

Denoising for Diffusion Tensor Imaging with Low Signal to Noise Ratios: Method and Monte Carlo Validation

KLAUS R. HAHN, SERGEI PRIGARIN, KARSTEN RODENACKER, & KHADER M. HASAN

ABSTRACT: Diffusion Tensor Imaging (DTI) is a magnetic resonance technique which enables the in vivo visualisation and characterisation of nerve fibers in the human brain indicating diseases like Multiple Sclerosis or Alzheimer. Increasing the sensitivity by a reduction of the voxel size or by application of measurement parameters with better diffusion differentiation induces low signal to noise ratios (SNR) in the measured diffusion weighted images (DWI) which collect the information for DTI. Due to nonlinear noise propagation within the DTI formalism the fiber properties are modelled by biased non Gaussian random fields. We present a new denoising method for DTI with low SNR and explore it by Monte Carlo simulations based on a phantom and on human brain data with high resolution 1 mm^3 voxels. The method combines voxelwise averaging, nonlinear spatial filtering and a modified Rician bias correction of the DWIs to reduce bias and noise in the estimated images of fiber properties. These procedures reduce the main part of the mean squared error, but due to sample size restrictions residual noise may remain. Therefore, (post) filtering is finally applied directly to those denoised images. The method is linked to the Delta Method formalizing the asymptotic Gaussian limit of nonlinear noise propagation. The simulations demonstrate the feasibility of a quantitative analysis of DTI data of the human brain with 1 mm^3 resolution, measured at a clinical field strength of 3T, if thermal noise is the source of the dominating artifacts. In addition, the results support the design of new high resolution experiments. The method is not limited to the standard tensor model.

AMS 2000 CLASSIFICATIONS: Statistics, data analysis 62-07, Biology and other natural sciences, biomedical imaging and signal processing 92C55, Stochastic systems and control, data smoothing 93E14, Stochastic systems and control, filtering 93E11.

1. INTRODUCTION

Diffusion Tensor Imaging (DTI) is a recent magnetic resonance technique to detect the structure and connectivity of nerve fibers in the human brain in vivo. Clinical investigations show that e.g. Multiple Sclerosis and Alzheimer disease can be indicated by this technique [1, 2].

Experimental and theoretical research to improve the accuracy of DTI, is in a stage of development. Thermal or Johnson noise is one of the main obstacles, introducing systematic errors in scalar and vectorial diffusion quantities [3]. DTI data with high thermal noise level or equivalently with low SNR can be a result of the selection of small voxel sizes [4], low field strengths, or high diffusion sensitizing b -values [5]. Current DTI protocols use voxel sizes of 8-27 mm^3 [4]. DTI with smaller voxels would reduce partial volume effects causing “phantom connections” between nerve fiber tracts and erroneous mixing of grey and white matter which modifies scalar diffusion quantities like e.g. the fractional anisotropy FA [6]. See **Appendix 1** for a definition of SNR and of FA and for a short description of the DTI formalism.

Klaus R. Hahn* & Karsten Rodenacker: Institute of Biomathematics and Biometry, German Research Center for Environmental Health - HMGU, D-85764 Neuherberg, Ingolst. Landstr. 1, Germany, *E-mail: hahn@gsf.de**

Sergei Prigarin: Institute of Computational Mathematics and Mathematical Geophysics, Novosibirsk, Russia.

Khader M. Hasan: Department of Diagnostic and Interventional Imaging, University of Texas Medical School at Houston, 6431 Fannin, Houston, Texas USA.

Focussing on noise, we find, that due to nonlinear noise propagation within the DTI formalism the distributions of the diffusion quantities become with decreasing SNR skewed, biased and heteroscedastic. This is exemplified in numerous Monte Carlo, bootstrap and perturbation theory studies which are based on the Rician noise model for the diffusion weighted images (DWI), see [7] for a review and **Appendix 2** for some properties of Rician distributions. Proposals for reduction of noise artefacts are in most cases restricted to moderate and high SNR > 5 , where bias effects in the DWIs can be ignored. An important role herein occupies denoising by filtering, where e.g. non-linear, edge preserving procedures [8-11] and spline applications [12] are in use.

Denoising methods of DTI data with low SNR are – to our knowledge – published at present in only few papers, see e.g. [13] Basu *et al.*, 2006 and [14] Fillard *et al.*, 2006. The main problem addressed there is the strong influence of the Rician mean value bias in the DWIs. Basu *et al.*, incorporated a Rician likelihood term via the Bayesian approach into the diffusion equation which is then applied as a spatial filter to the DWIs. By Fillard *et al.*, a variational approach incorporating the Rician distribution and an anisotropic regularization to estimate the tensor via a Log-Euclidian metrics is proposed. By Zhang *et al.*, 2008 [15] the Rician noise model is incorporated into a multi-channel wavelet-based denoising method, where wavelet transforms with anisotropic nonlinear diffusion are used. Martin-Fernandez *et al.*, 2009 [16] apply a multichannel Wiener filter with Rician bias correction to the DWIs and incorporate anisotropic neighbourhoods and sequential application of the filter. In all approaches inclusion of the Rician noise model improves the results for low SNR DTI compared to those calculated under the assumption of unbiased Gaussian noise. Similar results are published by Hahn *et al.*, 2005 [17], where a preliminary version of our method was presented. However, inclusion of the Rician noise model and the bias correction (BC) are not the only problems connected with denoising of low SNR DTI. The problems with residual noise effects at the noise floor (SNR = 0) and in tensor-derived fiber quantities, but also the need for a realistic gold standard model are further challenges.

We present a novel denoising procedure which enables numerical post-processing of DTI data with a minimum SNR ≈ 0.5 in the DWIs. The Monte Carlo validation of the procedure is based on a gold standard model derived from human brain DTI data with $1 \times 1 \times 1 \text{ mm}^3$ resolution. To ensure that the tensor-derived quantities can be estimated reliably, denoising must be performed on large samples. Consequently, we propose a multi-step procedure. First, the DWIs of NEX experimental replications are averaged voxelwise. As NEX is restricted by scan time limitations, mainly due to patient motion, the averaged DWIs are further denoised by nonlinear spatial filtering. Application of a modified voxelwise Rician BC to the denoised DWIs then follows. Due to the high slope of BC at low SNR and due to inherent sample size limitations, the achieved SNR enhancement in the fiber quantities is frequently not sufficient for a reliable quantitative analysis. Therefore, in a final step, nonlinear filtering is applied directly to those images of fiber quantities. We demonstrate that our method approximates the Gaussian limit of nonlinear noise propagation as is formalized by the statistical Delta Method [18]. Due to the problem of residual noise, the validations are performed on ensembles of Monte Carlo experiments with different random seeds for simulated noise. For spatial smoothing a robust nonlinear filter chain is adapted to spatially varying Rician noise and to high curvature, typical for DWIs. The Rician BC is regularized by a modification at the noise floor to prevent singular diffusion. For postfiltering a model of the standard deviation of residual noise is proposed, based on the error propagation for FA and mean diffusivity (MD).

2. METHODS AND MATERIALS

2.1 Statistical Aspects of low SNR DTI

Neglecting the influence of non-thermal noise, caused e.g. by scanner instabilities, patient motion or cardiac pulsation, a simplified stochastic model of noise in DTI can be formulated, where the DWIs have Rician

distributions. An illustration of nonlinear noise propagation and of our denoising strategy is given in Fig. 1 under the simplifying assumption that spatial filtering and voxelwise averaging are equivalent. Distributions for FA in a single voxel derived from Monte Carlo simulations with a minimal SNR ≈ 1 in the DWIs are shown. Fifteen icosahedral diffusion gradients [20, 21] measure in this model cigar shaped diffusion, with $\lambda_1 = 0.00155$, $\lambda_2 = \lambda_3 = 0.00027 \text{ mm}^2 \text{ s}^{-1}$, FA = 0.8 and $\text{SNR}_0 = 42$. Without DWI averaging skewness and a heavy tail is apparent, see distribution (a). A high risk to draw an outlier is evident. DWI averaging of 20 replications introduces an approximate Gaussian shape but also a bias (b). Application of BC to the DWIs produces a less perfect Gaussian shape but eliminates the mean value bias, see (c). It is clear from the Figure that (c) is best suited to estimate the true FA value (vertical bar) drawing a random number. The goal of the proposed denoising method is to achieve distributions like (c) for all relevant DTI quantities. The distribution (c) can be further narrowed and comes closer to Gaussian shape if more replications in the DWIs are applied or if, by a final denoising step, samples of (c) are averaged to replace (c) by its mean value distribution (postfiltering). The properties of the distributions in Fig. 1 are in line with the statistical Delta Method (DM). The DM predicts approximate Gaussian statistics for the DTI quantities, if noise effects in the DWIs are sufficiently reduced. This follows from an iterative application of the following Theorem, which we present in a simplified univariate manner: If the distributions of a sequence of random variables T_m approach with increasing m the Gaussian distributions $N(\Theta, \tau^2/m)$, where Θ is the expectation value and τ^2/m the variance, then, for a nonlinear mapping $f: T_m \rightarrow f(T_m)$ with $\dot{f}(\Theta) \neq 0$, the distributions of $f(T_m)$ tend to $N(f(\Theta), \dot{f}(\Theta)^2 \tau^2/m)$, see Lehmann [18] for formally exact formulations including the multivariate cases.

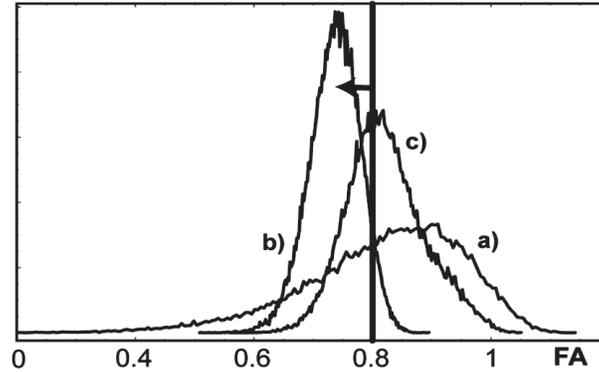


Figure 1: FA Distributions for Low SNR. DWI Averaging with NEX = 1, (a) and NEX = 20, (b). For Distribution (c) an Additional Bias Correction is Performed. Vertical Bar Gives the True FA Value Θ , The Arrow Indicates the Bias Shift of the Mean for (b); The Mean Values of (a) and (c) are Practically Unbiased

In our method, denoising is performed essentially on the DWIs $|S_j(\mathbf{x})|$. The advantage of DWIs relies in case of a high noise level on the fact that a noise model is formally known where bias and variance are parameterized, at least approximately, by measurable quantities [7]. Bias and variance in DTI are caused partly by hidden quantities, like e.g. by the angles between the unknown diffusion and the diffusion gradients, therefore the parameterization within the Rician model is an important advantage. On the other hand, spatial DWI filtering is technically difficult, as DWIs contain all gradient projected information about the measured diffusion process and carry correspondingly more “information-complexity” than specific fiber properties. Finally, an identification of the filtered DWIs with their Rician expectation values is necessary for the bias correction and enforces a modification of the Rician BC close to the noise floor (SNR = 0). One might argue that the unbiased complex signal $S_j(\mathbf{x})$ could be a better candidate for denoising as it is corrupted only by Gaussian noise in the real and imaginary channels. However, it was shown by Sijbers and Dekker [22] that an estimation of $S_j(\mathbf{x})$ is quite sensitive to instabilities in the phase shifts. As in practice phase artefacts caused by temporal instabilities of the scanner must be expected, DWI estimation is recommended.

