Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Local estimation of the noise level in MRI using structural adaptation

Karsten Tabelow[1], Henning U. Voss[2], Jörg Polzehl[1]

submitted: May 7, 2014

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: karsten.tabelow@wias-berlin.de
joerg.polzehl@wias-berlin.de

[2] Citigroup Biomedical Imaging Center
Weill Cornell Medical College
516 East 72nd Street
New York, NY 10021, USA
E-Mail: hev2006@med.cornell.edu

No. 1947
Berlin 2014

2010 *Mathematics Subject Classification.* 62G05, 62P10.

Key words and phrases. Magnetic Resonance Imaging, Noise estimation, Maximum Likelihood estimation.

This work was partially supported by the DFG Research Center MATHEON “Mathematics for key technologies”.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

We present a method for local estimation of the signal-dependent noise level in magnetic resonance images. The procedure uses a multi-scale approach to adaptively infer on local neighborhoods with similar data distribution. It exploits a maximum-likelihood estimator for the local noise level. The validity of the method was evaluated on repeated diffusion data of a phantom and simulated data using T1-data corrupted with artificial noise. Simulation results are compared with a recently proposed estimate. The method was applied to a high-resolution diffusion dataset to obtain improved diffusion model estimation results and to demonstrate its usefulness in methods for enhancing diffusion data.

1 Introduction

Noise in Magnetic Resonance Imaging (MRI) affects data analysis in neuroscientific problems or clinical applications. For example, in functional MRI it is directly related to the sensitivity of the experiment [Worsley et al., 2002]. In diffusion MRI it leads to variability and, even more important, to a bias of diffusion model parameter estimates [Pierpaoli and Basser, 1996, Basser and Pajevic, 2000, Jones and Basser, 2004]. Consequently, quantification of the noise is required to measure of the quality of MRI data. Furthermore, estimates of the standard deviation of the MR signal are directly used in a number of data enhancing methods to discriminate noise variation from structural differences, see e.g. Aja-Fernández et al. [2008], Coupé et al. [2008], Becker et al. [2012], Rajan et al. [2012], Becker et al. [2014], Haldar et al. [2013] or the citations in the latter paper for a more comprehensive list.

MR magnitude image reconstruction from Fourier transformed k-space data of a single receiver coil leads to Rician distributed data [Gudbjartsson and Patz, 1995]. For most parallel imaging methods the distribution depends on the reconstruction algorithm but is mostly approximated by the more general non-central χ-distribution [Aja-Fernández et al., 2013]. The scale parameter σ of these distributions is determined by the noise variation of the complex valued noise in k-space, the local coil sensitivities and signal correlations between coils [Aja-Fernández and Tristán-Vega, 2012, Aja-Fernández et al., 2011]. Accordingly, the noise level is generally not a global quantity over the MR image, but varies locally.

Almost all estimation methods for σ rely on the properties of the noise in the image background, i.e., in the absence of a signal, see e.g. Aja-Fernández et al. [2009a] for a comprehensive list of procedures. Only a few methods have been developed that are suitable for the estimation of the noise power also in the presence of an MR signal [Sijbers et al., 1998, Aja-Fernández et al., 2009b, 2013].

In this work we present a novel method for the local estimation of the noise power σ in the presence of a signal, i.e., within tissue regions. The procedure is based on the propagation

separation approach [Polzehl and Spokoiny, 2006] adapted for non-central χ-distributed three-dimensional data. The method searches locally for maximal neighborhoods of a voxel with similar data distribution. This information can be used to infer on the noise parameter through maximum-likelihood estimation, cf. Sijbers et al. [1998] for Rician distributed data.

We demonstrate the effectiveness of the method in a) diffusion weighted data of a diffusion phantom measured repeatedly for a single diffusion gradient direction, b) simulated T1-data corrupted with artificial noise, and c) a diffusion MRI dataset. For the latter we additionally demonstrate how the result of the local noise estimation can be used to obtain an improved and unbiased estimate for the parameters of the diffusion tensor model [Basser et al., 1994a,b]. Finally, we demonstrate how the local estimate of the noise power can be used to improve the results of a recently developed method for noise reduction in diffusion MRI data [msPOAS, Becker et al., 2014].

2 Theory

2.1 Noise distribution in multiple-coil MR acquisition

An MR image is acquired in frequency or k-space and has to be transferred to the common image domain via inverse Fourier transform [Callaghan, 1991]. The noise present in k-space data of a single coil can be modeled as a complex additive Gaussian variable with zero expectation and homogeneous standard deviation σ. For a single-coil acquisition the local signal of the magnitude image in the spatial domain then follows a Rician distribution [Gudbjartsson and Patz, 1995]. For a multiple-channel RF coils [Roemer et al., 1990] the data from all L coils is used in the reconstruction of the image data. In case of a sum-of-squares image reconstruction, the standardized signal S_i/σ at the spatial position x_i of voxel i is usually considered to be non-central χ distributed with $2L$ degrees of freedom and non-centrality parameter θ_i [Constantinides et al., 1997]. Another special reconstruction method SENSE1 [Sotiropoulos et al., 2013b] used in the Human Connectome Project leads to $L = 1$ and hence to a Rician distribution as for single coil acquisitions. With other parallel imaging methods like GRAPPA a non-central χ distribution with adjusted, location dependent distribution parameters serves as a valid approximation of the true data distribution [Aja-Fernández et al., 2011].

The probability density p_S for the distribution P_S of the signal S_i generally depends on three parameters θ_i, σ_i, L_i and is given by

$$p_S(S_i;\theta_i,\sigma_i,L_i) = \frac{S_i^{L_i}\theta_i^{(1-L_i)}}{\sigma_i^{(L_i+1)}} e^{-\frac{1}{2}\left(\frac{S_i^2}{\sigma_i^2}+\theta_i^2\right)} I_{L_i-1}\left(\frac{\theta_i S_i}{\sigma_i}\right), \tag{2.1}$$

where I_{L_i-1} denotes the (L_i-1)-th order modified Bessel function of the first kind. The mean μ and variance v of the distribution are given by

$$\mu(\theta_i,\sigma_i,L_i) = \sigma_i\sqrt{\frac{\pi}{2}}\,\mathbb{L}_{1/2}^{(L_i-1)}\left(-\frac{\theta_i^2}{2}\right) \tag{2.2}$$

$$v(\theta_i,\sigma_i,L_i) = \sigma_i^2\left(2L_i+\theta_i^2\right) - \mu^2(\theta_i,\sigma_i,L_i), \tag{2.3}$$

where $\mathbb{L}_{1/2}^{(L-1)}$ is a generalized Laguerre polynomial. In general, σ_i and L_i vary smoothly with location.

