J. Indian Inst. Sci.
A Multidisciplinary Reviews Journal
ISSN: 0970-4140 Coden-JIISAD

REVIEW ARTICLE

Recent Computational Advances in Denoising for Magnetic Resonance Diffusional Kurtosis Imaging (DKI)

Calvin B. Shaw[1*] *and Jens H. Jensen*[2,3]

Abstract | Magnetic resonance imaging (MRI) is widely used in clinical practice and medical research for the assessment of disease. Magnetic resonance diffusional kurtosis imaging (DKI) is a specific MRI technique that is useful for quantifying microstructural properties of biological tissues, particularly in brain. However, images derived with DKI can be sensitive to noise, as the MRI sequences needed for DKI strongly attenuate the signal. To mitigate this inherent noise sensitivity of DKI, advanced denoising methods maybe applied. Although a variety of denoising approaches have been considered in the broad context of MRI, the specific performance of these methods for DKI has not yet been thoroughly investigated. In this review, we examine three different denoising strategies for DKI—Gaussian filtering, non-local means filtering, and a local principal components analysis technique. These three denoising methods are compared qualitatively in terms of their abilities to increase image fidelity and to remove noise bias for the DKI-derived parametric maps.

Keywords: *Diffusional kurtosis imaging, Denoising, Non-local means, Principal component analysis, Gaussian filtering*

1 Introduction

Diffusion magnetic resonance imaging (MRI) is a non-invasive and non-ionizing imaging method that quantifies the diffusion process of water molecules in biological tissues.[1] Diffusion tensor imaging (DTI), the most commonly used diffusion MRI technique, relies on an assumption that the water diffusion displacement probability density function (dPDF) has a Gaussian form.[2] While this assumption holds quite accurately for homogeneous solutions, the PDF in many biological tissues, including brain, can deviate substantially from Gaussianity, as a result of diffusion barriers (e.g., plasma membranes) and of microscopic heterogeneity due to a diversity of cellular compartments.[3, 4] Diffusional kurtosis imaging (DKI) extends DTI in order to capture non-Gaussian diffusion effects by estimating the kurtosis of the dPDF, and it has thus emerged as an advanced diffusion MRI technique for characterizing tissue microstructure in brain, as well as in other organs.[5–7] With DKI, promising results in neurological applications have been obtained for brain tumors,[8, 9] trauma,[10, 11] Alzheimer's disease,[12, 13] attention deficit hyperactivity disorder,[14, 15] stroke,[16, 17] and epilepsy.[18, 19] As DKI is being increasingly applied, it is important to optimize the signal processing aspects of this approach.

Diffusion MRI typically uses echo planar imaging (EPI) MRI sequences that suffer from relatively poor signal-to-noise-ratio (SNR), because the required long echo times and strong diffusion encoding gradient pulses attenuate the signal.[20] Moreover, DKI tends to be more noise sensitive than DTI, as it uses higher maximum diffusion weightings (b values) and a more complicated signal model to fit the data.[5–7] Low SNR may substantially impact both the precision and accuracy diffusion metrics estimated with DKI, with a noise bias arising since the metrics depend nonlinearly on the signal.[7, 21] As a consequence, several sophisticated strategies have been proposed to reduce noise contamination in

Signal processing aspects: In this review, we refer to signal processing aspects as artifacts arising due to systematic noise such as low SNR, field inhomogeneities, etc.,

Kurtosis: The measure of non-Gaussianity is defined as the excess kurtosis, which reflects the heterogeneity of the diffusion environment.

[1] *Department of Radiology, University of California, Davis, Sacramento, USA.* [2] *Center for Biomedical Imaging, Medical University of South Carolina, Charleston, SC, USA.* [3] *Department of Radiology and Radiological Science, Medical University of South Carolina, Charleston, SC, USA.*
* *cbshaw@ucdavis.edu*

Springer IISc Press

DKI measurements. One approach is to optimize the experimental design of the DKI acquisition, thereby increasing the precision of the estimated diffusion parameters.[22, 23] Another approach is to repeat the data acquisition multiple times and increase the SNR by signal averaging.[24] However, as the SNR only increases in proportion to the square root of the total acquisition time, such optimization techniques may not be a viable option from a clinical perspective. Alternatively, several post-processing techniques have been proposed that aim at reducing noise, which for diffusion MRI is both non-Gaussian (e.g., Rician) and spatially varying (the non-Gaussian nature of the signal noise should not be confused with the non-Gaussian nature of diffusion, as these are two independent statistical processes).

Post-processing techniques: Post-processing techniques do not alter the scan-time and are computationally efficient, making it a feasible approach.

Denoising methods proposed by several different groups for diffusion MRI can be broadly classified into two categories. First, one may apply denoising directly to the complex *k*-space data available from the scanners. This approach generates diffusion-weighted (DW) images through appropriate image reconstruction algorithms that accounts for the statistical distribution of noise embedded in the complex *k*-space data.[25–27] Although this approach can be effective in terms of reducing noise in the DW images, it is computationally intensive. One of the main bottlenecks is with the reconstruction time associated in estimating DW image volumes from each coil/channel and then combining them.[26] To overcome this limitation, a second strategy aims to reduce noise only in DW images that have already been reconstructed by the MRI scanner. This approach is commonly called the post-processing denoising step in the calculation of diffusion parameters.

There are two main types of post-processing denoising techniques: (a) spatial-domain-based filters and (b) transform-based filters. The spatial-domain-based filters are intended to work directly on the voxels of raw DW images obtained from the scanner. Some of the commonly used spatial domain filters include the Gaussian filters, which use a Gaussian probability density function to average voxels.[28] An alternative approach developed by Rudin et al.[29] to minimize the L1-norm of the gradient of the image (total variation norm) is effective in terms of preserving the edges in the anatomy.[30] Yet another spatial domain noise reduction technique that has received considerable attention is non-local means (NLM) filtering, which exploits the similarities in structures within a slice (2D) or volume (3D) through weighted averaging.[31–33] On the other hand, transform-based filters use an appropriate transformation basis on the pixels or voxels to perform denoising. Two transformations that have yielded effective results are principal component analysis (PCA)[34–36] and wavelet transformation-based denoising.[37] These transformations help in describing the image features spatially or spectrally by employing wavelet coefficients or eigenvalues to filter unwanted signal.

Gaussian probability density function: This Gaussian probability density function in this context refers to the distribution of the signal noise and not the diffusion of water molecules.

It is important to note that these methods have primarily been applied to either just raw DW images or to DTI. Although one recent study has considered the NLM method for DKI,[33] the comprehensive assessment of these approaches in this context is still pending. Here we consider denoising for DKI using three different methods: (a) Gaussian filtering, (b) NLM filtering, and (c) PCA denoising based on the universal Marchenko–Pastur law. We discuss the underlying theory for these three techniques, including practical considerations, and we compare how they perform for in vivo human brain data by adding simulated noise.

2 DKI Signal Model

Diffusional kurtosis imaging is an extension of the familiar DTI technique that allows for the estimation of the diffusional kurtosis and related diffusion metrics. The signal model for DTI is simply monoexponential decay, [1, 2] given as

$$S(\mathbf{n},\ b) = ((S_0 e^{-bD_{\text{app}}})^2 + \eta^2)^{\frac{1}{2}}, \tag{1}$$

where $S(n,\ b)$ is the DW signal with a diffusion weighting b (the b value) along a diffusion encoding direction, n, and S_0 is the DW signal with $b = 0$. D_{app} is the apparent diffusion coefficient (ADC) and η^2 is the noise variance. Here a Rician noise correction has been added to the signal model according to the prescription of Gudbjartsson and Patz.[39] Equation (1) is typically cast as a linear tensor estimation problem, where D_{app} results in a directional diffusivity profile encoded by the 3×3 diffusion tensor.[2] Tensor-derived diffusion metrics for DTI include mean diffusivity (MD, average of directional diffusivities along all directions), axial diffusivity ($D_{\parallel}$, the maximum ADC), radial diffusivity ($D_{\perp}$, average diffusivity perpendicular to the direction of maximum ADC), and fractional anisotropy (FA, measure of anisotropy for the ADC).