We present in the following sections 2.2-2.4 the technical tools of the method and give some explanations concerning our human brain gold standard model.

2.2 The Spatial Filter

2.2.1 General Description

We adapt a filter chain designed for edge preserving smoothing of noisy images with low SNR proposed by Aurich [23, 24]. The basic version iterates nonlinear filters which calculate weighted averages of noisy scalar images $f(\mathbf{x})$ on a grid in space, $\mathbf{x} \in \mathbf{R}^3$, within regions or scales defined by spatial and functional Gaussian windows, Φ and Ψ . The window parameters μ , η control the properties of the filter, see Eq. [1]. Aurich's version is designed for images which have, apart from the edges, small local variability corrupted by a constant noise level. For $f(\mathbf{x}) = \text{DWI}(\mathbf{x})$ sharp edges are also apparent, but additionally we find connected regions with large spatial variability and with high curvature. Furthermore, the noise level is not constant for low SNR. The original version of the filter chain is therefore adapted to DWI smoothing.

The first iteration of the filter $F(\eta, \mu, \cdot)$ is defined by

$$F(\eta_{(1)}, \mu_{(1)}, f(\mathbf{x})) = \frac{\sum_{\mathbf{y} \in \text{Neighborhood of } \mathbf{x}} \Phi \Psi f(\mathbf{y})}{\sum_{\mathbf{y} \in \text{Neighborhood of } \mathbf{x}} \Phi \Psi} \quad (1)$$

with $\Phi = e^{-(\mathbf{x}-\mathbf{y})^2/2\eta_{(1)}^2}$ and $\Psi = e^{-(f(\mathbf{x})-f(\mathbf{y}))^2/2\mu_{(1)}^2}$

For $(k-1)$ iterations we define

$$f_{k-1}(\mathbf{x}) = F(\eta_{(k-1)}, \mu_{(k-1)}, \dots, F(\eta_{(2)}, \mu_{(2)}, F(\eta_{(1)}, \mu_{(1)}, f(\mathbf{x}))))$$

The last, k -th, iteration is defined by

$$f_{\text{smooth}}(\mathbf{x}) = F(\eta_{(k)}, \mu_{(k)}, f(\mathbf{x})), \text{ where } \Phi = e^{-(\mathbf{x}-\mathbf{y})^2/2\eta_{(k)}^2} \text{ and } \Psi = e^{-(f_{k-1}(\mathbf{x})-f_{k-1}(\mathbf{y}))^2/2\mu_{(k)}^2}.$$

The functional window $\Psi = \Psi(\mathbf{x}, \mathbf{y})$ reduces the weights of DWI(\mathbf{y}) in Eq. [1] for large distance from DWI(\mathbf{x}) and can produce very irregular volumes of voxels for averaging with weights well above zero. The modification in the last iteration – note that $f(\mathbf{x})$ is averaged but not $f_{k-1}(\mathbf{x})$ – reduces a tendency of the filter chain to underestimate curvature or to “flatten out” the signal [24], this filter caused bias can be further reduced by a convenient choice of k , see section 3.2.

2.2.2 The Parameters

The parameters for the spatial windows are given by $\eta(i) = \delta c^{i+1}$, where $\delta = \text{step}/(1.9\sqrt{\pi})$ and $c = \sqrt[3]{4}$, and $\text{step} = \text{minimum}(\text{step}_x, \text{step}_y, \text{step}_z)$, where $\text{step}_x, \text{step}_y, \text{step}_z$ are the side lengths of a voxel [25]. Anisotropic voxels are admissible but an isotropic grid achieves better results as it increases the sample size within the windows.

To introduce a spatially varying Rician noise for low SNR the filter chain is applied in a novel way. Assuming first a constant noise level σ_0 in the DWIs the parameters $\{\mu_{(1)}, \mu_{(2)}\} = \{3\sigma_0, \sigma_0\}$ and $\mu_{(i)} = \sigma_0/2^{i-2}$ for $i > 2$ are applied for a first estimation of $\mathbf{E}[|S(\mathbf{x})|]$. For σ_0 the asymptotic Rician noise is used, it can be derived e.g. from the standard deviation σ_R of the measured background noise by $\sigma_0 = \sigma_R/\sqrt{2 - \pi/2}$ [19]. The Rician standard deviation $\sigma(\mathbf{x})$ can now be approximated numerically via $\mathbf{E}[|S(\mathbf{x})|]$, see [19]. In the parameters of Ψ σ_0 is replaced by the approximate $\sigma(\text{SNR}(\mathbf{x}))$ for the second application of the chain to the measured DWIs. For Rician noise we have $\sigma(\text{SNR}) \rightarrow 0.655 \sigma_0$ with $\text{SNR} \rightarrow 0$, therefore the second filter application can improve the resolution at low SNRs.

For postfiltering or final reduction of noise in “denoised” images of $FA(\mathbf{x})$ and $MD(\mathbf{x})$, also a heteroscedastic $\sigma_o(\mathbf{x}) = qQ * F(\mathbf{x})$ is used. See **Appendix 2** for the explicit formula of $F(\mathbf{x})$ and for its justification. A numerical IDL-program of the filter algorithm within an environment for robust σ_R estimation and for data processing can be delivered from the authors, an implementation is in use in the Medical School of Houston³.

2.3 The Bias Correction

The expectation of the DWIs $\mathbf{E}[|S_j(\mathbf{x})|]$ become strongly biased for low SNR. Rician statistic offers a bias correction BC with $BC(\mathbf{E}[|S_j(\mathbf{x})|]/\sigma_0) = \text{SNR}_j(\mathbf{x})$. This transformation BC can be applied in an approximate way to an empirical mean $|\overline{S_j}|_{NEX}/\sigma_0 = \sum_{k=1}^{NEX} |S_{j,k}|/(NEX * \sigma_0)$ estimated by filtering or voxelwise averaging, where k counts the sample for averaging and j the diffusion gradient. In Fig. 2 panel (A) the bias correction BC (solid line) and a modification for $\text{SNR} < 0.5$ (broken line) are presented. In panels (B-C) for $\text{SNR} = 1$ and $NEX = 10, 30$ simulated densities of $|\overline{S_j}|_{NEX}/\sigma_0$ (MNR, green) and of $BC(\text{MNR})$ with modification (red) and without modification (blue) are shown. Close to the noise floor, $\text{MNR} \sqrt{\pi/2}$ or $\text{SNR} = 0$, we are confronted with three problems. First, due to statistical uncertainty some MNRs may fall below the range of BC , see the blue-shaded areas within the green MNR densities. In this case a Rician bias correction is not defined. If we define $BC(\text{MNR}) = 0$ for those MNRs, then the transformed density becomes singular, see the blue peaks near zero. Second, the slope of Rician BC increases for low SNR considerably (see panel A), enlarging the variance in the densities of $BC(\text{MNR})$. For an illustration, compare the blue and green densities in panels (B-C) and see section 2.1. Third, at the noise floor the derived diffusion can become singular due to the Stejskal-Tanner equation, see for the eigenvalues in Eq. A1.1 of **Appendix 1**. As possible solution, we propose a simple modification of the Rician BC by a smooth C^1 -extensions for $\text{SNR} < 0.5$. A convenient modification is defined by $BC(\mathbf{E}[|S_j(\mathbf{x})|]/\sigma_0) = (\mathbf{E}[|S_j(\mathbf{x})|]/1.44 \sigma_0)^{8.76}$, (see broken line in panel A). This modification reduces the variance, see the red densities in panels (B-C), regularizing the noise floor of the (blue) densities. See section 3.1 for quantitative implications of this modification.

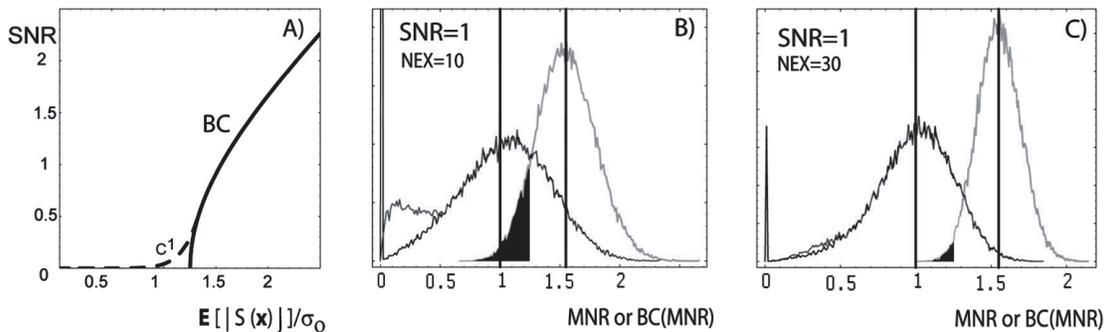


Figure 2: Panel (A), Graph of the Rician Bias Correction BC (Solid Line) and of a Modification for $\text{SNR} < 0.5$ (Broken Line). Panels (B) and (C) Show Densities of Mean to Noise Ratios MNR (Green), and of $BC(\text{MNR})$ without (Blue) and with C^1 Modification (Red). Blue Shading Indicates MNR Below the Noise Floor, the vertical bar at 1 indicates the SNR, the higher bar at 1.55 the $\mathbf{E}[\text{MNR}]$.

