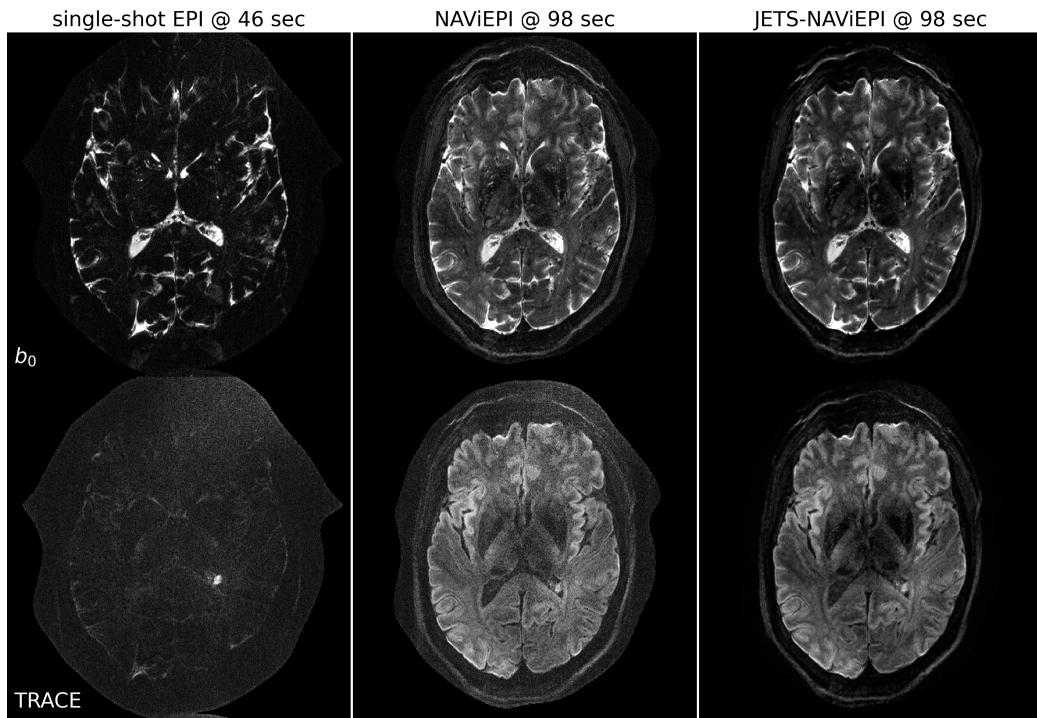


## Graphical Abstract

### Accelerated Diffusion Weighted Magnetic Resonance Imaging at 7 T: Joint Reconstruction for Shift-Encoded Navigator-based Interleaved Echo Planar Imaging (JETS-NAViEPI)

Zhengguo Tan, Patrick Alexander Liebig, Robin Martin Heidemann, Frederik Bernd Laun, Florian Knoll

**3-scan trace acquisition with voxel size 0.5 X 0.5 X 2.0 mm<sup>3</sup>**



## Highlights

### **Accelerated Diffusion Weighted Magnetic Resonance Imaging at 7 T: Joint Reconstruction for Shift-Encoded Navigator-based Interleaved Echo Planar Imaging (JETS-NAViEPI)**

Zhengguo Tan, Patrick Alexander Liebig, Robin Martin Heidemann, Frederik Bernd Laun, Florian Knoll

- Navigator-based interleaved EPI acquisition with minimal distortion mismatch between echoes
- Novel accelerated diffusion acquisition with shifted phase encoding among diffusion directions for complementary  $k$ - $q$ -space sampling at 7 T
- Generalized joint  $k$ - $q$ -slice diffusion-weighted image reconstruction with overlapping locally low-rank regularization
- Efficient simultaneous multi-slice (SMS) image reconstruction
- 3-scan trace acquisition with the voxel size  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$  and 60 slices at 1.5 min

# Accelerated Diffusion Weighted Magnetic Resonance Imaging at 7 T: Joint Reconstruction for Shift-Encoded Navigator-based Interleaved Echo Planar Imaging (JETS-NAViEPI)

Zhengguo Tan<sup>a</sup>, Patrick Alexander Liebig<sup>b</sup>, Robin Martin Heidemann<sup>b</sup>,  
Frederik Bernd Laun<sup>c</sup>, Florian Knoll<sup>a</sup>

<sup>a</sup>*Department Artificial Intelligence in Biomedical Engineering (AIBE),  
Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen, Germany*

<sup>b</sup>*Siemens Healthcare GmbH, Erlangen, Germany*

<sup>c</sup>*Institute of Radiology, University Hospital Erlangen, Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen, Germany*

---

## Abstract

The pursuit of high spatial-angular-temporal resolution for in vivo diffusion-weighted magnetic resonance imaging (DW-MRI) at ultra-high field strength (7 T and above) is important in understanding brain microstructure and function. Such pursuit, however, faces several technical challenges. First, increased off-resonance and shorter  $T_2$  relaxation require faster echo train readouts. Second, existing high-resolution DW-MRI techniques usually employ in-plane fully-sampled multi-shot EPI, which not only prolongs the scan time but also induces a high specific absorption rate (SAR) at 7 T. To address these challenges, we develop in this work navigator-based interleaved EPI (NAViEPI) which enforces the same effective echo spacing (ESP) between the imaging and the navigator echo. First, NAViEPI renders no distortion mismatch between the two echoes, and thus simplifies shot-to-shot phase variation correction. Second, NAViEPI allows for a large number of shots

(e.g.  $> 4$ ) with undersampled iEPI acquisition, thereby rendering clinically-feasible high-resolution sub-millimeter protocols. To retain signal-to-noise ratio (SNR) and to reduce undersampling artifacts, we developed the  $k_y$ -shift encoding among diffusion encodings to explore complementary  $k$ - $q$ -space sampling. Moreover, we developed a novel joint reconstruction with overlapping locally low-rank regularization generalized to the multi-band multi-shot acquisition at 7 T (dubbed JETS-NAViEPI). R1.6

*Keywords:* Diffusion-weighted magnetic resonance imaging, Echo planar imaging, Navigator, Ultra-high field, Joint reconstruction, Low rank, Simultaneous multi slice

---

<sup>1</sup> **1. Introduction**

<sup>2</sup> Diffusion-weighted magnetic resonance imaging (DW-MRI) ([Le Bihan et al., 1986; Merboldt et al., 1985](#)) is a non-invasive modality that is sensitive to the intravoxel Brownian motion of water molecules. DW-MRI forms the basis for diffusion tensor imaging (DTI) ([Basser et al., 1994; Mori et al., 1999](#)) and high angular resolution diffusion imaging (HARDI) ([Tuch et al., 2002](#)), and has been widely used in acute brain ischemia diagnosis, in tumor detection and staging, and in neuroscience ([Jones, 2010](#)).

<sup>9</sup> For DW-MRI acquisition, the commonly used pulse sequence is single-shot echo-planar imaging (SS-EPI) ([Mansfield, 1977](#)). SS-EPI is capable of rapidly acquiring one DW image per radio-frequency excitation at the order of 100 ms, and is thus motion robust. However, conventional SS-EPI, even with three-fold accelerated acquisition ([Bammer et al., 2001](#)) using parallel imaging ([Roemer et al., 1990; Ra and Rim, 1993; Pruessmann et al., 1999; Griswold et al., 2002](#)), still suffers from low spatial resolution and geometric distortions.

<sup>17</sup> In the quest for high spatial-angular-temporal-resolution and minimal-geometric-distortion DW-MRI, tremendous efforts have been made. Techniques on the correction of image distortions induced by off-resonances and eddy currents have been developed ([Andersson et al., 2003](#)). Furthermore, gSlider ([Setsompop et al., 2018](#)) with blipped-CAIPI ([Setsompop et al., 2012](#)) for simultaneous multi-slice (SMS) ([Maudsley, 1980; Breuer et al., 2005](#)) was proposed to achieve high-resolution DW-MRI. Advanced pulse sequences based on **multi-shot EPI** have also been developed, including but not limited to interleaved EPI (iEPI) ([Butts et al., 1993](#)), PROPELLER ([Pipe et al., 2002](#)), R3.7, R3.8 R3.15

26 and readout-segmented EPI (rsEPI) (Porter and Heidemann, 2009; Heide-  
27 mann et al., 2010).

28 Based on four-shot iEPI, multiplexed sensitivity encoding (MUSE) image  
29 reconstruction achieved DW-MRI with a sub-millimeter in-plane resolution  
30 and maximal  $b$ -value  $800 \text{ s/mm}^2$  at 3 T (Chen et al., 2013). The four-shot  
31 iEPI employed in MUSE acquired an in-plane fully-sampled  $k$ -space, except  
32 partial Fourier. Every shot (segment), corresponding to four-fold under-  
33 sampling, was then reconstructed via parallel imaging to obtain shot-to-shot  
34 phase variation. This indicates that increasing the number of shots in MUSE  
35 will result in higher undersampling per shot, and consequently, degrade shot  
36 phase estimation (Wu and Miller, 2017). On the other hand, the use of in-  
37 plane fully-sampled four-shot iEPI is challenging at ultra-high field (e.g. 7 T),  
38 because the SAR is linearly proportional to the square of the field strength.

39 Alternatively, navigator-based iEPI acquisition has been proposed (Jeong  
40 et al., 2013; Dai et al., 2017, 2018). These proposals may allow for a larger  
41 number of shots, and hence higher spatial resolution. However, due to the use  
42 of different ESP between the imaging echo and the navigator echo, these pro-  
43 posals suffered from geometric distortion mismatch between the two echoes  
44 and thus required specific compensation methods. In contrast, rsEPI (Porter  
45 and Heidemann, 2009; Heidemann et al., 2010) used the same readout seg-  
46 ment for both echoes, and thus required no correction of the navigator echo.

47 Beyond the MUSE-type parallel imaging reconstruction, compressed sens-  
48 ing (Lustig et al., 2007; Block et al., 2007) has been explored. For instance,  
49 multi-shot reconstruction techniques based on structured low-rank matrix  
50 completion (MUSSELS) (Mani et al., 2017; Bilgic et al., 2019) achieved 5-

shot DW-MRI with 9-fold undersampling per shot. Recently, JULEP (Dai et al., 2023) incorporated explicit phase mapping into MUSSELS. These reconstruction techniques, i.e., MUSE, MUSSELS and JULEP, targeted the reconstruction of one DW image from interleaved EPI acquisition, and did not explore joint- $k$ - $q$ -space undersampling or reconstruction.

Joint- $k$ - $q$ -space undersampling can be achieved via proper regularization along the diffusion encoding direction. Relevant examples are diffusion undersampling with Gaussian process estimated reconstruction (DAGER) (Wu et al., 2019) and magnitude-based spatial-angular locally low-rank regularization (SPA-LLR) (Hu et al., 2020). However, DAGER addressed the reconstruction problem of single-shot EPI acquisition. SPA-LLR focused on the reconstruction of single-band and fully-sampled iEPI acquisition.

In this work, we propose a Joint  $k$ - $q$ -slice rEconsTruction framework for Shift-encoded NAVigator-based interleaved EPI at 7 T (dubbed JETS-NAViEPI). Our pulse sequence, NAViEPI, differs from most existing techniques. First, NAViEPI builds upon interleaved EPI, thereby allowing for fast and efficient  $k$ -space coverage. Second, inspired by rsEPI, NAViEPI ensures the same effective ESP between the imaging and the navigator echo, thereby minimizing geometric distortion and allowing for the use of a larger number of shots. NAViEPI essentially integrates the advantages of both iEPI and rsEPI. Third, NAViEPI utilizes undersampled multi-shot iEPI, thereby alleviating the SAR problem at 7 T. Fourth, NAViEPI shifts the  $k$ -space in-plane sampling pattern along the phase encoding ( $k_y$ ) direction. This shifting creates complementary  $k$ - $q$ -space sampling, which leads to the possibility of our joint  $k$ - $q$ -slice reconstruction. Specifically, we employ spatial-diffusion

<sup>76</sup> overlapping LLR regularization to jointly reconstruct all diffusion encodings  
<sup>77</sup> and multi-band slices. In vivo experiments at 7 T and comparisons with other  
<sup>78</sup> techniques demonstrate the efficiency of our proposed method in achieving  
<sup>79</sup> high spatial resolution DW-MRI at ultra-high field.

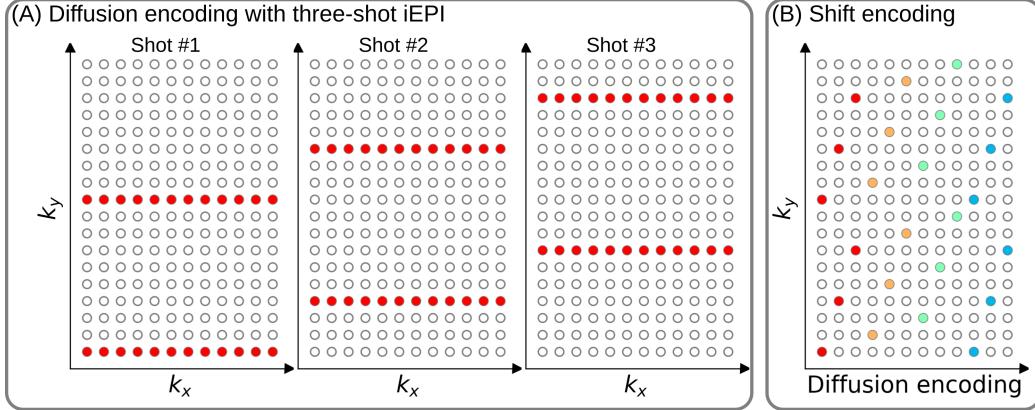


Figure 1: (A) An example DW-MRI acquisition with three-shot interleaved EPI acquisition. (B) The proposed  $k_y$  shifted diffusion encoding scheme. This example employs three shots per DW image. Therefore, every three columns have the same color.

## 80 2. Materials and methods

### 81 2.1. Multi-band shift-encoded iEPI acquisition

82 Fig. 1 (A) displays the diffusion-weighted image acquisition based on  
 83 three-shot interleaved EPI with three-fold in-plane undersampling. Conven-  
 84 tionally, such a sampling pattern is repeated for all diffusion directions. In  
 85 contrast, we propose the  $k_y$ -shifted diffusion encoding, as shown in Fig. 1 (B).  
 86 The interleaved EPI sampling pattern is shifted by one  $k_y$  line per diffusion  
 87 direction, with the cycling period being the in-plane undersampling factor.

