Hahn, Prigarin, Pütz:

Spatial Smoothing for Diffusion Tensor Imaging with low Signal to Noise Ratios

Sonderforschungsbereich 386, Paper 358 (2003)

Online unter: http://epub.ub.uni-muenchen.de/

Spatial smoothing for Diffusion Tensor Imaging with low Signal to Noise Ratios

Klaus Hahn, Sergei Prigarin[+] and Benno Pütz[++]

Institute of Biomathematics and Biometrics of the National Research Center
for Environment and Health –gsf- , Neuherberg, Germany
[+]Institute of Computational Mathematics and Mathematical Geophysics, Novosibirsk, Russia
[++]Max-Planck-Institute of Psychiatry, München, Germany

Abstract: Though low signal to noise ratio (SNR) experiments in DTI give key information about tracking and anisotropy, e.g. by measurements with very small voxel sizes, due to the complicated impact of thermal noise such experiments are up to now seldom analysed. In this paper Monte Carlo simulations are presented which investigate the random fields of noise for different DTI variables in low SNR situations. Based on this study a strategy for spatial smoothing, which demands essentially uniform noise, is derived. To construct a convenient filter the weights of the nonlinear Aurich chain are adapted to DTI. This edge preserving three dimensional filter is then validated in different variants via a quasi realistic model and is applied to very new data with isotropic voxels of the size 1x1x1 mm^3 which correspond to a spatial mean SNR of approximately 3.

Discussion Paper 358 SFB386 2003

1 Introduction

Thermal or Johnson noise in Diffusion Tensor data is one of the obstacles which obstruct the evaluation of experiments with low Signal to Noise Ratios (SNR). The Rician family of distributions for the Diffusion Weighted magnitude Signals (DWI), see **[1,2,3],** is modified non linearily to the noise distributions of the tensor coefficients and to those of further derived variables, like anisotropy or main diffusion directions, which describe anatomical details of nerve fiber bundles. Mainly due to these nonlinearities noise in derived variables has a very complicated structure.

To increase the SNR`s **[2]** of the variables, denoising is frequently performed voxel wise. To this end, experiments with the minimal number of necessary gradients and perhaps different b values **[4]** are repeated. The tensor coefficients are then derived by multivariate linear regression on the log linearized DWI`s via the Stejskal Tanner equations **[4.1].** In a more recent experimental setup, multi gradient experiments **[5,6,7]** with appreciably enhanced numbers of gradients, frequently uniformly oriented in space, are applied. The tensor is here derived by a voxel wise singular value decomposition **[7].** In both cases, the multitude of experiments reduces noise effects in the tensor and in derived variables. Scanning time and subject motion **[4]** limit these procedures.

A complementary smoothing technique is offered by spatial smoothing where, for a single data set, small samples of neighboring voxels are used to estimate the mean of a local variable within this sample. Spatial smoothing is rather restrictive, care must be taken to avoid the introduction of bias effects. Such bias or blurring effects can be introduced by mixing information from different objects, like e.g. by averaging noisy diffusion directions from spatially close but different axon bundles. As will be shown, the application of convenient nonliner filter constructions can reduce blurring essentially. Most nonlinear spatial filters, see for a recent review **[8]** presuppose that the noise distributions within the neighborhoods involved are homogenous and that they are at least close to a Gaussian shape. Only recently a nonlinear filter was developed, which can be applied to situations with varying variance [**9]**.

At present spatial smoothing or regularization in DTI is discussed for different random variables, like e.g. the magnitude signal or DWI **[10],** the tensor field **[11,12],** the Eigenvalues and anisotropy coefficients **[13,14,15]** or the principal diffusion directions **[16,17,18].** Applying smoothing methods directly to the variable of interest has the advantage to restrict on low dimensional fields of the variable minimizing computer time and to deal with interpretable quantities supporting special adapted regularization methods, like in **[16,17,18].** On the other hand, for some of these variables, like Eigenvalues, anisotropy coefficients or main directions already in case of voxel wise smoothing bias effects are reported. These are investigated in model calculations for SNR`s above 20 by perturbation theory **[19**] or above approximately 5 by Monte Carlo simulations **[13, 14, 20]**. The results indicate that the voxel wise statistics of random variables involved in a DTI analysis is in general at least for medium SNRs (approximately 5-15) and low SNRs (<5 , where the Rician properties become essential) no more Gaussian and may be better described by skewed distributions **[13].** The fact, that bias and variance in simulation studies increase with 1/SNR **[14,19,20],** further indicates that these distributions presumably change with DWI (or diffusion) from voxel to voxel and form a specific random field for every variable in the brain. This is also exemplified by Monte Carlo simulations of Skare **[7],** which show that bias and variance of anisotropy depend on the, in practice unknown, angles between local main diffusion and the diffusion gradients applied.

In spite of their difficult statistical properties, the analysis of DTI experiments with low SNR`s is important, as these experiments hide key information to several central topics. To reduce partial volume effects which can e.g. cause “phantom connections” **[4]** between anatomically separated fiber tracts, experiments with small voxels which produce, due to their low signal intensities, low SNR`s are desirable. Also recent experiments with high b values **[21]** which separate fast and slow diffusion fractions lead to variables with low SNR. In addition, scanning time and artifacts due to high numbers of experimental replications or in multigradient experiments could be reduced, when spatial smoothing of low SNR data could reliably complement voxel wise smoothing.

To this end, in the first chapter of this work DTI data are analysed for typical features which may help to construct convenient spatial filters. The main results can be summarized as follows: Discontinuities, e.g. in the anisotropy field, favour the application of nonlinear, edge preserving filters, like those proposed in Parker **[10]** to reduce blurring.

On the other hand, as the fields of the variables have, aside from the edges, also extended regions with appreciable smooth curvature, the linear approximation quality of the filter is also important. This is a non trivial observation, as most edge preserving filters, like the diffusion equation applied in **[10],** the M-smoother **[22]** or the Aurich chain **[23],** achieve optimal edge resolution only for a restricted class of functions which change stepwise between constant regions. These filters produce for a parameterization adapted to this situation only weak approximations in regions with appreciable smooth curvature. It is the main goal of this paper to establish a new spatial filter which combines edge detection with a good generalization quality for curved functions.

As this filter shall be applicable in low SNR situations, it is necessary to find out the variables with the most regular distributions close to Gaussian. Therefore, a special Monte Carlo study is performed to investigate for a range of DTI variables and for different experimental designs how their distributions comply with these basic statistical demands of spatial smoothing. Low and high b value experiments with minimal and multigradient equipment are involved. Most attention is payed to low SNR situations, however, the smoothing strategies found are also convenient for medium and high SNR. Starting with the fields of the DWIs, the random fields of the tensor and that of derived variables like Eigenvalues etc. are investigated for a simple diffusion model with realistic parameters. In case of minimal gradient experiments with constant b values the Rician distributed DWIs offer the most regular field of distributions in all cases studied. For multigradient experiments a more ambiguous situation is apparent. In contrast to other variables, the angle of the main diffusion shows a regular distribution field close to Gaussian even for low SNR due to the improved direction resolution. For replications of minimal experiments with different b values and for multigradient experiments a novel “back” transformation of the tensor, derived from the multitude of experimental DWIs, to a minimal set of virtual DWIs with approximately Gaussian statistics is discussed.

Based on this analysis a new three dimensional nonlinear spatial filter for scalar fields is proposed. Its construction is based on the Aurich chain of sigma filters **[23].** This chain, originally constructed for constant signals seperated by steps and corrupted by Gaussian noise, is modified to include the conditions of discontinous signals with high curvature and with Rician noise. The construction is applicable to the individual spatial DWI fields, but with minor modifications also to fields of other variables. Edges in the spatial fields of the variables are regarded as information about anatomic boundaries between different

fiber bundles or other organic unities. As the DWIs are derived by projections of the diffusion tensor via the gradient directions not necessarily all edges of the tensor field are present in every DWI field. Therefore, the weights of the filter for e.g. a DWI component are extended to incorporate this information preventing wrong information mixing or blurring. The filter application is performed fully automatic, the only input parameter is the standard deviation of noise, which can easily be derived from the background of the images. The filter is fast and both numerically and statistically robust. This filter is validated in low SNR situations by a quasi realistic "gold standard" model, which was achieved on the basis of measured DTI data which were consecutively smoothed by a method published in **[28].** Finally, to illustrate the effectiveness of the filter it is applied to very new data which are measured on an isotropic grid with a voxel volume of 1x1x1 mm^3 and a mean $SNR \approx 3$.

2 Theory

2.1 Noise, Edges and Curvature

To illustrate typical features in DTI data, we present in Fig. 1 some data for a region around the corpus callosum within an axial slice. To reduce noise, three replications of a minimal experiment with the gradients $\{[1,0,1],[-1,0,1],[0,1,1],[0,-1,1],[1,1,0],[-1,1,0]\}/\sqrt{2}$ under clinical conditions (1.5 T, voxel size: 1.9x1.9x4 mm^3) for the b value b=880 sec/mm^2 were performed. For b=0 two replications were applied. After averaging the replicated magnitude signals the tensor is derived via the Stejskal Tanner equation , see Eq. (2). In the left panel of Fig. 1 the tensor field component $d_{11}(\vec{x})$ is presented, see Eq.(1), in the right panel the mean DWI, $\left|S_j(\vec{x})\right|_{mean}$, for $\left|\vec{g}_j\right\rangle = [0,-1,1]/\sqrt{2}$. A closer inspection of the data still shows appreciable noise with spatially varying variance, compare e.g. in d_{11} the yellow valleys of the pyramidal tracts and the green bow of the corpus callosum region in front. Especially the region around the corpus callosum shows steep intensity variations due to the ventricles, see the red-blue maxima of water diffusion in d_{11}. Also in |S| the water diffusion is separated from that in white matter by edges. As in |S| large intensities correspond to small diffusion also the transition from corpus callosum to neighbouring white matter tissue is

frequently stepwise, see e.g. the steep red-blue bow of the corpus callosum DWI in front. This is related to a strong variation of the local anisotropy or Eigenvalues when the corpus callosum is left within white matter. In contrast, also regions with smoothly curved intensities are apparent in both panels. This feature may be more obvious in Fig. 5, where a smoothed DWI is presented.

