MoDL-MUSSELS: Model-Based Deep Learning for Multi-Shot Sensitivity Encoded Diffusion MRI

Hemant K. Aggarwal, *Member, IEEE*, Merry P. Mani, and Mathews Jacob, *Senior Member, IEEE*

***Abstract*—We introduce a model-based deep learning architecture termed MoDL-MUSSELS for the correction of phase errors in multishot diffusion-weighted echo-planar MRI images. The proposed algorithm is a generalization of existing MUSSELS algorithm with similar performance but with significantly reduced computational complexity. In this work, we show that an iterative re-weighted least-squares implementation of MUSSELS alternates between a multichannel filter bank and the enforcement of data consistency. The multichannel filter bank projects the data to the signal subspace thus exploiting the phase relations between shots. Due to the high computational complexity of self learned filter bank, we propose to replace it with a convolutional neural network (CNN) whose parameters are learned from exemplary data. The proposed CNN is a hybrid model involving a multichannel CNN in the k-space and another CNN in the image space. The k-space CNN exploits the phase relations between the shot images, while the image domain network is used to project the data to an image manifold. The experiments show that the proposed scheme can yield reconstructions that are comparable to state of the art methods while offering several orders of magnitude reduction in run-time.**

***Index Terms*—Diffusion MRI, Echo Planar Imaging, Deep Learning, convolutional neural network**

I. Introduction

Diffusion MRI (DMRI), which is sensitive to anisotropic diffusion processes in the brain tissue, has the potential to provide rich information on white matter anatomy [1] and hence have several applications including neurological disorders [2], the aging process [3], and acute stroke [4]. DMRI relies on large bipolar directional gradients to encode water diffusion, results in the attenuation of signals from diffusing molecules in the direction of the gradient. The diffusion encoded signal is often spatially encoded using single-shot echo planar imaging (ssEPI), which allows the acquisition of the entire k-space in a single excitation. While it can offer high sampling efficiency, the longer readout makes the acquisition vulnerable to B0 inhomogeneity induced distortions. Specifically, the recovered images often exhibit geometric distortions and signal drop-outs along the phase encoding direction. These artifacts essentially limit the extent of k-space coverage and thereby the spatial resolution that ssEPI sequences can achieve.

Multi-shot echo planar imaging (msEPI) methods were introduced to minimize the distortions related to the long readouts in ssEPI. This scheme segments the k-space over multiple excitation and shots as shown in (Fig. 1), which shortens the readout duration for each of shot. While multi-shot imaging can offer high resolution, a challenge is its vulnerability to inter-shot motion in the diffusion setting. Specifically, subtle physiological motion during the large bipolar gradients manifest as phase differences between different shots. The direct combination of the k-space data from these shots results in Nyquist ghost artifacts.

We recently introduced a multi-shot sensitivity-encoded diffusion data recovery algorithm using structured low-rank matrix completion (MUSSELS) [5], which allows the combination the k-space data from different shots. The method exploits the redundancy between the Fourier samples of the shots to jointly recover the missing k-space samples in each of the shots. The k-space data recovery is then posed as a matrix completion problem that utilizes a structured low-rank algorithm and parallel imaging to recover the missing k-space data in each shot. While this scheme can offer state of the art results, the challenge is the high computational complexity. The large data size and the need for matrix lifting make it challenging to reconstruct the high-resolution data from different directions and slices despite the existence of fast structured low-rank algorithms.

In this paper, we introduce a novel deep learning framework to minimize the computational complexity of MUSSELS [5]. This work is inspired by the network structure of MUSSELS and is similarly formulated in k-space to exploit the convolutional relations between the Fourier samples of the shots. The proposed scheme is also motivated by our recent work on model-based deep learning (MoDL) [6] and similar algorithms that rely on un-rolling of iterative algorithms [7]–[9]. The main benefit of MoDL is the ability to exploit the physics of the acquisition scheme, add multiple regularization priors [10], and improve performance. In addition, the unrolled and learned recovery scheme offers significantly reduced run time during image recovery/testing. The use of the conjugate-gradient algorithm within the network to enforce data consistency in MoDL provides improved performance for a specified number of iterations. The sharing of network parameters across iterations enables MoDL to keep the number of learned parameters decoupled from the number of iterations, thus providing good convergence without increasing the number of trainable parameters. A lesser number of trainable parameters translate to significantly reduced training data in data constrained medical imaging applications.

We first introduce an iterative reweighted least-squares algorithm (IRLS) [11] based approach to solve the MUSSELS cost function [5]. The original MUSSELS algorithm, which is based on iterative singular value shrinkage, alternates between a data-consistency block and a low-rank matrix recovery block. By contrast, the IRLS algorithm alternates between a data-consistency block and a residual multichannel convolution

This work is supported by 1R01EB01996101A1.

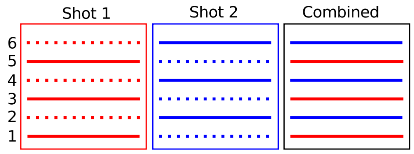


Fig. 1. Demonstration of multi-shot EPI acquisition employing multiple excitations and readouts. The first RF excitation and diffusion sensitization are followed by k-space readout by shot 1 that samples k-space lines 1,3, and 5. The second RF excitation and diffusion sensitization are followed by k-space readout by shot 2 capture lines 2,4, and 6. The combined data corresponds to the fully sampled k-space. Yet, the k-space samples in shot 1 and shot 2 are sensitized to different motion due to omnipresent inter-shot motion and hence will have a unique net phase. Thus, the image generated from the combined shot will show phase artifacts. MUSSELS attempts to recover the missing k-space samples in shot 1 and shot 2 thus bypassing the need to combine the k-space samples from different shots

block. The multichannel convolution block can be viewed as the projection of the data to the null-space of the multi-channel signals; the subtraction of the result from the original ones, induced by the residual structure, projects the data to the signal subspace, thus removing the artifacts in the signal. The IRLS MUSSELS algorithm learns the parameters of the denoising filter from the data itself, which requires several iterations. Motivated by [6], we propose to replace the multichannel linear convolution block in MUSSELS by a convolutional neural network (CNN). Unlike the self-learning strategy in MUSSELS, where the filter parameters are learned from the measured data itself, we propose to learn the parameters of the non-linear CNN from exemplar data. We hypothesize that the non-linear structure of the CNN will enable us to learn and generalize from examples; the learned CNN will facilitate the projection of each test dataset to the associated signal subspace. While the architecture is conceptually similar to MoDL, the main difference is the extension to multichannel settings and the learning in the Fourier domain (k-space) motivated by the MUSSELS IRLS formulation.

The proposed framework has similarities to recent k-space deep learning strategies [12]–[15], which also exploit the convolution relations in the Fourier domain. The main distinction of the proposed scheme with these methods is the model-based framework, along with the training of the unrolled network. Many of the current schemes [14] are not designed for the parallel imaging setting. The use of the conjugate gradient steps in our network allows us to account for parallel imaging in an efficient manner, requiring few iterations. We also note the relation of the proposed work with [16], which uses a self-learned network to recover parallel MRI data; the weights of the network are estimated from the measured data itself. Since we estimate the weights from exemplar data, the proposed scheme is significantly faster.

