

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

Zhengguo Tan<sup>1</sup> | Patrick A 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 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 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 selfgated self-supervised DeepDWI learning for multi-shot echo planar imaging with difusion 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 diffusion 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:** 3828

## 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

Part of this work has been presented at the ISMRM 2025, Honolulu, USA.

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.

Compressed sensing 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 multi-shot  $k$ -space. 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. Hu et al.<sup>15</sup> developed SPA-LLR, which employs locally-low rank (LLR) regularization in joint  $k$ - $q$ -space reconstruction. Further, Tan et al.<sup>24</sup> extended LLR to accelerated multi-band multi-shot multi-shell reconstruction.

Recently, emerging deep learning techniques show great promise in deep DWI reconstruction. Cho et al.<sup>25</sup> proposed to learn a  $k$ -space regularization function with ResNet<sup>26</sup> and zero-shot self-supervised learning<sup>27,28</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. 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 utilized as a  $q$ -space regularizer, in combination

with a total-variation spatial regularizer, in the joint  $k$ - $q$ -space reconstruction, as solved by the alternating direction method of multipliers (ADMM)<sup>31</sup>.

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 multi-band 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<sup>32,33,34</sup>. 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 | 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). Protocol #1 with 1 mm isotropic resolution serves as the reference with in-plane fully sampling and the multi-band factor 3. This reference 4-shot data is retrospectively undersampled to only 2 shots (i.e., 2-fold in-plane undersampling) and then trained and tested with the proposed self-supervised learning. Protocols #2 and #3 realize high resolution mesoscale DWI with isotropic resolution 0.7 mm. Two-fold acceleration is used in both the in-plane and in-slice directions. Every diffusion encoding is acquired by three shots in an interleaved manner and is shifted with respect to its former, resulting in a  $6 \times 2$ -fold acceleration per shot. It is noteworthy that the total scan time can be reduced to about 10 minutes (Protocol #3) when switching off navigator acquisition.

**TABLE 1** Acquisition protocols

Protocol <sup>1</sup>	#1 (1.0 mm)	200	#2 (0.7 mm NAV)   #3 (0.7 mm)
FOV (mm <sup>2</sup> )		200	
Matrix size	200 × 200 × 114		286 × 286 × 176
Voxel (mm <sup>3</sup> )	1.0 × 1.0 × 1.0		0.7 × 0.7 × 0.7
Shots	4		3
Acceleration	1 × 3		2 × 2
Partial Fourier	5/8		5/8
Bandwidth (Hz/Pixel)	1086		972
ESP (ms)	1.04		1.17
Navigator	No	Yes	No
TE (ms)	66	58/98.3	58
TR (ms)	5400	15000	8900
Acquisition (min)	7 : 52	16 : 27	9 : 57

<sup>1</sup> All protocols employed the MDDW diffusion acquisition mode with monopolar diffusion encoding gradients, 1  $b_0$  volume and 20 diffusion-weighted volumes with the  $b$ -value of 1000 s/mm<sup>2</sup>.

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).

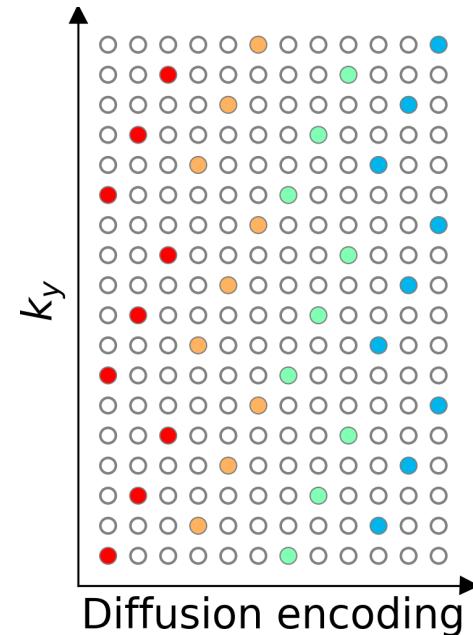
## 2.2 | Multi-Band Multi-Shot DWI with Diffusion-Shift Encoding

Our previous work<sup>24</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 line  $k_y$  for a diffusion encoding is shifted with respect to its adjacent line to create a 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,

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

Here, the images  $\mathbf{x}$  are point-wise multiplied with the precomputed shot-to-shot phase variation maps ( $\Phi$ ) and coil sensitivity maps ( $\mathbf{S}$ ). The output images are then converted to  $k$ -space via the two-dimensional fast Fourier transform ( $\mathbf{F}$ ), multiplied point-wise with the multi-band phases ( $\Theta$ ), summed along the slice dimension ( $\Sigma$ ), 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>4,35,5,36</sup> achieved the self-gating strategy, where the  $k$ -space data of each shot were used to reconstruct its corresponding shot image followed by a phase smoothing approach (i.e., the phase variation operator  $\Phi$ ).



**FIGURE 1** Three-shot DWI with diffusion shift encoding. This work employs three-shot per diffusion encoding and each shot has an in-plane undersampling factor of 6. Every three columns assemble one diffusion encoding and thus are colored the same. The starting  $k_y$  line is shifted between adjacent diffusion encoding to create complementary  $k$ - $q$ -space sampling.

Self-gated shot phase estimation does not require the acquisition of phase navigator data, thereby rendering short scan time. In the imaging scenario of submillimeter resolution, usually many shots are needed. As a result, this increases the acceleration factor per shot and thus necessitates the use of navigators. The drawback of

adding navigators is the increase of scan time. Therefore, this work aims to develop an efficient DWI protocol that can achieve submillimeter resolution while retaining short scan time.

