qModeL: A Plug-and-Play Model-based Reconstruction for Highly Accelerated Multi-shot Diffusion MRI Using Learned Priors

Merry Mani[1], *Vincent A Magnotta*[1], *Mathews Jacob*[2]

[1]*Department of Radiology, University of Iowa, Iowa City, Iowa*

[2]*Department of Electrical and Computer Engineering, University of Iowa, Iowa City, Iowa*

February 5, 2021

Correspondence to :

Merry Mani

L309 Papajohn Biomedical Discovery Building,

169 Newton Road

Iowa City, Iowa, 52242

email: merry-mani@uiowa.edu

phone number: (319) 335-9569.

Word count : about 5000

figures+ tables count : 10

Abstract

Purpose: To introduce a joint reconstruction method for highly under-sampled multi-shot diffusion weighted (msDW) scans.

Methods: Multi-shot EPI methods enable higher spatial resolution for diffusion MRI, but at the expense of long scan-time. Highly accelerated msDW scans are needed to enable their utilization in advanced microstructure studies, which require high q-space coverage. Previously, joint k-q under-sampling methods coupled with compressed sensing were shown to enable very high acceleration factors. However, the reconstruction of this data using sparsity priors is challenging and is not suited for multi-shell data. We propose a new reconstruction that recovers images from the combined k-q data jointly. The proposed qModeL reconstruction brings together the advantages of model-based iterative reconstruction and machine learning, extending the idea of plug-and-play algorithms. Specifically, qModeL works by pre-learning the signal manifold corresponding to the diffusion measurement space using deep learning. The pre-learned manifold prior is incorporated into a model-based reconstruction to provide a voxel-wise regularization along the q-dimension during the joint recovery. Notably, the learning does not require in-vivo training data and is derived exclusively from biophysical modeling. Additionally, a plug-and-play total variation denoising provides regularization along the spatial dimension. The proposed framework is tested on k-q under-sampled single-shell and multi-shell msDW acquisition at various acceleration factors.

Results:

The qModeL joint reconstruction is shown to recover DWIs from 8-fold accelerated msDW acquisitions with error less than 5% for both single-shell and multi-shell data. Advanced microstructural analysis performed using the under-sampled reconstruction also report reasonable accuracy.

Conclusion: qModeL enables the joint recovery of highly accelerated multi-shot dMRI utilizing learning-based priors. The bio-physically driven approach enables the use of accelerated multi-shot imaging for multi-shell sampling and advanced microstructure studies.

Keywords: Deep Learning, autoencoder neural network, machine learning, multi-shot diffusion, multi-shell, NODDI, k-q acceleration

Introduction

Diffusion weighted magnetic resonance imaging (dMRI) is a widely used neuroimaging technique that makes use of the microscopic diffusing motion of water molecules in the neuronal tissue as a signal source (1). This water diffusion, whose characteristic length scales are of the order of microns, are encoded using large bipolar gradients to form *diffusion-weighted* images (DWIs). Because of its sensitivity to the *micron*-scale water displacement, dMRI is ideally suited to probe tissue *micro*structure, as well as change in the *micro*structure in relation to pathologies, normal brain development and aging (2).

The dMRI signal is typically spatially encoded using single-shot echo-planar imaging (ssEPI) techniques, primarily due to its high time-efficiency. However, the ssEPI readout leads to undesirable artifacts arising from B0-related and T2*-related effects. To minimize artifacts, the ssEPI readout duration is often restricted. The resulting low k-space coverage leads to voxel sizes that are orders of magnitude larger than the size of the underlying microstructures being probed. Thus, a large portion of microscopic information is obscured due to the coarse resolution, which leads to partial volume confounds in microstructure studies (3).

One approach to overcome the confounds of low spatial resolution is to employ elaborate modeling schemes (4–7) to resolve the microstructural features. These biophysically inspired modeling approaches seek to estimate the features of neural tissue on the sub-resolution scale or the mesoscopic scale, intermediate between the microscopic cell size and the macroscopic imaging resolution. For example, mesoscopic features pertaining to the differently oriented neurites (fiber dispersion) can be studied using the NODDI model (8) from the macroscopic dMRI voxels. Similarly, several other biophysical models have been proposed for advanced microstructure analysis such as NODDIDA (5, 9, 10), ball-and-stick model (11), CHARMED (12). The mesoscopic analysis using these elaborate models are achieved by using extensive measurements in q-space. The high q-space coverage compensates for the low k-space coverage to overcome the resolution barrier. Typically, high q-space coverage involves sampling over multiple b-shells and entail extensive scan-time.

On the other hand, the low-resolution barrier arising from the ssEPI readout can be overcome by using multi-shot readouts, a strategy where the long k-space readout is split into multiple shorter segments (13). Multi-shot EPI (msEPI) methods have been shown to provide close to 5-fold improvement in spatial resolution (using 2D slice encoding and even more using 3D slab encoding) for diffusion studies (14, 15). However, since multiple repetition times (TRs) are required to fully encode a given dMRI gradient preparation using msEPI, this approach also entail extended scan-time.

An unexploited approach to maximize the resolution capability and hence accuracy of in-vivo microstructure studies, is to couple the high k-space coverage of multi-shot methods and concurrently utilize the mesoscopic

resolution capability of biophysical modeling. However, as noted before, this approach requires extensive measurements both in k-space and q-space. Fully sampling the combined k-q space is time consuming and demands high acceleration for practical imaging. Recently, several works have shown the capability of joint k-q under-sampling schemes to afford very high acceleration factors (16–19).

While joint k-q acceleration schemes can simultaneously improve k-space and q-space coverage, the reconstruction is challenging. Since each diffusion encoding is highly under-sampled, complementary information from k- and q-domains are utilized to design the reconstruction of the individual DWIs (18, 20). The works in (16, 17, 20, 21) proposed to jointly reconstruct all DWIs simultaneously by exploiting the complimentary information in both domains. Reconstruction methods that rely on large dictionaries of q-space signals are typically employed to provide high acceleration capabilities (16, 17, 22). However, joint reconstruction methods become computationally expensive as the size of the dictionary increases. This is especially true for compressed sensing (CS) reconstruction methods that rely on sparse optimizations. Moreover, sparse optimization is challenging in the presence of highly correlated dictionary basis, which becomes particularly problematic for multi-shell diffusion signals.

In this paper, we introduce a novel learning-based framework to enable efficient reconstruction for joint k-q accelerated dMRI. We address the computational complexity of previous joint reconstruction methods by designing a compact q-space regularization that is pre-learned using machine learning. Similar to previous CS methods, the proposed q-space prior also makes use of very large dictionaries of q-space signals; however, the proposed framework derives a compact representation of the high dimensional signal manifold. The joint reconstruction incorporating the proposed q-space regularization is computationally efficient because of the pre-learning of the signal manifold, which is detached from the actual reconstruction. This approach allows us to work with very large dictionaries which are inevitable with elaborate diffusion models. Moreover, unlike previous CS methods, we are not restricted to using sparsity-conforming models anymore, thus allowing to generalize the framework to arbitrary biophysical models.

We incorporate the pre-learned q-space regularization into a traditional iterative model-based reconstruction that accounts for the imaging physics. We formulate a new algorithm to solve the joint reconstruction problem utilizing the idea of plug-and-play alternating direction method of multipliers (ADMM) (23–25). The resulting formulation using learned priors, termed qModeL, can accommodate multi-channel multi-shot multi-shell diffusion MRI acquisition and jointly recovers the entire set of DWIs from highly under-sampled acquisitions.

Theory

The multi-shot EPI acquisition model

In a *fully sampled* msDW acquisition, several TRs, equaling to the number of shots (denoted by N_s), are required to sample the k-space of a given q-space point. The image acquisition model in the presence of Gaussian noise η is given by:

$$\widehat{y}_{i,j,q}(\mathbf{k}) = \int_{\mathbb{R}^2} s_q(\mathbf{r}) c_j(\mathbf{r}) \exp\left(\mathbf{i}\, \mathbf{k}^T \mathbf{r}\right) d\mathbf{r} + \eta_{i,j}(\mathbf{k}); \quad [1]$$
$$i \in [1, N_s],\ j \in [1, N_c],\ q \in [1, N_q]\ \forall \mathbf{k} \in \Phi_{i,q}.$$

where $\hat{y}$ denotes the measured k-space data and $s_q(\mathbf{r})$ is the DWI corresponding to point $\mathbf{q}$; $\mathbf{r}, \mathbf{k}$ denotes the spatial and the k-space co-ordinates respectively and $\Phi_{i,q}$, denotes the subset of the k-space samples that is acquired at the i^{th}-shot for the q^{th} q-space point. The sampling indices of the different shots are complementary for each q-space point such that $\bigcup_{i=1}^{N_s} \Phi_{i,q} = \Theta$, where Θ is the Fourier grid corresponding to the fully sampled image. Each individual DWI, $s_q(\mathbf{r})$ can recovered from the known k-space samples of the point $\mathbf{q}$. In standard single-shot dMRI, $Ns = 1$ and the above equation reduces to the single-shot parallel imaging acquisition model with N_c coils, where $\mathbf{c}_j(\mathbf{r})$ denotes the coil sensitivity of the j^{th} coil. In the multi-shot case, however, this reconstruction is not a straight-forward parallel imaging reconstruction and requires additional phase compensation steps to remove artifacts related to phase inconsistencies inherent to multi-shot acquisitions (26).

The image acquisition model in Eq [1] can be compactly represented as

$$\hat{y}_i = \mathcal{A}_i(s) + \eta, \quad [2]$$

where $\mathcal{A}_i$ is the acquisition operator that compactly represents the operation $\mathcal{M}_i \circ \mathcal{F} \circ \mathcal{C}_i$. Here, $\mathcal{F}$, $\mathcal{M}_i$ and $\mathcal{C}_i$ denotes the Fourier transform, the k-space sampling mask for a specific shot i, and weighting by coil sensitivities, respectively. To achieve phase-compensated reconstruction, $\mathcal{C}_i$'s are modified to account for the phase corresponding to the respective k-space shots and are referred to as composite sensitivities (26).

Note that when the number of q-space points are large (e.g: $\mathrm{N}_q > 30$, multi-shell sampling), a fully sampled msDW acquisition result in long scan times. To accelerate such acquisitions, joint k-q undersampled methods were proposed to provide high acceleration rates (16, 17, 27–29).

Review of joint reconstruction for k-q accelerated dMRI

In joint k-q accelerated dMRI, only a subset of the k-space samples are measured for each q-space point.

