Hahn, Prigarin, Heim, Hasan:

Random noise in Diffusion Tensor Imaging, its Destructive Impact and Some Corrections

Projektpartner

gsf

MAX-PLANCK-GESELLSCHAFT

TIIT

Random noise in Diffusion Tensor Imaging, its Destructive Impact and Some Corrections

Klaus R. Hahn,[1] Sergej Prigarin,[2] Susanne Heim,[3] and Khader Hasan[4]

[1] Institute of Biomathematics and Biometrics, National Research Center for Environment and Health, D-85764 Neuherberg, Ingolst. Landstr. 1, Germany
[2] Institute of Computational Mathematics and Geophysics, Novosibirsk, Russia
[3] Institute of Statistics, LM-University, Ludwigstr. 33, Munich, Germany
[4] Department of Radiology, University of Houston, Texas, USA

Summary. The empirical origin of random noise is described, its influence on DTI variables is illustrated by a review of numerical and in vivo studies supplemented by new simulations investigating high noise levels. A stochastic model of noise propagation is presented to structure noise impact in DTI. Finally, basics of voxelwise and spatial denoising procedures are presented. Recent denoising procedures are reviewed and consequences of the stochastic model for convenient denoising strategies are discussed.

1 Introduction

Though the theoretical and experimental basics of Diffusion Tensor Imaging DTI are still in a stage of development, it is well established, that magnetic resonance measurements of diffusing water molecules can reveal unique information about the architecture of normal and diseased brain tissues. See [1] for a recent survey on basic concepts, experimental methods, postprocessing procedures, and potential applications. An enumeration of the limitations of DTI at present would fill a long list. Some of them are caused by the 'artifacts' which comprise effects of subject motion, eddy currents, susceptibility variations, calibration errors, and noise [1, 2].

Random or Johnson noise is essentially white and has its origin in thermal Brownian motion of electrons. Johnson noise is superposed in DTI by two components: noise from the scanner apparatus and noise from the patient's body inside the scanner [3]. The measurement of the magnetization, carrying the anatomical information, results in complex valued data in k-space which give, after Fourier transformation, the signal in configuration or physical space [3]. Johnson noise in the data or signals can be approximated by a complex Gaussian distribution with mean zero, constant standard deviation and independend real and imaginary parts [4].

Consequently, the magnitude of the signal follows the family of Rician distributions [5], which comprises distributions with nearly Gaussian shape for low noise levels. In case of increasing noise however, the distributions become appreciably skewed and have a biased mean value. Hence the expectation value of a noisy signal magnitude is different from its noise-free or true value. Such magnitudes or diffusion weighted images DWI are mapped in the standard model of DTI via the Stejskal-Tanner equations ST to the diffusion tensor [6, 7] and then via several nonlinear transformations to detailed anatomical information of the brain. From a statistical point of view one should realize that in general any nonlinear transformation can transform a Gaussian distribution to a skewed and heavy tailed one with biased mean value. Consequently, a chain of such transformations can create a highly complicated stochastic situation. In fact, nonlinear noise propagation can lead to severe misinterpretations in DTI, which is still one of the central problems.

Several topics are addressed in this paper. First we present some formal results of stochastics to model noise propagation in DTI. Then, a survey of published results on noise artifacts and denoising methods is presented. The whole range of signal to noise ratios is covered, own results for high noise levels supplement the review. Final aim of this work is to structure the complicated field of noise impact in DTI and to support the application and development of convenient denoising methods.

2 Noise impact

First of all we introduce some fundamental concepts of DTI and of its statistics. In particular, the statistical Delta Method will be introduced. This method describes the large sample convergence to Gaussian distributions for variables which are derived by nonlinear transformations. Second, we present a survey of published results on noise artifacts. These studies comprise results achieved by Monte Carlo simulation and by bootstrap sampling.

2.1 Noise propagation model

In the following we restrict ourselves to statistical aspects caused by Johnson noise; the influence of non statistical distortions is excluded. In addition, only the standard diffusion tensor with rank=2 is considered, for extensions see Chapter 10 by Özarslan et al. The chain of nonlinear transformations, leading from measured quanitites to anatomically relevant observables proceeds as follows: {signals S} $\mapsto$ {magnitudes $|S|$} $\mapsto$ {tensor $\mathbf{d}$} $\mapsto$ {eigenvalues, eigenvectors} $\mapsto$ {anisotropy, tracks} $\mapsto$ {connectivity, etc.} Only for the first transformation, the statistics is completely formalized [5]. We present in the following a closer look at the second map in order to exemplify statistical peculiarities caused by nonlinear mappings. The central concept in standard DTI is the diffusion tensor $\mathbf{d}(\mathbf{x})$ for any voxel $\mathbf{x}$ in e.g. a brain. The three

