



Evaluation of the accuracy and precision of the diffusion parameter ESTimation with Gibbs and Noise removal pipeline

Benjamin Ades-Aron^{a,*}, Jelle Veraart^{a,*}, Peter Kochunov^b, Stephen McGuire^c, Paul Sherman^c, Elias Kellner^d, Dmitry S. Novikov^a, Els Fieremans^a

^a Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine, NY, USA

^b Department of Psychiatry, University of Maryland School of Medicine, MD, USA

^c U.S. Air Force School of Aerospace Medicine, Aeromedical Research Department, 2510 5th Street, Building 840, Wright-Patterson AFB, OH, 45433-7913, USA

^d Department of Diagnostic Radiology, University Medical Center Freiburg, Freiburg, Germany

ARTICLE INFO

Keywords:

Diffusion MRI
Denoising
Gibbs ringing
Image processing
Artifact correction

ABSTRACT

This work evaluates the accuracy and precision of the Diffusion parameter ESTimation with Gibbs and Noise Removal (DESIGNER) pipeline, developed to identify and minimize common sources of methodological variability including: thermal noise, Gibbs ringing artifacts, Rician bias, EPI and eddy current induced spatial distortions, and motion-related artifacts. Following this processing pipeline, iterative parameter estimation techniques were used to derive diffusion parameters of interest based on the diffusion tensor and kurtosis tensor. We evaluated accuracy using a software phantom based on 36 diffusion datasets from the Human Connectome project and tested the precision by analyzing data from 30 healthy volunteers scanned three times within one week. Preprocessing with both DESIGNER or a standard pipeline based on smoothing (instead of noise removal) improved parameter precision by up to a factor of 2 compared to preprocessing with motion correction alone. When evaluating accuracy, we report average decreases in bias (deviation from simulated parameters) over all included regions for fractional anisotropy, mean diffusivity, mean kurtosis, and axonal water fraction of 9.7%, 8.7%, 4.2%, and 7.6% using DESIGNER compared to the standard pipeline, demonstrating that preprocessing with DESIGNER improves accuracy compared to other processing methods.

1. Introduction

Diffusion MRI (dMRI) is a technique to non-invasively visualize the structure of microstructural tissue properties (Beaulieu, 2002; Jones, 2010; Le Bihan et al., 1986). A common drawback to dMRI, despite its use in clinical applications in brain (Moseley et al., 1990; Warach et al., 1995) and body (Haider et al., 2007; Kiselev, 2017; Koh and Collins, 2007), is the sparse signal relative to noise or artifacts. There is a need for robust dMRI processing pipelines to provide clinicians and researchers with confidence in diffusion outcome measurements. Such a pipeline should adequately identify and correct for thermal noise and for the various imaging artifacts typically present in dMRI, as described here in more detail.

Human dMRI data collected on clinical scanners suffers from low signal-to-noise ratio (SNR) typically due to a combination of strong diffusion gradients and long echo times (TE = 60–120 ms). The low SNR limits the precision and, albeit less intuitive, the accuracy of diffusion

parameter estimators (Aja-Fernandez et al., 2008; Veraart et al., 2011). Improved precision with the same hardware can only be achieved by lengthening the scan time, which is often clinically infeasible since precision increases slowly, as a square root of scan time. Otherwise, increasing SNR can be achieved by lowering the spatial resolution, which results in loss of anatomical detail; by shortening the echo time, which is often not possible due to hardware limitations; or by smoothing, which inherently decreases the effective resolution due to partial volume effects. Furthermore, since the noise level couples with the expected signal due to the generally non-Gaussian nature of the MRI noise in magnitude MR images (Gudbjartsson and Patz, 1995), one needs to employ an accurate noise model either in the parameter estimation framework (Sijbers and den Dekker, 2004) (e.g., maximum likelihood estimators), or correct the signal (e.g., using the method of moments), prior to any parameter fitting (Aja-Fernandez et al., 2008; Koay and Basser, 2006).

In addition to limited SNR, human dMRI data may also suffer from a plethora of particular artifacts that stem from various sources. Long scan

* Corresponding authors. 660 1st Ave, New York, NY, 10016, USA.

E-mail addresses: Benjamin.Ades-Aron@nyumc.org (B. Ades-Aron), Jelle.Veraart@nyumc.org (J. Veraart).

<https://doi.org/10.1016/j.neuroimage.2018.07.066>

Received 31 December 2017; Received in revised form 11 July 2018; Accepted 30 July 2018

Available online 2 August 2018

1053-8119/© 2018 Elsevier Inc. All rights reserved.

times intensify motion-based and physiological artifacts (Le Bihan et al., 2006), strong gradients can cause eddy current artifacts (Andersson and Skare, 2002; Andersson and Sotiropoulos, 2016) and these gradients coupled with magnetic field inhomogeneity and EPI read-out lead to geometric distortions (Andersson et al., 2003). In addition, radio frequency pulse inhomogeneity can lead to spatial signal bias (Smith et al., 2004) and insufficient spatial sampling can cause Gibbs ringing (Barker and Wheeler-Kingshott, 2001; Tournier et al., 2011; Veraart et al., 2016a). Gibbs ringing artifacts are of particular relevance in dMRI. They are produced when high-contrast boundaries are sampled with an insufficient spatial sampling frequency. This artifact has been identified as the source of inaccuracies in the parameterization of the dMRI data, including negative radial diffusivity or kurtosis values (Perrone et al., 2015; Veraart et al., 2016a), often referred to as the so-called “black voxels” in kurtosis maps (Kellner et al., 2016; Perrone et al., 2015; Veraart et al., 2016a).

To address the need of a processing pipeline that amends the low SNR and artifacts mentioned above, we developed the DESIGNER pipeline to restore the quality and integrity of dMRI data while preserving spatial resolution (Ades-Aron et al., 2016). The main DESIGNER steps are: (i) MP-PCA denoising (Veraart et al., 2016c) based on identifying Marchenko-Pastur distribution of noise-only principal components. This approach improves SNR (by factor of 2–4) without longer scan time or reduced spatial resolution; (ii) Gibbs correction based on shifting the zero-crossings of the sinc function in k -space (Kellner et al., 2016); (iii) Rician bias correction based on the estimated noise level (Veraart et al., 2016b); (iv) Corrections for EPI distortions, motion, and field inhomogeneity (Smith et al., 2004); (v) Parameter estimation. In this work, this pipeline combines preprocessing and analyses steps in a specific order, and then, as an example, estimates diffusion signal's cumulants (diffusion tensor imaging (DTI; Basser et al., 1994) and kurtosis tensor imaging (DKI; Jensen et al., 2005)) (Kisilev, 2010), as well as White Matter Tract Integrity (WMTI) metrics (Fieremans et al., 2011) for quantification of the axonal water fraction (AWF). However, this approach is compatible with parameter estimation for any signal representation and/or biophysical model of choice as the last step.

DESIGNER is currently available online in a fully integrated open-source framework implemented in both MatLab (The MathWorks, Inc.) and Python and can be found at: www.github.com/NYU-DiffusionMRI/Diffusion-Kurtosis-Imaging.

The outline of this work is as follows. We first characterize the nature of thermal noise and a number of specific artifacts relevant to dMRI. Next, we show how each component of DESIGNER is tailored to either remove noise, or model and correct artifacts. Finally, we demonstrate the robustness of the pipeline in terms of accuracy relative to an unbiased high-SNR reference dataset and parameter precision in human data.

The accuracy and precision of the DESIGNER pipeline were evaluated

using simulated and test-retest dMRI datasets. *Accuracy* was evaluated by synthesizing a reference dMRI signal based on averaging the dMRI signals from 36 subjects randomly taken from the Human Connectome Project data set. *Precision* was evaluated in terms of the scan-rescan reproducibility of diffusion metrics in normal human volunteers. We report that the DESIGNER pipeline which performs MP-PCA denoising and Gibbs correction prior to standard postprocessing steps demonstrated higher accuracy and stability as compared to typical dMRI analysis approaches that do not perform those steps or use smoothing instead.

2. Methods

2.1. DESIGNER pipeline

The flowchart shown in Fig. 1 describes the order of operations of DESIGNER and compares it to typical processing pipelines used in dMRI. The most distinct feature of the designer pipeline is the specificity in image quality improvement, including thermal noise reduction and Gibbs ringing correction. We will show that such targeted approaches outperform more brute force techniques, e.g. spatial smoothing, which are adopted by several dMRI analysis pipelines (Cui et al., 2013; Tabesh et al., 2011).

Our basic metric of the efficacy of denoising and Gibbs ringing removal is the power spectrum, $E(\mathbf{k}) = |\epsilon(\mathbf{k})|^2/A$, of normalized residuals $\epsilon(\mathbf{x}) = (S_{\text{post}}(\mathbf{x}) - S_{\text{pre}}(\mathbf{x}))/\sigma$ within a slice, where S is the original signal, S_{post} is the post correction signal, S_{pre} is the pre correction signal, and σ is standard deviation of noise, $\epsilon(\mathbf{k}) = \int d^2\mathbf{x} \epsilon(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}$, and A is the slice area (the number of voxels). We study this power spectrum as a function of the radial $k = |\mathbf{k}|$ within a slice, after angular-averaging over the shell-like bins. For the perfectly spatially-uncorrelated residuals, the power spectrum normalized in this way is a unit line $E(k) = 1$ for all k . We apply this metric to each step and show results for MPPCA and Gibbs correction steps.

We now describe each pipeline step in detail.

