

Imaging Neuroscience #203: Responses to Editors and Reviewers

Reviewer #248

Authors have partially addressed my previous comments. A list of pending issues is listed below:

Major:

- 1) *The experiment is generally satisfactory, but its description is a bit confusing (especially point b) below:*

- a) *L125: "Protocol #1 with six-shot iEPI and without in-plane undersampling was implemented" then (L131) "We then retrospectively reduced the four-shot data to only one shot", guess there is a misprint in L125 which should be four-shot rather than six-shot?*
- b) *L132: "with the proposed k_y shifting to simulate three-fold in-plane undersampling" → as there is no in-plane undersampling in Protocol #1 in the baseline data (according to Table 1) this sounds a bit confusing, if 1 out of 4 shots are used this would be four-fold in-plane undersampling, so unclear how you reach three-fold in-plane undersampling? Guess you mean something like "three-fold in-plane interleaving"? A few more details may help here.*
- c) *L133: "JETS reconstruction was performed on all data" → "all data" is a bit vague, please reword.*

- a) *Sorry about the misprint. We corrected it as four-shot.*
- b) *Thank you for the question. You are right that "1 out of 4 shots are used", so it is four-fold in-plane undersampling. We corrected this in the manuscript.*
- c) *Thank you. This sentence has been reworded.*

- 2) *Fine, only a small detail, L61: "acquisition. SPA-LLR" → "acquisition and SPA-LLR"*

Done.

- 3) *There are some pending issues:*

- a) *Authors could drop a line in text indicating why they use JULEPS rather than MUSSELS.*
- b) *Results of MUSE + Denoiser look stronger than with previous denoiser. Indeed, now they are quite comparable to JETS (bit noisier but also bit sharper). A question is whether complex denoiser is applied before (replacing TV) or after shot-combination of MUSE? May make sense to try both approaches. Also, authors may want to draw a few lines about potential synergies between their LLR and the denoiser.*
- c) *In previous version you also included comparisons for FA and different b-values but these seem missing here. I think that maintaining comparisons at different b-vals could be particularly useful.*

- a) Done.
- b) The complex denoiser is applied after the shot combination in MUSE. The local-PCA denoiser operates on spatial-diffusion matrices, whereas MUSE reconstructs one DWI at a time. Therefore, the denoiser is only applicable after the reconstruction of all DWIs, and is not possible before the completion of MUSE (e.g. before shot combination).
We added "potential synergies between their LLR and the denoiser" into Discussion.
- c) We added back the three-shell data. **TODO**

4) *Description of these experiments is misleading at times:*

- a) L300: *"According to (Cordero-Grande et al., 2019), it is suggested to keep the patch size roughly equal to the diffusion encoding length. In this 3-scan trace acquisition example, the diffusion encoding length is 4 (1 b0 plus 3 orthogonal diffusion directions). Among the tested block sizes, 3 is the closest to 4, and hence renders better denoising" → but in (Cordero-Grande et al., 2019), patch sizes are automatically estimated (have tried to find suggestion reported without success); whilst patch sizes inducing close to square matrices could be intuitively better for large matrices, that's not so clear for small matrices as here (4 columns); however, even in this case, patch sizes would be width*height, but here they are used interchangeably with widths. All of this needs detailed clarification or another thought.*
- b) L311: *"which is prone to checkerboard artifacts even with the use of random shifting (Saucedo et al., 2017) in each ADMM iteration" → please review, as message of (Saucedo et al., 2017) seems different: "This result shows the robustness of the LLR-IRPA strategy, which maintains improved computational efficiency compared to CLEAR without introducing block artifacts", i.e., their method shows similar levels of artifacts than overlapped approaches (they even report a reduction in some experiments) / P24(L308): "the increment from one local patch to the next" → perhaps replace increment by "step" or "distance" for clarity?*

- a) Please refer to your minor comment 9 for responses.
- b) Thank you for the suggestion. We agreed that LLR-IRPA improves the reconstruction for non-overlapping blocks. In our ablation experiments, the use of overlapped LLR better suppresses blocky artifacts. We rewrote this part.

6) *Abstract looks better now, but you may add a one-sentence summary of experimental results. Also, "developed the k_y " → "developed a k_y ".*

Thank you. We added the summary of experimental results in Abstract.

7) *Still unclear how do you select the slice / diffusion encoding of snapshots provided in Figs. 4-9. Also, caption of Fig. 5, "retrospectively" → "retrospectively"*

Figures 4 and 5 show a slice containing the globus pallidus with strong T_2 -weighted contrast. Figure 6 selects a slice with a suspicious lesion (the circular bright spot) within the left ventricle. Figure 8 displays three representative slices: (left) the lower brain region which identifies the B_1^+ field inhomogeneity, (middle) the middle brain slice which shows susceptibility artifacts in the frontal region, and (right) the upper brain slice which shows the ventricle.