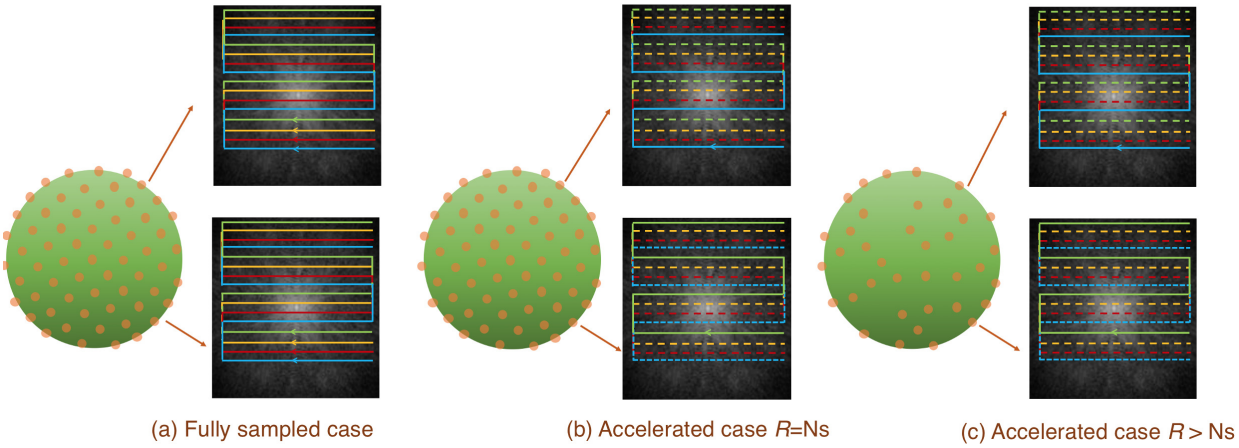


Figure 1: Illustration of random k-q undersampling using multi-shot trajectories. The brown points on the green background represents the sampling points in q-space. Each q-space point is associated with its own k-space. (a) illustrates the fully sampled case. Here, (N_s =) 4 k-space shots (denoted using 4 colors) are required to fully sample the k-space of each q-space point. In the k-q under-sampled case (b), only 1 shot (shown in bold line) out of the four shots are used to sample a given q-space point (corresponding to an acceleration factor of $R = N_s = 4$). Each q-space point is sampled with a different shot, which allows to distribute the aliasing patterns differently in each of the diffusion weighted images. The sampling pattern shown above also accommodates partial Fourier imaging. To achieve $R > N_s$, we skip the sampling of some q-space points randomly (illustrated in (c)). Note that every sampled q-space point employs 1 random shot out of the 4 k-space shots, as in (b).

The k-q acceleration can be easily implemented on MRI scanners using multi-shot trajectories with pseudo-random sampling of the k-space shot indices for each q-space point (16, 17) (see figure 1). Specifically, for an acceleration factor of R, only N_s/R k-space shots are utilized for a given q-space point, which is randomized for each q-space point. In such cases, for any given q-space point, the condition $\bigcup_{i=1}^{N} \Phi_{i,q} = \Phi_q \neq \Theta$. At high acceleration factors (e.g.: $R \approx N_s$), it becomes difficult to reconstruct each individual DWIs, $s_q(\mathbf{r})$, from the available k-space samples of point q.

On the other hand, there is high redundancy present in the combined measurements available from all q-space points, which can be exploited in a joint reconstruction setting. The joint reconstruction has the general form

$$\tilde{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{S}) - \widehat{\mathbf{Y}} \right\|_2^2 + \lambda_1 \left\| \mathcal{R}_{\mathbf{q}}\{\mathbf{S}\} \right\| + \lambda_2 \left\| \mathcal{R}_{\mathbf{r}}\{\mathbf{S}\} \right\|. \quad [3]$$

where, $\mathbf{S}(\mathbf{r})$ is of size $N_1 \times N_2 \times N_q$ and $\mathbf{Y}(\mathbf{k})$ is the k-space measurements from all N_q q-space points. $\mathcal{R}_{\mathbf{q}}\{\mathbf{S}\}$ is a generic q-space regularization that facilitates the recovery along the q-dimension. The q-space regularization can be further augmented with a spatial regularization $\mathcal{R}_{\mathbf{r}}$, thus maximally exploiting the complementary information in the q-space and spatial domains.

Sparsity-based q-space regularization methods

In the context of CS recovery, $\mathcal{R}_{\mathbf{q}}\{\mathbf{S}\}$ is an appropriate sparsifying transform and its ℓ_1-norm is penalized. When sparsity-conforming representations are available, the ℓ_1-based recovery provides the necessary regularization to recover the q-space signal in every voxel. Several sparse modeling schemes were previously explored for diffusion signals based on dictionaries derived using the Gaussian mixture model (30), the spherical ridge-lets basis (31), the spherical polar Fourier models (32) etc. However, several biophysical diffusion models often do not have a sparse representation. One approach to sparsify such models is by linearizing the non-linear models around the parameters of interest (33). Here, a large dictionary $\mathbf{Z}$ is generated, whose atoms are made from discretizing the various parameters of the model within a pre-defined range. While this approach allows ℓ_1 regularized recovery utilizing any biophysical model, the use of such large dictionaries in the joint reconstruction setting results in a reconstruction problem with high computational complexity.

In the joint k-q recovery setting, the utilization of the linearized dictionaries will result in the following reconstruction problem :

$$\tilde{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \left\| \mathcal{H}(\mathbf{F}) - \widehat{\mathbf{Y}} \right\|_2^2 + \lambda_1 \left\| \mathbf{F} \right\|_{\ell 1} + \lambda_2 \left\| \mathcal{T}\{\mathbf{F}\} \right\|_{TV}. \qquad [4]$$

Here, the goal is to reconstruct $\mathbf{F}$, the matrix of the coefficients of the linearized model, directly from the k-space measurements. For example, in the case of ball-and-stick (BAS) diffusion model (30), the pre-defined dictionary $\mathbf{Z}$ consist of atoms that corresponds to Gaussian basis function. $\mathbf{F}$ is the co-efficient of the Gaussian basis functions, the recovery of which will enable the recovery of the DWIs such that $\mathbf{ZF} = \mathbf{S}$ and $\mathcal{H}(\mathbf{F}) = \mathcal{A}(\mathbf{ZF}) = \mathcal{A}(\mathbf{S})$. Solving for $\mathbf{F}$ jointly in all the voxels using Eq [4] is highly computationally expensive. The complexity scales up with the size of the dictionary. Moreover, it is well known that the CS reconstruction becomes challenging when atoms of the dictionary are correlated. The correlation between dictionary atoms further increase in the multi-shell setting due to the signal decay with respect to the b-values.

Proposed Method

Our goal is to develop a more general q-space regularizer to achieve joint reconstruction of the k-q undersampled data. Below, we propose such a regularizer which can be used in conjunction with any diffusion model. The computational complexity of the joint reconstruction involving the proposed regularizer is not compromised by the size of the dictionary, and is unaffected by the correlated atoms of the dictionary, thus overcoming the disadvantages of the sparse priors.

A Learning based Biophysically Driven Q-space Prior

Below, we design a novel compact q-space regularizer $\mathcal{R}_{\mathbf{q}}\{\mathbf{S}\}$, fully exploiting the advantages of linearized dictionaries. However, instead of employing the huge dictionary directly in the reconstruction process, we propose a learning-based approach, where we learn the signal manifold corresponding to the dictionary. The learning provides a compact representation of the manifold, which can be used as a regularizer in a joint reconstruction setting. This framework, which detach the learning step from the actual reconstruction, allow to utilize dictionaries of any size and complicated biophysical models during the learning phase. During reconstruction, we utilize the learned manifold as a regularization.

We demonstrate the idea of deriving the learning-based priors by making use of a generalized diffusion model, which can enable detailed biophysical modeling. This multi-compartmental model, known in the diffusion modeling literature as the standard model (SM), incorporates multiple mesoscopic parameters of the neural tissue :

$$K(b,\zeta) = f_1 e^{-bD_a\zeta^2} + f_2 e^{-bD_e^{\perp} - b\left(D_e^{||} - D_e^{\perp}\right)\zeta^2} + f_{\text{iso}} e^{-bD_{\text{iso}}} \quad [5]$$

including compartmental volume fractions (f_i's) and compartmental diffusivities (D's) pertaining to the intra-neurite, extra-neurite, and isotropic compartment. The measured diffusion signal is modeled as a spherical convolution of the multi-compartmental diffusion signal K with the fiber orientation distribution function (ODF) $\mathcal{P}$ as:

$$S(b,\mathbf{g}) = S_0 \int_{\widehat{\mathbf{n}}} \mathcal{P}(\widehat{\mathbf{n}}) \circledast K(b, \widehat{\mathbf{g}}\cdot\mathbf{n})\, d\widehat{\mathbf{n}} \quad [6]$$

where b and $\mathbf{g}$ are respectively the diffusion gradient strength and direction vector, $S(b,\mathbf{g})$ and S_0 are the diffusion weighted and the reference non-diffusion weighted image. $\circledast$ denotes a spherical convolution operation. Note that this model generalizes several other diffusion models (5, 6) and hence, can provide an appropriate manifold that can approximate the in-vivo diffusion signal, (which are acquired using traditional single diffusion encoded PSGE acquisitions), to acceptable levels.

In the learning step, we construct a dictionary by generating q-space signals using the SM. This is achieved by discretizing all the free parameters in Eq [5-6] within their biophysical ranges. Note that the signal $S(b,\mathbf{g})$ is a non-linear function of the parameters and lie on a manifold in a high dimensional space. We propose to pre-learn this manifold using the dictionary of q-space signals generated from the above 3-compartment model. Note that even though we use linearized dictionaries, our approach is aimed at learning the non-linear manifold directly, whereas the CS methods aim to represent the discretized signals and not the non-linear

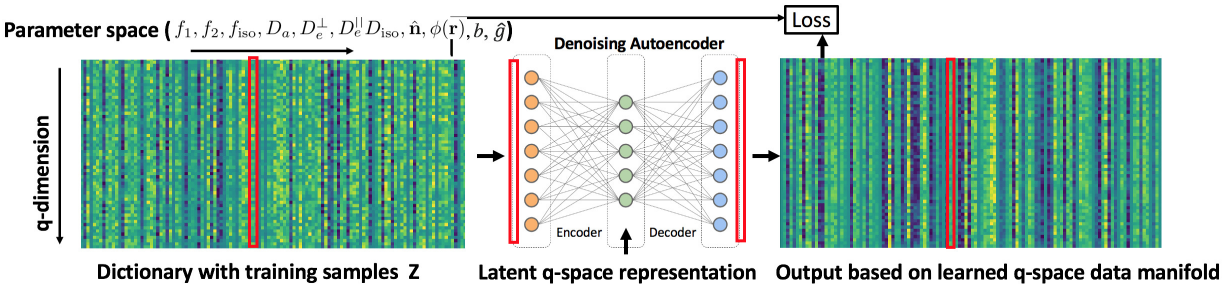


Figure 2: Illustration of the learning of the q-space manifold spanned by the 3-compartment model using a denoising autoencoder. The dictionary consists of data generated by discretizing the free parameters of the model. The denoising autoencoder is trained using the dictionary, which learns the manifold corresponding to the parameter space. The output of the autoencoder preserves the latent features of the input data.

manifolds.

Pre-learning of q-space manifold using a denoising autoencoder neural network

To pre-learn the signal manifold, we use machine learning and employ a denoising auto-encoder (DAE) neural network. DAEs have been previously employed as unsupervised schemes to learn a data manifold (34, 35). These networks can efficiently learn a compact latent space from noisy training data and are extensively used for denoising (36). Here, we propose to train the DAE using the linearized dictionary to learn the latent representation of the q-space signal manifold and employ the trained DAE in under-sampled data recovery. The dictionary generated from the diffusion model, denoted by $\mathbf{Z}$, is used as inputs to train the DAE. An illustration of this pre-learning is provided in figure 2. Once the DAE is trained using the dictionary, it learns the latent space corresponding to the training samples.