2.1.1. Denoising using MP-PCA

Data is first processed using our recently introduced MP-PCA technique for noise estimation and reduction (Veraart et al., 2016b, 2016c). This technique is enabled by exploiting data redundancy in the PCA domain using properties of the eigenspectrum of random covariance matrices. It is important that the denoising step is the *first stage* of the pipeline as it relies on noise being uncorrelated both spatially and among successive acquisitions (in the dMRI case, in q -space). Performing this step after processing steps that use interpolation to reconstruct images would result in correlated noise and failure of the basic assumptions underlying the random matrix theory-based approach to PCA denoising. MP-PCA removes only thermal noise without compromising anatomical

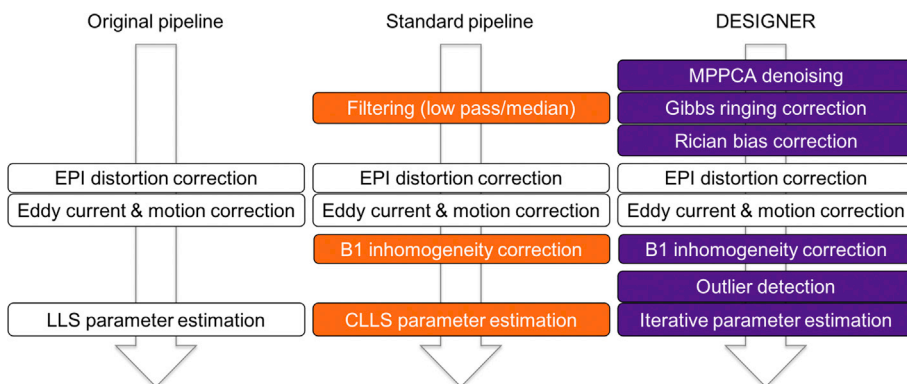


Fig. 1. Overview of processing pipelines for dMRI: Original DTI pipeline (left), standard DKI (middle) and DESIGNER (right). Differences between pipelines are highlighted, the main being that smoothing is present in the standard DKI pipeline which has been replaced with MP-PCA denoising (by exploiting the redundancy of signal in the dMRI dataset), followed by Gibbs artifact correction, and Rician bias correction in DESIGNER. Additionally the original pipeline uses linear least squares (LLS) fitting, standard pipeline uses constrained linear least squares (CLLS) fitting, and DESIGNER uses iteratively reweighted linear least squares (IRWLLS) fitting. The steps of DESIGNER (in this order) improve performance of the downstream artifact correction methods such as EPI/eddy distortion and motion corrections, and in our DKI example, can be followed by an *unconstrained* iterative WLLS DKI fit due to improved quality of the fit input.

features, requires about 1 min processing time for the whole brain, and provides an accurate estimate of the noise level that is needed for Rician bias correction in a later stage. MP-PCA software has been implemented in MatLab (<https://www.nitrc.org/projects/mpcca/>) and as a C++ library function, as part of the open-source MRtrix package (Tournier et al., 2012).

Here, we evaluate the performance of MP-PCA in terms of accuracy and precision, whereby we also investigate the effect of spatial redundancy by selecting different kernel windows and varying SNR-levels, as described below.

2.1.2. Gibbs ringing artifact correction

The MR image is reconstructed from finite sampling of the signal following the inverse Fourier transform. High-contrast boundaries such as the border between cerebrospinal fluid (CSF) and gray matter (GM) or white matter (WM) may produce an image artifact in the form of Gibbs rings due to insufficient sampling of the high frequency information. Diffusion parameters are significantly affected by Gibbs ringing, both qualitatively (Barker G and Wheeler-Kingshott, 2001; Perrone et al., 2015) and quantitatively, as our recent analytical calculations show (Veraart et al., 2016a). Consequently, correction techniques like Gegenbauer reconstruction (Amatur and Haacke, 1991) or extrapolation methods (Amatur et al., 1991; Archibald and Gelb, 2002) aim at recovering missing frequencies based on the analytical information available. We apply an alternative approach where the truncation in k -space can be modeled as a convolution with a sinc-function in image space. Hence, the severity of the artifacts depends on how this sinc-function is sampled. The method implemented here re-interpolates the image based on local subvoxel shifts to sample the ringing pattern at the zero-crossings of the oscillating sinc-function (Kellner et al., 2016). This sinc-function is found by computing the total variation in the neighborhood of each voxel in each orthogonal direction in an image defined on a square lattice and computing the shift in voxel position which minimizes the oscillation. The final voxel value is found by interpolating the shifted image based on the original image grid. This way, the artifact can be effectively and robustly removed with a minimal amount of filtering.

2.1.3. Rician bias correction

The noise in dMRI images is governed by a Rician or non-central χ distribution that creates a positive bias in low SNR (especially $\text{SNR} < 2$) images (Gudbjartsson and Patz, 1995). This bias may affect the fitting of a diffusion model at high b data (e.g., $b \gtrsim 2000 \text{ s/mm}^2$) (Gudbjartsson and Patz, 1995). We use MP-PCA denoising to estimate an unbiased standard deviation σ of the noise at each voxel (Veraart et al., 2016b) from low b (typically, up to $b = 1000 \text{ s/mm}^2$) data, and use the analytical paradigm described by (Koay and Basser, 2006) to estimate the expected value of true signal voxel intensity η :

$$\eta^2 = \langle M \rangle^2 + (\xi(\theta) - 2)\sigma^2$$

where $\langle M \rangle$ is a proxy of the expectation value of the measured magnitude signal intensity, and $\xi(\theta)$ is a correction factor with $\theta \equiv \frac{\eta}{\sigma} \equiv \text{SNR}$ (Koay and Basser, 2006). At $\text{SNR} > 2$, $\xi(\theta)$ goes to 1 and we make use of the approximation: $\eta = \sqrt{\langle M \rangle^2 - \sigma^2}$ to calculate the correct signal intensity. It has been shown that denoising using MP-PCA returns the proxy $\langle M \rangle$ for signal magnitude (Veraart, 2016b)

2.1.4. EPI + eddy current and motion correction

Strong and rapidly switching diffusion gradient fields cause scaling and shear effects in the phase-encoding direction, as well as image translations that vary with slice position. To account for geometric distortions, eddy currents, and motion artifacts, FSL's *TOPUP* (Andersson et al., 2003) and single target *eddy* (Andersson and Sotiropoulos, 2016) are used to model slice dependent geometric shearing, and correct for

motion differences by registering to a high SNR $b = 0$ image respectively. DESIGNER improves the accuracy of EPI and eddy current correction by preceding this correction with MP-PCA denoising. Here, we evaluate the extent of accuracy improvement in motion correction using two motion correction tools – *eddy* and TORTOISE (C. Pierpaoli et al., 2010). The DWI phantom fully described in the next section is used here to evaluate the improvement in accuracy in motion correction that results from MPPCA denoising. We generate 6 random sinusoids to add continuous rotation and translation to the phantom, then add noise at $\text{SNR} = 15$ and denoise with MPPCA. Motion correction is performed on the noiseless phantom, noisy phantom, and denoised phantom. The mutual information is computed between each volume of the original phantom and each corresponding volume of the three test datasets.

2.1.5. B_1 bias field correction

The signal intensity of MR image data can be spatially biased due to radiofrequency (RF) field inhomogeneities caused by high-density receiver coils. This is often referred to as a B_1 bias field. This variability in signal through tissue of the same type can affect the tensor fitting to diffusion data. In particular, in the case of multiple b -shell acquisitions spread over different scan series, the inhomogeneity spectrum can vary and propagate major artifacts to model parameters. For each DWI series, we estimate the bias field based on the mean $b = 0$ image and apply the field correction to all related volumes. This correction is performed using the FAST (Smith et al., 2004; Zhang et al., 2001) tool in FSL.

2.1.6. Outlier detection

The two types of outlier correction implemented here include the eddy signal dropout (“repol”) tool (Andersson et al., 2016), as well iterative parameter estimation during tensor estimation. Least-squares regression models assume that signal variability is due to thermal noise rather than artifacts such as ghosting or chemical shifts. While signal variability produced by thermal noise is approximately Gaussian distributed with $\text{SNR} > 2$, variability produced by artifacts, in general, cannot be modeled. Due to the likely presence of outliers in dMRI data, it is important to use an estimator that is not sensitive to artifacts. The method of outlier detection and rejection (Collier et al., 2015) employs iteratively reweighted linear least-squares (IRWLLS) regression to identify potential outliers, whereby the residuals of the fit determine the weighting applied to the next iteration, ensuring that outlying directions with larger residuals will receive correspondingly lower weights during estimation of the tensor (Collier et al., 2015).

The steps described here outline the major components of DESIGNER. After the completion of these steps, a model can be fit to the signal using a diffusion parameter estimator. Placing positivity constraints on diffusion or kurtosis parameters is a reasonable assumption for most biological applications of dMRI, however the disadvantage is that such constraints inevitably bias the estimated values. In the classical approach, these constraints are required. With DESIGNER, we typically obtain fairly stable and unbiased results even without constraints. In addition, constrained nonlinear fitting is often time consuming, especially as compared with linear fitting for DTI and DKI, and is prone to fitting outlier data points to boundary condition values.

The DESIGNER pipeline provides the advantage of targeting noise and artifacts as separate entities, increasing both accuracy and precision of estimated parameters. We will now describe our methods for evaluating accuracy and precision.

2.2. Accuracy

2.2.1. Diffusion brain software phantom used to assess accuracy

We measured the accuracy of DESIGNER by creating a diffusion brain phantom using data from the Human Connectome Project (Van Essen et al., 2013). 36 dMRI images with voxel size $1.25 \times 1.25 \times 1.25 \text{ mm}^3$, acquired in a matrix of $145 \times 174 \times 145$ were used to create an artificial dMRI brain phantom. These datasets were corrected for EPI and motion

artifacts on an individual basis prior to phantom construction.