- 8) *Understood regarding raw data. However, regarding source code, you say "we publish our source code (<https://github.com/ZhengguoTan/sigpy>)" but I see no reference to this work in that link. You may clarify by better documenting the repo / providing entry point of specific methods?*

Thank you for the comment. All our source codes have been pushed here: <https://github.com/ZhengguoTan/sigpy>. We will document this repo once the manuscript is accepted.

For an entry point of the proposed methods, please refer to the interactive demo at https://github.com/ZhengguoTan/demo_jets_diffusion_mri_7t. This earned 2nd Place in the MR-Pub Competition organized by the Reproducible Research Study Group in ISMRM 2023.

Minor:

- 6) *This comment is not satisfactorily addressed:*

a) *Sentence "firstly slides [...] matrices" (L202) unclear, requested change ("you should reword and explicitly mention what dimensions are concatenated in rows and columns of matrices") not attended.*

b) *Reasons for not using 3D patches are not convincing. Authors are referring to SMS as reason behind, but this pertains to acquisition whilst patches are defined on image space, so usage of 3D (for instance to exploit redundancies between adjacent slices) could in principle be possible.*

a) Thank you for the suggestion. We explicitly defined the spatial and diffusion dimension in the manuscript.

b) SMS acquires spatially separated slices, which means the slices in the joint reconstruction are not spatially continuous. Therefore, 2D patches are used for joint reconstruction. We removed the statement in the manuscript to avoid misunderstanding. But we agreed that after the reconstruction of all slices, 3D-patch-based denoising can be used as a postprocessing step. We put this into Discussion.

- 7) *Please, use a symbol rather than "input" (L215) for clarity of this expression.*

Thank you for the suggestion. We replaced "input".

- 8) *This is only partly attended, when you say "efficient implementation" (L214, L217) → claim on efficiency does not seem supported from description... inverse density weighting is well-known for reconstructing original data levels back when slide-windowing, so this is just the standard method, not something particularly efficient.*

We removed the "efficiency" statement.

- 9) *This comment is not satisfactorily addressed. In new Fig. 6b we still see higher noise for larger block sizes, so my previous comment, "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?" may still be relevant.*

Very good insight. This is connected to your major comment 4a) and we tried to attend them together.

Yes, you are right that patch sizes should be width \times height. We corrected this in the manuscript.

The automatic estimation of the block size is described here (<https://mrtrix.readthedocs.io/en/dev/reference/commands/dwidenoise.html>), which suggests selecting "the smallest isotropic patch size that exceeds the number of DW images". Therefore, the potential hidden factor is that although small block sizes aid with localization, one also need to construct a "balanced" matrix (block size \approx diffusion encoding length) for SVT.

- 11) *Sorry, I don't see that this has been attended.*

We added the row "slice thickness" in Table 1.

- 12) *(New content) Reg. Eq. 6, L. "where the index k denotes the iteration" → this iteration was not in previous manuscript, unclear what it refers to.*

It refers to the iteration step of the iterative phase smoothing procedure.

Reviewer #249

I was pleased to see that the Authors have put significant effort into this revision. However some major concerns I had regarding the description of the method and its potential reproducibility still remain.

Major

Number of subjects

I am happy that the Authors now state that they measured three subjects in the Methods section, but the presented data still only seem to show a single subject (e.g. in Fig. 9). Perhaps the Authors used a different subject for each separate experiment? Nevertheless, I would really like to see data from more than one subject for each experiment, even if this is put into an Appendix or Supplementary Information. Reproducing results in more than one subject is important to evaluate whether a method gives good results outside of a single (often "best") subject. If the main results do really each rely only on a single subject, I would strongly argue that each result should be reproduced in another subject before publication.

TODO

Checkerboard artefact suppression (pg. 15)

- (1) I thank the Authors for providing the code that let me understand what the T and T^H operators are doing. However I now think that expressing the operation of T and T^H as matrix multiplication is very misleading. I would urge the Authors to rethink how they express this part. Probably it is e.g. better to write $T(x)$ instead of Tx , unless the authors explicitly define what the multiplication means in this non-standard operation.*

Thank you for looking into the code. We changed to $T(x)$ instead.