To account for the non-Gaussianity of the diffusion process, an apparent excess kurtosis K_{app} may be added to Eq. (1) in the following way:

$$S(\mathbf{n}, b) = ((S_0 e^{-bD_{\text{app}} + \frac{1}{6}b^2 D_{\text{app}}^2 K_{\text{app}}})^2 + \eta^2)^{\frac{1}{2}}, \tag{2}$$

which is the standard signal model for DKI. By fitting to the experimental data, Eq. (2) can be used to estimate both D_{app} and K_{app}. The directional dependence of D_{app} is once again described by the 3×3 diffusion tensor, while the dependence of K_{app} is characterized by a $3 \times 3 \times 3 \times 3$ kurtosis tensor.[7, 38] In addition to the metrics listed above based on the diffusion tensor, DKI also allows for diffusion metrics that are defined in an analogous fashion in terms of the kurtosis tensor, such as mean kurtosis (MK), axial kurtosis ($K_{\parallel}$), radial kurtosis ($K_{\perp}$), and kurtosis fractional anisotropy (KFA). In order for the DKI signal model of Eq. (2) to be sufficiently sensitive to non-Gaussian diffusion effects so as to facilitate a precise estimate of K_{app}, higher maximum b values are typically needed with DKI than for DTI. Since stronger diffusion weightings decrease the DW signal, DKI is more prone to noise bias than DTI when the SNR is low.

3 Noise Characteristics in DKI

The complex raw DW data (k-space) acquired using the echo planar imaging (EPI) sequence is often influenced by various factors, namely bulk subject motion, eddy currents, inhomogeneities in the main magnetic field, susceptibility-induced distortion, signal dropouts, incomplete fat suppression, flow artifacts, etc.[2, 5–7, 40] However, in many cases the principal confounding factor that degrades the DW image quality, and consequently derived diffusion metrics, is thermal noise. The noise in k-space is generally assumed to be an uncorrelated Gaussian noise with zero mean and equal variance in real and imaginary part due to the linearity associated with the Fourier transform. However, the common practice is to discard the phase information and reconstruct DW images just from the magnitude information. The truncation of the phase data (a nonlinear transformation) results in altering the probability density function of the noise and is found to obey a Rician distribution in a single-channel coil MRI setup and non-central chi distribution in a multi-channel setup when the different channels are combined via sum of squares. Consequently, the noise in DWI is frequently modeled by a Rician distribution. If the DW image voxel intensity in the absence of noise is represented by X and the measured voxel intensity is Y, then the Rician probability distribution for Y, [39] is given by

$$p_Y(Y) = \frac{Y}{\sigma^2} e^{-\frac{X^2+Y^2}{2\sigma^2}} I_0\left(\frac{X \cdot Y}{\sigma^2}\right), \tag{3}$$

where I_0 represents the zeroth order Bessel function of the first kind and σ represents the standard deviation of the Gaussian noise contained in the real and imaginary parts of the k-space. For high SNR ($X/\sigma \gg 1$), $p_Y(Y)$ is well approximated by a Gaussian distribution. However, for low SNR ($X/\sigma \cong 1$), this Rician probability distribution differs markedly from a Gaussian, which is the reason for the noise correction factors in Eqs. (1) and (2). Since DKI requires higher diffusion weighting than DTI, and therefore uses DW images with lower SNR, non-Gaussian effects are typically more important for DKI.

4 Denoising Methods for DKI

4.1 Gaussian Filtering

Until recently, Gaussian filtering has been widely employed to denoise the raw DWI images in DKI, as it is a simple and fast denoising approach. Although the structure of the Gaussian filter is trivial, we still briefly introduce the probability distribution associated with it. The filter is essentially a smoothing 2-D or 3-D convolution operator that blurs the images and removes high-frequency components, using a kernel with a normal distribution of mean, $\mu = 0$ and in a 2D case it is given as,[39]

$$p(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \tag{4}$$

where x and y are the distance of the pixels from the center of the kernel and σ is the standard deviation of the Gaussian kernel. The width of the Gaussian kernel is measured in terms of the full width at the half maximum (FWHM) given by $2\sqrt{2\ln 2\sigma}$, with a common practice being to choose the FWHM to be equal to the pixel size of the image.[39] The Gaussian filter is most effective when the DW images have sufficient SNR so the statistics are Gaussian, and it can be a poor choice to denoise DW images with low SNR, for which the statistics may be Rician as noted earlier.

4.2 Non-local Means (NLM) Filtering

The NLM filter is based on taking advantage of the natural redundancy or spatial correlation present in the DW images to remove noise.[31] If $\mathrm{NLM}(v)(x_i)$ represents the restored intensity of a pixel or voxel size (x_i) after the application of NLM filter, then the definition of $\mathrm{NLM}(v)(x_i)$ is the weighted average of all voxel intensities in the image, given as

$$\mathrm{NLM}(v)(x_i) = \sum_{x_j \in I} w(x_i, x_j) v(x_j), \tag{5}$$

where $v(x_j)$ denotes the intensity of voxel x_j and $w(x_i, x_j)$ is the weight that is assigned to pixel, x_i to recover x_j. The weight is a similarity score between voxels x_i and x_j that ranges from 0 to 1. The similar voxels are searched over a neighborhood called the patch. The usual choice to calculate the weight is through finding the distance between x_i and x_j denoted as $d(v(N_i), v(N_j))$, where N_i and N_j are the neighborhoods around voxels x_i and x_j and $w(x_i, x_j)$ is defined as

$$w(x_i, x_j) = \frac{1}{Z(i)} e^{-\frac{d(v(N_i), v(N_j))}{(h\hat{\sigma})^2}}, \quad (6)$$

where $\hat{\sigma}$ is the estimation standard deviation of the noise, h is the filtering parameter and $Z(i) = \sum_j w(x_i, x_j)$ is the normalization constant.[32, 33] Finally, the distance d is typically the Euclidean distance given as

$$d(v(N_i), v(N_j)) = \sqrt{\frac{1}{N} \sum_k^N \Delta(v(y_k), v(z_k))}, \quad (7)$$

where y_k and z_k represent the kth voxels in the neighborhoods N_i and N_j. A recent work proposed a vector-based NLM (vNLM) that treated the set of DW images as multi-spectral image with each voxel being a $(n + 1)$-dimensional vector, Δ was defined as $\Delta(v(y_k), v(z_k)) = \left|\left|v^i(y_k) - v^i(z_k)\right|\right|_2^2$, where $v^i(\cdot)$ denotes the ith component of the vector $v(\cdot)$.

4.3 Marchenko–Pastur-Based PCA (MPPCA)

Unlike Gaussian and NLM filtering, PCA-based filtering is a transform-domain approach that maps the redundant data in the spatial domain to eigenvalues and eigenvectors.[34, 35] The objective of PCA is to discriminate between signal-carrying and noise-only components by thresholding the eigenvalues. However, one of the challenges associated with a PCA-based approach is determination of the threshold values to distinguish signal and noise. Veraart and coworkers, proposed an extension to the traditional PCA-based technique by estimating the noise level, in turn overcoming the limitation of choosing the threshold in an automated fashion.[36] The estimation of the noise level relies on exploiting the fact that noise-only eigenvalues are expected to obey the universal Marchenko–Pastur law, which is derived from the random matrix theory for noisy covariance matrices. The noise-level estimation also accounts for Rician distribution in the data to threshold the eigenvalues.

The proposed algorithm had two objectives, one is to threshold eigenvalues and the other to reconstruct the denoised matrix. The algorithm is initiated by considering a redundant $M \times N$ data matrix $\boldsymbol{X}$ that has $P(\ll \min(M, N))$ principal components calculated through the singular value decomposition of $\boldsymbol{X}$:

$$\boldsymbol{X} = \sqrt{N}\boldsymbol{U} \wedge \boldsymbol{V}^{\boldsymbol{T}}, \quad (8)$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are left and right singular unitary matrices, respectively, and $\wedge$ is the diagonal matrix with singular values.[36] Veraart and coworkers[36] have derived a closed form expression for estimating the denoised matrix based on Marchenko–Pastur (MP) distribution given as

$$\hat{\boldsymbol{X}} = \sqrt{N}\boldsymbol{U}\tilde{\wedge}\boldsymbol{V}^{\boldsymbol{T}}, \quad (9)$$

where $\tilde{\wedge}$ is obtained by nullifying the singular or eigenvalues below a threshold, $\lambda \leq \lambda_+$. The last step of the algorithm is to simultaneously estimate the noise level (σ) in the data matrix $\boldsymbol{X}$ and significant signal components P simultaneously that determines λ_+.

