

**RESEARCH ARTICLE**

# High-Resolution and Motion-Robust Diffusion-Weighted Imaging with Self-Gated Self-Supervised Unrolled Reconstruction

Zhengguo Tan<sup>1</sup> | Patrick Liebig<sup>2</sup> | Annika Hofmann<sup>3</sup> | Frederik B Laun<sup>4</sup> | Florian Knoll<sup>3</sup>

<sup>1</sup>Michigan Institute for Imaging Technology and Translation (MIITT), Department of Radiology, University of Michigan, Ann Arbor, Michigan, USA

<sup>2</sup>Ultra-High Field Predevelopment Team, Siemens Healthcare GmbH, Erlangen, Bavaria, Germany

<sup>3</sup>Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Bavaria, Germany

<sup>4</sup>Institute of Radiology, University Hospital Erlangen, Erlangen, Bavaria, Germany

**Correspondence**

Zhengguo Tan, Room 3111, Medical Science Unit I, 1301 Catherine St, Ann Arbor, MI 48109. Email: zgtan@med.umich.edu

**Funding Information**

German Research Foundation (DFG); projects 513220538, 512819079, project 500888779 in the Research Unit RU5534 for MR biosignatures at UHF. National Institutes of Health (NIH); grants R01EB024532 and P41EB017183.

**Abstract**

**Purpose:** High-resolution and motion-robust diffusion-weighted imaging (DWI) is clinically demanding. The purpose of this work is to develop an efficient self-supervised algorithm unrolling technique for submillimeter-resolution and motion-robust DWI.

**Methods:** We developed submillimeter DWI acquisition utilizing multi-band multi-shot EPI with diffusion shift encoding. We unrolled the alternating direction method of multipliers (ADMM) to perform scan-specific self-supervised DeepDWI learning for multi-shot echo planar imaging with diffusion shift encoding on a clinical 7T scanner.

**Results:** We demonstrate that (1) ADMM unrolling is generalizable across slices, (2) ADMM unrolling outperforms multiplexed sensitivity-encoding (MUSE) and compressed sensing with locally-low rank (LLR) regularization in terms of image sharpness, tissue continuity, and motion robustness, (3) ADMM unrolling enables clinically feasible inference time.

**Conclusion:** Our proposed ADMM unrolling enables whole brain DWI of 21 volumes at 0.7 mm isotropic resolution and 10 minutes scan, and shows higher signal-to-noise ratio (SNR), clearer tissue delineation, and improved motion robustness, which makes it plausible for clinical translation.

**KEYWORDS:**

diffusion weighted imaging, submillimeter resolution, image reconstruction, machine learning, self-supervised learning, algorithm unrolling

**WORD COUNT:** 3708

## 1 | INTRODUCTION

Diffusion-weighted imaging (DWI)<sup>1</sup> has been an important imaging modality in neuro-scientific research and clinical diagnosis and staging of tumors. However, clinical DWI, based on single-shot echo planar imaging (EPI)<sup>2</sup>,

poses challenges in the pursuit of high spatial, temporal, and angular resolution. Until now, the search for precise neuro imaging has fostered significant advances in DWI, including multi-shot EPI (interleaved<sup>3,4,5</sup> and readout-segmented<sup>6,7</sup>), field inhomogeneity and eddy current correction<sup>8</sup>, simultaneous multi-slice<sup>9</sup>, reconstruction techniques such as parallel imaging<sup>10,11,12</sup> and

compressed sensing<sup>13,14,15</sup>, as well as diffusion-weighted image denoising<sup>16,17</sup>.

To achieve submillimeter isotropic resolution DWI, advanced  $k$ -space encoding strategies have been proposed. For example, Setsompop et al.<sup>18</sup> developed generalized slice dithered enhanced resolution (gSlider) to boost SNR per slice. gSlider excites one thick slab multiple times with complementary slice encoding schemes (e.g., Hadamard encoding). The thin slices are then reconstructed by solving a linear least square problem given the complementary slab signal. However, it has been reported that gSlider has stricter requirements on  $B_0$  and  $B_1$  field homogeneities and shows residual slab boundary artifacts (appearing as striping artifacts)<sup>19</sup>. Another advanced  $k$ -space encoding strategy, rotating-view echo planar time-resolved imaging (Romer-EPTI)<sup>20</sup>, has recently been developed. Romer-EPTI acquires thick-slice volumes with different slice orientations. In addition, a rigid-motion transformation matrix is extracted from the multi-shot data and incorporated into the super-resolution linear least square problem to disentangle motion-free thin-slice images. Together with low-rank subspace modeling and reconstruction<sup>21,22,23</sup>, Romer-EPI achieves submillimeter TE-specific distortion-free DWI. However, Romer-EPTI requires relatively long scan time.

Advanced image reconstruction techniques have also been proposed to push the boundary of DWI. Mani et al.<sup>14</sup> developed MUSSELS, in which structural low rankness is enforced in  $k$ -space to reconstruct multi-shot images. After reconstruction, one diffusion-weighted image is calculated via root sum of square of all shot images. MUSSELS bypasses the shot-to-shot phase variation correction. The concept of structural low rank has been explored in the regime of deep learning MRI reconstruction<sup>24,25</sup>. Cho et al.<sup>26</sup> proposed to learn a  $k$ -space regularization function with ResNet<sup>27</sup> and zero-shot self-supervised learning<sup>28</sup>. This approach learns an unrolled algorithm utilizing only the acquired data itself, and thus requires no extra training data and is scan specific. Similar to MUSSELS, this approach addresses the multi-shot EPI reconstruction problem for a single diffusion encoding, and does not perform joint reconstruction that explores  $q$ -space redundancy. Supervised learning has also been proposed for DWI reconstruction. Mani et al.<sup>29</sup> proposed to learn a denoising autoencoder (DAE) model<sup>30</sup> from a physics-informed simulated dictionary. This learned DAE model is subsequently incorporated as a regularizer in the alternating direction method of multipliers (ADMM)<sup>31</sup> unrolling reconstruction. Although appealing, this approach is model specific and requires the construction of a large-scale dictionary.

To address the challenges of achieving submillimeter isotropic DWI at clinically feasible scan times, we propose a novel solution that leverages an ADMM unrolling method with self-supervised learning for multi-band, multi-shot DWI with diffusion shift encoding. Unlike previous methods that require extensive phase variation correction or long scan times (e.g., gSlider's strict field homogeneity requirements or Romer-EPTI's lengthy acquisition times), our approach incorporates complementary  $k$ - $q$ -space sampling and jointly reconstructs multiple slices and all diffusion-weighted images. Moreover, by training the ADMM unrolling model using a single slice, we enable self-gated joint reconstruction, significantly reduce training time, and avoid the need for large-scale dictionaries or extra training data, as seen in methods like supervised learning approaches. Our method not only achieves high-resolution DWI at 0.7 mm isotropic resolution with 21 diffusion-encoding directions but also does so in under 10 minutes of scan time and approximately 1 minute of reconstruction time per slice. This provides a clinically viable, efficient solution to the submillimeter resolution DWI challenge.