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>24</sup> or neural networks with nonlinear activation functions. ADMM was implemented in PyTorch to solve for Equation (2).

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

Instead of the two-step alternating minimization unrolling scheme as used in MoDL<sup>34</sup>, we employed the ADMM unrolling to solve the self-supervised learning reconstruction in Equation (2). The update rule of ADMM unrolling reads

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

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  $\mathbf{v}$ .

Every training epoch consists of 12 looping repetitions. In each repetition, 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  $\mathbf{V}$ , as shown in Figure 2. Each repetition has different masks. In each training epoch, the corresponding masks of the given repetition is used in order to update the ResNet parameters  $\omega$  (Figure 2 (B)). Plus, the validation step is performed after every training epoch to update the minimal validation loss. If the validation loss does not

---

**Algorithm 1** Self-Supervised ADMM Unrolling

---

```

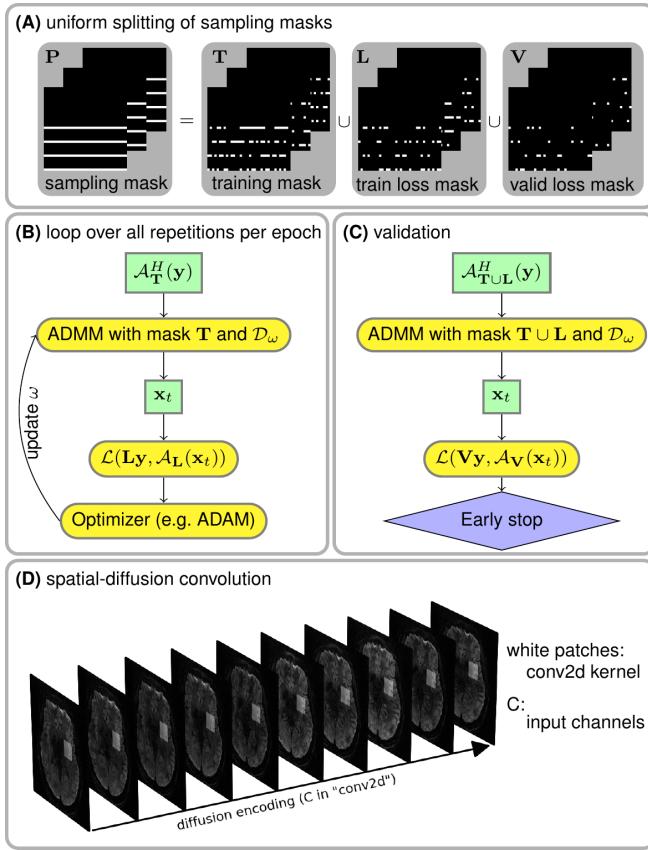
1: Initialization:
2:   split sampling mask  $\mathbf{P}$  into 12 repetitions, each of
   which consists of three disjoint sets  $\mathbf{T}$ ,  $\mathbf{L}$ , and  $\mathbf{V}$ 
3:    $p \leftarrow 0$  and  $N_{\text{epoch}} \leftarrow 100$ 
4:    $\mathcal{D}_\omega$  set as ResNet
5:    $\rho \leftarrow 0.05$  and  $\lambda \leftarrow 0.05$ 
6:    $\text{Loss}_{\text{valid}} \leftarrow \inf$  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

```

---

reduce for 12 consecutive epochs or if 100 epochs are reached, the training is terminated. The use of three disjoint masks is inline with the zero-shot self-supervised learning approach<sup>27,28</sup> for scan-specific parallel imaging reconstruction. In contrast, Self-supervised learning via data undersampling (SSDU)<sup>38</sup> splits the sampling mask into only two sub-masks, but requires multiple datasets for training.

The index  $k$  in Equation (3) denotes the unrolling iteration, and  $\mathcal{D}_\omega$  denotes the ResNet<sup>26</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.



**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>37</sup> during training, but remain fixed during the validation step. **(D)** A stack of diffusion-weighted images is input into ResNet during ADMM unrolling.

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.

## 2.4 | 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.

## 2.5 | 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.

## 01 | RESULTS

### 02 | Ablation Study

05 Figure 3 validates the proposed self-supervised ADMM  
 06 unrolling reconstruction method with the 4-shot fully-  
 07 sampled 1.0 mm dataset (Protocol #1 in Table 1 ).  
 08 MUSE on the 2-shot undersampled data exhibits notice-  
 09 able image quality degradations, as confirmed by the  
 10 visual inspections as well as the SSIM and PSNR quan-  
 11 tities. Both LLR and ADMM unrolling are capable of  
 12 reconstructing high quality diffusion-weighted images  
 13 without significant loss of image details and SNR. The  
 14 computed quantitative metrics show ADMM unrolling  
 15 performs slightly better than LLR.

