Research Article
Denoising Diffusion MRI via Graph Total Variance in Spatioangular Domain

Haiyong Wu ⬤ and Senlin Yan

School of Electrical Engineering, Nanjing Xiaozhuang University, Nanjing 211171, China

Correspondence should be addressed to Haiyong Wu; wu.haiyong@njxzc.edu.cn

Diffusion MRI (DMRI) plays an essential role in diagnosing brain disorders related to white matter abnormalities. However, it suffers from heavy noise, which restricts its quantitative analysis. The total variance (TV) regularization is an effective noise reduction technique that penalizes noise-induced variances. However, existing TV-based denoising methods only focus on the spatial domain, overlooking that DMRI data lives in a combined spatioangular domain. It eventually results in an unsatisfactory noise reduction effect. To resolve this issue, we propose to remove the noise in DMRI using graph total variance (GTV) in the spatioangular domain. Expressly, we first represent the DMRI data using a graph, which encodes the geometric information of sampling points in the spatioangular domain. We then perform effective noise reduction using the powerful GTV regularization, which penalizes the noise-induced variances on the graph. GTV effectively resolves the limitation in existing methods, which only rely on spatial information for removing the noise. Extensive experiments on synthetic and real DMRI data demonstrate that GTV can remove the noise effectively and outperforms state-of-the-art methods.

1. Introduction

DMRI is a unique way in the *in vivo* characterization of anatomical connectivity in the human brain [1]. However, it is frequently impeded by pronounced thermal noise due to its echo-planar acquisition strategies and low signal-to-noise ratio (SNR) [2]. One practical way to reduce the noise variance of DMRI data is averaging several acquisitions [3]. Therefore, it requires an extended period that is not suitable for clinical settings. On the other hand, many postprocessing algorithms have been applied in DMRI denoising to improve the data quality. Most of them can be roughly grouped into transform-based and spatial approaches.

Transform-based denoising methods are frequently used in signal processing because it is generally believed that noise and signal are more easily distinguished in the transform domain. Nowak [4] proposed the wavelet domain filter for reducing Rician noise in MRI. This method filters the squared magnitude of MRI in that distribution changes from Rician to noncentral Chi (nc-χ). Wood and Johnson [5] proposed an MRI denoised strategy based on the wavelet packet. This method exceptionally performs well on removing Rician noise with low SNR since wavelet packet provides a more compact signal representation than single wavelet decomposition. Other transform-based approaches, including curvelet transform [6] and block-matching and 3D (BM3D) [7] denoising schemes, have been proven to tackle Rician noise in MRI data efficiently.

Spatial denoising methods utilize the neighbor information around the pixel to reduce the variance caused by noise. Henkelman [8] presented the pioneering study on estimating Gaussian noise in MR images and smoothened them by convolution filtering, but this method blurs sharp edges that provide crucial clinical information. Gerig et al. [9] applied nonlinear anisotropic diffusion (AD) filters to 2D and 3D MRI denoising. AD filters can remarkably leverage the MRIs' quality in terms of persevering boundaries. However, AD methods assume that the processed MRI are piecewise constant and the noise is

Gaussian distributed with zero means. Both of those mentioned above are not applicable for MRI. Tong et al. [10] proposed an improved AD algorithm that used an adaptive scheme to select parameters to reduce Rician noise in MRI. To avoid tuning numbers of parameters carefully, Awate and Whitaker [11] proposed a nonparametric neighborhood statistics technique for MRI denoising. They inferred the prior from the degraded MRI and the knowledge of the Rician noise model. The uncorrupted MRI statistics are modeled in a nonparametric Markov Random Field (MRF) and estimated by the expectation-maximization algorithm. Experiments show that this method can well preserve important features in brain MRI. Fabio et al. [12] defined a 3D local Gaussian MRF (LGMRF) that allows tuning the filter in an unsupervised way. The LGMRF model can automatically implement the regularization to find the trade-off between noise reduction and detail preservation.

TV-based regularization [13] is one of the most influential denoising algorithms since it can effectively reduce the variance while preserving shape edges. Martin et al. [14] proposed a TV-Rician denoising model for MRI data by solving a semi-implicit numerical scheme. Liu et al. [15] extracted the noise map through a two-step wavelet-domain estimation method, then denoised the MRI data based on a generalized TV model. Later, the authors extended their work by combining none local mean (NLM) filters [16]. The aforementioned methods mainly process the data in the spatial domain (i.e., x-space) while ignoring that DMRI data consists of spatial and diffusion wavevector space. These denoising algorithms lead to new smoothing artifacts caused by averaging over differential oriented signals, particularly in the highly curved white matter structures.