2.2 Estimating a local smooth noise standard deviation using adaptive weights

We propose a method for estimation of the local scale parameter σ_i in two- or three dimensional MR images. We assume the effective number of coils L_i to be known and locally homogeneous. The parameter $\sigma_i = \sigma(x_i)$ also generally depends on the coils sensitivity maps, noise correlations between receiver coils' and spatially slowly varying parameters in the reconstruction algorithm, and is therefore assumed to be a smooth and slowly varying function of location.

The parameter function θ_i is supposed to be locally constant with x_i [Becker et al., 2012, 2014]. This assumption is motivated by the observation that the expected signal intensity or equivalently the non-centrality parameter θ_i relates to properties of the biological tissue. Then, θ_i is approximately constant in regions with one type of tissue, while values for different tissue may considerably differ. In the following the terms "homogeneity region" refers to this property of the non-centrality parameter.

We use the propagation-separation approach [Polzehl and Spokoiny, 2006] to iteratively infer on the non-centrality parameter θ and its homogeneity regions. The adaptive refinement of the homogeneity regions can then be used to estimate the local scale parameter σ_i using weighted maximum likelihood.

Let K_{loc} and K_{ad} be two non-decreasing kernel function supported on the interval $[0, 1)$, k^* a pre-specified number of iteration steps and $\{h^{(k)}\}_{k=0}^{k^*}$ a monotone sequence of bandwidths. Let $\lambda > 0$, $h_{\mathrm{med}} > 0$, N_0 denote further parameters of the procedure, which are explained in detail below.

We propose the following iterative algorithm:

- $k = 0$: Initialize $\tilde{\sigma}_i^{(0)} = \bar{\sigma}$ using a global initial estimate (or guess) for the noise standard deviation. Set the initial estimate for the non-centrality parameter constant, i.e., $\hat{\theta}_i^{(0)} = 1$, to enforce a non-adaptive weighting scheme $w_{ij}^{(1)}$ for the first iteration step, see below. Set $N_i^{(0)} = 1$ for all voxel i, cf. also below. Set $k = 1$.
- For each voxel i compute adaptive weights $w_{ij}^{(k)} =$

$$K_{\mathrm{loc}}\left(\frac{\|x_i - x_j\|}{h^{(k)}}\right) K_{\mathrm{ad}}\left(\frac{N_i^{(k-1)} \mathcal{KL}\left(P_S\left(\hat{\theta}_i^{(k-1)}, \tilde{\sigma}_i^{(k-1)}\right), P_S\left(\hat{\theta}_j^{(k-1)}, \tilde{\sigma}_i^{(k-1)}\right)\right)}{\lambda}\right) \tag{2.4}$$

  and their sum $N_i^{(k)} = \sum_j w_{ij}^{(k)}$ over all voxel j. Here, $\|\cdot\|$ denotes the Euclidean norm to calculate the spatial distance between two voxel i and j. The first term of $w_{ij}^{(k)}$ defines non-adaptive weights depending the location kernel K_{loc} and the spatial distance of voxel

i and j. The second term evaluates the statistical difference between the estimated non-centrality parameters $\hat{\theta}_i^{(k-1)}$ and $\hat{\theta}_j^{(k-1)}$ from the previous iteration step to infer on the homogeneity regions of θ_i.

The definition of $w_{ij}^{(k)}$ employs the Kullback-Leibler $\mathcal{KL}$ divergence between the probability distributions $P_S\left(\hat{\theta}_i^{(k-1)}, \tilde{\sigma}_i^{(k-1)}\right)$ and $P_S\left(\hat{\theta}_j^{(k-1)}, \tilde{\sigma}_i^{(k-1)}\right)$, where we dropped the dependence on the (homogeneous) L for brevity of the notation. Both distributions have the same estimated scale parameter from the previous iteration step $\tilde{\sigma}_i^{(k-1)}$ which reflects the assumption that $\sigma_i = \sigma(x_i)$ is smooth and slowly varying, i.e., $|\sigma_i - \sigma_j|$ is small for $\|x_i - x_j\| < h^{(k)}$.

λ is the adaptation bandwidth of the procedure. It controls the amount of spatial adaptation in the definition of the weights. In the extreme case of $\lambda \to \infty$ the second term equals to $K_{\mathrm{ad}}(0)$, such that the weighting schemes $w_{ij}^{(k)}$ are fully non-adaptive. For $\lambda = 0$ the weights $w_{ij}^{(k)}$ always vanish for $i \neq j$ due to the bounded support of the kernel function K_{ad}. For a proper choice of λ (see below), $w_{ij}^{(k)}$ will, with high probability and increasing k, approach 1 or 0, depending on voxel i and j belonging to the same homogeneity region or not, respectively. Thus, for large k, the weighting scheme $W_i^{(k)} = (w_{i1}^{(k)}, \cdots, w_{in}^{(k)})$ describes the homogeneity region of voxel i.

- If $N_i^{(k)} := \sum_j w_{ij}^{(k)} > N_0$ we obtain estimates for $\theta(x_i)$ and $\sigma(x_i)$ by weighted log-likelihood

$$\left(\hat{\sigma}_i^{(k)}, \hat{\theta}_i^{(k)}\right) = \mathrm{argmax}_{(\theta,\sigma)} \sum_j w_{ij}^{(k)} \log p_S(S_j; \theta, \sigma). \tag{2.5}$$

  Otherwise we set $\tilde{\sigma}_i^{(k)} := \tilde{\sigma}_i^{(k-1)}$ and $\hat{\theta}_i^{(k)} := \sqrt{\left(\sum_j w_{ij}^{(k)} S_j^2 / \sum_j w_{ij}^{(k)} - 2L_i(\tilde{\sigma}_i^{(k-1)})^2\right)_+}$ employing the moment equation (2.3). The adaptive weights enforce that observations S_j follow a distribution with significantly different non-centrality parameter θ are not utilized in the estimator above. Within the iteration process (for increasing k), the sum of all weights $N_i^{(k)}$ increases, the inference on the homogeneity regions will be more informed, and the estimates above stabilize.

- If $N_i^{(k)} > N_0$ we stabilize the estimate of $\sigma(x_i)$ using the median filter

$$\tilde{\sigma}_i^{(k)} = \mathrm{median}_{j:\|x_i - x_j\|_1 < h_{\mathrm{med}}} \hat{\sigma}_j^{(k)}$$

  over a cube of voxel centered in i. The median filter reduces the variability of the estimate. The robustness of the median filter avoids a bias that may be caused by insufficient adaptation in some voxel j.