The pre-learned compact representation of the manifold can be used as a q-space prior. The diffusion signal, $\mathbf{S}(\mathbf{r})$, at any given voxel $\mathbf{r}$ is constrained to lie on the q-space manifold. Denoting the encoder-decoder mapping by $\mathcal{D}_\Theta(\mathbf{S})$, this projection of the noisy input samples to the q-space manifold, provides a voxel-wise constraint along the q-dimension to guide the under-sampled data recovery. Thus, minimizing the projection error $\left\|\mathcal{D}_\Theta(\mathbf{S}) - \mathbf{S}\right\|$ provides the necessary regularization to recover the under-sampled data along the q-dimension.

Joint reconstruction using q-space prior

We propose to replace the q-space regularizer in Eq [3], with the new learned q-space regularization as

$$\mathbf{S} = \underset{\tilde{\mathbf{S}}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{S}) - \widehat{\mathbf{Y}} \right\|_2^2 + \lambda \left\| \mathcal{P}_\Theta(\mathbf{S}) \right\|_2^2 \qquad [7]$$

$$\text{where } \mathcal{P}_\Theta(\mathbf{S}) = \mathcal{D}_\Theta(\mathbf{S}) - \mathbf{S}.$$

The above reconstruction can be solved in an iterative fashion by alternating between data consistency enforcement and q-space signal projection using the neural networks as proposed in (29). Here, we derive a traditional iterative model-based reconstruction incorporating the projection into the reconstruction process. Conveniently, the DAE encoder-decoder mapping can be incorporated into a traditional model-based reconstruction setting with the knowledge of the weights learned from each layer of the DAE after training (37). With the known weights, we derive below an optimization strategy, extending the idea of plug-and-play ADMM. Since our problem deals with the recovery of highly under-sampled data, we include a TV regularization in the spatial domain in addition to the regularization in the q-dimension. The resulting qModeL reconstruction is given by

$$\mathbf{S}^{*}= \underset{\tilde{\mathbf{S}}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{S}) - \widehat{\mathbf{Y}} \right\|_2^2 + \lambda_1 \left\| \mathcal{P}_{\Theta}(\mathbf{S}) \right\|_2^2 + \lambda_2 \left\| \mathbf{S} \right\|_{TV} . \qquad [8]$$

In the next section, we develop a plug-and-play algorithm to solve the qModeL reconstruction.

Plug-and-play joint reconstruction algorithm

To solve Eq 8 using ADMM, we start by defining two auxiliary variables $\mathbf{V}$ and $\mathbf{N}$. We re-write the cost function $C1$ in Eq 8 using these auxiliary variables as:

$$C1 : \left\| \mathcal{A}(\mathbf{S}) - \widehat{\mathbf{Y}} \right\|_2^2 + \lambda_1 \left\| \mathcal{P}_{\Theta}(\mathbf{N}) \right\|_2^2 + \lambda_2 \left\| \mathbf{V} \right\|_{TV} , \text{s. t. } \mathbf{N} = \mathbf{S} \text{ and } \mathbf{V} = \mathbf{S}.$$

The augmented Lagrangian corresponding to the above constrained optimization problem is given by:

$$\begin{aligned} \mathcal{L}(\mathbf{S}, \mathbf{V}, \mathbf{N}) = \left\| \mathcal{A}(\mathbf{S}) - \widehat{\mathbf{Y}} \right\|_2^2 + \lambda_1 \left\| \mathcal{P}_{\Theta}(\mathbf{N}) \right\|_2^2 + \lambda_2 \left\| \mathbf{V} \right\|_{TV} + \frac{\beta_1}{2} ||\mathbf{S} - \mathbf{N}||^2 + \gamma_1^T (\mathbf{S} - \mathbf{N}) \\ + \frac{\beta_2}{2} ||\mathbf{S} - \mathbf{V}||^2 + \gamma_2^T (\mathbf{S} - \mathbf{V}) \end{aligned} \qquad [9]$$

where γ_1 and γ_2 are the Lagrange multipliers. To solve Eq [9], each variable is sequentially updated by solving a set of sub-problems, during which the other variables are kept fixed:

$$\mathbf{S}^{(k+1)}= \underset{\tilde{\mathbf{S}}}{\operatorname{argmin}} \left\| \mathcal{A}(\mathbf{S}) - \widehat{\mathbf{Y}} \right\|_2^2 + \frac{\beta_1}{2} \left\| \mathbf{S} - \underbrace{\mathbf{N} + \frac{\gamma_1}{\beta_1}}_{\bar{s_1}} \right\|^2 + \frac{\beta_2}{2} \left\| \mathbf{S} - \underbrace{\mathbf{V} + \frac{\gamma_2}{\beta_2}}_{\bar{s_2}} \right\|^2 \qquad [10.1]$$

$$\mathbf{N}^{(k+1)}= \underset{\tilde{\mathbf{N}}}{\operatorname{argmin}} \lambda_1 \left\| \mathcal{P}_{\Theta}(\mathbf{N}) \right\|_2^2 + \frac{\beta_1}{2} \left\| \mathbf{N} \underbrace{- \mathbf{S}^{(k+1)} - \frac{\gamma_1}{\beta_1}}_{\bar{N}^{(k)}} \right\|^2 \qquad [10.2]$$

$$\mathbf{V}^{(k+1)} = \underset{\tilde{\mathbf{V}}}{\operatorname{argmin}}\, \lambda_2 \|\mathbf{V}\|_{TV} + \frac{\beta_2}{2} \left|\left| \mathbf{V} - \underbrace{\mathbf{S}^{(k+1)} - \frac{\gamma_2}{\beta_2}}_{\bar{V}^{(k)}} \right|\right|^2 \quad [10.3]$$

$$\tilde{\gamma}_{\mathbf{1}}^{(k+1)} = \tilde{\gamma}_{\mathbf{1}}^{(k)} + (\mathbf{S}^{(k+1)} - \mathbf{N}^{(k+1)}) \text{ where } \tilde{\gamma}_{\mathbf{1}} = \frac{\gamma_{\mathbf{1}}}{\beta_1} \quad [10.4]$$

$$\tilde{\gamma}_{\mathbf{2}}^{(k+1)} = \tilde{\gamma}_{\mathbf{2}}^{(k)} + (\mathbf{S}^{(k+1)} - \mathbf{V}^{(k+1)}) \text{ where } \tilde{\gamma}_{\mathbf{2}} = \frac{\gamma_{\mathbf{2}}}{\beta_2} \quad [10.5]$$

As noted in previous plug-and-play ADMM approaches(23, 24), the steps 10.2, 10.3 can be interpreted as a "denoising" of the corresponding variable, i.e.,

$$\mathbf{N}^{(k+1)} = \mathcal{D}1_{\sigma_1}(\bar{\mathbf{N}}^k) \quad [11]$$

$$\mathbf{V}^{(k+1)} = \mathcal{D}2_{\sigma_2}(\bar{\mathbf{V}}^{(k)}), \quad [12]$$

where, $\mathcal{D}1$ and $\mathcal{D}2$ represent the denoising operations and the parameter σ_i controls the strength of the denoiser depending upon the noise level. Here, $\sigma_i \overset{\text{def}}{=} \sqrt{\lambda_i/\beta_i}$. The above interpretation enables us to replace steps 10.2 and 10.3 with the q-space denoising by the pre-trained DAE and TV denoising, respectively. The subproblem in 10.1 can be solved using conjugate gradients (CG), with phase compensation.

A schematic of the proposed phase-compensated reconstruction is given in figure 3. Note that the q-space denoising is ideally applied on real-valued data.The phase-compensated reconstruction provide images that are close to real-valued, on which the denoising can be applied. The q-space denoiser is trained to account for any residual phase that is present in the DWIs as explained in the next section. Figure 3(b) illustrates the idea of the projection step, which aims to minimize the projection error, while being consistent to the measured data.

The above algorithm provides a fast and simple implementation for the joint recovery. Depending upon the application, different denoisers can be interchangeably plugged in, adding flexibility to the framework. In the next section, we test the proposed joint recovery method on under-sampled in-vivo data. We also demonstrate the proposed learning-based method on two different diffusion models, which corresponds to two different manifolds/denoisers.

Methods

Training details

We trained two DAEs for demonstration of the plug-and-play approach. The first DAE was trained using the SM described in Eq 5-6 and the second DAE was training using the ball-and-stick (BAS) model (17, 30, 38)

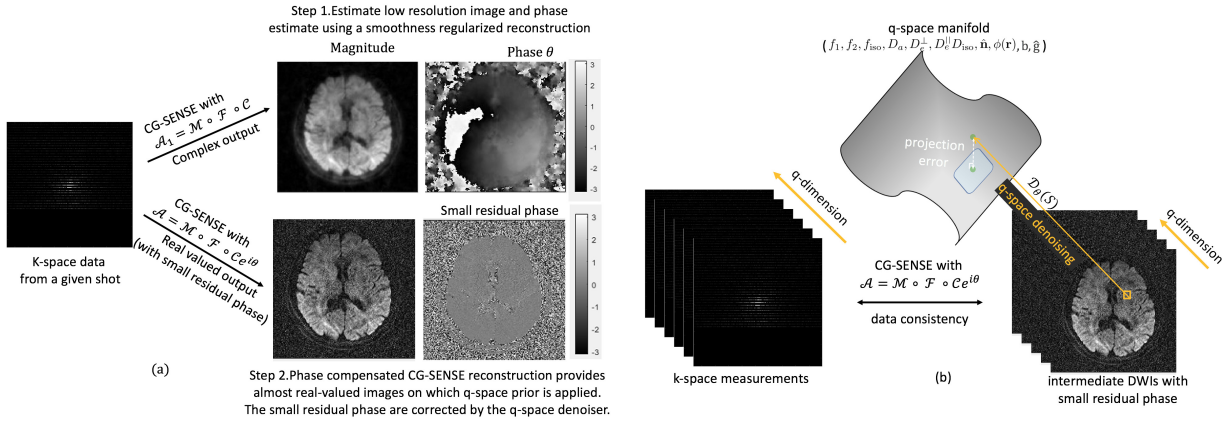


Figure 3: The different components in the proposed reconstruction. (a) shows the two steps involved in the phase compensated recovery. In the first step, the background phase, θ is estimated using a CG-SENSE reconstruction with smoothness regularization. In the second step, the estimated phase is incorporated into the forward operator $\mathcal{A}$ such that real-valued images are recovered. When phase compensation is not perfect, small residual phase errors can result in the images. The q-space denoiser is trained to correct for small phase errors. (b) shows the qModeL reconstruction which aims to minimize the projection error to a pre-learned q-space manifold. The pre-learned q-space manifold is a function of a large set of variables including the diffusion directions and the b-values used for the biophysical modeling. In the absence of noise, the q-space signal corresponding to every voxel will map to points on the manifold. In the noisy case, the reconstruction tries to minimize the projection error to this manifold subject to data-consistency to the measured data.

given by :

$$S(b, \mathbf{g}_k) = S_0 \sum_i f_i \; exp^{(-b\mathbf{g}_k^T \mathbf{D}_i \mathbf{g}_k)}, \qquad [13]$$

Each model has its own unique features. While the SM accounts for three compartments per voxel, BAS model do not account for the extra-neurite compartment, however multiple intra-neurite compartments are modeled. For training, we assumed two fiber populations per voxel for BAS model and one for SM. The free model parameters in each model were discretized within their biophysical ranges: f_i's $\in [0, 1]$, and D's $\in [0.1, 3]$(39). 60 unique fiber directions were considered for both models. The two fiber directions per voxel for the BAS model were chosen from the above 60 directions.