We first computed the nonlinear warp from a $b = 0$ image of each dataset to the $b = 0$ image of a single HCP subject using Elastix (Klein et al., 2010). 6th order spherical harmonic representations were then computed for $b = 1000, 2000$, and 3000 s/mm^2 shells of each dataset and then transformed to a common space by applying the warp computed based on $b = 0$ images. Spherical harmonics were rotated based on the Jacobian of each warp field in order to keep them correctly aligned after nonlinear transformation. Finally, transformed spherical harmonics for each shell were averaged and projected in 30 unique directions generated by finding the lowest Coulomb energy for 30 repelling charges on a sphere (Jones et al., 1999). This procedure allowed us to generate a new set of dMRI images with the same resolution but much higher SNR by averaging over both subjects and over diffusion directions. High SNR benchmark parametric maps of FA, MD, MK, and AWF were computed for the HCP phantom based on an unconstrained linear least-squares fit.

2.2.2. Evaluating accuracy of DESIGNER on the software phantom

The above software phantom was used as a reference by which to benchmark the efficacy of different steps of DESIGNER. By adding noise with known variance to a ground truth anatomical image, we can directly measure the value and SNR improvement gained after the implementation of each correction routine. Ideally this phantom would have no noise or artifacts that could potentially bias results; however, with 36 HCP datasets, we were able to produce a phantom with at most 1/6 the noise of the individual HCP images and $\text{SNR} > 100$.

Gibbs artifacts were added to the phantom through convolution of the diffusion signal with a sinc-function that has the same frequency as that caused by image truncation in k -space and then down-sampled by a factor of two. 50 realizations of noise at an SNR of 15 were also added to all shells ($b = 0, 1000, 2000, 3000 \text{ s/mm}^2$) of the phantom dMRI images to simulate thermal noise in 50 independent acquisitions. For all 50 noise realizations of the phantom, images were processed (see Fig. 1) by using A) the *original* pipeline with minimal preprocessing (EPI, eddy current, and motion correction only), B) a *standard* pipeline with a smoothing kernel of $1.25 \times$ voxel size, or C) DESIGNER, and then compared to the high SNR reference values. In addition, we also varied the SNR level from 5 to 50 by injecting Rician noise at varying σ into each volume and tested the effect of the window kernel for performing MP-PCA in the posterior limb of the internal capsule (PLIC) by either a) choosing a window of randomly sampled voxels through the whole brain, or b) choosing a window of voxels that belong exclusively to the single anatomic region of interest.

Furthermore, a region of interest (ROI) analysis was performed to test the accuracy of DESIGNER compared to standard pipelines in different white matter regions. For this analysis, the JHU (Hua et al., 2008) white matter atlas was used to delineate different regions. ROI values over the HCP phantom will be compared to the voxel-wise average of 50 noise realizations after processing with both DESIGNER and the standard pipeline.

2.3. Precision

2.3.1. Dataset used to assess to precision

We tested the precision of DESIGNER pipeline by collecting data in $N = 30$ subjects (22/8 males/females, average age 25.8 ± 6.4 range 18–41 years) who were imaged with MRI three times within a 5-day period. All subjects provided informed consent prior to participation. The study was reviewed and approved by the 59th Medical Wing, USAF, Institutional Review Board. Commencing seven days prior to the first MRI and continuing throughout the study duration all subjects were alcohol free, drug/medication free, and tobacco free. To minimize diurnal physiological fluctuations, the daily time of repeat scans within the same subject was consistent for all three scans.

The human dMRI was performed on a 3T Trio (Siemens, Erlangen Germany) and consisted of 12 shells of b -values ($b = 250, 500, 600, 700,$

$800, 900, 1000, 1250, 1500, 1750, 2000, 2500 \text{ s/mm}^2$; bipolar gradient duration $\delta = 47 \text{ ms}$, gradient separation $\Delta = 54 \text{ ms}$), and GRAPPA = 2 acceleration. Thirty-two isotropically distributed diffusion-weighted directions were collected per shell, in addition to a total of sixteen $b = 0$ images. The imaging data were collected using a single-shot, echo-planar, single refocusing spin-echo, diffusion-weighted sequence with a spatial resolution of $1.7 \times 1.7 \times 4.6 \text{ mm}$ and seven slices prescribed in sagittal orientation to sample the midsagittal band of the corpus callosum. The sequence control parameters were $\text{TE/TR} = 120/1500 \text{ ms}$ with the $\text{FOV} = 200 \text{ mm}$.

2.3.2. Evaluating precision of DESIGNER

Test-retest reproducibility data was used to estimate the precision of the parameter measurements. This dataset pays specific attention to the corpus callosum, which is a white matter tract that has been well studied in diffusion literature. The chosen orientation allowed us to get a clear sample of the genu and splenium of the corpus callosum without needing to worry about confounders like partial volume effects. Due to the unusual acquisition of this dataset, the original reproducibility data was down sampled in q -space to $b = 500, 1500$, and 2500 s/mm^2 with 30 directions per shell in order to better represent a typical clinical DKI scanning protocol, although with thick slices to help compensate for long TE and short TR. This data has an SNR of 12–15 at $b = 0$ in brain tissue, comparable to clinical diffusion protocols. Images were processed (see Fig. 1) using both (i) the standard pipeline with a smoothing kernel of $1.25 \times$ the voxel size and constrained tensor fitting, and (ii) DESIGNER.

Parametric maps including FA, MD, MK, and AWF (Fieremans et al., 2011) were computed using an unconstrained fit prior to alignment to the JHU-ICBM-FA-1mm template (Hua et al., 2008) and are used to illustrate the precision of the two pipelines. Technically, AWF derivation from kurtosis tensor was designed only for highly-aligned WM tracts. As applied to the whole brain, it loses its meaning as water fraction inside axons; here, we use this metric merely for the demonstration purpose of model parameter estimation. Not surprisingly, its performance improves, roughly, similarly to that of kurtosis tensor components. The metric used to assess precision is the coefficient of variability: σ_n/μ_n where σ_n is the standard deviation in the n th parameter over the three time points, and μ_n is its mean over the three time points. CoV can be favorably biased in the case where there are very few samples (Sokal and Rohlf, 1981), in order to compensate for the small number of test-retest datasets, we compute the unbiased CoV = $\left(1 - \frac{1}{4t}\right)\sigma_t/\mu_t$ where t is the sample size.

We registered the three datasets in each subject to the first time-point based on the within-subject FA images with a six-degree of freedom rigid body registration using FSL-FLIRT (Smith et al., 2004) software. A nonlinear warp was computed from the FA of each subject to the JHU-ICBM-FA-1mm template (Hua et al., 2008) and applied to the parametric maps. σ_n/μ_n were compared at each voxel of the brain in order to generate variability maps and show which areas demonstrate comparatively higher precision. For this analysis, the JHU (Hua et al., 2008) white matter atlas was used to delineate different regions in the corpus callosum and FSL-FAST was used to select total CSF, GM, and WM regions in template space. Mean and standard deviation of σ_n/μ_n were measured voxel-wise, over ROIs in the corpus callosum, and over total tissue type ROIs in order to summarize the precision of each pipeline.

3. Results

3.1. Effects of individual processing steps

3.1.1. Denoising using MP-PCA

Fig. 3a shows an example axial slice taken from the HCP phantom (with added noise as described in Sec. 2.3.2) through the human brain at $b = 1000 \text{ s/mm}^2$, where the image has been denoised using MP-PCA and using spatial smoothing using a Gaussian kernel with full width at half maximum (FWHM) = $1.25 \times 1 \text{ voxel}$ for each axis). We found that

smoothing partially removes anatomic signal while MP-PCA removes only signal related to thermal noise, which can be seen in the residual maps below. Power spectra of residuals (PSR) normalized to noise variance (Fig. 3) for MP-PCA show a constant energy equal to unity for all k meaning that mostly noise is being removed (independently at each k space frequency), however smoothing removes high frequency components of the image giving very large energy spectrum values. It is qualitatively evident that MP-PCA outcomes are less blurred as compared to smoothed images. Residual maps show that the smoothed dataset has lower tissue specificity; the presence of anatomy in residuals of Fig. 2a implies blurring of signal at these locations. It is more difficult to distinguish separate anatomy in the smoothed dataset compared to the

denoised dataset because of partial volume effects introduced by the smoothing operation.

We also compared the noise reduction due to MP-PCA to noise reduction due to fitting a model to the same data. A fit of an exact physical model by maximizing the likelihood of the *a priori* known data distribution, e.g. Rician, provides the best possible “denoising” and noise bias correction automatically. Fig. 4a (left) shows that the distribution of MD values from a set of 1000 voxels *randomly selected* over the entire brain are the same with and without MP-PCA applied to the data (without introducing any imaging artifacts). Conversely, ROI homogeneity (Fig. 4a, right) allows MP-PCA to sometimes remove more noise than even an ideal model fitting would enable. (In an extreme case of all