In the following sections, we will analyze these three denoising techniques discussed above with the help of a human dataset and discuss their relative performance.

5 DKI Data and Post-Processing

To compare how the three noise-removal approaches introduced above perform, we acquired a DKI dataset from a healthy volunteer (male, 30 years). The subject was imaged on a Siemens Tim Trio 3T system using a single-shot DW EPI sequence with a twice-refocused spin echo. The protocol used for this study was approved by the Medical University of South Carolina and informed consent was obtained from the subject prior to participation in the scan. Three b values of 0, 1000, and 2000 s/mm^2 and 64 uniformly distributed diffusion encoding directions were used to estimate the diffusion and kurtosis tensor. The acquisition parameters used were TR = 7200 ms, TE = 103 ms, voxel size = 2 × 2 × 2 mm^3, number of slices = 59, matrix size = 110 × 110, bandwidth = 1352 Hz/Px, parallel imaging factor of 2 and a 32-channel head coil with adaptive combine mode. To obtain a DKI dataset that has minimal bias of noise, we acquired DKI datasets with three averages, and 25 images with no diffusion weighting (b_0 images). The total acquisition time for three averages was 48.0 min.

All the DWIs were coregistered to their corresponding averaged b_0 images using SPM8

(Wellcome Trust Center for Neuroimaging, London, UK) 6-parameter rigid body transformation. An average DKI dataset was created by averaging the three DKI datasets. The averaged DKI dataset was processed with the freely available software Diffusional Kurtosis Estimator (DKE) implemented in MATLAB (https://www.nitrc.org/projects/dke/). The diffusivity-based metrics MD, $D_{\parallel}$, $D_{\perp}$, FA, and kurtosis-based metrics MK, $K_{\parallel}$, $K_{\perp}$, and KFA were then estimated.

The DKI dataset with three averages was reconstructed with the DKE software to generate parametric maps for the four diffusivity-based and four kurtosis-based diffusion metrics. From now on, we will refer to this dataset and its corresponding parametric maps as "gold standard". In order to compare the performances of the three denoising algorithms, Rician noise was added to the raw DW images. Rician noise was added to all the DW images such that the SNR with respect to DW images at $b = 2000$ s/mm^2 was progressively decreased. The DW images at $b = 2000$ s/mm^2 are used as a reference since they have the lowest SNR compared to $b = 0$, and 1000 s/mm^2. We evaluated performance by considering three different noise cases of SNR = 30, 15, and 8. The gold standard dataset was considered as ground truth for all the three cases, in analyzing the diffusion and kurtosis parametric maps. For the Gaussian filtering, we implemented the isotropic Gaussian filter that is available with the DKE software. The NLM filter described by Zhou and coworkers,[33] VNLM-d was implemented using the DWI denoising software developed by Coupe (https://sites.google.com/site/pierrickcoupe/softwares).[34] For the rest of the paper, we will refer to VNLM-d as just NLM filter. The search radius for NLM was chosen as $[2 \times 2 \times 2]$. Finally, the MPPCA filter was implemented by the MATLAB code provided by the authors (https://github.com/NYU-DiffusionMRI/dwidenoise).[36] The parameters for all the three methods were optimized based on the visual appearance for all the three cases of noise and are summarized in Table 1. Note that the patch size and window size for NLM and MPPCA, respectively, are indicated in terms of number of voxels, while the Gaussian kernal size is given in millimeters.

Table 1: Parameters used for the three denoising methods.

Gaussian	NLM		MPPCA
Kernel size	Patch size	Standard deviation	Sliding window size
[2.5 2.5 2.5]	[10 × 10 × 10]	0.5	[5 × 5 × 5]

6 Results

In Fig. 1, representative slices before and after denoising the DW images are shown for $b = 2000$ s/mm^2 using Gaussian, NLM, and MPPCA at three different SNRs, 30, 15, and 8. The first and second columns indicate the images before and after adding noise. The third, fourth, and fifth columns give the denoised images at varying SNRs after applying Gaussian, NLM, and MPPCA filters, respectively. From visual appearance, the capability of each denoising method is clearly visible, where MPPCA and NLM evidently outperforms Gaussian filtering (especially at SNR = 8). Between MPPCA and NLM, the efficacy of MPPCA is visible in the case of SNR = 8, for example, in the ventricles.

The goodness criterion of denoising lies in the ability to estimate reliable diffusion metrics from the denoised images. Representative maps of diffusion metrics (both diffusivity and kurtosis based) are given in Figs. 2 and 3. The first row gives the parametric maps estimated from the gold standard dataset, which is considered as ground truth. The second row indicates the diffusion and kurtosis maps obtained after adding noise (SNR = 30) and without any denoising. Finally, the last three rows give the results of diffusion metric maps after using Gaussian, MPPCA, and NLM methods. Similar to Figs. 2 and 3, Figs. 4, 5, 6, and 7 are representative diffusion and kurtosis maps for SNR = 15 and SNR = 8. Figures 2, 3, 4, 5, 6, and 7 reveal the superiority of the non-linear filters such as MPPCA and NLM compared to the traditional Gaussian filter. Moreover, these results also illustrate the higher sensitivity to noise of the kurtosis metrics (MK, $K_{\parallel}$, $K_{\perp}$, KFA) relative to diffusivity metrics (MD, $D_{\parallel}$, $D_{\perp}$, FA), presumably due to the kurtosis metrics stronger dependence on the higher diffusion weightings as suggested by Eq. (2).

As far as diffusivity metrics are concerned, the most noise-sensitive metric is FA, as is apparent in Figs. 2, 4, and 6. On the other hand, for the kurtosis metrics, the most sensitive is KFA, and the least sensitive is MK. The signal-preserving abilities of each method is reflected through the diffusivity and kurtosis metrics, in which, Gaussian filtering clearly suppresses the high-frequency components resulting in blurring of the edges

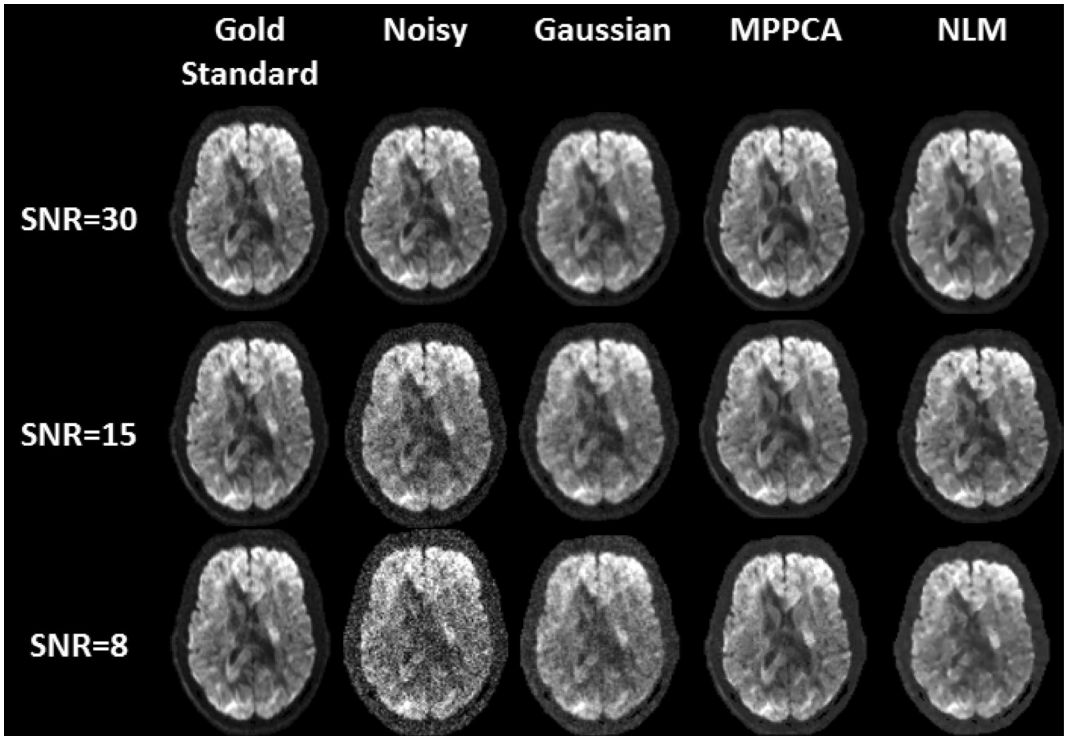

Figure 1: Performance comparison of the three denoising techniques at SNR levels 30, 15 and 8. *First column* represents the gold standard slices (reference dataset) at $b = 2000$ s/mm^2, *second column* represents slices after adding Rician noise at SNR levels, 30, 15, and 8 (SNR was measured with respect to $b = 2000$ s/mm^2). *Third*, *fourth* and *fifth columns* indicate the denoised images obtained by Gaussian, MPPCA, and NLM methods, respectively, at varying SNR levels.

in the anatomy. MPPCA and NLM on the other hand demonstrate the high accuracy of denoising and edge preservation. However, we find that MPPCA is robust even under low SNR conditions (SNR = 8); KFA in Fig. 7 is better estimated by MPPCA than by NLM, which appears noisy compared to gold standard in Fig. 1.