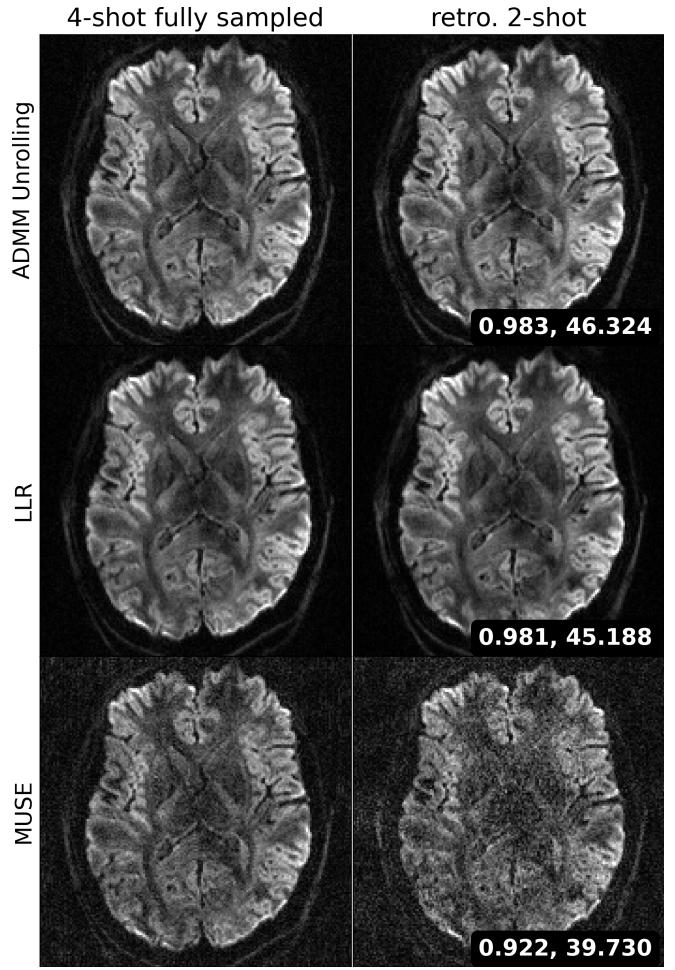
### 17 | Model Generalizability

19 Figure 4 demonstrates the generalizability of the pro-  
 20 posed ADMM unrolling approach, i.e., an unrolled  
 21 ADMM model trained on one single multi-band slice  
 22 is applicable to all remaining "unseen" slices. Single-  
 23 direction diffusion-weighted images from both the slice-  
 24 by-slice training and the single-slice training strategies  
 25 are displayed. The absolute difference between these  
 26 two images shows no residual structural information,  
 27 but mainly noise. Moreover, we plotted the mean and  
 28 standard deviation within the selected region-of-interest  
 29 (colored boxes in Figure 4 ) along all diffusion encoding  
 30 directions. This again proves the cross-slice generaliza-  
 31 tion of the proposed selfgated self-supervised ADMM  
 32 unrolling method. The plotted curves show quantita-  
 33 tively similar values between the two training strategies.  
 34 With this, the following results were obtained based upon  
 35 the single-slice training strategy.

### 38 | Self-Gated ADMM Unrolling

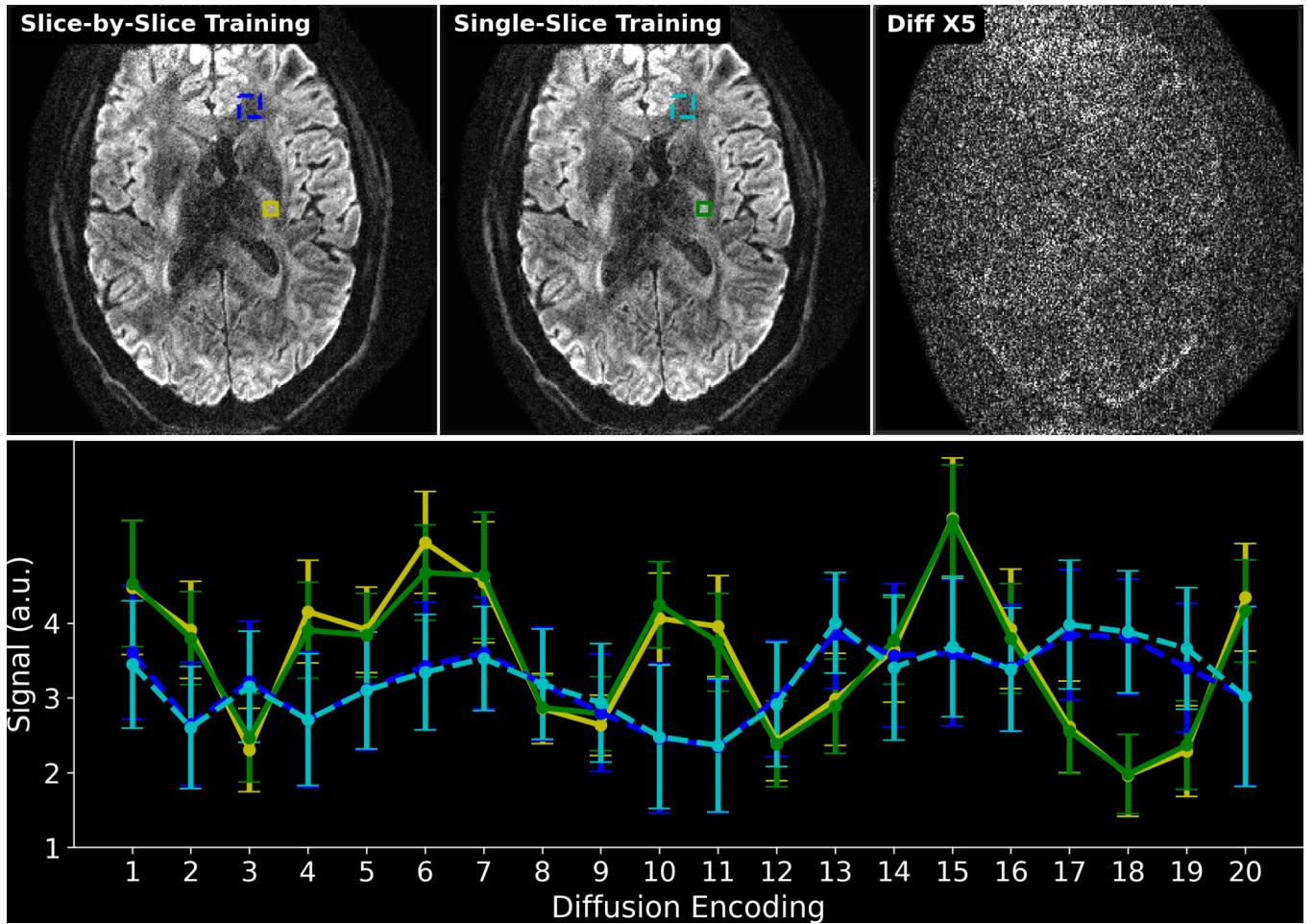
40 Figure 5 demonstrates the efficacy of the self-gated  
 41 self-supervised ADMM unrolling reconstruction by com-  
 42 paring with the navigated reconstruction on the first  
 43 volunteer. Both MUSE and ADMM unrolling recon-  
 44 structions were performed. Data were acquired using  
 45 the NAVIEPI sequence, as listed in Protocol #2 in  
 46 Table 1 . The single-direction diffusion-weighted images  
 47 are displayed.