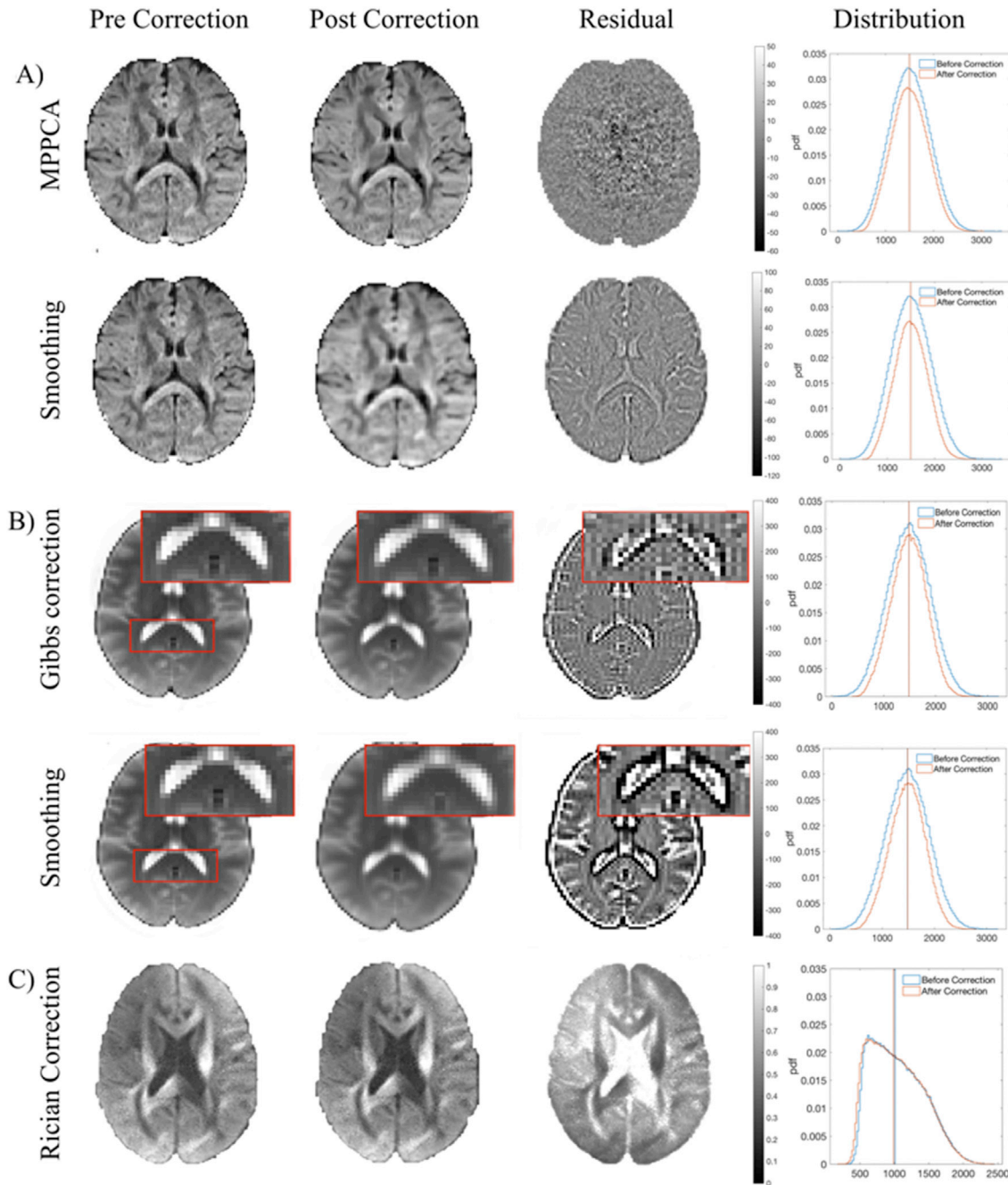


Fig. 2. Effects of DESIGNER steps on HCP phantom. A) Effects of MP-PCA denoising (first row) versus smoothing (second row) on a $b = 1000 \text{ s/mm}^2$ image. Residuals show that MP-PCA remove only noise while smoothing removes additional signal. B) Effect of the Gibbs artifact correction on a $b = 0$ image with Gibbs artifacts evident in the splenium of the corpus callosum before and after artifact correction, smoothing removes extra anatomy in addition to artifacts. C) Effect of the Rician bias correction on a $b = 2000 \text{ s/mm}^2$ image and the distribution of signal in the posterior limb of the internal capsule before and after correction.

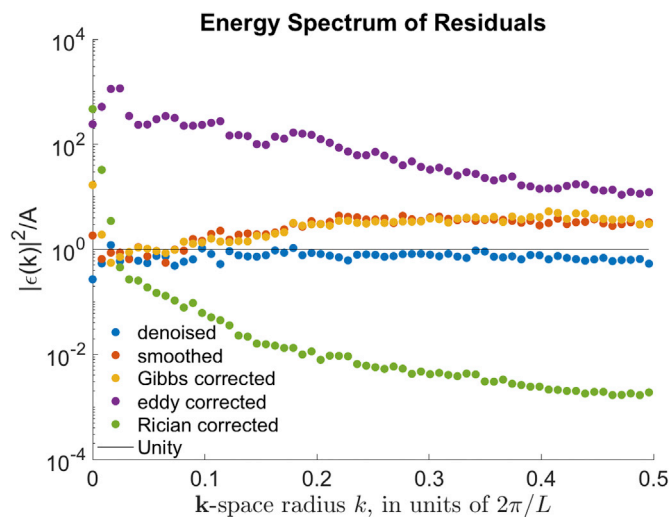


Fig. 3. Energy spectrum of residuals for $b = 1000$ s/mm² human data processed with DESIGNER. Residual maps were normalized by σ to give unitless energy values, therefore an energy of one implies that only noise with a standard deviation of σ was removed from the original image.

ROI voxels having the same ground truth parameters, there would be a single informative principal component for the whole ROI, irrespective of the model complexity, and the remaining components carry only the noise, which MP-PCA removes.)

We similarly found that MP-PCA denoising of an anatomically coherent region will minimize the variability in diffusivities irrespective of underlying SNR. We found that the number of significant principal components increases from 2 to 6 as phantom SNR increases from 5 to 50 (Fig. 4b). We also confirmed that the normalized variance of the mean diffusivity estimated from the HCP phantom depends linearly on the added noise variance σ^2 (i.e., $1/\text{SNR}^2$), in agreement with the Central limit theorem: $\sigma_D^2 = \sigma_{D,0}^2 + \alpha\sigma^2$. The slope α is greatest (2.67) for diffusivities estimated without any denoising, equivalent to independent voxel-wise model fit (as discussed above). Any improvement in precision (reduction in α) will be associated with the spatial redundancy. Smoothing captures this redundancy within the local Gaussian kernel, hence its α is notably lower (0.61); however, denoising with MP-PCA reduces α even further to 0.14, by maximizing the spatial redundancy effect.

In addition, we found that by varying the SNR it is possible to isolate

biological variability inherent to tissue in the HCP dataset in the PLIC from the variability that comes from added noise. From the intercept value $\sigma_{D,0}^2$ (extrapolating σ_D^2 to $\sigma = 0$), we found that in the ideal (noise free) phantom data, there is 4.47% relative variability of mean diffusivity in this region. Interestingly, we found 4.7% variability in the same phantom region in MD derived from the original phantom, meaning that the phantom is very close to the ideal noise free scenario.

3.1.2. Gibbs artifact correction

PSR for Gibbs correction (Fig. 3) increases along with k , meaning that the correction has an effect primarily in high contrast regions. Residuals depicted in Fig. 2b (brain images also from HCP phantom) demonstrate the effect of Gibbs artifact correction. Ideal residuals should show only periodic lines (period = 2 voxels) emanating from the boundaries between tissue types. Even though both the Gibbs correction used here and smoothing effectively remove artifacts, smoothing additionally removes relevant anatomic detail. This is particularly evident in the splenium of the corpus callosum, where ringing originating from the white matter-ventricle boundary causes an overestimation of white-matter intensities that are not smoothed away. Power spectra of residuals in Fig. 3 show that minimal additional signal was removed in addition to the high frequency ringing artifacts.

3.1.3. Rician bias correction

Power spectrum results for Rician correction exist only at low k , which makes sense since the Rician bias correction shifts intensities of the entire DWI downwards by σ . The effects of Rician bias correction on the HCP phantom and the relevant residual map through the corpus callosum in a $b = 2000$ s/mm² image are shown in Fig. 2c. This correction is most visible as a bias in raw tissue intensity, as illustrated by the histogram of $b = 2000$ s/mm² intensities in the posterior limb of the internal capsule and the negative shift in signal after bias correction.

3.1.4. EPI + eddy current and motion correction

PSR values for eddy current and motion correction have a very high amplitude at low frequencies because the operation is performed on the entire image rather than in specific regions. We find increased accuracy of motion correction (particularly at low SNR) using TORTOISE (C. Pierpaoli et al., 2010) and FSL-eddy (Smith et al., 2004) after data has been denoised. In Supplementary Fig. 1 we plot the mutual information between the HCP phantom and the same phantom after adding rotation, translation, and Rician distributed noise. We performed motion correction and MPPCA denoising on the phantom to measure the dependence of motion correction on image SNR. It is visible from Supplementary Fig. 1 that denoising increases the accuracy of motion correction at low

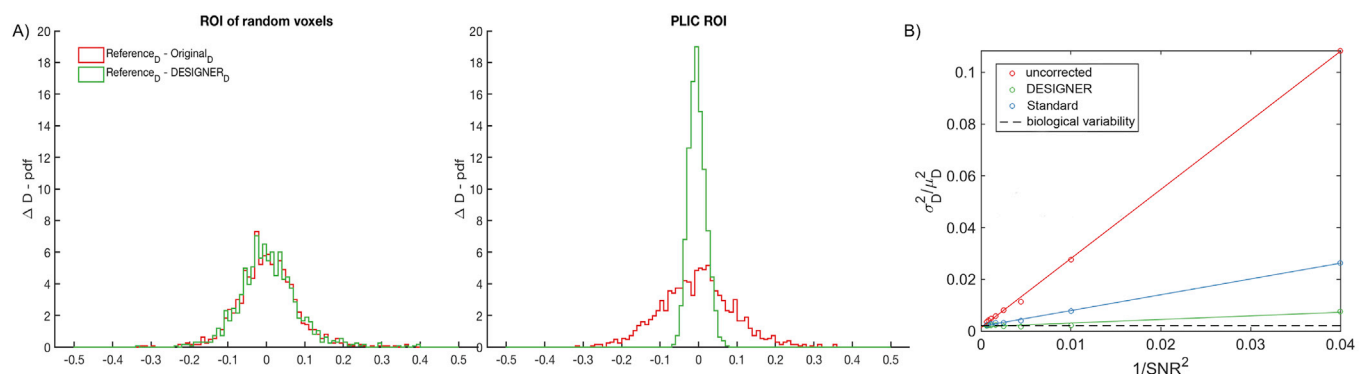


Fig. 4. A) The distribution of MD deviation from the ground truth in an ROI of randomly distributed voxels over the entire brain when denoising is applied prior to fitting and when no denoising is applied, compared to the distribution of MD in an ROI of the PLIC. This example illustrates that a correct signal model/representation works as well as denoising (i.e., in the perfect model case denoising should not yield extra benefit if the fitting is unbiased and no imaging artifacts are introduced). B) Variability of Diffusivity scales linearly with noise variance (PLIC of HCP phantom). The residual variance extrapolated to zero noise level provides an estimate for the inherent biological variability.

b-values (SNR = 2.7) by a factor of 2.