Recently, Chen et al. have proposed two novel NLM-based denoise methodologies in joint $x - q$ space. Before NLM denoising, for each point in $x - q$ space, they (1) defined a spherical patch from which they extracted the rotation-invariant features patch matching [17] and (2) performed graph framelet transforms to extract robust rotation-invariant features after encoding the q-space sampling domain using a graph [18]. Graphs have the ability to well modeling the data with irregular and complex structures [19]. In this paper, we represent the joint $x - q$ space DMRI data with a graph and associate signal coefficients with graph nodes. Then, we consider the denoising DMRI data on the graph as an optimization problem, which seeks to minimize a TV function based on graph edge signals. The main contribution of our GTV denoising framework includes threefold: (i) the proposed method avoids computing rotation-invariant features; (ii) the minimizing variation procedure harnesses not only over spatial space but also angular space, allowing information to be shared across DW images for effective denoising; (iii) the information in angular space represents structures oriented in different directions, so our method is expected to be more effective for denoising without introducing new artifacts.

The remainder of this paper is organized as follows. We describe the detail of the proposed method in Section 2 and evaluate the effectiveness on synthetic data and public real data in Section 3. Section 4 shows our further discussion of this work. Finally, we conclude this work in Section 5.

2. Method

2.1. Graph Representation.
Given a graph $G = (\mathcal{V}, w)$, where $\mathcal{V} = \{v_n\}_{n=1}^{N}$ represents the set of N nodes and $w \in \mathbb{C}^{N \times N}$ is an affinity matrix, which characterizes the relationships between every pair of nodes on graph. The graph signal is a mapping function $\mathcal{V} \longrightarrow \mathbb{C}$ that associates signal coefficients $S_n \in \mathbb{C}$ to each node $v_n \in \mathcal{V}$ of the graph.

We show how to represent DMRI data using a graph. Specifically, given a sampling point (i, j) in spatioangular domain, which is determined by a spatial location $\mathbf{x}_i \in \mathbb{R}^3$ and a gradient direction $\mathbf{q}_j \in \mathbb{R}^3$ and diffusion weighting $b_j \propto \|\mathbf{q}_j\|^2$, we consider it as a node of the graph and define the affinity weight with another node (i', j') as [20]

$$w_{i,j;i}{}',j' := \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_i{}'\|_2^2}{\sigma_x^2}\right) \times \exp\left(-\frac{1 - \left\langle \mathbf{q}\wedge_j, \mathbf{q}\wedge_j{}'\right\rangle^2}{\sigma_q^2}\right)$$
$$\times \exp\left(-\frac{\left(\sqrt{b_j} - \sqrt{b_{j'}}\right)^2}{\sigma_b^2}\right), \tag{1}$$

where σ_x, σ_q, and σ_b are three parameters controlling the commitments of spatial, angular, and diffusion weight, respectively [21].

2.2. GTV of DMRI.
We define the GTV of DMRI signal S with respect to graph $\mathcal{G}$ as

$$\sum_{(i,j)} \left\|\nabla_{\mathcal{G}} S_{i,j}\right\|_1 = \sum_{(i,j)} \sum_{(i',j') \in \mathcal{N}_{i,j}} \sqrt{w_{i,j;i}{}',j'} \left\|S_{i,j} - S_i{}',j'\right\|_1, \tag{2}$$

where $S_{i,j}$ is a diffusion signal associated with spatial location $\mathbf{x}_i$, gradient direction $\mathbf{q}_j$, and diffusion weighting b_j. $\mathcal{N}_{i,j}$ is the spatioangular neighborhood of $S_{i,j}$. We set $w_{i,j;i}{}'j' = 0$ where $(i', j') \notin \mathcal{N}_{i,j}$; then, (2) becomes the standard l_1-norm $\|\nabla_{\mathcal{G}}(S)\|_1$.

To denoise the DMRI data, we propose

$$\min_S \frac{1}{2}\left\|S - S'\right\|_2^2 + \lambda\|\nabla_{\mathcal{G}}(S)\|_1, \tag{3}$$

where the quadratic fidelity term is to force the underlying clean data S to stay close to the acquired noisy DMRI data S', and tuning parameter λ controls the trade-off between two parts of (3) [22]. Since (1) assumes the nearer nodes have a higher weight, (3) ensures they have lower dissimilarities. It is worth noting that our GTV takes the neighborhood region compared to traditional TV, which only considers the neighbors of a sample along Cartesian axes.