II. Background

A. Problem formulation

The long EPI readouts, which are needed for high-resolution diffusion MRI, are vulnerable to field inhomogeneity induced spatial distortions. In addition, the large rewinder gradients also make the achievable echo-time rather long, resulting in lower signal to noise ratio. To minimize these distortions, It is a common practice to acquire the data using multishot EPI schemes. These schemes acquire a highly undersampled subset of k-space at each shot; since the subsets are complementary, the data from all these shots are combined together to obtain the final image. The image acquisition of the i^{th} shot and the j^{th} coil can be expressed as

$$y_{i,j}[\mathbf{k}] = \int_{\mathbb{R}^2} \rho(\mathbf{r}) s_j(\mathbf{r}) \exp\left(\mathbb{i}\ \mathbf{k}^T\mathbf{r}\right) d\mathbf{r} + n_{i,j}[\mathbf{k}]; \quad \forall \mathbf{k} \in \Theta_i. \tag{1}$$

Here, $\mathbf{s}_j(\mathbf{r})$ denotes the coil sensitivity of the j^{th} coil and $\Theta_i; i = 1,..,N$ denotes the subset of the k-space that is acquired at the i^{th}-shot. Note that the sampling indices of the different shots are complementary; the combination of the data from the different shots will result in a fully sampled image. Specifically, we have $\bigcup_{i=1}^{N} \Theta_i = \Theta$, where Θ is the Fourier grid corresponding to the fully sampled image. The above relation, to acquire desired image $\rho(\mathbf{r})$ from N shots can be compactly represented as

$$\mathbf{y}_i = \mathcal{A}_i(\rho(\mathbf{r})) + \mathbf{n}, \quad i = 1,..,N \tag{2}$$

in the absence of phase errors. Here, $\mathbf{y}_i$ represents the undersampled multi-channel measurements of i^{th} shot acquired using acquisition operator $\mathcal{A}_i$ and $\mathbf{n}$ represents the additive Gaussian noise that may corrupt the samples during acquisition.

Diffusion MRI uses large bipolar diffusion gradients to encode the diffusion motion of water molecules. Unfortunately, subtle physiological motion between the bipolar gradients often manifests as phase errors in the acquisition. With the addition of the unknown phase function $\phi_i(\mathbf{r}); |\phi_i(\mathbf{r})| = 1$ introduced by physiological motion, the forward model gets modified as

$$\mathbf{y}_i = \mathcal{A}_i\Big(\underbrace{\rho\ (\mathbf{r})\phi_i(\mathbf{r})}_{\rho_i(\mathbf{r})}\Big) + \mathbf{n}, \quad i = 1,..,N \tag{3}$$

If the phase errors $\phi_i(\mathbf{r}); i = 1,..,N$ are uncompensated, the image obtained by the combination of $\mathbf{y}_i$, $i = 1,..,N$ will consist of Nyquist ghosting artifacts. Current multishot methods on GE scanners termed as MUSE [17] rely on the independent estimation of $\phi_i(\mathbf{r})$, $i = 1,..,N$ from low-resolution reconstructions of the phase corrupted images ρ_i. The forward model can be compactly written as $\mathbf{y} = \mathcal{A}(\boldsymbol{\rho})$, where $\boldsymbol{\rho} = \left[\boldsymbol{\rho}_1{}^T, \quad \ldots \quad \boldsymbol{\rho}_N{}^T\right]^T$ is the vector of multishot images. Once the phases are estimated, the reconstruction is posed as a phase aware reconstruction [17].

B. Brief Review of MUSSELS

MUSSELS relies on a structured low-rank formulation to jointly recover the phase corrupted images ρ_i from their under-sampled multichannel measurements, capitalizing on the multichannel nature of the measurements as well as annihilation relations between the phase corrupted images. The key observation is that the above images satisfy an image domain annihilation relation $\rho_i(\mathbf{r})\phi_j(\mathbf{r}) - \rho_j(\mathbf{r})\phi_i(\mathbf{r}) = 0, \forall \mathbf{r}$ [18]. This multiplicative relation translates to convolution relations in the Fourier domain:

$$\widehat{\rho_i}(\mathbf{k}) * \widehat{\phi_j}(\mathbf{k}) - \widehat{\rho_j}(\mathbf{k}) * \widehat{\phi_i}(\mathbf{k}) = 0 \quad \forall \mathbf{k}, \tag{4}$$

where $\widehat{x}$ denotes the Fourier transform of x. Since the phase images $\phi_j(\mathbf{r})$ are smooth, their Fourier coefficients $\widehat{\phi_j}(\mathbf{k})$ can be assumed to be support limited to a region Λ in the Fourier domain. This allows us to rewrite the convolution relations in (4) in a matrix form using block-Hankel convolution matrices $\mathbf{H}_\Lambda^\Gamma(\rho)$. The matrix product $\mathbf{H}_\Lambda^\Gamma(\rho)\ \mathbf{h}$ corresponds to the 2D convolution between a signal ρ supported on a grid Γ and the filter $\mathbf{h}$ of size Λ. Thus, the Fourier domain convolution relations can be compactly expressed using matrix matrices [5] as

$$\left[\mathbf{H}_\Lambda^\Gamma(\hat{\rho}_i)|\mathbf{H}_\Lambda^\Gamma(\hat{\rho}_j)\right]\begin{bmatrix}\hat{\boldsymbol{\phi}}_{\boldsymbol{j}}\\ -\hat{\boldsymbol{\phi}}_{\boldsymbol{i}}\end{bmatrix} = \mathbf{0}.$$

We note that there exists a similar annihilation relation between each pair of shots, which imply that the structured matrix

$$\mathbf{T}(\widehat{\boldsymbol{\rho}}) = \left[\mathbf{H}_\Lambda^\Gamma(\hat{\rho}_1)\ |\ \cdots\ |\ \mathbf{H}_\Lambda^\Gamma(\hat{\rho}_N)\right] \tag{5}$$

is low-rank. MUSSELS recovers the multi-shot images from their undersampled k-space measurements by solving:

$$\tilde{\rho} = \arg\min_{\boldsymbol{\rho}} \left\|\mathcal{A}(\boldsymbol{\rho}) - \mathbf{y}\right\|_2^2 + \lambda\left\|\mathbf{T}\left(\widehat{\boldsymbol{\rho}}\right)\right\|_*, \tag{6}$$

where $\|\cdot\|_*$ denotes the nuclear norm. The above problem is solved in [5] using iterative shrinkage algorithm.

III. Deep learned MUSSELS

A. IRLS reformulation of MUSSELS

To bring the MUSSELS framework to the MoDL setting, we first introduce an iterative reweighted least squares (IRLS) reformulation [11] of MUSSELS. Using an auxiliary variable $\mathbf{z}$, we rewrite (6) as

$$\arg\min_{\boldsymbol{\rho},\mathbf{z}}\ \left\|\mathcal{A}(\boldsymbol{\rho}) - \mathbf{y}\right\|_2^2 + \beta\|\widehat{\boldsymbol{\rho}} - \mathbf{z}\|_F^2 + \lambda\|\mathbf{T}(\mathbf{z})\|_* \tag{7}$$

We observe that (7) is equivalent to (6) as $\beta \to \infty$. An alternating minimization algorithm to solve the above problem yields the following steps:

$$\begin{aligned}\boldsymbol{\rho}_{n+1} &= \arg\min_{\boldsymbol{\rho}} \left\|\mathcal{A}(\boldsymbol{\rho}) - \mathbf{y}\right\|_2^2 + \beta\ \|\widehat{\boldsymbol{\rho}} - \mathbf{z}_n\|_F^2 && (8)\\ \mathbf{z}_{n+1} &= \arg\min_{\mathbf{z}} \|\widehat{\boldsymbol{\rho}}_{n+1} - \mathbf{z}\|_F^2 + \frac{\lambda}{\beta}\ \|\mathbf{T}(\mathbf{z})\|_* && (9)\end{aligned}$$

We now borrow from [11], [19] and majorize the nuclear norm term in (9) as

$$\left\|\mathbf{T}(\mathbf{z})\right\|_* \leq \left\|\mathbf{T}(\mathbf{z})\mathbf{Q}\right\|_F^2, \tag{10}$$