- If $k = k^*$ stop, else increase k by 1 and continue with the second step, i.e., with the definition of the weights.

The parameter N_0 defines a minimal sum of weights $N_i^{(k)}$ before the initial guess for σ_i is updated. $N_0 >> 1$ guarantees the identifiability in (2.5) using observations from a sufficiently large homogeneous vicinity of voxel i.

2.3 Maximum Likelihood estimator for non-central χ-distribution

Given a sample $\mathcal{S} = (S_1, \ldots, S_n)$ a local weighted log-likelihood as used in the iterative method described in the preceding section (up to terms that do not depend on the parameters and hence will be irrelevant in the optimization and dropping the dependence on the iteration step k and on i) takes the form, see Eq. (2.1),

$$\begin{aligned} l(\mathcal{S}; W_i; \theta, \sigma) =& \sum_j w_{ij} \log p_S\left(S_j; \theta, \sigma\right) \\ =& - N_i \left((L+1) \log \sigma + (L-1) \log \theta + \frac{\theta^2}{2} + \frac{\xi}{2\sigma^2} \right) \\ &+ \sum_j w_{ij} \log I_{L-1}\left(\frac{\theta S_j}{\sigma}\right) \end{aligned} \tag{2.6}$$

with $N_i = \sum_j w_{ij}$ and $\xi = \frac{1}{N_i} \sum_j w_{ij} S_j^2$. Here, we assumed the probability density p_s to be non-central χ with non-centrality parameter θ and $2L$ degrees of freedom. The locality of the estimates is then realized by using this estimator at each voxel i.

In order to solve the optimization problem in Eq. (2.5) we take the derivatives with respect to θ^2 and σ and set them to zero. This yields

$$\begin{aligned} &\frac{\sigma}{\theta} \frac{\partial l}{\partial \sigma}(\mathcal{S}; W_i; \theta, \sigma) = \\ &-N_i \left(\frac{L+1}{\theta} - \frac{\xi}{\theta \sigma^2} \right) - \sum_j w_{ij} \frac{I_L\left(\frac{\theta S_j}{\sigma}\right) + I_{L-2}\left(\frac{\theta S_j}{\sigma}\right)}{2 I_{L-1}\left(\frac{\theta S_j}{\sigma}\right)} \frac{S_j}{\sigma} \stackrel{!}{=} 0 \\ &\frac{\partial l}{\partial \theta}(\mathcal{S}; W_i; \theta, \sigma) = \\ &-N_i \left(\frac{L-1}{\theta} + \theta \right) + \sum_j w_{ij} \frac{I_L\left(\frac{\theta S_j}{\sigma}\right) + I_{L-2}\left(\frac{\theta S_j}{\sigma}\right)}{2 I_{L-1}\left(\frac{\theta S_j}{\sigma}\right)} \frac{S_j}{\sigma} \stackrel{!}{=} 0 \end{aligned}$$

By adding both equations and re-arranging the result we get

$$\hat{\theta}^2 = \frac{\xi}{\sigma^2} - 2L = \frac{1}{N_i \sigma^2} \sum_j w_{ij} S_j^2 - 2L \tag{2.7}$$

Substituting this into the local likelihood Eq. (2.6) and again removing the constant term that does not depend on σ yields

$$\begin{aligned} \breve{l}(\mathcal{S}; W_i; \sigma) =& - N_i \left(\xi/\sigma^2 + 2 \log \sigma + \frac{L-1}{2} \log\left(\xi - 2L\sigma^2\right) \right) \\ &+ \sum_j w_{ij} \log I_{L-1}\left(\frac{S_j}{\sigma^2} \sqrt{\xi - 2L\sigma^2} \right). \end{aligned} \tag{2.8}$$

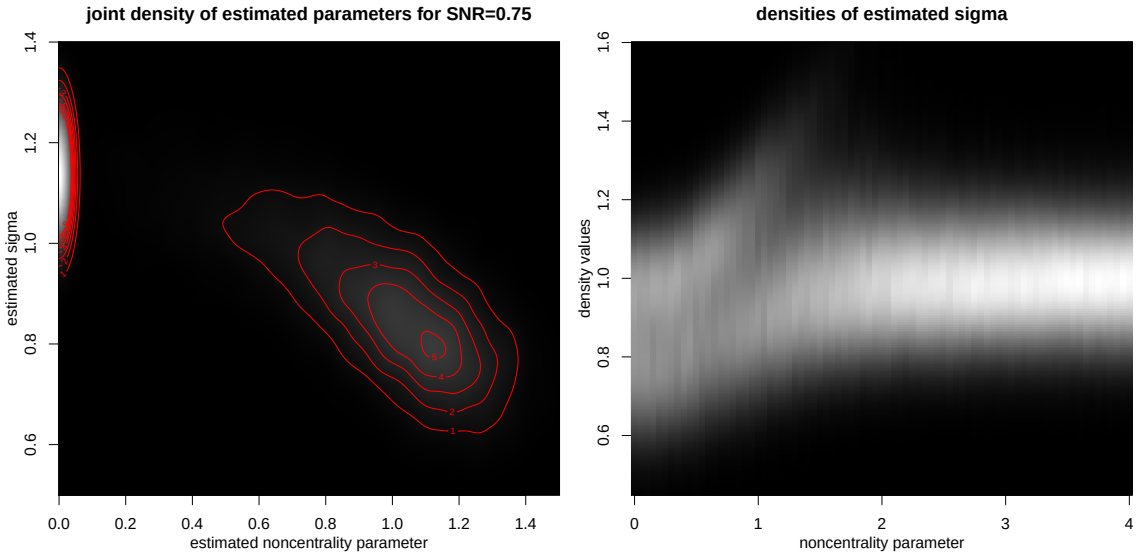

Figure 1 – *2D density of estimated parameters $\hat{\theta}$ and $\hat{\sigma}$ obtained from simulating 1000 samples of size 50 from a non-central χ-distribution with parameter $\theta = 0.75$ ($\sigma = 1$) maximizing the weighted likelihood in Eq.* (2.5) *(left) and marginal densities of estimates $\hat{\sigma}$ for varying non-centrality parameter $\theta \in (0, 4)$ (right).*

Finally we get an estimator for σ_i as a solution of an optimization problem in the univariate parameter σ:

$$\hat{\sigma}_i = \sqrt{\frac{N_i}{N_i - 1}}\, \mathrm{argmax}_\sigma\, \breve{l}(\mathcal{S}; W_i; \sigma). \tag{2.9}$$