2.3. Optimization.
Graph gradient $\nabla \mathcal{G}$ is a linear operator (hereafter denoted as L) which associates the DMRI signal to the corresponding node, and consequently, (3) is usually
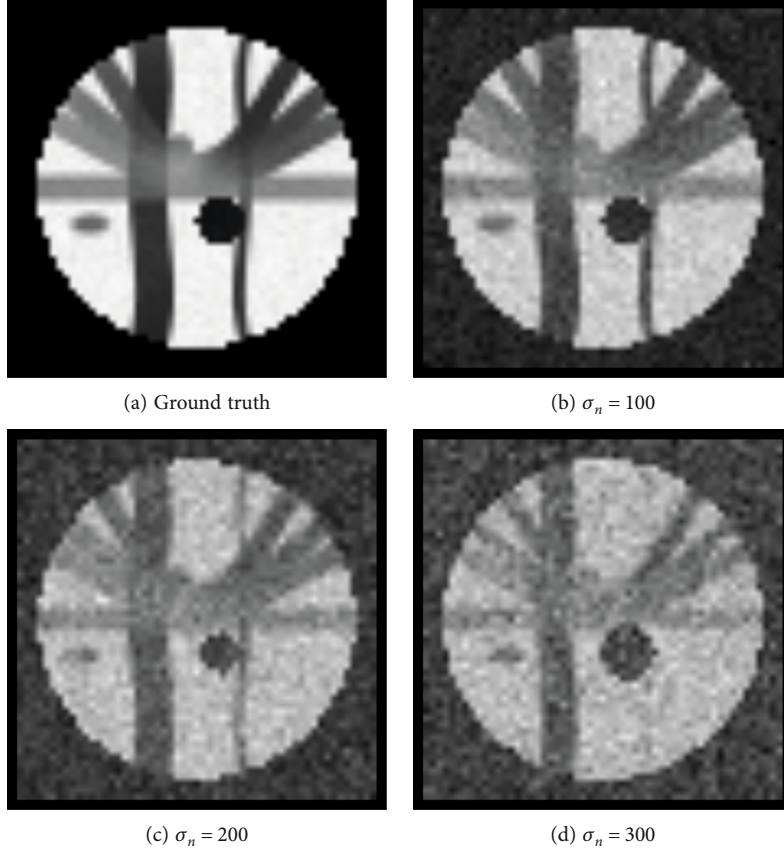
(a) Ground truth

(b) $\sigma_n = 100$

(c) $\sigma_n = 200$

(d) $\sigma_n = 300$

Figure 1: Some examples of synthetic data degraded by Gaussian noise with different deviations.

Input: $U_0 = S'$, $V_0 = \nabla_{\mathscr{G}}(S')$
Output: S
for $n = 0, 1, \cdots, N-1$ **do**
$\quad Z_n = U_n - \tau \nabla_f(U_n)$
$\quad T_n = Z_n - \tau L^T V_n$
$\quad Q_n = \text{prox}_{\tau, g^*}(V_n + \tau L T_n)$
$\quad P_n = Z_n - \tau L^T Q_n$
$\quad (U_{n+1}, V_{n+1}) = (U_n, V_n) + \tau(P_n, Q_n) - (U_n, V_n)$
$\quad$ **if** $\|U_{n+1} - U_n\|_2^2 / \|U_n\|_2^2 < \varepsilon$ and $\|V_{n+1} - V_n\|_2^2 / \|V_n\|_2^2 < \varepsilon$
$\quad\quad\quad$ **BREAK**
$S = U_{n+1}$

Algorithm 1: Forward-backward-based primal-dual for GTV optimization.

referred to the *primal problem* since it can be expressed as the following form:

$$\min_x f(x) + g(Lx), \qquad (4)$$

where $f(x)$ represents the fidelity term in (3) and it is convex and differential with a β-Lipschitz continuous gradient $\nabla_f(x)$. (4) is related to the following dual problem [23]:

$$\min_v f^*(-L^T v) + g^*(v), \qquad (5)$$

where f^* is the conjugate of function f. It is worth noting that only a small number of neighbors are taken into account in the proposed graph (i.e., the number of edges is much smaller than that of nodes); therefore, (5) is much easier to solve than (4). The second term $g(D) = \lambda \|D\|_1$ where $D = L(S)$ is convex and has proximity operator [24].

$$\text{prox}_{\tau, g}(D) = \text{sgn}(D) \circ \max(|D| - \tau \gamma, 0). \qquad (6)$$