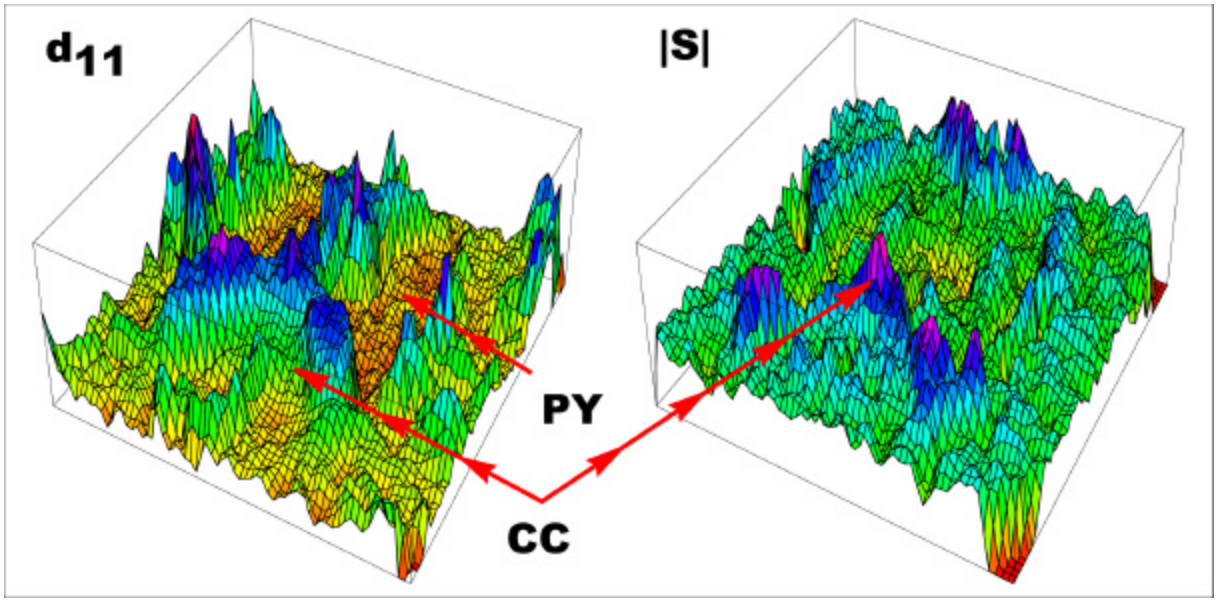

Fig. 1: Left panel shows d_{11} on an axial slice around corpus callosum. Red-blue maxima are mainly due to high diffusion within the ventricles. Right panel shows the mean DWI for the gradient $[0,1,-1]/\sqrt{2}$ in the same region, here the maxima are mainly due to the signal from the corpus callosum diffusion. The data are denoised voxel wise via DWI averaging based on two replications for the b=0 experiment and three replications for b=880 sec/mm^2, red arrows indicate parts of corpus callosum and of pyramidal tract, the few extremal voxels in the corners are outside the brain.

The presence of steep variations or edges in DTI data is further supported by the following observation: According to empirical knowledge about the streams of axon tracks **[24, 25]** it is evident that neighbouring nerve bundles can have strongly different main directions in white matter. This has the consequence that the Eigenvectors in regions with different fiber streams change in a discontinous way. As the tensor is real and symmetric, the tensor field can be formulated as a function of the Eigenvalue and Eigenvector fields $l_i(\vec{x})$ and $|i(\vec{x})\rangle$ in braket notatation, where the ket, $|...\rangle$, is a three dimensional vector, the bra, $\langle...|$, its transpose and $\langle...|...\rangle$ a scalar product:

$$d(\vec{x}) = \begin{pmatrix} d_{11}(\vec{x}) & d_{12}(\vec{x}) & d_{13}(\vec{x}) \\ d_{12}(\vec{x}) & d_{22}(\vec{x}) & d_{23}(\vec{x}) \\ d_{13}(\vec{x}) & d_{23}(\vec{x}) & d_{33}(\vec{x}) \end{pmatrix} = \sum_{i=1}^{3} l_i(\vec{x}) \left| i(\vec{x}) \right\rangle \left\langle i(\vec{x}) \right| \qquad (1)$$

This representation implies that spatial discontinuities or "steep" variations in the Eigenvalues or Eigenvectors map to the tensor field and to other derived quantities like the anisotropy coefficients. They will in general also transform to the DWIs which are connected to $d(\vec{x})$ by the Stejskal Tanner equations, see Eq. (2).

$$\left| S_j(\vec{x}) \right| = \left| S_0(\vec{x}) \right| * f * e^{-b \sum_{i=1}^{3} l_i(\vec{x}) \left\langle i(\vec{x}) \left| \vec{g}_j \right. \right\rangle^2} \qquad (2)$$

Where b is the b value of the experiment, f a volume fraction, which is only for slow diffusion components different from unity, $\left| \vec{g}_j \right\rangle$ a diffusion gradient and $\left| S_0(\vec{x}) \right|, \left| S_j(\vec{x}) \right|$ are the magnitude signals. However, as the so called ADCs , $\sum_{i=1}^{3} l_i(\vec{x}) \left\langle i(\vec{x}) \left| \vec{g}_j \right. \right\rangle^2$, are dependend on the angles between $\left| i(\vec{x}) \right\rangle$ and $\left| \vec{g}_j \right\rangle$, discontinuities of $d(\vec{x})$ are not necessarily present in all DWIs, $\left| S_j(\vec{x}) \right|$. Assume e.g. two diffusions in neighbouring voxels wich change discontinuously their main directions , $\left| 1(\vec{x}) \right\rangle$, but keep the same angle with $\left| \vec{g}_j \right\rangle$. In addition, assume that the two diffusion ellipsoids have identical shapes and are circular in the planes spanned by $\left| 2(\vec{x}) \right\rangle$ and $\left| 3(\vec{x}) \right\rangle$. Rotating for every diffusion $\left| 2(\vec{x}) \right\rangle$ around $\left| 1(\vec{x}) \right\rangle$ into the plane spanned by $\left| 1(\vec{x}) \right\rangle$ and $\left| \vec{g}_j \right\rangle$ produces identical ADC`s for both diffusions and due to $\left| S_0(\vec{x}) \right|$, DWIs which change continuously between both voxels. On the other hand, at least one $\left| S_j(\vec{x}) \right|$ must show any discontinuity which is present in $d(\vec{x})$, else the step could not be resolved by the projections of the gradients.

2.2 The random fields of the DTI variables and spatial smoothing

According to Jensen`s inequality for concave or convex transformations of random variables **[26],** bias effects like those discussed in the Introduction are to be expected for the nonlinearly transformed random variables of DTI. Case studies on the propagation of noise show that bias and variance in DTI variables increase with 1/SNR **[14,19,20]**.

Further, Skare et al demonstrate by model simulations that bias and variance of anisotropy coefficiants can depend on the angles between local diffusion and the diffusion gradients applied **[7].** These results indicate that the random properties of DTI variables should in general be described by locally varying random fields. This can be substantiated by the structure of the Stejskal Tanner equations, see Eq. (2): The exact magnitude signals without noise, DWI_{exact} , depend on the local angles between diffusion and gradients and on the local Eigenvalues or diffusion strengths. This dependence is mapped to the specific DWI distribution within the Rician family **[2]**, determined by the SNR (= DWI_{exact} / standard deviation of noise)**.** Superpositions of the logarithm of noisy DWIs determine in minimal experiments the tensor coefficients and so carry over their local dependence on diffusion and gradients to the statistical properties of derived random variables, like anisotropy. In contrast, a spatially uniform Gaussian statistics for the field of a variable would be optimal for an attempt to apply spatial smoothing methods without running into poor mean value estimates or uncontrollable bias effects. In this context Monte Carlo simulations are performed to quantify the distribution properties in dependence of the SNR. The final aim of this investigation is to find variables which are suitable for spatial filtering in low SNR situations in the frame of different experimental arrangements.

In the following Monte Carlo study noise propagation is simulated within a simplified model. The tensor describes a three dimensional cigar shaped diffusion which rotates with its main axis within the x-y plane of the lab system around the z axis, the tensor has the form:

$$d = \begin{pmatrix} d_{11} & d_{12} & 0 \\ d_{12} & d_{22} & 0 \\ 0 & 0 & \dfrac{d_{11}+d_{22}-\sqrt{(d_{11}-d_{22})^2+4{d_{12}}^2}}{2} \end{pmatrix}, \qquad (3)$$

where the element d_{33} is the smaller eigenvalue of the upper submatrix. As all unkown parameters are contained in the upper submatrix, the simulated measurements, quantified by Eq. (2), need to be performed only in the x-y plane. Consequently, for a minimal experiment only 3 diffusion gradients in this plane are necessary. The simplified model was chosen as neither Eq. (2) nor the tensor diagonalization depend essentially on the effective tensor dimension. Therefore, we can expect that the model is suited to reveal the typical statistical features of DTI variables. On the other hand, due to its simplicity the

model allows a more comprehensive study than more general models. In literature many three dimensional Monte Carlo case studies on the noise propagation in DTI can be found, see e.g. **[7, 13, 14, 19, 20].** They treat experiments with medium and high SNRs. Some of their results which are important for the present investigation are incorporated into the final summarizing discussion below.

For fixed diagonal tensor coefficients d_{11} and d_{22} the off diagonal d_{12} is changed in 40 constant steps within the interval $d_{12} \in \left[-0.8*\sqrt{d_{11}d_{22}}, 0.8*\sqrt{d_{11}d_{22}}\right]$. This interval is defined by the positive definiteness condition of the tensor and covers, without the factor 0.8, all possible off diagonal diffusions. The factor is introduced to limit the ratio of it`s sorted eigenvalues $l_1 > l_2$ to $l_1 / l_2 < 10$, preventing unrealistic high anisotropy. Low and high b value experiments are simulated via the Stejskal Tanner equation (2) . For the first group of experiments b=900 s/mm^2 and f=1 are applied with the diagonal components d11=.00156 mm^2/s and d22=.00084 mm^2/s to achieve a realistic mean diffusity of .00084 mm^2/s **[14].** For b=3500 s/mm^2 the procedure of Clark et al **[21]** was used. There, the slow tensor component is fitted by a monoexponential model to the signal decay within a high b-value interval. The fraction, f, of the slow component is achieved by an extrapolation of high b-value data. The slow diagonal diffusion for the Monte Carlo simulation is defined to be d_{11}=.0007 mm^2/s and d_{22}=.00035 mm^2/s with f= .31. This is in agreement with the fraction and the slow three dimensional mean diffusity of .00035 mm^2/s found in corpus callosum by Clark et al **[21].** For both models, the fractional anisotropy FA per voxel is approximately in the range $FA \in [.4,.9]$, the main diffusion directions scatter within a cone of 70 degrees around [1,0]. The b=0 s/mm^2 signal was assumed to be $|S_0|$=1000.