If in Eq. (2.7) $\xi_i/\sigma^2 - 2L \leq 0$ the optimization problem in Eq. (2.5) has no solution in the interior of the parameter domain and $\hat{\theta}_i = 0$, i.e., the χ-distribution is central. In Eq. (2.8) the two terms containing the logarithm diverge and have to be specially handled. Recalling that

$$\lim_{z \to 0} I_{L-1}(z) = \left(\frac{z}{2}\right)^{L-1} \Gamma(L)$$

we find that the problematic terms cancel and we get

$$\breve{l}(\mathcal{S}; W_i; \sigma) = -\, N_i \left(\xi_i/\sigma^2 + 2\log\sigma - \log\Gamma(L)\right) + (L-1) \sum_j w_{ij} \log\left(\frac{S_j}{2\sigma^2}\right).$$

The estimator in Eq. (2.9) is biased for small values of θ/σ [Sijbers et al., 1998]. In particular, its density has two modes, one arising from the estimates from the central χ-distribution case, and one for the case where $\xi_i/\sigma^2 - 2L > 0$. This effect is demonstrated in Fig. 1. Thus the estimator causes problems in regions of the MR image, where the signal is very low or zero. However, in this paper we are explicitly interested in the estimation of the noise parameter in the region containing brain tissue and hence with sufficiently large signal. There exist a large number of estimators for σ in the background, see e.g. Aja-Fernández et al. [2009a].

2.4 Choice of parameters of the procedure

The specific choice for the kernel functions K_{loc} and K_{ad} has only minor influence on the estimation results, see, e.g., Section 6.2.3 in Scott [1992]. We choose them as:

$$K_{\mathrm{loc}}(x) = \begin{cases} 1-x^2 & x<1 \\ 0 & x \geq 1 \end{cases} \quad \text{and} \quad K_{\mathrm{ad}}(x) = \begin{cases} 1 & x<0.5 \\ 2-2x & 0.5 \leq x < 1 \\ 0 & x \geq 1 \end{cases}$$

for computational efficiency. We choose $h^{(0)} = 1$. The monotone sequence $\{h^{(k)}\}_{k=0}^{k^*}$ of bandwidths is chosen using only the non-adaptive part of the weights defined in Eq. (2.4):

$$\tilde{\omega}_{ij}^{(k)} = K_{\mathrm{loc}}\left(\frac{\|x_i - x_j\|}{h^{(k)}}\right)$$

such that

$$\frac{\sum_j (\tilde{\omega}_{ij}^{(k-1)})^2 \left(\sum_j \tilde{\omega}_{ij}^{(k)}\right)^2}{\left(\sum_j \tilde{\omega}_{ij}^{(k-1)}\right)^2 \sum_j (\tilde{\omega}_{ij}^{(k)})^2} = c_h.$$

This ensures an approximately constant variance reduction for the non-adaptive estimator from step $k-1$ to k. We choose $c_h = 1.25$.

λ is the main parameter of the procedure as it controls the amount of adaptation of the method. A reasonable value is $\lambda = 5$, see Figure 5.

We recommend $k^\star = 20$, $N_0 = 2$ and $h_{\mathrm{med}} = 5$. The mean number of effective coils L (or $L(x_i)$) needs to be specified by the user.

2.5 Application to an unbiased estimation of the diffusion tensor for DTI data

Diffusion weighted imaging [dMRI; Jones, 2010] has become a widely used standard tool for structural in-vivo examination of the brain in recent years. There, the application of an additional diffusion weighting gradient in the magnetic field leads to a signal attenuation that is directly related to the diffusion properties of the water in tissue along the considered gradient direction. Additionally, at least one non-diffusion weighted image volume is acquired for comparison. Very often in the Gaussian diffusion approximation, the directional dependence of the diffusion properties is described within the diffusion tensor model exploited in diffusion tensor imaging [DTI; Basser et al., 1994a,b]. The physical model for DTI is formulated in terms of a noiseless situation. Let $\zeta_{p,i}$ denote the noiseless image value at voxel i for the p-th volume of the diffusion weighted MRI dataset, corresponding to the diffusion weighting gradient in direction $\vec{g}_p$ and the b-value b_p [Basser et al., 1994b]. This series also includes the non-diffusion weighted image volumes. The diffusion tensor model then describes the data by a symmetric, positive definite 3×3 matrix $\mathcal{D}_i$ and a parameter ζ_i^0 with

$$\zeta_{p,i} = \zeta_i^0 \exp(-b_p \vec{g}_p^\top \mathcal{D}_i \vec{g}_p) \quad \forall p. \tag{2.10}$$

Positive definiteness of the diffusion tensor $\mathcal{D}_i$ can be enforced by re-parametrizing $\mathcal{D}_i = R_i^\top R_i$ using an upper triangular matrix R_i, see Koay et al. [2006] or Ghosh et al. [2013].

Traditionally the diffusion tensor is estimated using nonlinear regression

$$(\hat{\zeta}_i^0, \hat{\mathcal{R}}_i) = \operatorname{argmin} \sum_p (S_{i,p} - \zeta_{i,p})^2 \tag{2.11}$$

thereby ignoring the difference between non-centrality parameter and expectation of a non-central chi-distribution. For small SNR this causes a bias in the tensor estimate.

The log-likelihood function for the diffusion tensor model assuming a non-central χ distribution with $2L_i$ degrees of freedom and non-centrality parameter $\zeta_{p,i}/\sigma_{p,i}$ for the standardized observed image intensities $S_{p,i}/\sigma_{p,i}$ is given by

$$\begin{aligned} l(\{S_{p,i}\}_p; \{\sigma_{p,i}\}_p, L_i; \zeta_i^0, R_i) = \\ \sum_p \left[\log \left(\frac{S_{p,i}^{L_i} \zeta_{p,i}^{(1-L_i)}}{\sigma_{p,i}^2} \right) - \frac{1}{2} \left(\frac{S_{p,i}^2 + \zeta_{p,i}^2}{\sigma_{p,i}^2} \right) + \log \left(I_{L_i - 1} \left(\frac{\zeta_{p,i} S_{p,i}}{\sigma_{p,i}^2} \right) \right) \right], \end{aligned} \tag{2.12}$$

where $\zeta_{p,i}$ is given by Eq. (2.10) and the re-parametrization in terms of the model parameters ζ_i^0, R_i. Estimates are then obtained by maximizing the log-likelihood.