48 The diffusion-weighted images from navigated recon-  
 49 structions show spatially varying phase. The reason is  
 50 that the shot-to-shot phase variations were estimated  
 51 from the second echo in NAVIEPI, i.e., the navigator,  
 52 whose echo time is different from the first echo. The echo  
 53 time difference results in residual phases in the combined



**FIGURE 3** Ablation study with the fully-sampled reference data acquired by Protocol #1. The first column displays one diffusion-weighted image from the 4-shot fully-sampled data reconstructed via (from top to bottom) the proposed self-supervised ADMM unrolling, LLR, and MUSE. The second column displays the diffusion-weighted image from the retrospectively undersampled 2-shot data reconstructed via the afore-mentioned methods. Two image metrics, structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR) are computed between the 4-shot and the 2-shot reconstructions.

93 diffusion-weighted images. On the contrary, selfgated  
 94 reconstructions show only subtle phase, because shot-  
 95 to-shot phase variations were estimated from the first  
 96 echoes themselves. The reduced phase variation in the  
 97 selfgated reconstruction leads to less phase ambiguity.  
 98 This is beneficial in the ADMM unrolling reconstruction,  
 99 where convolutions were performed in both the real and  
 100 imaginary channels. Reduced phase ambiguity fosters the  
 101 learning procedure. Consequently, compared to MUSE  
 102 with the MPPCA denoiser<sup>39</sup>, the selfgated ADMM  
 103 unrolling reconstruction achieves strong denoising and  
 104 resolves clear tissue details.



**FIGURE 4** 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 two sets of 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.

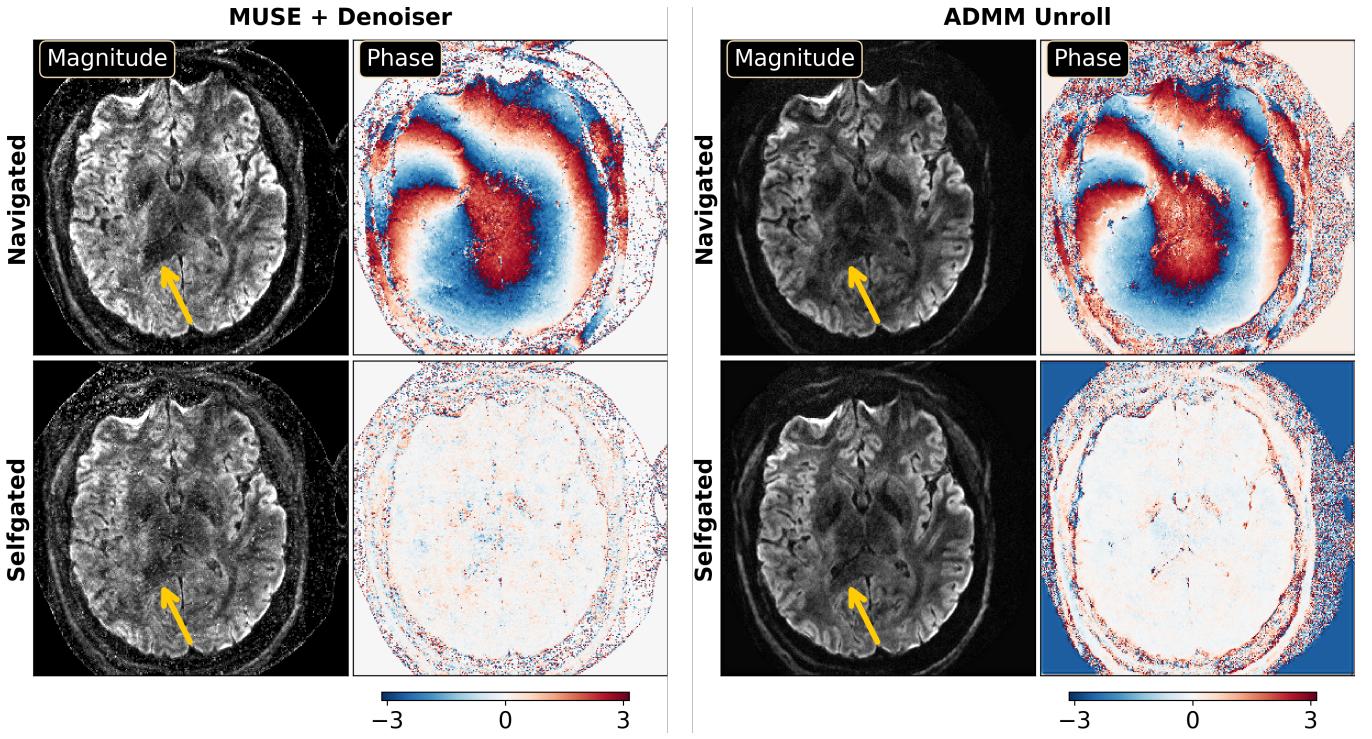
The advantage of the proposed ADMM unrolling for high resolution DWI with accelerated acquisition is further evident in Figure 6. The mean diffusion-weighted image from ADMM unrolling shows clear delineation of the claustrum, which is a thin sheet of neurons and is important to consciousness. In contrast, MUSE with the MPPCA denoiser shows noisy and blurred boundaries of the claustrum.