In recent experimental arrangements the frame of minimal experiments is extended to more gradients. In this context, Jones **[5],** Papadakis **[6]** and Skare **[7]** discuss the propagation of noise from the DWIs to the tensor coefficients for higher numbers and varying directions of gradients. Skare proposes a condition number, $cond$, depending only on the gradient directions, which controls an error in the tensor components due to noise in the signals. In our simulations uniformly distributed gradient directions are applied which were proposed by Jones **[5]** with a rather optimal condition number, $cond = 1.4$, for a plane measurement. All applied gradient sets include the direction $\vec{g}_{12} = [1,1]/\sqrt{2}$ which is used as reference projection in the graphical presentations.

To introduce Johnson noise, noise simulated by a complex Gaussian distribution with standard deviation σ_0 is added to the signals independently for every gradient direction, see e.g. Skare et al or Pierpaoli and Basser **[13, 15]** for a precise algorithmic definition of the procedure applied. Different noise levels with standard deviations σ_0 are involved in the simulations. Experiments for b=900 s/mm^2 are performed with σ_0 = 30*n, with n=1…6. The slow diffusion component was investigated with σ_0=15, 20 and 30. As a SNR=100/3 for the b=0 signal corresponds for a 1.5 T DTI scanner approximately to a voxel size of 2x2x4 mm^3, the voxel sizes investigated vary for b=900 s/mm^2 approximately from 2.5 to 16 mm^3 and for b=3500 s/mm^2 from 16 to 32 mm^3.

The Stejskal-Tanner equations are solved numerically for the tensor coefficients by the singular value decomposition, see **[7]** for details. This algorithm can treat both minimal and overdetermined experiments, where the number of gradients > number of tensor coefficients, with the same b values in all experiments. In case of different b values a multivariate linear regression, proposed by Basser et al **[4.1]** was applied. Finally the tensor was diagonalized to determine the eigenvalues and eigenvectors and to derive from them several variables.

For different random variables, q, measures of bias effects and higher moments, m_n, about the mean, $mean_q$, are discussed and presented in Fig. 2-3. In detail, the relative bias, $(mean_q - q_{exact}) / q_{exact}$, the bias uncertainty, $\sqrt{m_2} / |q_{exact}| = s_q / |q_{exact}|$, the skewness, $m_3 / s_q^{\,3}$, and the normalized standard deviation of noise quantifying heteroscedasticity, $s_q / \max(s_q)$, are presented and discussed for the variables $|S_{12}|$, d_{12}, the eigenvalues $l_1 > l_2 = l_3$, the trace $\frac{l_1 + l_2 + l_3}{3}$ and the fractional anisotropy FA, see Basser **[27].** The angles, α_{exact}, of the main exact Eigenvectors are defined as angles with [1,0] and are in the range $\left[-\frac{p}{2}, \frac{p}{2}\right]$, see below. The direction of a noisy main Eigenvector is defined only modulo π, therefore it is aligned to the exact Eigenvector before it`s angle, α, with [1,0] is calculated. To avoid singularities, the bias is quantified by $(mean_a - a_{exact}) / const$, see Figures, the other moments for q=α are calculated like above. All variables are plotted versus the SNR, $\left|S_{12}^{exact}\right| / s_0$, varying with the diffusion coefficient

d_{12} . The quantity $|S_{12}^{exact}|$ is derived by the gradient $\vec{g}_{12} = [1,1]/\sqrt{2}$ from the model diffusion not distorted by noise. A large number of Monte Carlo simulations was performed to explore the systematics in the distributions, typical results are presented and discussed.

In Fig. 2 A, B, C results of b=900 s/mm^2 experiments with three uniformly distributed gradients in the x-y plane (*cond=1.4*) are presented, Fig. 2D presents a minimal experiment for b=3500 s/mm^2. The moments are derived via Monte Carlo sampling with 50 000 iterations per parameter combination or model diffusion. Results for the voxel volumes 5 and 16 mm^3 are presented in Fig. 2A, C and Fig. 2B, D. The vertical bars at SNR=1.8, 3.8 and 11.3 indicate the transition d_{12} =0 mm^2/s. On the left sides of these transitions the exact main diffusion components increase their angles with [1,0] from 0^0 to 35^0 with decreasing SNR`s, on the right sides these angles decrease from $0^0$ to $-35^0$ with increasing SNR`s. Anisotropy, $FA \in [.4,.9]$, is roughly proportional to the magnitude of these angles.

In the panels a) of Fig. 2 the relative bias of the variables defined above and indicated in the Figure legend is presented. To facilitate the interpretion of the angular bias instead of the relative, the absolute bias is presented in units: *degree/constant*, the value of *constant* is indicated in the Figure. As real experiments are performed frequently with small numbers of replications, the uncertainty or the standard deviation of the relative bias is presented in panels b) of Fig. 2 A, B, D. In case of the angle the quantity *standard deviation(angle)/const* is presented. As with respect to spatial smoothing, symmetry and the homogeneity of variance are important, in panels c) skewness and in panels d) the heteroscedasticity is presented in Fig. 2 A, B, D. These plots present the lower moments of locally varying random fields for the variables, when different model diffusions, d_{12}, are identified with differently localized voxels. In Table I, for the sake of transparency, a numerical evaluation of these plots is presented. For the columns of variables the ranges of the plotted curves are given by : "minimum / maximum". Three consecutive rows in the Table correspond to one panel as indicated, these rows are ordered according to the labels A,B,D of Fig. 2.

Due to the convenient condition number for the gradient set the relative bias effects in Fig. 2A are quite small, see the first row in Table I. For a different gradient set, e.g.

{[1,0],[0,1],[1,1]}/$\sqrt{2}$, the bias in all variables is nearly doubled. However, the bias uncertainty, see 4.th row, deteriorates single measurements appreciably. The maximum uncertainty for FA is close to 50% and that of the angle close to 25°, which would produce severe errors in tracking algorithms, see Anderson **[19].** In Fig 2B, see second and 5.th row, both are reduced due to the enlarged SNR. In all variables of Fig. 2A, except $|S_{12}|$, skewness is appreciable and can become larger than that of a Rayleigh distribution, which has an approximate skewness of 0.6, see row 7. In Fig. 2B, see 8.th row, the skewness of d_{12}, λ_1, α and Trace comes close to 0.6. High skewness enlarges the risc of a non representative or outlying measurement appreciably, therefore spatial filters usually presuppose symmetric distributions. Constant variance is a further prerequisite of unbiased spatial smoothing. Heteroscedasticity in Fig. 2B can become even larger than that in Fig. 2A, see rows 10 and 11, the exception is again $|S_{12}|$.

Minimal experiments for b=3500 s/mm^2 were performed for the voxel volumes 16, 24 and 32 mm^3 . See Fig. 2D and Table 1 for results with the voxel volume 16 mm^3 . Due to the enhanced b value and the weighting factor the SNR scale is reduced compared to that of Fig. 2B and is close to that of Fig. 2A. The results in Fig. 2D and Fig. 2A show similarities in the uncertainty, skewness and heteroscedasticiy, see rows 4,7,10 and 6,9,12. The relative bias is however appreciably enhanced compared to Fig. 2A, see rows 3 and 1.

	$\lvert S_{12}\rvert$	d_{12}	l_1	l_2	*FA*	*a*	*Tr*
a)	.01/.2	−∞/+∞	−.01/.06	−.2/−.04	−.02/.18	−.18/.16	−.01/0
	0/.02	−∞/+∞	0/.01	−.02/−.01	0/.02	−.02/.04	0/0
	0/.8	−∞/+∞	−.3/−.03	−.15/.5	−.2/.08	−.78/.4	−.2/−.05
b)	.1/.6	.5/+∞	.29/.31	.3/1	.15/.5	.3/.8	.15/.2
	.05/.2	.08/+∞	.08/.1	.1/.3	.05/.2	.08/.23	.05/.07
	.2/1.2	.3/+∞	.2/.3	.3/.7	.15/.45	.4/1	.17/.18
c)	0/.3	−.8/1.3	1.3/1.7	−.8/.1	−.6/.1	−1/.1	.8/1.
	.01/.01	−.3/.6	.22/.62	−.35/.2	−.2/.05	−.4/.42	.2/.5
	0/.6	−.6/.8	1/1.3	−.4/.15	−.52/.15	−.5/.15	.78/.82
d)	.88/1	.55/1	.65/1	.8/1	.45/1	.35/1	.82/1
	.99/1	.48/1	.54/1	.7/1	.4/1	.35/1	.72/1
	.68/1	.68/1	.94/1	.65/1	.5/1	.38/1	.95/1

Table 1 : Ranges of the curves presented in Fig. 2A, B, D. Minimum/maximum is given, else see text.

The presentation for the 5 mm^3 and the 16 mm^3 voxel simulations in Fig. 2 A, B, D, suggests the following conclusion : At low and mean SNR (nearly) every variable shows the feature of a locally varying random field. The deviations from Gaussian distributions increase with 1/SNR, the quantitative ranges of bias, skewness and heteroscedasticity prevent reasonable spatial smoothing without controled corrections. This conclusion is further supported by results for other voxel volumes which are not presented graphically.

Only in case of the DWIs the distributions are known and simply parametrized via SNR **[2].** As is outlined above, for a general gradient arrangement the variables derived from the DWIs mix information of the logarithm of different noisy DWIs which leads to families of distributions with highly complex dependence on basic parameters, like local diffusion and gradient properties. In contrast, as the DWIs are distributed according to Rice, via the "SNR" of the mean signal, MNR, the true signal can be estimated **[2],** i.e. the bias in any DWI can in principle be corrected. For all other variables investigated neither an explicit shape of the distributions nor any bias correction is known for mean to low SNRs.