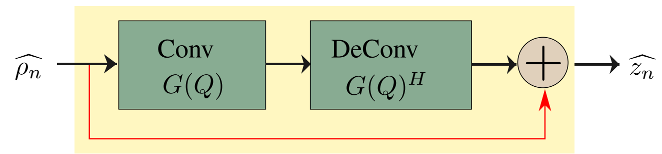
where the weight matrix is specified by

$$\mathbf{Q} = \left[\mathbf{T}^H(\mathbf{z})\mathbf{T}(\mathbf{z}) + \epsilon\mathbf{I}\right]^{-1/4} \tag{11}$$

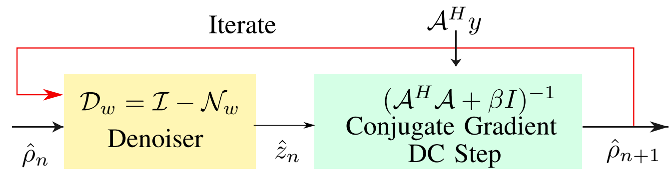
Here, $\mathbf{I}$ is the identity matrix. With the majorization (10), the $\mathbf{z}$-subproblem in (9) would involve the alternation between

$$\mathbf{z}_{n+1} = \arg\min_{\mathbf{z}} \|\widehat{\boldsymbol{\rho}}_{n+1} - \mathbf{z}\|_F^2 + \frac{\lambda}{\beta}\ \|\mathbf{T}(\mathbf{z})\mathbf{Q}\|_F^2 \tag{12}$$

and the update of the $\mathbf{Q}$ using (11). Thus the IRLS reformulation of MUSSELS scheme would alternate between (8), (12), and (11). The matrix $\mathbf{Q}$ may be viewed as a surrogate for the null-space of $\mathbf{T}(\mathbf{z})$. The update step (12) can be interpreted as finding an approximation of $\widehat{\boldsymbol{\rho}}_{n+1}$ from the signal subspace.



(a) Representation of Eq. (18) as MUSSELS Denoiser $\mathcal{D}_w$.



(b) The IRLS MUSSELS algorithm

Fig. 2. (a). The interpretation of Eq. (18) as a convolution-deconvolution network. (b) The IRLS MUSSELS iterates between (18), and (8). The data consistency (DC) step represents the solution of Eq. (8).

B. Interpretation of MUSSELS as an iterative denoiser

We note that the entries of the matrix $\mathbf{Q}$ can be split as

$$\mathbf{Q} = \begin{bmatrix}\mathbf{q}_{11} & \dots & \mathbf{q}_{N,1}\\ \dots & & \\ \mathbf{q}_{1N} & \dots & \mathbf{q}_{NN}\end{bmatrix} \tag{13}$$

such that

$$\mathbf{T}(\mathbf{z})\mathbf{q}_1 = \mathbf{H}_\Lambda^\Gamma(\mathbf{z}_1)\mathbf{q}_{11} + ..\mathbf{H}_\Lambda^\Gamma(\mathbf{z}_N)\mathbf{q}_{N1}.$$

Due to commutativity of convolution $h * g = g * h$, we have the relation

$$\mathbf{H}_\Lambda^\Gamma(\mathbf{g})\mathbf{h} = \mathbf{S}(\mathbf{h})\mathbf{g}, \tag{14}$$

where $\mathbf{S}(\mathbf{h})$ is an appropriately sized block Hankel matrix constructed from the zero-filled entries of $\mathbf{h}$. We use this relation to rewrite

$$\mathbf{T}\left(\mathbf{z}\right)\mathbf{Q} = \underbrace{\begin{bmatrix}\mathbf{S}(\mathbf{q}_{11}) & \mathbf{S}(\mathbf{q}_{12}) & \dots & \mathbf{S}(\mathbf{q}_{1N})\\ \vdots & \dots & \vdots & \\ \mathbf{S}(\mathbf{q}_{N1}) & \mathbf{S}(\mathbf{q}_{12}) & \dots & \mathbf{S}(\mathbf{q}_{NN})\end{bmatrix}}_{\mathbf{G}(\mathbf{Q})}\underbrace{\begin{bmatrix}\mathbf{z}_1\\ \vdots\\ \mathbf{z}_N\end{bmatrix}}_{\mathbf{z}} \tag{15}$$

We note that $\mathbf{G}(\mathbf{Q})\mathbf{z}$ correspond to the multichannel convolution of $\mathbf{z}_1, \dots, \mathbf{z}_N$ with the filterbank having filters $\mathbf{q}_{i,j}$. With this reformulation, (12) is simplified as

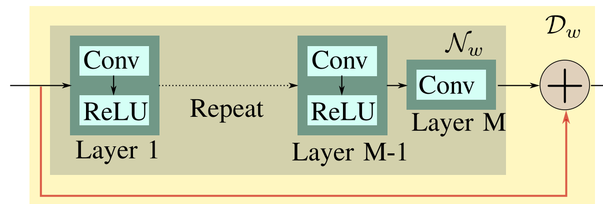
$$\mathbf{z}_{n+1} = \arg\min_{\mathbf{z}} \|\widehat{\boldsymbol{\rho}}_n - \mathbf{z}\|_F^2 + \frac{\lambda}{\beta}\left\|\mathbf{G}\left(\mathbf{Q}\right)\mathbf{z}\right\|_F^2 \tag{16}$$

Differentiating the above expression and setting it equal to zero, we get

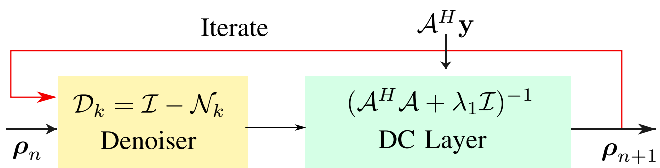
$$\mathbf{z}_{n+1} = \left(\mathbf{I} + \frac{\lambda}{\beta}\ \mathbf{G}\left(\mathbf{Q}\right)^H \mathbf{G}\left(\mathbf{Q}\right)\right)^{-1} \boldsymbol{\rho}_{n+1} \tag{17}$$

One may use a numerical solver to determine $\mathbf{z}_{n+1}$. An alternative is to solve this step approximately using the matrix inversion lemma, assuming $\lambda << \beta$:

$$\begin{aligned}\mathbf{z}_{n+1} &\approx \left[\mathbf{I} - \frac{\lambda}{\beta}\mathbf{G}\left(\mathbf{Q}\right)^H\mathbf{G}\left(\mathbf{Q}\right)\right]\ \widehat{\boldsymbol{\rho}}_{n+1}\\ &= \widehat{\boldsymbol{\rho}}_{n+1} - \frac{\lambda}{\beta}\mathbf{G}\left(\mathbf{Q}\right)^H\mathbf{G}\left(\mathbf{Q}\right)\ \widehat{\boldsymbol{\rho}}_{n+1} && (18)\end{aligned}$$



(a) The M-layer CNN based Denoiser



(b) Proposed k-space MoDL-MUSSELS architecture

Fig. 3. The block diagram of the proposed k-space network architecture to solve Eq. (19). (a) The $\mathcal{N}_w$ block represents deep learned noise predictor and $\mathcal{D}_w$ is a residual learning block. (b) Here, the denoiser D_k is the M-layer network $\mathcal{D}_w$ that performs k-space denoising.

We note that $G(\mathbf{Q})$ can be viewed as a single layer convolutional filter bank, while multiplication by $G(\mathbf{Q})^H$ can be viewed as flipped convolutions (deconvolutions in deep learning context) with matching boundary conditions. Note that both of the above layers do not have any non-linearities. Thus, (18) can be thought of as a residual block, which involves the convolution of the multishot signals $\widehat{\boldsymbol{\rho}}_n$ with the columns of $\mathbf{Q}$, followed by deconvolution as shown in Fig. 2(a). As discussed before, the filters specified by the columns of $\mathbf{Q}$ are surrogates for the null-space of $\mathbf{T}(\widehat{\boldsymbol{\rho}})$. Thus, the update (18) can be thought of as removing the components of $\widehat{\boldsymbol{\rho}}_n$ in the null-space and projecting the data to the signal subspace, which may be viewed as a *sophisticated denoiser* as shown in Fig. 2(a).