Note that τ is the step size in the iterative optimization process and is typically set to the inverse of the Lipschitz constant, and $\circ$ denotes the Hadamard product. With the
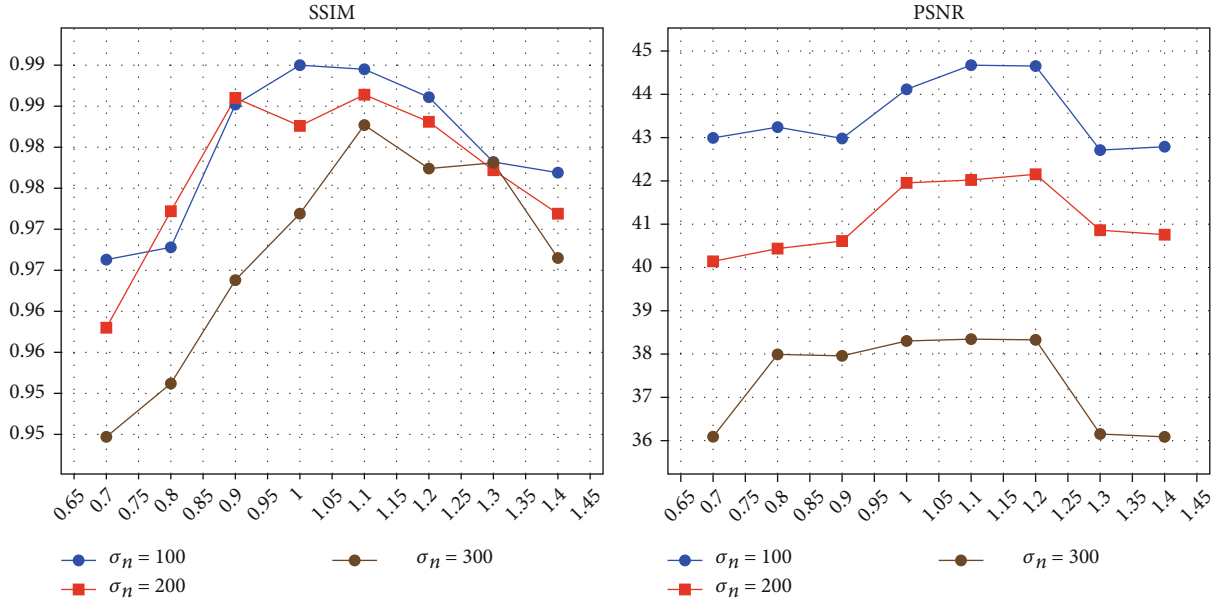
FIGURE 2: Selecting of σ_q via indicators including SSIM and PSNR.
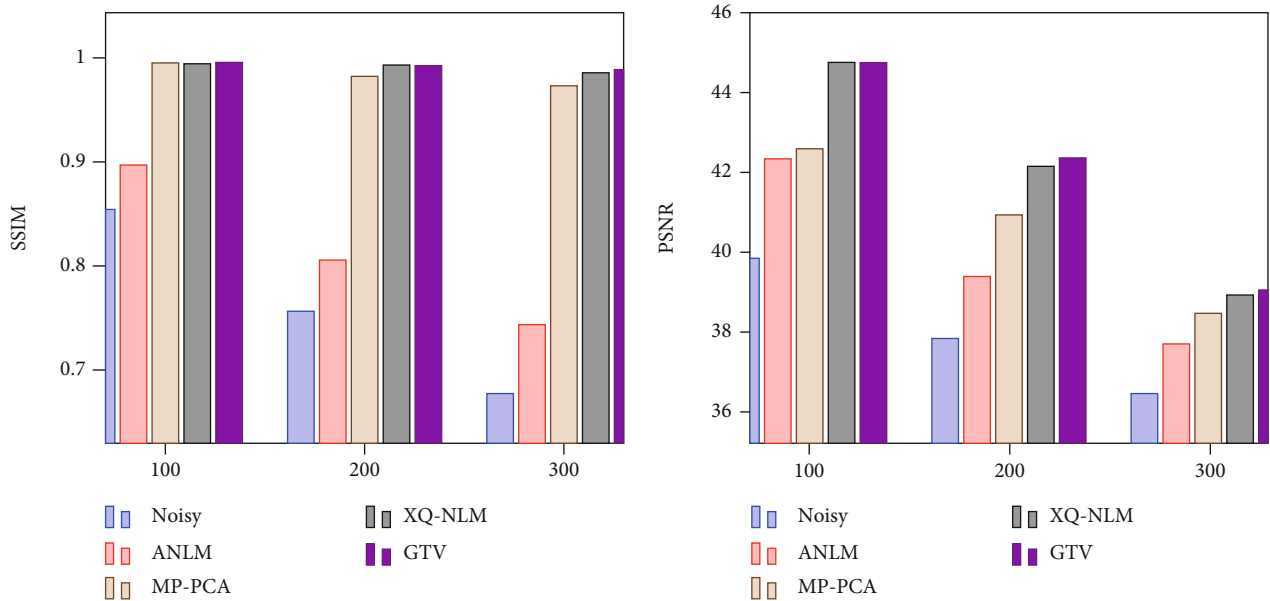
FIGURE 3: Quantitative comparison of denoising performance, where x-axis represents standard deviation of noise.

aid of these tools, we can recover the unique solution of (3) via a forward-backward-based primal-dual algorithm [23, 24].

3. Experiments

3.1. Datasets. We evaluate the proposed GTV based on a synthetic dataset and a subject from Traveling Human Phantom (THP) that was collected for a multisite neuroimaging reliability study [25]. The research was completed according to the endorsed guidelines.

In order to evaluate the denoising performance of the proposed method for various fiber structures, we use phantomαs to produce a synthetic dataset [26]. The fiber geometric configuration is completely identical to that in ISBI 2013 HARDI challenge, and the gradient direction files are consistent with the THP dataset, i.e., using $b = 1000$ s/mm^2 with a total of 30 noncollinear gradient directions. Image dimensions are $55 \times 55 \times 55$ with $2 \times 2 \times 2$ mm^3 resolution.