## 2 | METHODS

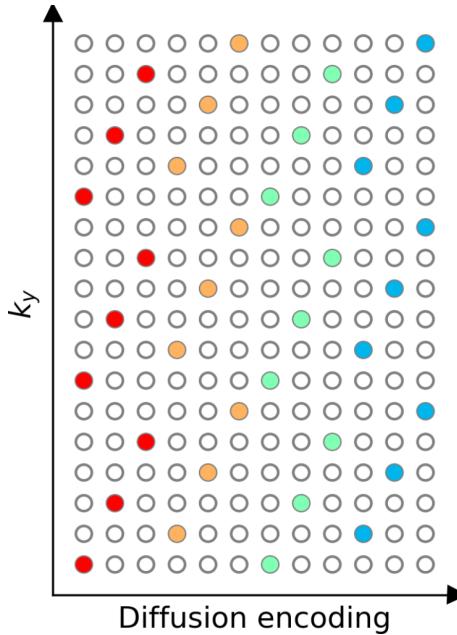
### 2.1 | Multi-Band Multi-Shot DWI with Diffusion-Shift Encoding

Our previous work<sup>32</sup> demonstrated the joint  $k$ - $q$ -slice reconstruction for multi-band multi-shot navigator-based interleaved EPI (NAViEPI) DWI acquisition with diffusion-shift encoding. As shown in Figure 1, the starting  $k_y$  line for a diffusion encoding is shifted with respect to its adjacent one to create complementary  $k$ - $q$ -slice sampling pattern. In the joint reconstruction, the forward model maps the multi-slice multi-diffusion-weighted images ( $\mathbf{x}$ ) to their corresponding  $k$ -space,

$$\mathcal{A}(\mathbf{x}) = \mathbf{P}\boldsymbol{\Sigma}\boldsymbol{\Theta}\mathbf{F}\mathbf{S}\Phi\mathbf{x} \quad (1)$$

Here, the images  $\mathbf{x}$  are point-wise multiplied with the pre-computed shot-to-shot phase variation maps ( $\Phi$ ) and coil sensitivity maps ( $\mathbf{S}$ ). The output images are then converted to  $k$ -space via two-dimensional fast Fourier transform ( $\mathbf{F}$ ), point-wise multiplied with the multi-band phases ( $\boldsymbol{\Theta}$ ), summed along the slice dimension ( $\boldsymbol{\Sigma}$ ), and then multiplied by the  $k$ -space undersampling mask ( $\mathbf{P}$ ).

In Equation (1), one challenge is to accurately estimate the shot-to-shot phase variation. Multiplexed sensitivity-encoding (MUSE) type reconstruction techniques<sup>4,33,5,34</sup> achieved the self-gating strategy, where the  $k$ -space data of each shot were used to reconstruct its



**FIGURE 1** Multi-shot DWI with diffusion shift encoding. This work employs 3-shot per diffusion encoding and each shot has an in-plane undersampling factor of 6. Therefore, every three columns have the same color. The starting  $k_y$  line is shifted between adjacent diffusion encoding to create complementary  $k$ - $q$ -space sampling.

corresponding shot image followed by a phase smoothing approach. Self-gated shot phase estimation does not require the acquisition of phase navigator data. However, it requires small undersampling factors per shot and fully-sampled DWI acquisition assembling all shots. Alternatively, undersampled DWI acquisition can be enabled via the acquisition of navigators for shot phase estimation<sup>32</sup>. This approach allows for mesoscale resolution DWI at 7T, but still requires a long scan time. As listed in Table 1, the total acquisition of Protocol #1 at 0.7 mm isotropic resolution takes 16 : 27 minutes with phase navigators. This scan time can be reduced to approximately 10 minutes by removing the phase navigators (Protocol #2 in Table 1).

With the operator  $\mathcal{A}$ , the joint reconstruction reads,

$$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathcal{A}(\mathbf{x})\|_2^2 + \lambda \mathcal{R}(\mathbf{x}) \quad (2)$$

where  $\mathbf{y}$  is the measured  $k$ -space data. The first term in Equation (2) presents the data-consistency term, and the second term presents the regularization function  $\mathcal{R}(x)$  with the regularization strength  $\lambda$ . When using the Tikhonov regularization, i.e.  $\mathcal{R}(\mathbf{x}) = \|\mathbf{x}\|_2^2$ , Equation (2) can be solved via the conjugate gradient (CG) method. For nonlinear regularization functions, such as the

locally-low rank (LLR) regularization<sup>32</sup> or neural networks with nonlinear activation functions. ADMM was employed in this work to solve for Equation (2).

## 2.2 | Image Reconstruction via Self-Supervised ADMM Unrolling

---

### Algorithm 1 Self-Supervised ADMM Unrolling

---

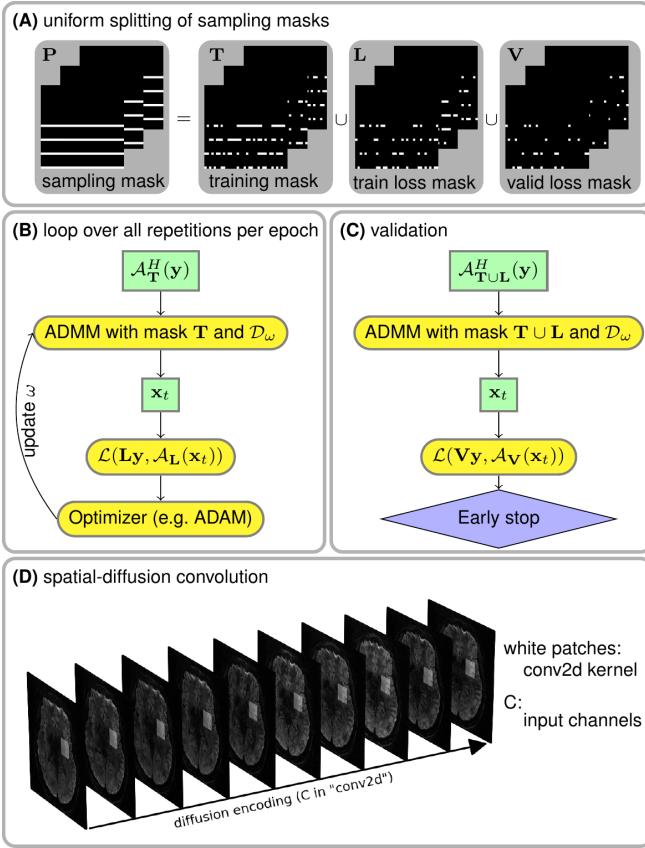
```

1: Initialization:
2:   split sampling mask  $\mathbf{P}$  into 12 repetitions, each of
   which consists of three disjoint sets  $\mathbf{T}$ ,  $\mathbf{L}$ , and  $\mathbf{V}$ 
3:    $p \leftarrow 0$  and  $N_{\text{epoch}} \leftarrow 100$ 
4:    $\mathcal{D}_\omega$  set as ResNet
5:    $\rho \leftarrow 0.05$  and  $\lambda \leftarrow 0.05$ 
6:    $\text{Loss}_{\text{valid}} \leftarrow \infty$  and  $\text{trace} \leftarrow 0$ 
7: function ADMM(mask)
8:    $\mathcal{A}_{\text{mask}} \leftarrow$  set the mask in the forward operator  $\mathcal{A}$ 
9:    $\mathbf{x}^{(0)} \leftarrow \mathcal{A}_{\text{mask}}^H(\mathbf{y})$ 
10:   $\mathbf{v}^{(0)} \leftarrow \mathbf{x}^{(0)}$  and  $\mathbf{u}^{(0)} \leftarrow \mathbf{0}$ 
11:   $k \leftarrow 0$  and  $N_{\text{unroll}} \leftarrow 8$ 
12:  while  $k < N_{\text{unroll}}$  do
13:     $\mathbf{x}^{(k+1)} \leftarrow$  conjugate gradient with 6 iterations
14:     $\mathbf{v}^{(k+1)} \leftarrow (\lambda/\rho) \cdot \mathcal{D}_\omega(\mathbf{x}^{(k+1)} + \mathbf{u}^{(k)})$ 
15:     $\mathbf{u}^{(k+1)} \leftarrow \mathbf{u}^{(k)} + \mathbf{x}^{(k+1)} - \mathbf{v}^{(k+1)}$ 
16:     $k \leftarrow k + 1$ 
17:  end while
18:  return  $\mathbf{x}^{(k+1)}$ 
19: end function
20: Training:
21: while  $p < N_{\text{epoch}}$  or  $\text{trace} \leq 12$  do
22:    $\mathbf{x}_t \leftarrow \text{ADMM}(\mathbf{T})$ 
23:    $\text{Loss}_{\text{train}} \leftarrow \mathcal{L}(\mathbf{Ly}, \mathcal{A}_{\mathbf{L}}(\mathbf{x}_t))$ 
24:   update  $\omega$  via ADAM
25: Validation:
26:    $\mathbf{x}_t \leftarrow \text{ADMM}(\mathbf{T} \cup \mathbf{L})$ 
27:    $\text{Loss}_{\text{temp}} \leftarrow \mathcal{L}(\mathbf{Vy}, \mathcal{A}_{\mathbf{V}}(\mathbf{x}_t))$ 
28:   if  $\text{Loss}_{\text{temp}} \leq \text{Loss}_{\text{valid}}$  then
29:      $\text{Loss}_{\text{valid}} \leftarrow \text{Loss}_{\text{temp}}$ 
30:      $\text{trace} \leftarrow 0$ 
31:   else
32:      $\text{trace} \leftarrow \text{trace} + 1$ 
33:   end if
34: end while