2.4 The Data and the Gold Standard Model

To produce a convenient gold standard model for validation a small voxel experiment was performed on a consented healthy volunteer (40 year old female). The isotropic $1 \times 1 \times 1 \text{ mm}^3$ diffusion weighted data were acquired on a General Electric 1.5 T MRT scanner using a dual spin echo prepared diffusion sequence with ramp sampling and fat suppression. The b -value is $b = 1400 \text{ s/mm}^2$, 21 icosahedral diffusion gradients were

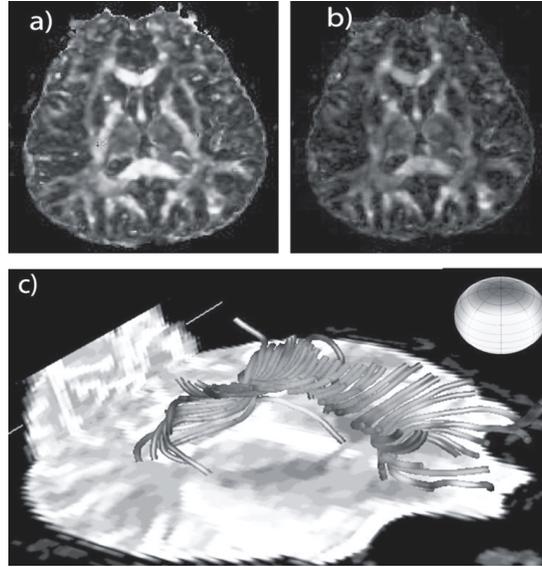


Figure 3: One Axial FA Map and FA-Weighted Directional Map of the Gold Standard Model in Panels (a-b); for Panel (c) Tracks of the Corpus Ccallosum are Calculated, Also the Colour Sphere for (b) is Shown, Mapping Direction and Colour

applied, a small volume of $256 \times 256 \times 28$ voxels was measured and a noise level $\sigma_0 = 130$ was determined from the background, see [17] for more details. The number of replications was $NEX = 4$. The voxelwise averaged DWIs were denoised. Due to the very low averaged $SNR \approx 2.2$, the method could not reduce the random effects sufficiently, also several spikes in the DWIs and in the reference due to field instabilities were observed. Therefore, to achieve a sufficiently smooth gold standard model, the denoised DWIs were further postprocessed by an edge preserving $3 \times 3 \times 3$ Median filter. See Fig. 3 for illustrations of this model. Note, that current DTI protocols for clinical and neurobiological applications use voxel sizes of 8-27 mm^3 or larger for the human brain [4] and may differ therefore in some details of fiber properties measured with 1 mm^3 resolution. Remaining spikes close to the cerebral spinal fluid (CSF) were removed from the model by a mask over the central CSF region, see Fig. 5, 8-11 for that mask.

3. MONTE CARLO EXPLORATIONS

In section 3.1 the mean squared error of DWIs is compared for denoising with different BCs. In section 3.2 the length of the filter chain is adapted and the inclusion of heteroscedastic noise into filtering is tested. In section 3.3 the denoising method is tested on a phantom of bifurcating nerve fibers. In section 3.4 the denoising method is explored on the gold standard model.

3.1 Comparison of Different Bias Corrections

To explore the effect of different BCs on the DWIs a Monte Carlo study is summarized in Fig. 4, comparing four denoising schemes ($M1$, $M2$, $M3$, ML). The simulations are performed within the Rician noise model. Method $M1$ is defined as estimation of $|S_{\text{without noise}}|$ by $|\bar{S}|_{NEX}$, the gradient index j is omitted. In $M2$ the estimator $M1$ is modified by additional application of the Rician BC including the projection to zero, see section 2.3. In $M3$ the Rician BC with the C^1 modification is applied. In ML a Maximum Likelihood estimator is involved, where the likelihood function is based on the density of the Rician noise model including implicitly a BC with zero projection [26]. The relative rooted mean squared error

$$rrmse = \sqrt{r \text{bias}^2 + \text{Variance}[\text{estimator}] / |S_{\text{without noise}}|^2}$$

with the relative bias $r \text{bias} = (\mathbf{E}[\text{estimator}] - |S_{\text{without noise}}|) / |S_{\text{without noise}}|$ is calculated by simulations. The sample sizes are $NEX = 10, 30, 50$ (red, blue, green), five different $SNRs = 0.25, 0.5, 1, 2, 3$ are applied.

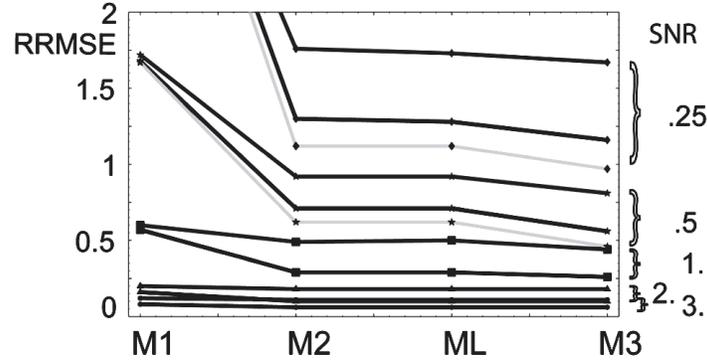


Figure 4: Simulated Relative Rooted Mean Squared Errors for $M1$ - $M3$, ML . One Solid Line Corresponds to One SNR, Colours to the Number of Replications $NEX = 10$ (Red), 30 (Blue), 50 (Green), Groups with the Same SNR are Collected by Brackets, for $SNR > 0.5$ Blue and Green Curves Coincide

The high values of $rmse$ in $M1$ are reduced by the BC included in $M2$, $M3$ and ML . For all fixed SNR and NEX we find the smallest $rmse$ for $M3$, the somewhat higher values for $M2$ and ML are close together. Therefore, in all further calculations the BC of $M3$ will be used. The $rmse$ of $M3$ for $SNR = 0.5$ and $NEX = 30$ (Blue) is still 56%, enlarging NEX to $NEX = 50$ (Green) produces only a small improvement. Together with the $rmse$'s for $SNR = 0.25$ this indicates that $SNR = 0.5$ may be at the lower limit for the bias correction of $M3$ with $NEX \approx 30$, though a quantitative answer depends on the implications for the fiber properties under consideration, see sections 3.3 and 3.4.

3.2 Adaption of the Filter Chain and Inclusion of Varying Rician Noise

For low SNR, practical restrictions of NEX and of the filter lead to residual randomness after denoising. Consequently, validations should be performed on ensembles of model data, so that the error analysis can be based on $rmse$'s for individual voxels. The ensemble size N_e , as a compromise between statistical stability and practicability, is put in the following to $N_e = 100$.

The window sizes $\eta_{(i)}$ and $\mu_{(i)}$, $i = 1 \dots k$, are defined in section 2.2.2. Their optimal functional dependence on the voxel size on the noise level was derived analytically for i.i.d noise by Mühlhaus [25]. These window sizes are tested in many applications with different noise, see [24] and showed a good robustness. They were also applied successfully to DWI filtering with medium SNR [27, 28] and therefore were not changed.

Less attention is paid to the influence of the length k of the filter chain. Chains with $k = 2$ -5 iterations are applied to a DWI of the human brain model. See Fig. 5 panel (a) for a slice of \mathbf{E} [DWI] covering parts of corpus callosum lying in regions with low expectation values. A mask around CSF, mentioned in section 2.4, is included in black colour. Rician noise with the noise level $\sigma_0 = 65$ was added to the DWIs, see for computational details **Appendix 1**. By this noise level a field strength of $3T$ is modelled. Else, the b -value $b = 1400 \text{ s/mm}^2$ and 21 icosahedral diffusion gradients are applied. In panels (b-c) lineplots for the yellow line of panel (a) are shown. The panels show \mathbf{E} [DWI]/ σ_0 of the model (black), the corresponding ensemble means of filtered noisy models (green: $k = 2$, orange: $k = 3$, red: $k = 4$, blue: $k = 5$) and one noisy model (oscillating black line). At the bottom of the panels b-c) $2 * rmse / \sigma_0$ in the same colouring is shown. Transition from $NEX = 1$ to $NEX = 4$ improves the results. At the extrema of \mathbf{E} [DWI]/ σ_0 filtering with higher k tends to “flatten out”, this is more apparent for $NEX = 1$. The Box Whisker Plots in panels (d-e) give the distributions of $rmse / \sigma_0$ for the slice including now all 21 icosahedral diffusion gradients, colouring like in (b-c). For $NEX = 1$ (d) the chain with $k = 2$ has, due to large residual variance, a higher median than the chain with $k = 3$. For $k = 4$ -5 the distributions show enhancement of $rmse$ due to more “flattening”. Panel (e) shows the same tendencies, but the differences are smaller. These results and the practical aspect to minimize computer time favour a chain with length $k = 3$ for applications to the high resolution human brain model and this be used from now on.

In panel (f) of Fig. 5 the inclusion of spatially varying Rician noise in DWI filtering is tested for $NEX = 4$ on the FA map. The Box Whisker Plots show the $rmse$ of FA for tissue voxels of slice (a). The distributions are restricted to $FA > 0.6$, as mainly high FA is correlated with low SNR. The yellow Box presents filtering with spatially varying noise, the cyan Box for constant noise. The inclusion of heteroscedastic noise $\sigma(\mathbf{x})$ induces an improvement for higher $rmse$ and some deterioration for $rmse < 0.03$. As high residual noise is more critical for a quantitative analysis of FA maps, we conclude that inclusion of $\sigma(\mathbf{x})$ into the DWI filtering process makes sense.