eigenvalues $\lambda_i(\mathbf{x})$ and eigenvectors $|i(\mathbf{x}) >$ describe the geometric properties of a diffusion ellipsoid along the fibers. As the tensor is real and symmetric, a convenient braket notation [8] is used, where the ket $|. >$ is a column vector, the bra $< .|$ a transposed ket, and $< .|. >$ a scalar product. In this notation we get

$$\mathbf{d}(\mathbf{x}) = \begin{pmatrix} d_{11}(\mathbf{x}) \; d_{12}(\mathbf{x}) \; d_{13}(\mathbf{x}) \\ d_{12}(\mathbf{x}) \; d_{22}(\mathbf{x}) \; d_{23}(\mathbf{x}) \\ d_{13}(\mathbf{x}) \; d_{23}(\mathbf{x}) \; d_{33}(\mathbf{x}) \end{pmatrix} = \sum_{i=1}^{3} \lambda_i(\mathbf{x})|i(\mathbf{x}) >< i(\mathbf{x})|. \qquad (1)$$

The ST equations are then

$$|S_j(\mathbf{x})| = |S_0(\mathbf{x})| \exp\left(-b \sum_{i=1}^{3} \lambda_i(\mathbf{x}) < i(\mathbf{x})|\mathbf{g}_j >^2 \right), \qquad (2)$$

incorporating the diffusion weighting b-value which is a function of scanning parameters [7], a normalized diffusion measuring gradient $|\mathbf{g}_j >$, the diffusion weighted image DWI $|S_j(\mathbf{x})|$, $j \geq 1$, and the reference $|S_0(\mathbf{x})|$.

Noise enters the system via the complex signals $S_k(\mathbf{x})$, by $S_{k,noisy}(\mathbf{x}) = \mathbf{Re}[S_k(\mathbf{x})] + \varepsilon_{Re} + i(\mathbf{Im}[S_k(\mathbf{x})] + \varepsilon_{Im})$, for $k \geq 0$, where ε_{Re} and ε_{Im} are independent and normally distributed, $\varepsilon \sim N(0, \sigma)$. The noise level σ is the Rayleigh corrected standard deviation of background noise [4, 5].

To quantify the signal to noise ratio, we define $SNR_k = |S_k(\mathbf{x})|/\sigma$ for k$\geq$0, where $|S_k(\mathbf{x})|$ is without noise [5]. As SNR_k determines, within the Rician family, the distribution of $|S_{k,noisy}(\mathbf{x})|$, these distributions change in space. In the same way, the statistical properties of derived variables affected by noise, like $\mathbf{d}(\mathbf{x})$, change with $\mathbf{x}$. In short notation, those variables build up random fields.

We introduce now several abbreviations to formulate the least square estimation of the tensor from the measured DWIs for the general case including $n \geq 6$ gradient directions $< \mathbf{g}_j| = (g_{1,j}, g_{2,j}, g_{3,j})$. To make the notation more transparent, the dependence on $\mathbf{x}$ and the label *noisy* is suppressed:

$$\mathbf{D^T} := (D_1, D_2, D_3, D_4, D_5, D_6) = (d_{11}, d_{22}, d_{33}, d_{12}, d_{13}, d_{23})$$
$$(\mathbf{A})_j := (g_{1,j}^2, g_{2,j}^2, g_{3,j}^2, 2g_{1,j}g_{2,j}, 2g_{1,j}g_{3,j}, 2g_{2,j}g_{3,j}) \qquad (3)$$
$$s_j := -log(|S_j|/|S_0|)/b,$$

where $(\mathbf{A})_j$ is a row of the $n \times 6$ matrix $\mathbf{A}$. The equation $\mathbf{AD} = \mathbf{s}$ for $\mathbf{D}$ is then solved by $\mathbf{D} = \mathbf{B}^{-1}(\mathbf{A}^T\mathbf{s})$, where $\mathbf{B} = \mathbf{A}^T\mathbf{A}$, see [9]. Introducing the weights $w_{il} = \sum_{k=1}^{6} B_{lk}^{-1} A_{ik}$, we can finally write $D_l = \sum_{i=1}^{n} w_{il}s_i$, and find for the expectation of D_l,

$$\mathbf{E}[D_l] = -\sum_{i=1}^{n} w_{il}\mathbf{E}[log(|S_i|)]/b + \mathbf{E}[log(|S_0|)]\sum_{i=1}^{n} w_{il}/b. \qquad (4)$$