Because the reconstruction involves the recovery of DWIs from noisy complex k-space data, the phase-compensated reconstruction can result in DWIs with small residual phases. To account for any residual phase modulations in the q-space signals, we treat phase also as a training parameter. The simulated diffusion signals were modulated with small random phase and further corrupted by adding complex white Gaussian noise at various levels. Both the real and corrupted data (with and without noise) were used for training. The training was performed using about 75,000 instances of diffusion signals for both models. Standard DAE training procedures were utilized using TensorFlow with stochastic gradient descent (SGD). The atoms were randomized during each epoch to help achieve faster convergence. 500 epochs were used with the mean-

squared error (MSE) loss function, and default Adam optimizer parameters. The training took about 90 minutes on a NVIDIA Quadro K5000 GPU card. After training, the weights were saved and were used to implement the DAE in Matlab. All reconstructions were implemented in Matlab.

The DAE was re-trained every time a different acquisition is used (b-value, sampling points etc.). The input layer to the DAE consist of as many layers as the sampled q-space points. The complex training data was separated into real and imaginary parts for training. A 3-layer fully connected network was used for the encoder and for the decoder with 15 neurons in the bottleneck layer each for the real and the imaginary part.

Datasets

We employed a 4-shot EPI acquisition with Stejskal-Tanner diffusion encoding to collect the dMRI data. All experiments were performed on healthy volunteers following the Institutional Review Board requirements at the University of Iowa and obtaining informed written consent. The experiments were performed on 3T field strength using scanners equipped with different gradients. Scanner 1 was equipped with gradients with G_{max} of 80 mT/m and SR_{max} of 200 T/m/s and a 44-channel head coil. Scanner 2 was equipped with a 32-channel head coil with gradients with G_{max} of 33 mT/m and SR_{max} of 120 T/m/s. All datasets were acquired in a fully sampled manner keeping all 4 shots for each diffusion direction. A fully sampled whole brain acquisition is highly time consuming. To avoid subject motion, the scan-time was limited to under 10 minutes, by imaging only a few slices.

*Dataset 1 (scanner 1):*1.1 mm in-plane resolution (FOV: 210×210 mm, sampling matrix: 192×136, slice thickness: 5 mm) with a b-value of 1200 s/mm^2. The dataset consist of 60 diffusion directions with TE = 55 ms and partial Fourier (pF) sampling at 71%. The TR was 2500 ms and the total acquisition time was 10 mins for 12 slices.

Dataset 2 (scanner 2): 0.82 mm in-plane resolution (FOV: 210×210 mm, sampling matrix: 256×152; slice thickness: 5 mm) at b-value of 700 s/mm^2. 60 diffusion encodings were collected with TE of 84 ms and pF at 59%. The TR was 2500 ms and the total acquisition time was 10 mins for 10 slices.

Dataset 3 (scanner1): multi-shell data at 1.7 mm isotropic resolution with a b-value of 1000, 2000, 3000 s/mm^2 (FOV: 210×210 mm, sampling matrix: 128×96). Each b-shell was sampled using 60 diffusion directions with pF at 71% giving a TE of 58.3 ms. With the above TE, the sampling pattern is fully exploited to give full k-space sampling for b-value of 1000, but only 71% for b-value of 2000 and 3000. The TR was 2500 ms and the total acquisition time was 10 mins for each shell for 12 slices.

Retrospective joint k-q under-sampling

The data were retrospectively under-sampled using pseudo-random under-sampling. Depending upon the

extent of acceleration considered, two different under-sampling schemes were employed. In the first case, all datasets were under-sampled by a factor of $R = N_s\ (= 4)$. This was achieved by using only 1 random shot for each q-space point. This sampling essentially reduces the 4-shot acquisition to one TR per q-space point, matching the time-efficiency of a single-shot acquisition. To achieve acceleration factors higher than the number of shots ($R > N_s$), we propose to skip the sampling of some q-space points altogether, in a randomized fashion. i.e., we only sample $N_q * N_s/R$ random q-space points. This approach further reduces the TRs to improve the time-efficiency. For example, in the multi-shell data, the fully sampled acquisition involves 180 q-space samples. At $R = 6$, we will only sample 120 diffusion directions. The total TRs were thus reduced to 120 compared to 720 (180 x 4) TRs in the fully sampled case. Using this strategy, we retrospectively under-sampled all datasets to achieve $R = 6, 8$.

Reconstruction

The k-q under-sampled data were reconstructed using qModeL. We also performed comparison of the qModeL reconstruction to previous CS-based joint recovery, specifically, the L1-TV method given in Eq 4 and the SNR-enhancing joint reconstruction (SER) (40, 41). All methods were re-implemented using the plug-and-play approach. We used an off-the-shelf TV denoiser (42) for all the experiments.

The data from the fully-sampled case was reconstructed using CG-SENSE, which may be treated as the "ground-truth" case to study the accelerated cases. To report the error in the under-sampled reconstructions, we compute the root means square error (RMSE) and the structure similarity index (SSIM) with respect to the CG-SENSE reconstruction. However, in the absence of additional regularizers, the CG-SENSE reconstruction provides noisy results compared to the qModeL reconstruction at $R = 1$, because of the additional denoising terms (note that CG-SENSE is simply the first term of the qModeL reconstruction). Hence, to avoid mischaracterization of the denoising achieved by qModeL as reconstruction error, we also report the RMSE and the SSIM with respect to qModeL at $R = 1$.

The data-consistency term for all of the reconstructions was implemented as a phase-compensated reconstruction (26), as illustrated in figure 3(a). The phase for the reconstruction were estimated from individual shots to derive a low-resolution phase map (43). The estimated phase were then combined with coil sensitivity maps to derive the composite sensitivity maps (26).

Results

The performance of the q-space regularizer

We first study whether the biophysically-derived q-space manifold prior is good enough to represent in-vivo diffusion data. In figure 4, we plot the measured q-space signal, $\mathbf{S}$, from the boxed region (fig 4a) from

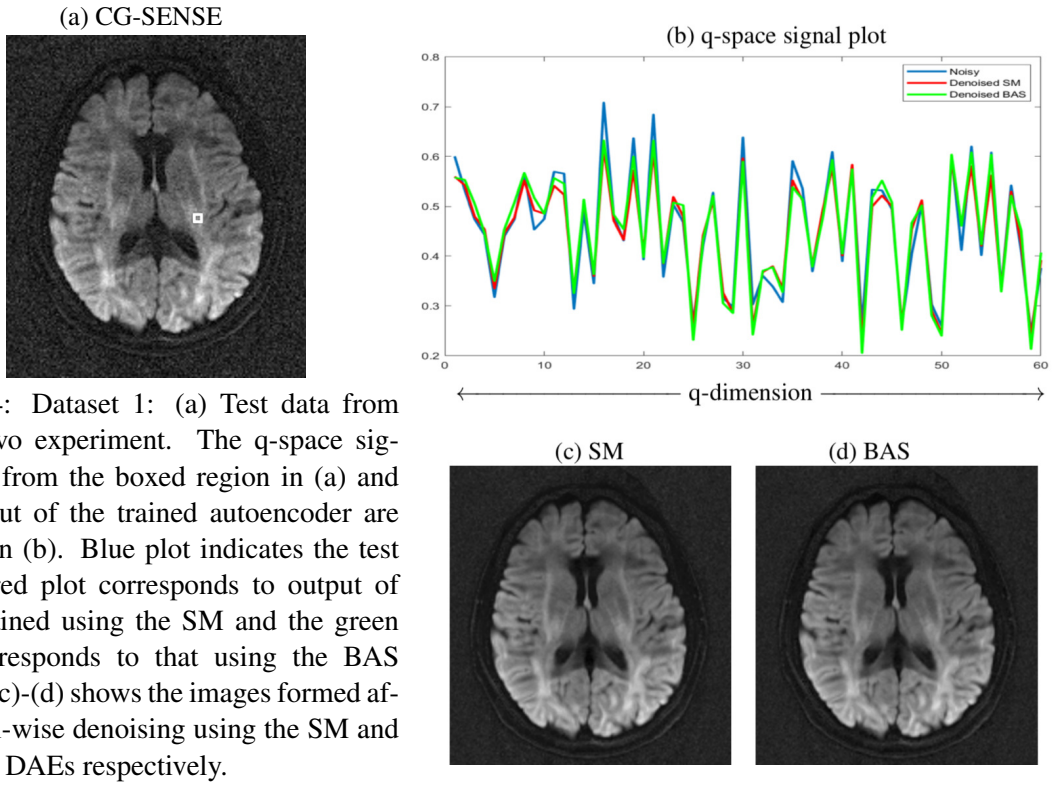


Figure 4: Dataset 1: (a) Test data from an in-vivo experiment. The q-space signal plot from the boxed region in (a) and the output of the trained autoencoder are plotted in (b). Blue plot indicates the test signal, red plot corresponds to output of DAE trained using the SM and the green plot corresponds to that using the BAS model. (c)-(d) shows the images formed after voxel-wise denoising using the SM and the BAS DAEs respectively.

the in-vivo dataset 1, along with the denoised signal, $\mathcal{D}_\Theta(\mathbf{S})$, from the trained DAE. The red and green plots (fig 4b) correspond to the denoising obtained corresponding to the SM-derived manifold and the BAS-derived manifold respectively. The denoised signals from both the DAEs predicts the in-vivo data reasonably well and motivates its utilization in the recovery of missing q-space samples from under-sampled in-vivo experiments. Figure 4a shows the DWI before denoising and fig 4c-d shows the DWI after voxel-wise denoising using the two DAEs.