Concerning skewness, above $SNR \approx 4$ the DWIs are symmetric according to Rice; below $SNR \approx 4$ the maximal skewness of $|S_{12}|$ in Fig. 2A is well below the corresponding maxima of the residual variables. In Fig. 2D this maximum is at least below those of λ_1, Trace and d_{12}. Heteroscedasticity is minimal for $|S_{12}|$ in Fig. 2A, in Fig. 2D λ_1, FA and α are more uniform. However, as in DWIs the variance scales with MNR, varying variance can be incorporated approximately into spatial filters for the DWIs as will be shown below.

Concluding, in low SNR situations the DWIs should be preferred for spatial smoothing, all other variables should be derived from these true value estimations. Even in situations with medium SNRs, see Fig. 2B, the regular Gaussian distributions of the DWIs with trivial bias shift make them to the preferable candidates for spatial smoothing.

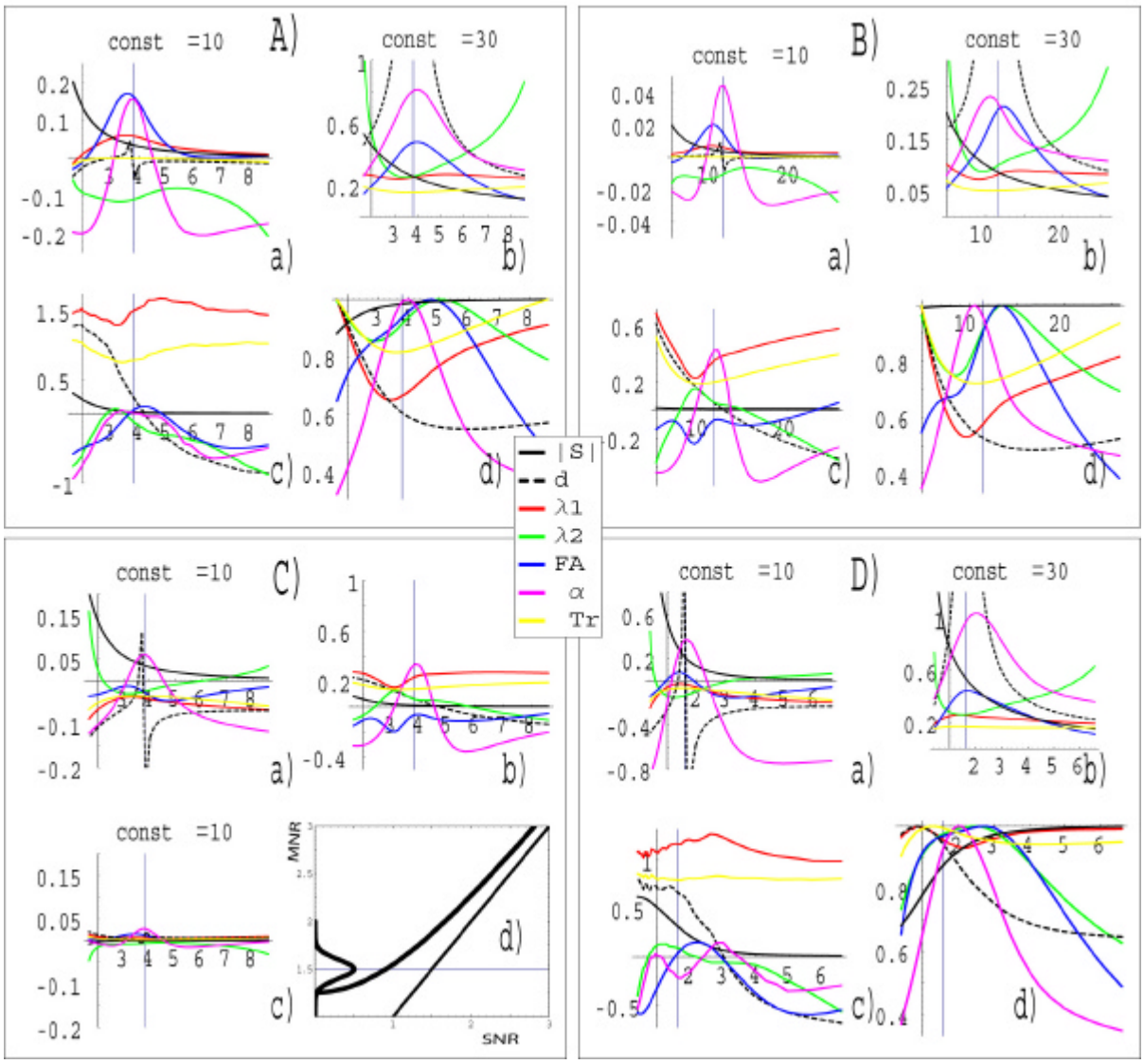

Fig. 2: Results of Monte Carlo simulations for b=900 s/mm^2 and voxel volumes of 5 mm^3, Fig. 2A, and of 16 mm^3, Fig. 2B. The relative bias (panel a), the uncertainty of relative bias (panel b), the skewness (panel c) and the normalized standard deviation or heteroscedasticity (panel d) for different diffusions, d_{12} , scaled by SNR=$|S_{12}^{exact}|/\sigma_0$ along the horicontal axes are presented. Results for the variables $|S_{12}|$, d_{12}, λ_1, λ_2, FA, α and Trace are shown, coloring is indicated in the legend. In case of the angular α , the absolut bias and standard deviation normalized by a constant given is plotted. Fig. 2D presents a simulation with b=3500 s/mm^2 and a voxel volume of 16 mm^3. Fig. 2C presents an experiment like in Fig.2A, but for n_R=15 replications. The random variables are derived from the averaged magnitude signals $|S_i| = \sum_{j=1}^{n_R} |s_{i,j}| / n_R$, i=1,2,3. Panel a) shows the relative bias, panel b) skewness; panel c) the relative bias after bias correction. The plot in panel d) characterizes the relation between the mean to noise ratio, MNR, and the SNR according to Rice. For the SNR the range $SNR \in [0, 2.8]$ is shown. The distance of this function to the diagonal straight line characterizes the Rician bias, the Gaussian distribution centered at MNR=1.5 illustrates the impact of uncertainty on the bias correction.

It is recommendable that voxel wise and spatial smoothing are applied complementary. As the edge resolution of nonlinear spatial filters depends on the noise level, see below, voxel wise denoising should precede spatial smoothing, when possible. Two variants of voxel wise denoising are in the trade. First, experiments with low numbers of gradients are repeated to reduce noise by voxel wise averaging **[*19*].** In the second variant, multigradient experiments **[*5,6,7*]** are applied. Concerning the first variant for constant b values, we assume that the experiments discussed above are repeated n_R times. The naive transition to individually averaged random variables (e.g. $\tilde{FA} = \sum_{i=1}^{n_R} FA_i / n_R$) reduces their uncertainty and approximately also skewness by a reduction factor, $1/\sqrt{n_R}$, but not bias and heteroscedasticity. When, however, averaging is performed only on the DWIs and when all residual variables are derived from these averaged magnitude signals we find that also bias is reduced in all variables. These results are only slightly modified when, instead averaging, the singular value decomposition on the $3*n_R$ DWIs is applied to derive voxel wise an averaged tensor. For DWI averaging with n_R =15 relative bias and skewness for an experiment with b=900 s/mm^2 and a voxel volume of 5 mm^3 are presented in panels a) and b) of Fig. 2C. The uncertainty of Fig. 2A is reduced by $1/\sqrt{n_R}$, heteroscedasticity remains unchanged. The bias effects shown in panel a) of Fig. 2C can be further reduced. Assume, we have after 15 replications reasonable estimates of $|S_{mean}|$, then the Rician bias correction relating MNR=$|S_{mean}|/\sigma_0 \sqrt{n_R}$ to SNR=$|S_{exact}|/\sigma_0$, see Fig. 2C panel d), can be applied via e.g. tabulated interpolation. All variables which are derived from these estimates of $|S_{exact}|$ should then asymptotically, like $|S_{exact}|$ itself, be essentially unbiased. See the relative bias in panel c) of Fig. 2C after an application of this bias correction. A precise bias correction is limited by the reduction factor, $\sqrt{n_R}$, of the standard deviation of the mean value distribution , $\sigma(|S_{mean}|)=\sigma/\sqrt{n_R}$, see Fig. 2C panel d). In case of low SNR`s and too small n_R this uncertainty can create poor mean value estimates which correspond to undefined signals with |S|<0. For a $SNR \approx 1$ at least a factor n_R=15 is necessary, see Fig. 2C panel d). The uncertainty of $|S_{mean}|$ can be reduced by additional spatial smoothing which produces an effective rise of n_R . To estimate the improvement, assume for simplicity strict homogenuos diffusion within the n_F-1 nearest voxels surrounding the voxel of interest in the signal and perform unweighted spatial

averaging on the mean value of the DWIs based on n_R experimental replications. The results would be equivalent to those for $n_{effective}=n_R*n_F$ experimental replications.

In another type of data acquisition, minimal experiments are repeated with different b values. This arrangement goes back to Basser **[4.1].** The system of Stejskal Tanner equations is solved for the mean tensor by multivariate linear regression, where a special weighting via the reciprocal error variance of the DWIs is proposed **[4.1].** Several simulations for different b value combinations with maximum b=900 s/mm^2 were performed. We found in all cases more regular distributions for the tensor fields when the regression was performed without the weighting proposed, the introduction of the weights increased in all cases bias and skewness. An example for a voxel size of 5 mm^3 and values b_1=500, b_2=700 and b_3=900 s/mm^2 without weighting is presented in Fig. 3A, the corresponding tensor components d_{11}, d_{12} and d_{22} are shown. To perform additional spatial smoothing a transformation to virtual DWIs based on this mean tensor is convenient, as the noise levels in these DWIs is reduced like in the mean tensor. This idea was first proposed in the Proceedings of MICCAI 2001 by Hahn et al **[28].** A transformation to virtual DWIs via Eq.(2) with $b=(b_1+b_2+b_3)/3$ and the same gradients like in the "real" experiment produced in all cases which were studied DWI distributions with essentially better smoothing prerequisites than those of the tensor components, see Fig 3A for a typical comparison between tensor and virtual DWIs.