```

---

Instead of the two-step alternating minimization unrolling scheme as used in Modl<sup>25</sup>, we employed the ADMM unrolling to solve the self-supervised learning reconstruction in Equation (2). The update rule of



**FIGURE 2** Illustration of the key components in ADMM unrolling. **(A)** The sampling mask  $P$  in Equation (1) was uniformly split into three disjoint sets: the training mask  $\mathbf{T}$  used for the data consistency term during training, the train loss mask  $\mathbf{L}$  used for the loss function calculation during training, and the validation loss mask  $\mathbf{V}$  used for the loss function calculation during validation. **(B)** and **(C)** show the flowchart for the training and the validation of an unrolled ADMM model, respectively. Note that the ResNet parameters  $\omega$  are updated via ADAM<sup>35</sup> during training, but remain fixed during the validation step. **(D)** A stack of diffusion-weighted images is input into ResNet during ADMM unrolling.

ADMM unrolling reads

$$\begin{cases} \mathbf{x}^{(k+1)} = \underset{\mathbf{x}^{(k)}}{\operatorname{argmin}} \|\mathbf{y} - \mathcal{A}(\mathbf{x}^{(k)})\|_2^2 \\ \quad + \frac{\rho}{2} \|\mathbf{x}^{(k)} - \mathbf{v}^{(k)} + \mathbf{u}^{(k)}\|_2^2 \\ \mathbf{v}^{(k+1)} = (\lambda/\rho) \cdot \mathcal{D}_\omega(\mathbf{x}^{(k+1)} + \mathbf{u}^{(k)}) \\ \mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \mathbf{x}^{(k+1)} - \mathbf{v}^{(k+1)} \end{cases} \quad (3)$$

ADMM updates the variables  $\mathbf{x}$ ,  $\mathbf{v}$ , and  $\mathbf{u}$  in an alternating scheme. It splits the unrolled reconstruction into three steps, as shown in Equation (3) and in the pseudo code of Algorithm 1. First, the updating step for  $\mathbf{x}$  is solved by conjugate gradient. Second, the variable  $\mathbf{v}$  is then updated via the forward pass of the neural network

$\mathcal{D}_\omega$  with the input as the sum of current estimates of  $\mathbf{x}$  and  $\mathbf{u}$ . Third, the variable  $\mathbf{u}$  is updated by adding its current estimate to the difference between  $\mathbf{x}$  and  $\mathbf{v}$ .

As shown in Figure 2, the data sampling mask  $\mathbf{P}$  is split into three disjoint sets: the training mask  $\mathbf{T}$  for the data consistency term, the training loss mask  $\mathbf{L}$  for the loss function calculation, and the validation loss mask  $\mathbf{V}$ . Each set consists of 12 repetitions constructed via random uniform sampling of the data mask  $\mathbf{P}$ . In each training epoch, every repetition is looped through in order to update the ResNet parameters  $\omega$ . Plus, the validation step is performed after every training epoch to update the minimal validation loss. If the validation loss does not reduce for 12 consecutive epochs or if 100 epochs are reached, the training is terminated.

The index  $k$  in Equation (3) denotes the unrolling iteration, and  $\mathcal{D}_\omega$  denotes the ResNet<sup>27</sup> parameterized by  $\omega$ . In this work, 2D convolution was employed to construct the ResNet layers. In PyTorch, 2D convolution requires four-dimensional tensors as input and output. For instance, a matrix with the size  $(N, C, H, W)$  is acceptable for the 'conv2d' function in PyTorch. Here,  $W$  and  $H$  denote the width and height of the convolution kernel,  $C$  denotes the number of channels, and  $N$  denotes the batch size. However, the diffusion-weighted images ( $\mathbf{x}$ ) to be reconstructed have the size  $(N_{\text{diff}}, N_Z, N_Y, N_X, 2)$ , where 2 stands for the real and imaginary part of the complex-valued diffusion-weighted images,  $N_X$  and  $N_Y$  are the width and the height of diffusion-weighted images,  $N_Z$  is the number of slices (identical to the multi-band factor), and  $N_{\text{diff}}$  is the number of diffusion encoding. To train a ResNet based on 2D convolution, the diffusion-weighted images were reshaped and permuted as  $(N_Z, 2 \cdot N_{\text{diff}}, N_Y, N_X)$ , as illustrated in Figure 2 (D). In this manner, 2D convolution kernels in combination with ReLU activation functions loop through the varying diffusion-weighted contrast to learn the key features of the high-dimensional data and to reduce noisy and aliasing artifacts in unrolled reconstruction.

## 2.3 | Model Generalizability

Volumetric whole brain DWI acquisition consists of many multi-band slices, and the training of algorithm unrolling models on all slices requires hundreds of GPU computing hours. To investigate the model generalizability and to accelerate reconstruction, we performed two training and inference strategies. First, we trained the ADMM unrolling model with only one multi-band slice data, and subsequently tested the model on all remaining multi-band slices. We called this approach "single-slice

training". Second, we trained and tested every multiband slice individually, which was referred to as "slice-by-slice training". The single-slice training strategy saves tremendous computing time, as its model is learned from one single slice and the inference time per slice is only about one minute. By comparing these two training strategies, we aim at demonstrating the model generalizability and its applicability to other slices which are "unseen" in single-slice training.

## 2.4 | Comparison of Regularization Techniques

This work compared the reconstruction performance of three different regularization techniques, Tikhonov  $\ell^2$  regularization (as used in MUSE), LLR regularization, and ADMM unrolling with a learned regularization. Note that MUSE is a simultaneous multi-slice (SMS) parallel imaging method and poses no regularization along the diffusion dimension, effectively solving each DWI reconstruction independently. In contrast, the other two regularized reconstructions fall into the joint reconstruction regime. They jointly reconstruct all diffusion-weighted images and impose regularization terms exploring spatial-diffusion redundancies. For example, LLR enforces the low rankness of local spatial-diffusion matrices from diffusion-weighted images, whereas ADMM unrolling learns a regularization function composed by neural networks based on spatial-diffusion convolution kernels while enforcing data consistency during the unrolled training process.

## 2.5 | In Vivo Acquisition and Reconstruction

Table 1 lists two acquisition protocols implemented on a clinical 7 T MR system (MAGNETOM Terra, Siemens Healthineers, Erlangen, Germany) equipped with a 32-channel head coil (Nova Medical, Wilmington, MA, USA) and the XR-gradient system (maximum gradient strength 80 mT/m and a peak slew rate 200 T/m/s). Protocols #1 and #2 realized mesoscale high-resolution DWI with 0.7 mm isotropic resolution. Two-fold acceleration is employed in both in-plane and slice directions. Every Diffusion encoding is acquired by three shots in an interleaved manner and is shifted with respect to its former one, which results in 6 × 2-fold acceleration per shot. Noteworthy, the total scan time can be reduced to about 10 minutes (Protocol #2) when switching off navigator acquisition.