Figure 7 shows coronal- and sagittal-view diffusion-weighted images with the same diffusion encoding as in Figure 5. As mentioned in Section 2.4, the unrolled ADMM model was trained using only one slice and then inferred on all remaining slices. Again, the single-slice 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

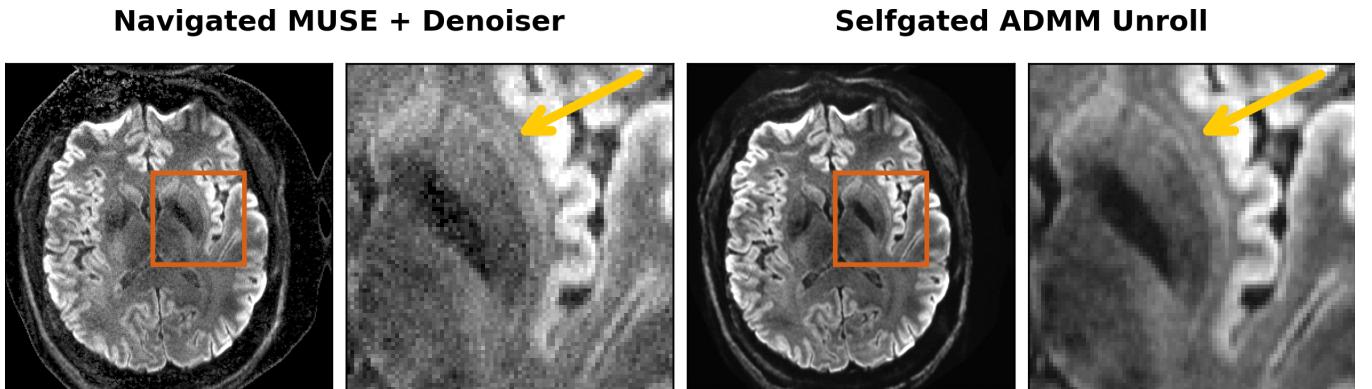
exhibits residual motion-induced stripping artifacts<sup>40</sup>, whereas the self-gated ADMM unrolling approach substantially removes these artifacts and supplies high-quality diffusion-weighted images without the need of navigators. Both reconstructions show  $B_1$  field inhomogeneities in the cerebellum region as well as off-resonance induced spatial distortion in the frontal brain region. These artifacts, however, are beyond the scope of this work.

### 3.4 | Diffusion Tensor Imaging (DTI)

Figures 4 and 8 utilize the Protocol #3 acquired data from the same volunteer. Here, Figure 8 displays the cFA maps based on the reconstructed diffusion-weighted images by MUSE with denoiser, LLR, and ADMM unrolling, respectively. Given the  $2 \times 2$ -fold acceleration