In case of a multigradient experiment with a constant b value the multitude of DWIs reduce noise voxel wise in the tensor calculated via the singular value decomposition. In addition, for every diffusion direction some of the gradients will project ADC`s with relatively high magnitude signals or high SNR's thus reducing bias effects. The multigradient approach for the diffusion model of Fig. 2A is illustrated in Fig. 3B, where 45 uniformly distributed gradients with *cond=1.4* are applied. The number of DWIs agrees with that of the experiment presented in Fig. 2C. The relative bias in panel a) is strongly reduced, compare panels a) in Fig. 2A,C and Fig. 3B. Skewness is in general comparable with that of the replicated minimal experiment, see panels c) in Fig. 2C and Fig. 3B, however the skewness of the angle is essentially reduced due the high angular resolution. A low relative bias could be achieved already by application of only 15 uniformly distributed gradients, however, for low skewness 45 gradients are necessary. The second moments, see panels b) and d), in Fig. 3B are close to those of the experiment presented in Fig. 2C, they are mainly determined by the number of DWIs.

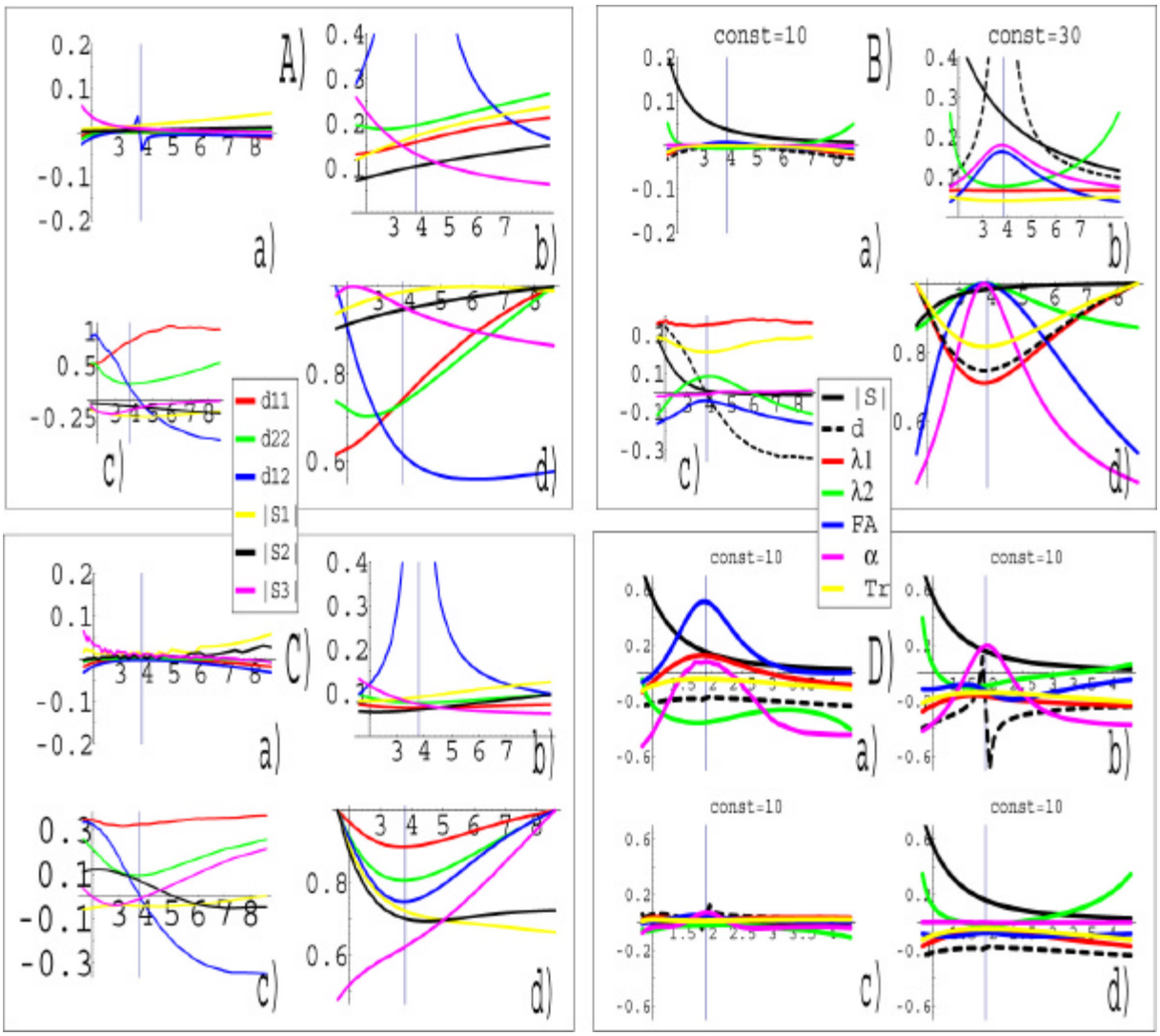

Fig. 3: Fig. 3A presents the tensor components for an experiment with b=700, 800 and 900 s/mm^2 and a voxel volume of 5 mm^3, for more details see text. Else the virtual DWIs , inverted with b=700 s/mm are shown. Fig. 3B presents the variables indicated in the legend for a multigradient experiment with 45 gradients, b=900 s/mm^2 and a voxel volume of 5 mm^3. Fig. 3C shows the tensor components derived by singular value decomposition and the virtual DWIs inverted with 3 gradinets and cond=1.4. In Fig. 3 A-C panels a)-d) show relative bias, uncertainty, skewness and heteroscedasticity. Fig. 3D presents in all panels the relative bias for simulations with very low SNR$\in [1,5]$, with b=900 s/mm^2 and a voxel volume of 2.5 mm$^{3.}$ Panel a) shows results for a minimal gradient set (cond=1.4) with 1 replication, panel b) for 15 replications, panel c) like b) but with consecutive bias correction. Panel d) presents multigradient experiments with 45 and more gradients (cond=1.4).

With respect to further spatial smoothing, the situation is ambiguous. Noise is only reduced in the tensor components and in the derived variables. Though their bias is practically zero the further conditions for mean value estimation via spatial smoothing are at least not optimal. Skewness for d_{12} and λ_1 can become effective, heteroscedasticity for α, FA, λ_1 and d_{12}, see panels c) and d). On the other hand, the relative noise level or inverse SNR of the individual variables, see panel b), is low, so that the filter bias caused by smoothing of any variable may be of second order.

The situation could be improved in principle by application of more gradients or by some replications of the whole experiment. An introduction of virtual DWIs with low level noise, similar to the method applied in multi b value experiments above, see Fig. 3A, only partially improves the situation. To demonstrate this, the tensor resulting from the singular value decomposition for the 45 DWIs is transformed to 3 DWIs by a virtual experiment with only 3 gradients, with *cond=1.4* and the same b value as in the "real" experiment. These DWIs have typically a low noise level, see panel b) in Fig. 3C, are only slightly biased like the tensor coefficients, see panel a), are more symmetric than the tensor, see panel c), but increase heteroscedasticity compared to the tensor, see panel d). possibly this would be reasonable prerequisites for an adaptive filter, like that of Polzehl **[9].**

To study the situation for very low SNR, in the panels of Fig. 3D the relative bias for a b=900 s/mm^2 experiment with a voxel volume of 2.5 mm^3 is presented, the corresponding SNR range is approximately 1-5. In panel a) one minimal experiment with *cond=1.4* is applied, the corresponding bias reduction in panel b) is due to 15 replications and DWI averaging. This residual bias is further reduced, see panel c), by the bias correction described above. Performance of a multigradiend experiment with 45 uniformly distributed gradients, *cond=1.4*, produces a less perfect bias reduction, see panel d). Remarkable is the fact, that this bias could not be further reduced by increasing the numbers of gradients.

In experiments with low SNR noise can produce large fractions of voxels where the tensor violates positive definiteness and can no more be interpreted as a quantity describing diffusion. Especially in regions with large ansiotropy the smaller eigenvalues are close to 0 and frequently become negative by noise distortions. It was proposed by Ahrens et al **[29]** to constrain the Eigenvalues positive within the least square algorithm which derives voxel wise a mean tensor. The presented combination of voxel wise and spatial

smoothing offers an alternative solution to this problem. In the simulations of experiments with b=900 s/mm^2 and voxel sizes 2.5, 3.3, 4, 5.3, 8 and 16 mm^3, the fractions of negative definite tensors were .23, .19, .15, .11, .05 and .03 when situations with a true diffusions FA>=0.8 were considered. After voxel wise denoising of the DWIs without any discrimination between "negative" or "positive" voxels by 10 (effective) replications, these fractions reduced well below 10^{-3}. A comparable reduction was achieved by application of multigradient experiments with 10 gradients.

SUMMARY: The aim of the presented study was to explore the possibility of spatial filtering on DTI variables in low SNR situations. It was exemplified, that for low SNR, the distributions of the DTI variables are those of spatially varying non Gaussian random fields which in general conflicts the basic demand of spatial smoothing for uniform noise close to Gaussian. This feature is characteristic for the different experimental arrangements which were simulated. This comprises the following arrangements: minimal experiments with low and high b values, replications of such experiments with constant and different b values and multigradient experiments. The local variability of the distributions depends in a complex way on basic system parameters like e.g. the angles between local diffusion and the measuring gradients of the individual experiment. According to the presented analysis and in agreement with earlier case studies published in several proceedings articles by Hahn et al **[28, 30, 31],** it seems that for low SNR only the DWIs offer a possibility to bring the advantages of spatial smoothing into the game. Their Rician distributions are well understood, scale with the experimentally well defined MNR, offer a simple bias and variance correction and smoothly fuse with increasing SNR into Gaussian distributions. These properties are not the optimal prerequisites for spatial smoothing, but it is possible to adapt nonlinear filters to DTI in such a way, that a reasonable quality for spatial DWI denoising can be achieved, see next chapter. For DWIs even with very low SNR a good bias reduction could be derived for model situations, see Figures 2C and 3D, where replications of minimal experiments with constant b values are coupled to ideal and simplified spatial smoothing, postprocessed by a bias correction. More realistic validations of this strategy will be presented in the next chapters. Note, that voxel wise denoising by experimental replications is an important first step as the quality of the edge resolution in spatial smoothing depends on the noise level. Sometimes such replications are performed with different b values which, by multivariate regression, lead to voxel wise denoised tensor coefficients. As is exemplified in Fig. 3A

and by numerous simulations within the present model, the prerequisites for spatial smoothing can be improved considerably by a backtransformation to virtual DWIs.