**TABLE 1** NAViEPI acquisition protocols

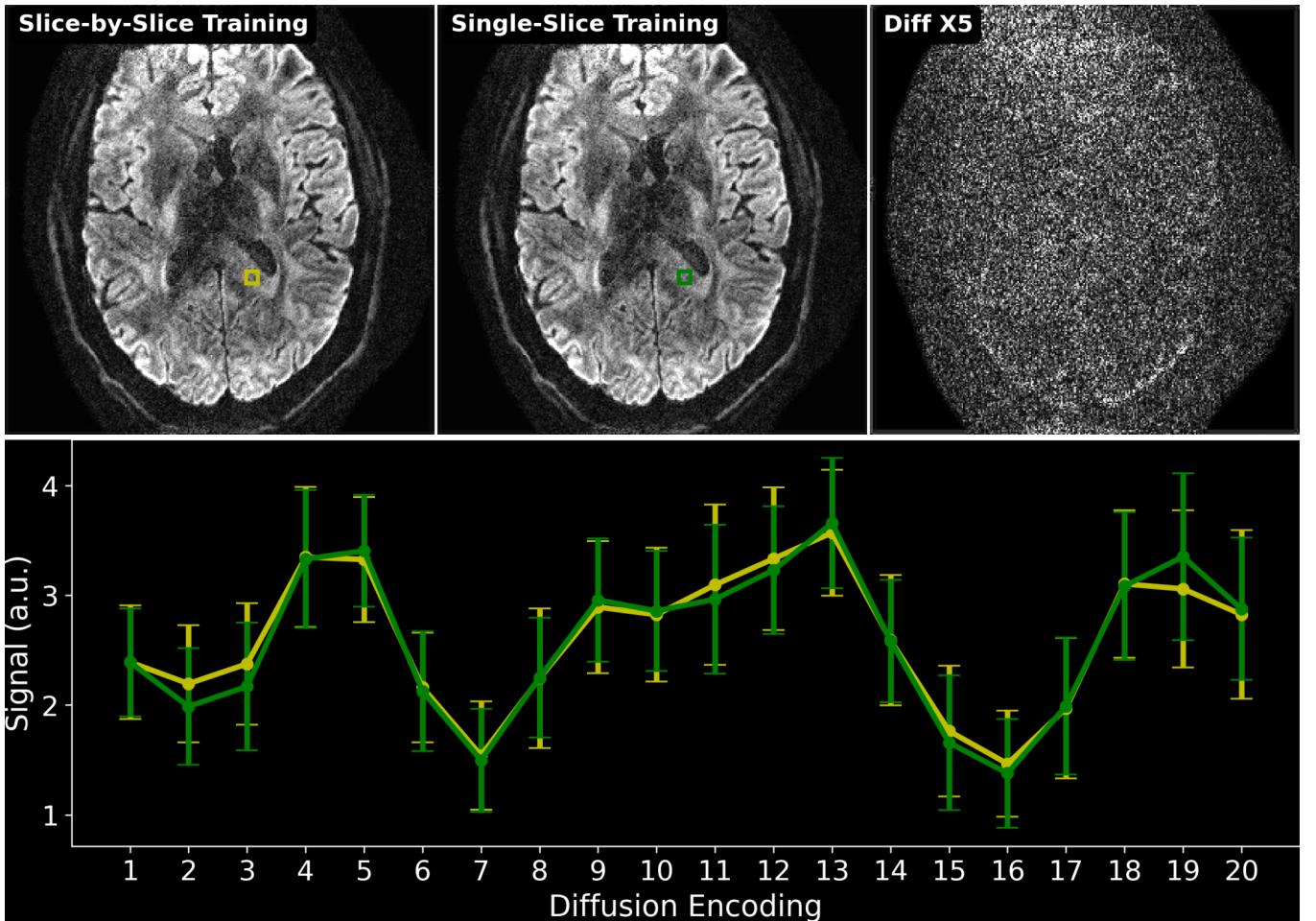
Protocol	#1	#2
Diffusion mode	MDDW	54
Diffusion scheme	monopolar	55
Diffusion direction	20	56
b-value (s/mm <sup>2</sup> )	1000	57
$b_0$	1	58
FOV (mm <sup>2</sup> )	200	59
Matrix size	286 × 286 × 176	60
Voxel (mm <sup>3</sup> )	0.7 × 0.7 × 0.7	61
Shots	3	62
Acceleration	2 × 2	63
Partial Fourier	5/8	64
Bandwidth (Hz/Pixel)	972	65
ESP (ms)	1.17	66
Navigator	Yes	No
TE (ms)	58/98.3	67
TR (ms)	15000	68
Acquisition (min)	16 : 27	69
		70
		71
		72
		73
		74
		75
		76
		77
		78
		79
		80
		81
		82
		83
		84
		85
		86
		87
		88
		89
		90
		91
		92
		93
		94
		95
		96
		97
		98
		99
		100
		101
		102
		103
		104
		105
		106

Three young healthy volunteers with written informed consent approved by the local ethics committee participated in this study. All reconstructions in this work were done on a single A100 SXM4/NVLink GPU with 80 GB memory (NVIDIA, Santa Clara, CA, USA).

## 3 | RESULTS

### 3.1 | Model Generalizability

Figure 3 demonstrates the generalizability of the proposed ADMM unrolling approach, i.e., an unrolled ADMM model trained on one single slice is applicable to all remaining "unseen" slices. Single-direction diffusion-weighted images from both the slice-by-slice training and the single-slice training strategies are displayed. The absolute difference between these two images shows no residual structural information, but mainly noise. Further, we plotted the mean and standard deviation within the selected region-of-interest (colored boxes in Figure 3 ) along all diffusion encoding. This again proves the cross-slice generalization of the proposed ADMM unrolling method. The plotted curves show quantitatively similar values between the two training strategies. With this, all following results were obtained utilizing the single-slice training strategy.



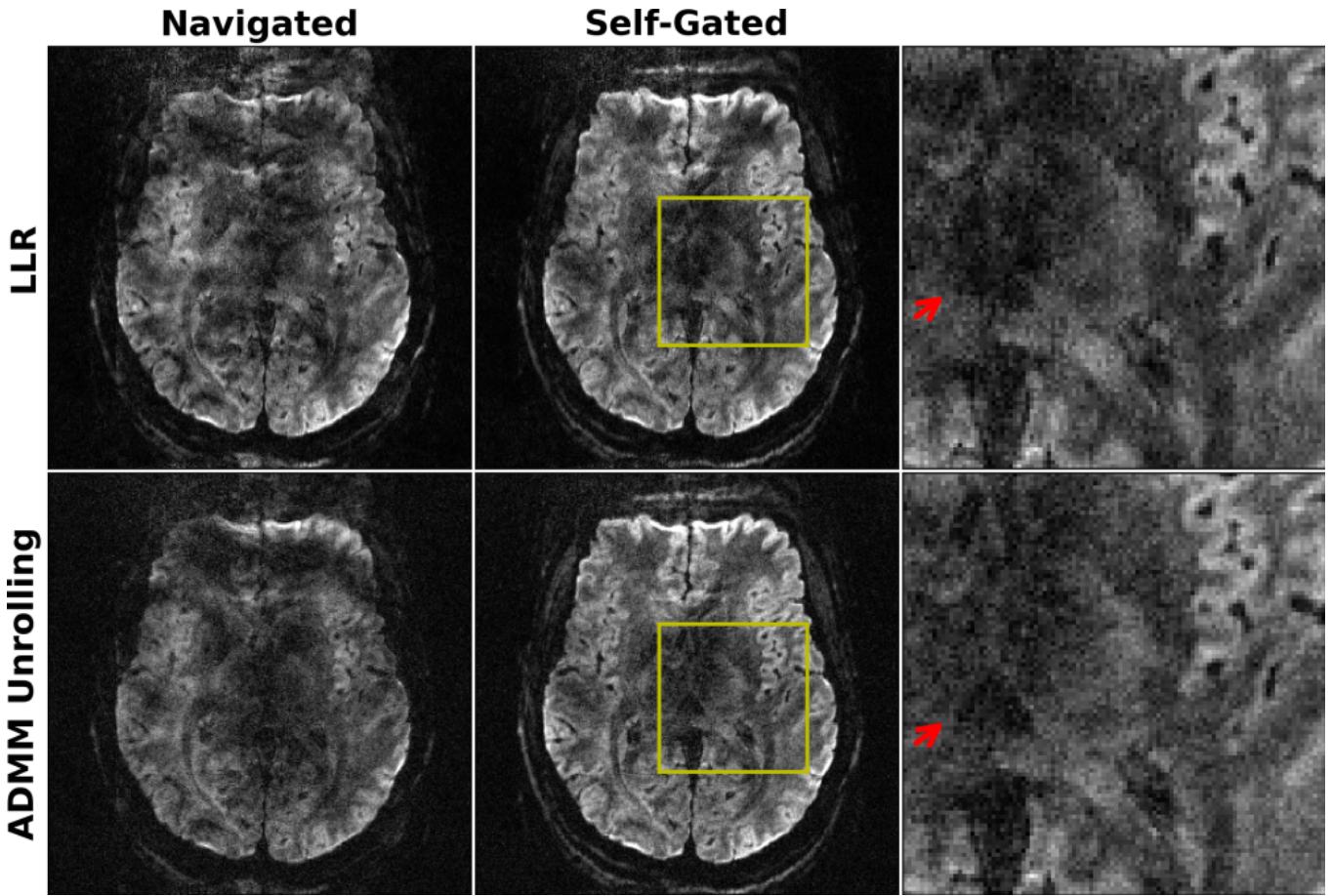
**FIGURE 3** Comparison of two training strategies: (1) slice-by-slice training, where every slice is trained and tested individually; (2) single-slice training, where the unrolled ADMM model is trained on only one slice and tested on all remaining slices. The top-right image shows the absolute difference between the reconstructed diffusion-weighted images at the 10th diffusion direction between (1) and (2). The bottom panel plots the mean and standard deviation of the signal within yellow and green rectangles in the slice-by-slice training and the single-slice training, respectively. No major qualitative or quantitative difference can be seen between the two training strategies.

