{s}_m \in \mathbb{C}^{N \times 1}$. $\odot$ 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, (15) was solved using an 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, (14) can be reformulated by imposing strict separability between the spatio-spectral and spatiotemporal variations [48], i.e.,

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

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 the 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 (15) can be formulated.

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 (15) or (16). For example, the learned network can be used to regularize individual FIDs at every time point, i.e.,

$$\begin{aligned} \hat{\mathbf{C}}_m, \hat{\rho} = \arg \min_{\{\mathbf{C}, \rho\}} & \|\mathbf{d} - \mathcal{A}\{\rho\}\|_2^2 + \lambda_1 \sum_{T=1}^{N_T} \sum_{n=1}^N \|\mathcal{C}(\rho_{n,T}) - \rho_{n,T}\|_2^2 \\ & + \lambda_2 R(\{\mathbf{C}_m\}), \\ \text{s.t. } \rho &= (\mathbf{s}_m \odot \mathbf{C}_m) \otimes \mathbf{V}_m \otimes \mathbf{G}_m. \end{aligned} \quad (17)$$

This formulation assumes that the desired FIDs at each voxel and each time point (i.e., $\rho_{n,T}$) 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 (16) offers the additional flexibility to learn and incorporate low-dimensional manifold constraints on molecule-specific temporal dynamics. Specifically, a set of molecule-dependent DAEs, denoted as $\mathcal{G}_m(\cdot)$, was trained (using both synthetic and experimental dynamic data) and integrated into the following low-rank tensor-based reconstruction problem [48]:

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

where $\mathbf{A}_m$ and $\mathbf{B}_m$ are matrix forms of $\{a_{m,l}(\mathbf{r})\}$ and $\{b_{m,s}(\mathbf{r})\}$ with $\mathbf{A}_m \in \mathbb{C}^{N \times L_m}$ and $\mathbf{B}_m \in \mathbb{C}^{N \times S_m}$. $\circ$ denotes the operation defined in (16), and λ_1 and λ_2 are regularization parameters controlling the tradeoff 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 (18) can be solved via alternating minimization of $\mathbf{A}_m$ and $\mathbf{B}_m$, which also involves the 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 the “Nonlinear Manifold Constrained Reconstruction” section) may be helpful in addressing this issue. A set of denoising dynamic ^2H -MRSI results is shown in Figure 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.

The motion issue

The formulations discussed previously 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].

Spectral quantification using learned models

With the ability to reconstruct high-SNR high-resolution spatiotemporal functions 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 (5) and perform parameter estimation voxel by voxel, e.g., through a nonlinear least squares (NLLS) fitting [53]

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

where $\hat{\rho}_n$ is the reconstructed FID at the n th 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, eddy current effects, etc. A Gaussian line shape 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 (19), but the problem is highly nonlinear and sensitive to noise and model mismatch. It can be time consuming to solve these voxel-wise NLLS equations 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 (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 $\{v_{lm}(t)\}$, i.e.,

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

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

$$\{\hat{\mathbf{U}}_m\} = \underset{(\mathbf{U}_m)}{\operatorname{argmin}} \left\| \hat{\rho} - \sum_{m=1}^M \mathbf{U}_m \mathbf{V}_m \right\|_F^2 + \sum_{m=1}^M \lambda_m R_m(\mathbf{U}_m) \quad (21)$$

where $\mathbf{U}_m$ is a molecule-specific spatial coefficient 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