Recently discussed multigradient experiments pose a more subtle situation. Such arrangements reduce bias effects in all variables, except DWIs, for low SNR quite effectively with increasing numbers of gradients, see Fig. 3B. But the higher moments are still not optimal for further spatial smoothing. In very low SNR situations however, see Fig. 3D panel d), these bias effects could not be arbitrarily reduced by an increasing number of gradients. Due to the high angular resolution in this arrangement the angle is an exception and has an unbiased and symmetric distribution which may be, together with the reduced noise level, sufficient for spatial smoothing. A backtransformation of the tensor coefficients to virtual DWIs, see Fig. 3C, improves the situation only partially, as heteroscedasticity is increased compared to the tensor.

In data with low SNR appreciable fractions of the voxels violate the condition of positive definiteness for the tensor, which is a prerequisite for its diffusion interpretation. Our model simulations indicate, that already for a moderate number of effective replications or multigradients the situation is improved essentially due to the convenient bias reduction in the Eigenvalues. In this context the possibility of a sorting bias for noisy Eigenvalues is frequently discussed in model simulations, see e.g**. [13, 4].** In our simplified model this effect could not be studied as the diffusion is effectively two dimensional and therefore the Eigenvalues are always sorted correctly. However, the mentioned bias reduction should also reduce these effects in true three dimensional situations.

For medium SNR the results presented in Fig. 2B exemplify that the proposed strategy for spatial smoothing is still convenient. According to Rice, the DWI distributions are, among all variables, optimal for voxel wise and spatial smoothing. This conclusion is supported by a different investigation from Anderson **[19]** who studied noise effects on bias and variance in Eigenvalues and Eigenvectors for SNRs from approximately 20 to 100 by perturbation theory and simulations. For SNRs around 20 averaging of DWIs by experimental replications is recommended to reduce noise effects. For higher SNRs Anderson's results imply an equivalence between DWI and tensor averaging. This is in line with a proposal by Pajevic et al **[12]** who apply B-splines to obtain a continuous

representation of the tensor coefficient fields for high SNR data. This method can be regarded as a special variant of spatial smoothing applied to the tensor fields.

2.3 A nonlinear Filter

Nonlinear spatial smoothing or edge preserving regularization is still an active field of research in statistics and image analysis, see Winkler et al **[8]** for a review and for a foundation of such filters on Bayesian concepts. We propose to apply a chain of Sigma filters, which goes back to Aurich **[23],** and estimates the mean of a noisy scalar intensity function $f(x),\ x \in R^n,\ n =$ dimension of space**.** The chain iterates nonlinear filters which combine a spatial window, Φ, and an intensity window, Ψ. The corresponding parameters vary according to special rules by which edge detection and generalization or smoothing quality are balanced. After the presentation of the standard version, several new modifications which are convenient for spatial DTI filtering will be introduced.

One filter step is defined by :

$$F_h^m \circ f(x) = \frac{\sum_{y \in Neighborhood\ of\ x} \Phi(x,y)\Psi(f(x), f(y))\, f(y)}{\sum_{y \in Neighborhood\ of\ x} \Phi(x,y)\Psi(f(x), f(y))} \qquad (4)$$

with $\Phi(x,y) = e^{-(x-y)^2/2h^2}$

and $\Psi(f(x), f(y)) = e^{-(f(x)-f(y))^2/2m^2}$

it`s k iterations are defined by :

$$f_k^{smooth}(x) = F_{h*c^{k-1}}^{m*d^{k-1}} \circ \ldots \circ F_{h*c}^{m*d} \circ F_h^m \circ f(x) \qquad (5)$$

The number of iterations, k, the spreading factors of the windows, c and d, and the initial width of the spatial window, η, are regarded as parameters. The initial intensity window width is m=3*$standard\ deviation\ of\ noise\ on\ f$.

To enable an adaptive application of the filter for the reader, a short presentation of the basic parameterization strategy will be given. Regard an intensity step H in an elsewhere constant function of space which is corrupted by Gaussian noise with standard deviation σ. The edge to noise ratio, ENR, is given by

$$ENR = H / 2s \qquad (6)$$

Iterating the chain k-times reduces σ by $1/a^k$ and allows for large samples an edge resolution down to

$$ENR = 1/a^k \qquad (7)$$

where $\alpha > 1$ is a speed factor determining the parameters c, d and η via :

$c = \sqrt[n]{a^2}$, $d = 1/a$ and $h = \frac{\sqrt[n]{4a^2}}{2\sqrt{p}} * smallest\ gridwidth$ where $n = \dim ension$ of space.

Frequently k ist chosen such high that the filter reaches its fixed point, i.e. that further iterations do no more change the results. Due to the nonlinearity of the chain, this parameterization could be justified only for stepwise constant signals with Gaussian noise, see Mühlhaus **[32]** for more details. As DTI magnitude signals combine discontinuities with curvature $\neq 0$, possibly hide edges of the diffusion tensor field, see chapter 2.1, and as their noise is Rician several modifications of this basic procedure are introduced.

In Rician distributions, for low SNR $s = s(MNR)$, therefore a heteroscedastic μ=3∗σ(*MNR*) is introduced into Ψ. Increasing k and α increases the edge resolution capacity, but, favoring estimates with curvature=0, it reduces the ability to approximate curved functions. Therefore a convenient balance suited for DTI data must be found by data or model validation. A further improvement of the generalization properties can be achieved by the use of $f(x)$ instead of $f_{k-1}^{smooth}(x)$ in the last iteration, see Winkler et al. **[8].** As the present filter produces finally an estimate of the mean value of the magnitude signal, the Rician bias correction should be applied for low SNR`s to the resulting $f_k^{smooth}(x)$.

In Fig 4 applications of this modified filter chain to simulated one dimensional noisy DWIs are presented in panels a), b) and c). Denoising of functions with two different shapes is

performed with the same parameters k=4 and $a = \sqrt{2}$ to demonstrate robustness of the method. The intermediate and final filter results of the iterations are shown. The final mean value estimate of the chain (yellow) is bias corrected (violet), the corresponding true model is indicated by black lines. In panel a) a sample size of 500 points is applied, in panel b) the size equals n_Rx500. Note the improvement of generalization for n_R=4. Practically the same results as in panel b) are achieved, when a sample of 500 points after voxel wise DWI averaging by n_R=4 is applied. Discontinuities are already well reproduced in panel a). Panel c shows an application to step wise constant signals. The transition to n_R=4 , not shown, improves the generalization again but not the resolution of the step at ENR=1/2. To achieve that, an essentially higher n_R would be necessary. In panel d) a linear Gaussian filter is applied to the n_R=4 situation. The blurring effects at the edges are inherent to linear spatial filters.

Due to the robustness of the filter construction, the weights can be further modified to model peculiarities of DTI data. E.g., the spatial windows, Φ, which are isotropic Gaussians in Eq (4), can be adapted in their shape to the diffusion ellipsoid or tensor of the central voxel, d(x), by polarized Gaussians.

$$\Phi(x, y) = e^{-(x-y)^T \Lambda^{-1}(x)(x-y)/2} \qquad (8)$$

The covariance matrix, Λ(x) , is derived from the raw or partially smoothed tensor, $d(x)$, by

$$\Lambda(x) = d(x)\ h^2 c^{2k-2} / \det(d(x))^{1/n} \qquad (9)$$

where the scaling factor is chosen in such a way, that the volumes of isotropic and anisotropic spatial windows are identical for every iteration, k, to conserve the maximal sample size. This modification improves the isotropic window of Eq. (4) in situations were the local diffusion ellipsoid reflects the alignment of the neighbouring fibers, like e.g. in the pyramidal tract.

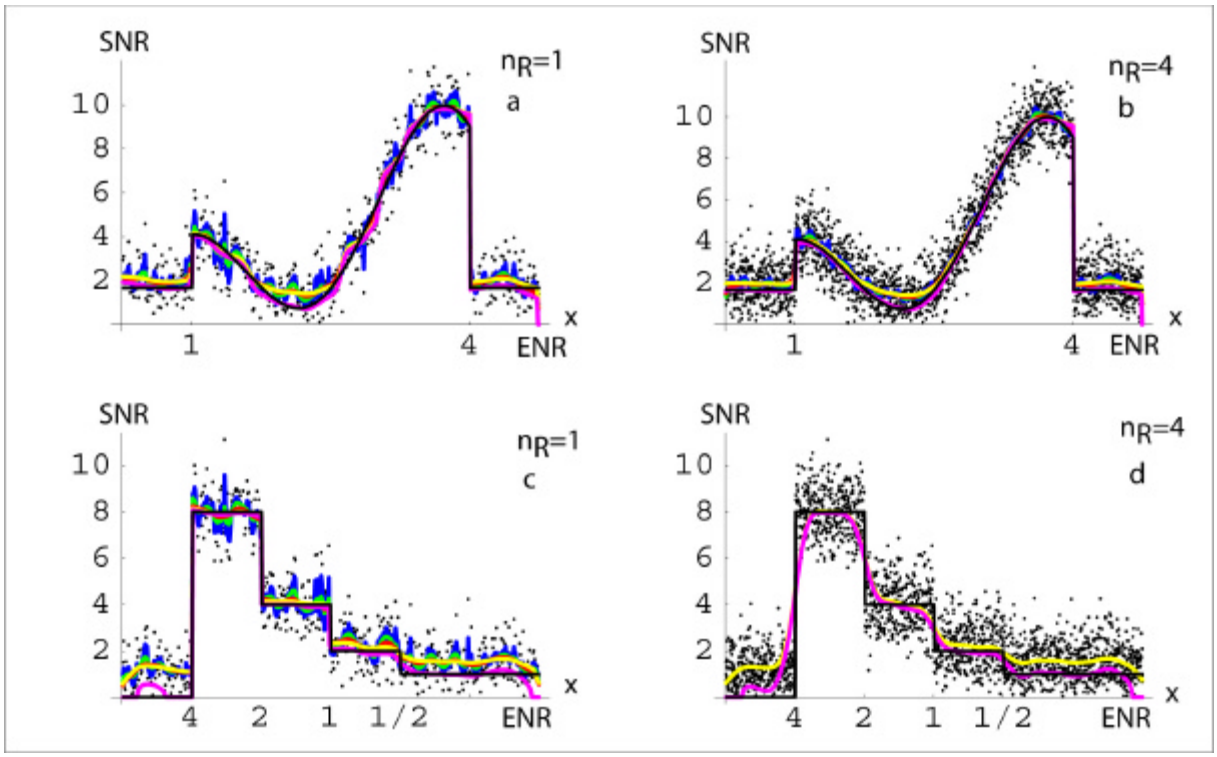

Fig.4: Denoising of magnitude signals is illustrated. On the data points of panels a),b),c) the nonlinear filter chain is applied. The results of the consecutive iterations (k=4) are given by blue, green, red and yellow curves, the final bias corrected approximation is violet. True signals are indicated by black lines. The individual graphs give signal to noise ratios (SNR) versus space coordinate (x) and edge to noise ratios (ENR) at discontinuities. In panel d) a linear Gaussian filter followed by the bias correction is applied. The sample sizes are n_Rx500.