### 3.2 | Retrospectively Self-Gated ADMM Unrolling

Figure 4 demonstrates the efficacy of the self-gated self-supervised ADMM unrolling reconstruction by comparing to the navigated and self-gated reconstructions on the first volunteer. Data were acquired using the NAViEPI sequence, as listed in Protocol #1 in Table 1. The single-direction diffusion-weighted images with accidental motion are displayed.

The selected diffusion encoding shows residual aliasing-like and severe motion-blurring artifacts in the navigated reconstructions, including both LLR regularization and ADMM unrolling. The main reason of these artifacts is that the acquisition of navigators increases the total scan time, resulting in higher sensitivity to accidental inter-shot motion. Admittedly, navigators are

valuable in the case of ultra high spatial resolution using many shots, e.g. 3-fold in-plane undersampling and 5-shot acquisition for the in-plane resolution of 0.5 mm<sup>32</sup>, which led to an in-plane acceleration of 15 per shot. In contrast, this experiment utilized 3 shots, yielding 6 × 2-fold acceleration per shot (refer to Protocol #1). Such an acceleration rate proves achievable in the self-gated approach. Both LLR regularized and unrolled ADMM reconstructions supply geometrically correct diffusion-weighted images without noticeable aliasing artifacts. This in turn indicates that motion corrupted the navigator data in this measurement. Further, self-gated ADMM unrolling exhibits much clearer tissue delineation in reconstructed diffusion-weighted images, as indicated by red arrows in the zoomed-in views in Figure 4, whereas self-gated LLR suffers from slightly blurry tissue boundaries and ambiguous signals.



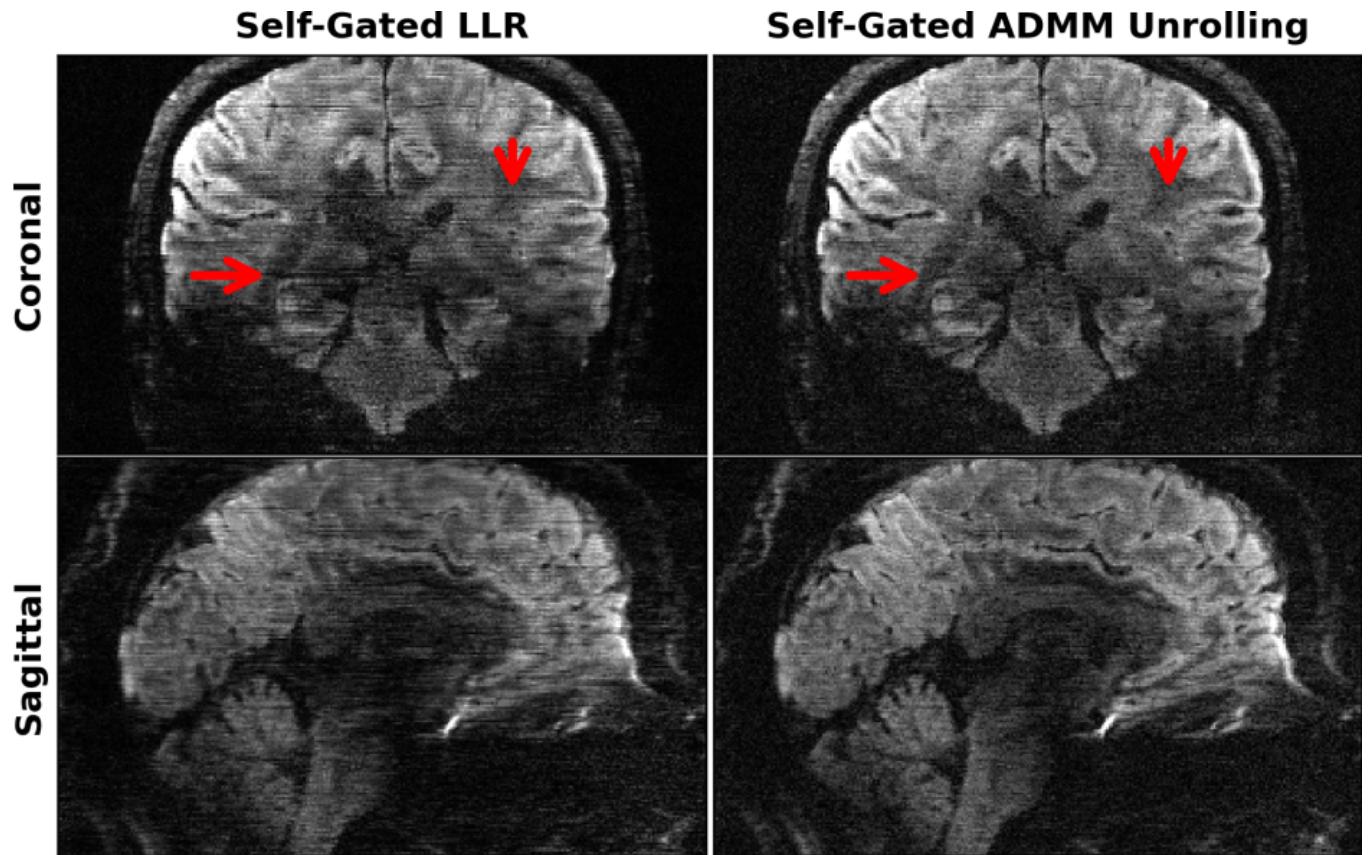
**FIGURE 4** Comparison of (top) LLR regularized and (bottom) ADMM unrolling reconstruction on 0.7 mm isotropic resolution DWI acquired by Protocol #1 with shot phase estimated from (left) navigators and (middle) imaging echoes, respectively. Zoomed views of the yellow boxes from the self-gated reconstruction are displayed in the right-most column. The use of navigators prolongs the total scan time, and thus increases the sensitivity to motion, as shown in the single-direction diffusion-weighted image reconstructed with navigated shot phase, where accidental motion occurred during navigator acquisition. The retrospectively self-gated reconstruction discards navigators, and renders sharper diffusion-weighted images. Compared to LLR, unrolled ADMM is advantageous in resolving clearer tissue boundaries in diffusion-weighted images, as indicated by red arrows.