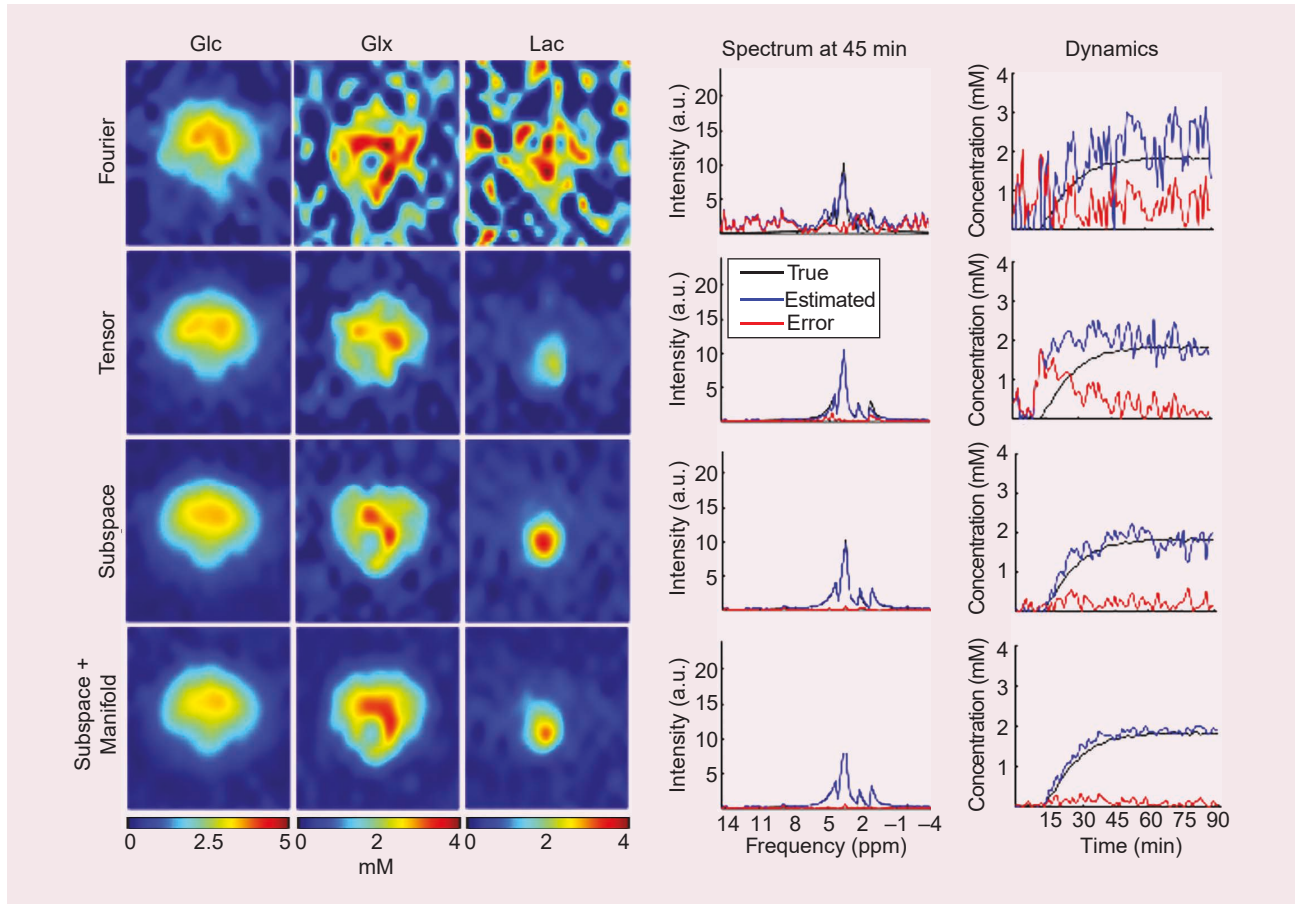


FIGURE 8. A set of dynamic ^2H -MRSI results from a rat brain by the standard Fourier reconstruction (Fourier); a direct low-rank tensor filtering of the noisy data (Tensor); a tensor reconstruction with learned subspaces (Subspace); and joint subspace model and manifold regularization (Subspace+Manifold) [48]. For each method, maps of three metabolites (columns 1–3), localized spectra at a particular time point (fourth column) and temporal dynamics for the same voxel (fifth 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. a.u.: arbitrary unit.

subspace model at individual voxels and less sensitivity to noise by transforming the nonlinear fitting into low-dimensional linear problems. It also allows for the easier incorporation of spatial priors on individual components through $R_m(\cdot)$. The 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 DOF in $\mathbf{U}_m$.