7 Discussion

In this review, we have demonstrated the application of advanced denoising algorithms in DKI with two state-of-the-art denoising techniques, MPPCA and NLM, and a traditional Gaussian filtering technique. Since DKI employs relatively high diffusion weightings, the SNR is typically low so that denoising is often crucial. The use of Gaussian-based filters can be ineffective, both because they do not take into account Rician noise statistics and because they blur edges. This leads to systematic bias in the anisotropy (both FA and KFA) in particular. We found that Gaussian filter was particularly inadequate for the $b = 2000$ set of images with the strongest departure from Gaussian noise. However, Gaussian filtering has the least computational complexity, generating a denoised image in about a minute for our implementation.

In contrast, both MPPCA and NLM account for Rician noise in DW images. However, the two methods differ in the way denoising is performed. MPPCA is a transform-based denoising method, while NLM is a spatial domain-based denoising method. NLM's performance is determined by the "similarity measure" (Eq. 6) that finds a similar patch in the neighborhood and estimates the standard deviation of the noise level, indicated as threshold in Table 1. In order to have an optimal performance, these two parameters need to be tuned by the user. The advantage of NLM lies with retaining the high-frequency components (edges) unlike Gaussian filtering. We found NLM filter to be on par with MPPCA at moderate SNR levels (SNR = 30 and 15), but NLM had a suboptimal performance for higher noise levels

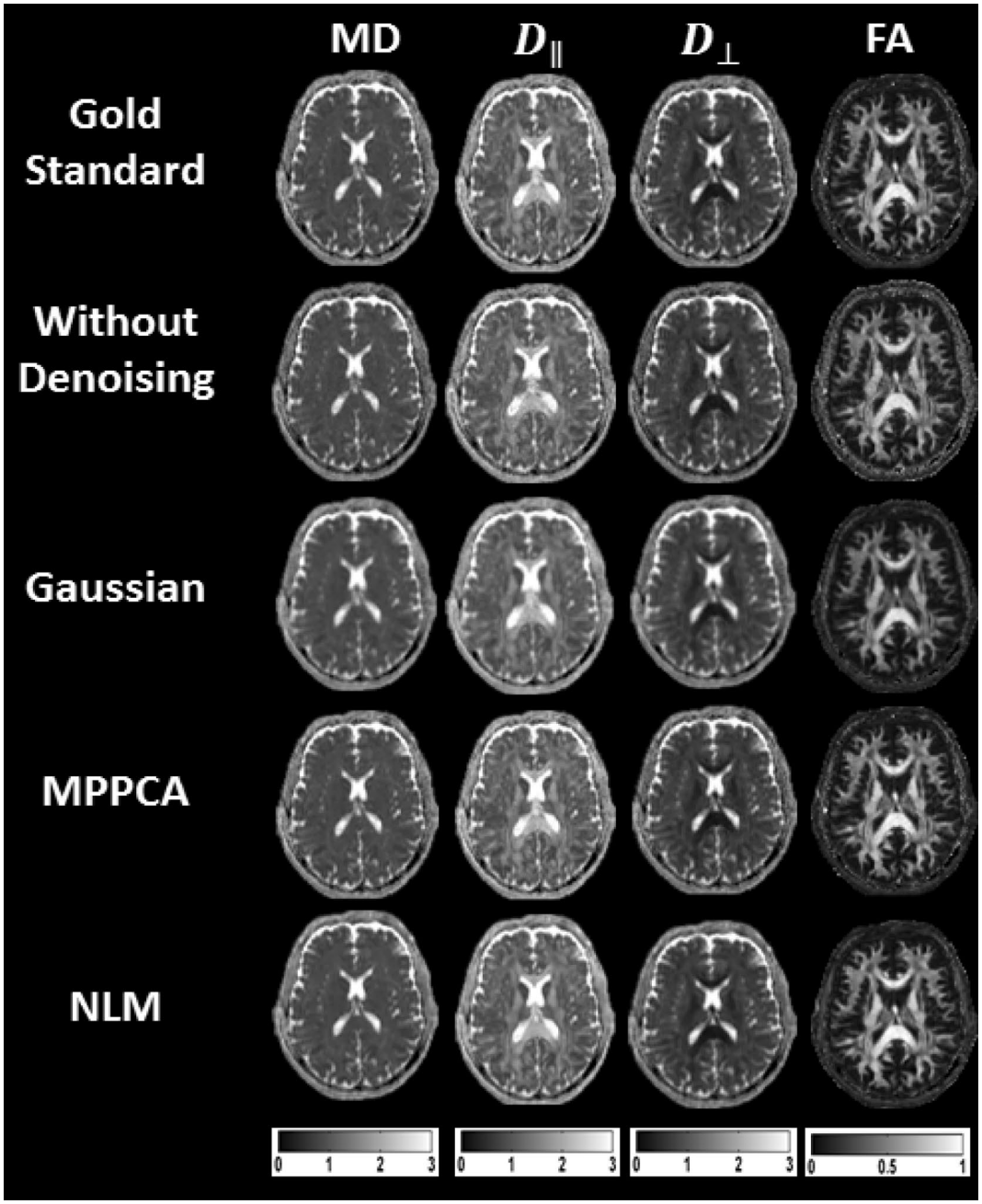

Figure 2: Parametric axial maps for MD, $D_{\parallel}$, $D_{\perp}$, and FA for SNR = 30. The *first* and *second rows* represent the parametric maps with the gold standard dataset and maps from the noisy dataset without any denoising, respectively. The *third*, *fourth*, and *fifth rows* indicate the maps after applying Gaussian, MPPCA, and NLM filters. The *scale bars* for the three diffusivities are in units of $\mu m^2/ms$, and FA is dimensionless.

compared to MPPCA. This may be because of heuristic selection of standard deviation (threshold parameter in Table 1). The computational time for the NLM is decided by the patch size (Table 1); for a patch size of [10 × 10 × 10], the total time was about 6 min.