Figure 5 shows coronal- and sagittal-view diffusion-weighted images with the same diffusion encoding as in Figure 4. As mentioned in Section 2.3, the unrolled ADMM model was trained using only one slice and then inferred on all remaining slices. The model generalizes well across slices. The inference of every slice takes only about one minute, whereas the LLR reconstruction takes about 48 minutes per slice. More importantly, the self-gated LLR reconstruction exhibits residual motion-induced stripping artifacts (refer to red arrows in Figure 5)<sup>36</sup>, whereas the self-gated ADMM unrolling approach substantially removes these artifacts and supplies high-quality diffusion-weighted images without the need of navigators. Both reconstructions show  $B_1$  field inhomogeneities in the cerebellum region and residual spatial distortion in the frontal brain region. These artifacts, however, are beyond the scope of this work.

### 3.3 | Prospectively Self-Gated ADMM Unrolling

Figure 6 compares the reconstruction results using the prospectively acquired iEPI data without navigators of the second volunteer (refer to Protocol #2 in Table 1). The snapshot single diffusion-direction diffusion-weighted images as well as mean diffusion-weighted images at three orthogonal views were displayed.

The MUSE reconstruction suffers from strong noise at such mesoscale voxel size. Its corresponding mean diffusion-weighted images show improved the visibility of brain tissues, but the overall image quality is not sufficient. The LLR regularized reconstruction largely reduces noise, but still exhibits suspicious dark signal in the axial view, which appears as striping artifacts in the



**FIGURE 5** Single-direction diffusion-weighted images at 0.7 mm isotropic resolution as reconstructed by retrospectively self-gated (left) LLR and (right) ADMM unrolling in (top) the coronal and (bottom) the sagittal views, respectively. The same diffusion direction as in Figure 4 is chosen for display. ADMM unrolling reduces phase ambiguities in the shot-combined reconstruction, thereby rendering clearer tissue delineation and reducing stripping artifacts (as indicated by the red arrows).

coronal and the sagittal views (refer to the red arrows in Figure 6). On the other hand, the LLR regularized reconstruction shows amplified noise in the cerebellum region (refer to the blue arrow in the sagittal view of Figure 6), which could be caused by the  $B_1$  excitation field inhomogeneity at 7T.

The above-mentioned striping artifacts are nearly gone in the unrolled ADMM reconstruction. Moreover, the diffusion-weighted images from ADMM unrolling in the axial view show clearer diffusion contrasts and thus better tissue delineation and continuity in the coronal and the sagittal views. Further, as indicated by the blue arrows in Figure 6, the diffusion-weighted image in the sagittal view shows more homogeneous signal distribution and reduced noise surrounding the cerebellum.

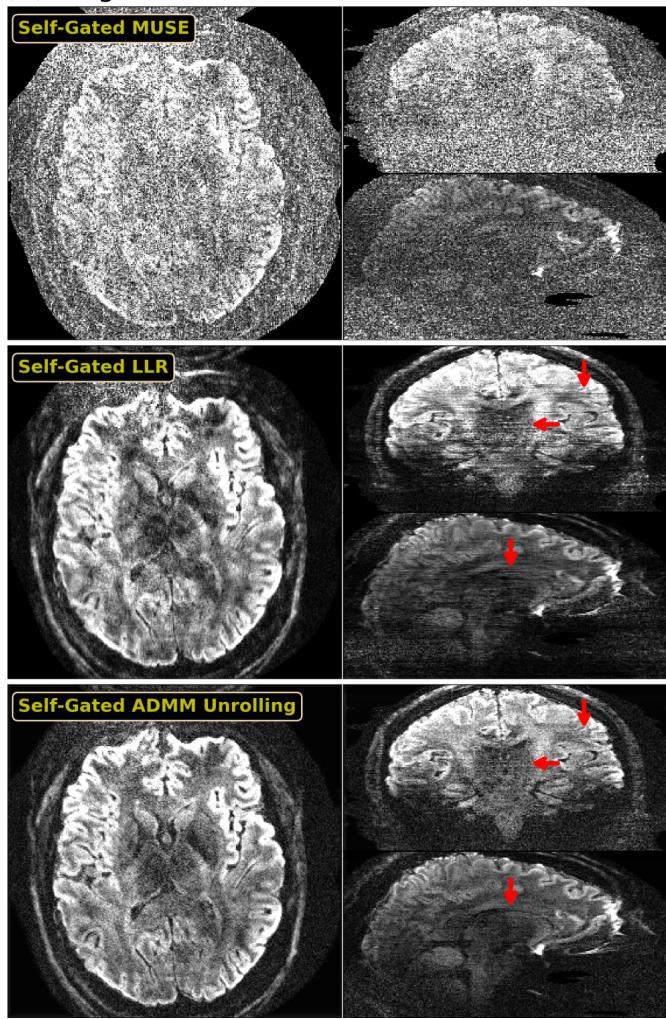
Figure 7 displays the training and validation loss as well as the learned regularization strength along epochs for the results shown in Figure 6. It can be seen that 100 epochs are sufficient for the convergence of ADMM unrolling. The model converges well along epochs, and does not show any over-fitting behavior (The validation

loss decays similarly as the training loss). In addition, the regularization strength converges to the value of about 0.027.

Figure 8 shows the reconstructed diffusion-weighted images at four different diffusion directions based on the iEPI data acquired from the third volunteer (the same subject as in Figure 3). In this experiment, the volunteer was instructed to keep still during scan. Again, the proposed self-gated ADMM unrolling reconstruction with spatial-diffusion convolution illustrates superior tissue structure delineation and diffusion contrasts to the LLR regularized reconstruction. The LLR reconstruction suffers from amplified noise in the frontal brain region. In contrast, the unrolled ADMM approach generally illustrates more homogeneous signal and noise distribution across the field-of-view. While LLR builds upon one single linear transformation (singular-value decomposition, SVD) and one nonlinear operation (soft thresholding)<sup>37</sup>, the ResNet in ADMM unrolling builds upon multiple two-dimensional convolutions and nonlinear activation functions. Therefore, deep neural networks

## 0.7 mm mesoscale DWI with 21 volumes @ 10 minutes

(A) Single-dir. DWI



(B) Mean DWI from 20 directions

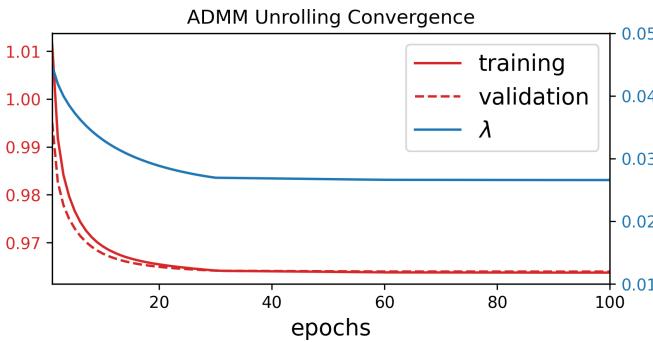
enable more in-depth exploration of key features in the high-dimensional data.

## 4 | DISCUSSION

This work reported a novel self-gated self-supervised learning approach based on ADMM unrolling for multi-shot undersampled iEPI acquisition and high-resolution DWI reconstruction. The self-gated ADMM unrolling achieved whole brain DWI with 20 diffusion-encoded directions and a  $b$ -value of  $1000 \text{ s/mm}^2$  at 0.7 mm

isotropic resolution, all within a scan time of less than 10 minutes. For comparison, the compressed sensing reconstruction with locally-low rank regularization was implemented also with the generic ADMM algorithm. Thus, our work assured fair comparison among different regularization methods.

The proposed self-gated ADMM unrolling approach is well-suited for online reconstruction deployment. Firstly, it requires much shorter acquisition time than the conventional MUSE approach with fully-sampled iEPI and our previous NAViEPI method. Secondly, it does



**FIGURE 7** Convergence analysis along the ADMM unrolling training and validation epochs for the results in Figure 6. Displayed curves are (red solid) the training loss, (red dashed) the validation loss, and (blue solid) the learned regularization strength  $\lambda$ , respectively. All parameters converge sufficiently and show no over-fitting.

not require large-scale fully-sampled data for training. Instead, its training is scan specific and requires only one slice. The trained ADMM unrolling model is applicable to different slices. Third, the inference time of the trained model is much shorter compared to the LLR regularization approach.