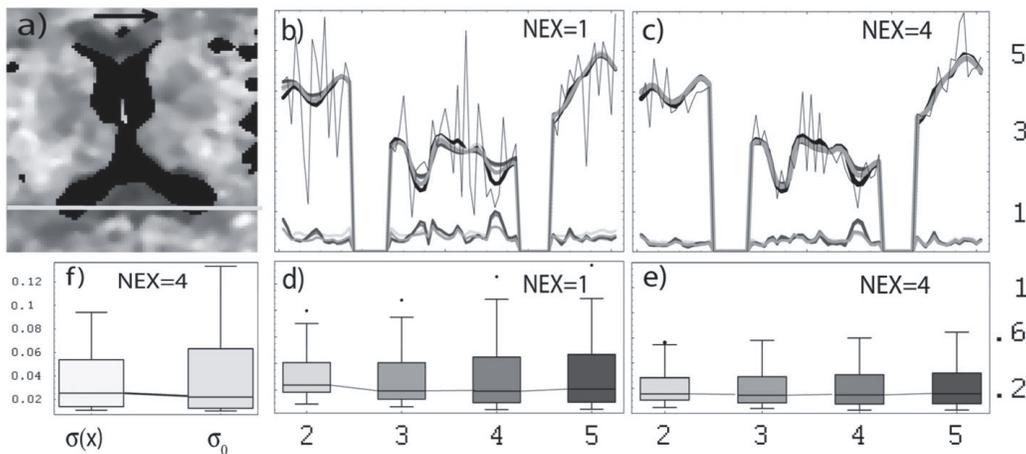


Figure 5: Adaption of the Filter. Panel (a) Shows an Axial Slice of $E [DWI]/\sigma_0$ for One Gradient (Arrow), Darker Grey Indicates Lower MNR, A Mask Around the Cerebrospinal Fluid is Included in Black. Panels (b-c) Show Profiles Along the Yellow Line in (a). Black: Model, Black Oscillating: One Noisy Model, Green - Orange - Red - Blue : Means of Filtered Models with Chain Lengths $k = 2 - 3 - 4 - 5$. In Same Colouring at Bottom of Panels, $2*rmse/\sigma_0$. Panels (d-e), Box Whisker Plots (Quantile = 0.45, In all Figures) of $rmse/\sigma_0$ versus k for Slice (a), All 21 Gradients Included. Panel (f) Shows Box Whisker Plots for $rmse$ of FA in Slice (a) After Denoising with $NEX = 4$ and with $FA_Model > 0.6$. Yellow Box Indicates Inclusion of Rician Noise into the Filter, Cyan Box Refers to Filtering with Constant Noise

3.3 Test of the Method on a Phantom

The performance of the optimized filter chain is demonstrated first on an artificial phantom containing a bifurcating tube of nerve fibers within a volume of fibers. The fibers outside the tube are orthogonal to those inside, see Fig. 6 panel (a) for a schematic representation. White matter diffusion with $\lambda_2 = \lambda_3$ is modelled, inside the tube for $FA = 0.82$, $MD = .00097 \text{ mm}^2/\text{s}$ and with $SNR > 0.5$ in the DWIs, outside for $FA = 0.56$, $MD = .00086 \text{ mm}^2/\text{s}$ and with $SNR > 1.5$. The parameters field strength, b -value and number of gradients are like in section 3.2.

In panel (b) the SNR distribution is presented, 50% of the DWI voxels are affected by biased Rician noise. For the coloured rectangle in panel (a) the main diffusion directions are shown for every voxel panel (c), colouring indicates different FA . In panel (d) the noisy ensemble ($NEX = 1$) is presented, mixed colouring in one cone indicates FA -variability due to noise. For panel (e) denoising is performed with $NEX = 1$, filtering and application of BC . For panel (f), $NEX = 4$ replications are averaged in the denoising procedure. Already in (e) every voxel is correctly classified, the cones of uncertainty are reduced, at the edges the reduction is somewhat smaller due to smaller regions of homogeneity for filter. In panel (f) even at the edges small variance is achieved.

In Fig. 7 the FA map of the same phantom slice is analysed. In panel (a) the noise free FA map, in panel (b) one noisy map of the ensemble ($NEX = 1$) are shown. Panel (c) gives the ensemble mean after filtering ($NEX = 1$), BC application and FA postfiltering. The model of the noise level for postfiltering is explained in **Appendix 2**, the parameter $q_{FA} = 12$ is applied for $NEX = 1$. Note that in panel (c) the individual

edges are exactly reproduced and that the mean is essentially unbiased. Strongly biased individual FA values are the result of DWI filtering alone, see (d) for the $rmse$ and (h) for a profile plot along the black line in (a). If BC is applied to the filtered DWIs the $rmse$ is reduced, see (e) for the $rmse$ and (i) for the profile.

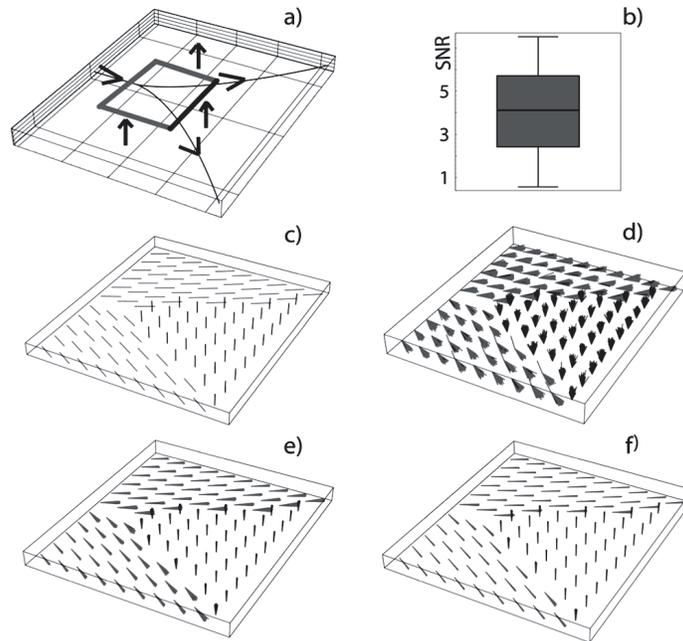


Figure 6: Filter Performance at Edges for a Bifurcating Tube of Fibers in a Phantom. Panel (a) Shows Schematically the Phantom Volume, Arrows Indicate Directions of Main Diffusion, Coloured Rectangle Indicates the Slice Shown in (c-f). Panel (b), Box Whisker Plot of SNR. Panel (c), Main Diffusion Directions of the Model without Noise, Red: $FA = 0.82$, Blue $FA = 0.56$. Panel (d), Noisy Ensemble; Panel (e), Denoised Ensemble (NEX = 1, Filtering, BC); Panel (f), Denoised Ensemble (NEX = 4, Filtering, BC).

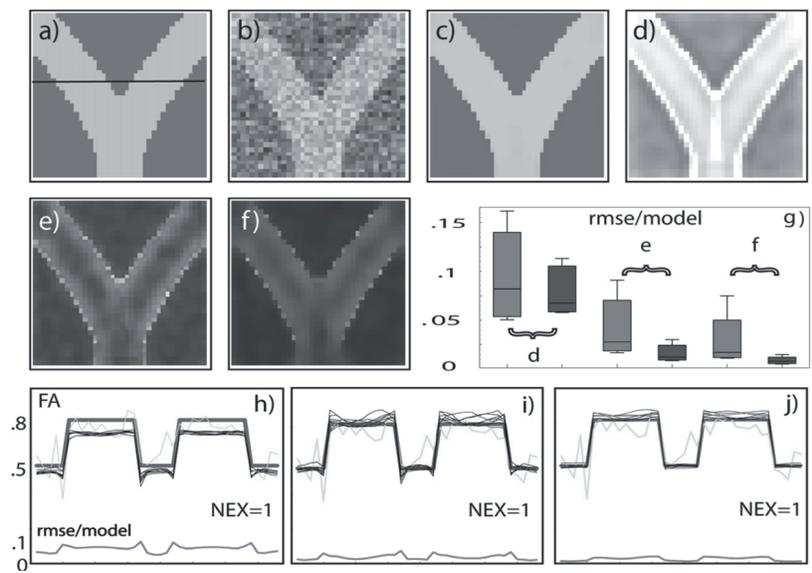


Figure 7: Effect of Denoising on FA Maps. Panel (a), FA Map of the Phantom, Line Indicates Position of the Profiles (h-j); Panel (b), One Noisy Model; Panel (c), Ensemble Mean of FA Map After Denoising (Filtering, BC , Postfiltering). Panels (d-f), $rmse$ of FA Maps; (d), Filtered; (e), Filtered and BC ; (f), Like (e) and Postfiltering. Panels (a-f) are Based on NEX = 1. Panel (g), Box Whisker Plots of the $rmse$ of FA . Red Boxes Correspond to Panels (d-f), Blue Like Red Ones but with NEX = 4. Panels (h-j) Show FA Profiles, Blue : Model, Green : Noise, Black : 10 Random Elements of the Denoised Ensemble, Red : $rmse$. Panel (h), Filtering; Panel (i), Filtering and BC ; Panel (j), Filtering and BC and Postfiltering

Note that the variance in the denoised FA ensemble increases from (h) to (i) Mainly for Low SNR, see **Appendix 2** for the error propagation of FA . If finally postfiltering is applied to the denoised FA images, an improvement also at the edges is achieved, see panels (f) and (j). In panel (g) Box Whisker plots of the $rmse$ of panels (d-f) are given in red for $NEX = 1$, in blue for $NEX = 4$ (postfiltered with $q_{Fa} = 6$). After postfiltering, the “blue” $rmse$ of (f) in panel (g) is below 1.5% at all voxels of the slice, this corresponds to a $SNR = 1/rmse$ of FA above 67.

3.4 Validation of the Method on the Human Brain Model

Finally, the method is validated on the human brain model described in section 2.4 for the diffusion properties MD , FA and main diffusion direction $|i(\mathbf{x})\rangle$, which are frequently used biomarkers [29]. Experiments with field strength $3T$, $b = 1400$ s/mm², 21 gradients and $NEX = 1, 4$ and 8 are simulated, the noise level of section 3.2 is applied.

A volume around the corpus callosum is used as gold standard for denoising. For quantitative error analysis an axial slice in the middle of the volume is chosen, it covers 80×110 voxels. In Fig. 8 results for MD are presented. To guarantee a reliable error analysis, the region around CSF is faded out by the mentioned mask to remove measurement artefacts, see section 2.4. Panel (a) shows the model MD , panel (b) shows MD affected with noise for $NEX = 1$, and panel (c) shows the ensemble mean after full denoising (filtering, BC , postfiltering) with $NEX=1$. A good agreement between (a) and (c) is apparent, indicating an essentially unbiased averaged image of MD after denoising. In panel (d) the minimum SNR of the DWIs is presented,