88 It is worth noting that, as shown in Fig. 1 (A), the undersampling factor  
 89 of one segment is  $R_{\text{in-plane}} \times N_{\text{shot}}$  (ignore multi-band undersampling here),  
 90 yielding nine-fold in-plane undersampling in this example. In other words,  
 91 the undersampling factor per segment linearly scales up with the number  
 92 of shots. Consequently, conventional self-gating reconstruction techniques,  
 93 e.g. MUSE, suffer from degraded shot-to-shot phase estimation, which in

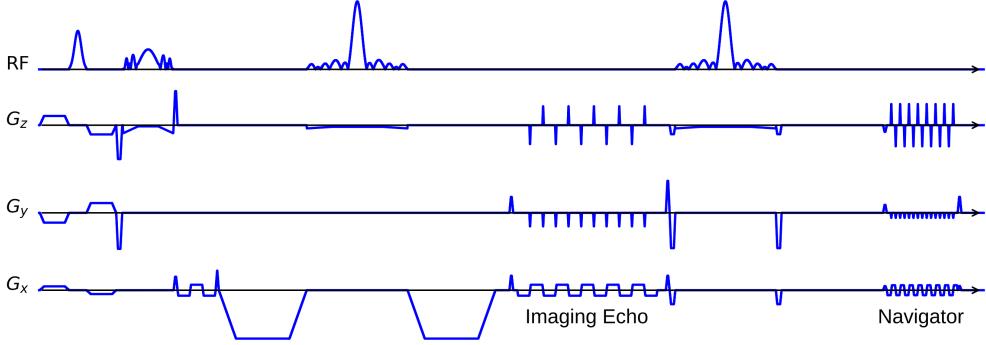


Figure 2: The NAViEPI sequence diagram. SMS is utilized for the acquisition of both imaging and navigator echoes. While the acceleration factor per navigator is the same as listed in Table 1, the acceleration factor per imaging echo is in addition linearly scaled by the number of shots.

94 turn limits the number of shots and spatial resolution.

95 *2.2. NAViEPI: Navigator-based iEPI with consistent effective ESP between  
96 the imaging and the navigator echo - where iEPI meets rsEPI*

97 Instead of the self-gated MUSE with in-plane fully-sampled iEPI and  
98 a limited number of shots, We propose NAVigator-based interleaved EPI  
99 (NAViEPI), as illustrated in Fig. 2. Moreover, inspired by rsEPI (Porter and  
100 Heidemann, 2009), NAViEPI enforces a consistent effective ESP between the  
101 imaging and the navigator echo, thereby minimizing distortion mismatch  
102 between the two echoes.

103 Since one imaging echo presents one segment in multi-shot EPI acquisi-  
104 tion, its effective ESP is defined as

$$\text{ESP}_{\text{eff}} = \frac{\text{ESP}}{R_{\text{in-plane}} \times N_{\text{shot}}} \quad (1)$$

105 Here, a larger number of shots (segments) increases the undersampling factor  
106 per segment (see Fig. 1), but decreases the effective ESP. Since the navigator

107 echo is acquired for each segment, its in-plane undersampling factor equals  
108  $R_{\text{in-plane}}$ . Therefore, the effective ESP of the navigator echo must match that  
109 of the imaging echo, as given in Eq. (1). With a matching effective ESP, the  
110 base resolution of the navigator echo can then be determined.

111 *2.3. In vivo acquisition protocols*

112 We implemented multiple in-vivo acquisition protocols at a clinical 7 T  
113 MR system (MAGNETOM Terra, Siemens Healthineers, Erlangen, Ger-  
114 many) equipped with a 32-channel head coil (Nova Medical, Wilmington,  
115 MA, USA) and the XR-gradient system (**maximum gradient strength 80 mT/m**  
116 **with a peak slew rate of 200 T/m/s**). To calibrate coil sensitivity maps, refer- R3.22  
117 ence scans employed a gradient-echo (GRE) sequence. Spectral fat saturation  
118 and mono-polar diffusion-encoding gradients were used. The phase-encoding  
119 direction was selected as anterior-to-posterior.

Table 1: NAViEPI acquisition protocols

Protocol	1.0 mm isotropic		sub-millimeter	
	#1	#2	#3	#4
Diffusion mode	MDDW <sup>1</sup>		3-scan trace	
Diffusion scheme			monopolar	
Diffusion direction	20	30	3	
<i>b</i> -value (s/mm <sup>2</sup> )			1000	
<i>b</i> <sub>0</sub>	0			1
Voxel size (mm <sup>3</sup> )		1.0	0.5 × 0.5 × 2.0	
Slices	141	124	60	
FOV (mm <sup>2</sup> )		200	220	
Base resolution		200	440	
Navigator	No	No	Yes	No
Shots	4	2	5	1
TR (ms)	7700	8500	4400	8000
TEs (ms)	67	56	58/95.1	143
ESP (ms)	1.02	1.06	1.52	1.48
Bandwidth (Hz/Pixel)		1086	758	
Partial Fourier		6/8		
Acceleration <sup>2</sup>	1 × 3	3 × 2	3 × 2	
TA (min) <sup>3</sup>	10 : 42	9 : 12	1 : 38	0 : 46

<sup>1</sup> MDDW: Multi-direction diffusion weighting;<sup>2</sup> Acceleration: Both in-plane and slice undersampling can be employed, denoted as ( $R_{\text{in-plane}} \times R_{\text{slice}}$ );<sup>3</sup> TA: Total acquisition time.

121 This study was approved by the local ethics committee. Three volunteers R3.1  
122 with informed consent obtained before scanning participated in this study.  
123 Detailed acquisition protocols are listed in Table 1.

124 *2.3.1. 20-diffusion-direction acquisition at 1.0 mm<sup>3</sup> isotropic resolution*

125 As listed in Table 1, Protocol #1 with six-shot iEPI and without in- R1.1,  
126 plane undersampling was implemented. This protocol represents the acquisi-  
127 tion scheme employed in many existing multi-shot reconstruction techniques,  
128 (e.g., MUSE, SPA-LLR, and JULEP). The acquired data from this protocol R1.1,  
129 served as ground truth. Different reconstruction methods, i.e., JETS, MUSE, R1.2  
130 and JULEP were compared.

131 We then retrospectively reduced the four-shot data to only one shot per R1.1  
132 diffusion encoding without and with the proposed  $k_y$  shifting to simulate  
133 three-fold in-plane undersampling. JETS reconstruction was performed on  
134 all data to validate the proposed  $k_y$ -shifted acquisition.

135 *2.3.2. 30-diffusion-direction acquisition at 1.0 mm<sup>3</sup> isotropic resolution*

136 Protocol #2 in Table 1 was implemented for diffusion tensor imaging R1.1  
137 (DTI) (Basser et al., 1994). Given the use of 30 diffusion encodings, we  
138 reduced the number of shots to two and switched off the navigator such as  
139 to shorten the total scan time. This is reasonable because two segments with  
140 three-fold in-plane undersampling correspond to six-fold undersampling per  
141 segment, which is acceptable for the self-gated phase estimation.

142 *2.3.3. 3-scan trace acquisition at 0.5 × 0.5 × 2.0 mm<sup>3</sup> voxel size*

143 As listed in Table 1, Protocol #3 was implemented based on NAViEPI R1.1  
144 with five shots per diffusion encoding. This protocol was compared against

<sub>145</sub> single-shot EPI (Protocol #4) with the same spatial resolution and acceleration,  
<sub>146</sub> such as to demonstrate the sampling efficiency of NAViEPI.

<sub>147</sub> *2.4. Forward modeling*

<sub>148</sub> Our proposed acquisition method yields multi-dimensional **multi-band**  
<sub>149</sub>  $k$ -space data  $\mathbf{y}_{c,q,s}$ , where  $c, q, s$  denotes the index of the coil sensitivity R1.25  
<sub>150</sub> map, the diffusion encoding, and the shot, respectively. **Acquisition model-**  
<sub>151</sub> **ing needs to consider several aspects.** R1.26

<sub>152</sub> First, the acquired  $k$ -space data  $\mathbf{y}$  is mapped from individual shot images  
<sub>153</sub>  $\mathbf{x}_{q,s,z}$  via the forward model,

$$\begin{aligned} \mathbf{y}_{c,q,s} &= \mathbf{P}_{q,s} \boldsymbol{\Sigma} \boldsymbol{\Theta}_z \mathbf{F} \mathbf{S}_c \mathbf{x}_{q,s,z} \\ \mathbf{y} &:= \mathbf{E}_1 \mathbf{x} \end{aligned} \quad (2)$$

<sub>154</sub> Here, the encoding matrix  $\mathbf{E}_1$  comprises a chain of linear operators. Every  
<sub>155</sub> shot image  $\mathbf{x}$  is point-wise multiplied by a set of coil sensitivity maps ( $\mathbf{S}$ ) and  
<sub>156</sub> Fourier transformed ( $\mathbf{F}$ ). The output is then point-wise multiplied by the  
<sub>157</sub> multi-slice phase map ( $\boldsymbol{\Theta}$ ) with  $z$  the slice index in simultaneously excited  
<sub>158</sub> slices. This operator shifts individual slice along the phase-encoding direction  
<sub>159</sub> via varying phase modulation (Breuer et al., 2005). The SMS  $k$ -space data  
<sub>160</sub> is then summed (collapsed,  $\boldsymbol{\Sigma}$ ) along the slice dimension and masked (point-  
<sub>161</sub> wise multiplied,  $\mathbf{P}$ ) by the sampling pattern of **each** diffusion encoding and  
<sub>162</sub> shot. R1.27

<sub>163</sub> Second, for diffusion MRI based on multi-shot EPI, **multiple shots ac-**  
<sub>164</sub> **quired for a given** diffusion encoding need to be combined as one DW image R1.28  
<sub>165</sub> ( $\tilde{\mathbf{x}}$ ). A possibility is to perform magnitude average (Chen et al., 2013) or R1.29

166 root-sum-squares (RSS) (Mani et al., 2017) of shot images. This method  
 167 is robust to in-plane motion, but sub-optimal concerning SNR (Guhaniyogi R3.27  
 168 et al., 2016). Alternatively, shot combination can be done via shot-to-shot R1.30  
 169 phase variation correction (Liu et al., 2005; Chen et al., 2013). This can  
 170 be incorporated to our formulation as point-wise multiplication between the R1.31  
 171 shot-to-shot phase variation ( $\Phi$ ) and the DW image ( $\tilde{\mathbf{x}}$ ),

$$\mathbf{x}_{q,s,z} = \Phi_{q,s,z} \tilde{\mathbf{x}}_{q,z} \quad (3)$$

172 Note that  $\tilde{\mathbf{x}}$  can be obtained by applying the adjoint of  $\Phi$  to  $\mathbf{x}$ . In MUSE,  
 173  $\Phi$  is obtained by parallel imaging reconstruction of all shots with subsequent  
 174 phase smoothing of every shot image. Based on this phase correction, the  
 175 complete forward model follows

$$\mathbf{y} := \mathbf{E}_2 \tilde{\mathbf{x}} = \mathbf{E}_1 \Phi \tilde{\mathbf{x}} \quad (4)$$

176 where the encoding matrix  $\mathbf{E}_2$  comprises the chain of the shot-to-shot phase  
 177 variation  $\Phi$  and the encoding matrix  $\mathbf{E}_1$ . We implemented these two encoding  
 178 matrices in SigPy (Ong and Lustig, 2019).

### 179 2.5. Joint $k$ - $q$ -slice reconstruction

180 Based on the generalized forward models in Eqs. (2) and (4), our proposed  
 181 joint  $k$ - $q$ -slice reconstruction can be formulated as a three-step approach.

182 **I. Navigator echo reconstruction.** The acquisition of navigator echoes  
 183 follows the forward model in Eq. (2), so the reconstruction of navigator  
 184 echoes can be formulated as:

$$\operatorname{argmin}_{\mathbf{x}} \|\mathbf{y} - \mathbf{E}_1 \mathbf{x}\|_2^2 + \lambda \mathbf{R}(\mathbf{x}) \quad (5)$$

185 where  $\mathbf{R}(\mathbf{x})$  denotes the regularization functional with the regularization  
 186 strength  $\lambda$ . In this work,  $\ell^2$  regularization was used, i.e.,  $\mathbf{R}(\mathbf{x}) =$   
 187  $\|\mathbf{x}\|_2^2$ . In the case of self-navigating (i.e., no navigator acquired) as Pro-  
 188 tocol #2, the central  $k$ -space region (i.e., 1/4 of the full image matrix)  
 189 of each segment is used as  $\mathbf{y}$  in Eq. (5).

190 **II. Iterative phase smoothing.** Shot-to-shot phase variation was ex-  
 191 tracted from the reconstructed navigator echo phases. Assuming that R1.13  
 192 phase images are spatially smooth (Chen et al., 2013; Dai et al., 2023),  
 193 we employed the iterative approach to smooth phase,

$$\mathbf{x}^{(k+1)} = \mathbf{F}^{-1} \mathcal{H} \mathbf{F} \mathbf{x}^{(k)} \quad (6)$$

194 where the index  $k$  denotes the iteration, and  $x^{(0)}$  is then the recon-  
 195 structed navigator image from Step I.  $\mathcal{H}$  is the Hanning window.

196 **III. Shot-combined reconstruction.** Joint reconstruction of all DW im-  
 197 ages using the shot-combined forward model  $\mathbf{E}_2$  with shot-to-shot phase  
 198 variation from Step II reads:

$$\operatorname{argmin}_{\tilde{\mathbf{x}}} \|\mathbf{y} - \mathbf{E}_2 \tilde{\mathbf{x}}\|_2^2 + \lambda \|\mathbf{T} \tilde{\mathbf{x}}\|_* \quad (7)$$

199 Here, LLR regularization was employed in the local spatial-diffusion  
 200 matrices, based on the theory of partially separable functions (Liang, R1.35  
 201 2007; Trzasko and Manduca, 2011; Zhang et al., 2015).  $\mathbf{T}$  represents a  
 202 linear operator that firstly slides a local patch window through all DW  
 203 images and then flattens every set of local patches to two-dimensional R1.36  
 204 (2D) matrices, comprising spatial and diffusion dimensions. The nuclear  
 205 norm regularization is enforced via singular value thresholding (SVT)

of all flattened 2D matrices (Cai et al., 2010). We implemented this regularization term as a proximal operator (Beck, 2017). R3.28

This work employed blipped-CAIPI SMS (Setsompop et al., 2012), in which spatially separated slices are simultaneously excited and acquired. Therefore, 2D instead of 3D patches were used to construct the spatial-diffusion matrices. R1.14

It has been reported that LLR is prone to checkerboard artifacts when  $\lambda$  is too large (Hu et al., 2020). We overcame this problem by utilizing overlapping blocks and providing an efficient implementation. If the blocks overlap,  $\mathbf{T}^H \mathbf{T}$  input  $\neq$  input, where  $\mathbf{T}^H$  denotes the Hermitian adjoint operator of  $\mathbf{T}$ . This is because overlapping values are summed in the output of  $\mathbf{T}^H$ . Our efficient implementation was to scale  $\mathbf{T}^H$  as  $(1/\text{divisor})\mathbf{T}^H$ , where the divisor matrix was obtained by  $\mathbf{T}^H \mathbf{T} \mathbf{1} \cdot \mathbf{1}$  denoted the matrix of all ones with the same shape as the input. R3.11, R3.30  
R3.31

## 2.6. Reconstruction

The acquired raw data was read in by twixtools (<https://github.com/pehses/twixtools>). Ramp-sampling regridding and FOV/2-ghost correction were also performed in twixtools. Subsequently, coil sensitivity maps were computed from reference scans using ESPIRiT (Uecker et al., 2014) in SigPy (Ong and Lustig, 2019).