THP data was acquired using the Siemens 3T TrioTim MR scanner with the following imaging protocol: 128×128 imaging matrix, $2 \times 2 \times 2$ mm^3 resolution, TE = 92 ms,
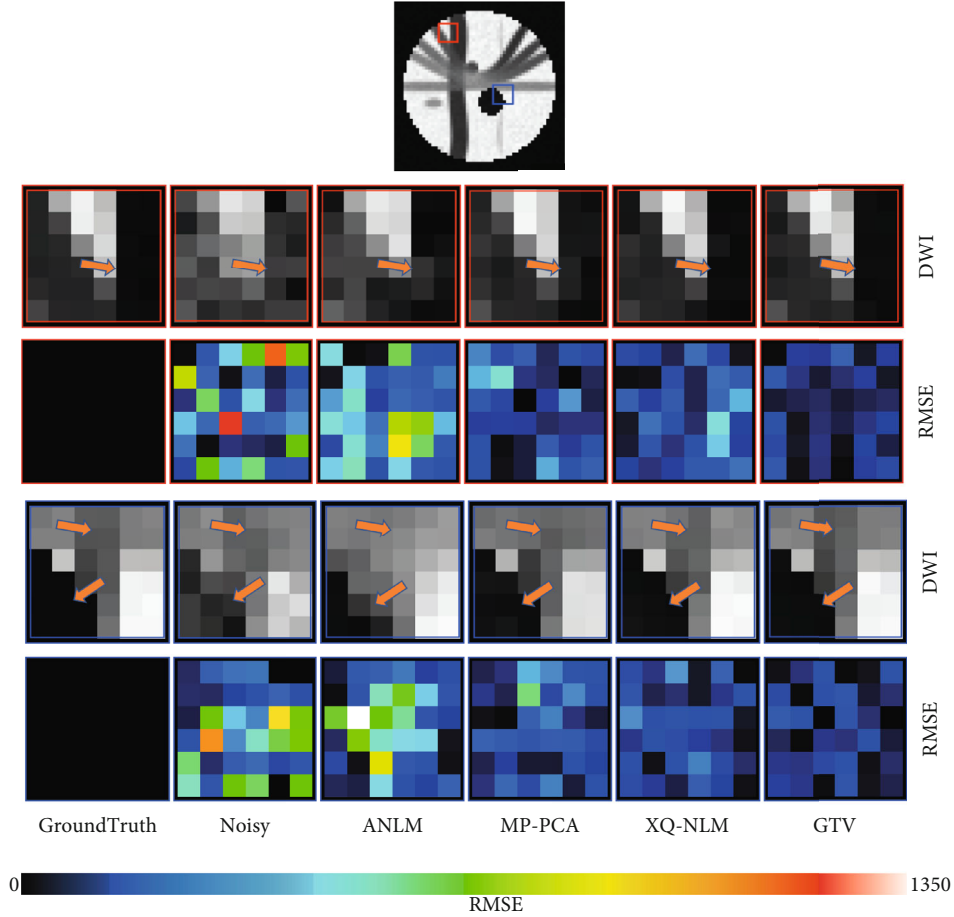
FIGURE 4: Regional close-up views of DW images and RMSE maps.

and TR = 1,200 ms. b values and gradient direction are the same as the synthetic data.

To evaluate the performance of our proposed approach, we add Gaussian distributed noise with zero mean and different deviation of $\sigma_n = 100, 200$ and 300 to the synthetic data. Some example images are shown in Figure 1.

3.2. Parameter Settings. We have implemented Algorithm 1 with the help of open-source Python packages including PyGSP [27] and PyUNLocBox [28], which, respectively, facilitate a number of operations on graph and solvers on nondifferentiable convex optimization problems using proximal splitting methods [29]. The relative tolerance stopping criterion ε is set to 10^{-3}, and the maximum number of iterations is set to $N = 200$ [24]. Both of them are set automatically by PyUNLocBox. Larger values of tuning parameter λ lead to smoother solutions of (3) whereas smaller values emphasize better fitting to the required measurements. Hence, we set $\lambda = 0.1$.

The proposed approach regards each sampling point in the spatioangular domain as a node. Therefore, DMRI data will lead to a large graph, requiring colossal memory, and GTV is difficult to converge. To overcome this weakness, we construct the graph by dividing the DMRI data into overlapping patches in spatial space, then optimize each patch simultaneously. Although a larger patch can capture more
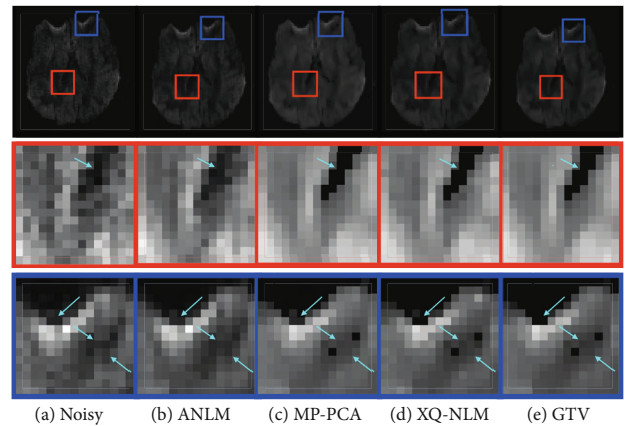

FIGURE 5: Comparison of edge-preserving performances among different denoising algorithms.

context information, we set the size of the patch to $5 \times 5 \times 5$ and the overlap step to 2 according to our compute unit with 12 GB memory.