Finally, MPPCA exploits the multi-directional redundancy in DWI patterns than using local spatial image patches (e.g., NLM). MPPCA approach eliminates the need for searching similar patches that can result in faster processing. A local 3D patch is selected from the multi-directional data

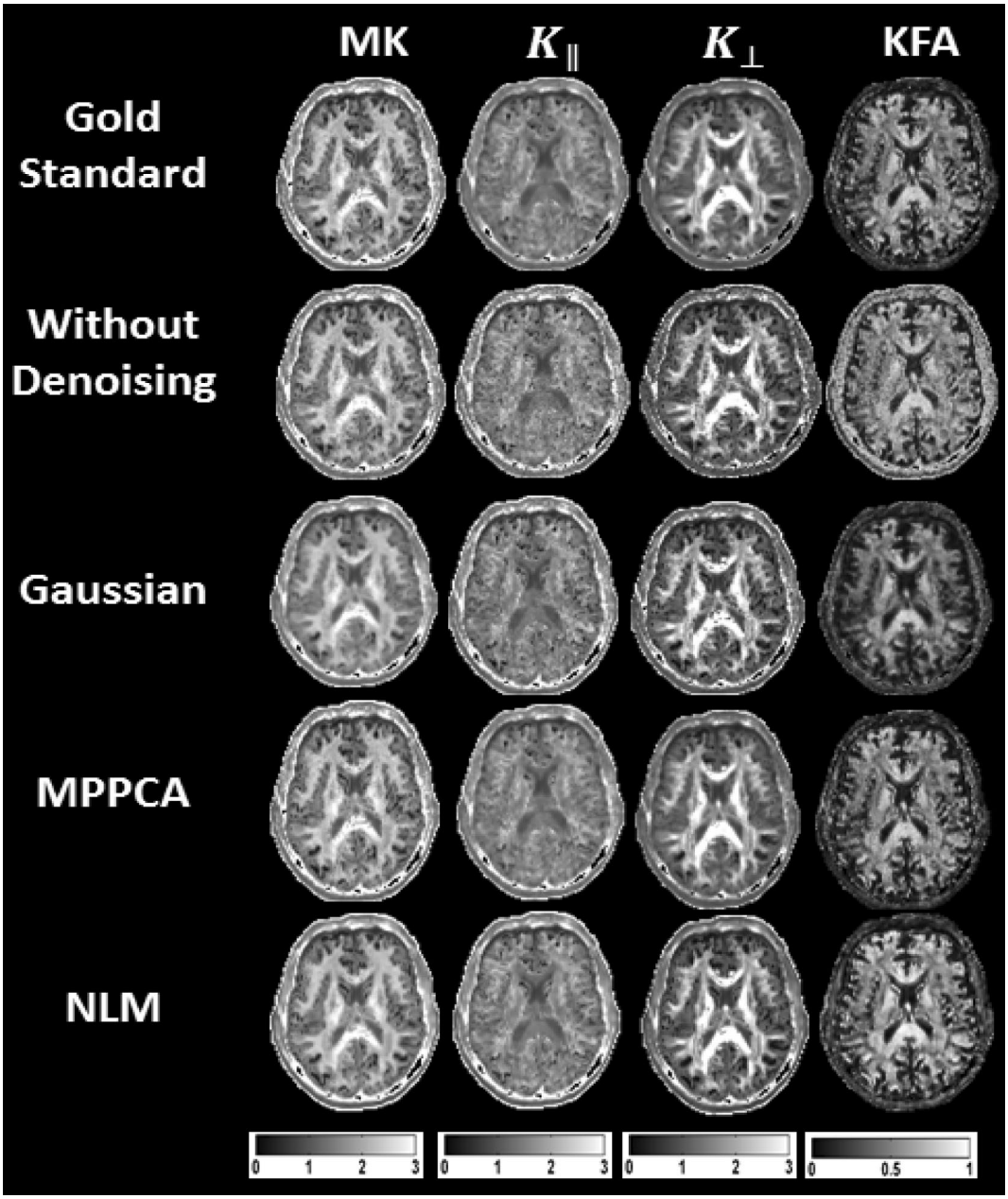

Figure 3: Parametric axial maps for MK, $K_{\parallel}$, $K_{\perp}$, and KFA for SNR = 30. The *first* and *second rows* represent the parametric maps with the gold standard dataset and maps from the noisy dataset without any denoising, respectively. The *third*, *fourth*, and *fifth rows* indicate the maps after applying Gaussian, MPPCA, and NLM filters. The *scale bars* for kurtosis metrics are all dimensionless.

and the default size of [5 × 5 × 5] was used for all the noise cases.[36] Further, the eigenvalues of this 3D patch are computed, from the eigenvalues, signal-carrying principal components and noise-carrying components are separated. The noisy-only eigenvalues are thresholded using the Marchenko–Pastur (MP) law, which is essentially nothing but estimating the standard deviation of the noise distribution. Once the thresholding is performed, an inverse transformation is applied to project the 3D patch back to the spatial domain. An attractive feature of MPPCA is just providing the size of the local patch, and we found that the default choice of [5 × 5 × 5]

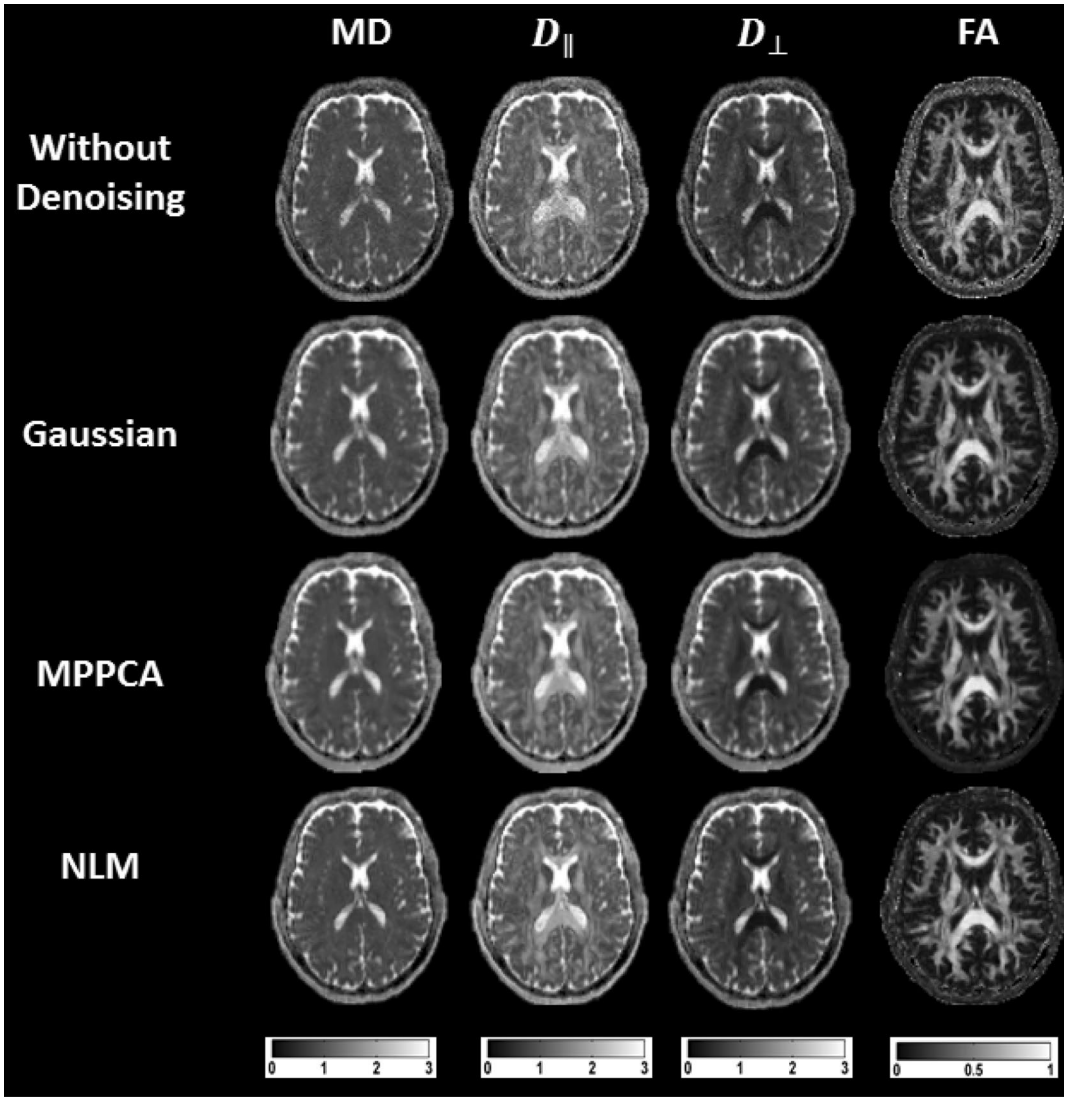

Figure 4: Parametric axial maps for MD,$D_{\parallel}$, $D_{\perp}$, and FA for SNR = 15. The *first* and *second rows* represent the parametric maps with the gold standard dataset and maps from the noisy dataset without any denoising, respectively. The *third*, *fourth*, and *fifth rows* indicate the maps after applying Gaussian, MPPCA, and NLM filters. The *scale bars* for the three diffusivities are in units of $\mu m^2/ms$ and FA is dimensionless.

was robust for all the noise cases considered and outperformed the other methods. The drawback of MPPCA approach is that the noise level is assumed to be constant within the kernel. However, if parallel imaging is adopted the noise might be spatially varying and MPPCA can have suboptimal performance. The computational time for MPPCA, at about 8 min, was a little longer than for NLM.