With this pre-processing as well as the implemented forward models and proximal operator, the inverse problem in Eq. (7) was solved by the alternating direction method of multipliers (ADMM) (Boyd et al., 2010). R1.39

229 ADMM solves the minimization problems in an alternating update scheme,

$$\begin{cases} \mathbf{x}^{(k+1)} := \underset{\mathbf{x}}{\operatorname{argmin}} \| \mathbf{y} - \mathbf{E}(\mathbf{x}) \|^2 + \rho/2 \| \mathbf{T}\mathbf{x} - \mathbf{z}^{(k)} + \mathbf{u}^{(k)} \|_2^2 \\ \mathbf{z}^{(k+1)} := \mathcal{T}_{\lambda/\rho}(\mathbf{T}\mathbf{x}^{(k+1)} + \mathbf{u}^{(k)}) \\ \mathbf{u}^{(k+1)} := \mathbf{u}^{(k)} + \mathbf{T}\mathbf{x}^{(k+1)} - \mathbf{z}^{(k+1)} \end{cases} \quad (8)$$

230 where  $k$  denotes the ADMM iteration.  $\mathbf{z}$  is the auxiliary variable ( $\mathbf{z} = \mathbf{T}\mathbf{x}$ ),  
231 and  $\mathbf{u}$  is the Lagrangian multipliers. Importantly, when solving Eq. (2),  $\mathbf{x}$   
232 denotes shot images and  $\mathbf{E}$  denotes  $\mathbf{E}_1$  in Eq. (8). In contrast,  $\mathbf{x}$  denotes shot-  
233 combined images and  $\mathbf{E}$  denotes  $\mathbf{E}_2$  when solving Eq. (4).  $\mathbf{x}$  can be solved  
234 using linear least square algorithms, e.g. conjugate gradients (Hestenes and R3.35  
235 Stiefel, 1952), while  $\mathbf{z}$  is updated via singular value thresholding ( $\mathcal{T}$ ) with R2.8,  
236 the thresholding parameter  $\lambda/\rho$ . The coupling parameter  $\rho$  is effective in R3.32  
237 both the update of  $\mathbf{x}$  and  $\mathbf{z}$ . It acts as Tikhonov regularization strength  
238 when updating  $\mathbf{x}$ , but also inversely scales the thresholding strength when  
239 updating  $\mathbf{z}$ , as shown in Supporting Information Figures S1 and S2.

240 In this work, 15 ADMM iterations with  $\rho = 0.05$  and  $\lambda = 0.04$ , and a  
241 block size of 6 for LLR (refer to Supporting Information Figure S3) were  
242 used. All reconstructions were done on a single A100 SXM4/NVLink GPU  
243 with 40 GB memory (NVIDIA, Santa Clara, CA, USA). R3.37

244 We compared our proposed joint reconstruction with established multi-  
245 shot reconstruction techniques, specifically, MUSE (Chen et al., 2013) and  
246 JULEP (Dai et al., 2023), hosted on GitHub by Dr. Dai (Dai et al., 2023).  
247 Further, we performed the local-PCA denoising (Cordero-Grande et al., 2019)  
248 as implemented in MRtrix (Tournier et al., 2019) on the MUSE reconstructed  
249 complex DW images. R1.3

250 With reconstructed DW images from Protocol #2 in Table 1, color-coded

251 fractional anisotropy (cFA) maps ([Basser et al., 1994](#)) were fitted using DiPy  
252 ([Garyfallidis et al., 2014](#)),

253 **3. Results**

254 *3.1. Iterative smoothing of shot-to-shot phase variation*

255 Navigators were acquired with the acceleration rate as listed in Table 1.  
256 Besides, the base resolution of navigators (e.g. 32 in Protocol #3 in Table 1)  
257 was smaller than imaging echoes. As a result, reconstructed navigator phases  
258 (refer to the first column in Fig. 3) from Step I in Section 2.5 are not spatially  
259 smooth. Such phases, when used in the shot-combined reconstruction, result  
260 in signal void artifacts in DW images. To address this problem, we utilized  
261 the iterative smoothing procedure. As show in Fig. 3, the ripple-like phase  
262 artifact disapper after five iterations. It can also be seen that such iterative  
263 procedure retains the shot-to-shot phase variation.

264 *3.2. Comparison to MUSE and JULEP with four-shot iEPI acquisition*

265 The iterative phase smoothing was also applicable to MUSE-type self- R1.1,  
266 navigating reconstruction, where shot phases were reconstructed from imag- R1.2  
267 ing echoes. Fig. 4 compares our proposed JETS with MUSE (Chen et al.,  
268 2013), MUSE with complex-valued local-PCA denoiser (Cordero-Grande et al.,  
269 2019), and JULEP (Dai et al., 2023). The residual noise from MUSE can be  
270 largely removed by the denoiser. However, when compared to JETS, the de-  
271 noiser shows residual noise patterns within the globus pallidus (indicated by  
272 the red arrow). JETS also shows better denoising than JULEP. The reason  
273 is that JETS enforces spatial-diffusion regularization, whereas JULEP for-  
274 mulates structured low-rank regularization of the four shots for one diffusion  
275 encoding.

### Iterative smoothing of shot-to-shot phase variation

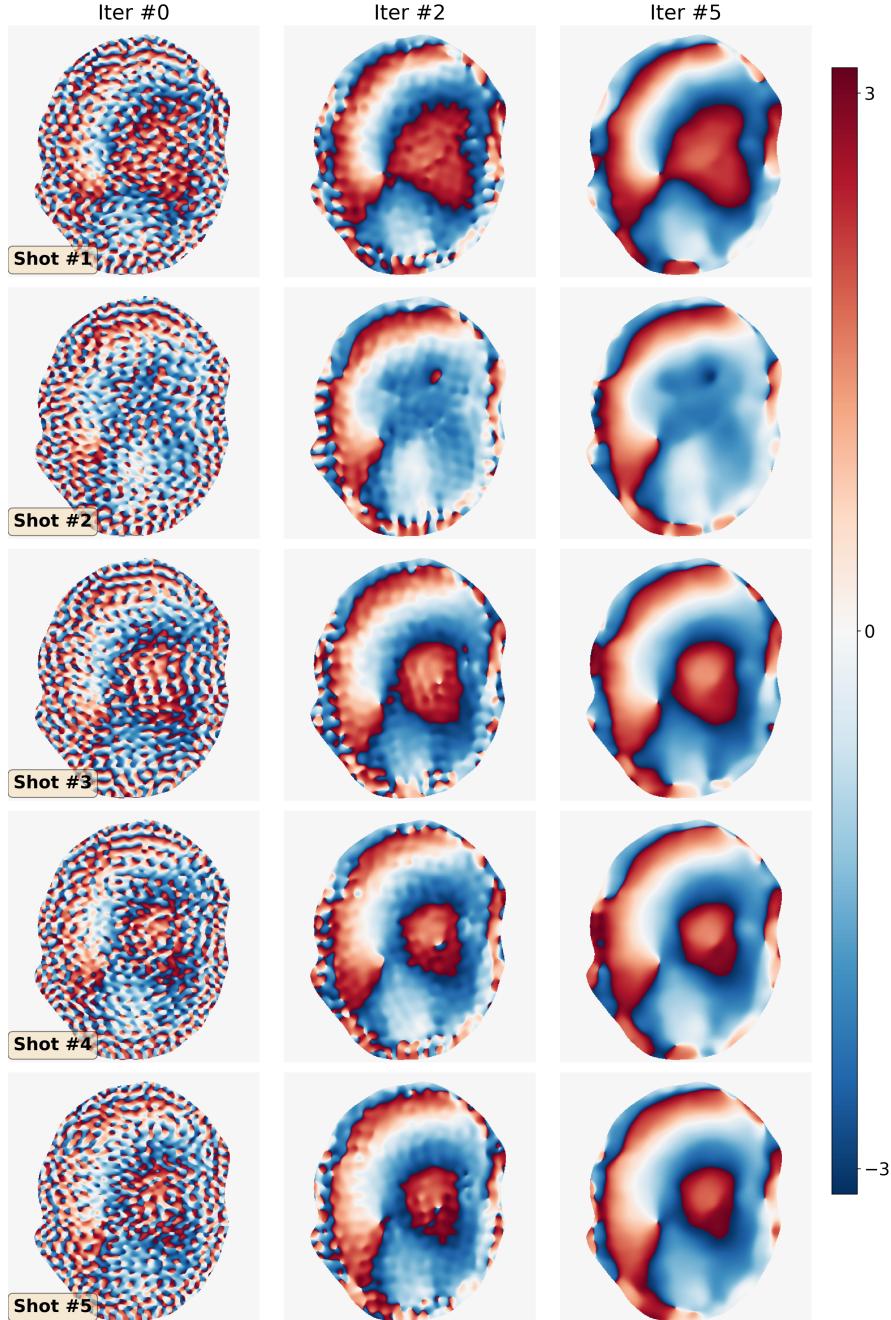


Figure 3: Iterative smoothing of shot-to-shot phase variation according to Eq. (6). Navigators from Protocol #3 were reconstructed based on Step I in Section 2.5 and then used as the input (iter #0, left column).

**8th DW image from 4-shot iEPI @ 1 mm ISO**

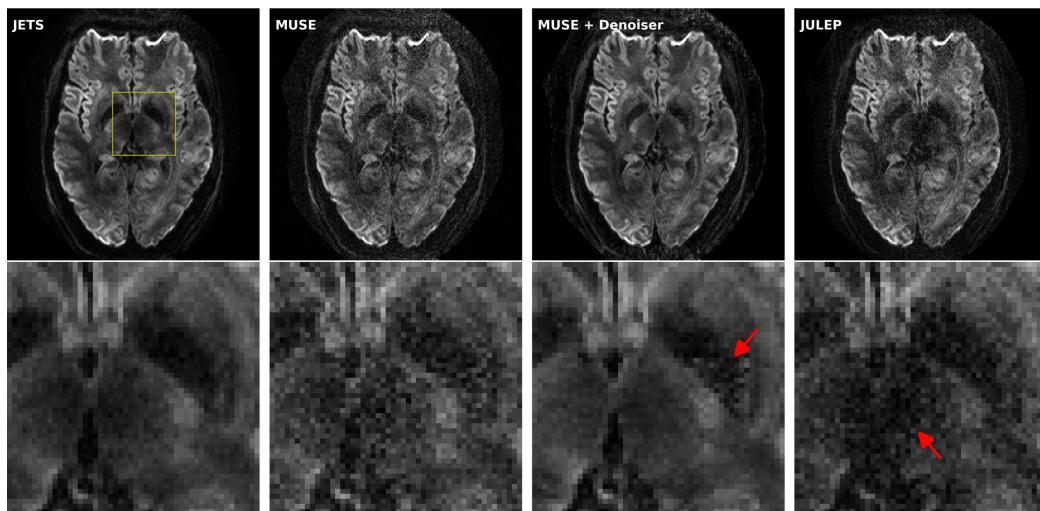


Figure 4: Reconstructed DW images (the 8th diffusion encoding) based on 4-shot iEPI acquisition with 1 mm isotropic resolution (Protocol #1 in Table 1). Four reconstruction methods are compared (from left to right): JETS, MUSE, MUSE with denoiser, and JULEP. The 2nd row displays the magnified views of the yellow square.

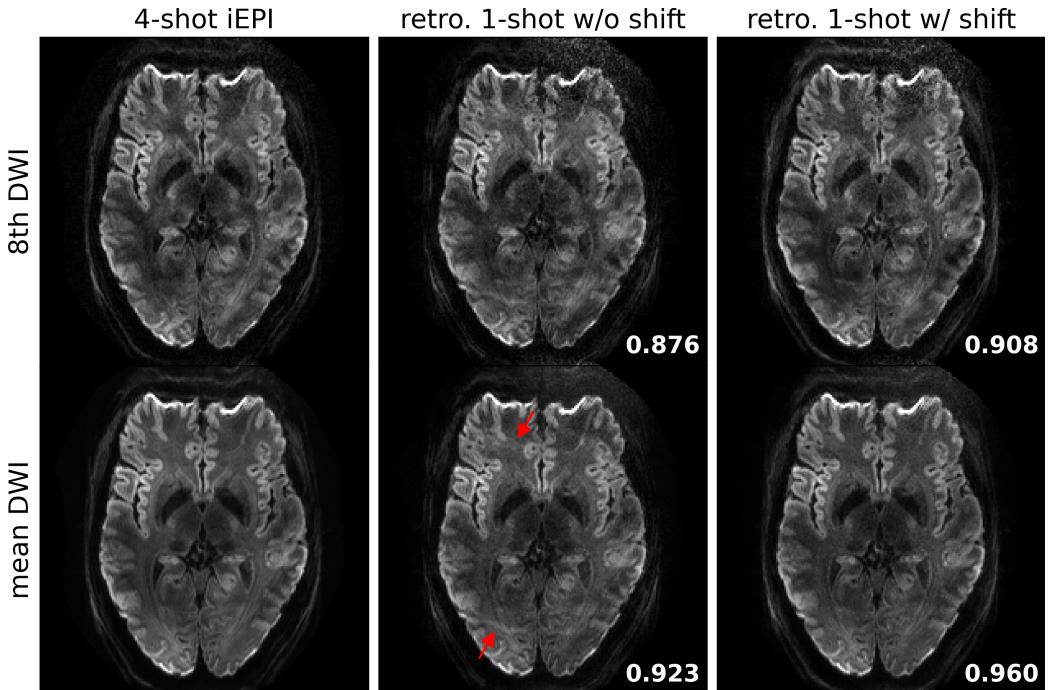


Figure 5: Quantitative validation of the proposed  $k_y$ -shift encoding sampling pattern based on 4-shot iEPI acquisition with 1 mm isotropic resolution (Protocol #1 in Table 1). (Top) the 8th diffusion encoding and (bottom) mean DWI over 20 diffusion encodings. (1st column) JETS reconstruction of 4-shot iEPI acquisition is used as the ground truth. The 2nd and the 3rd column displays JETS reconstruction of retrospectively undersampled 1-shot acquisition without and with  $k_y$  shifting, respectively. structural similarity (SSIM) values are computed and displayed in the bottom right corners.