The MUSSELS scheme as represented in Fig. 2 provides state of the art results [5]. However, note that the filters specified by the columns of $\mathbf{Q}$ are estimated for each diffusion direction by alternating between (8), (18), and (11). The computational complexity of the structured low-rank algorithm is high, especially in the context of diffusion-weighted imaging where several directions need to be estimated for each slice.

C. MoDL-MUSSELS Formulation

To minimize the computational complexity of MUSSELS, we propose to learn a non-linear *denoiser* from exemplar data rather than learning a custom denoising block specified by $\left[\mathbf{I} - \frac{\lambda}{\beta}\mathbf{G}\left(\mathbf{Q}\right)^H\mathbf{G}\left(\mathbf{Q}\right)\right]$ for each direction and slice. We hypothesize that the non-linearities in the network as well as the larger number of filter layers can facilitate the learning of a generalizable model from exemplar data. This framework may be viewed as a multi-channel extension of MoDL [6]. The cost function associated with the network is

$$\arg\min_{\boldsymbol{\rho}} \|\mathcal{A}(\boldsymbol{\rho}) - \mathbf{y}\|_2^2 + \lambda_1\|\mathcal{N}_k(\boldsymbol{\rho})\|_2^2 \tag{19}$$

Here, $\mathcal{N}_k(\boldsymbol{\rho})$ is a non-linear residual convolutional filterbank working in the Fourier domain, with

$$\mathcal{N}_k(\boldsymbol{\rho}) = \boldsymbol{\rho} - \mathcal{D}_k(\boldsymbol{\rho}) \tag{20}$$

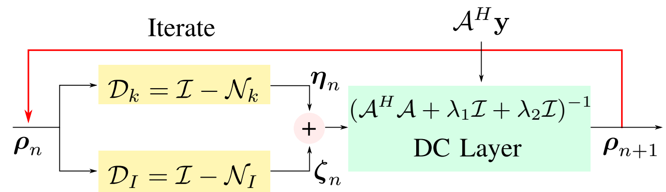


Fig. 4. The proposed hybrid MoDL-MUSSELS architecture architecture resulting from alternating scheme shown in in (22)-(24). Here $\mathcal{D}_k$ and $\mathcal{D}_I$ blocks represents k-space and image-space denoising networks respectively. Both the $\mathcal{D}_k$ and $\mathcal{D}_I$ networks have identical structure as in Fig 3(a). The learnable convolution weights are differnt for both the networks $\mathcal{D}_k$ and $\mathcal{D}_I$ but remains constant across iterations.

$\mathcal{D}_k(\boldsymbol{\rho})$ can be thought of as a multichannel CNN in the Fourier domain such that the image domain input $\boldsymbol{\rho}$ is first transformed to k-space as $\widehat{\boldsymbol{\rho}}$ then passes through the k-space model and then transformed back to image domain. Figure 3(a) shows the proposed M-layer CNN architecture. The overall k-space MoDL-MUSSELS network archtecture is shown in Fig. 3(b) that solves Eq. (19). Unlike MUSSELS in Fig. 2, the parameters of this network are not updated within the iterations and is learned from exemplar data.

D. Hybrid MoDL-MUSSELS Regularization

A key benefit of the MoDL framework over direct inversion methods is the ability to exploit different kinds of priors, as shown in our prior work [10]. The MUSSELS and the MoDL-MUSSELS scheme exploits the multichannel convolution relations between the k-space data. By contrast, we relied on an image domain convolutional neural network in [6] to exploit the structure of patches in the image domain. Note that this structure is completely complementary to the multichannel convolution relations. We now propose to jointly exploit both the priors as follows:

$$\arg\min_{\boldsymbol{\rho}} \|\mathcal{A}(\boldsymbol{\rho}) - \mathbf{y}\|_2^2 + \lambda_1\|\mathcal{N}_k(\boldsymbol{\rho})\|_2^2 + \lambda_2\|\mathcal{N}_I(\boldsymbol{\rho})\|_2^2, \tag{21}$$

here, $\mathcal{N}_k$ is the same prior as in (19), while $\mathcal{N}_I$ is an image space residual network of the form $\mathcal{N}_I(\boldsymbol{\rho}) = \boldsymbol{\rho} - \mathcal{D}_I(\boldsymbol{\rho})$. Here, $\mathcal{D}_I$ is a image domain CNN as in [6]. The problem (21) can be rewritten as

$$\arg\min_{\boldsymbol{\rho}} \|\mathcal{A}(\boldsymbol{\rho}) - \mathbf{y}\|_2^2 + \lambda_1\|\boldsymbol{\rho} - \mathcal{D}_k(\boldsymbol{\rho})\|_2^2 + \lambda_2\|\boldsymbol{\rho} - \mathcal{D}_I(\boldsymbol{\rho})\|_2^2.$$

By substituting $\boldsymbol{\eta} = \mathcal{D}_k(\boldsymbol{\rho})$, and $\boldsymbol{\zeta} = \mathcal{D}_I(\boldsymbol{\rho})$, an alternating minimization based solution to the above problem iterates between following steps:

$$\boldsymbol{\rho}_{n+1} = (\mathcal{A}^H\mathcal{A} + \lambda_1\mathcal{I} + \lambda_2\mathcal{I})^{-1}(\mathcal{A}^H\mathbf{y} + \lambda_1\boldsymbol{\eta} + \lambda_2\boldsymbol{\zeta}) \tag{22}$$

$$\boldsymbol{\eta}_{n+1} = \mathcal{D}_k(\boldsymbol{\rho}_{n+1}) \tag{23}$$

$$\boldsymbol{\zeta}_{n+1} = \mathcal{D}_I(\boldsymbol{\rho}_{n+1}). \tag{24}$$

The above solution results in the hybrid MoDL-MUSSELS architecture shown in Fig. 4. Note that this alternating minimization scheme is similar to the plug-and-play priors [20] widely used in inverse problems. The main exception is that we train the resulting network in an end-to-end fashion. Note that unlike the plug-and-plug denoisers that learn the image manifold, the network $\mathcal{D}_k$ is designed to exploit the redundancies between the multiple shots resulting from the phase relations. This non-linear network are expected to project the multichannel k-space data orthogonal to the null-spaces of the multichannel Hankel matrices.

IV. Experiments

We perform several experiments to validate different aspects of the proposed model such as benefits of the recursive network, impact of regularization, comparison with existing deep learning model such as U-NET [21], and comparison with recent a model-based technique termed as MUSE [17].

A. Dataset Description

In-vivo data were collected from healthy volunteers at the University of Iowa in accordance with the Institutional Review Board recommendations. The imaging was performed on a GE MR750W 3T scanner using a 32-channel head coil. A Stejskal-Tanner spin-echo diffusion imaging sequence was used with a 4-shot EPI readout. A total of 60 diffusion gradient direction measurements were performed with a b-value of 700 s/mm^2. The relevant imaging parameters were FOV= 210×210 mm, matrix size = 256×152 with partial Fourier oversampling of 24 lines, slice thickness= 4 mm and TE = 84 ms. Data were collected from 7 subjects. The acquisition was repeated twice for 5 subjects while two subjects had only one set of measurements.

The training dataset constituted a total of 68 slices, each having 60 directions and 4-shots, from 5 subjects. The validation was performed on 6 slices of the 6th subject whereas testing was carried out on 5 slices of the 7th subject. Thus, a total of 4080, 360, and 300 complex images each having size $256 \times 256 \times 4$ (rows $\times$ columns $\times$ shots) were used for training, validation, and testing respectively.