3.1.5. Outlier detection and correction

Fig. 5 shows a $b = 1000 \text{ s/mm}^2$ image containing a ghosting artifact, voxels that were shown to have large residuals during the fit, and a parametric map in which the effect of the artifact has been diminished. The bottom row shows the results of a fit routine using constrained linear least squares (CLLS) regression and how artifacts can potentially corrupt parametric maps if unaccounted. Designer and IRWLLS clearly removes the ghosting artifact depicted in Fig. 5.

3.2. Comparison of DESIGNER with standard pipelines

3.2.1. Accuracy

We found that DESIGNER provides the most accurate results, particularly in high contrast regions where there exists greater potential for partial volume effects, indicating that processing techniques that include smoothing decrease the overall contrast of diffusion parameters, which may bias the results of statistical analyses.

Fig. 6 shows that smoothing tends to bias parametric maps based on histograms of signal over WM, GM, and CSF in the HCP phantom. Parameters derived after processing with DESIGNER hold true to the ground truth signal distribution, and smoothing causes changes in the distribution shape due to partial volume effects. DESIGNER keeps extremely close to the same mean and median values as the ground truth, while other pipelines have large discrepancies. We also show the distribution of parameter values after minimal preprocessing with only EPI and eddy current correction (Original pipeline) to demonstrate that processing the signal prior to tensor fitting removes the bias caused by thermal fluctuations.

Diffusion parameters show decreased contrast when smoothed compared to DESIGNER. Fig. 7 shows the decrease in contrast present in FA, MD, MK, and AWF in actual subjects from HCP, rather than in the simulated brain phantom.

Results of an ROI analysis based on JHU-ICBM regions (Hua et al., 2008) indicate that on average, DESIGNER provides more accurate values than the standard pipeline that uses smoothing. A two-tailed t -test, comparing the mean value of DESIGNER and standard pipeline to that of the ground truth gave corresponding results. FA showed statistical difference from reference data in 0/48 and 42/48 regions with on average 3.36% and 12.84% relative difference from the reference for DESIGNER and standard pipeline, respectively. MD showed statistical difference from reference in 3/48 and 46/48 regions with average 1.56% and 10.22% relative difference from the reference for DESIGNER and

standard pipeline, respectively. MK showed statistical difference from reference data in 10/48 and 35/48 regions with on average 1.42% and 5.57% relative difference from the reference for DESIGNER and standard pipeline, respectively. Finally, AWF showed statistical difference from reference data in 9/48 and 43/48 regions with on average 1.65% and 9.22% relative difference from the reference for DESIGNER and standard pipeline, respectively. These data are tabulated in [supplementary material](#), and Fig. 8 provides boxplots that show the differences between groups for FA, MD, MK, and AWF.

3.2.2. Precision

The inherent variability of each pipeline is indicated in Fig. 9 by voxel-wise mean σ_n/μ_n maps over all 30 subjects in standard space. When no denoising is applied to raw data, there is up to 40% variability between the three scans, and σ_{FA}/μ_{FA} was 22% for the standard pipeline with smoothing, and 20% for DESIGNER. This means that DESIGNER achieves at least as good reproducibility as smoothing with $\text{FWHM} = 1.25 \times 1 \text{ voxel}$, albeit without the loss of anatomical detail, and without parameter biases due to partial volume. It is expected that a certain degree of smoothing can achieve reduction in the noise that matches that of MP-PCA denoising. At the same level of precision, the advantage of using DESIGNER is improved accuracy, as shown in Section 3.2.1. We also found that variability in data processed with only EPI and motion correction can be greater than that of DESIGNER by up to a factor of 2. Table 1 shows mean and standard deviation CoV values in GM, WM, and CSF for each diffusion parameter.

An ROI analysis over local white matter regions found that the CoV varies depending on the white matter regions analyzed. The corpus callosum was found to show the least variability compared to other white matter regions for FA, MD, and AWF, while MK showed greater variability due to the influence of outliers. We found that DESIGNER showed the lowest CoV ($\text{MD } 0.084 \pm 0.128$) in the genu compared to Standard ($\text{MD } 0.146 \pm 0.134$) and Original MD (0.162 ± 0.136) and these results are representative of the whole corpus callosum.

4. Discussion

The accuracy and precision analyses demonstrate that using DESIGNER to process diffusion images improves the quality of diffusion parameters compared to existing dMRI analysis approaches. DESIGNER maximizes the accuracy of diffusion metrics by modeling and correcting for several artifacts common in diffusion data collected with clinical scanners. When comparing HCP phantom parameters processed using DESIGNER and standard pipelines, we found that DESIGNER helps retain SNR and image contrast comparable to that of the reference values and that spatial smoothing lead to lower contrast. The distribution of FA, MD, MK, and AWF values were most similar to reference values and suffered from fewer outliers (particularly in MK) following DESIGNER processing. Smoothing led to systematic underestimation of FA and overestimation of MD values due to induced partial voluming with gray matter or CSF. Finally, we found in an ROI analysis that even in small white matter regions DESIGNER holds to the same distribution and ground-truth values, while the standard pipeline sees many significant differences. We also observed that pipeline accuracy depends on spatial location in the brain. We used a voxel wise, FWE corrected one-sample t -test to assess differences between the parameters obtained from the noisy versions of the phantom processed with each of the pipelines. We found that pure white matter regions are more accurate independent of pipeline choice, however DESIGNER does the best job preserving parameter accuracy in regions with ambiguous tissue types. These results are shown in [Supplementary Fig. 3](#).

DESIGNER maximizes the precision and reliability of DTI, DKI and white matter microstructural modeling parameters by reducing the effect of random noise on parametric maps. We observed that the use of the standard pipeline based on smoothing with $1.25 \times \text{voxel}$ size FWHM lead to higher variability and/or lower precision than that of DESIGNER,

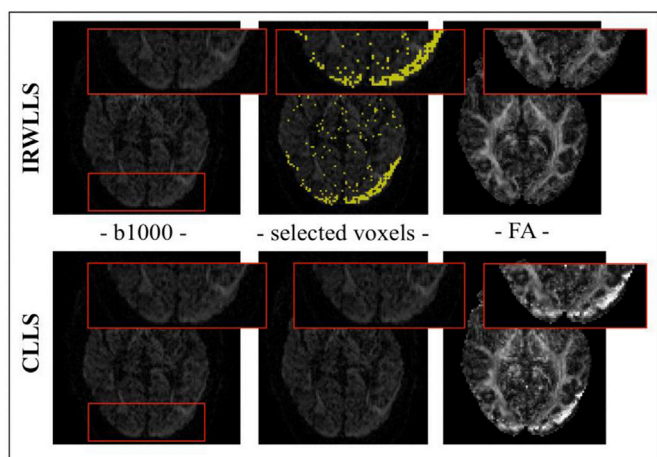


Fig. 5. IRWLLS detection of outlier voxels due to a chemical shift ghosting artifact in an example $b = 1000 \text{ s/mm}^2$ image. These voxels are reweighted during fitting to produce the corrected parametric maps.