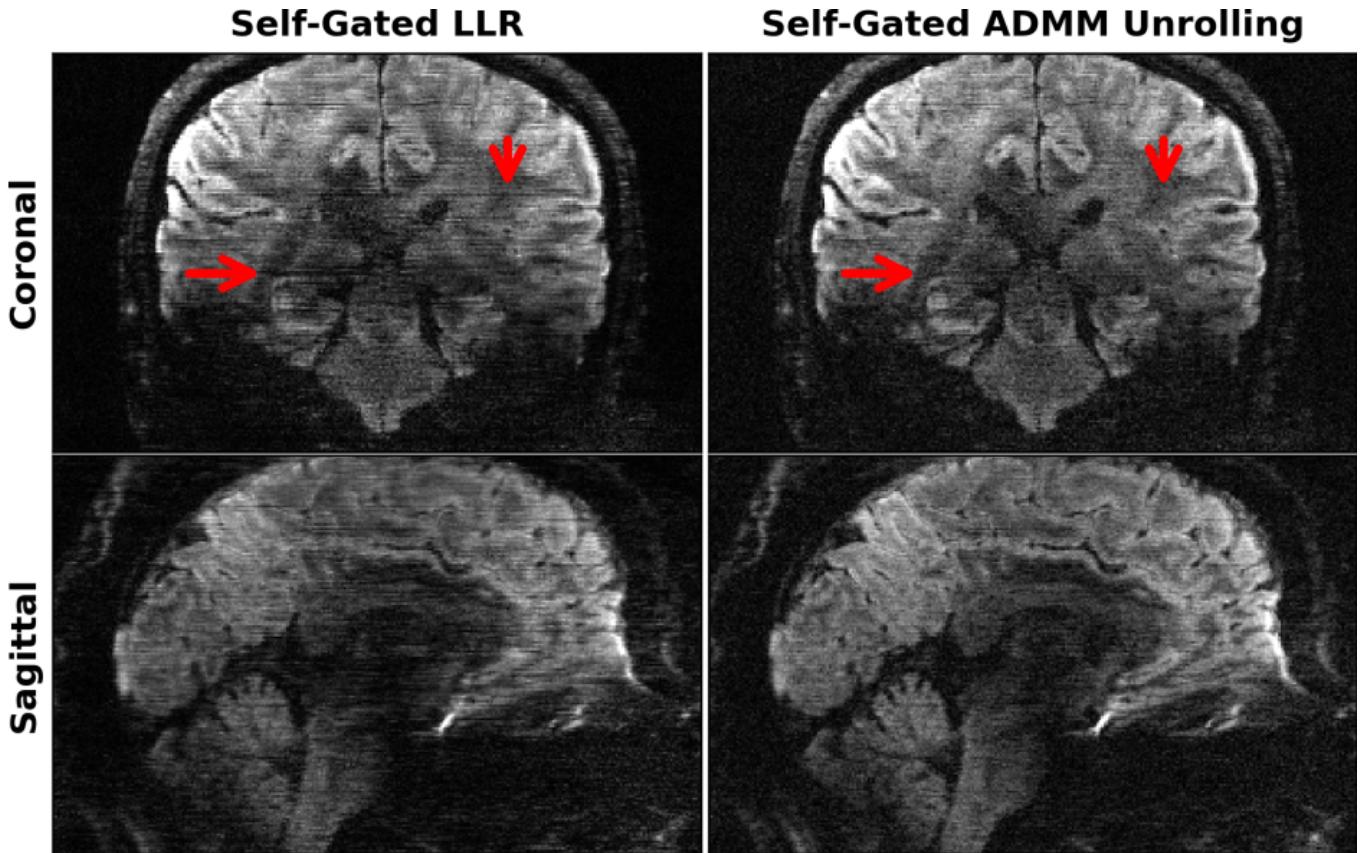
**FIGURE 5** Validation of the proposed self-gated ADMM unrolling reconstruction with data acquired by Protocol #2 in Table 1. Both MUSE and ADMM unrolling were performed with navigated and selfgated shot-to-shot phase maps, respectively. Compared to MUSE, the ADMM unrolled reconstruction excels in denoising while maintaining structural details. Selfgated ADMM unrolling shows improved image quality in terms of tissue delineation than navigated reconstruction.



**FIGURE 6** 0.7 mm isotropic resolution DWI with the proposed selfgated ADMM unrolling enables the visualization of the tiny structure claustrum, whereas the MUSE reconstruction shows only blurred appearance. Displayed images are the mean diffusion-weighted image from 20 directions and its zoomed-in region.

and the submillimeter spatial resolution (0.7 mm), the MPPCA denoiser applied onto MUSE is insufficient to supply sharp fiber orientations. Although LLR shows improvements when compared to the MUSE approach, but still shows overall blurring in the cFA map, especially within the gray matter region. The proposed selfgated self-supervised ADMM unrolling is able to resolve thin fibers within gray matters, as pointed by the color arrows.

Figure 9 displays the training and validation loss as well as the learned regularization strength along epochs. 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. Note that the validation loss is slightly low than the training loss. This is because more



**FIGURE 7** 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 5 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).

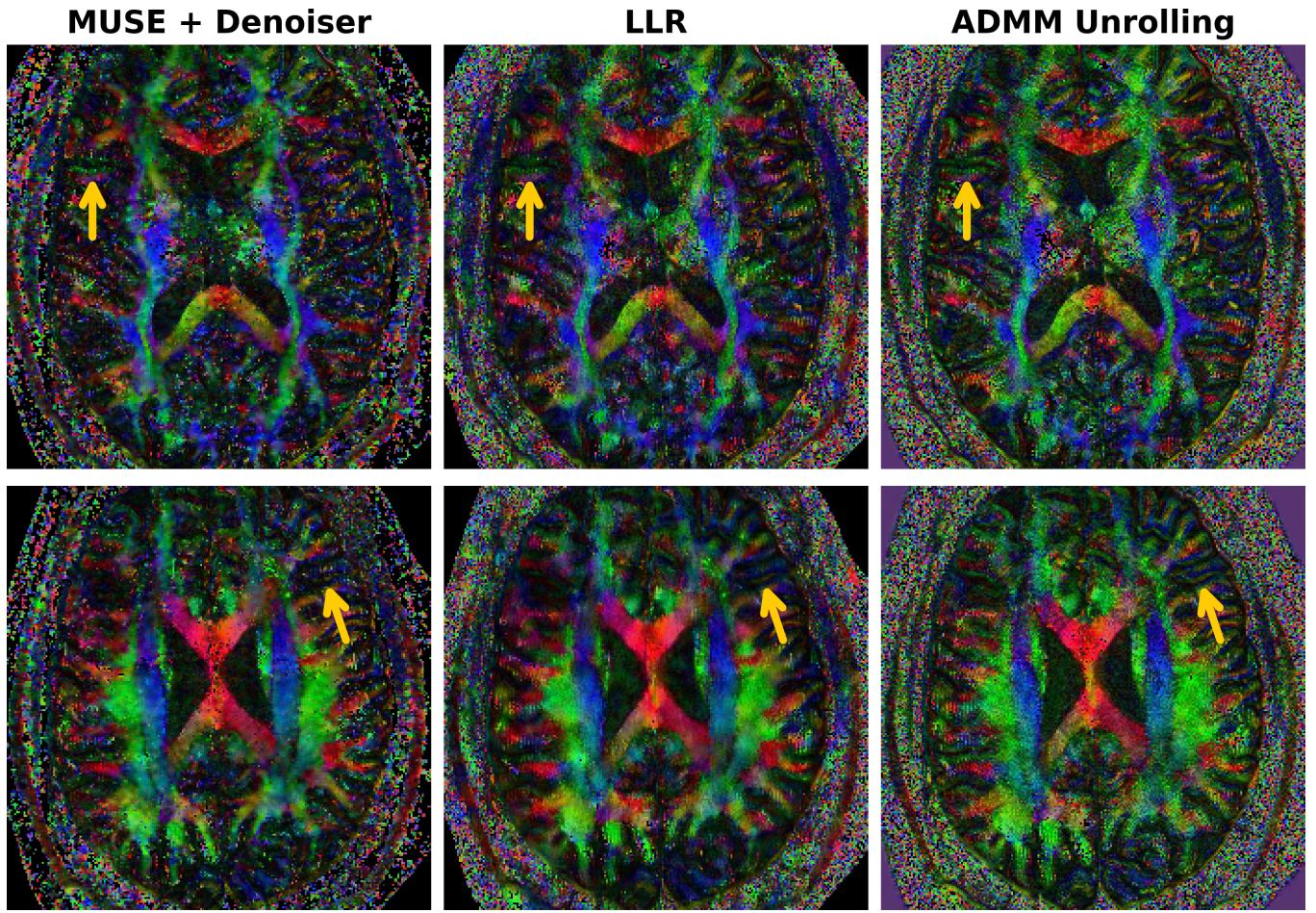
data is split into the training mask than the validation mask.