276    3.3. *Retrospectively undersampling from the four-shot iEPI acquisition*

277    JETS reconstruction results on the four-shot prospectively fully-sampled R1.1  
278    data from Protocol #1 in Table 1, as well as on the retrospectively under-  
279    sampled one-shot data without and with the proposed  $k_y$  shift are displayed  
280    in Fig. 5. Residual aliasing artifacts are visible in the reconstruction without  
281     $k_y$  shifting, as pointed by the red arrows. On the contrary, the reconstruction  
282    with the proposed  $k_y$  shifting among diffusion encodings shows much reduced  
283    aliasing, reduced noise, and higher SSIM.

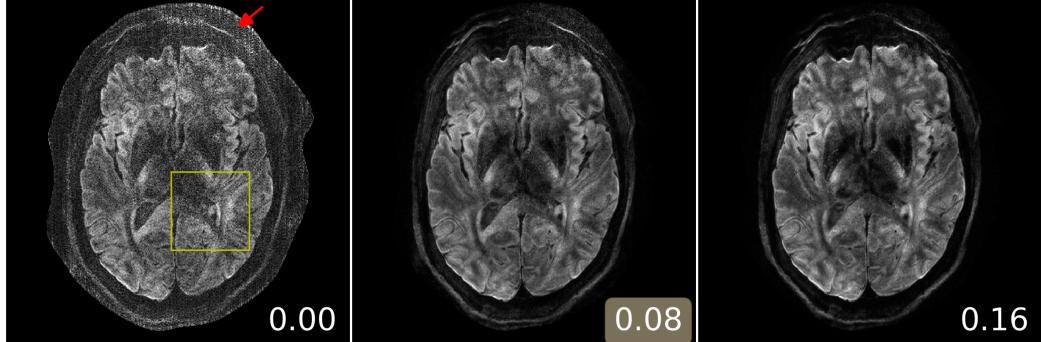
284    3.4. *Analysis of reconstruction parameters*

285    Here we provide a systematic analysis of the proposed JETS reconstruc- R1.4  
286    tion with LLR regularization applied to the spatial-diffusion dimension, as  
287    shown in Fig. 6.

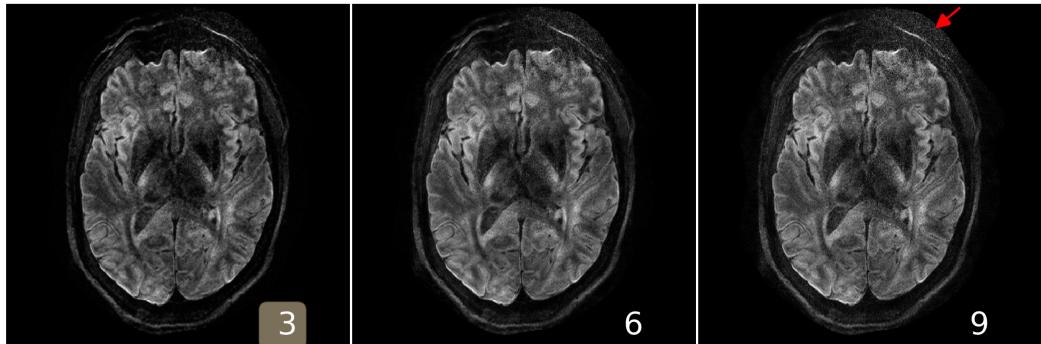
288    First, we varied the regularization strength  $\lambda$  from 0, to 0.08, and to  
289    0.16. The reconstruction with  $\lambda = 0$  in Eq. (7) corresponds to parallel  
290    imaging reconstruction without LLR regularization. It is worth noting that  
291    the proposed NAViEPI sequence demonstrates high-quality sub-millimeter  
292    DW images ( $0.5 \times 0.5 \times 2.0 \text{ mm}^3$  in this example). The DW images can be  
293    further improved with the use of LLR regularization, i.e., reduced noise, as  
294    seen in the reconstruction with  $\lambda = 0.08$ . Increasing  $\lambda$  (e.g. 0.16) further  
295    reduces noise, but at the cost of increased blurring. Therefore,  $\lambda = 0.08$  was  
296    selected in this work.

297    Second, besides the regularization strength, the block size (i.e., the width  
298    of square 2D patches) also plays a role in denoising. Here, the block size of 3  
299    shows the best denoising as compared to 6 and 9, especially in the peripheral  
300    brain region. According to (Cordero-Grande et al., 2019), it is suggested

**(A) varying  $\lambda$ , keeping block as 6 and stride as 1**



**(B) varying block size, keeping  $\lambda$  as 0.08 and stride as 1**



**(C) varying stride, keeping  $\lambda$  as 0.08 and block as 6**

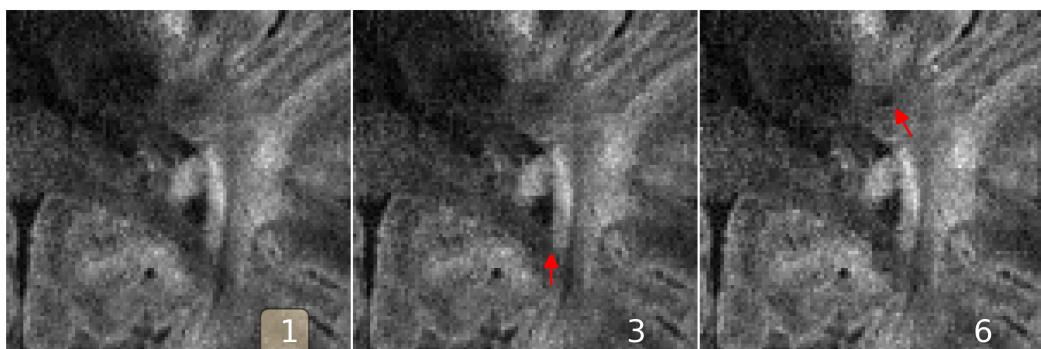


Figure 6: Analysis of reconstruction parameters based on the 3-scan trace acquisition with  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$  (Protocol #3 in Table 1). Displayed are JETS reconstructed single-direction DW images. **(A)** Varying the regularization strength  $\lambda$  from 0 to 0.08 and 0.16. **(B)** Varying the block size from 3 to 6 and 9. **(C)** Varying the stride size from 1 to 3 and 6 (non-overlapping).

301 to keep the patch size roughly equal to the diffusion encoding length. In  
302 this 3-scan trace acquisition example, the diffusion encoding length is 4 (1  $b_0$   
303 plus 3 orthogonal diffusion directions). Among the tested block sizes, 3 is the  
304 closest to 4, and hence renders better denoising, whereas the other block sizes  
305 may require stronger regularization strength. In practice, we also observed  
306 that smaller block sizes construct smaller matrices for SVT and thus lead to  
307 shorter computation time.

308 Third, we varied the stride, i.e., the increment from one local patch to  
309 the next. When the block size equals the stride (6 in this example), it refers  
310 to non-overlapping LLR, which is prone to checkerboard artifacts even with  
311 the use of random shifting ([Saucedo et al., 2017](#)) in each ADMM iteration,  
312 as indicated by the red arrows in the magnified views of Fig. 6. Therefore,  
313 we utilized the overlapping LLR with the stride of 1, which requires the  
314 implementation of the scaling matrix in Section 2.5 to avoid the summation  
315 of overlapped pixels. The drawback of overlapping LLR, however, is the  
316 increased number of matrices for SVT and hence the increased computation  
317 time ([Saucedo et al., 2017](#)).

318 *3.5. Sampling efficiency of NAViEPI*

319 As shown in Fig. 7, NAViEPI achieves sub-millimeter resolution (voxel R3.5  
320 size  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$ ) with the use of 5-shot acquisition. When compared  
321 to the single-shot acquisition with the same voxel size, the acquisition time  
322 of NAViEPI is about two times longer, but the image quality of NAViEPI is  
323 remarkably improved.

324 In the sub-millimeter imaging scenario, the increased base resolution re-  
325 quires longer TE (143 ms) in the single-shot acquisition, which results in

**3-scan trace acquisition with voxel size  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$**

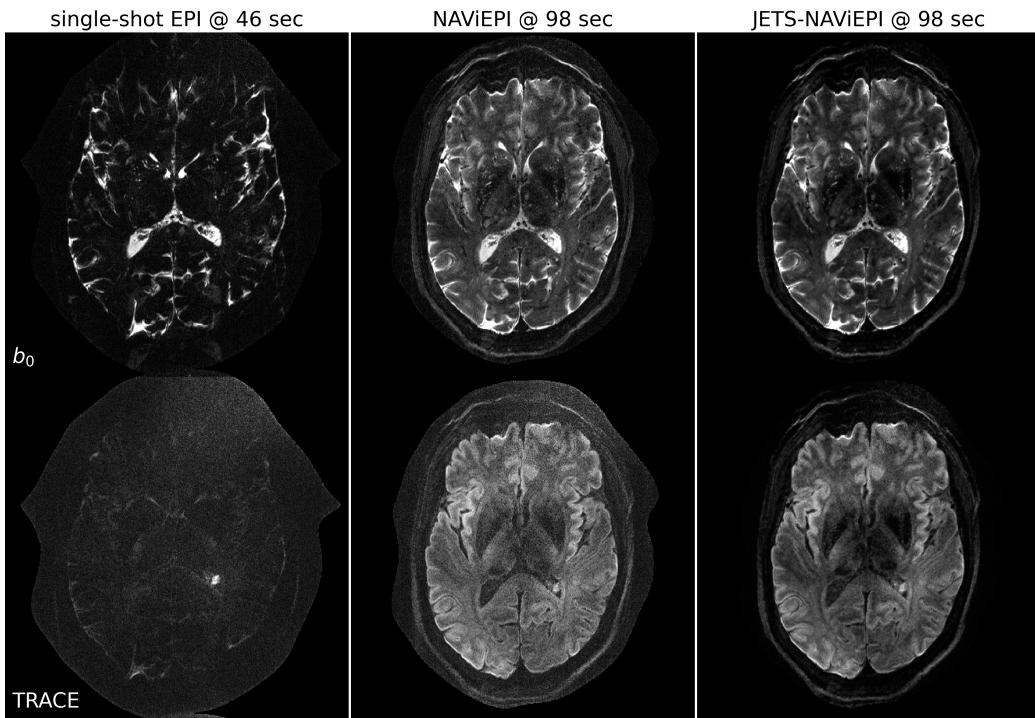


Figure 7: Sampling efficiency of the proposed NAViEPI sequence. 5-shot NAViEPI acquisition with the voxel size  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$  (Protocol #3) was compared with single-shot EPI acquisition (Protocol #4). Both the 1st and the 2nd columns were reconstructed via parallel imaging without LLR regularization, whereas the 3rd column was reconstructed via JETS.

**3-scan trace acquisition with voxel size  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$**

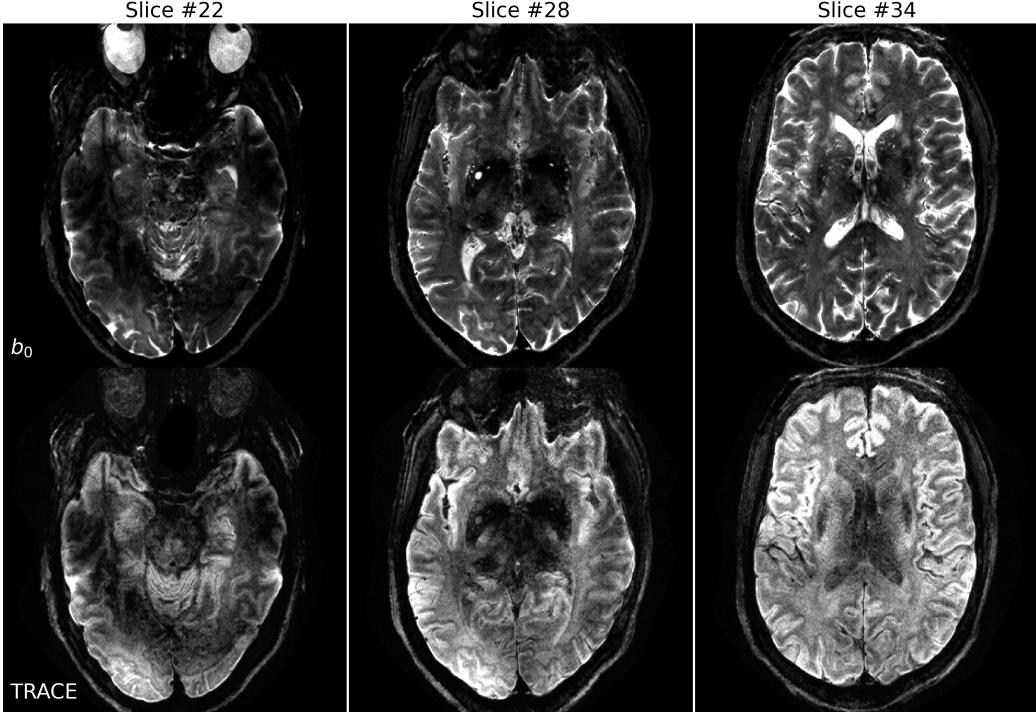


Figure 8: Reconstruction of the 3-scan trace acquisition with the voxel size  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$  (Protocol #3) at different slices.

326 significant signal loss due to  $T_2$  relaxation. Therefore, sub-millimeter DWI  
327 necessitates multi-shot acquisition, which is subject to shot-to-shot phase  
328 variation and long scan time. However, NAViEPI solves both challenges. The  
329 5-shot acquisition reduces TE to 58 ms, and thus retains SNR significantly  
330 compared to the single-shot acquisition. Moreover, the JETS reconstruction  
331 can help to reduce noise and improve structural visibility.

332 Fig. 8 shows the JETS reconstructed  $b_0$  and TRACE images in different R1.7  
333 slice locations. Admittedly, the lower brain region (e.g. slice #22) exhibits in-  
334 homogeneous and lower signal intensity than the upper slices. Such inhomogeneity

335 geneity can be alleviated with the use of multi-channel parallel transmission  
336 (Katscher et al., 2003; Grissom et al., 2010).

337 *3.6. Diffusion tensor imaging*

338 Since 30 diffusion encodings were acquired in Protocol #2, the block size  
339 in LLR regularization was lifted to 6, such that the spatial-diffusion matrix  
340 for SVT has similar width and height. The other parameter were kept the  
341 same as found in Fig. 6.