This equation shows the origin of a possible bias in D_l, hence in general

$$\mathbf{E}[D_l] \neq -\sum_{i=1}^{n} w_{il} log(|\mathbf{E}[S_i]|)/b + log(|\mathbf{E}[S_0]|)\sum_{i=1}^{n} w_{il}/b. \qquad (5)$$

The right hand side of Eq(5) describes the tensor components without noise, as the signals S_i are normally distributed. When the DWIs $|S_i|$ are essentially free from bias, $SNR_i > 3$, the function log introduces a (possibly) small bias in D_l, as log is a concave map. This effect is enhanced if the DWIs are biased for $SNR_i \leq 3$. Concave or convex mappings of random variables produce bias effects due to the Jensen inequality [10]. Therefore, any further nonlinear transformation on D_l can, in principle, cause additional bias in the derived variable.

A further important aspect is the shape of the distributions. For high noise level the DWIs, as well as derived variables are not normally distributed. However, as noise is sufficiently reduced, the Delta Method [11] predicts approximate Gaussian statistics for all variables of the DTI chain. This follows from an iterative application of the following Theorem and of its generalizations : 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 any nonlinear transformation $f : T_m \rightarrow f(T_m)$ 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 [11] for an exact but less descriptive formulation and for extensions to multivariate cases. Thus, reduction of noise greatly simplifies the structure of the mentioned random fields. However, as Θ depends on the spatial coordinate, variance is still varying in space. The Delta Method does pose only weak restrictions on the distributions of T_m and thus extends the usefulness of the Central Limit Theorems CLT [11] as prerequisites for an application of this method. Appropriate T_m can in DTI experiments be achieved by performing m replications of experiments and consecutive averaging of the m magnitudes $|S_{i,noisy}(\mathbf{x})|$. For low m, due to practical limitations, this is one of the standard procedures in DTI to denoise data.

2.2 Noise artifacts

We review only a selection from the huge number of articles on noise artifacts and emphasize the diversity of artifacts, more technical papers are not considered. The artifacts are investigated by numerical modelling via Monte Carlo simulation MCS [12, 13], perturbation theory [8] and bootstrap sampling [1, 14].

The degree of anisotropy in diffusion is connected to the homogeneitiy of the fiber directions in a measured voxel, as diffusion propagates mainly along the fiber direction, see Chapter 7 by Vilanova et al. for illustrations of anisotropy and nerve fibers in the human brain. In [12] different quantitative indices of anisotropy are investigated. The authors show that rotational variant indices suffer from non random orientational artifacts and

can make highly anisotropic white matter structures appear isotropic in vivo. Therefore, rotationally variant indices depending on the eigenvalues are included in their study, like e.g. the fractional anisotropy $FA = (3\sum_{k=1}^{3}(\lambda_k - \lambda)^2/2\sum_{k=1}^{3}\lambda_k^2)^{1/2}$, where $\lambda = \sum_{k=1}^{3}\lambda_k/3$, see [12, 13] for more indices depending on eigenvalues. It is shown by MCS that those indices are biased in the presence of noise. This bias enhances artificially the mean anisotropy and can make isotropic diffusion appear anisotropic. Two sources of this error are detected : a) the mean eigenvalues are biased, where the largest eigenvalue is typically enhanced by noise, the smaller ones are reduced; b) noise introduces sorting bias, i.e. due to overlapping statistical distributions of neighboring eigenvalues, magnitude sorting fails. These bias effects increase with decreasing SNR, the study covers a range of $SNR_0 > 5$. Similar findings are reported in [13], in addition, the eigenvalue distributions are investigated. The dependence of skewness on the angles between main diffusion and laboratory system or diffusion gradients is apparent. This exemplifies that DTI distributions build up random fields. Also negative eigenvalues are detected for higher noise level preventing an interpretation of the tensor as a quantity describing diffusion. Perturbation theory is applied [15] for $SNR_0 > 20$, to calculate power series expansions of the eigenvalues and of eigenvectors of the tensor for different model diffusions. The results for the bias in eigenvalues and in FA of [12, 13] are essentially confirmed. Noise in the eigenvector orientation produces random walk trajectories which should model the nerve fiber pathes. The mean position error of the calculated tracks and the standard deviation are calculated for a total of 256 path steps. Both increased, in different manner, with the step number and the noise level. This may indicate fundamental limits in accuracy for tracking, though only a very simple tracking algorithm is applied [15]. The studies discussed so far deal with $SNR_0 > 5$ for b $\approx$ 1000 s mm^{-2} and focus more on even higher SNR_0, relevant for clinical investigations. Recently, experiments with higher b-values (b$\gg$1000 s mm^{-2}) to measure non-Gaussian diffusion [16] or with high spatial resolution (e.g. 1 mm^3) to reduce partial volume effects are performed. Such data include DWIs with $SNR_k < 3$ (henceforth with $k > 0$), which are strongly influenced by peculiarities of the Rician statistics, and consequently we may find different noise artifacts. The first systematic MCS for higher b-values was published recently [17]. Just one interesting result may be reported. In contrast to the findings of [12, 13, 15], the mean FA can now be essentially unbiased for b $\approx$ 3000 s mm^{-2}, or underestimated for b > 5000 s mm^{-2}.