This approach has been already considered in the literature for the estimation of the diffusion tensor [Landman et al., 2009] or for the extended diffusion kurtosis model [Veraart et al., 2011a,b, Ghosh et al., 2013, AndrÃľ et al., 2014]. The procedure avoids the bias in the parameter estimation that is caused by the skewness of the χ distribution causing the deviation of its expectation value from the non-centrality parameter in a noisy situation, see Eq. (2.2).

An alternative is provided by quasi-maximum likelihood, i.e., by minimizing the negative Gaussian log-likelihood

$$\mathcal{R}(\{S_{p,i}\}_p; \{\sigma_{p,i}\}_p, L_i; \zeta_i^0, R_i) = \sum_p \left[\frac{(S_{p,i} - \mu(\zeta_{p,i}/\sigma_{p,i}, \sigma_{p,i}, L_i))^2}{\nu(\zeta_{p,i}/\sigma_{p,i}, \sigma_{p,i}, L_i)} \right] \tag{2.13}$$

using (2.2), (2.3), (2.10) and $\mathcal{D}_i = R_i^\top R_i$, and effectively approximating the non-central χ distribution by a Gaussian distribution with appropriate moments.

3 Materials and methods

3.1 Experimental data (Diffusion weighted images of a diffusion phantom)

Data

Data were obtained from a DTI phantom with straight and crossing fibers assembled from parts of the phantom published by Pullens et al. [2010]. Diffusion weighted images were acquired on

a 7.0 Tesla 70/30 Bruker Biospec small animal MRI system with $450\mathrm{mT/m}$ maximum gradient amplitude and $4500\mathrm{T/m/s}$ maximum slew rate. Radiofrequency power was transmitted by a $72\mathrm{mm}$ diameter linear coil and picked up by a quadrature rat brain coil. We used a single-shot spin-echo EPI sequence with echo and repetition times of $\mathrm{TE} = 46\mathrm{ms}$ and $\mathrm{TR} = 2500\mathrm{ms}$, respectively. Four slices in the x-y plane with matrix size 128×128 (resolution $0.2344 \times 0.2344\mathrm{mm}^2$) and $2\mathrm{mm}$ thickness were scanned without skip. The acquisition matrix size was 128×91 (partial Fourier overscan factor of 1.4). We acquired 10 images without diffusion weighting and 1000 diffusion weighted images with diffusion gradient direction along the x-direction. The diffusion gradient width was $\delta = 4.5\mathrm{ms}$ and the spacing $\Delta = 9.2\mathrm{ms}$, yielding a b-value of $1000\mathrm{s/mm}^2$. Effective b-values for the non-diffusion weighted and the diffusion weighted scans were $2\mathrm{s/mm}^2$ and $1031\mathrm{s/mm}^2$, respectively. The total scan time for this scan was $42\mathrm{min}$.

Analysis

We obtained, using Eqs. (2.2) and (2.3) and adjusting for trends over time, a local reference estimates $\tilde{\sigma}_i$ from the 1000 replications. We then estimated σ_i for all images using the proposed method, leading to estimates $\hat{\sigma}_{i,j}$ in location x_i and image j. The effective number of coils was specified as $L = 1$ which is correct in this case. The other parameters of the procedure were $\lambda = 5$, $k^\star = 20$, and $h_{\mathrm{med}} = 5$.

3.2 Simulated data (T1 image)

Data

For the simulation experiments we used a 3D synthetic BrainWeb MR volume [Collins et al., 1998]. Specifically, we generated a noise-free T1 image with 1mm slice thickness and 0% intensity non-uniformity in the provided 12bit short raw format. The image intensity had a range from 0 to 4096 with a median value of approximately 1800 for gray matter areas and 2400 for white matter regions. To simulate a parallel imaging process we defined 8 artificial linear sensitivity maps along the x- and y-axis and the diagonals, and created pseudo image acquisition for each of the 8 simulated coils, see Fig. 2. We added complex Gaussian noise with standard deviation $\sigma_K = 50, 100, 200, 400, 800$ in k-space, specifying a correlation between coils k and l of $\rho = 0.5^{(d_{kl})}$ where d_{kl} refers to a distance between the spherically arranged coils. The final magnitude image was obtained using a SENSE1 reconstruction [Sotiropoulos et al., 2013b], which is the standard imaging protocol used in the Human Connectome Project [Sotiropoulos et al., 2013a]. The parameter σ_i in the resulting magnitude image ranges from $0.86\sigma_K$ to $2.73\sigma_K$ with maximum values in the center of the image, see Fig. 4g.

Analysis

We applied the method described in the Theory part of this paper to determine the noise parameter over the image within a brain mask covering most of the white and gray matter. The

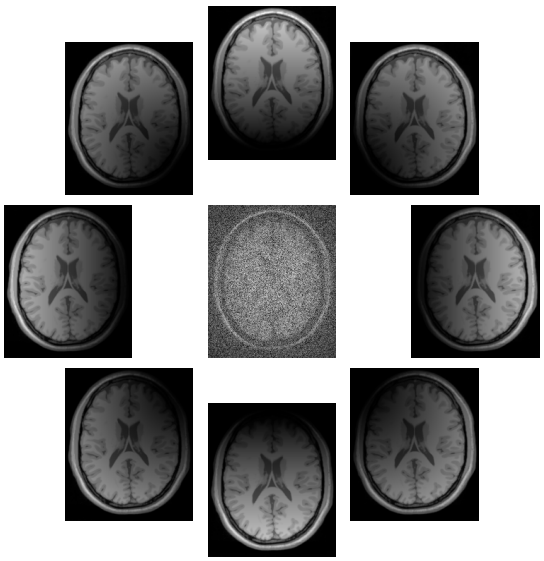

Figure 2 – *For the generation of the simulated data a noise-free T1 image from the BrainWeb [Collins et al., 1998] was used. By the use of 8 artificial linear sensitivity maps and adding complex Gaussian noise to each image from the artificial coils and using the reconstruction method SENSE1 [Sotiropoulos et al., 2013b], we obtained noisy data with known local noise standard deviation.*

parameters of the procedure were $\lambda = 5$, $k^\star = 20$, $h_{\mathrm{med}} = 5$ and $L = 1$. The analysis was restricted to a white/gray matter mask.

We also determined an estimate of σ_i by the method proposed in Aja-Fernández et al. [2013] using a cube of $5 \times 5 \times 5$ voxel centered in x_i as a local neighborhood.

Additionally we investigated the quality of results depending on parameters λ, h_{med} and k^*.