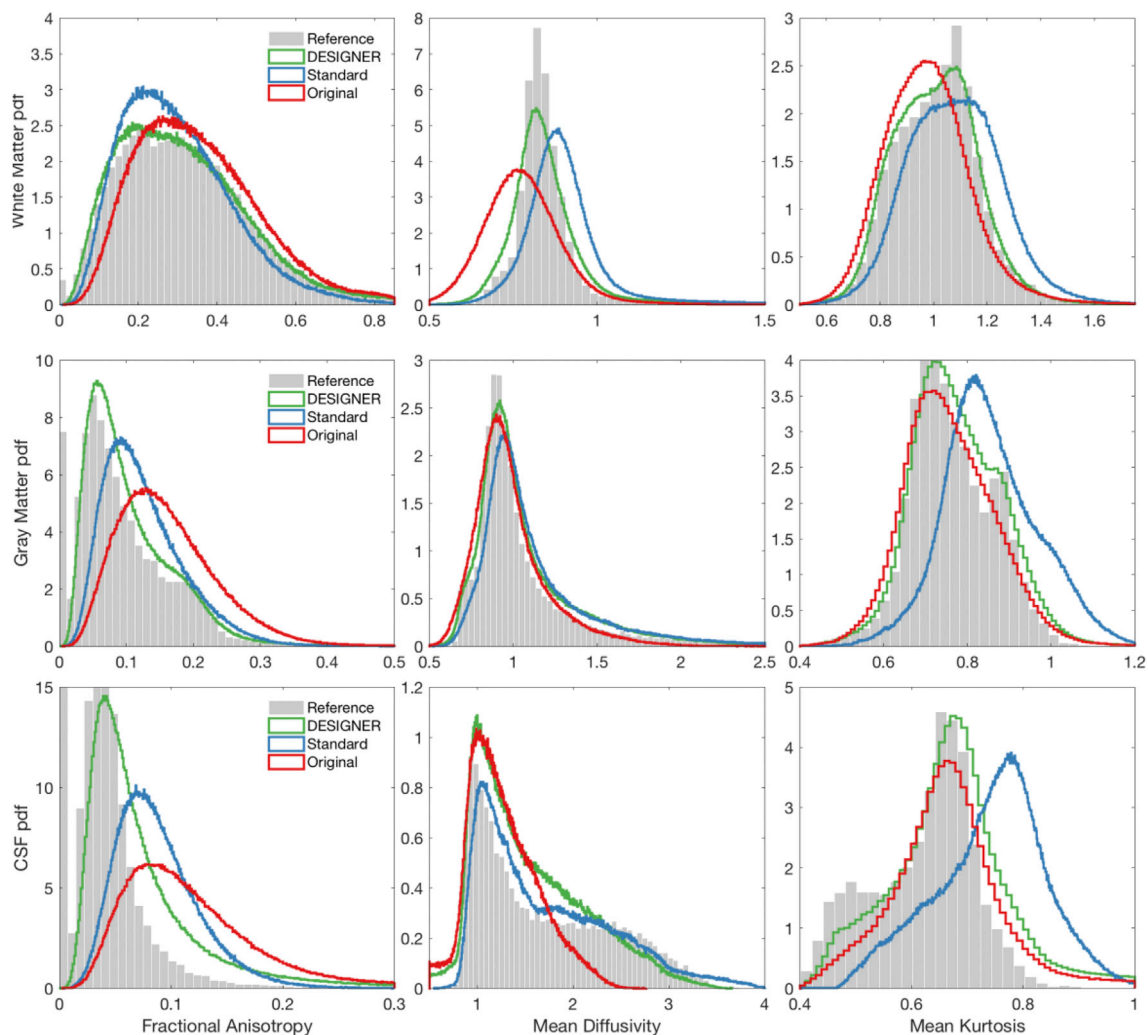


Fig. 6. Bias inherent to each pipeline and comparison to the ground truth. These images are based on ROIs of total white matter, gray matter, and CSF of the HCP brain phantom. Histograms are of FA, MD, MK, and AWF from left to right respectively. The distribution of parameters processed with DESIGNER hold more closely to that of the ground truth compared to processing using methods performing either smoothing (Standard) or no denoising (Original) prior to EPI, eddy and motion correction.

similarly, smoothing with greater than the proposed FWHM will result in lower variability but also lower anatomic specificity. We also found that variability tends to vary depending on spatial location in brain tissue and on the diffusion parameter in question. The WM tends to have the most stable CoV values compared to other regions, likely due to spatially varying physiological noise and how tissue is distributed in the brain. In addition to physiological noise, the number of outliers created during tensor estimation and the degree of smoothing affect the magnitude of CoV values, which is one reason why FA (due to eigenvalue repulsion) and MK (low radial diffusivity outlier effects) tend to have higher CoV values. While increasing the amount of smoothing may in principle increase the precision of results without limit, it also results in misleading statistical outcomes. Therefore, any pipeline should be evaluated both with respect to accuracy and precision.

We also show here that the benefits of spatial redundancy depend on the window kernel for MP-PCA denoising. When the underlying data matrix includes different tissue types, such that the underlying tissue parameters within the data matrix are sufficiently different, the lack of spatial redundancy might lower the benefit of the denoising step if the diffusion and noise model are known *a priori*. Indeed, to take this idea to a logical extreme by applying MP-PCA to a random set of voxels (Fig. 4a left), we see very similar variability between MP-PCA and tensor fitting

alone. This is because at b -values in the range of 1000–3000 s/mm^2 , the cumulant expansion (DTI, DKI) provides adequate representation of the signal's functional form, whereas spatial redundancy is cancelled by the random voxel selection. Alternatively, when denoising is applied on a coherent set of voxels from a local ROI (e.g., PLIC in Fig. 4a right), we see a very strongly decreased variance in the distribution of MD, indicating that we have removed notably *more* thermal noise than what would have been possible through voxel-wise DTI or DKI tensor fitting alone, with or without MP-PCA applied prior to that. The subsequent steps of DESIGNER applied after MP-PCA denoising benefit greatly from this boost to image SNR (see Supplementary Fig. 1: increased accuracy in motion correction after MPPCA denoising). Future work might focus on adaptive denoising kernels to fully exploit spatial redundancy, thereby enabling consistently improved precision of parameter estimators such as the maximum likelihood estimator.

In addition to improved accuracy and precision, DESIGNER also improves the performance of individual artifact correction techniques. In particular, the estimation of motion or distortion parameters is very challenging in case of low SNR and prone to local minima, and residual image misalignments or artifacts will lower the accuracy and precision of diffusion parameter estimators. As shown in Supplementary Fig. 1, preceding motion correction with the denoising step provides for more

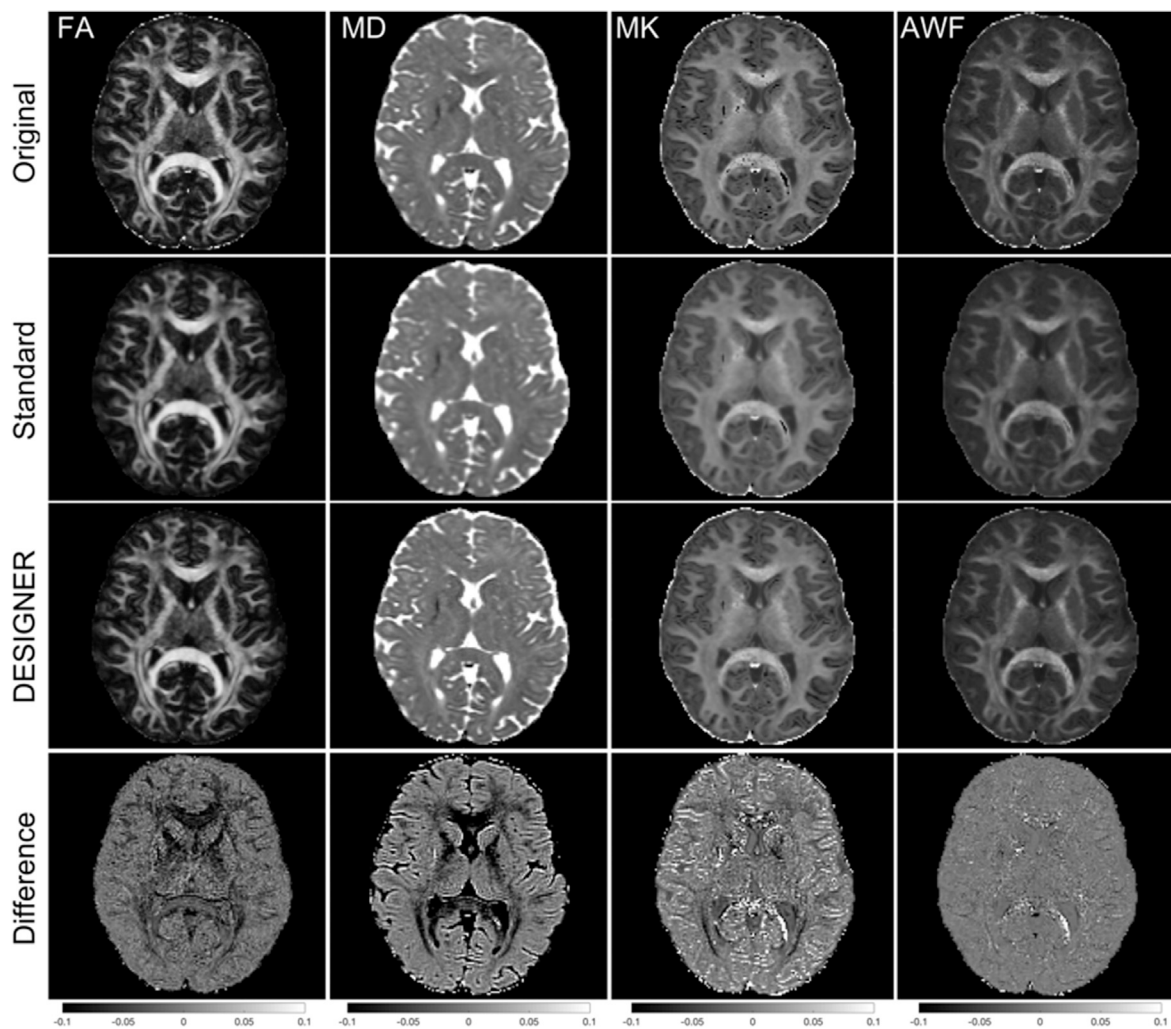


Fig. 7. Example, real-subject dataset - difference in contrast for FA, MD, MK, and AWF for DESIGNER, original pipeline (EPI, eddy current and motion correction), standard pipeline (smoothing prior to EPI, eddy current, motion correction), and the difference between DESIGNER and original pipeline.

stable correction and therefore likely benefits all subsequent processing steps.

DESIGNER is not necessarily tied to diffusion tensor or diffusion kurtosis imaging. We reasonably expect that the reduction of noise and artifacts in DWI datasets will also increase the accuracy and precision of any kind of biophysical model parameter estimation (e.g. based on the recently introduced rotational invariants (Novikov et al., 2018b)) as well as estimating parameters of popular signal representations, e.g. based on spherical harmonics, and subsequently the accuracy of fiber tractography results as well. In the future, we intend to show that this pipeline is capable of reducing the amount of spurious fibers generated during tractography.

One limitation of this study is that even though DESIGNER reduces the amount of image smoothing, it is not fully smoothing-free, especially due to steps involving image interpolation which include Gibbs correction, EPI distortion correction, and eddy current/motion correction. While it is impossible to avoid all sources of partial volume effects in signal processing, we believe we may have come as close as practically possible. We expect that the DESIGNER pipeline will help bridge the gap between diffusion microstructural imaging and clinical practice, by improving robustness, reproducibility, accuracy and precision, as well as by enabling separation of acquisition/modeling and biological variability (Novikov et al., 2018a). One of the fundamental reasons for the development of DESIGNER is to enable the further clinical feasibility of diffusion imaging in general, and DESIGNER accomplishes this by

boosting confidence in statistical analyses through corresponding confidence in parameter accuracy.