Non-parametric bootstrap BS analysis offers a more empirical approach to error analysis allowing a better inclusion of non statistical distortions. These distortions modify the statistical distributions but it is hard to model them in the frame of MCS. The BS draws inferences about some features of unknown distributions by generating multiple replications. The replications are achieved by iterated random drawings with replacement out of a pool of experimental data. See Fig. 1 for a BS scheme, applied to a typical DTI acquisition. In this setup k replications of an experiment with 6 diffusion gradients and one
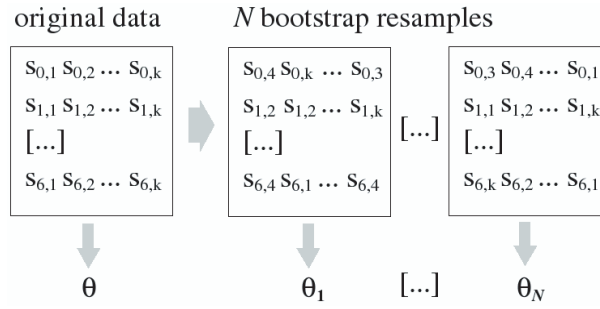
original data N bootstrap resamples

$$s_{0,1}\ s_{0,2}\ \cdots\ s_{0,k}$$
$$s_{1,1}\ s_{1,2}\ \cdots\ s_{1,k}$$
$$[\ldots]$$
$$s_{6,1}\ s_{6,2}\ \cdots\ s_{6,k}$$

$$s_{0,4}\ s_{0,k}\ \cdots\ s_{0,3}$$
$$s_{1,2}\ s_{1,2}\ \cdots\ s_{1,k}$$
$$[\ldots]$$
$$s_{6,4}\ s_{6,1}\ \cdots\ s_{6,4}$$

$[\ldots]$

$$s_{0,3}\ s_{0,4}\ \cdots\ s_{0,1}$$
$$s_{1,1}\ s_{1,2}\ \cdots\ s_{1,k}$$
$$[\ldots]$$
$$s_{6,k}\ s_{6,2}\ \cdots\ s_{6,1}$$

θ θ_1 $[\ldots]$ θ_N

Fig. 1. General bootstrap resampling scheme adapted to a typical DTI acquisition of a reference and six diffusion weighted signals, each k times. By random drawing with replacement N bootstrap resamples are obtained. The distribution of a statistic of interest Θ is determined by the $N+1$ samples.

reference are performed. A box indicates a complete data set with pool size k, the left box is the purely empirical starting point of the procedure. To the right, artificial resamples created by random drawings are shown. The information in all samples together defines the distribution of interest and allows to study, in an approximate way, the statistical properties of a random variable in a voxel.

By BS in [18] the uncertainty in main diffusion directions is analyzed in vivo. Applying the formalism of dyadic tensors 95% confidence intervals for the angles between the mean and the random directions are calculated. A correlation between this uncertainty and the anisotropy index $C_l = (\lambda_1 - \lambda_2)/(\lambda_1 + \lambda_2 + \lambda_3)$ for $\lambda_1 \geq \lambda_2 \geq \lambda_3$ is found. The uncertainty grows with decreasing C_l qualitatively like C_l^{-1}. Even for high C_l an uncertainty of about 2.5 degrees remains. In [19] BS is applied to distribution specific parameters which can serve as quality measures for DTI data, this could help to detect e.g. data which are corrupted by some machine error in the scanner. For this purpose, the confidence intervals of FA in white matter are determined and submitted to histogram analysis. The mean, modus and height are extracted as quality descriptors. The study particularly investigates the impact of noise and of denoising, as well as motion of the patient on those parameters. In [14] MCS and BS are applied in conjunction. By MCS it is investigated how good a multivariate Gaussian distribution can describe noise in the tensor. Marginal distributions of the tensor and the distributions of the squared rational anisotropy RA, see Chapter 7, are compatible with this assumption. In particular the marginals are well normal distributed already for $SNR_0 > 2$, when only 6 diffusion gradients are applied, and for clinical b-values. The estimation of the covariance matrix is less robust, for a linear regression model the diagonal elements are underestimated by about 20%. In addition, BS for DTI data is introduced in this paper. Its reliability is shown by MCS on simulated data. BS is also applied to human data under

approximate clinical conditions, one result is that in the majority of voxels the statistical properties of the tensor components are compatible with the Gaussian assumption.