As a word of caution, it is important to avoid performing denoising as a first step in the post-processing pipeline as coregistration, eddy current correction, etc., can change the noise characteristics. We noticed a significant bias and potentially misleading diffusion metric maps if the denoising was performed subsequent to any other correction schemes. Finally, there are a number of factors that can potentially influence the denoising results of the three filters such as image artifacts, incorrect image registration or Gibbs ringing. Veraart and coworkers[41] have investigated Gibbs ringing in diffusion MRI

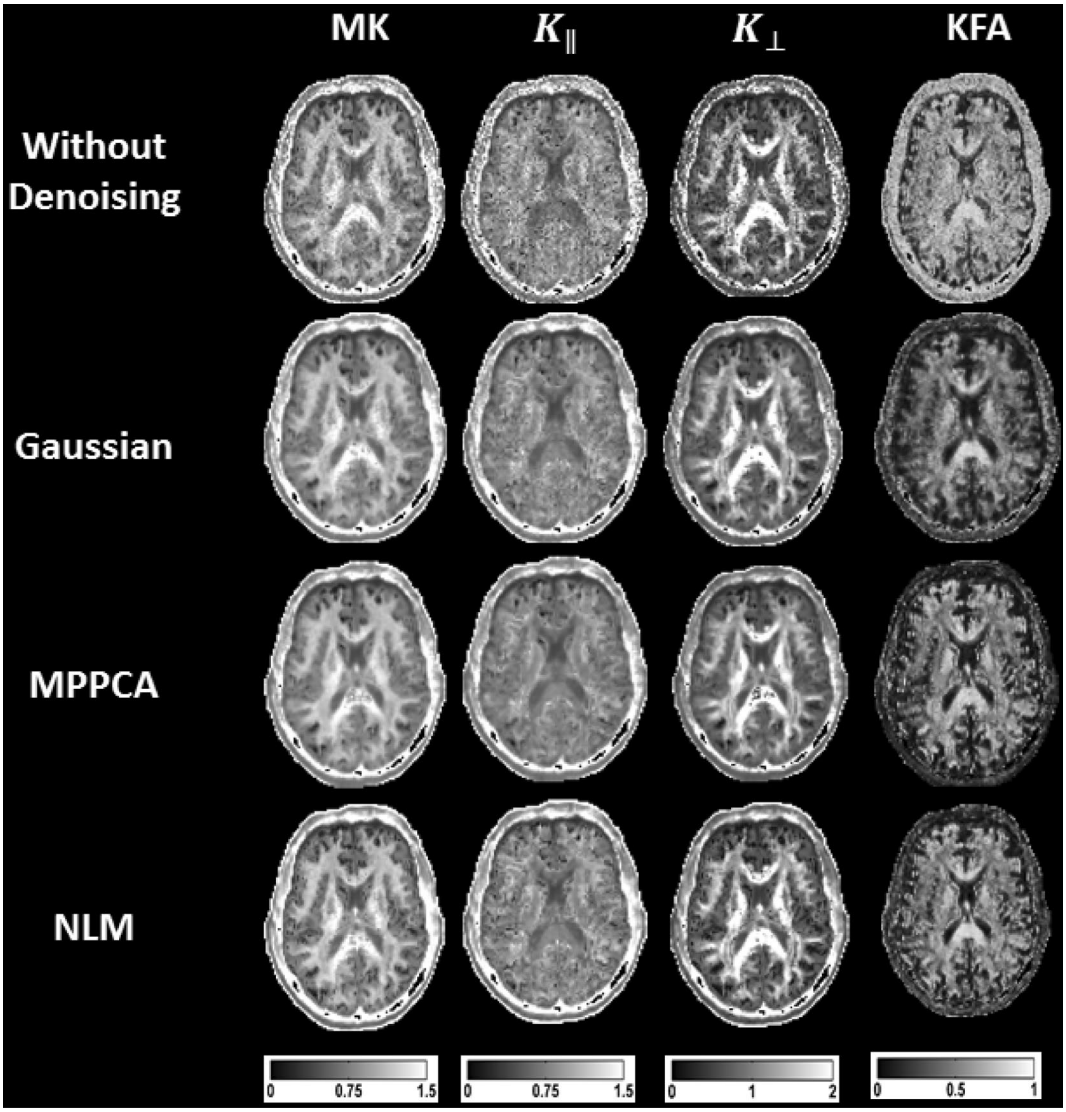

Figure 5: Parametric axial maps for MK, $K_{\parallel}$, $K_{\perp}$, and KFA for SNR = 15. The *first* and *second rows* represent the parametric maps with the gold standard dataset and maps from the noisy dataset without any denoising, respectively. The *third*, *fourth*, and *fifth rows* indicate the maps after applying Gaussian, MPPCA, and NLM filters. The *scale bars* for kurtosis metrics are all dimensionless.

due to the truncation of high-frequency components (because of limited bandwidth) using total variation (TV) based denoising. The use of TV resulted in suppressing ringing near sharp edges such as boundaries of tissues, thereby increasing the anatomical accuracy of the reconstructed images. The limitations of this review also include not considering the local and global effects of the discussed denoising algorithms within the context of DKI tractography.

8 Conclusion and Future Work

The effectiveness for estimating diffusivity and kurtosis metrics of the two state-of-the-art denoising techniques of NLM and MPPCA has