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

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \|\hat{\rho} - f(t; \mathcal{C}(\hat{\rho}; \theta))\|_2^2 \quad (22)$$

where the spectral parameter α was 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 the 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, the 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 line shape-corrected spectra from noisy data with macromolecule and baseline “contaminations” and line shape distortion. Estimation of metabolite concentrations from the corrected metabolite-only spectra can then be easily done using a linear least-squares fit to a metabolite basis. All these methods are voxel wise and thus difficult to incorporate spatial constraints. The learned mapping can be sensitive to changes in acquisition parameters, such as the spectral bandwidth, resolution, SNR, etc.

Future directions

Recent advances in computational MRSI using physics-based machine learning have demonstrated a good potential for addressing high-dimensional MRSI reconstruction and quantification problems, achieving impressive performance in the combination of SRS 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], [57], [58], [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 the “Integrating Learned Spatial and Spectral Priors” section 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 toward quantitative MRSI for clinical and basic science applications remains under investigated. A number of sophisticated parametric models have been presented, mostly for SVS, but they often yield significant voxel-to-voxel estimation variations and are computationally demanding for high-resolution data. The notion of voxel-by-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 a reduced number of parameters), should be developed with potentially significant practical impact.

We have focused this review on MRSI, but the methodologies discussed here are also adaptable to other high-dimensional

The notion of voxel-by-voxel fitting should continue to be challenged by new ways to integrate spatial constraints with deep learning-based solutions.

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. The 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 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 the 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 progress in connecting and fusing imaging information from different modalities across spatiotemporal scales by integrating physics-based modeling and data-driven learning.

Acknowledgment

This work was supported in part by NSF Grant NSF-CBET 1944249 and NIH Grant 1R35GM142969-01. We thank Yahang Li and Yudu Li for their help on the figures.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Illinois Urbana-Champaign Institutional Review Board under protocol number 20132 and University of Minnesota Institutional Animal Care and Use Committee under protocol number 1903-36898A.

Authors

Fan Lam (fanlam1@illinois.edu) received his Ph.D. degree in electrical and computer engineering from the University of Illinois Urbana-Champaign (UIUC) in 2015. He is currently an assistant professor in the Department of Bioengineering at UIUC, Urbana, IL 61801 USA, and a full-time faculty member with the Beckman Institute for Advanced Science and Technology, Urbana, IL 61801 USA. His awards and recognitions include the Best Student Paper Award from the IEEE-ISBI; Robert T. Chien Memorial Award from UIUC; NSF CAREER Award; NIH-NIBIB Trailblazer Award; and NIH-NIGMS MIRA R35 Award. His research focuses on developing advanced MR-based molecular imaging and multimodal brain mapping methods and their applications to the study of brain function. He is a Senior Member of IEEE.

Xi Peng (peng.xi@mayo.edu) received his Ph.D. degree in electrical engineering in 2012 from Wuhan University, Hubei, China. He was an associate professor at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, and a principal investigator on several major grants. He is currently an assistant professor at the Mayo Clinic, Rochester, MN 55905 USA. He received the Best Paper Award and Outstanding Research Award from ISMRM and OCSMRM, respectively. He is working on accelerated methods for

ultrahigh resolution magnetic resonance spectroscopic imaging, magnetic resonance elastography, and machine learning-based approaches for parallel imaging and quantitative susceptibility mapping. His research focuses on the innovation of MRI technologies in the development of acquisition and computational tools. He is a Member of IEEE.

Zhi-Pei Liang is the Franklin W. Woeltge Professor of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. His work has been recognized by a number of prestigious awards, including the Sylvia Sorkin Greenfield Award (Medical Physics); Otto Schmitt Award (IFMBE); Technical Achievement Award (IEEE-EMBS); and Gold Medal of ISMRM. He was elected to the International Academy of Medical and Biological Engineering in 2012 and to the U.S. National Academy of Inventors in 2021. His research focuses on the general areas of magnetic resonance imaging, signal processing, and machine learning. He is a Fellow of IEEE, ISMRM, and AIMBE.