342 The mean DWIs in Fig. 9 illustrate high spatial resolution and high SNR.  
343 In line with Fig. 8, we can notice the signal loss in the cerebellum region, due  
344 to the use of single-channel transmission in this work. On the other hand,  
345 the reconstructed cFA maps in Fig. 9 show clear fiber orientation in all ori-  
346 entations. Moreover, tiny fiber structures can be visualized in the zoomed-in  
347 cFA maps. Because of the low signal sensitivity surrounding the cerebellum,  
348 residual artifacts are visible in the zoomed-in sagittal view. To enhance the  
349 DTI fitting performance, one possibility is to acquire more diffusion encod-  
350 ings.

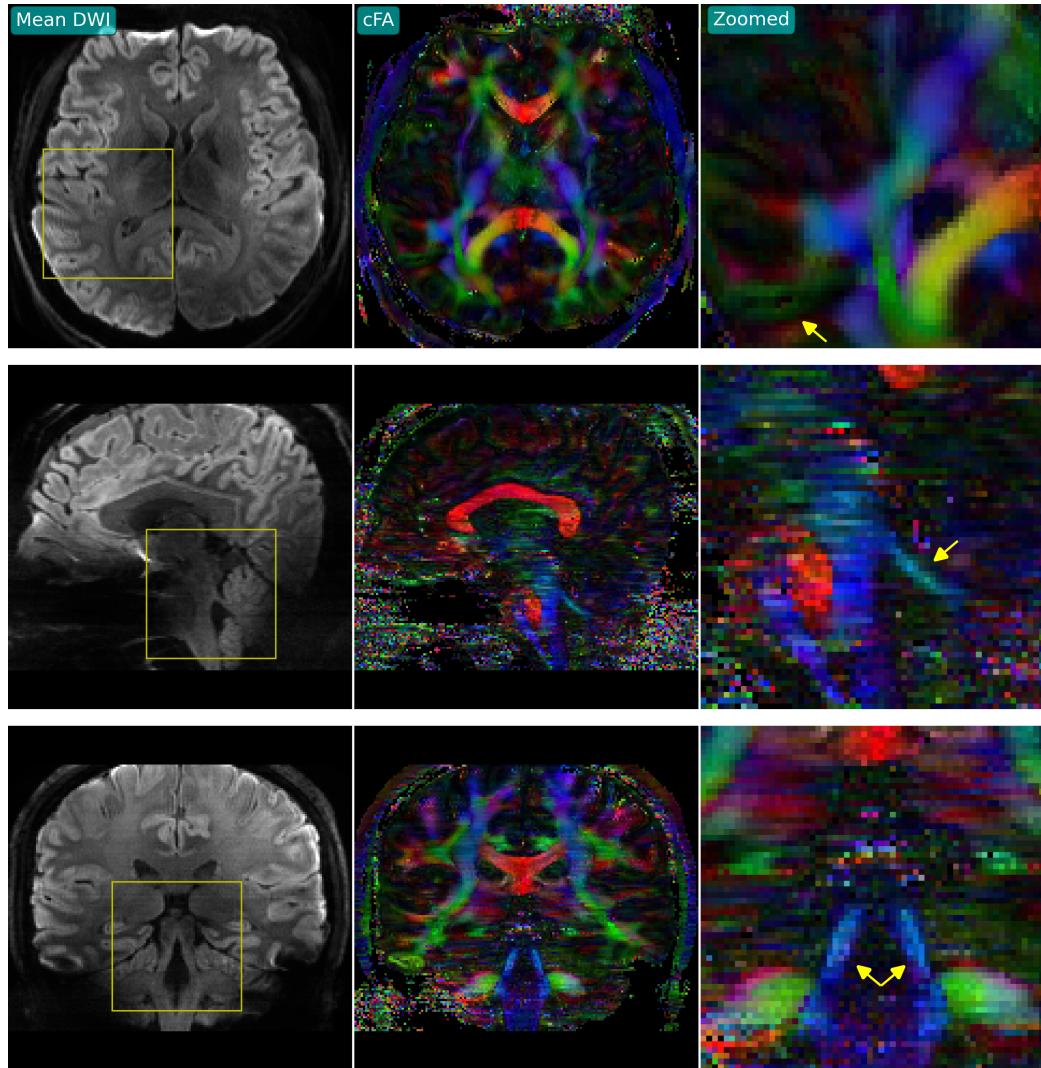


Figure 9: (Left) Mean DWI, (middle) cFA, and (right) cFA in the zoomed-in region based on JETS reconstructed DWI from Protocol #2. Three orthogonal slices (transversal, sagittal, and coronal) are displayed from top to bottom.

351    **4. Discussion**

352    This work reports a novel DW-MRI technique, JETS-NAViEPI. NAViEPI  
353    (1) achieves the fast and efficient acquisition of both imaging and navigator  
354    echoes, (2) enforces consistent effective ESP between the two echoes, and (3)  
355    allows for undersampled iEPI as well as a large number of shots. Moreover,  
356    compared to the single-shot acquisition, joint  $k$ - $q$ -slice reconstruction with  
357     $k_y$ -shift encoding on NAViEPI retains SNR and reduces aliasing artifacts  
358    in DW images. As a result, JETS-NAViEPI renders high spatiotemporal  
359    resolution diffusion MRI protocols in 7 T, e.g., 3-scan trace acquisition with  
360    the voxel size  $0.5 \times 0.5 \times 2.0 \text{ mm}^3$  at 1.5 min.

361    One limitation of JETS-NAViEPI is the long reconstruction time due to R3.48  
362    the simultaneous reconstruction of all DW images and the use of overlapping  
363    locally low-rank regularization. The reconstruction for the Protocol #3 in  
364    Table 1 on an A100 GPU takes about 2 min per multi-band slice. To reduce R3.48  
365    the computation time, coil compression algorithms (Buehrer et al., 2007;  
366    Huang et al., 2008) can be employed to reduce the number of coils for image  
367    reconstruction. Moreover, one may deploy multi-GPU distributed computing  
368    or modern optimization algorithms (e.g. stochastic gradient descent) (Ong  
369    et al., 2020) to speed up the reconstruction.

370    Neither the signal modeling in Eqs. (2) and (4) nor the LLR regulariza- R1.5  
371    tion considers the subject motion. In the presence of motion, the regularized  
372    reconstruction may degrade. To overcome this problem, scout-informed mo-  
373    tion estimation and reconstruction (Polak et al., 2022) could be integrated  
374    into the framework.

375    Another potential extension of this work is to incorporate distortion cor-

376 rection. The standard distortion correction method is known as TOPUP  
377 (Andersson et al., 2003), which acquires two scans with opposing phase-en-  
378 coding directions to obtain the field inhomogeneity map and then performs  
379 conjugate phase reconstruction to correct for distortion.

380 This work employed a single regularization weight  $\lambda$  to enforce low rank-  
381 ness along the spatial-diffusion direction. However, SNR may be heteroge-  
382 neous within the FOV. Therefore, one single regularization scalar may be  
383 inadequate to cover the whole FOV. Beyond this SVT-based machine-learn-  
384 ing reconstruction, one may seek to learn a  $q$ -space prior as the regularizer  
385 (Hammernik et al., 2018; Lam et al., 2019; Mani et al., 2021).

386 Although NAViEPI employs navigators for the acquisition of shot-to- R1.13  
387 shot phase variation, it is worth noting that phase behavior depends on  
388 several hard-to-control factors such as pulsatile motion, bulk motion, loca-  
389 tions within the brain, and diffusion sensitization strength. Therefore, more  
390 comprehensive modeling or post-processing such as image registration may  
391 be considered in future work.

392 While this work reconstructs all DW images and then performs model  
393 fitting, an alternative approach is to directly estimate  $b_0$  and diffusion tensors  
394 from measured  $k$ - $q$ -space data using model-based reconstruction (Knoll et al.,  
395 2015; Dong et al., 2018; Shafieizargar et al., 2023). Compared to DW image  
396 reconstruction, model-based reconstruction solves for a fewer number of un- R3.49  
397 knows, but requires strict diffusion tensor modeling and the use of nonlinear  
398 least square solvers.

399 **5. Conclusions**

400 We demonstrated the JETS-NAViEPI technique, which integrates a  $k_y$ -  
401 shifted encoding interleaved EPI sequence and a joint reconstruction with  
402 overlapping locally low-rank regularization for high spatial-angular-temporal  
403 resolution DW-MRI at 7 T. This technique allows for high-quality DW image  
404 reconstruction with accelerated acquisitions.

405 **Funding**

406 This work was supported by Deutsche Forschungsgemeinschaft (DFG) –  
407 Project Number 513220538, as well as National Institutes of Health (NIH)  
408 R01 EB024532 and P41 EB017183.

409 **Data and code available statement**

410 In the spirit of reproducible and open science, we publish our source  
411 code (<https://github.com/ZhengguoTan/sigpy>) as well as the raw  $k$ -space  
412 data (<https://doi.org/10.5281/zenodo.7548595>). We also provide inter-  
413 active demonstrations of the reconstruction procedure ([https://github.com/ZhengguoTan/demo\\_jets\\_diffusion\\_mri\\_7t](https://github.com/ZhengguoTan/demo_jets_diffusion_mri_7t)).  
414

415 **Acknowledgments**

416 The authors gratefully acknowledge the scientific support and HPC re-  
417 sources provided by the Erlangen National High Performance Computing  
418 Center (NHR@FAU) of Friedrich-Alexander-Universität Erlangen-Nürnberg  
419 (FAU) under the NHR project b143dc. NHR funding is provided by federal

420 and Bavarian state authorities. NHR@FAU hardware is partially funded by  
421 the German Research Foundation (DFG) – 440719683.

422 The authors thank Dr. Peter Neher for the discussion on MITK-Diffusion. R1.2  
423 The authors thank Dr. Berkin Bilgic for making the MUSSELS source code  
424 (<https://bit.ly/2QgBg9U>) publically available, Dr. Erpeng Dai for sharing  
425 the JULEP source code (<https://github.com/daiep/JULEP>) on GitHub,  
426 and Dr. Zhiyong Zhang for sharing the SPA-LLR source code (<https://github.com/ZZgroupSJTU/PMcmsDTI>) on GitHub. The authors also thank  
427 Dr. Philipp Ehses for the discussion on the use of twixtools (<https://github.com/pehses/twixtools>).  
428

430 **References**

431 Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct suscepti-  
432 bility distortions in spin-echo echo-planar images: application to diffusion  
433 tensor imaging. NeuroImage 20, 870–888. doi:[10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7).  
434

435 Bammer, R., Keeling, S.L., Augustin, M., Pruessmann, K.P., Wolf, R., Stoll-  
436 berger, R., Hartung, H.P., Fazekas, F., 2001. Improved diffusion-weighted  
437 single-shot echo-planar imaging (EPI) in stroke using sensitivity encoding  
438 (SENSE). Magn. Reson. Med. 46, 548–554. doi:[10.1002/mrm.1226](https://doi.org/10.1002/mrm.1226).

439 Basser, P.J., Mattiello, J., Le Bihan, D., 1994. MR diffusion tensor  
440 spectroscopy and imaging. Biophys. J. 66, 259–267. doi:[10.1016/S0006-3495\(94\)80775-1](https://doi.org/10.1016/S0006-3495(94)80775-1).  
441

- <sup>442</sup> Beck, A., 2017. First-order methods in optimization. Society for In-  
<sup>443</sup> dustrial and Applied Mathematics, Philadelphia, PA. doi:[10.1137/1.9781611974997](https://doi.org/10.1137/1.9781611974997).
- <sup>445</sup> Bilgic, B., Chatnuntawech, I., Manhard, M.K., Tian, Q., Liao, C., Iyer, S.S.,  
<sup>446</sup> Cauley, S.F., Huang, S.Y., Polimeni, J.R., Wald, L.L., Setsompop, K.,  
<sup>447</sup> 2019. Highly accelerated multishot echo planar imaging through synergistic  
<sup>448</sup> machine learning and joint reconstruction. Magn. Reson. Med. 82, 1343–  
<sup>449</sup> 1358. doi:[10.1002/mrm.27813](https://doi.org/10.1002/mrm.27813).
- <sup>450</sup> Block, K.T., Uecker, M., Frahm, J., 2007. Undersampled radial MRI with  
<sup>451</sup> multiple coils. Iterative image reconstruction using a total variation con-  
<sup>452</sup> straint. Magn. Reson. Med. 57, 1186–1098. doi:[10.1002/mrm.21236](https://doi.org/10.1002/mrm.21236).
- <sup>453</sup> Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2010. Distributed  
<sup>454</sup> optimization and statistical learning via the alternating direction method  
<sup>455</sup> of multipliers. Foundations and Trends in Machine Learning 3, 1–122.  
<sup>456</sup> doi:[10.1561/2200000016](https://doi.org/10.1561/2200000016).
- <sup>457</sup> Breuer, F.A., Blaimer, M., Heidemann, R.M., Mueller, M.F., Griswold, M.A.,  
<sup>458</sup> Jakob, P.M., 2005. Controlled aliasing in parallel imaging results in higher  
<sup>459</sup> acceleration (CAPIRINHA) for multi-slice imaging. Magn. Reson. Med.  
<sup>460</sup> 53, 684–691. doi:[10.1002/mrm.20401](https://doi.org/10.1002/mrm.20401).
- <sup>461</sup> Buehrer, M., Pruessmann, K.P., Boesiger, P., Kozerke, S., 2007. Array com-  
<sup>462</sup> pression for MRI with large coil arrays. Magn. Reson. Med 57, 1131–1139.  
<sup>463</sup> doi:[10.1002/mrm.21237](https://doi.org/10.1002/mrm.21237).