3 Corrections of noise effects

The different denoising methods can be divided into voxelwise and spatial procedures. In the first group, frequently experiments with a small number of gradients close to the minimum number $n_g = 6$ and sometimes with different b-values per experiment are repeated, to average the DWIs or to derive the tensor via regression methods [7]. In multigradient experiments the gradient number is enhanced, $n_g \gg 6$. Different ways to construct the spatial orientation of such gradients were proposed, see [20] for a review and a comparison. Finally, the tensor can be derived from the DWIs by the least squares fit described in Section 2.1. Both acquisition schemes reduce noise in the system voxelwise. A complementary technique is offered by spatial denoising, where samples of neighboring voxels are used to estimate the variable of interest. This technique relies on the fact, that anatomical units occupy at least several neighboring voxels in a brain, and that it is possible to detect those regions ROI of " homogeneity". Such methods are applied to reduce the sorting bias of tensor eigenvalues [21, 22], or to filter the spatial DWI fields [23, 24], more global assumptions are involved in the denoising methods [25, 26, 27, 28].

3.1 Voxelwise denoising

For $SNR_k > 3$, voxelwise averaging of DWIs derived from repeated measurements introduces, according to the classical CLT [11], unbiased normally distributed mean values with small variance. If, by a high number of replications, the variance is reduced sufficiently one can estimate the derived DTI variables practically without noise influence, due to the Delta Method. In agreement with that it was shown by MCS [15], that for $SNR_0 \approx 20$ bias in the eigenvalues is minimized best by averaging the DWIs before the tensor is derived. For medium $SNR_0 \approx 50$ the results imply an equivalence between DWI and tensor averaging, as the bias in the tensor, see Eq(4), is no more relevant. Only at high SNR_0 direct eigenvalue averaging is equivalent to the other methods, as the whole system is now close to the Gaussian limit.

Different orientations in multigradient systems led to the introduction of the condition number $\kappa = \kappa(\mathbf{A})$, which gives an error bound by $\kappa \geq$ (relative error in $\mathbf{D}$)/(relative error in $\mathbf{s}$) [29, 30], for notations see section 2.1. In [30] it is shown, that for icosahedral gradients κ is small and independend of rotations of $\mathbf{A}$ in the laboratory system. Within this bound, however, the tensor distributions and the bias effects do depend on SNR_k or on the gradient directions, see Eq(2) for an explanation. The number of gradients n_g has also an important influence on bias effects. The MCS in [31] shows, that
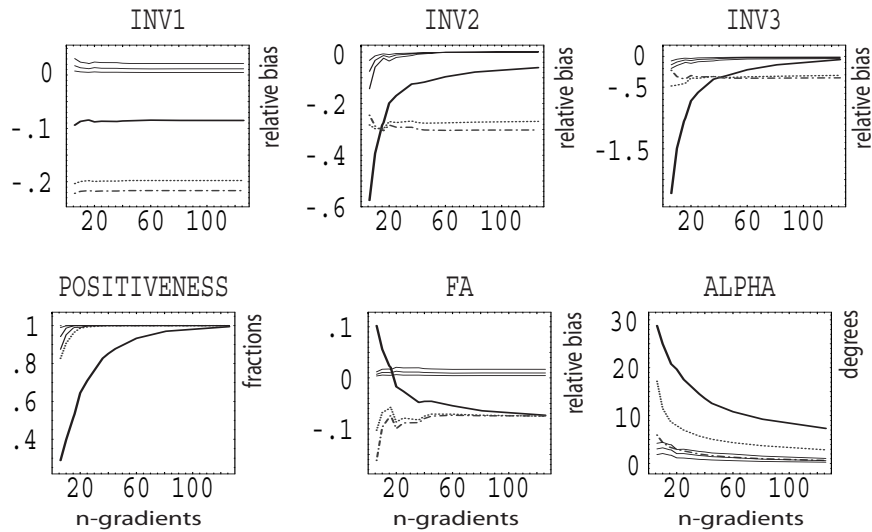
Fig. 2. Impact of Rician statistics on DTI variables for $SNR_0 = 4.2$. The model diffusion is "cigar shaped", $FA = 0.8$ and $MD = 0.0007\ mm^2 s^{-1}$. Relative bias versus number of applied gradients is presented for the invariants $INV1$, $INV2$, $INV3$ and FA, also the fractions of positive definite voxels and the angles between main eigenvector of the model and the averaged main eigenvector of noisy diffusion. Different line styles correspond to different levels of voxelwise averaging : *thick line* (one experiment), *dotted line* (5 replications and DWI averaging DA), *dashed dotted line* (30 replications and DA), *thin lines close together* (30, 50, 100 replications, DA and application of bias correction according to Rician statistics).