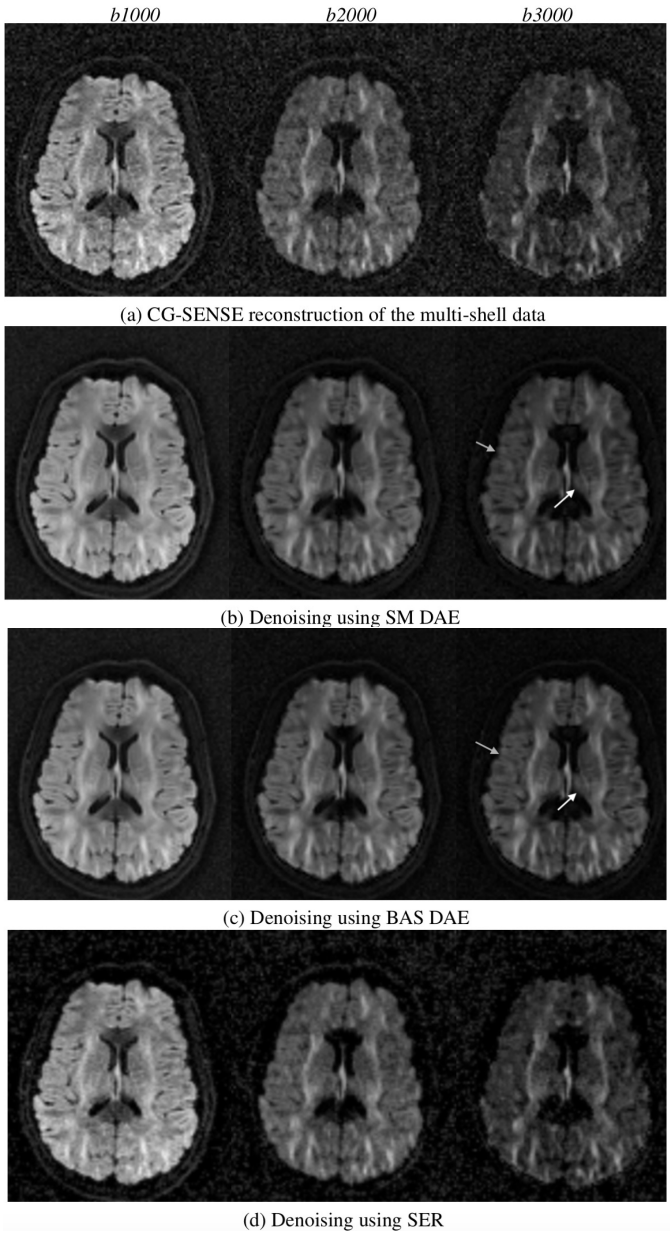
Figure 5 shows the testing of the DAEs on the multi-shell in-vivo data. Here, both the SM and the BAS DAEs were re-trained for the multi-shell acquisition setting. Figure 5(a) shows the CG-SENSE reconstruction of three DWIs from their fully sampled measurements. A representative DWI from each shell is shown. Notably, the DWIs from the higher shells are more noisy. As observed from fig 5(b)-(c), the noisy diffusion data at the higher b-shells are noticeably denoised by the voxel-wise application of the DAEs. The reconstruction essentially removes the partial Fourier blurring present in the high b-shells by translating the high SNR available from the low b-shell samples. The deblurring restores the anatomical features of the DWIs at high b-values, even in the absence of any spatial regularization. The denoising achieved by the pre-learned denoisers is compared to SER denoising also in fig 5(d), which is another joint denoising method exploiting the redundancy of structure in the DWIs. The superior capability

of the learning-based manifold prior to denoise the multi-shell data is evident from this experiment.

Figures 4-5 confirm the utility of the DAEs that were learned from biophysical modeling, to predict the q-space signals for in-vivo experiments. Next, we utilize the DAEs as a q-space prior in an iterative reconstruction for under-sampled data recovery.

Joint recovery for the case $R = \mathbf{N_s}$

Figure 6 shows the result from dataset 1 on two representative DWIs for the case of $R = 4$. In comparison to the individual DWI reconstruction using CG-SENSE (fig 6c), the qModeL joint reconstruction (fig 6d) exploiting the q-space and the TV denoisers are shown to provide significantly better reconstruction quality. The reconstruction from $R = 1$ using both the methods are also included for comparison. Figure 6(e)-(f) show the error in the qModeL reconstruction at $R = 4$ with respect to the fully sampled reconstructions. Table 1 lists the reconstruction errors from all the datasets at $R = 4$. In all cases, the qModeL joint reconstruction is shown to recover the DWIs with acceptable accuracy, with RMSE less than 5% with respect to the CG-SENSE reference. (The RMSE is approximately 1% compared to the qModeL fully sampled reconstruction).



(a) CG-SENSE reconstruction of the multi-shell data

(b) Denoising using SM DAE

(c) Denoising using BAS DAE

(d) Denoising using SER

Figure 5: Multi-shell data: (a) Three representative DWIs from b1000, b2000 and b3000 shells are shown. The images correspond to the CG-SENSE reconstruction from their fully sampled k-space measurements. (b)-(c) shows the DWIs after voxel-wise denoising using the SM and the BAS denoisers respectively. (d) shows the denoising using the SER method.

Comparison with other joint recovery methods

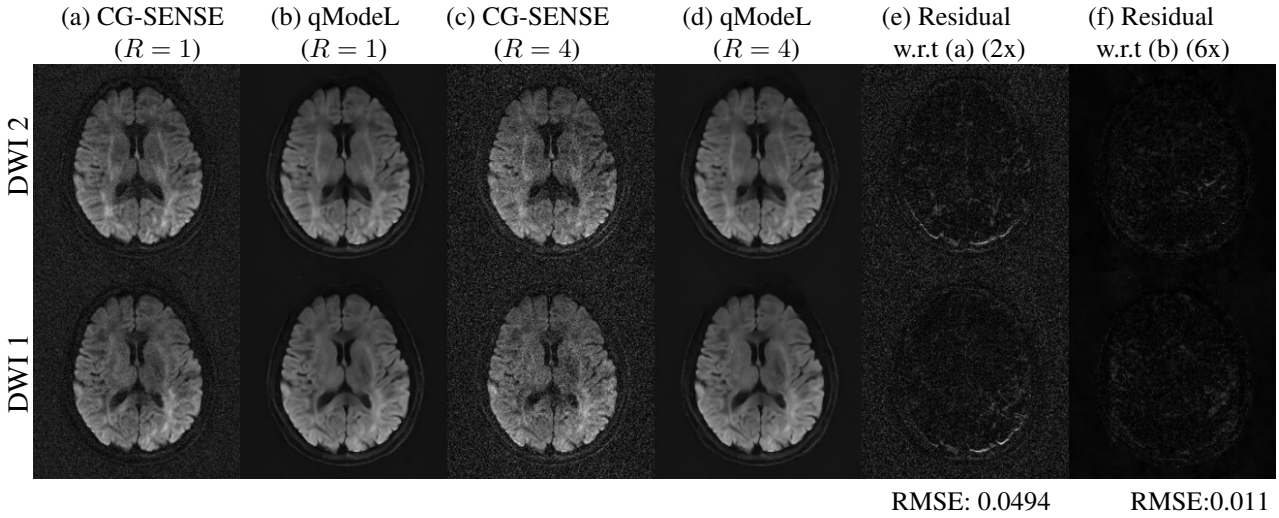


Figure 6: Under-sampled reconstruction at $R = 4$ from dataset 1. (a)-(b) corresponds to the CG-SENSE and the qModeL reconstruction of the fully sampled data (with partial Fourier), (c)-(d) shows the CG-SENSE and the qModeL reconstruction at $R = 4$. (e)-(f) shows the error in the qModeL reconstruction at $R = 4$, with respect to the CG-SENSE and the qModeL reconstruction at $R = 1$, magnified by 2 times and 6 times respectively. All images are shown on the same display scale [0, 1]. The RMSE for the full set of DWIs reconstructed at $R = 4$ is reported at the bottom with respect to the CG-SENSE and the qModeL references.

The qModeL reconstruction is compared to the L1-TV joint reconstruction given in Eq [4], which uses the dictionary based on Gaussian basis functions and the BAS model. Table 1 reports the RMSE and SSIM using L1-TV reconstruction at $R = 4$ on datasets 1, 2. Notably, the L1-TV reconstruction fails to reconstruct the multi-shell data accurately both at $R = 1$ & 4. Specifically, the method fails to retain the unique contrasts present in the different shells and converges to the same contrast for all the shells (figure S1 in Supporting Information). However, it successfully preserves the diffusion weighting corresponding to the gradient directions. This is due to the highly correlated atoms in the dictionary which presents a challenge for ℓ_1 reconstruction to find a sparse combination that can represent a multi-shell diffusion signal using the same

Table 1: Comparison of qModeL and L1-TV reconstruction at $R = 4$.

	Dataset 1 (192 x 192 x 60)		Dataset 2 (256 x 256 x 60)		Dataset 3 (128 x 128 x 180)	
	qModeL	L1-TV	qModeL	L1-TV	qModeL	L1-TV
(CG-SENSE at $R = 1$ as reference)						
RMSE :	.0494	.0487	.03	.0303	.034	-
SSIM :	.92	.924	.91	.92	.93	-
(The respective algorithms' $R = 1$ reconstruction as reference)						
RMSE :	.011	.0146	.0075	.0118	.009	-
SSIM :	.992	.990	.987	.9792	.995	-
Time (sec) :	480	1660	789	3463	376	-

basis. Nevertheless, the L1-TV method can jointly reconstruct the DWIs from the individual shells if the data from each shell is separately reconstructed.

Table 1 also reports the time to reconstruct the DWIs per slice for both methods. The L1-TV method was re-implemented using the plug-and-play ADMM. A total of 60 iterations were used in both the methods. For the L1-TV method, the dictionary size was limited to 60 atoms based on which sparse recovery was performed. In comparison, the qModeL dictionary had 900 unique atoms and their various permutations, from which the manifold was learned. Notably, the utilization of such big dictionary will result in significant computational bottleneck for the L1-TV reconstruction. The qModeL reconstruction is faster than the L1-TV reconstruction and at the same time provides better accuracy as noted from the RMSE and SSIM values. Figure S2-S3 in the supporting information shows the fractional anisotropy (FA) maps and direction of primary eigen vector (DEC maps) after a tensor fitting on the qModeL and L1-TV reconstructed data, using dataset 2. The SER prior, which was originally designed for denoising applications, was not able to provide good reconstructions for the highly under-sampled cases. Hence, we have not included the results. This behavior is expected since the SER makes use of the structural similarity of the DWIs which may not provide a strong enough regularization to remove the severe aliasing present in the highly under-sampled reconstructions.