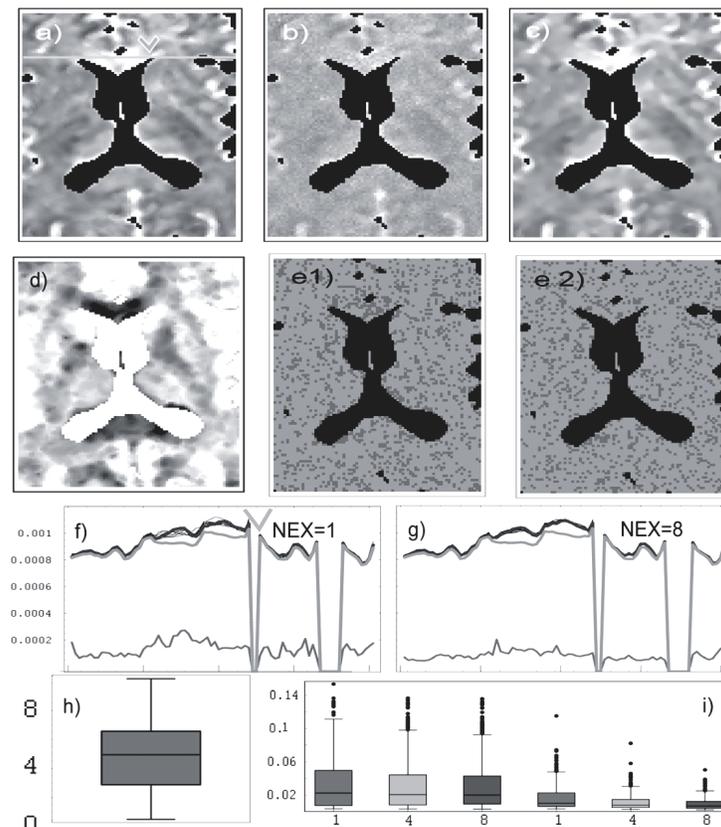


Figure 8: Denoising for the MD Map. Panel (a) The Model, Panel (b) Noisy MD ($NEX = 1$), Panel (c) Mean of Fully Denoised Ensemble ($NEX = 1$). Panel (d) Indicates Minimum SNR per Voxel. Panels (e1-e2) Present P-Values of Shapiro-Wilk Tests for Normality of MD Distributions. Red Voxels $P < .1$, Green Voxels $P > .1$. Panels (f-g) Show Profiles of MD (mm²/s), Blue: Model; Black: 10 Random Elements of Fully Denoised Ensemble, Green: $E[MD]$; Red: $rmse/100$. Panel (h), Box Whisker Plot of SNR for the Slice (a). Panel (i) Box Whisker Plots of $rmse$ of MD Before BC (Red, $NEX = 1$; Orange, $NEX = 4$; Blue, $NEX = 8$) and After Full Denoising.

the lowest SNRs are mainly clustered in the two regions of corpus callosum ($\text{SNR} \sim 0.5$). Panels (e1) and (e2) show significance levels (uncorrected) of Shapiro-Wilk Tests for Normality of the MD distributions in every tissue voxel. Red voxels indicate the significance levels $P < 0.1$, or a significant difference from Gaussian distributions, green voxels have $P \geq 0.1$ and are compatible with the Gaussian assumption [30]. Panel (e1) is calculated for $\text{NEX} = 1$ followed by full denoising. The red voxels are slightly clustered in low SNR regions, the ratio of the number of red to all tissue voxels in the slice is 0.18. Panel (e2) differs from panel (e1) by $\text{NEX} = 8$. The clusters have now essentially disappeared, the ratio reduces to 0.16. The fact that the number of Gaussian distributions is growing with NEX is compatible with the DM , the speed of convergence to the Gaussian limit seems to be low, see [31] for examples of such distributions. In panels (f-g) profiles of MD are shown for $\text{NEX} = 1$ and 8, the profiles are taken from the yellow line with a position indicating arrow shown in panel (a). Blue lines in panels (f-g) indicate the model, green lines the biased mean of the ensemble after filtering, black lines are 10 randomly chosen members of the ensemble after full denoising and red lines show $\text{rmse}/100$. Note the improvement of rmse from $\text{NEX} = 1$ to $\text{NEX} = 8$, which is again in line with the DM . In panel (h) the SNR distribution of the slice is presented. In panel (i) the Box Whisker Plots of the rmse of MD are given for full denoising with $\text{NEX} = 1, 4$ and 8. The left group (red, orange, blue) is calculated before application of BC , the right group after BC and postfiltering. Summarizing, we find after denoising with $\text{NEX} = 1$ that the SNR of MD is already above 50 for nearly

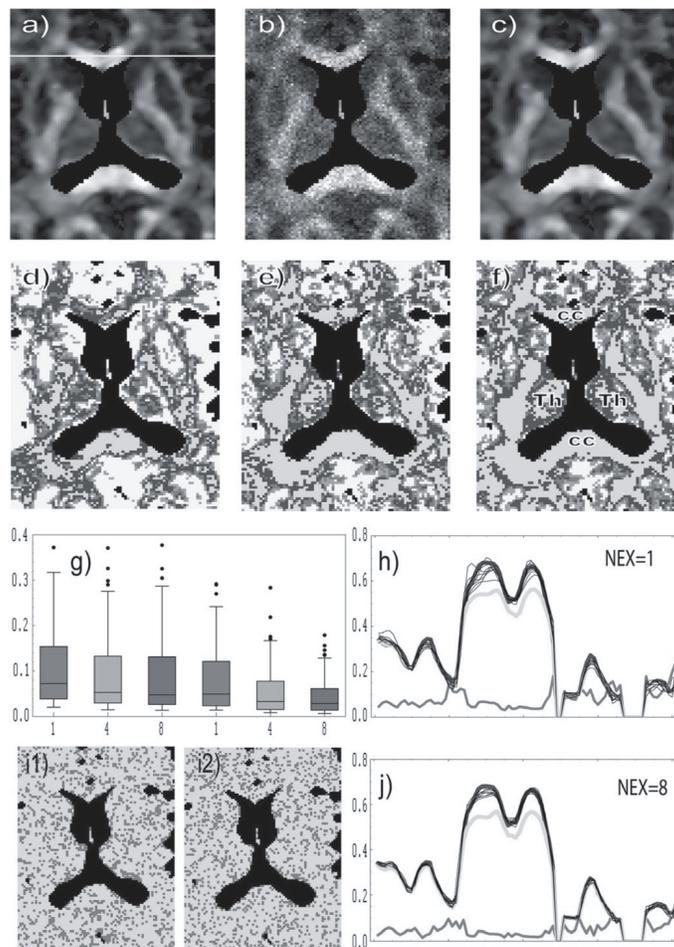


Figure 9: Denoising of the FA Map. The Model is Shown in Panel (a), One Noisy Model in Panel (b), The Mean of the Fully Denoised Ensemble ($\text{NEX} = 4$) in Panel (c). Panels (d-f), Local rmse for FA After Full Denoising with $\text{NEX} = 1$ (d), $\text{NEX} = 4$ (e) and $\text{NEX} = 8$ (f). Colouring: Green: $<4\%$, Bblue: $[4\%, 6\%)$, Red: $[6\%, 8\%)$, Orange: $[8\%, 10\%)$, Yellow: $>10\%$. In panel (f) The Corpus Callosum and the Thalamus are Indicated. Panel (g), Box Whisker Plots of rmse for $FA > 0.3$, Else Like in Fig. 8. Panels (h) and (j) Show Lineplots for yellow line in the FA map of Fig. 9 a), colouring like in Fig. 8, the rmse (red) is unscaled. Panels (i1-i2) indicate Shapiro-Wilk significance levels for FA distributions, colouring like in Fig. 8.

95% of the voxels, the SNRs further increase with NEX. Few outliers may be caused by remaining artefacts in the experimentally based gold standard model. Negative eigenvalues, which corrupt the interpretation of MD are detected before filtering in tissue voxels. For noisy images with $NEX = 1$ and application of BC approximately 200 voxels with negative eigenvalues are found, for $NEX = 4$ or 8 and BC 10 or 3 voxels. Already for denoising based on $NEX = 1$ all negative eigenvalues are removed.

In Fig. 9 the same slice is analyzed for FA . Panel (a) of Fig. 9 shows the model, panel (b) one noisy FA map and panel (c) the ensemble mean after full denoising, both with $NEX = 4$. Again we find an essentially unbiased average, compare (a) and (c). Panels (d-f) show color coded maps of the $rmse$ for $NEX = 1$ (d), 4(e) and 8(f) after full denoising. The error situation is less straightforward than in case of MD . $NEX = 1$ gives for white matter only partly a $rmse$ below 4% (green). This improves with $NEX = 4$. For $NEX = 8$ also parts of deep grey matter are well denoised, see for the Thalamus indicated in panel (f). In panel (g) Box Whisker Plots for $rmse$ of FA after denoising with $NEX = 1, 4$ and 8 (red, orange and blue) are given for $FA > 0.3$. The left group shows the distributions before application of BC , the right group after full denoising.

We find for the right group that 50% of the voxels have $20 < SNR < 100$ for $NEX = 1$, $33 < SNR < 100$ for $NEX = 4$ and $40 < SNR < 150$ for $NEX = 8$. In panels (h) and (j) profiles of FA , positioned and coloured like in Fig. 8 are given. A reduction of the $rmse$ (red) with increasing NEX from 1 to 8 is apparent, also a considerable bias (green). In panels (i1) and (i2), like for MD , significance levels of the Shapiro-Wilk Test for Normality for the FA distributions are presented after full denoising with $NEX = 1$ (i1) and $NEX = 8$ (i2). In (i1) 20% of the distributions differ significantly from Normality, in (i2) 19%.

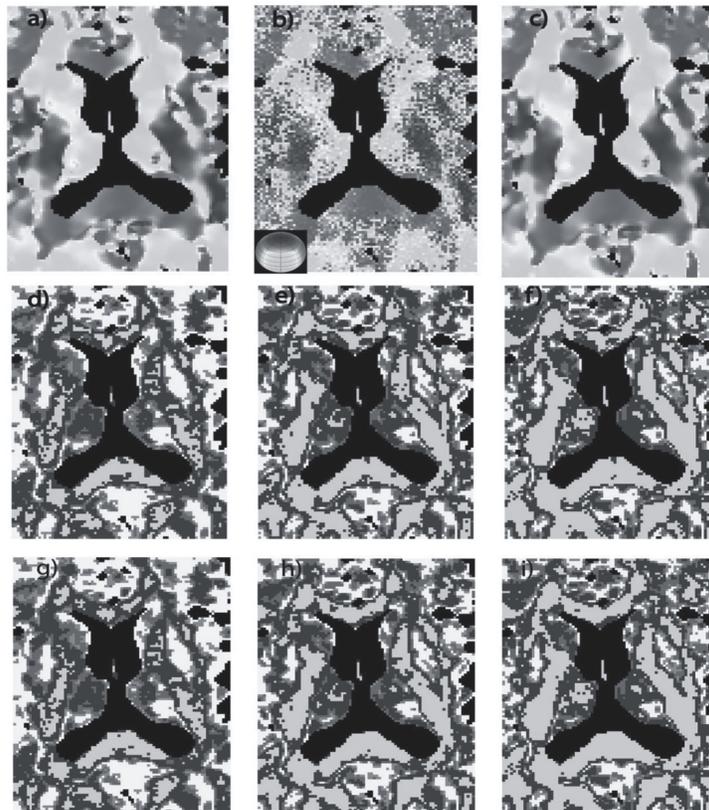


Figure 10: Denoising of Main Diffusion Directions. Panel (a) Color Coded Directional Map of the Model. Panel (b) Noisy Directions ($NEX = 1$), at Left Lower Corner the Colour Sphere is Shown. Panel (c) Ensemble Averaged Directions After Denoising ($NEX = 8$). Panels (d-f) Show Mean Angular Deviations from the Model in Degrees, Filtering Based on $NEX = 1$ (d), $NEX = 4$ (e) and $NEX = 8$ (f); Green : $<2^\circ$, Blue : $[2^\circ, 4^\circ)$, Red : $[4^\circ, 6^\circ)$, Orange : $[6^\circ, 8^\circ)$, Yellow : $>8^\circ$. Panels (g-i) Like (d-f) But with BC .