the fractional anisotropy FA, the mean diffusion $MD = \sum_{k=1}^{3} \lambda_k / 3$ and the direction of main diffusion depend on the number n_g of uniformly distributed gradients. Increasing n_g reduces and stabilizes those bias effects, for $SNR_0 = 15$ at least 20 gradients are necessary to achieve reliable anisotropy maps and 30 gradients for reliable directions and MD.

To include $SNR_k \leq 3$ we performed MCS at $SNR_0 \approx 4$. Several three dimensional models of realistic diffusion tensors are explored, with b= 1000 s mm^{-2} and $|S_0| = 1000$. The gradients are icosahedral, twelve different gradient sets are used, $n_g \in \{6, 10, 15, 16, 20, 25, 36, 40, 45, 60, 81, 126\}$. In Fig. 2 typical results for a "cigar shaped" diffusion with d_{11}, d_{22}, d_{33} =.00155, .000354, .000191 $mm^2 s^{-1}$, else *zero* are presented. To minimize SNR_k every gradient set is rotated, such that at least one gradient direction is parallel to the main diffusion direction of the model. This produces maximal Rician bias in the corresponding DWI. The relative bias of the three invariants of the tensor $INV1 = \sum_{k=1}^{3} \lambda_k$, $INV2 = \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3$, $INV3 = \lambda_1 \lambda_2 \lambda_3$ are shown in the upper panels. Below one finds the fractions of positive definite voxels, the relative bias of FA and the angle between the true main diffusion

direction and the averaged noisy main diffusion direction. Before averaging, the main directions were calculated by dyadic sorting, see below, and were aligned to the model. In contrast to [31], increasing n_g does not eliminate in all cases the bias effects, see thick lines in panels $INV1$, $INV2$, FA and $ALPHA$. This is due to a strong bias in the DWIs, with minimum $SNR_k \approx 1$. When DWI averaging is performed, the bias effects even increase, see dotted and dashed-dotted lines. Due to the classical CLT averaging of DWIs transforms the Rician distributions to nearly normal shape with small variance, centered around the biased Rician mean values. After bias correction [5], the relative bias of the invariants and of FA is practically *zero* for 100 replications, the angle bias is below 2 degrees for $n_g \geq 6$. If only 50 or 30 experiments are applied the results deteriorate only slightly (thin lines). Positive definiteness of the tensor is violated drastically before averaging, but denoising by DWI averaging improves the situation considerably, see Fig. 2.

3.2 Spatial denoising

In [12] the so called " lattice" index of anisotropy is proposed, combining eigenvalue and eigenvector information in a ROI. This index shows enhanced robustness in the presence of noise for low anisotropy, compared to intravoxel indices, like FA. Also eigenvalue sorting is improved by considering ROIs. In [21] vector sorting is introduced, where the maximum coherence of main eigenvectors in a ROI is involved. By MCS it is shown for isotropic diffusion that this method is superior to magnitude sorting. Dyadic sorting [22] is another improvement. In this method first magnitude sorting is applied to a spatially averaged tensor used as reference, then, by a dyadic overlap measure for tensors, the eigenvalues of the unknown tensor are sorted.

Nonlinear filtering is applied in [23, 24] to DWIs. In those filters the ROIs, or better effective windows, where smoothing is performed, are not only defined by spatial conditions, but also by a distance measure for the DWI to enable edge preservation. Edges are typical features of spatial DTI variables as e.g. anisotropy and fiber directions can change drastically between two voxels. In [23] the diffusion equation by Perona Malik is applied, in [24] a chain of nonlinear Gaussian filters is used. Both methods include only few assumptions about the structure of the signals and seem to be convenient denoising tools for DTI data with $SNR_k > 3$.