## 4 | DISCUSSION

This work reports a novel self-gated self-supervised learning approach based on ADMM unrolling for multi-shot multi-band undersampled iEPI acquisition and high-resolution DWI reconstruction. The self-gated ADMM unrolling achieves whole brain DWI with 21 diffusion volumes 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. Our proposed ADMM unrolling approach has several advantages. (1) Inline with the previous approaches for single image recovery<sup>27</sup> and parallel imaging<sup>28</sup>, our approach trains an unrolled reconstruction network with only one dataset utilizing the concept of data splitting<sup>27,38,28</sup>. Therefore, our approach is scan-specific and does not require large-scale datasets for training. (2) Our approach explores the joint  $k$ - $q$ -space redundancy with the use of spatial-diffusion convolutions and is

also constrained by the physics-based data consistency. Therefore, our approach is versatile to downstream diffusion model analysis (e.g., DTI). (3) We observe that the ADMM unrolling model can be trained from one single multi-band slice and is generalizable to other "unseen" multi-band slices. This substantially reduces the required training time. Furthermore, given that unrolled reconstructions require much shorter inference time than conventional iterative regularized reconstructions such as compressed sensing, our approach is feasible for clinical translation.

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>24</sup>. Therefore, the utilization of navigator acquisition and advanced deep



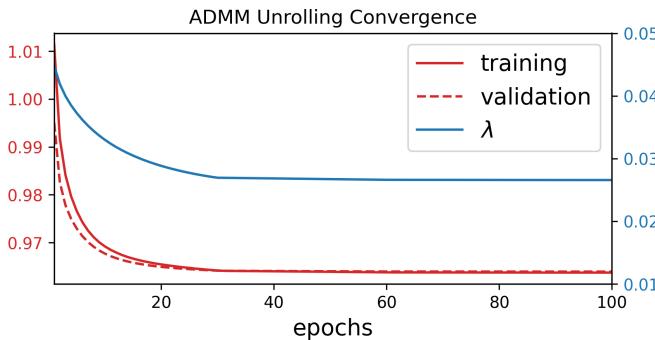
**FIGURE 8** Diffusion tensor imaging (DTI) model derived colored fractional anisotropy (cFA) maps based on the diffusion-weighted images as reconstructed by MUSE, LLR, and ADMM unrolling, respectively. Two slices are displayed in the top and the bottom row, respectively. Even with limited scan time (about 10 minutes at 0.7 mm spatial resolution) and limited diffusion directions (20), the proposed self-gated self-supervised ADMM unrolling reconstruction delivers clearer fiber orientations, as indicated by the maize-color arrows.

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>.

We observed that stripping-type motion artifacts occurred more frequently in the sub-millimeter isotropic resolution DWI regime. This makes sense, as scans with reduced slice thickness are more susceptible to shot-to-shot phase variations. In addition, sub-millimeter isotropic voxel resulted in higher noise in diffusion-weighted images. 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, because 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 does have several limitations. (1) 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>41</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>42</sup>. This approach does not require the



**FIGURE 9** Convergence analysis along the ADMM unrolling training and validation epochs. 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.

pre-determination of  $B_0$  field, but poses higher computational demand in the inversion course of phase increments from every echo. (2) As this work primarily focused on the development of selfgated self-supervised unrolled reconstruction for high-resolution DWI, only three volunteers were recruited. A pilot study with a large number of volunteers and even patients is beyond the scope of this work. (3) Given the small sample size, it is unlikely that we compare our proposed approach with other semi-self-supervised approaches such as SSDU<sup>38</sup>. Only MUSE and LLR are chosen for comparison in this work. However, we believe that MUSE and LLR are competitive and representative methods to be compared with, as the former one has already been translated to clinical practice and the latter one has been widely used for multi-contrast compressed sensing image reconstruction.