In Fig. 10 the effect of denoising on the main diffusion directions is analyzed. Upper row gives color coded directional maps of the model (a), of one noisy slice (b) and of the ensemble averaged directions after full denoising (NEX = 8) without postfiltering (c). Postfiltering was not applied as a convenient filter for vector fields was not available. Note in (a) and (c) the discontinuous transitions of the colors, indicating clearly discontinuous changes of main diffusion directions. Again, the remaining bias is very small, see (a) and (c). For quantitative error analysis, the ensemble averaged angle between model direction and denoised model direction is presented before application of BC in the second row, and after application of BC in the last row of Fig. 10. Panels (d, e, f) and (g, h, i) are calculated for NEX = (1, 4, 8). In both rows the error is reduced with increasing NEX. The error maps resemble those of FA , for NEX = 4 and 8 white matter and parts of the Thalamus have a mean angular deviation below 2° . Application of BC improves the error map in high FA regions of corpus callosum, in those regions the SNR is lowest.

4. SUMMARY AND DISCUSSION

If high numbers NEX of replications would be experimentally available, the thermal noise problem in low SNR DTI would not exist. The expectations $\mathbf{E}[|S_j(\mathbf{x})|]$ could be estimated reliably and after application of BC the diffusion properties would have unbiased Gaussian distributions with small variance, as is proven in the DM . Due to practical scan time limitations, mainly caused by patient motion, spatial filtering is needed to fill this gap at least partially. Residual noise in denoised DWIs after filtering necessitates a modification of the Rician BC at the noise floor. The proposed regularization produces a lower mse than the standard Rician BC and a Maximum Likelihood estimator down to $\text{SNR} \approx 0.5$. For filtering, we adapt a chain of nonlinear Gaussian filters to peculiarities of DWI smoothing for data with small voxels. The averaged filter results agree generally well with $\mathbf{E}[|S_j(\mathbf{x})|]$, critical are regions with high curvature. This behaviour improves by a transition from NEX = 1 to higher NEX for voxelwise averaging before filtering. The filter performance for scalar and vectorial diffusion properties is then tested on a phantom. The filter shows artefact free edge detection for orthogonal and bifurcating main diffusion directions. Residual noise effects in FA and MD maps can be reduced by application of postfiltering, which also improves the separation

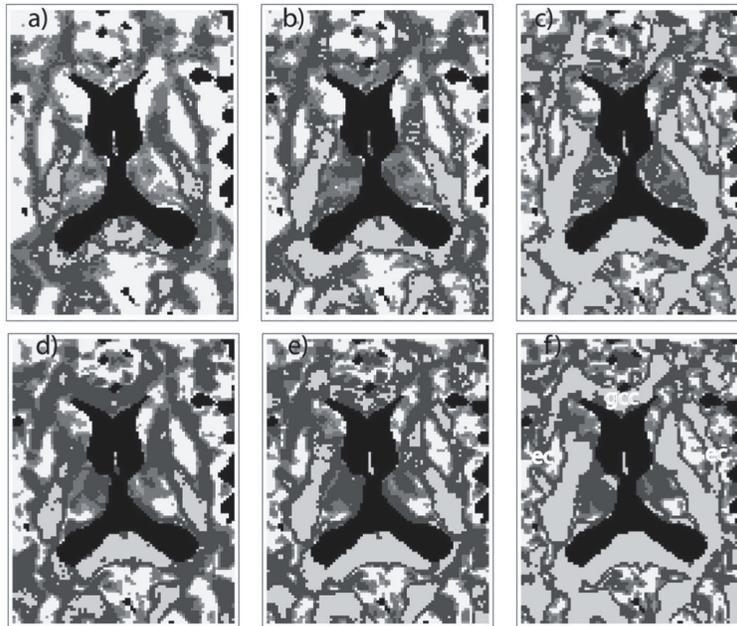


Figure 11: Denoising via Voxelwise Averaging and Application of BC . Upper Row, $rrmse$ of FA . Panel (a) NEX = 20, (b) NEX = 30, (c) NEX = 50. Colouring Like in Fig. 9. Lower Row Mean Angular Deviation from Model Diffusion. Panel (d) NEX = 20, (e) NEX = 30, (f) NEX = 50. Colouring Like in Fig. 10. In panel (f) The Genu of Corpus Callosum (gcc) and the External Capsule (ec) are Indicated

at the edges. For postfiltering a model of *FA* and *MD* noise levels after denoising of the DWIs is introduced on the basis of the error propagation formalism. For application of the full denoising method only two parameters, the level of background noise for DWI filtering and of residual noise for postfiltering, must be deduced from the data. As is mentioned in the introduction, a list of competing filters exist, most of them are tested however only for high and moderate SNRs. A comparison of these proposals for low SNR DTI is beyond the scope of this work, but our study shows that such validations should be performed on ensembles of noisy models to properly include residual randomness, which could easily be misinterpreted as filter blurring. To quantify the effectiveness of our spatial filter, the error maps for simulated voxelwise DWI averaging with $NEX = 20, 30, 50$ (*BC* included) are presented in Fig. 11. Panels (a-c) give the *rrmse* for *FA* and panels (d-f) the mean angular deviations. The results are roughly similar to those of panels (d-f) in Fig. 9 and panels (g-i) of Fig. 10, indicating close equivalence to spatial filtering and in general a high effectivity of the filter. In special regions the filter acts differently due to its dependence on the size of the spatial homogeneous regions. So is the filter more effective in the compact genu of corpus callosum (*gcc* in panel (f) of Fig. 11) and mainly for *FA* less effective in the thin external capsule (*ec*).

The Monte Carlo exploration of the method on the human brain model for *MD*, *FA* and the main diffusion directions is the central result of this study. It is connected to the problem of quantitative analysis and to the question for convenient designs of small voxel experiments. Quantitative statistical analysis of DTI properties is usually performed in three categories. First, descriptive statistical procedures are used, which need in many cases unbiased mean values and small residual errors. Also, inferential testing is applied on distributions estimated from spatially different regions of interest (ROIs) for one measured data set. Finally, group comparisons are tested, where e.g. significant differences in mean values of fiber properties for groups of experiments are deduced in individual voxels, see [32] for a more detailed discussion. In our study, the denoising method was applied to a simulated experiment with field strength $3T$, $b = 1400 \text{ s/mm}^2$, 21 icosahedral gradients and voxel size $1 \times 1 \times 1 \text{ mm}^3$. For such experiments our denoising method achieves for *MD* already for $NEX = 1$ a $SNR = (1/rrmse) > 50$ for about 95% of the tissue voxels. This SNR increases with NEX , see Fig. 8. The assumption of Normality, preventing outliers in *MD* distributions, cannot be rejected for $NEX = 1$ for 82% and for $NEX = 8$ for 84% of the voxels (based on uncorrected tests). Negative eigenvalues are removed completely by denoising. This signals positive conditions for a descriptive quantitative analysis of *MD* already after denoising with $NEX = 1$. For inferential testing, based on different ROIs of one measured data set, the answer is in our opinion negative for all NEX . This is due to the fact that spatial filtering induces spatial correlations in the fiber properties and destroys the necessary independence within the random samples, which is a prerequisite for the application of standard testing procedures. For group comparisons however, this independence is conserved, as filtering is performed independently for every member of the group. As the Gaussian condition is not fulfilled everywhere, non parametric testing may be useful. For *FA* we find a less straightforward error situation. The SNR of *FA* is only for preaveraging with $NEX \geq 4$ above 25 in the main parts of white matter, for $NEX = 8$ also in parts of the deep grey matter of the Thalamus, see Fig. 9. For 50% of the voxels with $FA > 0.3$ the SNR is above 20 for denoising with $NEX = 1$, above 33 for denoising with $NEX = 4$ and above 40 for denoising with $NEX = 8$. The Gaussian shape cannot be rejected for approximately 80% of the voxels for all three cases. Again, descriptive analysis and inferential group comparisons should be feasible, if we restrict to white matter and to convenient deep grey matter for sufficiently high NEX . For the main diffusion directions the error maps are similar to *FA*, see Fig. 10. For $NEX = 8$ white matter and most of Thalamus have a mean angular deviation below 2° , which may allow application of simple tracking methods [29] for descriptive analysis and for inferential group comparisons.

The human brain data for the gold standard model with $1 \times 1 \times 1 \text{ mm}^3$ resolution and field strength $1.5T$ revealed scanner artefacts which could not be removed sufficiently by our denoising method alone, see section 2.4. A thorough analysis of the experimental problems with high resolution is beyond our scope.

Small voxel DTI for reduced volumes seems to be more stable, see [33] for an application of our method to the early postnatal development of rat brains. One important constraint for technical parameters like e.g. field strength, b-values and NEX to measure human brain DTI with small voxels, however, is the need for a reliable quantitative statistical analysis. Our study indicates that experiments with parameters similar to those of section 3.4 could be in this sense reasonable candidates.

5. CONCLUSION

A denoising method for low SNR DTI is presented. The method is parameterized and validated in a statistically solid way for a human small voxel model. The simulation results indicate the feasibility of standard DTI with appreciably smaller voxels than are in use at present, if thermal noise in the DWIs is the source of the dominating artifacts. The principles of the presented method go beyond the standard DTI model, it may be applied to any diffusion model based on DWIs with low SNR, like Multiple-tensor fitting or Q -ball imaging [29].