B. Multichannel forward model

All of the model based schemes used in this study (MUSE, MUSSELS, MoDL-MUSSELS) rely on a forward model that mimics the image formation. We implement this forward model as described in (1) and (3). The raw dataset consists of 32-channels. We reduce the data to four virtual channels using singular value decomposition (SVD) of the non-diffusion weighted (b0) image. The coil sensitivity maps of these four virtual channels are estimated using ESPIRIT [22]. The same channel combination weights are used to reduce the diffusion weighted MRI data to four coils.

C. Network architecture and training

In this work, we trained a 8-layer CNN having convolution filters of size 3×3 in each layer. Each layer comprises of a convolution, followed by ReLU, except the last layer which consists of 1×1 convolution as shown in Fig. 5. The real and imaginary components of the complex 4-shots data were considered as channels in the residual learning CNN architecure whereas the data-consistency block worked explicitly with complex data.

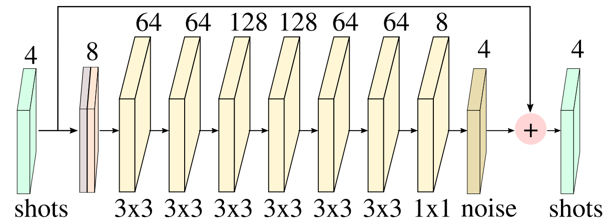


Fig. 5. The specific M=7 layer residual leanring convolutional neural network (CNN) architecture used as $\mathcal{D}_k$ and $\mathcal{D}_I$ blocks in the experiments. The 4-shot complex data is the input and output of the network. The first layer concatenates the real and imaginary part as 8 input features. The numbers on top of each layer represents the number of feature maps learned at that layer. We learn 3×3 filters at each layer except the last where we learn 1×1 filter.

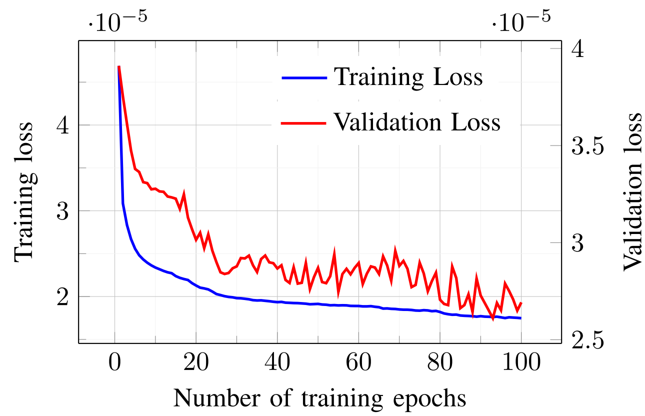


Fig. 6. The decay of training and validation errors with epochs. Each epoch represents one sweep through entire dataset. We note that both the losses decay with iterations. This suggests that the amount of training data is sufficient to train the parameters of the model. Our previous work [6] suggests that the re-use of the network weights across iterations significantly reduces the training data demand.

The proposed network architecture, as shown in Fig. 4, was unfolded for 3 iterations and the end-to-end training was performed for 100 epochs. The input to the unfolded network is the zero-filled complex data from the four shots, which corresponds to $A^H\mathbf{y}$, while the network outputs the fully sampled complex data for the four shots. The proposed MoDL-MUSSELS architecture combines the data from the four shots using sum-of-squares approach. The network weights were randomly initialized using Xavier initialization and shared between iterations. The network was implemented using TensorFlow library in Python 3.6 and trained using NVIDIA P100 GPU. The conjugate-gradient optimization in DC step was implemented as a layer operation in TensorFlow library as described in [6]. The total network training time of the network was around 37 hours.

The plot in Fig. 6 shows training loss decays smoothly with epochs. It can be noted that the loss on the validation dataset also has overall decaying behavior, which implies that the trained model did not over-fit the dataset. The model-based framework has considerably fewer parameters than direct inverse methods and hence requires far fewer training data to achieve good performance, as seen from the experiments in [6].

TABLE I
THE AVERAGE PSNR (DB) VALUES OBTAINED ON THE REAL TEST DATASET WITH K-SPACE ALONE MODEL AND HYBRID MODEL.

	$A^H b$	K-space model	Hybrid model
PSNR	18.89	24.31	27.96

D. *Quantitative metrics used for comparison*

The reconstruction quality is measured using peak signal-to-noise ratio (PSNR).

$$\mathrm{PSNR}(\mathbf{x}, \mathbf{y}) = 10 * \log_{10}\left(\frac{\max(\mathbf{x})^2}{\mathrm{MSE}(\mathbf{x}, \mathbf{y})}\right)$$

where MSE is the mean-square-error between $\mathbf{x}$ and $\mathbf{y}$. The final PSNR value is estimated by the average of the PSNR of individual shots.

E. *Algorithms used for comparison*

We compare the proposed scheme against MUSSELS [5], MUSE [17], and a UNET based solution [21] . The MUSSELS scheme iteratively solves (8), (12), and (11). MUSE [17] is a two-step algorithm which first estimates the motion induced phase using SENSE [23] reconstruction and total-variation denoising. With the knowledge of the phase errors, it recovers the images using a regularized optimization using (3) as the forward model. We extend the U-NET [21] model for the multi-shot diffusion MR image reconstruction, which is the extension of [24] to the multishot setting. The number of convolution layers, feature maps in each layer, and filter size are kept the same as in [21]. This k-space based formulation is similar to the one used in [24]. The input to the extended U-NET model is the 8-channel data obtained by concatenating the real and the imaginary parts of phase corrupted complex 4-shots. The MUSSELS [5] reconstruction were used as the ground truth for the training of the deep-learning models. We trained the network in image domain with 1000 epochs for 13 hours using ADAM [25] optimizer.

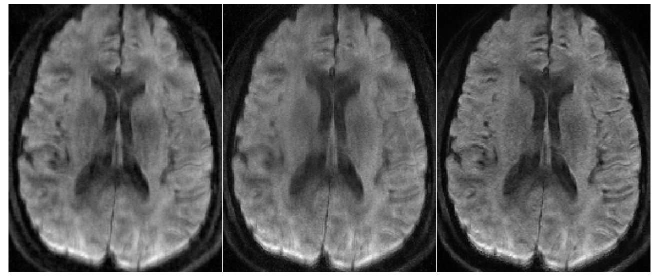
F. *Validation using numerical simulations*

To perform quantitative comparisons, we simulate the image formation numerically using the forward model in Eq. (3). In particular, we multiply one of the recovered shot images from the MUSSELS reconstruction with synthetically generated random bandlimited phase errors to generate the multishot data with the same undersampling patterns as in the real experiments. Gaussian noise of varying amount of standard deviation σ was added to the phase corrupted images.

V. RESULTS

A. *Benefit of multiple regularization priors*

We study the ability of the k-space network in minimizing the phase errors in Fig. 7(a). The experiments show the strength of the k-space network to compensate for phase errors, in comparison to the uncorrected combination of multishot data. Note that the results agree visually with the MUSE



(a) MUSE, 24.48 dB (b) K-space, 25.7 dB (c) Hybrid, 28.51 dB

Fig. 7. Comparison of k-space and hybrid models on testing data. (b) only the k-space model was used (b) the hybrid model. The numbers in the sub-captions are showing PSNR (dB) and SSIM values respectively. The k-space alone network is only designed to exploit the phase relation between the different shots. The results show the utility of the hybrid network, which also includes the image domain network, which provides additional regularization.