K-q under-sampling with missing volumes ($R > \mathbf{N_s}$)

Table 2 reports the results of the k-q under sampling experiments from all the datasets for R=6,8. In datasets 1-2, this corresponds to 15 ($R = 6$) and 30 ($R = 8$) missing volumes, and 60 ($R = 6$) and 90 ($R = 8$) missing volumes for dataset 3. In all the experiments, qModeL is able to recover all volumes with reasonable accuracy. The RMSE of the reconstructed DWIs are noted to be within 5% of CG-SENSE reconstruction (and $< 2\%$ compared to qModeL reference). Figure S4 in the supporting information and figure 7 shows the reconstruction results from dataset 2 and 3 from 8-fold under-sampling. In figure 7, representative DWIs from the three shells are shown. The CG-SENSE and the qModeL reconstruction from the fully sampled

Table 2: qModeL reconstruction errors at $R = 6, 8$

	Dataset 1 (192 x 192 x 60)		Dataset 2 (256 x 256 x 60)		Dataset 3 (128 x 128 x 180)	
	$R = 6$	$R = 8$	$R = 6$	$R = 8$	$R = 6$	$R = 8$
CG-SENSE at $R = 1$ as reference						
RMSE :	.0503	.0512	.032	.033	.038	.04
SSIM :	.915	.913	.91	.89	.93	.92
qModeL reconstruction at $R = 1$ as reference						
RMSE :	0.0145	0.0181	.0098	.0135	.009	.01
SSIM :	0.989	0.984	.981	.969	.994	.991

case are shown, along with the qModeL reconstruction from the accelerated case and its residual compared to the fully sampled case. The figure also shows the q-space signal plot that illustrates the recovery of the

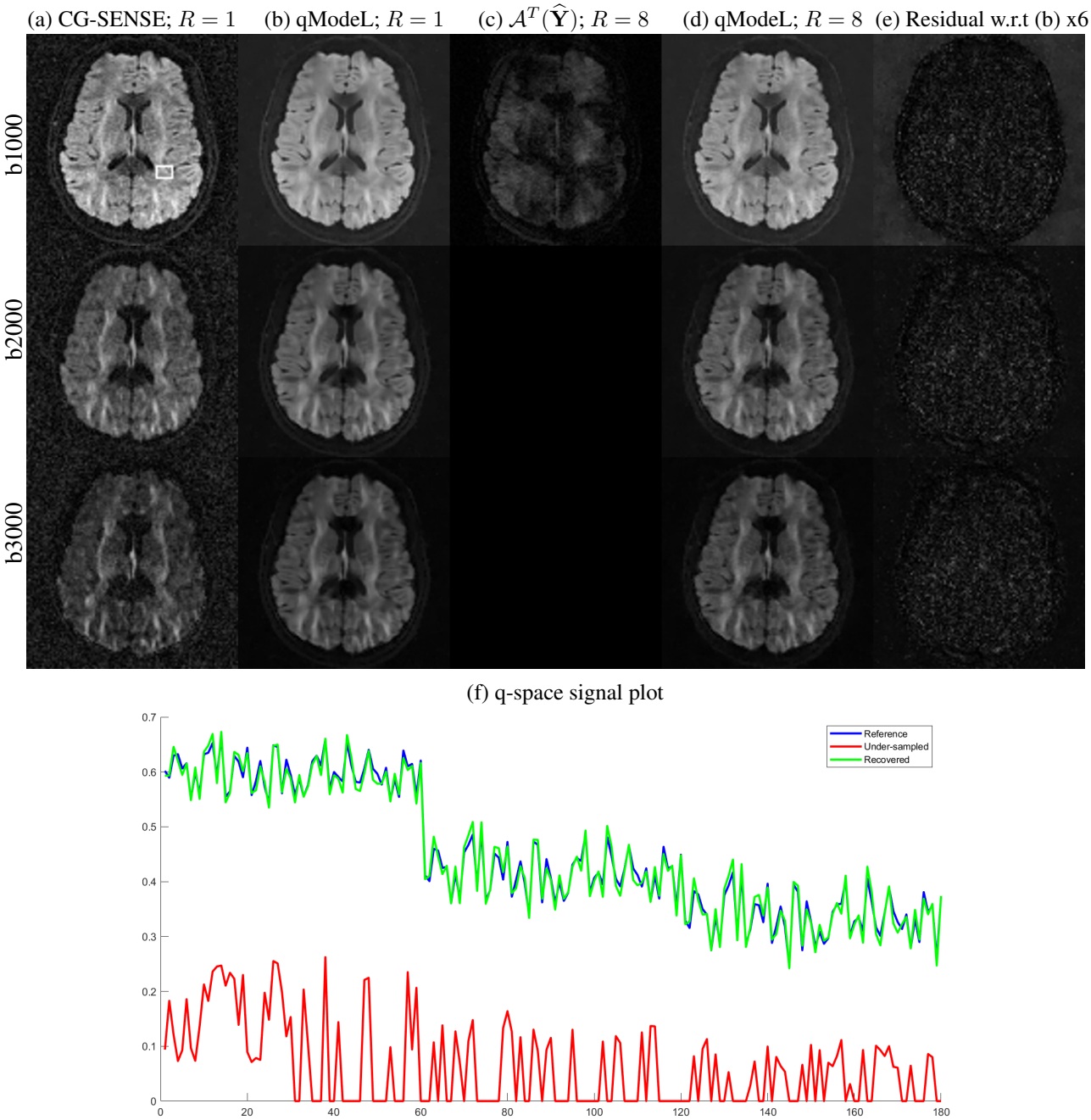


Figure 7: Multi-shell experiment at $R = 8$. (a)-(b) shows the fully sampled reconstruction using the CG-SENSE and the qModeL joint reconstruction. (c)-(d) shows the initial guess ($\mathcal{A}^T(\widehat{\mathbf{Y}})$) and the qModeL reconstruction at $R = 8$. Column (c) shows empty volumes (row 2,3) corresponding to q-space points that were not sampled. Column (d) shows the qModeL reconstruction, which recovers all the DWIs, including the missing volumes. Column (e) shows the residual error with respect to the qModeL reference, magnified 6 times. The signal plot from the boxed region is shown in (f). The red curve plots the q-space signal from $\mathcal{A}^T(\widehat{\mathbf{Y}})$, which shows missing volumes. The blue and the green curve plots the signal from the qModeL reconstruction at $R = 1$ and 8 respectively. The green plot is shown to recover the missing points and is noted to match the reference signal from the fully sampled case.

missing q-space volumes.

Figure 8 shows the result of microstructural analysis using a compartmental model that was performed on the multi-shell data. Specifically, we used the NODDI model and computed the intra-cellular volume fraction (v_{ic}) and the orientation dispersion (ODI) (44). Both the fully sampled reference data and the DWIs reconstructed using qModeL from the 4-, 6- and 8-fold accelerated data were used for estimating the NODDI parameters. The NODDI parameter maps computed at each acceleration as well as the error in the parameter estimates are also provided in comparison to the reference dataset. We note that the average error in the v_{ic} is less than .04 in the whole brain and the average error in the ODI is less than .07.

Contribution of TV denoiser and the q-space denoiser

We also analyzed the contribution of the two denoisers used in the qModeL reconstruction. Table S1 in the supporting information lists the reconstruction errors at $R = 4, 6$ & 8 when the TV denoiser not utilized in the reconstruction. The reconstruction errors are noted to be slightly higher when TV denoising is not employed, confirming that both denoisers contribute to the accuracy of the reconstruction. Figure S5 in the supporting information show example images from the joint reconstruction with and without the TV denoiser. The reasonable reconstruction results even in the absence of the TV denoiser confirm that the joint reconstruction is primarily driven by the q-space denoiser. However, the TV denoiser does improve the results as confirmed by the lower RMSEs with the TV denoiser.

Discussion

In this work, we developed a novel q-space prior exploiting a learning-based framework and demonstrated its utility in highly accelerated dMRI. Specifically, having a strong q-space regularization enables to undersample the combined k-q space of dMRI such that very high acceleration factors can be achieved. A joint recovery scheme exploiting the q-space prior allows to apply a voxel-wise constraint along the q-dimension to recover the q-space signals. This constraint can be applied along with other spatial priors to further improve the reconstruction quality. Our experiments show the utility of the above framework for joint recovery of diffusion data from highly accelerated acquisitions and also for multi-shell sampling. Specifically, we were able to achieve 8-fold acceleration using the proposed method without deteriorating the reconstruction quality.

The proposed learning-based q-space regularizer provides a highly adaptable framework for advanced microstructure studies. It overcomes several limitations associated with previously employed q-space priors that relied on sparse reconstructions. The new prior allows the use of any diffusion models that may not

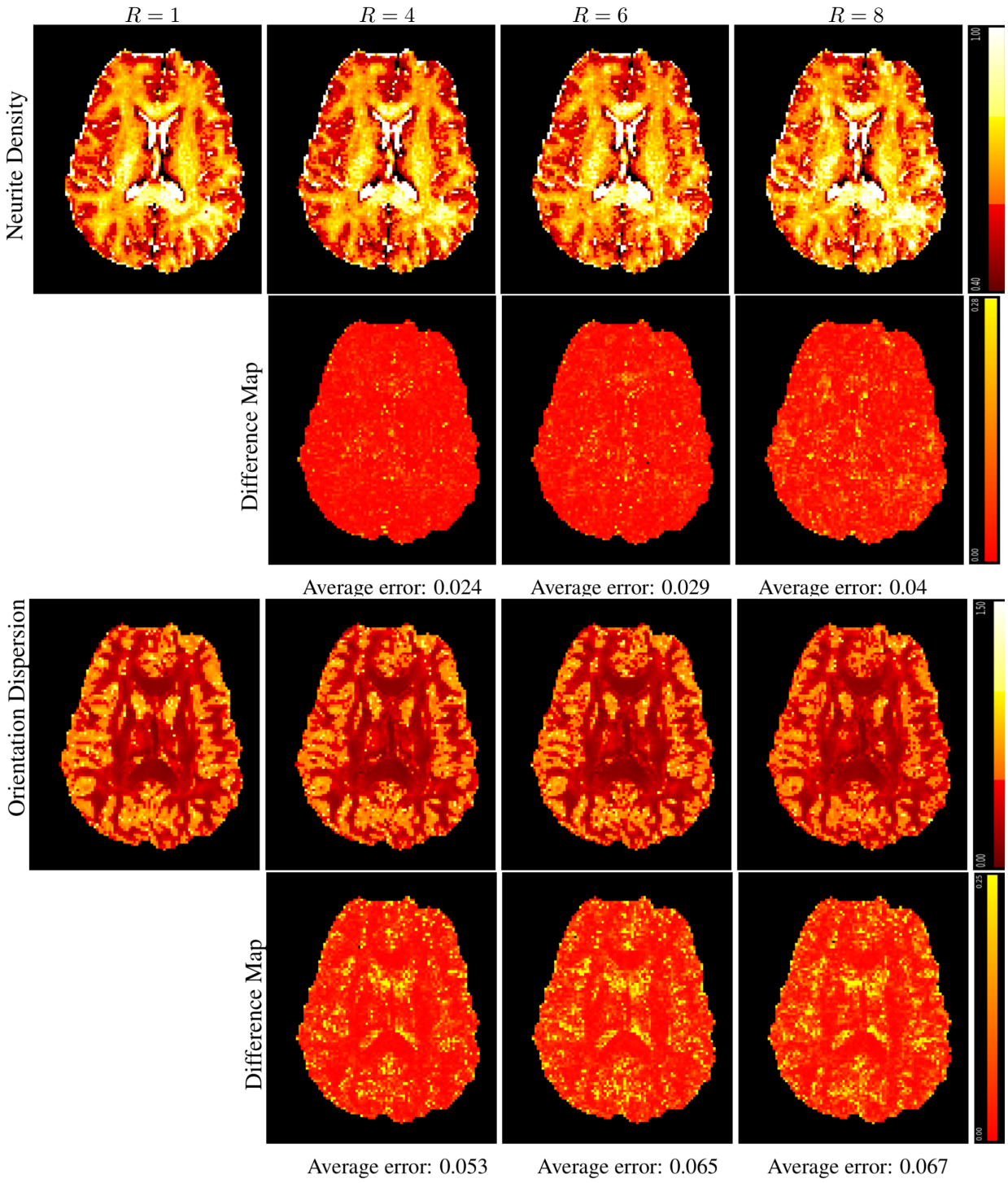


Figure 8: The microstructural studies using the accelerated multi-shell imaging using a compartmental model. The NODDI parameters, neurite density and the orientational dispersion, are computed on the fully sampled multi-shell data as well as the accelerated case at R=4,6 and 8. The maps and the error in the maps compared to the reference case are also shown. The average error is the maps are also noted.

conform to sparse representations. The learning-based approach utilizes very large dictionaries to derive a compact representation of the high-dimensional manifold. It is not affected by the correlations in the atoms. Because of the compactness of the learned representation, the reconstruction employing such priors are fast. We demonstrated improved performance of the new q-space prior compared to previous ℓ_1-based prior, in terms of the reconstruction quality, the reconstruction time and its ability to reconstruct multi-shell data.