It was demonstrated in chapter 2.1 that DWIs can hide edges which separate different information units. To reduce this possible cause of blurring also the intensity windows can be modified. For example one could use those of Eq. (10), instead of those in Eq. (4),

$$\Psi(f(x),\, f(y), \mathrm{var}) = \Psi(f(x),\, f(y)) * Min[e_1, e_2, e_3, e_4, e_5, e_6]$$

$$with \quad e_i = 1 - 1/(1 + e^{-b\,(|\mathrm{var}_i(x) - \mathrm{var}_i(y)| - step)}) \qquad (10)$$
$$b = \text{steepness of sigmoid } e_i$$
$$step = \text{edge defining distance}$$

for the functions var_i e.g. the DWIs can be used. For convenient steepnes, β, the function $Min[e_1, e_2, e_3, e_4, e_5, e_6]$ is close to unity for $|\mathrm{var}_i(x) - \mathrm{var}_i(y)| < step$ and close to zero else. The modification in Eq. (10) introduces information about the localization of the edges in all DWIs when a special DWI is denoised.

In both modifications information from smoothed data is convenient. Therefore, in realistic situations a two step procedure should be performed: first, apply the filter chain with Eq. (4), in a second run use Eq. (8) and (10) where $d(x)$ and $\mathrm{var}_i(x)$ from the first run are involved.

To cover the whole brain, in DTI experiments usually anisotropic voxels with an enlarged grid in axial direction are applied. It would be more convenient for spatial smoothing to use isotropic voxels which increase the effective space dimension, n, and lead to an improvement of the generalization quality. This is due to the more symmetric grouping of the data enabling a reduced spreading or better localization of the spatial windows, see c and h .

2.4 Some Results

To validate the filter and the strategy proposed in chapter 2.2 a quasirealistic model is introduced. It is based on an experiment with 1.5 T, using a diffusion weighted EPI sequence with 6 noncolinear gradients and 4 b-values, b=400, 500, 700 and 880 s/mm^2 . The voxel size is 1.9x1.9x4 mm^3 to cover the whole brain, see **[28]** for further details. The smoothed and postprocessed tensor is used as “gold standard” for further filter tests. In a region around the corpus callosum of the size 71x41x12 voxels the model tensor is transformed to the DWIs by Eq. (2), noise is introduced like in chapter 2.2 and three dimensional versions of the filter are applied in this region. To measure the difference between model and noisy or smoothed DWI the spatially averaged standard deviation with respect to the model is calculated and presented in Table 2. To exclude boundary effects of the filter only 5 axial slices are involved in this analysis, their middle slice is presented in Fig. 5. Also voxels with negative Eigenvalues and within the ventricles are omitted from the error analysis, see black voxels in left panel of Fig.5.

At present, the filter is tested in two variants. First, Rician noise and the bias correction are introduced and coupled with Eq. (4), see σ_{M1} in Tab. 2.

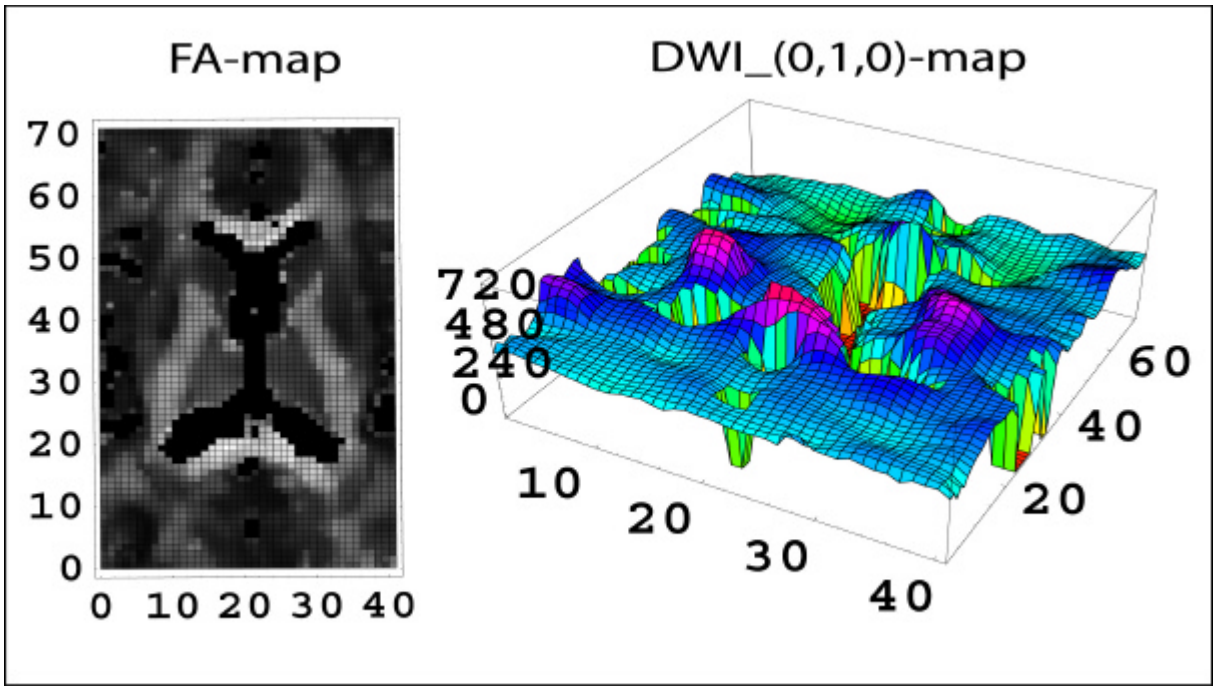

Fig. 5: Left panel: for orientation the FA-map for an axial slice of the gold standard model without noise is presented; voxels within the liquid and with negative Eigenvalues are excluded (black). Right panel shows a corresponding DWI-map for the gradient [0,1,0]; excluded voxels are projected to zero.

Then, Eq. (4) is replaced by Eq. (8) and Eq. (10) based on $d(x)$ and $\mathrm{var}_i = DWI$ of the gold standard model, see σ_{M2} in Tab.2. The noise levels σ=60, 90 and 120 are applied to cover medium and low SNR situations. In Tab. 2 a summary of the filter tests for every DWI is presented.

	σ=60, σ_{M1}	σ=60, σ_{M2}	σ=90, σ_{M1}	σ=90, σ_{M2}	σ=120, σ_{M1}	σ=120, σ_{M2}
DWI_(1,0,0)	27	23	36	31	42	41
DWI_(0,1,0)	28	24	36	32	44	41
DWI_(0,0,1)	29	24	40	32	47	41
DWI_(1,1,0)	29	23	35	31	42	41
DWI_(1,0,1)	29	23	37	31	43	41
DWI_(0,1,1)	30	24	37	31	44	41
Mean n_R	4.4	6.5	6	8	7.5	8.6

Table 2: Spatial means of the deviations between true and noisy (σ) or true and smoothed DWIs (σ_M) . Two filter variants are applied : M1={k=3, α=2, Eq. (4), Rician variance, bias correction}, M2={M1, but instead of Eq. (4) Eq. (8) and Eq. (10) based on the true model are applied}. In the last row mean numbers of experimental replications which produce equivalent denoising effects are given.

As is evident from the last row in Table 2, spatial filtering is quite effective and can replace on average 4-8 replications of experiments in the SNR region investigated. The improvement

between σ_{M1} and σ_{M2} is mainly due to Eq. (10) which reduces blurring between anatomically separated regions.

An application of the filter to very recent low SNR data is partially presented in Fig. 6. The experiment was performed by members of a group guided by P.Narayana **[33].** The data were achieved by a 1.5 T scanner, with a b value of b=1000 sec/mm^2 , the voxel size is 1x1x1 mm^3 . The data cover a volume of 256x256x28 slices. The gradients applied are {[0,.526,-.851], [0,-.526,-.851], [.526,-.851,0], [.526,.851,0], [-.851,0,.526] and [-.851,0,-.526]}, only one replication for b=0 and b>0 is available up to now. The local SNR of the data is partially very low, the spatial mean SNR is approximately $SNR \approx 3$. To avoid the introduction of approximation artifacts the three dimensional filter was applied to the DWIs in the version based on Eq. (4), k=3 and α=2. Rician heteroscedastic variance and the bias correction are included. The FA maps in Fig. 6 demonstrate the high efficiency of the filter for isotropic grids even at very low SNR. The structures in the right panels are essentially in agreement with anatomic knowledge. It is evident from the Figures that the results could not have been achieved by a direct nonlinear smoothing of the raw FA maps, shown in the left panels. Also, an enormous bias would be the result of such a strategy.

The present investigation is still somewhat preliminary. More test results are desirable. As Eq. (8) and Eq. (10) improve filtering of different features in DWI data an investigation of their specific properties would be interesting. Also the effectiveness of Eq. (10) without any model knowledge remains to be tested. Further, is the bias correction a critical point. When the mean value estimate is below the -in practice- unknown true value, due to small non representative noisy samples in the filter windows, the bias correction can underestimate the true signal appreciably. Convenient strategies are to be worked out.