Denoising of tensor fields is described in [25, 26], see also Chapter 18 by Pajevic et al, Chapter 19 and 25 by Weickert et al, and Chapter 24 by Westin et al. In [25] B-splines are applied to a discrete set of noisy DT-MRI measurements to obtain a continuous representation of the tensor field, see Chaper 18 for edge preserving representations by NURBS. In such representations differential geometric quantities, like curvature or torsion of fiber tracts, but also the tangent field could be derived directly. In [25] noise affected templates are denoised with good accuracy, except where the field is not homogenous, e.g. in regions where fiber tracts cross. In [26] the Stejskal Tanner Equations ST

for complex signals are used; to achieve them formally, replace the DWIs in Eq(2) by the corresponding complex signals or complex DWIs. Assuming that sufficiently many complex DWIs are given by the measurements, a smoothed tensor field $\mathbf{d}(\mathbf{x})$ and $S_0(\mathbf{x})$ are derived by a variational principle ensuring positive definiteness of the tensor. The minimization of the variational integral under ST constraints is achieved by an iterative procedure. The method is edge preserving and is tested in model and real data applications.

The main diffusion directions are smoothed in [27, 28]. By a variational regularization, in [27], coherent vector fields are estimated from noisy data. This method conserves discontinuities and reduces the regularization for small anisotropy. In model calculations those properties are verified, for higher noise level an influence of sorting bias leading to orthogonal directional artifacts is observed. The estimated vector field is then used as a prior to estimate, in a second step, also the eigenvalue fields. To this end a diffusion equation including the " flow" tensor is applied. This " flow" tensor includes information about the diffusion tensor and controls smoothing and edge preservation. For real data, denoised tensor, eigenvalue and FA fields are discussed. In [28] the regularized main directions are estimated by the Bayesian approach. The estimated maximum of the posterior probability for the main direction field relies on a trade-off between DTI data and a priori assumption regarding the low curvature of the nerve fibers. The a priori probability includes information about the behavior of the modelled direction field in the neighborhood or clique around the voxel of interest, leading to a Gibbs random field with interaction, the likelihood includes only voxelwise calculated probabilities. The regularized direction field is finally used to apply a new tracking algorithm to simulated and real data, allowing the treatment of diverging fibers.

Finally we mention a method proposed especially for very high noise levels, $SNR_k < 3$ [32]; DWI averaging is combined with nonlinear DWI filtering [24] and a bias correction, see Fig. 2. Human brain data with 1 mm^3 resolution, this is roughly a factor 10 below the clinical voxel volume which produces severe partial volume deficiencies [1], could be denoised successfully.

Both denoising principles discussed in Section 3 have inherent shortcomings. Voxelwise denoising involves many experiments, therefore patient motion introducing partial volume effects and distortions, as well as temporal instabilities in the scanner are the main limiting factors. Spatial smoothing suffers often from a trade off between blurring and bias caused by the applied method. Blurring occurs when different 'objects' cannot be discriminated by the denoising method and when consequently anatomically separated information is mixed together, like e.g. in the case of two neighbored fiber bundles with different directions. Bias is mainly caused by too strong priors; e.g. edge preserving filters can be tuned for very high quality in edge finding or 'object' discrimination, the price is usually a decreasing flexibility in the linear behavior or a decreasing ability to model curvature.

4 Conclusion

What can we learn from our analysis to find convenient strategies for spatial denoising ? For very high noise level, $SNR_k < 3$, denoising and bias correction of the DWIs is a suitable procedure [32]. For reduced noise the tensor distributions approach normality [14] and become also reasonable candidates for smoothing. Our own calculations indicate a limit for tensor smoothing well above $SNR_0 \approx 4$ due to the strong bias effects shown in Fig. 2. Tensor denoising is particularly important for multigradient designs where DWI averaging of replications, a convenient preprocessing step before spatial denoising, is seldom feasible. For higher SNR_0, when a bias due to the nonlinear DTI chain can be neglected, eigenvector and eigenvalue fields may be convenient variables. In [15] a similar SNR dependent denoising strategy for the reduction of eigenvalue bias is derived by perturbation theory. For $SNR_0 \approx 20$ denoising of the DWIs is recommended, for $SNR_0 \approx 50$ tensor denoising is shown to be equally good, and only for higher SNRs direct eigenvalue denoising is proposed. Additional priors in the denoising method may help to correct effects of minor skewness or kurtosis in the distributions, or may even correct intrinsic partial volume defects, e.g. in the main diffusion directions [28]. But, a spatial dependence in the (co)variance is predicted by the Delta Method already at very low noise levels and may be included in the denoising procedures.