- 464 Butts, K., Riederer, S.J., Ehman, R.L., Thompson, R.M., Jack, C.R., 1993.  
465 Interleaved echo planar imaging on a standard MRI system. Magn. Reson.  
466 Med. 31, 67–72. doi:[10.1002/mrm.1910310111](https://doi.org/10.1002/mrm.1910310111).
- 467 Cai, J.F., Candès, E.J., Shen, Z., 2010. A singular value thresholding  
468 algorithm for matrix completion. SIAM. J. Optim. 20, 1956–1982.  
469 doi:[10.1137/080738970](https://doi.org/10.1137/080738970).
- 470 Chen, N.K., Guidon, A., Chang, H.C., Song, A.W., 2013. A robust multi-  
471 shot scan strategy for high-resolution diffusion weighted MRI enabled by  
472 multiplexed sensitivity-encoding (MUSE). NeuroImage 72, 41–47. doi:[10.1016/j.neuroimage.2013.01.038](https://doi.org/10.1016/j.neuroimage.2013.01.038).
- 473
- 474 Cordero-Grande, L., Christiaens, D., Hutter, J., Price, A.N., Hajnal, J.V.,  
475 2019. Complex diffusion-weighted image estimation via matrix recovery  
476 under general noise models. NeuroImage 200, 391–404. doi:[10.1016/j.neuroimage.2019.06.039](https://doi.org/10.1016/j.neuroimage.2019.06.039).
- 477
- 478 Dai, E., Ma, X., Zhang, Z., Yuan, C., Guo, H., 2017. Simultaneous multislice  
479 accelerated interleaved EPI DWI using generalized blipped-CAIPI acqui-  
480 sition and 3D K-space reconstruction. Magn. Reson. Med. 77, 1593–1605.  
481 doi:[10.1002/mrm.26249](https://doi.org/10.1002/mrm.26249).
- 482
- 483 Dai, E., Mani, M., McNab, J.A., 2023. Multi-band multi-shot diffu-  
484 sion MRI reconstruction with joint usage of structured low-rank con-  
485 straints and explicit phase mapping. Magn. Reson. Med. 89, 95–111.  
486 doi:[10.1002/mrm.29422](https://doi.org/10.1002/mrm.29422).

- 486 Dai, E., Zhang, Z., Ma, X., Dong, Z., Li, X., Xiong, Y., Yuan, C., Guo, H.,  
487 2018. The effects of navigator distortion and noise level on interleaved EPI  
488 DWI reconstruction: a comparison between image- and *k*-space method.  
489 Magn. Reson. Med. 80, 2024–2032. doi:[10.1002/mrm.27190](https://doi.org/10.1002/mrm.27190).
- 490 Dong, Z., Dai, E., Wang, F., Zhang, Z., Ma, X., Yuan, C., Guo, H., 2018.  
491 Model-based reconstruction for simultaneous multislice and parallel imag-  
492 ing accelerated multishot diffusion tensor imaging. Med. Phys. 45, 3196–  
493 3204. doi:[10.1002/mp.12974](https://doi.org/10.1002/mp.12974).
- 494 Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., van der Walt, S.,  
495 Descoteaux, M., Nimmo-Smith, I., Contributors, D., 2014. DIPY, a library  
496 for the analysis of diffusion MRI data. Front. Neuroinform. 8, 1–17. doi:[10.3389/fninf.2014.00008](https://doi.org/10.3389/fninf.2014.00008).
- 498 Grissom, W.A., Sacolick, L., Vogel, M.W., 2010. Improving high-field MRI  
499 using parallel excitation. Imaging Med. 2, 675–693. doi:[10.2217/IIM.10.62](https://doi.org/10.2217/IIM.10.62).
- 501 Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V.,  
502 Wang, J., Kiefer, B., Haase, A., 2002. Generalized autocalibrating par-  
503 tially parallel acquisitions (GRAPPA). Magn. Reson. Med. 47, 1202–1210.  
504 doi:[10.1002/mrm.10171](https://doi.org/10.1002/mrm.10171).
- 505 Guhaniyogi, S., Chu, M.L., Chang, H.C., Song, A.W., Chen, N.K., 2016. Mo-  
506 tion immune diffusion imaging using augmented MUSE for high-resolution  
507 multi-shot EPI. Magn. Reson. Med. 75, 639–652. doi:[10.1002/mrm.25624](https://doi.org/10.1002/mrm.25624).

- 508 Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock,  
509 T., Knoll, F., 2018. Learning a variational network for reconstruction of  
510 accelerated MRI data. *Magn. Reson. Med.* 79, 3055–3071. doi:[10.1002/mrm.26977](https://doi.org/10.1002/mrm.26977).
- 512 Heidemann, R.M., Porter, D.A., Anwander, A., Feiweier, T., Heberlein, K.,  
513 Knösche, T.R., Turner, R., 2010. Diffusion imaging in humans at 7 T  
514 using readout-segmented EPI and GRAPPA. *Magn. Reson. Med.* 64, 9–  
515 14. doi:[10.1002/mrm.22480](https://doi.org/10.1002/mrm.22480).
- 516 Hestenes, M.R., Stiefel, E., 1952. Methods of conjugate gradients for solving  
517 linear systems. *J. Res. Natl. Bur. Stand.* 49, 409–436. doi:[10.6028/jres.049.044](https://doi.org/10.6028/jres.049.044).
- 519 Hu, Y., Wang, X., Tian, Q., Yang, G., Daniel, B., McNab, J., Hargreaves, B.,  
520 2020. Multi-shot diffusion-weighted MRI reconstruction with magnitude-  
521 based spatial-angular locally low-rank regularization (SPA-LLR). *Magn.*  
522 *Reson. Med.* 83, 1596–1607. doi:[10.1002/mrm.28025](https://doi.org/10.1002/mrm.28025).
- 523 Huang, F., Vijayakumar, S., Li, Y., Hertel, S., Duensing, G.R., 2008. A soft-  
524 ware channel compression technique for faster reconstruction with many  
525 channels. *Magn. Reson. Imaging* 26, 133–141. doi:[10.1016/j.mri.2007.04.010](https://doi.org/10.1016/j.mri.2007.04.010).
- 527 Jeong, H.K., Gore, J.C., Anderson, A.W., 2013. High-resolution human  
528 diffusion tensor imaging using 2-D navigated multishot SENSE EPI at 7  
529 T. *Magn. Reson. Med.* 69, 793–802. doi:[10.1002/mrm.24320](https://doi.org/10.1002/mrm.24320).

- 530 Jones, D.K., 2010. Diffusion MRI: Theory, methods, and applications. Oxford  
531 University Press. doi:[10.1093/med/9780195369779.001.0001](https://doi.org/10.1093/med/9780195369779.001.0001).
- 532 Katscher, U., Börnert, P., Leussler, C., van den Brink, J.S., 2003. Transmit  
533 SENSE. Magn. Reson. Med. 49, 144–150. doi:[10.1002/mrm.10353](https://doi.org/10.1002/mrm.10353).
- 534 Knoll, F., Raya, J.G., Halloran, R.O., Baete, S., Sigmund, E., Bammer, R.,  
535 Block, T., Otazo, R., Sodickson, D.K., 2015. A model-based reconstruction  
536 for undersampled radial spin-echo DTI with variational penalties on the  
537 diffusion tensor. NMR Biomed 28, 353–366. doi:[10.1002/nbm.3258](https://doi.org/10.1002/nbm.3258).
- 538 Lam, F., Li, Y., Peng, X., 2019. Constrained magnetic resonance spectro-  
539 scopic imaging by learning nonlinear low-dimensional models. IEEE Trans.  
540 Med. Imaging 39, 545–555. doi:[10.1109/TMI.2019.2930586](https://doi.org/10.1109/TMI.2019.2930586).
- 541 Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., Laval-  
542 Jeantet, M., 1986. MR imaging of intravoxel incoherent motions: appli-  
543 cation to diffusion and perfusion in neurologic disorders. Radiology 161,  
544 401–407. doi:[10.1148/radiology.161.2.3763909](https://doi.org/10.1148/radiology.161.2.3763909).
- 545 Liang, Z.P., 2007. Spatiotemporal imaging with partially separable functions,  
546 in: 2007 4th IEEE International Symposium on Biomedical Imaging: From  
547 Nano to Macro, pp. 988–991. doi:[10.1109/ISBI.2007.357020](https://doi.org/10.1109/ISBI.2007.357020).
- 548 Liu, C., Moseley, M.E., Bammer, R., 2005. Simultaneous phase cor-  
549 rection and SENSE reconstruction for navigated multi-shot DWI with  
550 non-Cartesian  $k$ -space sampling. Magn. Reson. Med. 54, 1412–1422.  
551 doi:[10.1002/mrm.20706](https://doi.org/10.1002/mrm.20706).

- 552 Lustig, M., Donoho, D., Pauly, J.M., 2007. Sparse MRI: The application of  
553 compressed sensing for rapid MR imaging. Magn. Reson. Med. 58, 1182–  
554 1195. doi:[10.1002/mrm.21391](https://doi.org/10.1002/mrm.21391).
- 555 Mani, M., Jacob, M., Kelley, D., Magnotta, V., 2017. Multi-shot sensitivity-  
556 encoded diffusion data recovery using structured low-rank matrix comple-  
557 tion (MUSSELS). Magn. Reson. Med. 78, 494–507. doi:[10.1002/mrm.26382](https://doi.org/10.1002/mrm.26382).
- 559 Mani, M., Magnotta, V.A., Jacob, M., 2021. qModeL: A plug-and-play  
560 model-based reconstruction for highly accelerated multi-shot diffusion MRI  
561 using learned priors. Magn. Reson. Med. 86, 835–851. doi:[10.1002/mrm.28756](https://doi.org/10.1002/mrm.28756).
- 563 Mansfield, P., 1977. Multi-planar image formation using NMR spin echoes.  
564 J Phys C 10, 55–58. doi:[10.1088/0022-3719/10/3/004](https://doi.org/10.1088/0022-3719/10/3/004).
- 565 Maudsley, A.A., 1980. Multiple-line-scanning spin density imaging. J. Magn.  
566 Reson. 41, 112–126. doi:[10.1016/0022-2364\(80\)90207-3](https://doi.org/10.1016/0022-2364(80)90207-3).
- 567 Merboldt, K.D., Hanicke, W., Frahm, J., 1985. Self-diffusion NMR imag-  
568 ing using stimulated echoes. J. Magn. Reson. 64, 479–486. doi:[10.1016/0022-2364\(85\)90111-8](https://doi.org/10.1016/0022-2364(85)90111-8).
- 570 Mori, S., Crain, B.J., Chacko, V.P., Van Zijl, P.C.M., 1999. Three-  
571 dimensional tracking of axonal projections in the brain by mag-  
572 netic resonance imaging. Ann. Neurol. 45, 265–269. doi:[10.1002/1531-8249\(199902\)45:2<265::AID-ANA21>3.0.CO;2-3](https://doi.org/10.1002/1531-8249(199902)45:2<265::AID-ANA21>3.0.CO;2-3).

- 574 Ong, F., Lustig, M., 2019. SigPy: A Python package for high performance  
575 iterative reconstruction, in: Proceedings of the 27th Annual Meeting of  
576 ISMRM, Montréal, CAN, p. 4819. doi:[10.5281/zenodo.5893788](https://doi.org/10.5281/zenodo.5893788).
- 577 Ong, F., Zhu, X., Cheng, J.Y., Johnson, K.M., Larson, P.E.Z., Vasanawala,  
578 S.S., Lustig, M., 2020. Extreme MRI: Large-scale volumetric dynamic  
579 imaging from continuous non-gated acquisitions. Magn. Reson. Med. 84,  
580 1763–1780. doi:[10.1002/mrm.28235](https://doi.org/10.1002/mrm.28235).
- 581 Pipe, J.G., Farthing, V.G., Forbes, K.P., 2002. Multishot diffusion-weighted  
582 FSE using PROPELLER MRI. Magn. Reson. Med. 47, 42–52. doi:[10.1002/mrm.10014](https://doi.org/10.1002/mrm.10014).
- 583 Polak, D., Splitthoff, D.N., Clifford, B., Lo, W.C., Huang, S.Y., Conklin, J.,  
584 Wald, L.L., Setsompop, K., Cauley, S., 2022. Scout accelerated motion  
585 estimation and reduction (SAMER). Magn. Reson. Med. 87, 163–178.  
586 doi:[10.1002/mrm.28971](https://doi.org/10.1002/mrm.28971).
- 587 Porter, D.A., Heidemann, R.M., 2009. High resolution diffusion-weighted  
588 imaging using readout-segmented echo-planar imaging, parallel imaging  
589 and a two-dimensional navigator-based reacquisition. Magn. Reson. Med.  
590 62, 468–475. doi:[10.1002/mrm.22024](https://doi.org/10.1002/mrm.22024).
- 591 Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., 1999.  
592 SENSE: Sensitivity encoding for fast MRI. Magn. Reson. Med. 42, 952–  
593 962. doi:[10.1002/\(SICI\)1522-2594\(199911\)42:5<952::AID-MRM16>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1522-2594(199911)42:5<952::AID-MRM16>3.0.CO;2-S).

- 596 Ra, J.B., Rim, C.Y., 1993. Fast imaging using subencoding data sets from  
597 multiple detectors. Magn. Reson. Med. 30, 142–145. doi:[10.1002/mrm.1910300123](https://doi.org/10.1002/mrm.1910300123).
- 599 Roemer, P.B., Edelstein, W.A., Hayes, C.E., Souza, S.P., Mueller, O.M.,  
600 1990. The NMR phased array. Magn. Reson. Med. 16, 192–225. doi:[10.1002/mrm.1910160203](https://doi.org/10.1002/mrm.1910160203).
- 602 Saucedo, A., Lefkimiatis, S., Rangwala, N., Sung, K., 2017. Improved com-  
603 putational efficiency of locally low rank MRI reconstruction using iterative  
604 random patch adjustments. IEEE Trans. Med. Imaging 36, 1209–1220.  
605 doi:[10.1109/TMI.2017.2659742](https://doi.org/10.1109/TMI.2017.2659742).
- 606 Setsompop, K., Fan, Q., Stockmann, J., Bilgic, B., Huang, S., Cauley, S.F.,  
607 Nummenmaa, A., Wang, F., Rathi, Y., Witzel, T., Wald, L.L., 2018. High-  
608 resolution *in vivo* diffusion imaging of the human brain with generalized  
609 slice dithered enhanced resolution: Simultaneous multislice (gSlider-SMS).  
610 Magn. Reson. Med. 79, 141–151. doi:[10.1002/mrm.26653](https://doi.org/10.1002/mrm.26653).
- 611 Setsompop, K., Gagoski, B.A., Polimeni, J.R., Witzel, T., Wedeen, V.J.,  
612 Wald, L.L., 2012. Blipped-controlled aliasing in parallel imaging for si-  
613 multaneous multislice echo planar imaging with reduced *g*-factor penalty.  
614 Magn. Reson. Med. 67, 1210–1224. doi:[10.1002/mrm.23097](https://doi.org/10.1002/mrm.23097).
- 615 Shafieizargar, B., Jeurissen, B., Poot, D.H.J., Klein, S., Van Audekerke,  
616 J., Verhoye, M., den Dekker, A.J., Sijbers, J., 2023. ADEPT: Accurate  
617 diffusion echo-planar imaging with multi-contrast shTs. Magn. Reson.  
618 Med. 89, 396–410. doi:[10.1002/mrm.29398](https://doi.org/10.1002/mrm.29398).