As mentioned in (1), construction of the graph involves three parameters: σ_x, σ_q, and σ_b. We normalize each exponent in exponential functions to $[0, 1]$ as it was suggested in [21]. The maximum distance in a $5 \times 5 \times 5$ patch equals to 64, and therefore, we set $\sigma_x = 8$. We fix parameters $\sigma_x =$

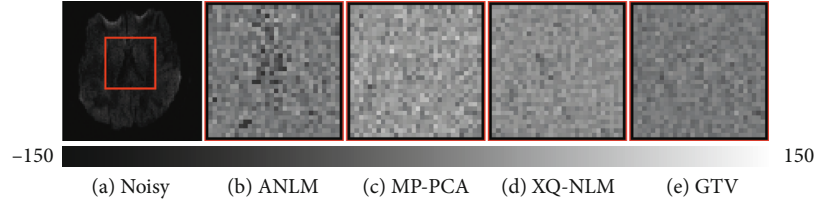(a) Noisy   (b) ANLM   (c) MP-PCA   (d) XQ-NLM   (e) GTV

FIGURE 6: Residual maps. Representation of whether structural information is removed after denoising.

8 and $\sigma_b = 1$, then evaluate the influence of parameter σ_q by measures of peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [30]. As indicated in Figure 2, the proposed method will achieve the best performance when $\sigma_q = 1.1$.

3.3. Methods for Comparison. We compared GTV with the following state-of-the-art methods:

(1) Adaptive nonlocal means (ANLM): ANLM [31] is an improved version of the NLM algorithm, which is adaptive soft coefficient matching. Based on [31], we set the patch radius to 2

(2) Marcenko-Pastur PCA (MP-PCA): MP-PCA [32] classifies the principal components of the observed DMRI signal based on Marcenko-Pastur distribution, then removes them as noise. Both ANLM and MP-PCA are implemented by DIPY [33]

(3) XQ nonlocal means (XQ-NLM): XQ-NLM [17] denoises the signal via weighted averaging of self-similar information, which is defined in the spatial-angular domain. The parameters are set to the default values as suggested in [17]

We transform the Rician signal to its Gaussian-distributed counterpart using Özarslan et al.'s method [34]. For a fair comparison, we estimate the nonstationary noise by MP-PCA for ANLM, XQ-NLM, and GTV. Then, we determine the standard deviation of stationary noise via PIESNO [35].

3.4. Results. We first quantitatively show the denoise performance through measures including SSIM and PSNR in Figure 3 where the x-axial represents the noise level. Although GTV has achieved slightly higher SSIM than MP-PCA and XQ-NLM, all of them are close to 1, meaning that denoised data structures closely resemble the original synthetic data. Regarding PSNR, GTV and XQ-NLM have outperformed MP-PCA and ANLM, which indicates filtering in spatial-angular space can remove more noise.

We randomly select two regions of interest (ROI) marked by red and blue rectangles, respectively. Close-up views of DW images and RMSE maps are presented in Figure 4. The first and third rows of DW images show that MP-PCA, XQ-NLM, and GTV have outperformed ANLM. With the help of RMSE maps in the below rows, we can observe that the reconstruction error of GTV is smaller than MP-PCA and XQ-NLM, which indicates GTV can restore the image more precisely.
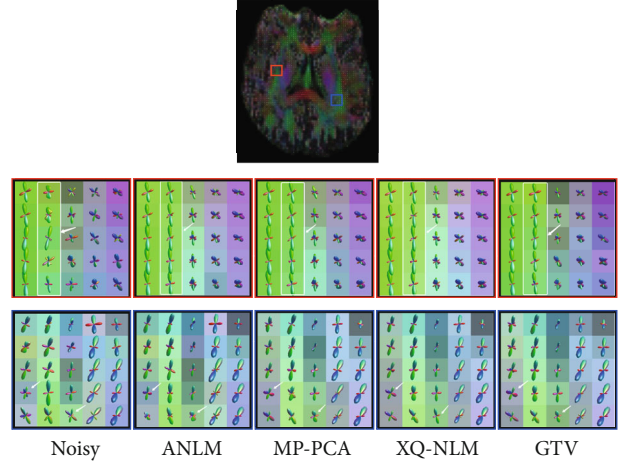

FIGURE 7: Fiber ODF comparison of various denoising algorithms.

One disadvantage of NLM-based methods is that the denoising process may add method noise which usually blurs the outputs. The region marked by the red square in Figure 5 is a part of the boundaries between white matter and the ventricles. Due to the consideration of information in the q-space, XQ-NLM and GTV have recovered the edges as MP-PCA does. In contrast, ANLM has removed some noise, but the remaining is far from perfect. The region marked by the blue square in Figure 5 includes a contingent of boundaries, and it shows that GTV reconstructs the internal details while its boundary is the clearest.

Residual maps are used to evaluate whether structural information has been removed after denoising. In Figure 6, residual maps of MP-PCA, XQ-NLM, and GTV have less structure information than that of ANLM. This observation provides evidence that ANLM has lower edge-preserving abilities in Figure 5 from another perspective.