3.3 Experimental data (diffusion weighted imaging)

MRI

We re-analyzed a dataset described already in Becker et al. [2012] and Becker et al. [2014]. Data were acquired from a whole body 7T MAGNETOM scanner (Siemens Healthcare) with a maximum gradient amplitude of $70\mathrm{mT/m}$ and a maximum slew rate of $200\mathrm{T/m/s}$ (SC72, Siemens Healthcare, Erlangen, Germany). The scan was performed using a single channel transmit, 24-channel receive phased array head coil (Nova Medical, Wilmington, MA, USA). An optimized monopolar Stejskal-Tanner sequence according to Morelli et al. [2010] together with the ZOOPPA approach described in Heidemann et al. [2012] was used with $\mathrm{TR}\,14.1\mathrm{s}$, $\mathrm{TE}\,65\mathrm{ms}$, $\mathrm{BW}\,1132\mathrm{Hz/pixel}$, and ZOOPPA acceleration factor of 4.6. A total of 91 slices with 10% overlap were acquired at a field-of-view (FoV) of $143 \times 147\mathrm{mm}^2$ resulting in an isotropic high resolution of $800\mu\mathrm{m}$. Diffusion weighting gradients were applied along 60 different directions at a b-value of $1000\mathrm{s/mm}^2$. 7 interspersed non-diffusion weighted images were

acquired. The scan was repeated 4 times. The subject was a healthy adult volunteer after obtaining written informed consent in accordance with the ethical approval from the University of Leipzig. Total acquisition time was $65\mathrm{min}$.

Analysis

We applied the procedure introduced in the Theory part to determine the local noise parameter $\sigma_{p,i}$ over all $p = 1, \ldots, 268$ images (with and without diffusion weighting) within a brain mask. The parameters of the procedure were $\lambda = 5$, $k^\star = 16$ (for reduced computational costs), $h_{\mathrm{med}} = 5$ and $L = 1$.

The data were then smoothed using a version of the msPOAS algorithm [Becker et al., 2014] that has been adapted to use local estimates of σ as indicated there as well as using a global estimate for σ.

Diffusion tensors where estimated using nonlinear regression (2.11) and the quasi-likelihood (2.13) for both the original and smoothed data. Quasi-likelihood was used instead of the likelihood (2.12) since it is suitable for both the original and the smoothed data.

3.4 Software

The new method for estimation the local noise parameter proposed in this paper is implemented within our R-package **dti** [Tabelow and Polzehl, 2014] (version 1.2-0). This package is freely available on CRAN (`http://cran.r-project.org`) and on NITRC (`http://www.nitrc.org`). The implementation uses FORTRAN and native R-code. The package also provides an implementation of the maximum-likelihood and quasi-maximum-likelihood estimates of the diffusion tensor model parameters. We used the package for all calculations in this paper.

4 Results

4.1 Experimental data (Diffusion weighted images of a diffusion phantom)

In Fig. 3 we summarize the results obtained for the repeated diffusion weighted images of the phantom. The relative mean absolute error

$$MRAE(x_i) = \frac{1}{1000} \sum_j \frac{|\hat{\sigma}_{ij} - \tilde{\sigma}_i|}{\tilde{\sigma}_i},$$

the mean relative bias

$$RBias(x_i) = \frac{1}{1000} \sum_j \frac{\hat{\sigma}_{ij} - \tilde{\sigma}_i}{\tilde{\sigma}_i}$$