The proposed scheme has significant differences with deep-learning based q-space acceleration techniques (45). These schemes rely on supervised learning to learn the mapping from the diffusion signal to the parameters of a specific model (e.g. NODDI) from training data (fully sampled q-space images). Once the training is achieved, the model parameters are recovered from accelerated q-space data. By contrast, our focus is to recover the diffusion weighted data itself, from accelerated k-q acquisition. With high spatial and q-space resolution of the recovered data, this allows the fitting of any desired diffusion model using traditional model-fitting routines. Further, in-vivo training data is completely absent in our setting, in contrast to methods exploiting u-NET trained priors (46, 47). The training is performed using numerical dictionaries computed from biophysical models. Hence, qModeL reconstruction makes use of biophysically-driven learning rather than data-driven learning. Nevertheless, our framework has the flexibility to learn signal manifolds from any given appropriate training data. Thus, q-space data manifolds can also be learned from in-vivo data samples. A data-driven approach may provide further opportunities to augment the learning based on biophysical modeling. In future works, we will explore if pre-learning using data-driven approaches offers further advantages compared to biophysical modeling approaches.

The biophysically driven training involved in qModeL framework has several advantages. First, the training involves a simple process which is computationally inexpensive. Hence re-training of the DAE is not an issue. Second, because the method does not require in-vivo training data, reproducibly of the method is robust and is not affected by the quality of the available data, nor restricted by the acquisition settings of the in-vivo training data. Moreover, highly complex acquisition settings can be simulated without needing extreme performance hardware and the long scan time required to accumulate exquisite SNRs. Translation between various acquisition settings can be easily achieved because of the ability to re-train the DAE in a short time. Third, the storing of the pre-learned DAE weights is cheap and can be easily stored on MRI scanners for various diffusion acquisition protocols similar to storing the q-space sampling patterns. These DAEs can be used for denoising purposes as well as for accelerated reconstructions on the scanners.

In the current work, we did not pursue any optimization of the k-q under-sampling patterns. The under-sampling was achieved in a purely random fashion and using existing multi-shot EPI sequences. Hence

there are several inefficiencies in the present setting which can be optimized to achieve improved performance. For example, the partial Fourier restricted the flexibility to maximize the random under-sampling employed in the acquisition. Alternatively, a variable density k-space sampling will have a higher SNR advantage. Non-Cartesian trajectories will also significantly improve the acceleration capability. Further experiments are needed to determine how far the spatial and the angular resolution can be pushed with the proposed method. Such studies are crucial to determine the added benefits brought forth by the proposed approaches on advanced microstructure studies. The demonstrated utility of the proposed method for accelerated multi-shell reconstruction further motivates sampling pattern optimizations (48) for such advanced microstructure studies. Further speed up of the proposed method can be obtained in conjunction with simultaneous multi-slice acquisition (49, 50).

In the present work, we have not implemented any comprehensive artifact correction such as correction of eddy-current induced effects, into the reconstruction. Existing post-processing routines (51) can be applied to the qModeL-reconstructed images. The presence of subject motion can have a detrimental effect on the qModeL joint reconstruction. In future works, joint motion correction and reconstruction strategies (52–54) will be pursued to address this limitation.

Conclusion: We presented a new method for the joint reconstruction of highly under-sampled multi-shot diffusion MRI data. The proposed method relies on a pre-learned regularizer that learns the signal manifold to predict the signal from missing q-space points. The q-space manifold is learned using a denoising autoencoder neural network. The learned encoder-decoder mapping is employed as a plug-and-play prior in an iterative model-based reconstruction setting. The method was tested on in-vivo datasets at various acceleration factors. The reconstructed images and the microstructural indices derived from these methods are shown to provide good accuracy to be used for routine microstructural studies.

Acknowledgements

Financial support for this study was provided by grants NIH 5 R01 EB022019, 5 R01 MH111578 and NIH 1R01EB019961-01A1. This work was conducted on MRI instruments funded by 1S10OD025025-01 and 1S10RR028821-01.

Legends:

Fig 1: Illustration of random k-q undersampling using multi-shot trajectories. The brown points on the green background represents the sampling points in q-space. Each q-space point is associated with its own k-space. (a) illustrates the fully sampled case. Here, (N_s =) 4 k-space shots (denoted using 4 colors) are required to fully sample the k-space of each q-space point. In the k-q under-sampled case (b), only 1 shot (shown in bold line) out of the four shots are used to sample a given q-space point (corresponding to an acceleration factor of $R = N_s = 4$). Each q-space point is sampled with a different shot, which allows to distribute the aliasing patterns differently in each of the diffusion weighted images. The sampling pattern shown above also accommodates partial Fourier imaging. To achieve $R > N_s$, we skip the sampling of some q-space points randomly (illustrated in (c)). Note that every sampled q-space point employs 1 random shot out of the 4 k-space shots, as in (b).

Fig 2: Illustration of the learning of the q-space manifold spanned by the 3-compartment model using a denoising autoencoder. The dictionary consists of data generated by discretizing the free parameters of the model. The denoising autoencoder is trained using the dictionary, which learns the manifold corresponding to the parameter space. The output of the autoencoder preserves the latent features of the input data.

Fig 3: The different components in the proposed reconstruction. (a) shows the two steps involved in the phase compensated recovery. In the first step, the background phase, θ, is estimated using a CG-SENSE reconstruction with smoothness regularization. In the second step, the estimated phase is incorporated into the forward operator $\mathcal{A}$ such that real-valued images are recovered. When phase compensation is not perfect, small residual phase errors can result in the images. The q-space denoiser is trained to correct for small phase errors. (b) shows the qModeL reconstruction which aims to minimize the projection error to a pre-learned q-space manifold. The pre-learned q-space manifold is a function of a large set of variables including the diffusion directions and the b-values used for the biophysical modeling. In the absence of noise, the q-space signal corresponding to every voxel will map to points on the manifold. In the noisy case, the reconstruction tries to minimize the projection error to this manifold subject to data-consistency to the measured data.

Fig 4: Dataset 1: (a) Test data from an in-vivo experiment. The q-space signal plot from the boxed region in (a) and the output of the trained autoencoder are plotted in (b). Blue plot indicates the test signal, red plot corresponds to output of DAE trained using the SM and the green plot corresponds to that using the BAS model. (c)-(d) shows the images formed after voxel-wise denoising using the SM and the BAS DAEs respectively.

Fig 5: Multi-shell data: (a) Three representative DWIs from b1000, b2000 and b3000 shells are shown. The images correspond to the CG-SENSE reconstruction from their fully sampled k-space measurements. (b)-(c) shows the DWIs after voxel-wise denoising using the SM and the BAS denoisers respectively. (d) shows the denoising using the SER method.

Fig 6: Under-sampled reconstruction at $R = 4$ from dataset 1. (a)-(b) corresponds to the CG-SENSE and the qModeL reconstruction of the fully sampled data (with partial Fourier), (c)-(d) shows the CG-SENSE and the qModeL reconstruction at $R = 4$. (e)-(f) shows the error in the qModeL reconstruction at $R = 4$, with respect to the CG-SENSE

and the qModeL reconstruction at $R = 1$, magnified by 2 times and 6 times respectively. All images are shown on the same display scale [0, 1]. The RMSE for the full set of DWIs reconstructed at $R = 4$ is reported at the bottom with respect to the CG-SENSE and the qModeL references.

Fig 7: Multi-shell experiment at $R = 8$. (a)-(b) shows the fully sampled reconstruction using the CG-SENSE and the qModeL joint reconstruction. (c)-(d) shows the initial guess ($\mathcal{A}^T(\widehat{\mathbf{Y}})$) and the qModeL reconstruction at $R = 8$. Column (c) shows empty volumes (row 2,3) corresponding to q-space points that were not sampled. Column (d) shows the qModeL reconstruction, which recovers all the DWIs, including the missing volumes. Column (e) shows the residual error with respect to the qModeL reference, magnified 6 times. The signal plot from the boxed region is shown in (f). The red curve plots the q-space signal from $\mathcal{A}^T(\widehat{\mathbf{Y}})$, which shows missing volumes. The blue and the green curve plots the signal from the qModeL reconstruction at $R = 1$ and 8 respectively. The green plot is shown to recover the missing points and is noted to match the reference signal from the fully sampled case.

Fig 8: The microstructural studies using the accelerated multi-shell imaging using a compartmental model. The NODDI parameters, neurite density and the orientational dispersion, are computed on the fully sampled multi-shell data as well as the accelerated case at R =4,6 and 8. The maps and the error in the maps compared to the reference case are also shown. The average error is the maps are also noted.

Fig S1: (a) shows the multi-shell reconstruction using L1-TV method where separate dictionaries are used for each shell. In this case, the DWIs from each shell retains their uniques contrasts both for $R = 1$ and $R = 4$. (b) shows the multi-shell reconstruction using a single dictionary where the atoms corresponding to different b-shells are present. Because of the highly correlated atoms in this dictionary, the joint reconstruction fails to retain the contrasts of the different b-shells. All shells converges to the same contrast. The diffusion gradient contrasts along the different direction vectors are retained.

Fig S2: The FA maps and the primary eigen vector maps reconstructed from dataset 2 at $R = 4$ using qModeL and L1-TV reconstruction. The reference maps computed using the fully sampled reconstruction are given in the top row. The error in the maps are displayed in the bottom row. The average error in the whole brain in the FA maps and DEC maps are reported for both methods with the average angular error within the white matter given in the brackets.

Fig S3: Zoomed in view of the FA maps given in figure S2 to enable better visualization of the cortical microstructures.

Fig S4: Reconstruction results at $R = 8$ from dataset 2. The top row left image shows a sample DWI volume #60 from the fully sampled case, reconstructed using qModeL. At $R = 8$, this DWI belonged to the set of 30 volumes that were not sampled (i,e, the k-space data from this volume is completely missing). The image on the right shows the reconstruction of this missing volume using the qModeL joint recovery. The bottom row shows the FA maps and the primary eigen vector maps estimated from qModeL reconstruction of all 60 volumes, from the 8-fold under-sampled data. The average angular error in the data is 7.3°.

Fig S5: The reconstruction of the multi-shell data at various acceleration factors using the qModeL reconstruction with and without the TV denoiser.