APPENDIX 1

BASICS OF THE DTI FORMALISM

Formal definitions of the Gaussian second-rank standard DTI model will be given to clarify some notations in the text. The positive definite and symmetric diffusion tensor $\mathbf{D}(\mathbf{x})$ comprises six real valued independent components $D_{ij}(\mathbf{x})$ for $(i, j) = (1, 1), (2, 2), (3, 3), (1, 2), (1, 3), (2, 3)$ at every voxel \mathbf{x} . The tensors can be decomposed into their eigenvalues $\lambda_i(\mathbf{x})$ and eigenvectors $|i(\mathbf{x})\rangle$, $i = 1, 2, 3$. Since the tensor is hermitian, a convenient bracket notation is introduced, where the ket $|\cdot\rangle$ is a column vector, the bra $\langle\cdot|$ a transposed ket, and $\langle\cdot|\cdot\rangle$ a scalar product. We can then write $\mathbf{D}(\mathbf{x}) = \sum_{i=1}^3 \lambda_i(\mathbf{x}) |i(\mathbf{x})\rangle \langle i(\mathbf{x})|$. One measured DWI $|S_j(\mathbf{x})|$ is linked to this model by the Stejskal Tanner equation

$$|S_j(\mathbf{x})| = |S_0(\mathbf{x})| \exp\left(-b \sum_{i=1}^3 \lambda_i(\mathbf{x}) \langle i(\mathbf{x}) | \mathbf{g}_j \rangle^2\right), \quad \text{for } j > 0, \quad [\text{A1.1}]$$

where b is the diffusion weighting b -value, $S_0(\mathbf{x})$ is the reference, and $|\mathbf{g}_j\rangle$ is a diffusion measuring gradient. For the high noise levels, applied in this paper, an unweighted least squares fit [34] is performed voxelwise to calculate for a sufficient number of gradients $j = 1 \dots N_g \geq 6$ the tensor via Eq. [A1.1], see e.g. [7, 20] for more details.

A basic parameter to quantify noise effects is the SNR. Different concepts of the SNR are in use. We restrict in most cases to $\text{SNR}_j(\mathbf{x}) = |S_j(\mathbf{x})|/\sigma_0$, $j \geq 0$, where the DWI or the reference is from Eq. A1.1 without noise and σ_0 is the standard deviation of the Rayleigh-corrected background noise [19].

Thermal noise effects in $|S_j(\mathbf{x})|$ are simulated via the complex signals $S_j(\mathbf{x})$, by $S_{j,\text{noisy}}(\mathbf{x}) = \text{Re}[S_j(\mathbf{x})] + \varepsilon_{Re} + i(\text{Im}[S_j(\mathbf{x})] + \varepsilon_{Im})$, where ε_{Re} , ε_{Im} are independent and normally distributed, $\varepsilon \sim \mathcal{N}(0, \sigma_0)$, and where $\text{Im}[S_j(\mathbf{x})] = 0$ is usually assumed in simulations based on a given tensor model. Rician bias effects in the mean values of DWIs are not reduced by voxelwise DWI averaging. Therefore, in the present paper $\text{SNR}_j(\mathbf{x})$ before averaging via NEX replications is used, in few exceptions the SNR concept is modified explicitly in the text.

Of central importance for neuroscience [29] are the scalar fiber properties mean diffusivity and the fractional anisotropy $MD(\mathbf{x}) = \sum_{i=1}^3 \lambda_i(\mathbf{x})/3$ and the fractional anisotropy $FA(\mathbf{x}) = \sqrt{3 \sum_{i=1}^3 (\lambda_i(\mathbf{x}) - MD(\mathbf{x}))^2 / 2 \sum_{i=1}^3 \lambda_i^2(\mathbf{x})}$. The main diffusion directions indicating the course of nerve fibers in the brain are given by $|i(\mathbf{x})\rangle$, they are coupled to the largest eigenvalue $\lambda_i(\mathbf{x})$.

APPENDIX 2

BASICS OF RICIAN STATISTICS

First we give some formal definitions and illustrations of Rician distributions, see [19] for additional information. Note in [19] a printing error in Table 1 concerning the standard deviations, see panel B of

Figure A2 for a correction. The Rician probability density for the variable $y = DWI_{\text{noisy}}(\mathbf{x})$ is determined by the parameter $A = DWI(\mathbf{x})$ without noise and by the standard deviation σ_0 . We have:

$$p(y) = (y/\sigma_0^2) \exp(-y^2 + A^2)/2\sigma_0^2) I_0(yA/\sigma_0^2) \quad [\text{A2.1}]$$

This implies for the expectation to noise ratio ENR

$$E[y]/\sigma_0 = \sqrt{\pi/8} \exp(-A^2/4\sigma_0^2) (A^2/\sigma_0^2 + 2) (I_0(A^2/4\sigma_0^2) + A^2/\sigma_0^2) I_1(A^2/4\sigma_0^2) \quad [\text{A2.2}]$$

where I_0, I_1 are the modified Bessel functions of the first kind of order 0 and 1, see [35].

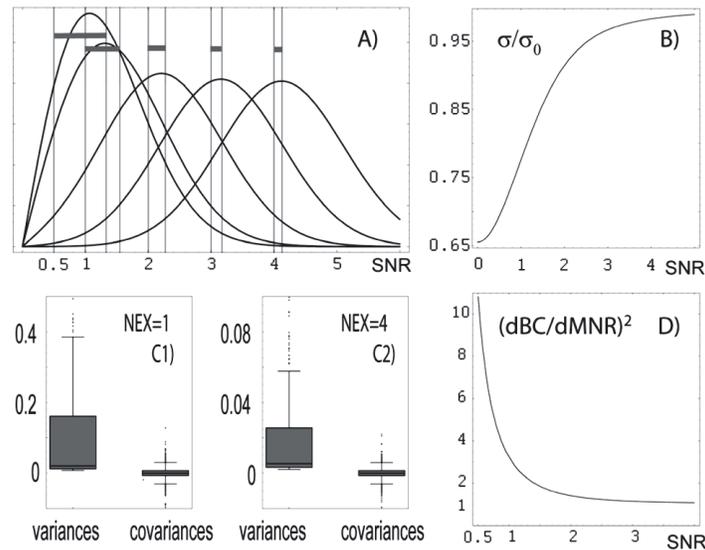


Figure A2: Panel (A), Rician Densities for SNR = 0.5, 1, 2, 3, 4. Pairs of Vertical Green Bars Connected by the Red Bar Indicate the SNR (Left) and the ENR (Right) per Density. Panel (B), Scaled Rician Standard Deviation versus SNR. Panels (C1, C2), Box-Whisker Plots (Quantile = 0.45) for Variances and Covariances of BC(MNR) for all Voxels and Diffusion Gradients in the Phantom, Depicted in Panel (c) of Fig. 6. Panel (D), Squared Derivative of BC versus SNR

In panel (A) of Figure A2 the densities for $\text{SNR} = A/\sigma_0 = 0.5, 1, 2, 3, 4$ are shown. The SNR and ENR are included by vertical green bars, the red horizontal bars connect SNR and ENR to quantify the mean value bias per density. In panel (B) the standard deviation of Rician noise divided by σ_0 is plotted versus SNR.

Error Propagation and a Model for Postfiltering

In this section we discuss the error propagation for FA and MD to motivate a model of the noise level for postfiltering. The following part uses notations from section 2.3.

For $Q(\mathbf{x}) = FA(\mathbf{x})$ or $MD(\mathbf{x})$ the variance of $Q(\mathbf{x})$ at fixed position \mathbf{x} can be approximated on the basis of a Taylor expansion truncated after the first derivatives, if the noise level is low [18]. In the following we assume therefore that the denoising procedure was applied, producing small residual noise in $\text{MNR}_j, BC(\text{MNR}_j)$

and also in Q . The variable Q can be written as a nonlinear function $Q = Q(BC(MNR_1), \dots, BC(MNR_{N_g}))$ for N_g diffusion measuring gradients. Omitting for simplicity the coordinate \mathbf{x} and using the short notation $v_i = (BC(MNR_i))$, we get for the variance σ_Q^2

$$\sigma_Q^2 = \sum_{i=1}^{N_g} \sigma_{BC(MNR_i)}^2 \left(\frac{\partial Q}{\partial v_i} \Big|_{E[BC(MNR_i)]} \right)^2 + 2 \sum_{\substack{i,j=1 \\ i < j}}^{N_g} \sigma_{BC(MNR_i), BC(MNR_j)} \left(\frac{\partial Q}{\partial v_i} \Big|_{E[BC(MNR_i)]} \right) \left(\frac{\partial Q}{\partial v_j} \Big|_{E[BC(MNR_j)]} \right) \quad [A2.3]$$

Eq. A2.3 combines a variance and a covariance term. For a comparison of the variances $\sigma_{BC(MNR_i)}^2$ and the covariances $\sigma_{BC(MNR_i), BC(MNR_j)}$ see the Box-Whisker plots in panels (C1, C2) of Fig. A2. The Box-Whiskers are given for all voxels of the phantom presented in panel (c) of Fig. 6. In contrast to the variances are the covariances symmetrically distributed around zero with small magnitudes. The magnitudes of the corresponding Pearson-correlations (not shown) are below 0.25 for 98% of all cases considered in (C1-C2). This indicates no or at least small correlations between the random variables $BC(MNR_i)$ and $BC(MNR_j)$. This result is compatible with the denoising procedure applied. The method induces spatial correlations in $BC(MNR_i)$ by the spatial filtering process, but no correlations between $BC(MNR_i)$ and $BC(MNR_j)$ at the same position \mathbf{x} , as denoising does not mix information from different gradients. We assume therefore that the covariance term is not important and can be omitted.

Calculating the error propagation for $BC(MNR_i)$, we get for $u_i = MNR_i$:

$$\sigma_{BC(MNR_i)}^2 = \sigma_{MNR_i}^2 \left(\frac{dBC}{du_i} \Big|_{E[MNR_i]} \right)^2 \quad [A2.4]$$

The squared derivative in Eq. A2.4 is presented in panel (D) of Fig. A2, it increases strongly with decreasing SNR. Substituting Eq. A2.4 into Eq. A2.3 without covariance term we get a similar dependence of σ_Q^2 on the SNR. This can be observed e.g. in the numerical results shown in section 3.3. Note in panels (e) and (i) of Fig. 7 the enhanced variance of $Q = FA$ in the region within the tube, compared to the region outside.