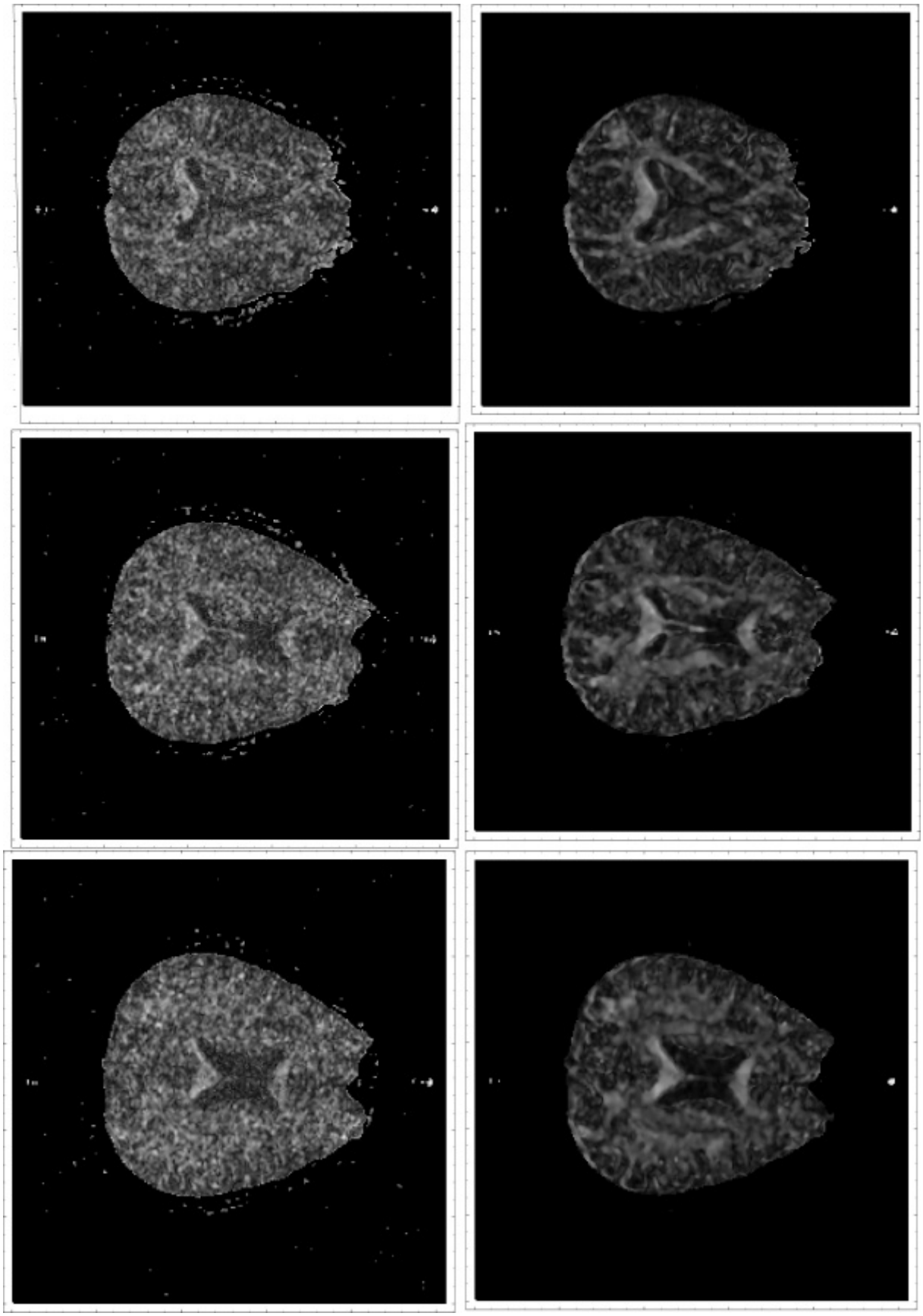

Fig. 6: Axial FA-maps of the data with isotropic voxels (see text), voxel volume=1x1x1 mm^3. Left panels, FA of raw data, right panels, FA based on smoothed DWIs, the filter process is performed with heteroscedastic Rician variance and bias correction and Eq. (4).

References:

1) M.Henkelmann, Measurement of signal intensities in the presence of noise in MR images,Med.Phys. 12/2, 232-233, 1985.
2) H. Gudbjartsson, S. Patz, The Rician Distribution of Noisy MRI Data, Mag. Res. Med. 34, 910-914, 1995
3) O. Dietrich, S. Heiland, K. Sartor, Noise Correction for the Exact Determination of Apparent Diffusion Coefficients at Low SNR, Mag. Res. Med. 45, 448-453, 2001
4) P.J. Basser, D.K. Jones, Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review, NMR Biomed. 15, 456-467, 2002.

4.1) P.Basser, J.M. Mattiello, D.L.Bihan, Estimation of the effective self-diffusion tensor from the nmr spin echo, Journ. of magn. Resonance, B103, 247-254, 1994.

5) D.K. Jones, M.A. Horsfield, A. Simmons, Optimal Strategies for Measuring Diffusion in Anisotropic Systems by Magnetic Resonance Imaging, Mag. Res. Med. 42, 515-525, 1999
6) N. Papadakis et al, A Comparative Study of Acquisition Schemes for Diffusion Tensor Imaging Using MRI, Jour. Mag. Res. 137, 67-82, 1999
7) S. Skare et al, Condition Number as a Measure of Noise Performance of Diffusion Tensor Data Acquisition Schemes with MRI, Jour. Mag. Res. 147, 340-353, 2000.
8) G. Winkler et al, Noise Reduction in Images: Some Recent Edge-Preserving Methods, Jour. of Pattern Recognition and Image Analysis 9(4), 769-766, 1999.
9) J. Polzehl, V. G. Spokoiny, Adaptive weights smoothing with applications to image restoration, Jour. of the Royal Statist. Society B 62(2), 335-354, 2000.
10) G. Parker et al, Nonlinear Smoothing for Reduction of Systematic and Random Errors in Diffusion Tensor Imaging, Jour. Mag. Res. Med. 11, 702-710, 2000.
11) P.J.Basser et al, In Vivo Tractography Using DT-MRI Data, Mag. Res. Med. 44, 625-632, 2000.
12) S. Pajevic et al, A Continuous Tensor Field Approximation of Discrete DT-MRI Data for Extracting Microstructural and Architectural Features of Tissue, Jour. Mag. Res. 154, 85-100, 2002.
13) S. Skare et al, Noise Considerations in the Determination of Diffusion Tensor Anisotropy, Mag Res Imaging 18, 659-669, 2000
14) M.E. Bastin et al, A Theoretical Study of the Effect of Experimental Noise on the Measurement of Anisotropy in Diffusion Imaging, Mag. Res. Imaging 16/7, 773-785, 1998.
15) C. Pierpaoli, P.J. Basser, Toward a Quantitative Assessment of Diffusion Anisotropy, Mag. Res. Med. 36, 893-906, 1996.
16) C. Poupon et al, Regularization of Diffusion-Based Direction Maps for the Tracking of Brain White Matter Fascicles, NeuroImage 12, 184-195, 2000.
17) C.R. Tench et al, Improved White Matter Fiber Tracking Using Stochastic Labeling, Mag. Res. Med. 48, 677-683, 2002.
18) C.R. Tench et al, White Matter Mapping Using Diffusion Tensor MRI, Mag. Res. Med. 47, 967-972, 2002.
19) A.W. Anderson, Theoretical Analysis of the Effects of Noise on Diffusion Tensor Imaging, Mag. Res. Med. 46, 1174-1188, 2001.

20) P.J. Basser, S. Pajevic, Staistical Artifacts in Diffusion Tensor MRI (DT-MRI) Caused by Background Noise, Mag. Res. Med. 44, 41-50, 2000.
21) C.A. Clark et al, In Vivo Mapping of the Fast and Slow Diffusion Tensors in Human Brain, Mag. Res. Med. 47, 623-628, 2002.
22) C.K. Chu et al, Edge-Preserving Smoothers for Image Pocessing, JASA 93, 526-541, 1998.
23) V. Aurich, J. Weule, Non-linear Gaussian Filters performing Edge Preserving Diffusion, Proc. of 17-th DAGM Symposium, Bielefeld, 538-545, Springer, 1995.
24) R. Nieuwenhuys, J. Voogd et al., The Human Central Nervous System, Springer Berlin, 1983.
25) M.R. Wiegell, H.B.W. Larsson and V.J. Wedeen, Fiber crossing in human brain depicted with DTI, Radiology, 897-903, 2000.
26) W. Feller, An Introduction to Probability Theory and its Application, vol. II, 1971 John Wiley & Sons.
27) P.J. Basser, Inferring microstructural features and the physiological state of tissues from diffusion-weighted images, NMR Biomed. 8, 333-344, 1995.
28) Hahn K., Prigarin S. und Pütz B., Edge preserving regularization and tracking for diffusion tensor imaging. Proc. MICCAI-Medical Image Computing and Computer Assisted Intervention, 195-203, 2001.
29) Ahrens E.T., Laidlaw D.H. et al, MR Microscopy of Transgenic Mice that Spontaneously acquire experimental allergic Encephalomyelitis, Mag. Res. Med. 40, 119-132, 1998.
30) Hahn K., Prigarin S. and Pütz B. , Glättung von Diffusionsfeldern und Modellierung von Diffusionspfaden für die Sichtbarmachung von Nervenbahnen im menschlichen Gehirn. Proc. Bildverarbeitung für die Medizin. 55-58, 2002.
31) Hahn K., Prigarin S. and Pütz B. , The Hierarchy of Random Variables in Diffusion Tensor Imaging and the Problems : Spatial Smoothing, Resolution and Tracking, Proc. of 8-th Human Brain Mapping Conference, Neuroimage On-Line, 2002.
32) E. Mühlhaus, „Die Sprungerhaltende Glättung verrauschter harmonischer Schwingungen“,Heinrich-Heine Universität Düsseldorf, Dissertation 1997.
33) P.A. Narayana, private communication, University of Texas, Houston Medical School, 2003.