References

- [1] P. C. Lauterbur, D. M. Kramer, W. V. House, and C.-N. Chen, "Zeugmatographic high resolution nuclear magnetic resonance spectroscopy: Images of chemical inhomogeneity within macroscopic objects," *J. Amer. Chem. Soc.*, vol. 97, no. 23, pp. 6866–6868, Nov. 1975, doi: 10.1021/ja00856a046.
- [2] T. R. Brown, B. M. Kincaid, and K. Ugurbil, "NMR chemical shift imaging in three dimensions," *Proc. Nat. Acad. Sci.*, vol. 79, no. 11, pp. 3523–3526, 1982, doi: 10.1073/pnas.79.11.3523.
- [3] B. Ross and T. Michaelis, "Clinical applications of magnetic resonance spectroscopy," *Magn. Reson. Q.*, vol. 10, no. 4, pp. 191–247, Dec. 1994.
- [4] R. A. de Graaf, *In Vivo NMR Spectroscopy: Principles and Techniques*. Hoboken, NJ, USA: Wiley, 2018.
- [5] Y. Li et al., "Fast high-resolution metabolic imaging of acute stroke with 3D magnetic resonance spectroscopy," *Brain*, vol. 143, no. 11, pp. 3225–3233, Dec. 2020, doi: 10.1093/brain/awaa264.
- [6] W. Bogner, R. Otazo, and A. Henning, "Accelerated MR spectroscopic imaging—A review of current and emerging techniques," *NMR Biomed.*, vol. 34, no. 5, p. e4314, May 2021, doi: 10.1002/nbm.4314.
- [7] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Arlington, VA, USA, 2007, pp. 988–991, doi: 10.1109/NFBI-ICFBI.2007.4387720.
- [8] F. Lam, Y. Li, R. Guo, B. Clifford, and Z.-P. Liang, "Ultrafast magnetic resonance spectroscopic imaging using SPICE with learned subspaces," *Magn. Reson. Med.*, vol. 83, no. 2, pp. 377–390, Feb. 2020, doi: 10.1002/mrm.27980.
- [9] C. Ma, B. Clifford, Y. Liu, Y. Gu, F. Lam, X. Yu, and Z.-P. Liang, "High-resolution dynamic ^3P -MRSI using a low-rank tensor model," *Magn. Reson. Med.*, vol. 78, no. 2, pp. 419–428, 2017, doi: 10.1002/mrm.26762.
- [10] Z. Iqbal, D. Nguyen, G. Hangel, S. Motyka, W. Bogner, and S. Jiang, "Super-resolution 1H magnetic resonance spectroscopic imaging utilizing deep learning," *Frontiers Oncology*, vol. 9, p. 1010, Oct. 2019, doi: 10.3389/fonc.2019.01010.
- [11] X. Qu, Y. Huang, H. Lu, T. Qiu, D. Guo, T. Agback, V. Orekhov, and Z. Chen, "Accelerated nuclear magnetic resonance spectroscopy with deep learning," *Angew. Chem. Int. Ed. English*, vol. 59, no. 26, pp. 10,297–10,300, Jun. 2020, doi: 10.1002/anie.201908162.
- [12] Y. Li, F. Lam, B. Clifford, and Z. Liang, "A subspace approach to spectral quantification for MR spectroscopic imaging," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 10, pp. 2486–2489, Oct. 2017, doi: 10.1109/TBME.2017.2741922.
- [13] S. S. Gurbani, S. Sheriff, A. A. Maudsley, H. Shim, and L. A. Cooper, "Incorporation of a spectral model in a convolutional neural network for accelerated spectral fitting," *Magn. Reson. Med.*, vol. 81, no. 5, pp. 3346–3357, May 2019, doi: 10.1002/mrm.27641.
- [14] H. Lee and H. Kim, "Bayesian deep learning-based ^1H -MRS of the brain: Metabolite quantification with uncertainty estimation using monte Carlo dropout," in *Proc. Int. Soc. Magn. Reson. Med.*, 2021, p. 2014.
- [15] Z.-P. Liang and P. C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. New York, NY, USA: IEEE Press, 2000.
- [16] H. Barkhuysen, R. de Beer, and D. van Ormondt, "Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals," *J. Magn. Reson.*, vol. 73, no. 3, pp. 553–557, Jul. 1987, doi: 10.1016/0022-2364(87)90023-0.