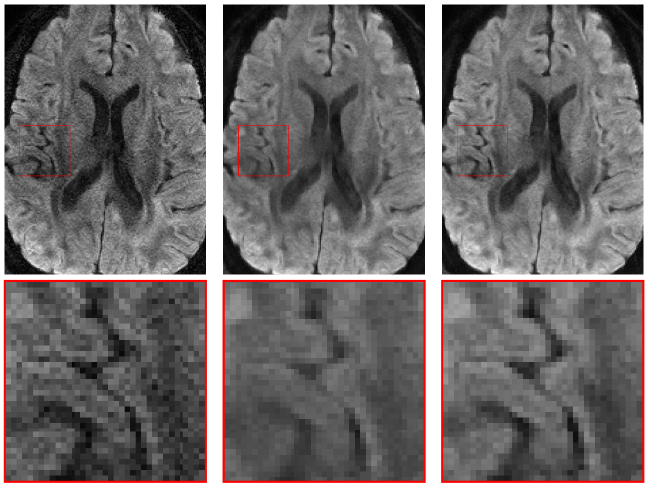
TABLE II
TESTING TIME TO RECONSTRUCT ALL 5 SLICES OF THE TEST SUBJECT. EACH SLICE HAD 60 DIRECTIONS AND 4-SHOTS. MUSSELS WAS RUNS ON CPU WITH PARALLEL PROCESSING.

Algorithm:	U-NET	MUSE	MUSSELS	MoDL-MUSSELS
Time (sec) :	7	110	1150	47

reconstructions. We study the ability of the hybrid network, with the addition of image domain regularization as in Fig. 4, to further improve the reconstructions in Fig. 7(c). We note that the image domain network exploits the manifold structure of patches, which serves as a strong prior which the k-space network has difficulty capturing. Note that the data was acquired using partial Fourier acceleration, where one side of k-space data was not acquired. The limited ability of the local k-space network is the reason for the blurring of the images. The additional image domain prior brought in by the hybrid scheme hence can reduce the blurring. Both the deep learned networks (k-space alone as well as the hybrid) network were trained independently. The quantitative comparison of the methods is shown in Table I, where we report the average PSNR and SSIM values obtained on the testing dataset by the two models. On average, the hybrid scheme offers more than 2 dB improvement in average PSNR values, which agrees with the visual improvement seen in Fig. 7.

B. *Impact of iterations on image quality*

In Fig 8, we study the impact of the number of iterations in the iterative algorithm described in (22)-(24). Specifically, we unroll the iterative algorithm for different number of iterations and compare the performance of the resulting networks. We use the hybrid model due to its improved performance as seen from the previous paragraph. The parameters of both the k-space and image-space networks are assumed to be constants with iterations; they are shared across iterations. The images in Fig. 8 correspond to a specific direction and slice in the testing dataset. We note that the contrast and details in the image improve with iterations, and improved visualization of some features as shown by zoomed portions.



(a) MUSSELS (b) One iteration PSNR=24.84 dB (c) Three iteration PSNR=27.00 dB

Fig. 8. Effect of iterations on image quality. We observe that the quality of the reconstructions with the proposed MoDL-MUSSELS scheme improve with iterations. Specifically, the sharpness of the image and the contrast seem to improve with more iterations.

C. Comparison with existing methods on experimentatal data

Figure 9 shows reconstruction quality on a testing dataset for two of the four shots a particular slice and direction. The proposed algorithm provides reconstructions with qualitatively similar magnitude and phase components. It is noted that the alias artefacts are significantly reduced compared to the uncorrected data.

In Fig. 10, we compare the proposed method with a deep learning based method named U-NET [21], as well as traditional model-based methods named as MUSE [17] and MUSSELS. We empirically find the best parameters of MUSE as $\lambda_1 = 1.5, \lambda_2 = 0.05,$ iter $= 50$. We extended and implemented the U-NET model for the multi-shot diffusion MRI as described above in Section IV (E). Figure 10 shows the reconstructions offered by the different algorithms. Note that the proposed scheme used MUSSELS as the ground truth for training. Visually, the MUSE results in a comparatively blurred image than proposed MoDL-MUSSELS scheme and MUSSELS. The UNET reconstructions appear less blurred, but it seems to miss some key features highlighted by boxes and arrows.

To further validate the reconstruction accuracy of all the DWIs corresponding to the test slice, we performed a tensor fitting using all the DWIs and compared the resulting fractional anisotropy (FA) maps and the fiber orientation maps. For this purpose, the DWIs reconstructed using various methods from the test dataset were fed to a tensor fitting routine (FDT Toolbox, FSL). FA maps were computed from the fitted tensors and the direction of the primary eigenvectors of the tensors was used to estimate the fiber orientation. The FA maps generated using the various reconstruction methods are shown in Fig. 11, which has been color-coded based on the fiber direction. It is noted that these fiber directions reconstructed by the MUSSELS method and the MoDL-MUSSELS match the true anatomy known for this brain region from a DTI white matter atlas[1].

Table II compares the time taken to reconstruct the entire testing dataset using the four compared methods. It is noted that the computational complexity of MoDL-MUSSELS is around 28 fold lower than MUSSELS. Note that MUSSELS estimates the optimal linear filter bank from the measurements itself, which requires significantly many iterations. By contrast, since the non-linear network is pre-learned, three alternations between the data consistency step and the projections provided by the deep learned network is sufficient for the proposed scheme to yield good recovery; the quite significant speedup follows directly from the significantly fewer number of iterations. Note that we rely on a conjugate gradient algorithm to enforce data consistency specified by (22). Note that solving (22) exactly as opposed to the use of steepest gradient steps at each iteration would require more unrolling steps, thus diminishing the gain in speedup. The greatly reduced runtime is expected to facilitate the deployment of the proposed algorithm on clinical scanners.

TABLE III
THE PSNR (DB) VALUES OBTAINED BY FOUR METHODS ON THE TESTING DATASET WITH SIMULATED PHASEES OF DIFFERENT BANDWIDTHS AND ADDED GAUSSIAN NOISE OF VARYING STANDARD DEVIATION σ.

Bandwidth, σ	U-NET	MUSE	MUSSELS	Proposed
3x3,0.00	23.24	27.68	27.99	31.13
5x5,0.00	22.5	27.32	27.45	30.63
7x7,0.00	21.81	26.44	26.93	30.11
3x3,.001	23.23	27.51	29.15	31.02
5x5,.001	22.5	27.17	28.25	30.68
7x7,.001	21.81	26.31	27.47	30.17
3x3,.003	22.97	26.29	28.51	29.36
5x5,.003	22.31	25.69	27.92	29.18
7x7,.003	21.67	24.97	27.36	28.89

D. Quantitative comparisons using simulated data

Table III summarizes the quantitative results from the simulated data in Section IV-F. Specifically, we quantitatively compare the reconstructions provided by the four algorithms, while varying the noise levels and bandwidths for phases; larger bandwidths imply larger spatial variations in the phase errors corresponding to large motion. The deep learning methods (UNET or MoDL-MUSSELS), which were trained using the true MUSSELS reconstructions, were used on the simulated experiments; no re-training was performed with synthetic data. Figure 12 shows the visual comparisons of the four methods in the less challenging (low noise and low bandwidth) and most challenging (high-noise and high bandwidth) settings.