We observed that stripping-type motion artifacts occurred more frequently in the sub-millimeter isotropic resolution DWI regime. In addition, sub-millimeter isotropic voxel resulted in higher noise in diffusion-weighted images. This makes sense, as scans with reduced slice thickness are more susceptible to shot-to-shot phase variations. Since the primary aim of this work is to develop an efficient self-supervised learning technique for sub-millimeter DWI, we did not explore other advanced sampling strategies such as gSlider. However, since unrolled algorithms are flexible to MR physics modeling (e.g., the forward operator  $\mathcal{A}$ ), the proposed ADMM unrolling is extendable to incorporate with the gSlider encoding model for enhanced SNR performance.

This work demonstrated the capability of self-gated ADMM unrolling in reconstructing 0.7 mm isotropic resolution 3-shot iEPI DWI with  $(6 \times 2)$ -fold acceleration per shot. However, we also observed that the self-gated approach failed to recover aliasing-free diffusion-weighted images in the case of higher acceleration factors (e.g. the  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$  DWI data with an acceleration of  $15 \times 2$  per shot). To address this issue, acquiring shot-to-shot phase navigators helps with the shot-combined DWI reconstruction<sup>32</sup>. Therefore, the utilization of navigator acquisition and advanced deep learning reconstruction should be application oriented. Ultra-high spatial resolution, which requires many shots, necessitates the use of navigator shots. For the 0.7 mm

resolution with 3 shots, as shown in this work, the self-gated acquisition is beneficial of reducing scan time, given the superior performance of the proposed ADMM unrolling reconstruction. Alternatively, employing optimized trajectories with a more densely-sampled  $k$ -space central region could help better estimate shot phase variations<sup>4</sup>.

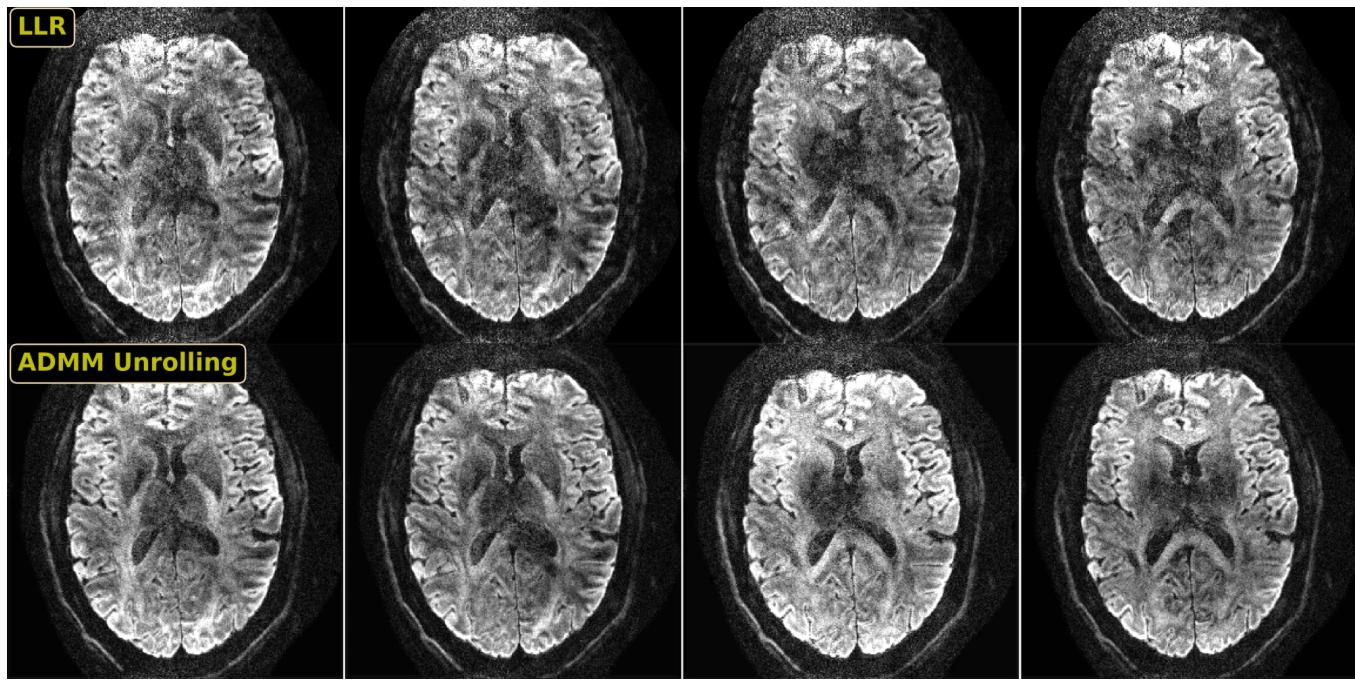
This work did not incorporate off-resonance correction in the reconstruction. As a logic extension, the multi-shot sequence can be modified to encode dynamic  $B_0$  field variation, which can then be employed in the SENSE-based forward operator and reconstruction. An established approach is known as the blip-up/down encoding<sup>38</sup>. This approach requires the acquisition of two images with opposing phase-encoding polarities (i.e., blip-up and blip-down) for the computation of  $B_0$  field maps. An alternative approach is to iteratively update  $B_0$  field based on the phase difference among acquired multiple echoes<sup>39</sup>. This approach does not require the pre-determination of  $B_0$  field, but poses higher computational demand in the inversion course of phase increments from every echo.

## 5 | CONCLUSIONS

In this work, we proposed a self-gated self-supervised learning reconstruction framework based on ADMM unrolling for high-resolution and motion-robust diffusion-weighted imaging at ultra-high field. Based on the mechanism of data splitting (cross validation), our proposed ADMM unrolling requires only one single slice for training and is generalized cross-slice. Plus, ADMM unrolling renders ultra-short inference / reconstruction time, and is thus feasible for clinical translation.

## ACKNOWLEDGMENTS

This work was supported in part by German Research Foundation (DFG) under projects 513220538 and 512819079, project 500888779 in the Research Unit RU5534 for MR biosignatures at UHF, and by the National Institutes of Health (NIH) under grants R01EB024532 and P41EB017183. The authors are grateful to scientific support and HPC resources provided by the Erlangen National High Performance Computing Center (NHR) of Friedrich-Alexander-University Erlangen-Nuremberg (FAU) under the NHR project b143dc. NHR is funded by federal and Bavarian state authorities. NHR@FAU hardware is partially funded by DFG under project 440719683. The authors are thankful to Dr. Vikas Gulani for insightful discussions. The



**FIGURE 8** Prospectively self-gated DWI reconstruction results at 0.7 mm isotropic resolution. Displayed images are one axial slice at four different diffusion-encoding directions. ADMM unrolling enables much cleaner delineations of diffusion contrasts than LLR regularized reconstruction.

authors thank ChatGPT for the revision of the Introduction section.

## DATA AVAILABILITY STATEMENT

In the spirit of open science and reproducible research, source codes of this work are available in <https://github.com/ZhengguoTan/DeepDWI>. The presented 0.7 mm DWI raw  $k$ -space data is available in <https://doi.org/10.5281/zenodo.10781347> and <https://doi.org/10.5281/zenodo.13864504>.