- [17] F. Lam and Z.-P. Liang, "A subspace approach to high-resolution spectroscopic imaging," *Magn. Reson. Med.*, vol. 71, no. 4, pp. 1349–1357, Apr. 2014, doi: 10.1002/mrm.25168.
- [18] H. M. Nguyen, X. Peng, M. N. Do, and Z.-P. Liang, "Denoising MR spectroscopic imaging data with low-rank approximations," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 78–89, Jan. 2013, doi: 10.1109/TBME.2012.2223466.
- [19] J. Kasten, A. Klauser, F. Lazeyras, and D. Van De Ville, "Magnetic resonance spectroscopic imaging at superresolution: Overview and perspectives," *J. Magn. Reson.*, vol. 263, pp. 193–208, Feb. 2016, doi: 10.1016/j.jmr.2015.11.003.
- [20] F. Lam, C. Ma, B. Clifford, C. L. Johnson, and Z.-P. Liang, "High-resolution (1) H-MRSI of the brain using SPICE: Data acquisition and image reconstruction," *Magn. Reson. Med.*, vol. 76, no. 4, pp. 1059–1070, Oct. 2016, doi: 10.1002/mrm.26019.
- [21] X. Peng, F. Lam, Y. Li, B. Clifford, and Z.-P. Liang, "Simultaneous QSM and metabolic imaging of the brain using SPICE," *Magn. Reson. Med.*, vol. 79, no. 1, pp. 13–21, Aug. 2018, doi: 10.1002/mrm.26972.
- [22] R. Guo, Y. Zhao, Y. Li, T. Wang, Y. Li, B. Sutton, and Z.-P. Liang, "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, Feb. 2021, doi: 10.1002/mrm.28459.
- [23] L. Parsons, E. Haque, and H. Liu, "Subspace clustering for high dimensional data: A review," *SIGKDD Explor. Newslett.*, vol. 6, no. 1, pp. 90–105, Jun. 2004, doi: 10.1145/1007730.1007731.
- [24] H. H. Lee and H. Kim, "Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain," *Magn. Reson. Med.*, vol. 82, no. 1, pp. 33–48, Jul. 2019, doi: 10.1002/mrm.27727.
- [25] Y. Li, Z. Wang, R. Sun, and F. Lam, "Separation of metabolites and macromolecules for short-TE ¹H-MRSI using learned component-specific representations," *IEEE Trans. Med. Imag.*, vol. 40, no. 4, pp. 1157–1167, Apr. 2021, doi: 10.1109/TMI.2020.3048933.
- [26] F. Lam, Y. Li, and X. Peng, "Constrained magnetic resonance spectroscopic imaging by learning nonlinear low-dimensional models," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 545–555, Mar. 2020, doi: 10.1109/TMI.2019.2930586.
- [27] P. Tiwari, M. Rosen, and A. Madabhushi, "Consensus-locally linear embedding (C-LLE): Application to prostate cancer detection on magnetic resonance spectroscopy," in *Proc. 2008 MICCAI*. Berlin, Heidelberg: Springer-Verlag, pp. 330–338.
- [28] G. Yang, F. Raschke, T. R. Barrick, and F. A. Howe, "Manifold learning in MR spectroscopy using nonlinear dimensionality reduction and unsupervised clustering," *Magn. Reson. Med.*, vol. 74, no. 3, pp. 868–878, Sep. 2015, doi: 10.1002/mrm.25447.
- [29] Y. Li, Z. Wang, and F. Lam, "SNR enhancement for multi-TE MRSI using joint low-dimensional model and spatial constraints," *IEEE Trans. Biomed. Eng.*, early access, 2022, doi: 10.1109/TBME.2022.3161417.
- [30] A. Klauser, S. Courvoisier, J. Kasten, M. Kocher, M. Guerquin-Kern, D. Van De Ville, and F. Lazeyras, "Fast high-resolution brain metabolite mapping on a clinical 3T MRI by accelerated 1H-FID-MRSI and low-rank constrained reconstruction," *Magn. Reson. Med.*, vol. 81, no. 5, pp. 2841–2857, May 2019, doi: 10.1002/mrm.27623.