We further evaluated the denoising performances through fiber ODFs. THP [25] is single-shell (with only a single $b = 0$ image) DMRI data. Response functions for single-fiber white matter (WM), as well as gray matter (GM) and cerebrospinal fluid (CSF), were estimated from the denoised data using an unsupervised method [36]. Then, we performed Single-Shell 3-Tissue Constrained Spherical Deconvolution (SS3T-CSD) [37] to obtain WM-like FODs as well as GM-like and CSF-like compartments in all voxels using MRtrix3Tissue, which is a fork of MRtrix3 [38]. The first row in Figure 7 shows a slice of FOD-based directionally encoded color (DEC) maps [39]. The fiber ODFs in the region marked by the red rectangle indicate more coherence after denoising. The region marked by a blue rectangle

indicates that the proposed GTV has more effectively reduced spurious fiber peaks that result from noise.

4. Discussion

GTV can remove the noise more effectively, which may be attributed to the following aspects:

(1) DMRI uses q-space information to characterize the direction and scale of the diffusion for water molecules in the tissues. The proposed GTV takes advantage of neighborhood similarity information in the spatioangular domain while constructing a graph

(2) Graph representation is a versatile model where nodes are associated with DMRI signal intensity, and edges reflect structural information

Although the proposed algorithm is very effective for denoising, it is not superior to the recent novel DMRI denoising method Patch2Self [40] which denoises DMRI with a self-supervised deep learning strategy. Furthermore, there may be two limitations in this study. First, taking each sample in spatioangular space as an independent node leads to a big graph, which requires a large amount of memory for storing graph properties, including affinity weighted matrix and Fourier basis. It can be relieved by dividing the data into patches [21]. Second, optimization of total variation takes a relatively long time. We speed up the denoising by using the multicore CPUs to optimize each patch simultaneously. Possible solutions include refactoring the software using C++ and combining a two-step optimization approach [16] developed to solve the resulting convex denoising GTV model.

5. Conclusion

In this study, we formulated denoising processing as an optimization problem, finding the DMRI data with minimal graph total variation. Both spatioangular information of DMRI data were incorporated to construct the graph, which significantly contributed to this paper. The performances of our proposed method were assessed via experiments on synthetic and real data. Numerical results demonstrate that GTV outperforms various current state-of-the-art approaches in terms of preserving edges and removing noise. Future works may extend GTV to patch GTV that associates values in a patch with a graph node.

Data Availability

The real data is downloaded from https://openneuro.org/datasets/ds000206/versions/1.0.0/download.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

References

[1] O. Sporns, G. Tononi, and R. Kötter, "The human connectome: a structural description of the human brain," *PLoS Computational Biology*, vol. 1, no. 4, article e42, 2005.

[2] E. Kirilina, A. Lutti, B. A. Poser, F. Blankenburg, and N. Weiskopf, "The quest for the best: the impact of different EPI sequences on the sensitivity of random effect fMRI group analyses," *NeuroImage*, vol. 126, no. 11, pp. 49–59, 2016.

[3] H. Johansen-Berg and T. E. J. Behrens, *Diffusion MRI: From Quantitative Measurement to In Vivo Neuroanatomy*, Academic Press, 2013.

[4] R. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," *IEEE Transactions on Image Processing*, vol. 8, no. 10, pp. 1408–1419, 1999.

[5] J. C. Wood and K. M. Johnson, "Wavelet packet denoising of magnetic resonance images: importance of rician noise at low snr," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 41, no. 3, pp. 631–635, 1999.

[6] J. Ma and G. Plonka, "Combined curvelet shrinkage and nonlinear anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2198–2206, 2007.

[7] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[8] R. Mark and Henkelman, "Measurement of signal intensities in the presence of noise in mr images," *Medical Physics*, vol. 12, no. 2, pp. 232-233, 1985.

[9] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of mri data," *IEEE Transactions on Medical Imaging*, vol. 11, no. 2, pp. 221–232, 1992.

[10] C. Tong, Y. Sun, N. Payet, and S.-H. Ong, "A general strategy for anisotropic diffusion in MR image denoising and enhancement," *Magnetic Resonance Imaging*, vol. 30, no. 10, pp. 1381–1393, 2012.

[11] S. P. Awate and R. T. Whitaker, "Feature-preserving mri denoising: a nonparametric empirical Bayes approach," *IEEE Transactions on Medical Imaging*, vol. 26, no. 9, pp. 1242–1255, 2007.

[12] F. Baselice, G. Ferraioli, and V. Pascazio, "A 3d mri denoising algorithm based on Bayesian theory," *Biomedical Engineering Online*, vol. 16, no. 1, pp. 1–19, 2017.

[13] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[14] A. Martin, J.-F. Garamendi, and E. Schiavi, "Mri tv-rician denoising," in *International Joint Conference on Biomedical Engineering Systems and Technologies*, pp. 255–268, Springer, 2013.

[15] R. W. Liu, L. Shi, W. Huang, J. Xu, S. C. H. Yu, and D. Wang, "Generalized total variation-based MRI Rician denoising

model with spatially adaptive regularization parameters," *Magnetic Resonance Imaging*, vol. 32, no. 6, pp. 702–720, 2014.