For postfiltering a simple model of the noise level is necessary for the filter. Based on the discussion above we propose the following parameterization:

$$\sigma_Q(x) = q_Q \sqrt{\sum_{i=1}^{N_g} \left(\frac{dBC}{du_i} \Big|_{E[MNR_i(x)]} \right)^2 \left(\frac{\partial Q}{\partial v_i} \Big|_{E[BC(MNR_i(x))]} \right)^2} \quad [A2.5]$$

For $Q = MD$ or FA in explicit functions $Q = Q(BC(MNR_1), \dots, BC(MNR_{N_g}))$ are proposed and validated for DTI measurements with icosahedral diffusion gradients (used also in our study) [36]. Consequently, the partial derivatives in Eq. A2.5 can be calculated analytically. For $Q = MD$ we have

$$MD(x) = \sum_{i=1}^{N_g} \log(|S_0(x)| / BC(MNR_i(x))) / (b * N_g)$$

and for

$$Q = FA$$

we have

$$FA(x) = \sqrt{\frac{3.75}{2.5 + \frac{1}{m(x)/MD(x)^2 - 1}}}$$

with

$$m(x) = \sum_{i=1}^{N_g} \log^2(|S_0(x)| / BC(MNR_i(x))) / (b^2 * N_g)$$

For postfiltering of real data only one realization of the random variable $Q(\mathbf{x})$ is available, therefore, the following procedure to fit q_Q is applied: First, calculate the sum in Eq. A2.5 for all \mathbf{x} after denoising of the DWIs, where the arguments $\mathbf{E}[MNR_i(\mathbf{x})]$ and $\mathbf{E}[BC(MNR_i(\mathbf{x}))]$ are approximated by $MNR_i(\mathbf{x})$ and $BC(MNR_i(\mathbf{x}))$. Then, take a homogeneous region of $Q(\mathbf{x})$ around a convenient point \mathbf{x}_0 and estimate there $\sigma_Q(\mathbf{x}_0)$ from the denoised image. From this determine q_Q via Eq. A2.5. The heteroscedastic standard deviation is then used in the window Ψ of the spatial filter, as is described in section 2.2.2. The effectivity of this model for postfiltering is exemplified for $Q = FA$ in panels (f), (g) and (j) of Fig. 7, see section 3.3 for the value of q_{FA} . For $Q = MD$ similar results (not shown) with the parameters $q_{MD} = 10$ (for $NEX = 1$) and $q_{MD} = 6$ ($NEX = 4$) were achieved for the phantom.

An extension of our nonlinear spatial filter to postfiltering of vector fields is not available, therefore we do not discuss a noise model for the main diffusion directions.

REFERENCES

- [1] K. Terajima, H. Matsuzawa, K. Tanaka, M. Nishizawa, and T. Nakada, (2007), Cell-Oriented Analysis in Vivo Using Diffusion Tensor Imaging for Normal-Appearing Brain Tissue in Multiple Sclerosis, *NeuroImage*, **37**, 1278-1285.
- [2] S. J. Teipel, R. Stahl, O. Dietrich, S. O. Schoenberg, R. Perneczky, A. L. W. Bokde, M. F. Reiser, H. Möller, and H. Hampel, (2007), Multivariate Network Analysis of Fiber Tract Integrity in Alzheimer's Disease, *NeuroImage*, **34**, 985-995.
- [3] P. J. Basser, and D. K. Jones, (2002), Diffusion-Tensor MRI: Theory, Experimental Design and Data Analysis – A Technical Review, *NMR. Biomed*, **15**, 456-467.
- [4] Hasan K. M., Rodenacker K., and Hahn K. R., (2006), Evaluation of SNR Performance and Utility of High Spatial and Angular Resolution Denoised 1 mm³ Isotropic DTI of Entire Human Brain at 3T, *Proc. Intl. Soc. Mag. Reson. Med.*, **14**, 344.
- [5] C. A. Clark, M. Hedehus, and E. Moseley, (2002), *In Vivo Mapping of the Fast and Slow Diffusion Tensors in Human Brain*, *Magn. Reson. Med.*, **47**, 623-628.
- [6] M. Kim, I. Ronen, K. Ugurbil, and Dae-S. Kim, (2006), Spatial Resolution Dependence of DTI Tractography in Human Occipito-Callosal Region, *NeuroImage*, **32**, 1243-1249.
- [7] K. R. Hahn, S. Prigarin, S. Heim, and K. M. Hasan, (2006), Random Noise in Diffusion Tensor Imaging, Its Destructive Impact and Some Corrections, In: J. Weickert and H. Hagen, Editors, *Visualization and Processing of Tensor Fields*, Berlin, Springer, 107-117.
- [8] G. J. M. Parker, J. A. Schnabel, M. R. Symms, D. J. Werring, and G. J. Barker, (2000), Nonlinear Smoothing for Reduction of Systematic and Random Errors in Diffusion Tensor Imaging, *J. Magn. Reson. Imaging*, **11**, 702-710.
- [9] Z. Ding, J. C. Gore, and A.W. Anderson, (2005), Reduction of Noise in Diffusion Tensor Imaging Using Anisotropic Smoothing, *Magn. Reson. Med.*, **53**, 485-490.
- [10] J. E. Lee, M. K. Chung, and A. L. Alexander, (2006), Evaluation of Anisotropic Filters for Diffusion Tensor Imaging, In: *Proc. of 3rd IEEE Intern., Symposium on Biomedical Imaging Nano to Macro*, 77-78.
- [11] K. Tabelow, J. Polzehl, V. Spokoiny, and H. U. Voss, (2007), Diffusion Tensor Imaging: Structural Adaptive Smoothing, *NeuroImage*.
- [12] S. Heim, L. Fahrmeir, P. Eilers, and B. Marx, (2007), 3D space-varying coefficient models with Application to Diffusion Tensor Imaging. *Comput. Statistics & Data Analysis*, **51**, 6212-6228.
- [13] S. Basu, T. Fletcher, and R. Whitaker, (2006), Rician Noise Removal in Diffusion Tensor MRI, *Proc MICCAI 2006*, Springer, 117-125.
- [14] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache, (2006), Clinical DT-MRI Estimation, Smoothing and Fiber Tracking with Log-Euclidian Metrics, In: *Proc. Third IEEE International Symposium on Biomedical Imaging*, 786-789.
- [15] X. Zhang, H. Ye, and H. Zhang, (2008), Multi-Channel Wavelet-Based Diffusion Method for Denoising DTI Images, In: *Proceedings of International Conference on Biomedical Engineering and Information*, 178-182.
- [16] M. Martin-Fernandez, E. Munzo-Moreno, L. Cammoun, J. P. Thirian, C. F. Westin, and C. Alberola-Lopez, (2009), Sequential Anisotropic Multichannel Wiener Filtering with Rician Bias Correction Applied to 3D Regularization of DWI Data, *Medical Image Analysis*, **13**(1), 19-35.
- [17] K. R. Hahn, S. Prigarin, and K. M. Hasan, (2005), The Feasibility of Diffusion Tensor Imaging for the Human Brain at 1 mm³ Resolution, In: *Proceedings of the 13th Annual Meeting of the ISMRM*, Miami, 161.

- [18] E. L. Lehmann, (1999), *Elements of Large-Sample Theory*, New York, Springer.
- [19] H. Gudbjartsson, and S. Patz, (1995), The Rician Distributions of nNoisy MRI Data, *Magn. Reson. Med.*, **34**, 910-914.
- [20] P. G. Batchelor, D. Atkinson, D. L. G. Hill, F. Calamante, and A. Connely, (2003), Anisotropic Noise Propagation in Diffusion tTensor MRI Sampling Schemes, *Magn. Reson. Med.*, **49**, 1143-1151.
- [21] K. M. Hasan, D. L. Parker, and A. L. Alexander, (2001), Comparison of Gradient Encoding Schemes for Diffusion Tensor MRI, *J. Magn. Reson. Imaging*, **13**, 769-780.
- [22] J. Sijbers, and A. J. Dekker, (2004), Maximum Likelihood Estimation of Signal Amplitudes and Noise Variance from MR Data, *Magn. Reson. Med.*, **51**, 586-594.
- [23] V. Aurich, J. Weule, 1995. Non-linear Gaussian filters performing edge preserving diffusion, in: Proc. of 17th DAGM Symposium, Bielefeld, Springer, 538-545.
- [24] G. Winkler, V. Aurich, K. R. Hahn, A. Martin, and K. Rodenacker, (1999), Noise Reduction in Images: Some Recent Edge-Preserving Methods, *J. Patt. Recog. Image Analysis*, **9**(4), 749-766.
- [25] E. Mühlhaus, (1997), Die Sprungerhaltende Glättung Verrauschter, Harmonischer Schwingungen, PHD Dissertation, University Düsseldorf, Department for Mathematics and Natural Science.
- [26] J. Sijbers, A. J. denDekker, P. Scheunders, and D. VanDyck, (1998), Maximum-Likelihood Estimation of Rician Distribution Parameters, *IEEE Trans. Med. Imaging*, **17**(3), 357-361.
- [27] K. R. Hahn, S. Prigarin, and B. Pütz, (2001), Edge Preserving Regularization and Tracking for Diffusion Tensor Imaging, In: Proc. of 4th International Conference on Medical Image Computing and Computerassisted Intervention MICCAI 2001, Utrecht, Berlin, Springer, 195-203.
- [28] S. Heim, K. Hahn, D. G. Saemann, L. Fahrmeir, and D. P. Auer, (2004), Assessing DTI Data Quality Using Bootstrap Analysis, *Magn. Reson. Med.*, **52**, 582-589.
- [29] S. Mori, (2007), *Introduction to Diffusion Tensor Imaging*, Elsevier.
- [30] J. Dufner, U. Jensen, and E. Schuhmacher, (1992), *Statistik Mit SAS*, Teubner Stuttgart.
- [31] K. R. Hahn, S. Prigarin, and K. M. Hasan, (2007), Can We Expect Reproducible and Unbiased Information from Denoised DTI with Low SNR?, In: *Proceedings of the 15th Annual Meeting of the ISMRM*, Berlin, 1187.
- [32] D. K. Jones, M. R. Symms, M. Cercignani, and R. J. Howard, (2005), The Effect of Filter Size on VBM Analyses of DT-MRI Data, *NeuroImage*, **26**, 546-554.
- [33] K. Bockhorst, P. A. Narayana, R. Liu, P. Ahobila-Vijula, J. Ramu, M. Kamel, J. Wosik, T. Bockhorst, K. Hahn, K. M. Hasan, and J. R. Perez-Polo, (2008), Early Postnatal Development of Rat Brain: In Vivo Diffusion Tensor Imaging, *Journ. Neuroscience Research*, 1520-1528.
- [34] A. Shen, and M. Srivastava, (1990), *Regression Analysis*, Springer.
- [35] W. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, (1992), *Numerical Recipes*, Cambridge University Press.
- [36] K. M. Hasan, and P. A. Narayana, (2003), Computation of the Fractional Anisotropy and Mean Diffusivity Maps without Tensor Decoding and Diagonalization: Theoretical Analysis and Validation, *Magn. Reson. Med.*, **50**, 589-598.