- [31] Y. Chen, Y. Li, and Z. Xu, "Improved low-rank filtering of MR spectroscopic imaging data with pre-learned subspace and spatial constraints," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 8, pp. 2381–2388, Aug. 2020, doi: 10.1109/TBME.2019.2961698.
- [32] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: K-T SLR," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011, doi: 10.1109/TMI.2010.2100850.
- [33] J. D. Trzasko, "Exploiting local low-rank structure in higher-dimensional KRI applications," in *Wavelets and Sparsity XV*, vol. 8858, D. V. D. Ville, V. K. Goyal, and M. Papadakis, Eds. Bellingham, WA, USA: SPIE, 2013, pp. 551–558.
- [34] T. Zhang, J. M. Pauly, and I. R. Levesque, "Accelerating parameter mapping with a locally low rank constraint," *Magn. Reson. Med.*, vol. 73, no. 2, pp. 655–661, Feb. 2015, doi: 10.1002/mrm.25161.
- [35] E. Milshteyn, C. Morze, G. D. Reed, H. Shang, P. J. Shin, P. E. Z. Larson, and D. B. Vigneron, "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, May 2018.
- [36] J. P. Haldar, "Low-rank modeling of local k-space neighborhoods (LORAKS) for constrained MRI," *IEEE Trans. Med. Imag.*, vol. 33, no. 3, pp. 668–681, Mar. 2014, doi: 10.1109/TMI.2013.2293974.
- [37] L. Vizioli, S. Moeller, L. Dowdle, M. Akcakaya, F. De Martino, E. Yacoub, and K. Ugurbil, "Lowering the thermal noise barrier in functional brain mapping with magnetic resonance imaging," *Nature Commun.*, vol. 12, no. 1, p. 5181, Aug. 2021, doi: 10.1038/s41467-021-25431-8.
- [38] Y. Li, L. Ruhm, A. Henning, and F. Lam, "LeaRning nonlinear representation and projection for fast constrained MRSI reconstruction (RAIISE)," in *Proc. Int. Soc. Magn. Reson. Med.*, 2022, p. TBD.
- [39] K. Gong, P. K. Han, T. Marin, G. E. Fakhri, Q. Li, and C. Ma, "High resolution MR spectroscopic imaging using deep image prior constrained subspace modeling," in *Proc. Int. Soc. Magn. Reson. Med.*, 2020, p. 388.
- [40] D. Ulyanov, A. Vedaldi, and V. S. Lempitsky, "Deep image prior," in *Proc. 2018 IEEE/CVF CVPR*, pp. 9446–9454.
- [41] R. Zhao, Z. Wang, and F. Lam, "Generative image prior constrained subspace reconstruction for high-resolution MRSI," in *Proc. Int. Soc. Magn. Reson. Med.*, 2022, p. 3500.
- [42] G. Daras, J. Dean, A. Jalal, and A. Dimakis, "Intermediate layer optimization for inverse problems using deep generative models," in *Proc. 38th Int. Conf. Mach. Learn.*, 2021, vol. 139, pp. 2421–2432.
- [43] W. Xia, Y. Zhang, Y. Yang, J.-H. Xue, B. Zhou, and M.-H. Yang, "GAN inversion: A survey," 2021, *arXiv:2101.05278*.
- [44] Y. Liu, Y. Gu, and X. Yu, "Assessing tissue metabolism by phosphorous-31 magnetic resonance spectroscopy and imaging: A methodology review," *Quantitative Imag. Med. Surgery*, vol. 7, no. 6, pp. 707–726, Dec. 2017, doi: 10.21037/qims.2017.11.03.
- [45] A. Santos-Diaz and M. D. Noseworthy, "Phosphorus magnetic resonance spectroscopy and imaging (³¹P-MRS/MRSI) as a window to brain and muscle metabolism: A review of the methods," *Biomed. Signal Process. Control*, vol. 60, p. 101967, Jul. 2020, doi: 10.1016/j.bspc.2020.101967.
- [46] J. R. Brender et al., "Dynamic imaging of glucose and lactate metabolism by ¹³C-MRS without hyperpolarization," *Scientific Rep.*, vol. 9, no. 1, p. 3410, Mar. 2019, doi: 10.1038/s41598-019-38981-1.