- (2) Further, I think T^H needs to be better explained and justified. Where is it used? Are all the blocks reconstructed separately and then recombined for computational efficiency? Or is it just used when calculating the nuclear norm? It seems that T^H is notation from the python package used by the Authors, but I think it is not very useful notation for those unfamiliar with that package. If I followed everything correctly, then what is really needed is an inverse of T (i.e. T^{-1}), but this is non-unique for overlapping blocks (one could have any combination of weights to recombine the overlapping elements as long as they sum to one). The Authors decided to take equal weights for each repeated elements, i.e. to take the mean, and I agree that this is probably the best choice to avoid checkerboard artefacts. However it would be much clearer (at least to me) if the authors expressed this in words, making clear that they take the mean of the overlapping elements, rather than in terms of an operator " T^H " and "multiplication" operations which are not defined by the Authors. The Authors could then point explicitly to the scripts they have written for details.*

First, the singular value thresholding (nuclear norm regularization) is calculated from the blocks, i.e., $T(x)$. Second, after the thresholding, T^H is used to recombine the blocks to

obtain the updated DW images.

- (3) *If the operations really do need to be expressed in such detail, then it would perhaps be good to be explicit that the operation to get the "divisor" just gives the element-wise denominator for the mean.*

Thank you for the suggestion. We reworded the description.

- (4) *The claim that this way of recombining blocks is efficient is not proven. I would suggest removing this claim.*

OK. We removed this statement.

Minor

My minor comments are mainly regarding the wording of some of the new text, which doesn't read as natural English to me. There are also some points which I didn't quite understand which the Authors could consider elaborating on further.

Sincerely thank you for your correction on the language. We have corrected all of them and highlighted the correction in the manuscript. Therefore, no response was provided in the following comments concerning language issues.

- (5) *pg. 3: Techniques on the correction of... ← Techniques for the correction of...*
- (6) *pg. 11: Different reconstruction methods, i.e., JETS, MUSE, and JULEP were compared. ← Different reconstruction methods, specifically JETS, MUSE, and JULEP, were compared.*
- (7) *pg. 12: A possibility is to perform magnitude average... ← One possibility is to perform magnitude averaging...*
- (8) *pg. 13: This can be incorporated to our formulation... ← This can be incorporated into our formulation...*
- (9) *pg. 14: Please define the nuclear norm in Eq. (7). This would probably also help alleviate some of the problems with the description of the checkerboard artefact suppression mentioned above.*

Thank you. We defined the nuclear norm.

- (10) *pg. 15: "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." I don't understand how it follows from using multiband that 2D patches are used. Wouldn't this apply to all 2D sequences?*

Thank you for pointing this out. Here we reconstructed multiband slices jointly, but the multiband slices are spatially separated. This prohibits the use of 3D patches for the LLR regularization. We removed this part to avoid this confusion.

- (11) *pg. 18: As show in Fig. 3... ← As shown in Fig. 3...*
- (12) *pg. 18: ...the ripple-like phase artifact disapper after five iterations. ← ...the ripple-like phase artifact disappears after five iterations.*
- (13) *Fig. 4: Please define what the arrows represent in the image caption.*
 Done.
- (14) *pg. 21: retrospectively ← retrospectively*
- (15) *pg. 21: structural ← Structural*
- (16) *Fig. 5: Please define what the arrows represent in the image caption.*
 Done.
- (17) *pg. 22: as pointed ← as indicated*
- (18) *pg. 22: On the contrary ← In contrast*
- (19) *pg. 22: First, we varied the regularization strength λ from 0, to 0.08, and to 0.16. ← First, we varied the regularization strength λ . We tested values of 0, 0.08, and 0.16.*
- (20) *pg. 22: "i.e., reduced noise" What does this mean? Perhaps you need to be more explicit?*
 TODO
- (21) *pg. 24: ...with the use of 5-shot acquisition. ← ...with the use of a 5-shot acquisition.*
- (22) *pg. 24: When compared to the single-shot acquisition... ← When compared to a single-shot acquisition*
- (23) *Fig. 8: The TRACE image at the bottom right seems to show aliasing near the bottom. Do you have an explanation for this?*
 TODO
- (24) *pg. 29: ...MRI protocols in 7 T, e.g., 3-scan trace acquisition... ← ...MRI protocols at 7 T, e.g., a 3-scan trace acquisition...*
- (25) *pg. 30: "The standard distortion correction method is known as TOPUP (Andersson et al., 2003), which acquires two scans with opposing phase-encoding directions to obtain the field inhomogeneity map and then performs conjugate phase reconstruction to correct for distortion." The B0 map does not need to come from TOPUP. Why not just collect a B0 map from multiecho data and incorporate it into the reconstruction? Often the multiecho data used to create SENSE maps for coil combination is used for this purpose.*

Thank you for the suggestion. We added this into Discussion.

(26) pg. 30: *"Beyond this SVT-based machine-learning reconstruction, one may seek to learn a q -space prior as the regularizer..." I don't understand what this sentence is trying to say. Is it a suggestion that machine learning could be used to learn a q -space prior?*

Yes, it is a suggestion that machine learning could be used to learn a q -space prior.

Reviewer #250

All points of my critique have been addressed adequately.

Thank you for your review.