- 619 Tournier, J.D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch,  
620 M., Christiaens, D., Jeurissen, B., Yeh, C.H., Connelly, A., 2019. MRtrix3:  
621 A fast, flexible and open software framework for medical image processing  
622 and visualisation. NeuroImage 202, 116137. doi:<https://doi.org/10.1016/j.neuroimage.2019.116137>.
- 624 Trzasko, J., Manduca, A., 2011. Local versus global low-rank promotion in  
625 dynamic MRI series reconstruction, in: Proceedings of the 19th Annual  
626 Meeting of ISMRM, Montréal, CAN, p. 4371.
- 627 Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen,  
628 V.J., 2002. High angular resolution diffusion imaging reveals intravoxel  
629 white matter fiber heterogeneity. Magn. Reson. Med. 48, 577–582. doi:[10.1002/mrm.10268](https://doi.org/10.1002/mrm.10268).
- 631 Uecker, M., Lai, P., Murphy, M.J., Virtue, P., Elad, M., Pauly, J.M.,  
632 Vasanawala, S.S., Lustig, M., 2014. ESPIRiT – an eigenvalue approach  
633 to autocalibrating parallel MRI: Where SENSE meets GRAPPA. Magn.  
634 Reson. Med. 71, 990–1001. doi:[10.1002/mrm.24751](https://doi.org/10.1002/mrm.24751).
- 635 Wu, W., Koopmans, P.J., Andersson, J.L., Miller, K.L., 2019. Diffusion  
636 Acceleration with Gaussian process Estimated Reconstruction (DAGER).  
637 Magn. Reson. Med. 82, 107–125. doi:[10.1002/mrm.27699](https://doi.org/10.1002/mrm.27699).
- 638 Wu, W., Miller, K.L., 2017. Image formation in diffusion MRI: A review  
639 of recent technical developments. J. Magn. Reson. Imaging 46, 646–662.  
640 doi:[10.1002/jmri.25664](https://doi.org/10.1002/jmri.25664).

641 Zhang, T., Pauly, J.M., Levesque, I.R., 2015. Accelerated parameter mapping  
642 with a locally low rank constraint. Magn. Reson. Med. 73, 655–661. doi:[10.1002/mrm.25161](#).

## NIMG-23-236: Responses to Editors and Reviewers

### **Reviewer 1**

*Authors propose an integrated acquisition and reconstruction methodology for accelerated multi-shell diffusion weighted imaging (DWI) based on an interleaved phase-encoding (PE) shifting and joint regularization with local low-rankness (LLR). The motivation is clear. The approach is original, although novelty is only incremental with respect to the state of the art. The methods are generally well presented, although certain details about the reconstruction algorithm are missing. The experimental section is weak, as quantitative comparisons or ablation/sensitivity analyses are missing, although the provided images illustrate the potential of the approach. The discussion is appropriate although some practically relevant points may be missing. The conclusions are generally well supported. Therefore, I would recommend a revision attending the points below.*

*Major:*

- (1) *No quantitative validation is provided. Authors should include quantitative comparisons of alternative methods using ground truth (GT) reconstructions (GT) built either from simulations or from retrospective subsampling of a long-enough DWI scan (or ideally from both).*

Thank you for the suggestion. We acquired another dataset using 4-shot interleaved EPI and then retrospectively undersampled the data to only one shot. We compared the reconstructed DW images both qualitatively and quantitatively. Please refer to R1.1 in the annotated manuscript.

- (2) *Authors do not compare with JULEP, DAGER or SPA-LLR but these are cited as state of the art methods (more recent than MUSE and MUSSELS). Therefore, authors should also compare with these methods or else precisely indicate why are these left aside.*

Thank you for the suggestion. We compared with MUSE and JULEP. We didn't compare with DAGER because DAGER was proposed for single-shot EPI acquisition. We did not compare with SPA-LLR because the SPA-LLR paper addressed the reconstruction problem of single-band acquisition. We thank all researchers who share their codes in the section Acknowledgments.

- (3) *For MUSSELS, authors should explicitly state if they base their implementation on Mani 2017 or Bilgic 2019 (whilst briefly motivating why). For local PCA, authors should more explicitly state whether they are comparing with Manjón 2013 or Veraart 2019.*

We based our MUSSELS implementation on Bilgic 2019 ([10.1002/mrm.27813](https://doi.org/10.1002/mrm.27813)) for two reasons. First, Dr. Bilgic shares his MUSSELS implementation (see Acknowledgments). Second, Dr. Bilgic's implementation allows for the reconstruction of multi-band multi-shot EPI data utilizing the readout extended FOV concept. However, in this revision, we compared with JULEP instead.

For local PCA, considering the major comment from Reviewer #3 (R3.6), we decided to compare with Cordero-Grande, et al. Neuroimage (2019) [10.1016/j.neuroimage.2019.06.039](https://doi.org/10.1016/j.neuroimage.2019.06.039).

- (4) A sensitivity analysis based on the GT should be included considering these factors: LLR regularization weight, LLR block size, LLR overlap factor, PE interleave configuration (i.e., why 2 shots/b?). In case some of these factors are/need to be left aside, authors should precisely justify why. Ablation experiments comparing full proposal versus removing interleaved PE / removing LLR would also be very interesting.

Thank you for the suggestion. We have conducted the sensitivity analysis and ablation experiments. Please refer to Figures 5 and 6.

- (5) LLR regularization performance and reliability may degrade in the presence of motion. Also, often DWI is performed with alternating PEs for distortion correction. SNR is heterogeneous over the FOV, which may not be appropriately covered by a single regularization weight. Please, add these aspects to discussion. See also minor point 5.

Thank you for the suggestion. We have incorporated motion into the discussion section.

- (6) Abstract is focused on quite general motivating aspects, but paper methods and results are described in one/two sentences each. Authors should rebalance to provide abstract readers a better/quick understanding of paper content. Namely, shift-encoding and LRR ideas should be more clearly described in the abstract.

Thank you for the suggestion. We have clearly described in the abstract the ideas of shift encoding and LLR.

- (7) Fig. 2 / Fig. 3 →you mention slice / diffusion direction included, but this is not very relevant, what matters is to specify the reasons for including these and not others. Also, it would be important to include snapshots at different slice locations within the brain, particularly at inferior locations, where reconstruction may become more challenging. Analogously, please provide rationale for snapshots selected in Fig. 4 and Fig. 5.

Thank you for the suggestion. We included slices from inferior locations.

- (8) <https://github.com/ZhengguoTan/sigpy> seems a link to a generic tool. <https://doi.org/10.5281/zenodo.7548595> is not available yet. Please, remember to provide paper-specific links before acceptance or otherwise remove these links from the manuscript.

<https://github.com/ZhengguoTan/sigpy> was forked from the original sigpy (<https://github.com/mikgroup/sigpy>) for the development of our proposed reconstruction.

As shown in Figure 1, <https://doi.org/10.5281/zenodo.7548595> is a link we reserved for the upload of all raw data in this work upon the publication of our manuscript. Therefore, we would like to keep this link in the manuscript.

Minor:

- (9) Reference to 7T included in title does not seem relevant enough to me. Similar challenges for DWI can manifest for high resolution / high b-value lower field scans as well, so I'd recommend removing reference to 7T in title.

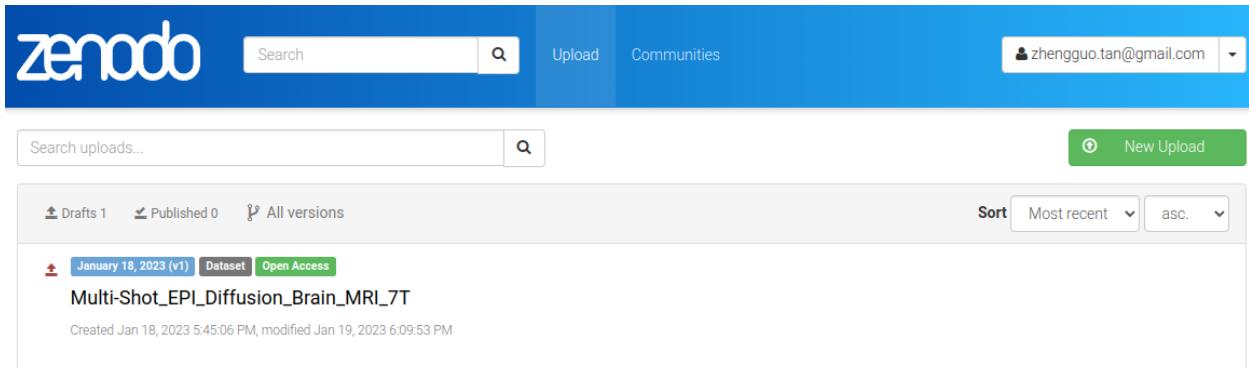


Figure 1: Reserved Zenodo link for the host of all raw data in this work.

Since the primary aim of this work is to develop an efficient data sampling and image reconstruction technique for diffusion MRI at 7 T, we would like to keep 7 T in the title.

- (10) *L55: "DAGER requires many diffusion directions" →JETS also seems to require many diffusion directions, so guess this should be formulated differently?*

We reformulated this paragraph.

- (11) *Section 2.2.1 and 2.2.2 →please, explicitly specify slice thickness.*

Both protocols employed isotropic resolution. We listed slice thickness in the table.

- (12) *L105: "acceleration" →would use "undersampling" as this is EPI and there's no 1-1 correspondence between undersampling and acceleration (several occurrences).*

Done.

- (13) *L160: "As phase images are spatially smooth" →this may be arguable, would suggest replacing by "Assuming that phase images are spatially smooth". Think you should add a bit more on this topic to discussion, as phase behaviour depends on several hard-to-control factors such as pulsatile motion, its impact at different locations within the brain, diffusion sensitization strength, bulk motion,... so phase correction may sometimes be really challenging? Some additional lines of discussion on integration of navigators / cardiac gating could also be worthwhile.*

Thank you for the suggestion. We have rephrased the sentence and added the topic of phase estimation to the discussion.

- (14) *L172: "firstly slides [...] matrices" →not very clear, think you should reword and explicitly mention what dimensions are concatenated in rows and columns of matrices. Also, reasons for not using 3D patches should be discussed.*

Thank you for the suggestion. 3D patches were not used because this is a SMS reconstruction. In SMS, spatially separated slices are simultaneously excited and acquired. This spatial separation of slices prohibits the use of 3D patches.

(15) L176: " $T^H T$  input neq input" →unclear, maybe a typo?

Thank you for the question. Please refer to R2.1 as well as the notebook for detailed explanation: [https://github.com/ZhengguoTan/demo\\_jets\\_diffusion\\_mri\\_7t](https://github.com/ZhengguoTan/demo_jets_diffusion_mri_7t).

(16) L181: "efficient implementation" →claim on efficiency does not seem supported from description... inverse density weighting is well-known for reconstructing original data levels back when slide-windowing. You should provide more details on efficiency or articulate description differently. Importantly, overlap ratio does not seem to be reported, but this may have a dramatic effect in computational cost.

Thank you for the comments. The scaling matrix (inverse density weighting) is needed in the case of overlapping patches (i.e. when stride < block), in order to remove accumulated values during SVT. The overlap ratio was also reported (i.e. stride = 1).

(17) Supporting Figure S3. "Small block size (i.e., 2) suffers from image blurring, whereas increasing block size gradually leads to increased noise" →may appear counter-intuitive as small block sizes should aid with localization and therefore prevent blurring, at the price of less information for denoising? Can you clarify on reasons / potential hidden factors for this behaviour?

Thank you. We removed the supporting figures, and included the sensitivity analysis as a main figure in the manuscript.

(18) L239: "blurring" →leading artifact resembles Rician bias rather than blurring in my opinion, could you clarify?

You could be right the DWI at high  $b$ -values resembles Rician bias.

(19) Seems Fig. 4 could be replaced by encompassing Supp. Fig. 6?

Thank you. The results were replaced.

Typos, suggestions...

(20) L33: "needed" →"used".

Done.

(21) L83: "benefits" →"benefits" to be stated experimentally, in methods better to refer to "properties"?

Done.

(22) L112: "as Section 2.2.1" →"as in Section 2.2.1".

Done.

(23) L119: "DW acquisition" →"DW acquisition volumes".

Done.

(24) L120: "demonstrates" → "will be used to demonstrate".

Done.

(25) L123: "slice collapsed k-space data" → not sure this expression is clear, you may consider rewording.

Thank you. Reworded.

(26) L125: "Such acquisition can be modeled in two ways." → "Acquisition modeling needs to consider several aspects".

Done.

(27) L134: "at every" → "of each".

Done.

(28) L136: "shot images per" → "multiple shots acquired for a given".

Done.

(29) L137: "One method" → "A possibility".

Done.

(30) L141: "is done" → "can be done".

Done.

(31) L142: "This method can be written" → "This can be incorporated to our formulation".

Done.

(32) L147: "(e.g. Hanning window)" → would remove as there are many other possibilities, so this reference may not contribute to clarity of description?

Done.

(33) L148: "phase correction method" → "phase correction".

Done.

(34) L152: ", utilizing the concept of object-oriented linear operator abstraction" → not sure this is adding much, would remove.

Done.

(35) L168: "Intuitively, low rankness comes from contrast variation feature of DW images" → meaning unclear to me, you may reword.

Thank you. Done.

(36) L173: "patchs" → "patches".

Done.

(37) L176: "as an" → "as a".

Done.

(38) L185: "The acquired raw data was read by *twixtools* (URL)" → not adding much, could be specified in code repo.

Yes, we agree.

(39) L191: "the proximal operator" → "proximal operator".

Done.

(40) L196: "x E" → better to separate rather than concatenate description for both.

Done.

(41) L221: "GFA" → acronym does not seem defined.

Thank you. We added the definition.

(42) L247: "desnoing" → "denoising".

Done.

(43) L252: "within the rectangular regions in Fig. 4" → these may be difficult to appreciate, consider enhancing a bit perhaps?

Thank you. We switched to the yellow box when it came to magnified views.

(44) L259: "smooth patterns" → "smoothness".

Done.

(45) L261: "spatial-angular" → really angular or b-vals are fused? Perhaps more accurate to say "spatial-diffusion"?

Done.

(46) L271: "achieves" → "uses".

Done.