References

1. Basser, P. J., Jones, D. K. (2002) Diffusion Tensor MRI: theory, experimental design and data analysis - a technical review. NMR Biomed. **15**, 456–467
2. Conturo, T. E., McKinstry, R. C. et al (1995) Diffusion MRI: Precision, accuracy and flow effects. NMR Biomed. **8**, 307–332
3. Vlaardingerbroek, M. T., den Boer, L. A. (1996) Magnetic Resonance Imaging. Springer, Berlin Heidelberg
4. Henkelmann, R. M. (1985) Measurement of signal intensities in the presence of noise in MR images. Med. Phys. **12/2**, 232–233
5. Gudbjartsson, H., Patz, S. (1995) The Rican distribution of noisy MRI data. Mag. Res. Phys. **34**, 910–914
6. Basser, P. J., Mattiello, J., LeBihan, D. (1994) MR diffusion tensor spectroscopy and imaging. Biophys. J. **66**, 259–267
7. Basser, P. J., Mattiello, J., LeBihan, D. (1994) Estimation of the effective self-diffusion tensor from the NMR spin echo. Journ. Mag. Res. **B 103**, 247–254
8. Messiah, A. (1972) Quantum Mechanics I. North-Holland Publishing
9. Press, W. P., Vetterling, W. T. et al (1992) Numerical Recipes. University Press, Cambridge
10. Renyi, A., (1962) Wahrscheinlichkeitsrechnung. VEB Deutscher Verlag der Wissenschaften, Berlin

11. Lehmann, E. L., (1999) Elements of Large-Sample Theory. Springer, New York
12. Pierpaoli, C., Basser, P., (1996) Toward a quantitative assessment of diffusion anisotropy. Mag. Res. Med. **36**, 893–906
13. Skare, S., Tie-Qiang, L. et al (2000) Noise considerations in the determination of diffusion tensor anisotropy. Mag. Res. Im. **18**, 659–669
14. Pajevic, S., Basser, P., (2003) Parametric and non-parametric statistical analysis of DT-MRI data. Journ. Mag. Res. **161**, 1–14
15. Anderson, A. W. (2001) Theoretical analysis of the effects of noise on diffusion tensor imaging. Mag. Res. Med. **46**, 1174–1188
16. Clark, C. A., Hedehus, M. et al, (2002) In vivo mapping of the fast and slow diffusion tensors in human brain. Mag. Res. Med. **47**, 623–628
17. Jones, D., Basser, P., (2004) ”Squashing peanuts and smashing pumpkins”: How noise distorts diffusion-weighted data. Mag. Res. Med. **52**, 979–993
18. Jones, D., (2003) Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. Mag. Res. Med. **49**, 7–12
19. Heim, S., Hahn, K. R. et al (2004) Assessing DTI data quality using bootstrap analysis. Mag. Res. Med. **52**, 582–589
20. Hasan, K. M, Parker, D. L. et al (2001) Comparison of gradient encoding schemes for diffusion-tensor MRI. Journ. Mag. Res. Im. **13**, 769–780
21. Martin, K. M., Papadakis, N. G. et al (1999) The reduction of the sorting bias in the eigenvalues of the diffusion tensor. Mag. Res. Im. **17**, 893–901
22. Basser, P. J., Pajevic, S. (2000) Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. Mag. Res. Med. **44**, 41–50
23. Parker, G. J.M., Schnabel, J. A. et al (2000) Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging. Journ. Mag. Res. Im. **11**, 702–710
24. Hahn, K. R., Prigarin, S. et al (2001) Edge preserving regularization and tracking for diffusion tensor imaging. In: Niessen W.J., Viergever, M.A.(Eds.) Medical image computing and computer-assisted intervention-MICCAI2001, Springer, Berlin, 195–203
25. Pajevic, S., Aldroubi, A. et al (2002) A continous tensor field approximation of discrete DT-MRI data for extracting microstructural and architectural features of tissue. Journ. Mag. Res. **154**, 85–100
26. Wang, Z., Vemuri, B. C. et al (2004) A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. IEEE Trans. Med. Im. **23**, 930–939
27. Coulon, O., Alexander, D. C. et al (2004) Diffusion tensor magnetic resonance image regularization. Med. Im. Anal. **8**, 47–67
28. Poupon, C., Clark, C. A. et al (2002) Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. NeuroImage **12**, 184–195
29. Skare, S., Hedehus, M. et al (2000) Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. Journ. Mag. Res. **147**, 340–352
30. Batchelor, P. G., Atkinson, D. et al (2003) Anisotropic noise propagation in diffusion tensor MRI sampling schemes. Mag. Res. Med. **49**, 1143–1151
31. Jones, D., (2004) The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo Study. Mag. Res. Med. **51**, 807–815
32. Hahn, K. R., Prigarin, S. et al (2004) A novel denoising technique for very noisy DTI data. In: Conf. Proc. of 12th Annual Meeting of ISMRM. Kyoto , 1208