VI. DISCUSSION

The proposed model-based deep learning method was trained using reconstructions obtained using the MUSSELS scheme. Once the training is completed, the testing was performed on phase corrupted images that were not included

[1]http://www.dtiatlas.org

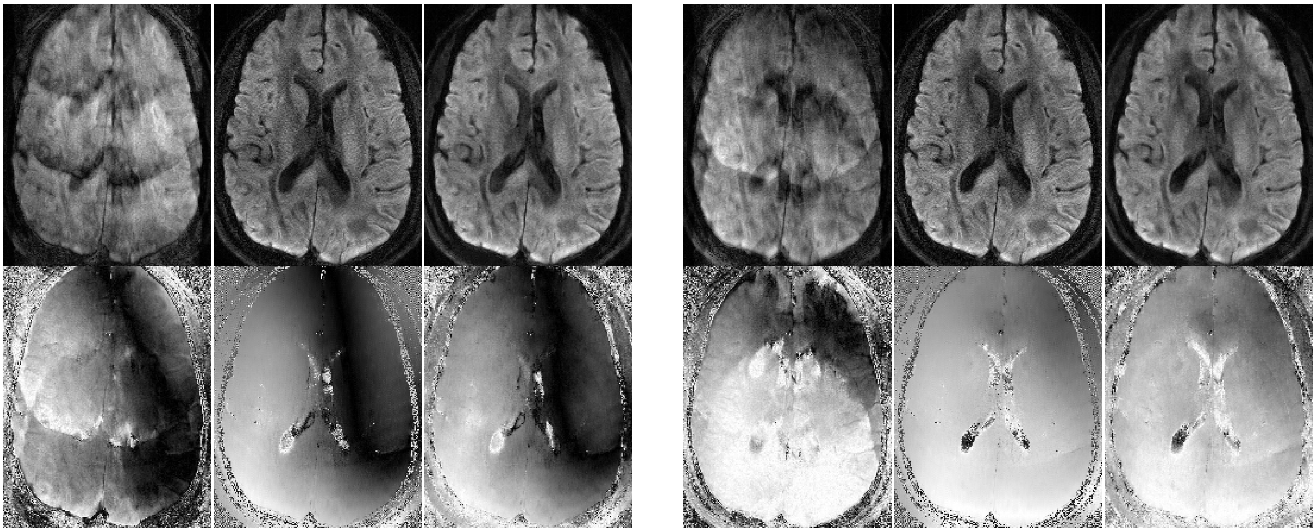


Fig. 9. Magnitude and phase images recovered by MUSSELS and MoDL-MUSSELS. The inputs to both the networks are the zero-filled shot images from the four shot acquisition, while the outputs are the recovered shot images. These images are then combined using sum of squares reconstruction. (a) and (b) correspond to shot 2 and shot 3, respectively. Note that the magnitude of the recovered magnitude images of the shots are roughly similar, while the phases are very different.

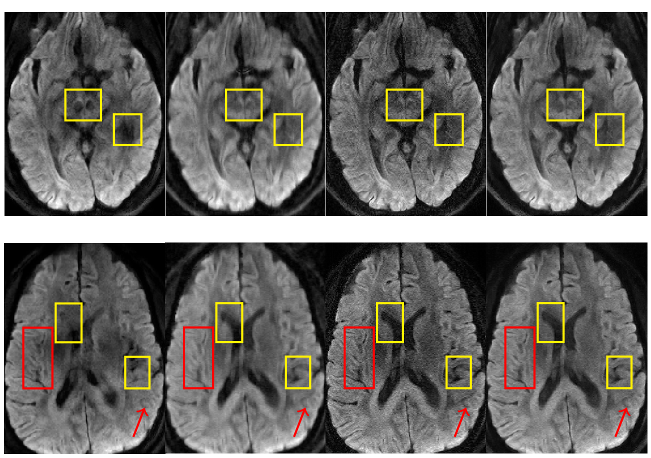


Fig. 10. Reconstructions obtained using different algorithms. The columns correspond to the reconstructions using UNET, MUSE, MUSSELS, and MoDL-MUSSELS, respectively. The rows correspond to two of the diffusion directions from two different slices. The red and yellow boxes highlights the differences.

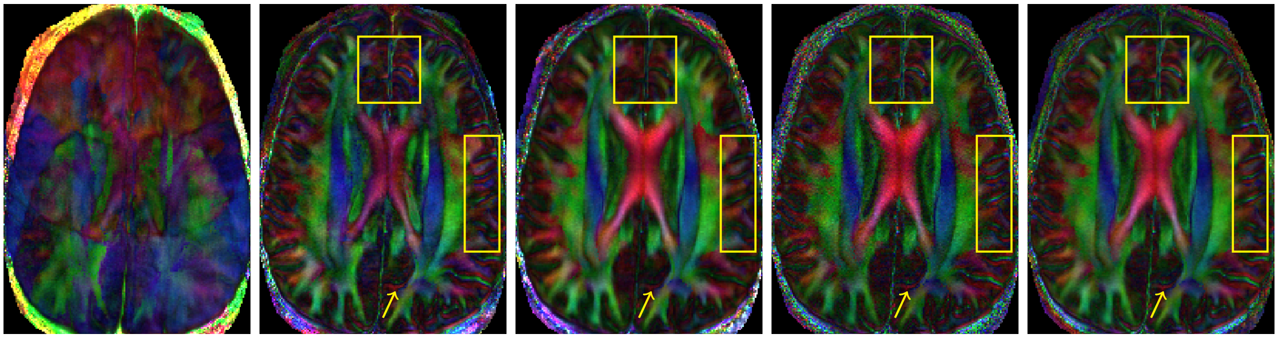
in the training data. We note that the MUSSELS scheme does not use any spatial regularization; the recovered images are not completely free in the inner brain regions, where the diversity of the coils are not high. We note from Fig. 10 that the reconstructions provided by MoDL-MUSSELS appear less noisy and visually more appealing compared to MUSSELS, even though similar noisy images recovered using MUSSELS were used for training. This behavior may be attributed to the convolutional structure of the network, which is known to offer implicit regularization [26].

In this work, we utilized a 8-layer neural network as shown in Fig. 5. However, the proposed MoDL-MUSSELS architecture in Fig. 4 method is not constrained by the choice of the network. Any network architecture (e.g. U-NET) may be used instead. It is possible that the results can improve by utilizing more sophisticated network architecture. Further, it can be noted that the proposed model architecture is flexible to allow different network architectures for image-space and k-space models. However, for the proof of concept, we used the same network architecture for both k-space and image-space. The results can further improve by incorporating more complex image-domain network architecture.

To avoid overfitting the model and reduce the training time, the proposed network in Fig. 4 was unfolded for 3 iterations before performing the joint training. The sharing of network parameters allow the network to be unfolded for any number of iterations without increasing the number of trainable parameters.

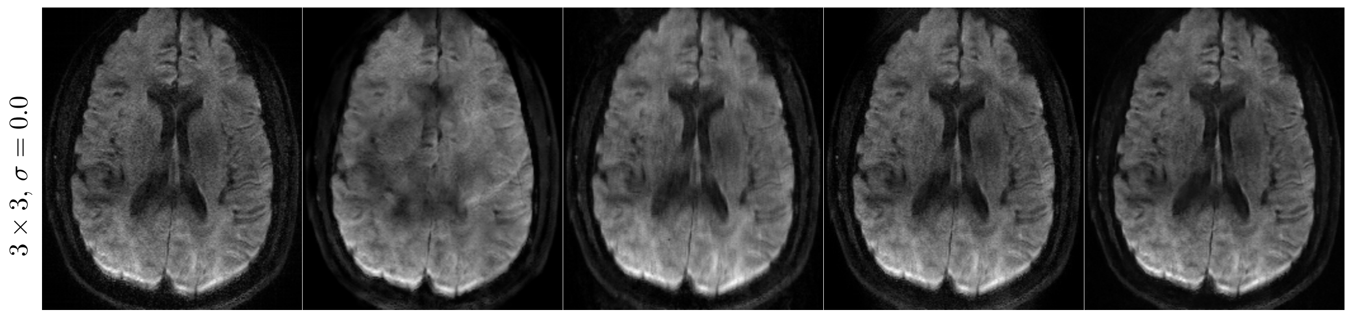
VII. Conclusions

We introduced a model based deep learning framework termed MoDL-MUSSELS for the compensation of phase errors in multishot diffusion-weighted MRI data. The proposed algorithm alternates between a conjugate gradient optimization algorithm to enforce data consistency and multichannel convolutional neural networks (CNN) to project the data to appropriate subspaces. We rely on a hybrid approach involving a multichannel CNN in k-space and another one in image space. The k-space CNN exploits the phase relations between the shot images, while the image domain network is used to project the data to an image manifold. The weights of the deep network, obtained by unrolling the iterations in the iterative optimization scheme, are learned from exemplary data in an end-to-end fashion. The experiments show that the proposed scheme can yield reconstructions that are comparable to state