- [47] H. M. D. Feyter et al., "Deuterium metabolic imaging (DMI) for MRI-based 3d mapping of metabolism in vivo," *Scientific Rep.*, vol. 4, no. 8, p. eaat7314, Aug. 2018, doi: 10.1126/sciadv.aat7314.
- [48] Y. Li et al., "Machine learning-enabled high-resolution dynamic deuterium MR spectroscopic imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3879–3890, Dec. 2021, doi: 10.1109/TMI.2021.3101149.
- [49] J. He, Q. Liu, A. G. Christodoulou, C. Ma, F. Lam, and Z.-P. Liang, "Accelerated high-dimensional MR imaging with sparse sampling using low-rank tensors," *IEEE Trans. Med. Imag.*, vol. 35, no. 9, pp. 2119–2129, Apr. 2016, doi: 10.1109/TMI.2016.2550204.
- [50] A. G. Christodoulou, J. L. Shaw, C. Nguyen, Q. Yang, Y. Xie, N. Wang, and D. Li, "Magnetic resonance multitasking for motion-resolved quantitative cardiovascular imaging," *Nature Biomed. Eng.*, vol. 2, no. 4, pp. 215–226, Apr. 2018, doi: 10.1038/s41551-018-0217-y.
- [51] B. Clifford et al., "High-resolution dynamic ³¹P-MR spectroscopic imaging for mapping mitochondrial function," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 10, pp. 2745–2753, Oct. 2020, doi: 10.1109/TBME.2020.2969892.
- [52] W. Bogner, A. T. Hess, B. Gagoski, M. D. Tisdall, A. J. van der Kouwe, S. Trattnig, B. Rosen, and O. C. Andronesi, "Real-time motion- and B0-correction for LASER-localized spiral-accelerated 3D-MRSI of the brain at 3 T," *NeuroImage*, vol. 88, pp. 22–31, Mar. 2014, doi: 10.1016/j.neuroimage.2013.09.034.
- [53] L. Vanhamme, A. van den Boogaart, and S. Van Huffel, "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, Nov. 1997, doi: 10.1006/jmre.1997.1244.
- [54] C. Ma and G. E. Fakhri, "MRSI spectral quantification using linear tangent space alignment (LTSA)-based manifold learning," in *Proc. Int. Soc. Magn. Reson. Med.*, 2022, p. 243.
- [55] F. Turco, I. Zubak, and J. Slotboom, "Deep learning based MRS metabolite quantification: CNN and RESNET versus non linear least square fitting," in *Proc. Int. Soc. Magn. Reson. Med.*, 2021, p. 2018.
- [56] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2Noise: Learning image restoration without clean data," in *Proc. 2018 35th Int. Conf. Mach. Learn. (ICML)*, vol. 7, pp. 4620–4631.
- [57] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Ugurbil, and M. Akcakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," *Magn. Reson. Med.*, vol. 84, no. 6, pp. 3172–3191, Dec. 2020, doi: 10.1002/mrm.28378.
- [58] J. Liu, Y. Sun, C. Eldeniz, W. Gan, H. An, and U. S. Kamilov, "RARE: Image reconstruction using deep priors learned without groundtruth," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 1088–1099, Oct. 2020, doi: 10.1109/JSTSP.2020.2998402.
- [59] M. Akcakaya, B. Yaman, H. Chung, and J. C. Ye, "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, Mar. 2022, doi: 10.1109/MSP.2021.3119273.
- [60] L. Chen et al., "In vivo imaging of phosphocreatine with artificial neural networks," *Nature Commun.*, vol. 11, p. 1072, Feb. 2020, doi: 10.1038/s41467-020-14874-0.
- [61] B. Spengler, "Mass spectrometry imaging of biomolecular information," *Analytical Chem.*, vol. 87, no. 1, pp. 64–82, Jan. 2015, doi: 10.1021/ac504543v.