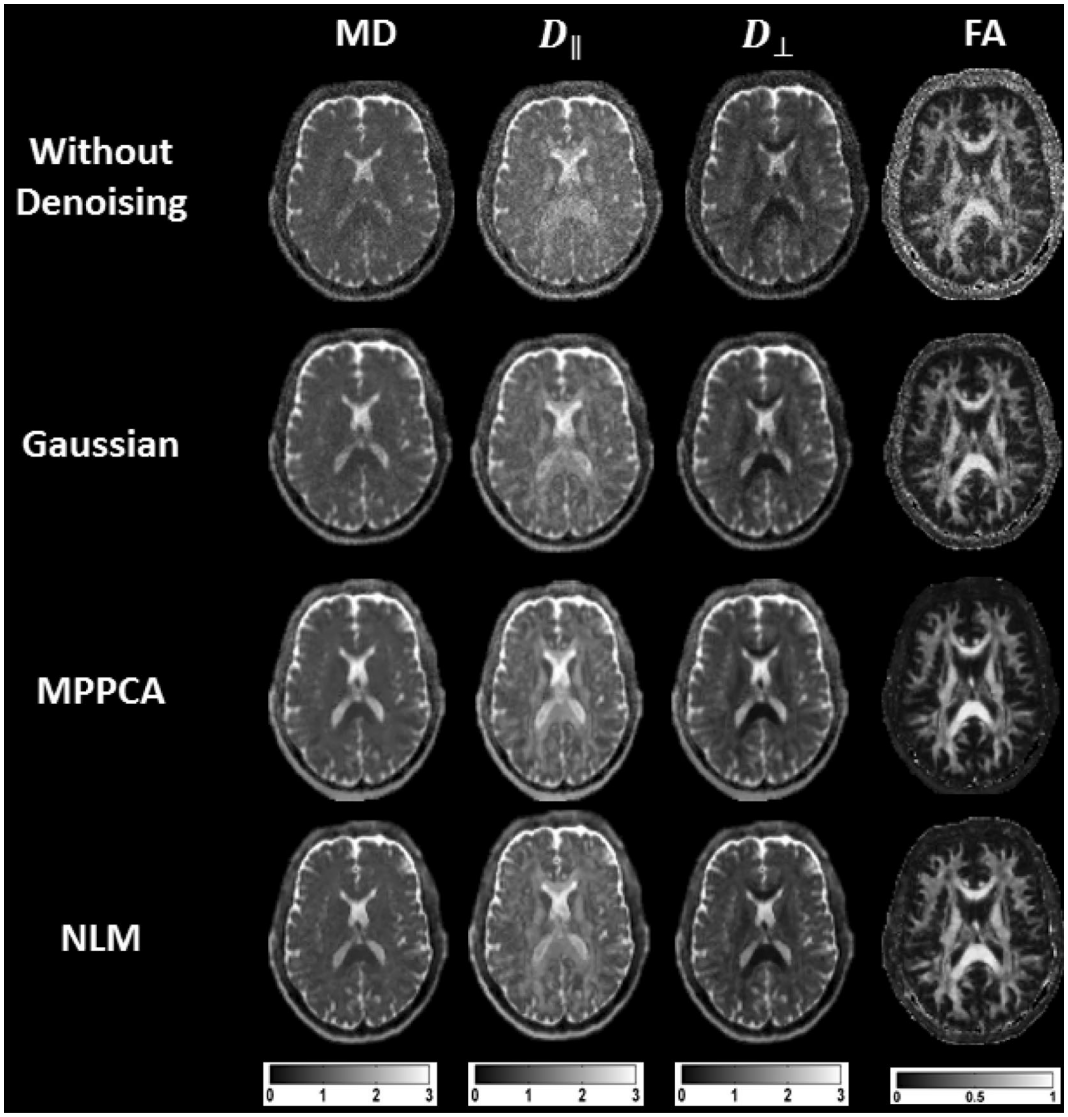

Figure 6: Parametric axial maps for MD, $D_{\parallel}$, $D_{\perp}$, and FA for SNR = 8. The *first* and *second rows* represent the parametric maps with the gold standard dataset and maps from the noisy dataset without any denoising, respectively. The *third*, *fourth*, and *fifth rows* indicate the maps after applying Gaussian, MPPCA, and NLM filters. The *scale bars* for the three diffusivities are in units of $\mu m^2/ms$ and FA is dimensionless.

been investigated here and compared with the performance of conventional Gaussian filtering. For an in vivo human DKI dataset, we have demonstrated that both NLM and MPPCA generally outperform conventional Gaussian filtering in improving the fidelity of diffusion metric maps. Moreover, our results indicate MPPCA to be the overall better choice, particularly for low SNR data.

Recent DTI studies have introduced alternative denoising techniques that exploit rank and edge constraints to enhance the apparent SNR of the DW images.[42] A total Kullback–Leibler principle has also been proposed for simultaneous denoising and estimation of diffusion tensors, in which denoising is performed directly on the tensor instead of on the images.[43] Alternatively, a recent method, collaborative filtering, also termed as Block matching and 4D filtering, has been shown to be effective in retaining fine structures in MR images.[44] Further study in the context of DKI focusing on comparing these other

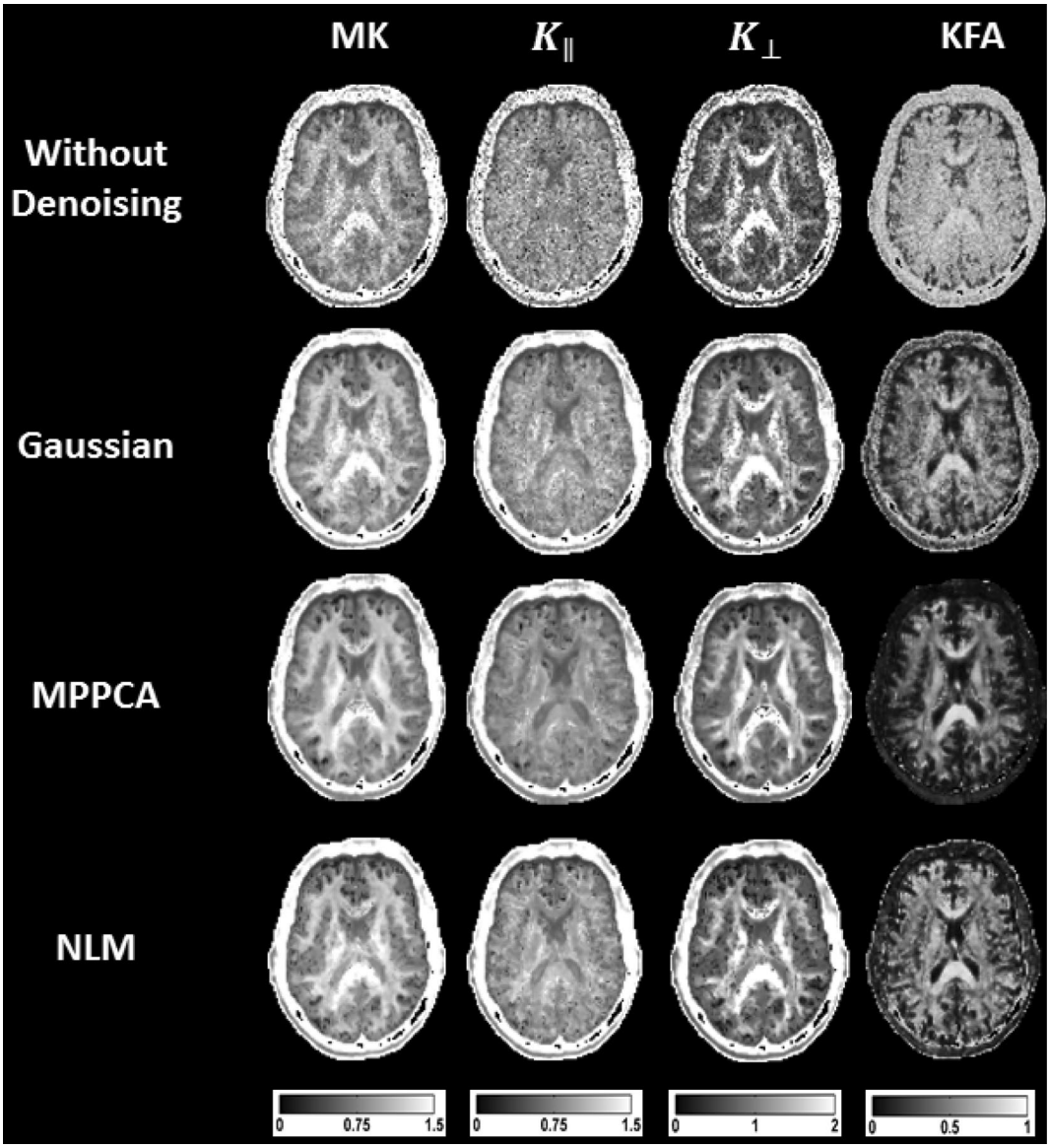

Figure 7: Parametric axial maps for MK, $K_{\parallel}$, $K_{\perp}$, and KFA for SNR = 8. The *first* and *second rows* represent the parametric maps with the gold standard dataset and maps from the noisy dataset without any denoising, respectively. The *third*, *fourth*, and *fifth rows* indicate the maps after applying Gaussian, MPPCA, and NLM filters. The *scale bars* for kurtosis metrics are all dimensionless.

denoising techniques with those investigated here could provide a better understanding of their relative advantage for improving DKI data analysis.

Received: 1 January 2017 Accepted: 19 July 2017
Published online: 18 September 2017

References

1. Basser PJ, Mattiello J, LeBihan D (1994) MR diffusion tensor spectroscopy and imaging. Biophys J 66:259–267
2. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G (1996) Diffusion tensor imaging of the human brain. Radiology 1996(201):637–648
3. Alexander DC, Barker GJ, Arridge SR (2002) Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. Magn Reson Med 48:331–340
4. DeCarlo LT (1997) On the meaning and use of kurtosis. Psychol Methods 2:292–307
5. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K (2005) Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med 53:1432–1440