## **Reviewer 2**

*This paper describes an image reconstruction method for diffusion-weighted segmented echo-planar imaging called JETS, which jointly processes all segments (shots) and diffusion weightings to explore local similarities between differently weighted images. This is achieved by adding a regularization term to the cost function guiding the image fitting that favors low-rankness of matrices created from image patches with different weightings. Additionally, inter-shot phase variations, which are a major problem of segmented diffusion MRI, are evaluated by separate pre-reconstructions of single segments and included in the signal model.*

*Although the elements of the presented method have described before (low-rank modelling of image patches is used in SVD-filtering methods, and the single-shot phase mapping in MUSE, both duly cited), combining them in the reconstruction procedure is novel, and visibly provides superior results. The paper should meet the interest of MRI methodology-oriented readers even though the proposed strategy is not immediately applicable for routine scanning due to prohibitive computation times. I only have some minor remarks regarding the clarity of method description and the comparisons with established methods:*

- (1) *It is not clear how the operator  $T$  in Eq. 6 treats the multiple patches of the image. A graphical illustration of how the patches are combined to a 2D matrix whose nuclear norm is then taken would be helpful.*

Thank you for the suggestion. We made an interactive demo to illustrate the operator  $T$ :  
[https://github.com/ZhengguoTan/demo\\_jets\\_diffusion\\_mri\\_7t](https://github.com/ZhengguoTan/demo_jets_diffusion_mri_7t).

- (2) *When the EPI segments are initially reconstructed to obtain the phase errors (Eq. 5), the related images surely represent single diffusion weighting. What is the  $T$  operator doing then in this equation? There is no diffusion dimension yet to impose a low rank approximation.*

Very good question. We reformed the method description.

- (3) *How was the patch size and the overlap selected (on what basis)?*

Thank you for the question. Please refer to R1.4 and the sensitivity analysis figure in the manuscript.

- (4) *The proposed strategy involves a shift of the sampling pattern between diffusion weightings. The advantage of this shift is intuitively understandable, and it is reported as a result (lines 279-280) but not shown directly (comparisons with MUSE and MUSSELS involve more differences than the shift alone). It would be great to see the proposed joint reconstruction without the shift for comparison.*

Thank you. We compared the reconstruction without and with the  $k_y$  shift in the manuscript.

- (5) *Is there any importance in the way the sampling shift is distributed among diffusion weightings?*

In this work, multi-band and in-plane undersampled interleaved EPI was proposed. With in-plane undersampling, it is indeed important to have  $k_y$  shifting. Please refer to the above-mentioned comparison in the manuscript.

- (6) *The proposed method seems very similar to MUSE with the PCA denoiser (the difference is that in JETS the LLR is promoted during the image fit, while in MUSE it is imposed afterwards) but delivers surprisingly better results. In the comparison of both strategies (Fig. 2) MUSE+PCA seems to reduce the noise stronger while some details are lost. If the PCA filter were made less aggressive, would MUSE still look worse than JETS?*

We compared with the complex-valued DWI denoising method (Cordero-Grande, et al. Neuroimage (2019) [10.1016/j.neuroimage.2019.06.039](https://doi.org/10.1016/j.neuroimage.2019.06.039)), in which the aggressiveness (strength) of the PCA filter is automatically determined. It is beyond the scope of this work to optimize the published denoising method.

- (7) *In the same figure, MUSSELS shows similar details to JETS, but with more noise. Here, conversely, could not MUSSELS be improved by adding a moderate degree of PCA filtering?*

In this revision, we switched to JULEP for comparison, because in principle JULEP is an improved version of MUSSELS.

- (8) *The symbol  $\mathcal{T}_{\lambda/\rho}$  is not defined in Eq.7.*

Thank you. Defined.

### **Reviewer 3**

*The authors present a novel approach to accelerate diffusion imaging acquisitions by jointly reconstructing highly accelerated diffusion-weighted images recorded with different diffusion weightings and k-space sampling patterns. The underlying assumption is that by encoding complementary k-spaces in the different diffusion images, they can acquire a smaller k-space for each individual image, thus accelerating the overall acquisition. They show that their algorithm is able to reconstruct this undersampled data, whereas other contemporary reconstruction algorithms (which do not jointly reconstruct the data) perform less well.*

*The principle is interesting, but I feel that the Authors are missing some important validation aspects, which are detailed as Major comments below.*

#### *Major comments*

- (1) *How many subjects were actually scanned? The Materials and Methods refers to "healthy volunteers", but I could not find an explicit number. The data presented in the paper seems to come from only a single subject. This is obviously inadequate to properly evaluate the performance of the method.*

Three subjects were scanned in this work.

- (2) *Was there any correction for different eddy currents and motion between volumes? If not, why not? Eddy currents and (out-of-plane) motion could break the assumption that the low-rank patches reflect the same underlying anatomy. Perhaps the joint reconstruction method developed by the Authors actually mitigates these effects (and that would represent an additional advantage of the method), but correction would definitely need to be done in the case of the other reconstruction methods in order to make a fair comparison.*

*As an aside, the Authors do refer to "motion robustness" several times, but this can surely only be robustness to in-plane motion. Out-of-plane motion in 2D imaging is harder as it is accompanied by a true loss of information. Model-based methods like that implemented in eddy (Andersson, et al. Neuroimage (2017) <https://doi.org/10.1016/j.neuroimage.2017.02.085>) can partially compensate for it, but don't seem to have been used here; and techniques like gSlider can correct for motion within the thick acquired slab during reconstruction (Wang, et al. Magn. Reson. Med. (2018) <https://doi.org/10.1002/mrm.27196>), but require specialised acquisitions.*

*I would suggest that the Authors properly discuss the issues of eddy currents and intra- and inter-volume motion, and perhaps consider how they could properly incorporate a consideration of these effects into their framework.*

Thank you. We included all these issues into the discussion.

- (3) *Similarly to 2.: was there any correction for susceptibility distortions? I understand that the segmented EPI will have less distortions than conventional single shot EPI, but they still need to be corrected for to get anatomically correct images. Was the GRE scan perhaps*

*used to make a B0 map which was included in the reconstruction?*

This work focused on the development of the NAViEPI sequence and the joint reconstruction method. No susceptibility distortion correction was done in this work, but we put this issue into the discussion.

- (4) *It did not become clear to me why this study is being done at 7T. The increased sensitivity to B0 distortions and SNR loss with long TEs which this study proposes to overcome are 7T problems that are much less pronounced at 3T. In part it is for this reason that most DWI studies are still performed at 3T. What benefit are the Authors aiming to get from running DWI studies at 7T?*

Thank you for the comment. We agree with you about the challenges of DWI at 7 T. However, we believe that there is an SNR advantage at 7 T. According to Figure 20 in Uğurbil et al. ([10.1016/j.neuroimage.2013.05.012](https://doi.org/10.1016/j.neuroimage.2013.05.012)), the SNR benefit that 7 T generally possesses over 3 T should outweigh the shorter  $T_2$  times at 7 T, even if the bandwidth is adjusted to compensate for the larger field inhomogeneities.

- (5) *Relatedly, the acquisition scheme doesn't seem that fast. A similar scheme at 3T could even be shorter as T2 signal loss is less pronounced and so each (multiband) slice could be acquired in a single shot.*

Thank you for the comment. We compare our proposed sequence NAViEPI with single-shot EPI for the acquisition of sub-millimeter ( $0.5 \times 0.5 \times 2 \text{ mm}^3$ ) diffusion-weighted images. Please refer to Figure 7 in the manuscript.

- (6) *The denoising approach used as comparison does not reflect the state of the art. It is generally recommended to perform denoising on complex data (see Cordero-Grande, et al. Neuroimage (2019) <https://doi.org/10.1016/j.neuroimage.2019.06.039>), as this will avoid the noise floor issues that are apparent in the MUSE + Denoiser panel in Fig. 3. This should be completely possible for the Authors, as they reconstruct the data themselves. I would say the only reason not to do complex denoising at this point is if you are stuck with only scanner reconstructed magnitude data. Complex denoising is available in openly available toolboxes, e.g. MRtrix3 (<https://mrtrix.readthedocs.io/en/latest/reference/commands/dwidenoise.html>).*

Thank you for the literature and the MRtrix tool. In fact, all reconstructions shown in this work were done off-line, including MUSE. We installed the MRtrix tool, and were able to use the denoising approach on complex MUSE images. Please refer to Figure 4 in the manuscript.

- (7) *While we are on the topic of "state of the art": All the papers in the Introduction demonstrating "conventional" SS-EPI are from over 20 years ago. There have been important technological developments, e.g. Connectom scanners, gradient inserts, sequence developments,*

e.g. multiband, gSlider, and image correction technique developments, e.g. the continued development of "topup" and "eddy", since then.

Thank you for the suggestion. We have included the techniques on multiband, gSlider, and topup in the Introduction.

- (8) *The Introduction is in general very confusingly written, and I would suggest considering how to put the introduced topics in a more logical order. As I see it, the fundamental problem is the trade-off between minimising distortions and maximising SNR while minimising acquisition time and sensitivity to motion, and this message does not come through clearly. For instance it is very confusing that it is suggested that navigators should be avoided because they increase the acquisition time, but segmented EPI – which literally (at least) doubles the acquisition time while also introducing motion sensitivity – is introduced as necessary without a consideration of the trade-offs.*

Very good insights. We removed the statement on navigators in the Introduction.

- (9) *How were the number of directions in the three shell case determined? Generally 30 directions are recommended even for  $b = 1000 \text{ s/mm}^2$ , but here only 20 are used.*

Thank you for the question. For the three-shell acquisition, we used the internal diffusion vector sets (DVS) available as the MDDW diffusion mode in Siemens scanners. The DVS with 20, 30, and 64 directions were used for the  $b$ -value of 1000, 2000, and  $3000 \text{ s/mm}^2$ , respectively.

- (10) *The "efficient implementation" to correct checkerboard artefacts (lines 177–183) seems very underspecified. Perhaps it would be better developed in an appendix and just mentioned in the main text? Specific points:*

Please refer to the open-source interactive demo on the explanation of our implementation: [https://github.com/ZhengguoTan/demo\\_jets\\_diffusion\\_mri\\_7t](https://github.com/ZhengguoTan/demo_jets_diffusion_mri_7t).

- *is  $(1/\text{divisor})$  a matrix inverse of  $(T' * T * 1)$ ? Or is it rather a scalar derived by solving the linear equation  $(T' * T * 1) = (\text{something})$ ?*

"divisor" is a matrix, but not the inverse of  $T^H T 1$ .

- *is  $1$  the matrix of all ones, or the identity matrix?*

$1$  is the matrix of all ones.

- *why would including this divisor prevent (or mitigate) checkerboard artefacts at all?*

The use of overlapping blocks prevents checkerboard artefacts, and overlapping blocks require the removal repeated values. Here we avoid this repeating by using the divisor.

#### *Minor comments and typos*

(11) *title should probably say "Diffusion \*Weighted\* Magnetic Resonance Imaging"*

Done.

(12) *typo in graphical abstract (METHODS (2): "reconsturction" instead of "reconstruction")*

Thank you. Done.

### *Abstract*

(13) *high b-values do not "increase ... noise"; this should probably be rephrased to make clear that more strongly diffusion weighted images have lower SNR, thus increasing the \*sensitivity to\* noise.*

Thank you. This has been rephrased.

(14) *"inplane" should be "in-plane" for consistency with the rest of the document.*

Done.

### *Introduction*

(15) *line 22: "spiral" is not a multi-shot EPI technique.*

Thank you. We removed "spiral".

(16) *line 37: should be "single shot images", not just "shot images"?*

Thank you. Corrected.

(17) *line 57: should be "in-plane".*

Done.

(18) *line 58: should be "still require long acquisition times".*

Done.

(19) *line 68: "i.e." should be "specifically".*

Done.

(20) *line 69: "the established DW image denoising algorithm, i.e., local PCA [REFS]" would be better written as "an established local PCA-based DW image denoising algorithm [REFS]" (though see the major comment above regarding whether this is state of the art).*

Thank you. We rephrased this sentence.

### *Materials and methods*

(21) *line 73: should be "Materials and methods" (not "Material").*

Done.

(22) *line 92: the non-standard way of representing maximum gradient strength and slew rate should be written more explicitly.*

Done.

(23) *line 107: why are the  $b=0$  images referred to as  $b = 50 \text{ s/mm}^2$  images here, but  $b_0$  acquisitions for the three shell protocol? This should be standardised.*

Thank you. The acquisition protocol has been standardised. Please refer to Table 1 in the manuscript.

(24) *line 107, 119: why is minutes'seconds" notation used here but "min" elsewhere? Recommend standardising, especially since the used of minutes'seconds" is fairly old-fashioned...*

Thank you. Corrected.

(25) *were the different b-values in the second acquisition scheme interleaved, or collected one after each other?*

The different  $b$ -values were collected one after the other. In other words, we first acquired 20 diffusion encoding at the  $b$ -value of  $1000 \text{ s/mm}^2$ , then 32 diffusion encoding at the  $b$ -value of  $2000 \text{ s/mm}^2$ , and finally 64 diffusion encoding at the  $b$ -value of  $3000 \text{ s/mm}^2$ .

(26) *line 119: should be "acquisitions".*

Done.

(27) *line 139: presumably this should be "This method is robust to in-plane motion".*

Thank you. Corrected.

(28) *line 176: should be "a proximal operator".*

Thank you. Corrected.

(29) *line 177: "Noteworthy" is not usually used in this sense. Here it seems redundant and can be deleted.*

Thank you. Deleted.

(30) *line 180: I suggest "Hermitian adjoint" rather than just "adjoint".*

Thank you. Added.

(31) *line 183: should be "the input".*

Thank you. Corrected.

(32) *line 194: the meaning of the  $T$  symbol should be defined.*

Done.

(33) *line 196: again, "Noteworthy" is an odd choice here. I would suggest "Importantly".*

Thank you. Replaced.

(34) *line 196: missing "and" between "x E".*

Thank you. Corrected.

(35) *line 199: should be "conjugate gradients".*

Thank you. Corrected.

(36) *line 203: should be "as" not "ss".*

Thank you. Corrected.

(37) *line 206: NVIDIA seems to write the name of the GPU as "NVLink".*

Thank you. Corrected.

(38) *line 209: again, "i.e." should be "specifically".*

Thank you. Corrected.

(39) *line 212: "implementation" should be "implementation" and quotes are backwards (should be " ).*

Thank you. Corrected.

(40) *lines 217–218: which shell or shells were used to compute the fODFs?*

We used all shells to compute the fODFs.

(41) *line 221: "GFA" abbreviation should be defined.*

Defined.

## *Results*

(42) *line 228: should be "loses".*

Thank you. Corrected.

(43) *line 241: "allows to resolve" isn't a standard English construction. Should be something like "allows the resolution of".*

Thank you. Corrected.

(44) *line 247: should be "denoising".*

Thank you. Corrected.

## *Discussion*

(45) *line 265: "dubbed as" is redundant; can just be "dubbed".*

Done.

(46) *line 273: should be "diffusion-directions".*

We change to "per diffusion-weighted image".

(47) *line 281: instead of "as" I would suggest "as used by".*

Done.

(48) *line 289: instead of "on GPU A100" I would suggest "on an A100 GPU".*

Done.

(49) *line 305: should be "solves for a fewer number of".*

Done.