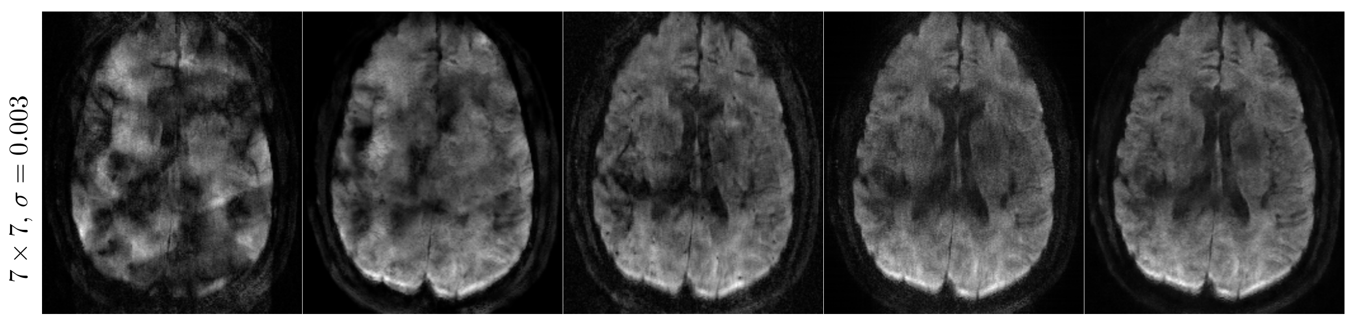


(a) Uncorrected (b) UNET (c) MUSE (d) MUSSELS (e) MoDL-MUSSELS

Fig. 11. The fractional-anisotropy maps of the slice corresponding to the second row of Fig. 10. These images are computed from the sixty directions of the slices, recovered using the respective algorithms. We note the the proposed scheme provide less blurred reconstructions than MUSE, which are comparable with MUSSELS.



(a) Ground Truth (b) U-NET, 23.08 (c) MUSE, 27.92 (d) MUSSELS, 30.27 (e) MoDL-MUSSELS, 31.73



(f) Simulated $A^H b$, 17.65 (g) U-NET, 22.22 (h) MUSE, 25.49 (i) MUSSELS, 26.80 (j) MoDL-MUSSELS, 28.01

Fig. 12. Simulation results: the least challenging case with phase errors of bandwidth 3x3 and noise standard deviation $\sigma = 0$ are on the top row, while the most challenging setting (7x7 phase errors and $\sigma = .003$). The quantitative results are shown in Table III. These results show that the proposed scheme provide the most accurate results at all parameter settings.

of the art methods, while offering several orders of magnitude reduction in run-time.

References

[1] D. Le Bihan, “Looking into the functional architecture of the brain with diffusion mri,” *Nature Reviews Neuroscience*, vol. 4, no. 6, p. 469, 2003.

[2] O. Ciccarelli, M. Catani, H. Johansen-Berg, C. Clark, and A. Thompson, “Diffusion-based tractography in neurological disorders: concepts, applications, and future developments,” *The Lancet Neurology*, vol. 7, no. 8, pp. 715–727, 2008.

[3] R. A. Charlton, T. Barrick, D. McIntyre, Y. Shen, M. O’sullivan, F. Howe, C. Clark, R. Morris, and H. Markus, “White matter damage on diffusion tensor imaging correlates with age-related cognitive decline,” *Neurology*, vol. 66, no. 2, pp. 217–222, 2006.

[4] M. Moseley, Y. Cohen, J. Mintorovitch, L. Chileuitt, H. Shimizu, J. Kucharczyk, M. Wendland, and P. Weinstein, “Early detection of regional cerebral ischemia in cats: comparison of diffusion-and t2-weighted mri and spectroscopy,” *Magnetic resonance in medicine*, vol. 14, no. 2, pp. 330–346, 1990.

[5] M. Mani, M. Jacob, D. Kelley, and V. Magnotta, “Multi-shot sensitivity-encoded diffusion data recovery using structured low-rank matrix completion (MUSSELS),” *Mag. Reson. in Med.*, vol. 78, no. 2, pp. 494–507, 2016.

[6] H. K. Aggarwal, M. P. Mani, and M. Jacob, “MoDL: Model based deep learning architecture for inverse problems,” *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, 2019.

[7] J. Adler and O. Öktem, “Learned primal-dual reconstruction,” *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, 2018.

[8] y. yang, J. Sun, H. Li, and Z. Xu, “Deep admm-net for compressive sensing mri,” in *Advances in Neural Information Processing Systems 29*, 2016, pp. 10–18.

[9] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, “Learning a Variational Network for Reconstruction of Accelerated MRI Data,” *Magnetic resonance in Medicine*, vol. 79, no. 6, pp. 3055–3071, 2017.

[10] S. Biswas, H. K. Aggarwal, and M. Jacob, "Dynamic mri using model-based deep learning and storm priors: MoDL-SToRM," *Magnetic Resonance in Medicine*, vol. 0, no. 0, 2019.

[11] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2008, pp. 3869–3872.

[12] J. Lee, Y. Han, and J. C. Ye, "k-space deep learning for reference-free epi ghost correction," *arXiv preprint arXiv:1806.00153*, 2018.

[13] R. Souza and R. Frayne, "A hybrid frequency-domain/image-domain deep network for magnetic resonance image reconstruction," *arXiv preprint arXiv:1810.12473*, 2018.

[14] S. Wang, Z. Ke, H. Cheng, S. Jia, Y. Leslie, H. Zheng, and D. Liang, "Dimension: Dynamic mr imaging with both k-space and spatial prior knowledge obtained via multi-supervised network training," *arXiv preprint arXiv:1810.00302*, 2018.

[15] T. Eo, Y. Jun, T. Kim, J. Jang, H.-J. Lee, and D. Hwang, "Kiki-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images," *Magnetic resonance in medicine*, vol. 80, no. 5, pp. 2188–2201, 2018.

[16] M. Akcakaya, S. Moeller, S. Weingartner, and K. Ugurbil, "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magnetic Resonance in Medicine*, vol. 81, no. April 2018, pp. 439–453, 2018.

[17] N.-k. Chen, A. Guidon, H.-C. Chang, and A. W. Song, "A robust multi-shot scan strategy for high-resolution diffusion weighted mri enabled by multiplexed sensitivity-encoding (muse)," *Neuroimage*, vol. 72, pp. 41–47, 2013.

[18] R. L. Morrison, M. Jacob, and M. N. Do, "Multichannel estimation of coil sensitivities in parallel mri," in *IEEE Int. Symp. Bio. Imag.* IEEE, 2007, pp. 117–120.

[19] G. Ongie, S. Biswas, and M. Jacob, "Convex recovery of continuous domain piecewise constant images from non-uniform Fourier samples," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 236–250, 2017.

[20] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-Play ADMM for Image Restoration: Fixed Point Convergence and Applications," *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 84–98, 2017.

[21] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).* Springer, 2015, pp. 234–241.

[22] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT - An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magnetic Resonance in Medicine*, vol. 71, no. 3, pp. 990–1001, 2014.

[23] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "Sense: sensitivity encoding for fast mri," *Magnetic resonance in medicine*, vol. 42, no. 5, pp. 952–962, 1999.

[24] Y. Han and J. C. Ye, "k-space deep learning for accelerated mri," in *arXiv:1805.03779*, 2018.

[25] D. P. Kingma and J. L. Ba, "Adam: a Method for Stochastic Optimization," *International Conference on Learning Representations 2015*, pp. 1–15, 2015.

[26] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9446–9454.