Table 1: Comparison of qModeL and L1-TV reconstruction at $R = 4$.

Table 2: qModeL reconstruction errors at $R = 6, 8$.

Table S1: qModeL reconstruction errors without the TV denoiser.

References

1 Pierpaoli C, Jezzard P, Basser P, Barnett A, and Chiro G. Diffusion Tensor MR Imaging of the Human Brain. *Radiology*, 201:637–648, 1996.

2 Horsfield MA and Jones DK. Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases - a review. *NMR in Biomedicine*, 15(7-8):570–577, 2002.

3 Reisert M, Kellner E, Dhital B, Hennig J, and Kiselev VG. Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach. *NeuroImage*, 147:964–975, 2017.

4 Veraart J, Fieremans E, and Novikov DS. On the scaling behavior of water diffusion in human brain white matter. *NeuroImage*, 185:379–387, 2019.

5 Novikov DS, Kiselev VG, and Jespersen SN. On modeling. *Magnetic Resonance in Medicine*, 79(6):3172–3193, 2018.

6 Novikov DS, Fieremans E, Jespersen SN, and Kiselev VG. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR in Biomedicine*, 32(4):e3998, 2019.

7 Jelescu IO, Palombo M, Bagnato F, and Schilling KG. Challenges for biophysical modeling of microstructure. *Journal of Neuroscience Methods*, 344:108861, 2020.

8 Zhang H, Schneider T, Wheeler-Kingshott CA, and Alexander DC. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 2012.

9 Jelescu IO, Veraart J, Adisetiyo V, Milla SS, Novikov DS, and Fieremans E. One diffusion acquisition and different white matter models: How does microstructure change in human early development based on WMTI and NODDI? *NeuroImage*, 107:242–256, 2015.

10 Mozumder M, Pozo JM, Coelho S, and Frangi AF. Populationbased Bayesian regularization for microstructural diffusion MRI with NODDIDA. *Magnetic Resonance in Medicine*, 82(4):1553–1565, 2019.

11 J Behrens TE, W Woolrich M, Jenkinson M, Johansen-Berg H, G Nunes R, Clare S, M Matthews P, M Brady J, and M Smith S. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50(5):1077–1088, 2003.

12 Assaf Y and Basser PJ. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage*, 27(1):48–58, 2005.

13 Wu W and Miller KL. Image formation in diffusion MRI: A review of recent technical developments. *Journal of Magnetic Resonance Imaging*, 46(3):646–662, 2017.

14 Mani M, Aggarwal HK, Magnotta V, and Jacob M. Improved MUSSELS reconstruction for highresolution multi-shot diffusion weighted imaging. *Magnetic Resonance in Medicine*, 83(6):2253–2263, 2020.

15 Bruce IP, Chang HC, Petty C, Chen NK, and Song AW. 3D-MB-MUSE: A robust 3D multi-slab, multi-band and multi-shot reconstruction approach for ultrahigh resolution diffusion MRI. *NeuroImage*, 159:46–56, 2017.

16 Mani M, Jacob M, Guidon A, Liu C, Song A, Magnotta V, and Zhong J. Acceleration of high angular and spatial resolution diffusion imaging using compressed sensing. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 326–329, 2012. ISBN 978-1-4577-1858-8.

17 Mani M, Jacob M, Guidon A, Magnotta V, and Zhong J. Acceleration of high angular and spatial resolution diffusion imaging using compressed sensing with multichannel spiral data. *Magnetic Resonance in Medicine*, 73 (1):126–138, 2015.

18 Wu W, Koopmans PJ, Andersson JL, and Miller KL. Diffusion Acceleration with Gaussian process Estimated Reconstruction (DAGER). *Magnetic Resonance in Medicine*, 82(1):107–125, 2019.

19 RamosLlordén G, Ning L, Liao C, Mukhometzianov R, Michailovich O, Setsompop K, and Rathi Y. Highfidelity, accelerated wholebrain submillimeter in vivo diffusion MRI using gSliderspherical ridgelets (gSliderSR). *Magnetic Resonance in Medicine*, 84(4):1781–1795, 2020.

20 Shi X, Ma X, Wu W, Huang F, Yuan C, and Guo H. Parallel imaging and compressed sensing combined framework for accelerating high-resolution diffusion tensor imaging using inter-image correlation. *Magnetic Resonance in Medicine*, 73(5):1775–1785, 2015.

21 Schwab E, Vidal R, and Charon N. (k, q)-Compressed Sensing for dMRI with Joint Spatial-Angular Sparsity Prior. pages 21–35. Springer, Cham, 2018.

22 Bilgic B, Chatnuntawech I, Setsompop K, Cauley SF, Yendiki A, Wald LL, and Adalsteinsson E. Fast Dictionary-Based Reconstruction for Diffusion Spectrum Imaging. *IEEE Transactions on Medical Imaging*, 32(11):2022–2033, 2013.

23 Chan SH, Wang X, and Elgendy OA. Plug-and-Play ADMM for Image Restoration: Fixed-Point Convergence and Applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, 2016.

24 Venkatakrishnan SV, Bouman CA, and Wohlberg B. Plug-and-Play priors for model based reconstruction. In *2013 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2013 - Proceedings*, pages 945–948, 2013. ISBN 9781479902484.

25 Ahmad R, Bouman CA, Buzzard GT, Chan S, Liu S, Reehorst ET, and Schniter P. Plug-and-Play Methods for Magnetic Resonance Imaging: Using Denoisers for Image Recovery. *IEEE Signal Processing Magazine*, 37(1): 105–116, 2020.

26 Liu C, Moseley ME, and Bammer R. Simultaneous phase correction and SENSE reconstruction for navigated multi-shot DWI with non-cartesian k-space sampling. *Magnetic Resonance in Medicine*, 54(6):1412–22, 2005.

27 Gao H, Li L, Zhang K, Zhou W, and Hu X. PCLR: Phase-constrained low-rank model for compressive diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 72(5):1330–1341, 2014.

28 Chao TC, Chiou JyG, Maier SE, and Madore B. Fast diffusion imaging with high angular resolution. *Magnetic Resonance in Medicine*, 77(2):696–706, 2017.

29 Mani MP, Aggarwal HK, Ghosh S, and Jacob M. Model-Based Deep Learning for Reconstruction of Joint k-q Under-sampled High Resolution Diffusion MRI. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2020.

30 Landman BA, Bogovic JA, Wan H, El Zahraa ElShahaby F, Bazin PL, and Prince JL. Resolution of crossing fibers with constrained compressed sensing using diffusion tensor MRI. *NeuroImage*, 59(3):2175–86, 2012.

31 Michailovich O, Rathi Y, and Dolui S. Spatially Regularized Compressed Sensing for High Angular Resolution Diffusion Imaging. *IEEE Transactions on Medical Imaging*, 30(5):1100–1115, 2011.

32 Descoteaux M, Angelino E, Fitzgibbons S, and Deriche R. Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510, 2007.

33 Daducci A, Canales-Rodríguez EJ, Zhang H, Dyrby TB, Alexander DC, and Thiran JP. Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. *NeuroImage*, 105:32–44, 2015.

34 Hinton GE and Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science*, 313(5786): 504–507, 2006.

35 Raza K and Singh NK. A Tour of Unsupervised Deep Learning for Medical Image Analysis. 2018.

36 Goodfellow I, Bengio Y, and Courville A. *Deep learning: The MIT Press*, volume 19. 2017.

37 Vincent P, Larochelle H, Lajoie I, Bengio Y, and Manzagol PA. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research*, 11: 3371–3408, 2010.

38 Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, and Woolrich MW. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1):144–55, 2007.

39 Jelescu IO, Veraart J, Fieremans E, and Novikov DS. Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR in Biomedicine*, 2016.

40 Haldar JP, Wedeen VJ, Nezamzadeh M, Dai G, Weiner MW, Schuff N, and Liang ZP. Improved diffusion imaging through SNR-enhancing joint reconstruction. *Magnetic Resonance in Medicine*, 69(1):277–289, 2013.

41 Haldar JP, Liu Y, Liao C, Fan Q, and Setsompop K. Fast submillimeter diffusion MRI using gSliderSMS and SNRenhancing joint reconstruction. *Magnetic Resonance in Medicine*, 84(2):762–776, 2020.

42 Lourakis M. TV-L1 Image Denoising Algorithm.

43 Chen NK, Guidon A, Chang HC, and Song AW. A robust multi-shot scan strategy for high-resolution diffusion weighted MRI enabled by multiplexed sensitivity-encoding (MUSE). *NeuroImage*, 72:41–47, 2013.

44 Harms RL, Fritz FJ, Tobisch A, Goebel R, and Roebroeck A. Robust and fast nonlinear optimization of diffusion MRI microstructure models. *NeuroImage*, 155:82–96, 2017.

45 Golkov V, Dosovitskiy A, Sperl JI, Menzel MI, Czisch M, Samann P, Brox T, and Cremers D. q-Space Deep Learning: Twelve-Fold Shorter and Model-Free Diffusion MRI Scans. *IEEE Transactions on Medical Imaging*, 35(5):1344–1351, 2016.

46 Bilgic B, Chatnuntawech I, Manhard MK, Tian Q, Liao C, Iyer SS, Cauley SF, Huang SY, Polimeni JR, Wald LL, and Setsompop K. Highly accelerated multishot echo planar imaging through synergistic machine learning and joint reconstruction. *Magnetic Resonance in Medicine*, 82(4):1343–1358, 2019.

47 Hu Y, Xu Y, Tian Q, Chen F, Shi X, Moran CJ, Daniel BL, and Hargreaves BA. RUN-UP: Accelerated multishot diffusion-weighted MRI reconstruction using an unrolled network with U-Net as priors. *Magnetic Resonance in Medicine*, 2020.

48 Aggarwal HK and Jacob M. J-MoDL: Joint Model-Based Deep Learning for Optimized Sampling and Reconstruction. 2019.

49 Barth M, Breuer F, Koopmans PJ, Norris DG, and Poser BA. Simultaneous multislice (SMS) imaging techniques. *Magnetic resonance in medicine*, 75(1):63–81, 2016.

50 Mani M, Jacob M, McKinnon G, Yang B, Rutt B, Kerr A, and Magnotta V. SMS MUSSELS: A navigator-free reconstruction for simultaneous multi-slice-accelerated multi-shot diffusion weighted imaging. *Magnetic Resonance in Medicine*, 83(1):154–169, 2020.

51 Andersson JL and Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, 125:1063–1078, 2016.

52 Guhaniyogi S, Chu ML, Chang HC, Song AW, and Chen Nk. Motion immune diffusion imaging using augmented MUSE for high-resolution multi-shot EPI. *Magnetic Resonance in Medicine*, 75(2):639–652, 2016.

53 Sommer K, Saalbach A, Brosch T, Hall C, Cross NM, and Andre JB. Correction of Motion Artifacts Using a Multiscale Fully Convolutional Neural Network. *AJNR. American journal of neuroradiology*, 41(3):416–423, 2020.

54 Poddar S and Jacob M. Dynamic MRI Using SmooThness Regularization on Manifolds (SToRM). *IEEE Transactions on Medical Imaging*, 35(4):1106–1115, 2016.