DESIGNER is currently available online and can be found at: [www.github.com/NYU-DiffusionMRI/Diffusion-Kurtosis-Imaging/designer](https://github.com/NYU-DiffusionMRI/Diffusion-Kurtosis-Imaging/designer) (<https://www.nitrc.org/projects/mppca/>). The current implementation is dependent on MRtrix preprocessing toolbox (Tournier et al., 2012) (v. 3.0, www.mrtrix.org) and FMRIB's diffusion toolbox (Andersson and Sotiropoulos, 2016) (*eddy*, FSL v. 5.0.10, www.fmrib.ox.ac.uk/fsl), as well as tensor estimation software (Veraart et al., 2013).

5. Conclusions

DESIGNER enables the robust estimation of parametric diffusion maps with improved accuracy and without compromising precision compared to standard processing routines based on spatial smoothing or without any form of denoising. The specific pre-processing steps used in DESIGNER (MP-PCA denoising, Gibbs artifact correction, Rician bias correction, EPI + eddy current and motion correction, IRWLLS fitting) alleviate the need for smoothing. Our results suggest that diffusion parameter estimation benefits most from noise and artifact specific processing techniques, and that further effort should be spent in analyzing specific sources of noise and artifacts in MRI and accounting for them individually, rather than by one-size-fits-all smoothing which potentially damages datasets.

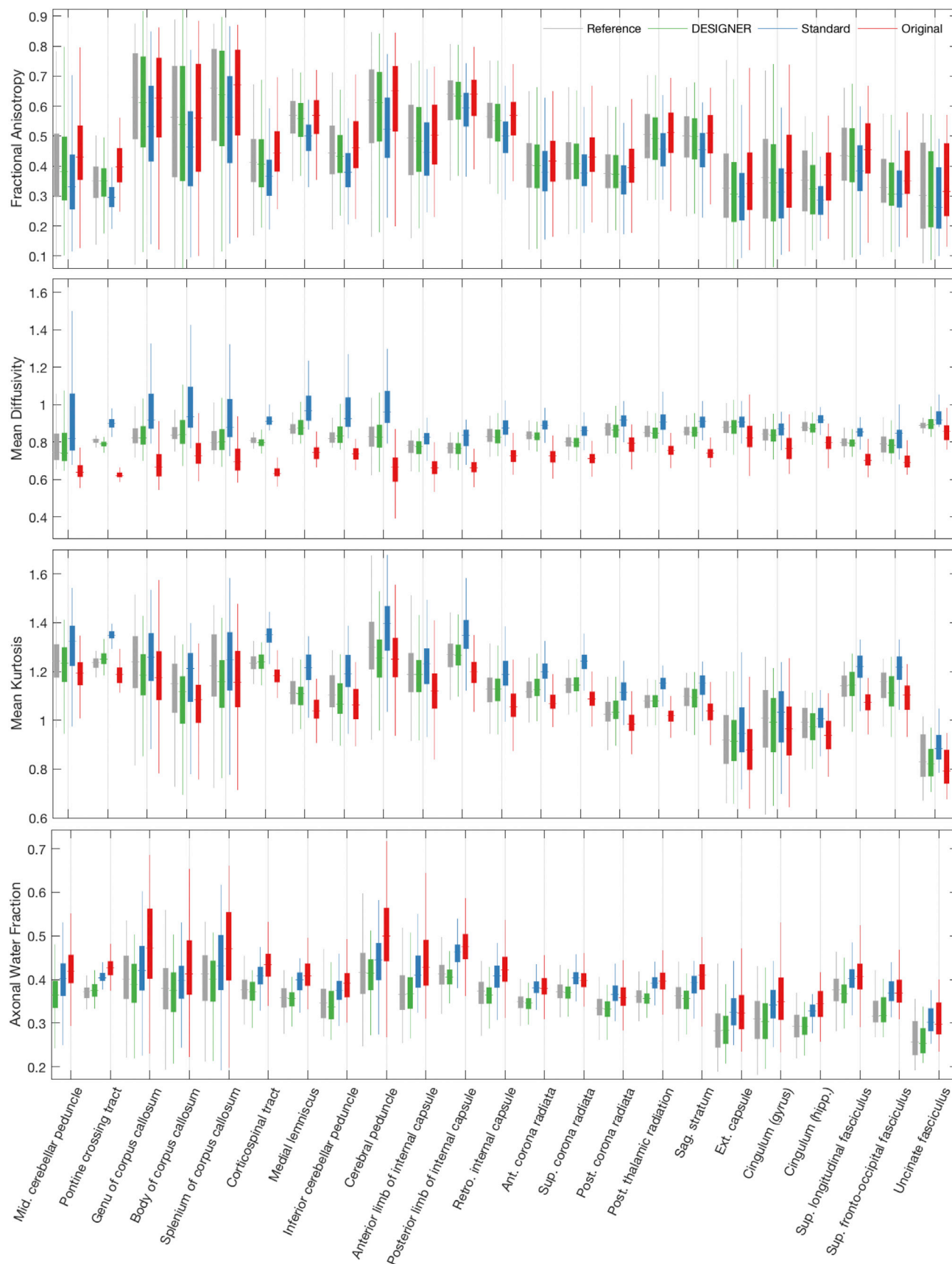


Fig. 8. Boxplots that show FA, MD, MK, and AWF values in 23 ROIs from the JHU White Matter atlas. ROIs over the left and right hemispheres have been averaged since there are very few differences across hemispheres that pertain to this analysis. Reference boxplots for HCP phantom data come from an ROI of the original phantom after averaging 50 noise realizations, boxplots of DESIGNER and standard and original pipelines represent the ROI after preprocessing and averaging over noise realizations.

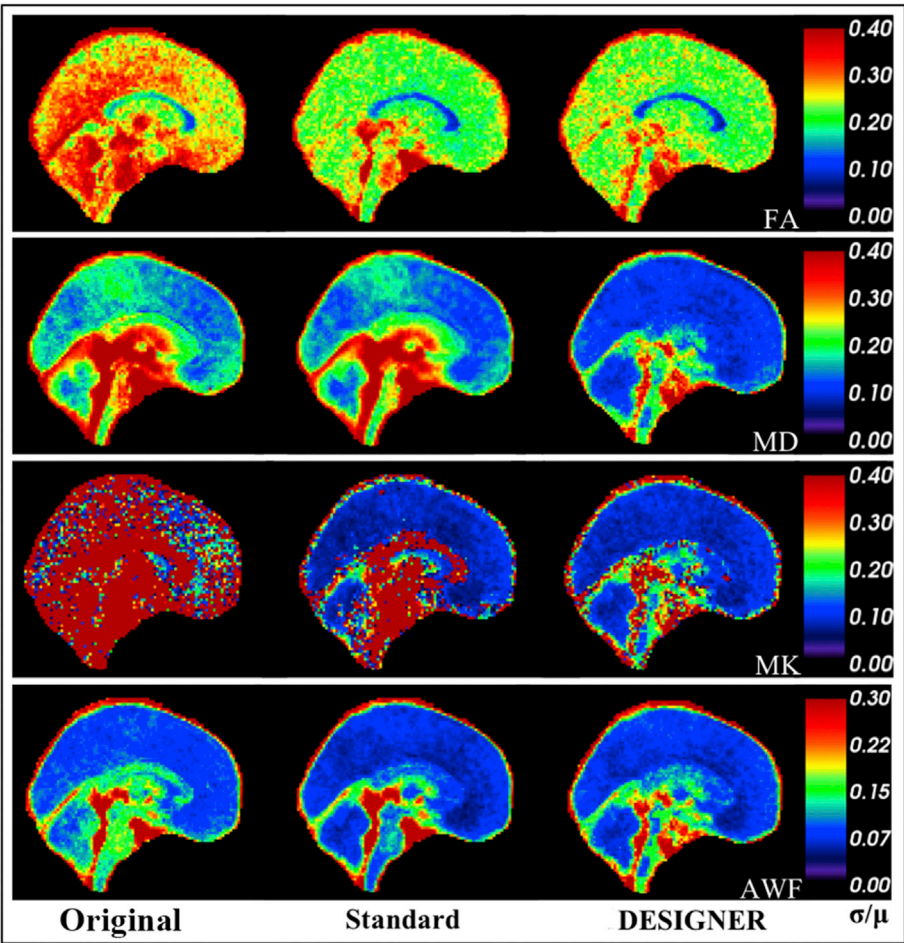


Fig. 9. Comparison of mean σ_n/μ_n over all 30 subjects for three cases: From left to right, the images show within subject variability when the original pipeline with only EPI distortion, eddy current and motion correction is applied, when DESIGNER is applied, and when the standard pipeline with smoothing (1.25xVS) is applied. The coefficient of variation is lowest in parametric maps processed with DESIGNER compared to processing with both topup + eddy and with smoothing.

Table 1
Mean and standard deviation σ_n/μ_n values for FA, MD, MK, and AWF over all CSF, white matter, and gray matter, over all sets of subjects. We compare the coefficient of variation in white matter, gray matter, and CSF between the three pipelines – Original, Standard with smoothing at a FWHM of 1.25 x voxel size, and DESIGNER.