6. Lu H, Jensen JH, Ramani A, Helpern JA (2006) Three-dimensional characterization of non-Gaussian water diffusion in humans using diffusion kurtosis imaging (DKI). NMR Biomed 19:236–247
7. Jensen JH, Helpern JA (2010) MRI quantification of non-Gaussian water diffusion by kurtosis analysis. NMR Biomed 23:698–710
8. Tietze A, Hansen MB, Østergaard L, Jespersen SN, Sangill R, Lundb TE, Geneser M, Hjelm M, Hansen B (2015) Mean diffusional kurtosis in patients with glioma: Initial results with a fast imaging method in a clinical setting. AJNR 36:1472–1478
9. Jiang R, Jiang J, Zhao L, Zhang J, Zhang S, Yihao Y, Yang S, Jingjing S, Nanxi S, Changliang S, Ju Z, Wenzhen Z (2015) Diffusion kurtosis imaging can efficiently assess the glioma grade and cellular proliferation. Oncotarget 6:42380–42393
10. Grossman EJ, Jensen JH, Babba JS, Chena Q, Tabesh A, Fieremans E, Xiaa D, Inglesea M, Grossman RI (2013) Cognitive impairment in mild traumatic brain injury: a longitudinal diffusional kurtosis and perfusion imaging study. AJNR 34:951–957
11. Stokum JA, Sours C, Zhuo J, Kane R, Shanmuganathan K, Gullapalli RP (2015) A longitudinal evaluation of diffusion kurtosis imaging in patients with mild traumatic brain injury. Brain Inj 29:47–57
12. Fieremans E, Benitez A, Jensen JH, Falangola MF, Tabesh A, Deardorff RL, Spampinato MV, Babb JS, Novikov DS, Ferris SH, Helpern JA (2013) Novel white matter tract integrity metrics sensitive to Alzheimer disease progression. AJNR Am J Neuroradiol 34:2105–2112
13. Benitez A, Fieremans E, Jensen JH, Falangola MF, Tabesh A, Ferris SH, Helpern JA (2014) White matter tract integrity metrics reflect the vulnerability of late-myelinating tracts in Alzheimer's disease. NeuroImage Clin 4:64–71
14. Helpern JA, Adisetiyo V, Falangola MF, Hu C, Di Martino A, Williams K, Castellanos FX, Jensen JH (2011) Preliminary evidence of altered gray and white matter microstructural development in the frontal lobe of adolescents with attention deficit hyperactivity disorder: a diffusional kurtosis imaging study. J Magn Reson Imaging 33:17–23
15. Adisetiyo V, Tabesh A, Martino AD, Falangola MF, Castellanos FX, Jensen JH, Helpern JA (2014) Attention-deficit/hyperactivity disorder without comorbidity is associated with distinct atypical patterns of cerebral microstructural development. Hum Brain Mapp 35:2148–2162
16. Hui ES, Fieremans E, Jensen JH, Tabesh A, Feng W, Bonilha L, Spampinato MV, Adams R, Helpern JA (2012) Stroke assessment with diffusional kurtosis imaging. Stroke 43:2968–2973
17. Zhang S, Yao Y, Shi J, Tang X, Zhao L, Zhu W (2016) The temporal evolution of diffusional kurtosis imaging in an experimental middle cerebral artery occlusion (MCAO) model. Magn Reson Imaging 34:889–895
18. Lee CY, Tabesh A, Benitez A, Helpern JA, Jensen JH, Bonilha L (2013) Microstructural integrity of early-versus late-myelinating white matter tracts in medial temporal lobe epilepsy. Epilepsia 54:1801–1809
19. Bonilha L, Lee CY, Jensen JH, Tabesh A, Spampinato MV, Edwards JC, Breedlove J, Helpern JA (2015) Altered microstructure in temporal lobe epilepsy: a diffusional kurtosis imaging study. AJNR Am J Neuroradiol 36:719–724
20. Nana R, Zhao T, Hu X (2008) Single-shot multiecho parallel echo-planar imaging (EPI) for diffusion tensor imaging (DTI) with improved signal-to-noise ratio (SNR) and reduced distortion. Magn Reson Med 60:1512–1517
21. Veraart J, Poot DHJ, Van Hecke W, Blockx I, Van der Linden A, Verhoye M, Sijbers J (2011) More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging. Magn Reson Med 65:138–145
22. Poot DHJ, Dekker AJD, Achten E, Verhoye M, Sijbers J (2010) Optimal experimental design for diffusion kurtosis imaging. IEEE Trans Med Imaging 29:819–829
23. Gilani N, Malcolm PN, Johnson G (2016) Parameter estimation error dependency on the acquisition protocol in diffusion kurtosis imaging. Appl Magn Reson 47:1229–1238
24. Glenn GR, Tabesh A, Jensen JH (2015) A simple noise correction scheme for diffusional kurtosis imaging. Magn Reson Imaging 33:124–133
25. Haldar JP, Wedeen VJ, Nezamzadeh M, Dai G, Weiner MW, Schuff N, Liang ZP (2013) Improved diffusion imaging through SNR-enhancing joint reconstruction. Magn Reson Med 69:277–289
26. Wirestam R, Bibic A, Latt J, Brockstedt S, Stahlberg F (2006) Denoising of complex MRI data by wavelet-domain filtering: application to high-b-value diffusion-weighted imaging. Magn Reson Med 56:1114–1120
27. Haldar JP, Hernando D, Song SK, Liang ZP (2008) Anatomically constrained reconstruction from noisy data. Magn Reson Med 59:810–818
28. Tabesh A, Jensen JH, Ardekani BA, Helpern JA (2011) Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. Magn Reson Med 65:823–836
29. Rudin LI, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Phys D 60:259–268
30. Block KT, Uecker M, Frahm J (2008) Suppression of MRI truncation artifacts using total variation constrained data extrapolation. Int J Biomed Imaging 2008:184123
31. Buades A, Coll B, Morel JM (2005) A non-local algorithm for image denoising. IEEE Comput Soc Conf Comput Vis Pattern Recognit 2:60–65
32. Manjón JV, Carbonell-Caballero J, Lull JJ, García-Martí G, Martí-Bonmatí L, Robles M (2008) MRI denoising using non-local means. Med Image Anal 12:514–523
33. Zhou M-X, Yan X, Xie H-B, Zheng H, Xu D, Yang G (2015) Evaluation of non-local means based denoising filters for diffusion kurtosis imaging using a new phantom. PLoS One 10:e0116986

34. Manjón JV, Coupé P, Buades A (2015) MRI noise estimation and denoising using non-local PCA. Med Image Anal 22:35–47
35. Manjón JV, Coupé P, Concha L, Buades A, Collins DL, Robles M (2013) Diffusion weighted image denoising using overcomplete local PCA. PLoS One 8:e73021
36. Veraart J, Novikov DS, Christiaens D, Ades-aron B, Sijbers J, Fieremans E (2016) Denoising MRI using random matrix theory. NeuroImage 142:394–406
37. Martin-Fernandez M, Villullas S (2015) The EM method in a probabilistic wavelet-based MRI denoising. Comput Math Methods Med 2015:182659
38. Steven AJ, Zhuo J, Melhem ER (2014) Diffusion kurtosis imaging: an emerging technique for evaluating the microstructural environment of the brain. Am J Roentgenol 202:W26–W33
39. Gudbjartsson H, Patz S (1995) The Rician distribution of noisy MRI data. Magn Reson Med 34:910–914
40. Jones DK, Basser PJ (2004) Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. Magn Reson Med 52:979–993
41. Veraart J, Fieremans E, Jelescu IO, Knoll F, Novikov DS (2016) Gibbs ringing in diffusion MRI. Magn Reson Med 76:301–314
42. Lam F, Babacan SD, Haldar JP, Weiner MW, Schuff N, Liang ZP (2014) Denoising diffusion-weighted magnitude MR images using rank and edge constraints. Magn Reson Med 2014(71):1272–1284
43. Liu M, Vemuri BC, Deriche R (2013) A robust variational approach for simultaneous smoothing and estimation of DTI. NeuroImage 67:33–41
44. Maggioni M, Katkovnik V, Egiazarian K, Foi A (2013) Nonlocal transform-domain filter for volumetric data denoising and reconstruction. IEEE Trans Image Process 22:119–133

Calvin B. Shaw received his M.Sc. (Engg) and Ph.D. degrees in engineering from the Indian Institute of Science, Bangalore, in 2012 and 2014, respectively. He is a Postdoctoral Research Scholar at the Department of Radiology, University of California Davis, USA. His research interests include biomedical optics, photoacoustic imaging, inverse problems, diffusion MRI, real-time MRI, diffusional kurtosis imaging, fiber tractography and neuroimage analysis.

Jens Jensen received a BS degree from Caltech and a Ph.D. from Princeton University, both in physics. For the past 20 years, his research has focused on developing MRI technology, and he has co-authored over 100 peer-reviewed articles. He is currently a Professor of Radiology at the Medical University of South Carolina.