[16] R. W. Liu, L. Shi, S. C. H. Yu, and D. Wang, "A two-step optimization approach for nonlocal total variation-based Rician noise reduction in magnetic resonance images," *Medical Physics*, vol. 42, no. 9, pp. 5167–5187, 2015.

[17] G. Chen, W. Yafeng, D. Shen, and P.-T. Yap, "Noise reduction in diffusion MRI using non-local self-similar information in joint x–q space," *Medical Image Analysis*, vol. 53, pp. 79–94, 2019.

[18] G. Chen, B. Dong, Y. Zhang, W. Lin, D. Shen, and P.-T. Yap, "Denoising of diffusion MRI data via graph framelet matching in x-q space," *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2838–2848, 2019.

[19] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[20] G. Chen, W. Yafeng, D. Shen, and P.-T. Yap, "XQ-NLM: denoising diffusion MRI data via x-q space non-local patch matching," in *In Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 587–595, Springer, 2016.

[21] Y. Hong, J. Kim, G. Chen, W. Lin, P.-T. Yap, and D. Shen, "Longitudinal prediction of infant diffusion MRI data via graph convolutional adversarial networks," *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2717–2725, 2019.

[22] W. Lu, J. Duan, Z. Qiu, Z. Pan, R. W. Liu, and L. Bai, "Implementation of high-order variational models made easy for image processing," *Mathematical Methods in the Applied Sciences*, vol. 39, no. 14, pp. 4208–4233, 2016.

[23] N. Komodakis and J.-C. Pesquet, "Playing with duality: an overview of recent primal?dual approaches for solving large-scale optimization problems," *IEEE Signal Processing Magazine*, vol. 32, no. 6, pp. 31–54, 2015.

[24] F. Mahmood, N. Shahid, U. Skoglund, and P. Vandergheynst, "Adaptive graph-based total variation for tomographic reconstructions," *IEEE Signal Processing Letters*, vol. 25, no. 5, pp. 700–704, 2018.

[25] V. A. Magnotta, T. Joy, D. L. Matsui et al., *DWI Traveling Human Phantom Study*, OpenNeuro, 2018.

[26] E. Caruyer, A. Daducci, M. Descoteaux, J. C. Houde, J. P. Thiran, and R. Verma, *Phantomas: a flexible software library to simulate diffusion MR phantoms*, In ISMRM, 2014.

[27] M. Defferrard, L. Martin, R. Pena, and N. Perraudin, *Pygsp: graph signal processing in python*, 2017, https://github.com/epfl-lts2/pygsp.

[28] M. Defferrard, R. Pena, and N. Perraudin, *Pyunlocbox: Optimization by Proximal Splitting*, 2017, https://github.com/epfl-lts2/pyunlocbox/.

[29] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212, Springer, 2011.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[31] P. Coupé, J. V. Manjón, M. Robles, and D. L. Collins, "Adaptive multiresolution non-local means filter for three-dimensional magnetic resonance image denoising," *IET Image Processing*, vol. 6, no. 5, pp. 558–568, 2012.

[32] J. Veraart, E. Fieremans, and D. S. Novikov, "Diffusion MRI noise mapping using random matrix theory," *Magnetic Resonance in Medicine*, vol. 76, no. 5, pp. 1582–1593, 2016.

[33] E. Garyfallidis, M. Brett, B. Amirbekian et al., "Dipy, a library for the analysis of diffusion MRI data," *Frontiers in Neuroinformatics*, vol. 8, p. 8, 2014.

[34] E. Özarslan, C. G. Koay, and P. J. Basser, "A signal transformational framework for breaking the noise floor and its applications in MRI," *Journal of Magnetic Resonance*, vol. 197, no. 2, pp. 108–119, 2009.

[35] C. G. Koay, E. Özarslan, and C. Pierpaoli, "Probabilistic Identification and Estimation of Noise (PIESNO): a self- consistent approach and its applications in MRI," *Journal of Magnetic Resonance*, vol. 199, no. 1, pp. 94–103, 2009.

[36] T. Dhollander, R. Mito, D. Raffelt, and A. Connelly, "Improved white matter response function estimation for 3-tissue constrained spherical deconvolution," *In ISMRM*, vol. 27, p. 555, 2019.

[37] T. Dhollander and A. Connelly, "A novel iterative approach to reap the benefits of multi-tissue CSD from just single-shell (+b=0) diffusion MRI data," *In ISMRM*, vol. 3010, p. 5, 2016.

[38] J.-D. Tournier, R. Smith, D. Raffelt et al., "_MRtrix3_ : a fast, flexible and open software framework for medical image processing and visualisation," *NeuroImage*, vol. 202, article 116137, 2019.

[39] T. Dhollander, R. E. Smith, J. D. Tournier, B. Jeurissen, and A. Connelly, "Time to move on: an FOD-based DEC map to replace DTI's trademark DEC FA," *In ISMRM*, vol. 23, p. 1027, 2015.

[40] S. Fadnavis, J. Batson, and E. Garyfallidis, "Patch2self: denoising diffusion MRI with self-supervised learning," in *Advances in Neural Information Processing Systems, Volume 33*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., pp. 16293–16303, Curran Associates Inc., 2020.