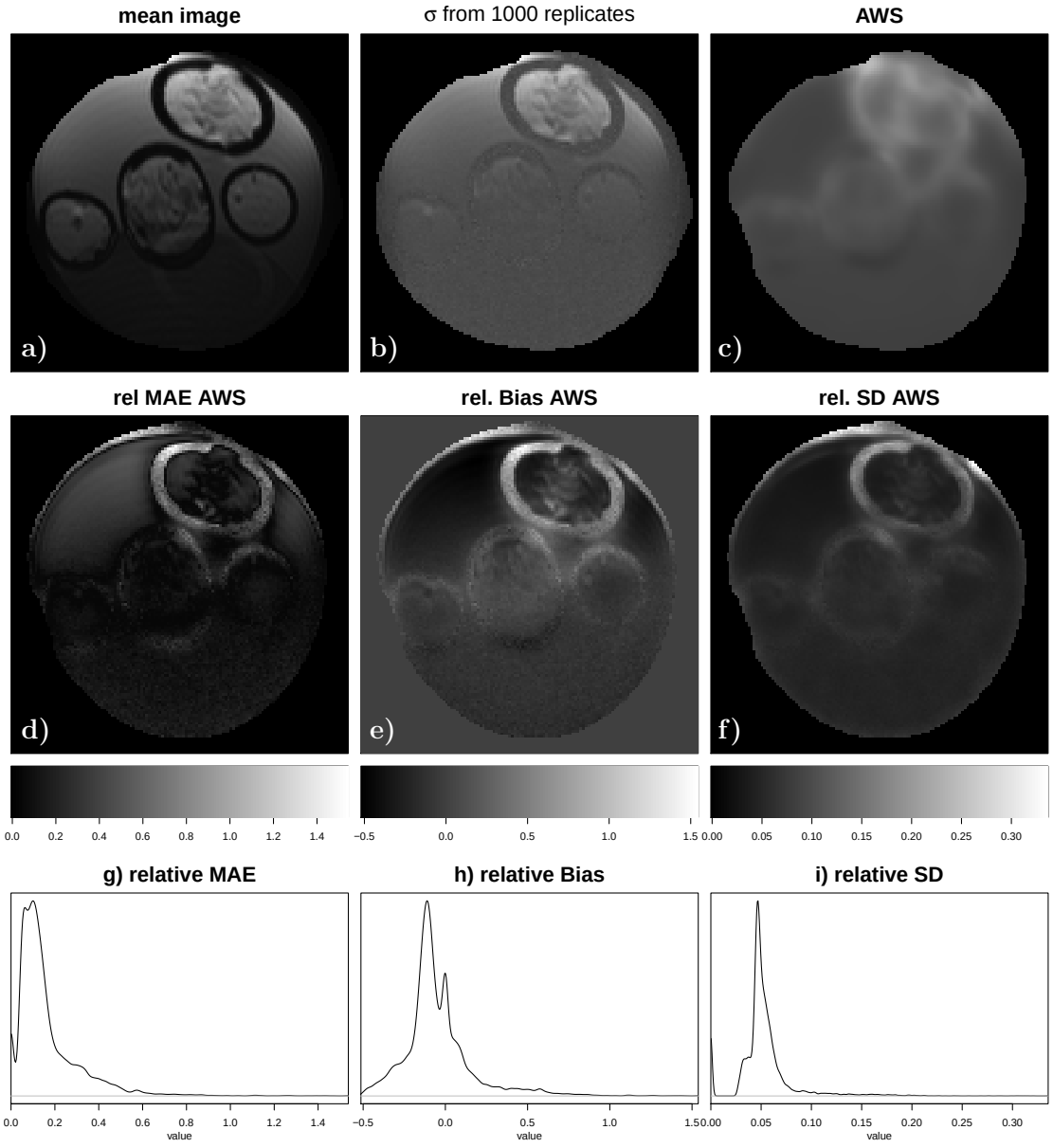

Figure 3 – *Results for repeated diffusion weighted image. a) mean of 1000 dMRI images, b) voxelwise estimate of σ obtained from 1000 replicates, c) mean (over 1000 replicates) of σ estimated using the proposed method, d) relative mean absolute error (MAE) of estimated σ, e) relative bias of estimated σ, f) relative standard deviation of estimated σ. d)-f) are associated by scale information and density plots in g)-h), respectively.*

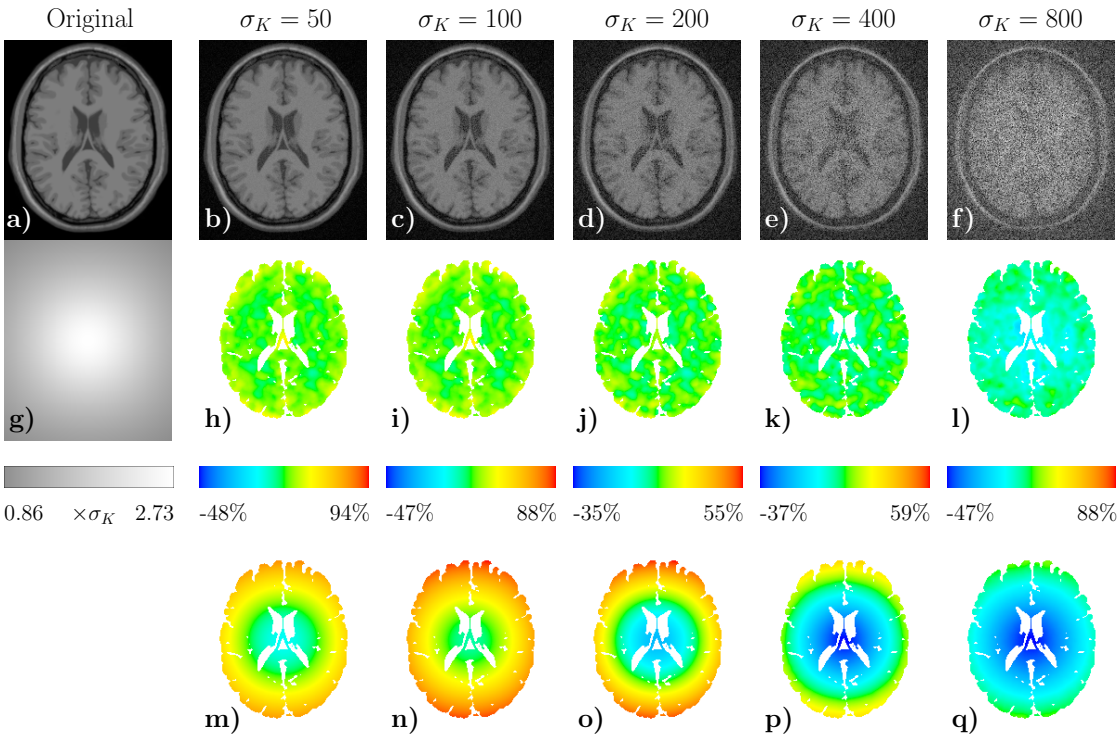

Figure 4 – *Simulation results for a slice of the BrainWeb MR volume. a) Original slice. b)-f) Slice after adding complex Gaussian noise with varying standard deviation using a SENSE1 Sotiropoulos et al. [2013b] reconstruction. g) image of locally varying effective σ_i parameters, h)-l) relative error of local estimates $\hat{\sigma}_i$ using the method proposed in this paper. m)-q) relative error of local estimates $\hat{\sigma}_i$ using the method described in Aja-Fernández et al. [2013]. The errors are given on a log-scale.*

and the relative standard deviation

$$RSD(x_i) = \frac{\mathrm{sd}_j\, \hat{\sigma}_{ij}}{\tilde{\sigma}_i}$$

are defined considering the estimates $\tilde{\sigma}_i$ obtained from the 1000 replicates as ground truth. Results refer to the region inside the phantom.

4.2 Simulated data (T1 image)

In Fig. 4 summarizes the results of the noise estimation for an arbitrarily selected slice (Fig. 4a) of the simulated T1 data. Figs. 4b)-f) show the same slice corrupted with complex Gaussian noise with standard deviation $\sigma_K = 50, 100, 200, 400, 800$ using artificial sensitivity maps, correlation between coils and a SENSE1 reconstruction [Sotiropoulos et al., 2013a], cf. Fig. 2. This leads to a location dependent parameter σ as shown in Fig. 4g. In Figs. 4h)-l) we show the logarithmic ratio of the local estimated noise parameter and its theoretical value (Fig. 4g) for all five noise levels. The color scale is defined such that green refers to the ideal log-ratio of 0. The range of the log-ratio varies with the noise level and is given below the color scale.

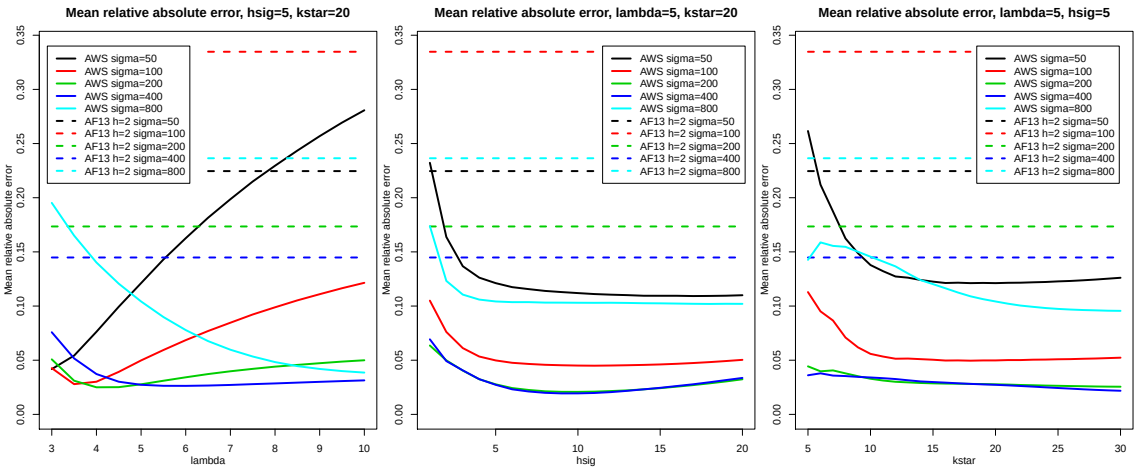

Figure 5 – *Influence of parameters on the estimation quality for the BrainWeb MR volume. Mean relative absolute error of estimated σ as function of parameters λ, h_{med} and k^*.*