## ORCID

Zhengguo Tan 0000-0002-4322-5109  
 Patrick Liebig 0000-0001-7342-3715  
 Frederik B Laun 0000-0002-9269-5609  
 Florian Knoll 0000-0001-5357-8656

## REFERENCES

1. Jones Derek K. *Diffusion MRI: Theory, methods, and applications*. Oxford University Press; 2010.
2. Mansfield Peter. Multi-planar image formation using NMR spin echoes. *J Phys C*. 1977;10:55-58.
3. Butts Kim, Riederer Stephen J, Ehman Richard L, Thompson Richard M, Jack Clifford R. Interleaved echo planar imaging on a standard MRI system. *Magn Reson Med*. 1993;31:67-72.
4. Liu Chunlei, Bammer Roland, Kim Dong-hyun, Moseley Michael E. Self-navigated interleaved spiral (SNAILS): Application to high-resolution diffusion tensor imaging. *Magn Reson Med*. 2004;52:1388-1396.
5. Chen Nan-Kuei, Guidon Arnaud, Chang Hing-Chiu, Song Allen W. A robust multi-shot scan strategy for high-resolution diffusion weighted MRI enabled by multiplexed sensitivity-encoding (MUSE). *NeuroImage*. 2013;72:41-47.
6. Porter David A, Heidemann Robin M. High resolution diffusion-weighted imaging using readout-segmented echo-planar imaging, parallel imaging and a two-dimensional navigator-based reacquisition. *Magn Reson Med*. 2009;62:468-475.
7. Heidemann Robin M, Porter David A, Anwander Alfred, et al. Diffusion imaging in humans at 7 T using readout-segmented EPI and GRAPPA. *Magn Reson Med*. 2010;64:9-14.
8. Andersson Jesper L R, Skare Stefan, Ashburner John. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*. 2003;20:870-888.
9. Setsompop Kawin, Gagoski Borjan A, Polimeni Jonathan R, Witzel Thomas, Wedeen Van J, Wald Lawrence L. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced  $g$ -factor penalty. *Magn Reson Med*. 2012;67:1210-1224.
10. Pruessmann Klaas P, Weiger Markus, Scheidegger Markus B, Boesiger Peter. SENSE: Sensitivity encoding

- for fast MRI. *Magn Reson Med.* 1999;42:952-962.
11. Griswold Mark A, Jakob Peter M, Heidemann Robin M, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med.* 2002;47:1202-1210.
  12. Bammer Roland, Keeling Stephen L, Augustin Michael, et al. Improved diffusion-weighted single-shot echo-planar imaging (EPI) in stroke using sensitivity encoding (SENSE). *Magn Reson Med.* 2001;46:548-554.
  13. Lustig Michael, Donoho David, Pauly John M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med.* 2007;58:1182-1195.
  14. Mani Merry, Jacob Mathews, Kelley Douglas, Magnotta Vincent. Multi-shot sensitivity-encoded diffusion data recovery using structured low-rank matrix completion (MUSSELS). *Magn Reson Med.* 2017;78:494-507.
  15. Hu Yuxin, Wang Xiaole, Tian Qiyuan, et al. Multi-shot diffusion-weighted MRI reconstruction with magnitude-based spatial-angular locally low-rank regularization (SPA-LLR). *Magn Reson Med.* 2020;83:1596-1607.
  16. Veraart Jelle, Novikov Dmitry S, Christiaens Daan, Ades-aron Benjamin, Sijbers Jan, Fieremans Els. Denoising of diffusion MRI using random matrix theory. *NeuroImage.* 2016;142:394-406.
  17. Tian Qiyuan, Li Ziyu, Fan Qiuyun, et al. SDnDTI: Self-supervised deep learning-based denoising for diffusion tensor MRI. *NeuroImage.* 2023;253:119033.
  18. Setsompop Kawin, Fan Qiuyun, Stockmann Jason, et al. High-resolution in vivo diffusion imaging of the human brain with generalized slice dithered enhanced resolution: Simultaneous multislice (gSlider-SMS). *Magn Reson Med.* 2018;79:141-151.
  19. Dai Erpeng, Liu Simin, Guo Hua. High-resolution whole-brain diffusion MRI at 3T using simultaneous multi-slab (SMSlab) acquisition. *NeuroImage.* 2021;237:118099.
  20. Dong Zijing, Reese Timothy G, Lee Hong-Hsi, et al. Romer-EPTI: Rotating-view motion-robust super-resolution EPTI for SNR-efficient distortion-free in-vivo mesoscale diffusion MRI and microstructure imaging. *Magn Reson Med.* 2024;.
  21. Liang Zhi-Pei. Spatiotemporal imaging with partially separable functions. In: *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'4)*. 2007;988-991.
  22. Zhang Tao, Pauly John M, Levesque Ives R. Accelerated parameter mapping with a locally low rank constraint. *Magn Reson Med.* 2015;73:655-661.
  23. Dong Zijing, Wang Fuyixue, Reese Timothy G, Bilgic Berkin, Setsompop Kawin. Echo planar time-resolved imaging with subspace reconstruction and optimized spatiotemporal encoding. *Magn Reson Med.* 2020;84:2442-2455.
  24. Hammernik Kerstin, Klatzer Teresa, Kobler Erich, et al. Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med.* 2018;79:3055-3071.
  25. Aggarwal Hemant K, Mani Merry P, Jacob Mathews. MoDL: Model-based deep learning architecture for inverse problems. *IEEE Trans Med Imaging.* 2018;38:394-405.
  26. Cho Jaejin, Jun Yohan, Wang Xiaoqing, Kobayashi Caique, Bilgic Berkin. Improved Multi-shot Diffusion-Weighted MRI with Zero-Shot Self-supervised Learning Reconstruction. In: *26th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'26)*. 2023;457-466.
  27. He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'16)*. 2016;770-778.
  28. Yaman Burhaneddin, Hosseini Seyed Amir Hossein, Akçakaya Mehmet. Zero-shot self-supervised learning for MRI reconstruction. In: *10th International Conference on Learning Representations (ICLR'10)*. 2022.
  29. Mani Merry, Magnotta Vincent A, Jacob Mathews. qModeL: A plug-and-play model-based reconstruction for highly accelerated multi-shot diffusion MRI using learned priors. *Magn Reson Med.* 2021;86:835-851.
  30. Hinton G. E., Salakhutdinov R. R.. Reducing the dimensionality of data with neural networks. *Science.* 2006;313:504-507.
  31. Boyd Stephen, Parikh Neal, Chu Eric, Peleato Borja, Eckstein Jonathan. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach Learn.* 2010;3:1-122.
  32. Tan Zhengguo, Liebig Patrick Alexander, Heidemann Robin Martin, Laun Frederik Bernd, Knoll Florian. Accelerated diffusion-weighted magnetic resonance imaging at 7 T: Joint reconstruction for shift-encoded navigator-based interleaved echo planar imaging (JETS-NAViEPI). *Imaging Neuroscience.* 2024;2:1-15.
  33. Uecker Martin, Karaus Alexander, Frahm Jens. Inverse reconstruction method for segmented multishot diffusion-weighted MRI with multiple coils. *Magn Reson Med.* 2009;62:1342-1348.
  34. Merrem Andreas, Hofer Sabine, Hosseini Ali Seif Amir, et al. Diffusion-weighted MRI of the prostate without susceptibility artifacts: Undersampled multi-shot turbo-STEAM with rotated radial trajectories. *NMR Biomed.* 2019;32:e4074.
  35. Kingma Diederik P, Ba Jimmy. ADAM: A Method for Stochastic Optimization. In: *3rd International Conference on Learning Representations (ICLR'3)*. 2015.
  36. Chang Wei-Tang, Huynh Khoi Minh, Yap Pew-Thian, Lin Weili. Navigator-Free Submillimeter Diffusion MRI using Multishot-encoded Simultaneous Multi-slice (MUSIUM) Imaging. In: *Proceedings of the 29th Annual Meeting of ISMRM*. 2021;1322.
  37. Cai Jian-Feng, Candès Emmanuel J, Shen Zuowei. A singular value thresholding algorithm for matrix completion. *SIAM J Optim.* 2010;20:1956-1982.
  38. Zahneisen Benjamin, Aksoy Murat, Maclare Julian, Wuerslin Christian, Bammer Roland. Extended hybrid-space SENSE for EPI: Off-resonance and eddy current corrected joint interleaved blip-up/down reconstruction. *NeuroImage.* 2017;153:97-108.
  39. Tan Zhengguo, Unterberg-Buchwald Christina, Blumenthal Moritz, et al. Free-breathing liver fat,  $R_2^*$  and  $B_0$  field mapping using multi-echo radial FLASH and regularized model-based reconstruction. *IEEE Trans Med Imaging.* 103

01 2023;42:1374-1387.



54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106