## 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 multi-band 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 and Dr. Xiaoqing Wang for insightful discussions. The 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 A Liebig 0000-0001-7342-3715
- Annika Hofmann 0009-0002-5723-1769
- 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

- 01 multiplexed sensitivity-encoding (MUSE). *NeuroImage*. 2013;72:41-47.
- 02
- 03 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.
- 04
- 05 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.
- 06
- 07 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.
- 08
- 09 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
- 11 10. Pruessmann Klaas P, Weiger Markus, Scheidegger Markus B, Boesiger Peter. SENSE: Sensitivity encoding for fast MRI. *Magn. Reson. Med.*. 1999;42:952-962.
- 12
- 13 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.
- 14
- 15 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.
- 16
- 17 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.
- 18
- 19 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.
- 20
- 21 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.
- 22
- 23 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.
- 24
- 25 17. Tian Qiyuan, Li Ziyu, Fan Qiuyun, et al. SDnDTI: Self-supervised deep learning-based denoising for diffusion tensor MRI. *NeuroImage*. 2023;253:119033.
- 26
- 27 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.
- 28
- 29 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.
- 30
- 31 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; 54:55-56.
- 32
- 33 21. Liang Zhi-Pei. Spatiotemporal imaging with partially separable functions. In: :988-991; 2007. 57:58-59.
- 34
- 35 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. 60:61-62.
- 36
- 37 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. 63:64-65.
- 38
- 39 24. 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. 67:68-70.
- 40
- 41 25. Cho Jaejin, Jun Yohan, Wang Xiaoqing, Kobayashi Caique, Bilgic Berkin. Improved Multi-shot Diffusion-Weighted MRI with Zero-Shot Self-supervised Learning Reconstruction. In: :457-466; 2023. 73:74-75.
- 42
- 43 26. He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. Deep residual learning for image recognition. In: :770-778; 2016. 76:77-78.
- 44
- 45 27. Quan Yuhui, Chen Mingqin, Pang Tongyao, Ji Hui. Self2Self With Dropout: Learning Self-Supervised Denoising From Single Image. In: ; 2020. 79:80.
- 46
- 47 28. Yaman Burhaneddin, Hosseini Seyed Amir Hossein, Akçakaya Mehmet. Zero-shot self-supervised learning for MRI reconstruction. In: ; 2022. 81:82-83.
- 48
- 49 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. 84:85-86.
- 50
- 51 30. Hinton G. E., Salakhutdinov R. R.. Reducing the dimensionality of data with neural networks. *Science*. 2006;313:504-507. 88:89.
- 52
- 53 31. Boyd Stephen, Parikh Neal, Chu Eric, Peleato Borja, Eckstein Jonathan. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*. 2010;3:1-122. 90:91-92.
- 54
- 55 32. Wang Shanshan, Su Zhenghang, Ying Leslie, et al. Accelerating magnetic resonance imaging via deep learning. In: :514-517; 2016. 95:96.
- 56
- 57 33. 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. 97:98-99.
- 58
- 59 34. 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. 101:102-103.
- 60
- 61 35. Uecker Martin, Karaus Alexander, Frahm Jens. Inverse reconstruction method for segmented multishot diffusion-weighted MRI with multiple coils. *Magn. Reson. Med.*. 104:105-106.

- 01 2009;62:1342-1348.
- 02 36. Merrem Andreas, Hofer Sabine, Hosseini Ali Seif Amir,  
03 et al. Diffusion-weighted MRI of the prostate without  
04 susceptibility artifacts: Undersampled multi-shot turbo-  
05 STEAM with rotated radial trajectories. *NMR Biomed.*.  
06 2019;32:e4074. 54
- 07 37. Kingma Diederik P, Ba Jimmy. ADAM: A Method for  
08 Stochastic Optimization. In: ; 2015. 55
- 09 38. Yaman Burhaneddin, Hosseini Seyed Amir Hossein,  
10 Moeller Steen, Ellermann Jutta, Uğurbil Kâmil,  
11 Akçakaya Mehmet. Self-supervised learning of physics-  
12 guided reconstruction neural networks without  
13 fully sampled reference data. *Magn. Reson. Med.*.  
14 2020;84:3172-3191. 61
- 15 39. Cordero-Grande Lucilio, Christiaens Daan, Hutter Jana,  
16 Price Anthony N, Hajnal Jo V. Complex diffusion-  
17 weighted image estimation via matrix recovery under  
18 general noise models. *NeuroImage*. 2019;200:391-404. 62
- 19 40. Pietsch Maximilian, Christiaens Daan, Hajnal Joseph V,  
20 Tournier J-Donald. dStripe: Slice artefact correction in  
21 diffusion MRI via constrained neural network. *Medical  
22 Image Analysis*. 2021;74:102255. 63
- 23 41. Zahneisen Benjamin, Aksoy Murat, Maclare Julian,  
24 Wuerslin Christian, Bammer Roland. Extended hybrid-  
25 space SENSE for EPI: Off-resonance and eddy current  
26 corrected joint interleaved blip-up/down reconstruction.  
27 *NeuroImage*. 2017;153:97-108. 64
- 28 42. Tan Zhengguo, Unterberg-Buchwald Christina, Blumen-  
29 thal Moritz, et al. Free-breathing liver fat,  $R_2^*$  and  $B_0$   
30 field mapping using multi-echo radial FLASH and reg-  
31 ularized model-based reconstruction. *IEEE Transactions  
32 on Medical Imaging*. 2023;42:1374-1387. 65
- 33 82
- 34 83
- 35 84
- 36 85
- 37 86
- 38 87
- 39 88
- 40 89
- 41 90
- 42 91
- 43 92
- 44 93
- 45 94
- 46 95
- 47 96
- 48 97
- 49 98
- 50 99
- 51 100
- 52 101
- 53 102
- 54 103
- 55 104
- 56 105
- 57 106