		FA	MD	MK	AWF
Original	CSF	0.262 ± 0.174	0.164 ± 0.207	0.147 ± 0.231	0.105 ± 0.156
	WM	0.271 ± 0.158	0.188 ± 0.206	0.141 ± 0.218	0.107 ± 0.132
	GM	0.271 ± 0.157	0.189 ± 0.201	0.137 ± 0.214	0.107 ± 0.128
Standard	CSF	0.207 ± 0.175	0.141 ± 0.209	0.094 ± 0.193	0.079 ± 0.163
	WM	0.218 ± 0.163	0.167 ± 0.208	0.094 ± 0.181	0.083 ± 0.142
	GM	0.217 ± 0.160	0.168 ± 0.203	0.093 ± 0.177	0.082 ± 0.137
Designer	CSF	0.193 ± 0.165	0.101 ± 0.185	0.097 ± 0.186	0.080 ± 0.147
	WM	0.203 ± 0.143	0.114 ± 0.175	0.100 ± 0.176	0.084 ± 0.108
	GM	0.202 ± 0.142	0.113 ± 0.172	0.100 ± 0.172	0.084 ± 0.106

Disclosure

EF, JV and DSN are co-inventors in technology related to this research; a patent application has been filed and is pending. EF and DSN are shareholders and hold advisory roles at Microstructure Imaging, Inc.

Acknowledgements

Research was supported by The National Institute of Neurological Disorders and Stroke of the NIH under award number R01 NS088040, and was performed at the Center of Advanced Imaging Innovation and Research (CAI2R, www.cai2r.net), and NIBIB Biomedical Technology Resource Center P41 EB017183. The authors would like to thank Dr.

Jesper Andersson and Dr. Steve Smith for their helpful discussion and correspondence during the production of thus work. We would also like to thank Dr. Carlo Pierpaoli and the TORTOISE team for their invaluable help. Research support was also provided by the Human Brain Mapping Project, which is jointly funded by NIMH and NIDA (P20 MH/DA52176), by General Clinical Research Core (HSC19940074H). JV is a Postdoctoral Fellow of the Research Foundation - Flanders (FWO; grant number 12S1615N).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.neuroimage.2018.07.066>.

References

- Ades-Aron, B., Veraart, J., Kellner, E., Lui, Y.W., Novikov, D.S., Fieremans, E., 2016. Diffusion Parameter Estimation with Gibbs and Noise Removal (DESIGNER). ISMRM, Singapore.
- Aja-Fernandez, S., Alberola-Lopez, C., Westin, C.F., 2008. Noise and signal estimation in magnitude MRI and Rician distributed images: a LMMSE approach. *IEEE Trans. Image Process.* 17, 1383–1398.
- Amartur, S., Haacke, E.M., 1991. Modified iterative model based on data extrapolation method to reduce Gibbs ringing. *J. Magn. Reson. Imag.* 1, 307–317.
- Amartur, S., Liang, Z.P., Boada, F., Haacke, E.M., 1991. Phase-constrained data extrapolation method for reduction of truncation artifacts. *J. Magn. Reson. Imag.* 1, 721–724.
- Andersson, J.L.R., Graham, M.S., Zsoldos, E., Sotiropoulos, S.N., 2016. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage* 141, 556–572.
- Andersson, J.L.R., Skare, S., 2002. A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. *Neuroimage* 16, 177–199.
- Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* 20, 870–888.
- Andersson, J.L.R., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063–1078.
- Archibald, R., Gelb, A., 2002. A method to reduce the Gibbs ringing artifact in MRI scans while keeping tissue boundary integrity. *IEEE Trans. Med. Imag.* 21, 305–319.
- Barker, G.P., Wheeler-Kingshott, C., 2001. Gibbs ringing and negative ADC values. In: *Proceedings of the 9th Meeting of the International Society for Magnetic Resonance in Medicine*, Glasgow (UK), p. 1546.
- Basser, Peter J., Mattiello, James, LeBihan, Denis, 1994. MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66 (1), 259–267.
- Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed.* 15, 435–455.
- C. Pierpaoli, L.W., Irfanoglu, M.O., Barnett, A., Basser, P., Chang, L.-C., Koay, C., Pajevic, S., Rohde, G., Sarlls, J., Wu, M., 2010. TORTOISE: an Integrated Software Package for Processing of Diffusion MRI Data. ISMRM, Stockholm, Sweden.
- Collier, Q., Veraart, J., Jeurissen, B., den Dekker, A.J., Sijbers, J., 2015. Iterative reweighted linear least squares for accurate, fast, and robust estimation of diffusion magnetic resonance parameters. *Magn. Reson. Med.* 73, 2174–2184.
- Cui, Z.X., Zhong, S.Y., Xu, P.F., He, Y., Gong, G.L., 2013. PANDA: a pipeline toolbox for analyzing brain diffusion images. *Front. Hum. Neurosci.* 7.
- Fieremans, E., Jensen, J.H., Helpen, J.A., 2011. White matter characterization with diffusional kurtosis imaging. *Neuroimage* 58, 177–188.
- Gudbjartsson, H., Patz, S., 1995. The rician distribution of noisy mri data. *Magn. Reson. Med.* 34, 910–914.
- Haider, M.A., van der Kwast, T.H., Tanguay, J., Evans, A.J., Hashmi, A.T., Lockwood, G., Trachtenberg, J., 2007. Combined T2-weighted and diffusion-weighted MRI for localization of prostate cancer. *Am. J. Roentgenol.* 189, 323–328.
- Hua, K., Zhang, J.Y., Wakana, S., Jiang, H.Y., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C.M., Mori, S., 2008. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage* 39, 336–347.
- Jensen, Jens H., et al., 2005. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Mag. Reson. Med.: Official J. Int. Soc. Mag. Reson. Med.* 53 (6), 1432–1440.
- Jones, D.K., 2010. *Diffusion MRI: Theory, Methods, and Application*. Oxford University Press, Oxford; New York.
- Jones, D.K., Horsfield, M.A., Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn. Reson. Med.* 42, 515–525.
- Kellner, E., Dhital, B., Kiselev, V.G., Reiser, M., 2016. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn. Reson. Med.* 76, 1574–1581.
- Kiselev, Valerij G., 2010. The cumulant expansion: an overarching mathematical framework for understanding diffusion NMR. *Diffus. MRI* 152–168.
- Kiselev, V.G., 2017. *Fundamentals of diffusion MRI physics*. NMR Biomed. 30.
- Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P., 2010. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imag.* 29, 196–205.
- Koay, C.G., Basser, P.J., 2006. Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. *J. Magn. Reson.* 179, 317–322.
- Koh, D.M., Collins, D.J., 2007. Diffusion-weighted MRI in the body: applications and challenges in oncology. *Am. J. Roentgenol.* 188, 1622–1635.
- Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., Laval-Jeantet, M., 1986. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 161, 401–407.
- Le Bihan, D., Poupon, C., Amadon, A., Lethimonnier, F., 2006. Artifacts and pitfalls in diffusion MRI. *J. Magn. Reson. Imag.* 24, 478–488.
- Moseley, M.E., Cohen, Y., Mintonovitch, J., Chileuit, L., Shimizu, H., Kucharczyk, J., Wendland, M.F., Weinstein, P.R., 1990. Early detection of regional cerebral ischemia in cats: comparison of diffusion- and T2-weighted MRI and spectroscopy. *Magn. Reson. Med.* 14, 330–346.
- Novikov, D.S., Kiselev, V.G., Jespersen, S.N., 2018a. On modeling. *Magn. Reson. Med.* 79, 3172–3193.
- Novikov, D.S., Veraart, J., Jelescu, I.O., Fieremans, E., 2018b. Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. *Neuroimage* 174, 518–538.
- Perrone, D., Aelterman, J., Pizurica, A., Jeurissen, B., Philips, W., Leemans, A., 2015. The effect of Gibbs ringing artifacts on measures derived from diffusion MRI. *Neuroimage* 120, 441–455.
- Sijbers, J., den Dekker, A.J., 2004. Maximum likelihood estimation of signal amplitude and noise variance from MR data. *Magn. Reson. Med.* 51, 586–594.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y.Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219.
- Sokal, R.R., Rohlf, F.J., 1981. *Biometry: the Principles and Practice of Statistics in Biological Research*, 2d ed. W. H. Freeman, San Francisco.
- Tabesh, A., Jensen, J.H., Ardekani, B.A., Helpen, J.A., 2011. Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. *Magn. Reson. Med.* 65, 823–836.
- Tournier, J.D., Calamante, F., Connelly, A., 2012. MRtrix: diffusion tractography in crossing fiber regions. *Int. J. Imag. Syst. Technol.* 22, 53–66.
- Tournier, J.D., Mori, S., Leemans, A., 2011. Diffusion tensor imaging and beyond. *Magn. Reson. Med.* 65, 1532–1556.
- Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.U.-M.H., 2013. The WU-minn human connectome project: an overview. *Neuroimage* 80, 62–79.
- Veraart, J., Fieremans, E., Jelescu, I.O., Knoll, F., Novikov, D.S., 2016a. Gibbs ringing in diffusion MRI. *Magn. Reson. Med.* 76, 301–314.
- Veraart, J., Fieremans, E., Novikov, D.S., 2016b. Diffusion MRI noise mapping using random matrix theory. *Magn. Reson. Med.* 76, 1582–1593.
- Veraart, J., Novikov, D.S., Christiaens, D., Ades-Aron, B., Sijbers, J., Fieremans, E., 2016c. Denoising of diffusion MRI using random matrix theory. *Neuroimage* 142, 384–396.
- Veraart, J., Rajan, J., Peeters, R.R., Leemans, A., Sunaert, S., Sijbers, J., 2013. Comprehensive framework for accurate diffusion MRI parameter estimation. *Magn. Reson. Med.* 70, 972–984.
- Veraart, J., Van Hecke, W., Sijbers, J., 2011. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a rician noise model. *Magn. Reson. Med.* 66, 678–686.
- Warach, S., Gaa, J., Siewert, B., 1995. Acute human stroke studied by whole-brain echo-planar diffusion-weighted magnetic-resonance-imaging (Vol 37, Pg 231, 1995). *Ann. Neurol.* 37, 688–690.
- Zhang, Y.Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.* 20, 45–57.