In Figs. 4m)-q) we show the corresponding results for the method proposed in Aja-Fernández et al. [2013].

Fig. 5 provides the mean relative absolute error as a function of the parameters λ, h_{med} and k^* and the five noise levels. The mean relative absolute error for the method proposed in Aja-Fernández et al. [2013] is given as a comparison. The profiles indicate a stable behavior of the procedure with respect to the chosen parameters. Large values of h_{med} and k^* provide slight improvements at the cost of significantly higher numerical effort. Improvements are mainly due to a reduced variability of the estimates. The parameter λ steers the adaptation to the structure in the non-centrality parameter of the χ-distribution leading to a balance between excluding observations from the tails of the noncentral χ-distribution when the non-centrality parameters coincide (same homogeneity region), and including observations with different non-centrality parameters (distinct regions). The optimal balance depends on the SNR and the homogeneity structure, with inadequate choice resulting in a bias of the estimates. Our recommendation $\lambda = 5$ serves as a good compromise, with slightly smaller / larger values being preferable in high / low SNR situations.

4.3 Experimental data (diffusion weighted imaging)

In Fig. 6 we illustrate the results for the diffusion weighted dataset. In a) and d) we show the mean, over all gradients, estimated σ_i for one slice of the non-diffusion weighted images and the diffusion weighted images, respectively. Figs. 6 b) and e) provide the ratio of the standard deviation $\mathrm{sd}_j\, \hat{\sigma}_{ij}$ over all gradients, and the mean estimated σ_i while c) and f) show densities of the values in b) and e). Note that the standard deviation $\mathrm{sd}_j\, \hat{\sigma}_{ij}$ reflects both the variability of the true σ_i over the consecutively acquired images and the variability of the individual estimates. Images in b) and e) and densities in c) and f) therefore provide upper bounds of the relative error.

Our last result concerns the effect using the locally estimated σ_i on estimates of the Fractional

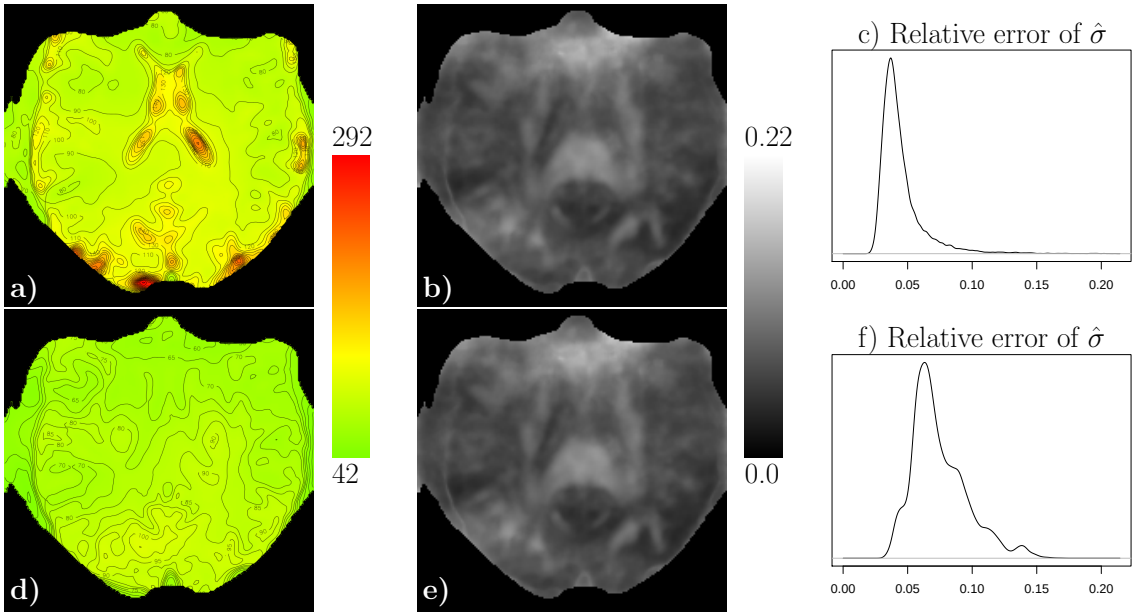

Figure 6 – *Results for the diffusion weighted dataset. Mean estimated σ over all a) non-diffusion weighted and d) diffusion weighted images. Ratio of the standard deviation of the estimated σ over all b) non-diffusion weighted and e) diffusion weighted images to the corresponding mean estimate. c) and f) show the corresponding densities of this relative error.*

Anisotropy (FA) in a tensor model. The locally estimated σ_i is employed in two different occasions: for suitable standardization of the signal in our msPOAS smoothing algorithm [Becker et al., 2014], and as a parameter in the quasi-likelihood function employed for estimating the tensor parameters, see Eq. (2.13). Fig. 7 provides results in terms of estimated FA and estimated main diffusion direction based on different estimates of the diffusion tensor, $\mathcal{D}_{nlreg}^{orig}$ and $\mathcal{D}_{nlreg}^{locPOAS}$ employing Eqn. (2.11) on the original data and data smoothed by msPOAS [Becker et al., 2014] (16 steps) using local estimates σ_i, respectively, and $\mathcal{D}_{qlike}^{orig}$ and $\mathcal{D}_{qlike}^{locPOAS}$ where the quasi-likelihood Eq. (2.13) was used instead of non-linear regression. For comparison we also computed an estimate $\mathcal{D}_{qlike}^{glPOAS}$ using Eq. (2.13) and data smoothed by msPOAS (16 steps) employing a global estimate of σ.

Fig. 7a) provides the estimated FA obtained from $\mathcal{D}_{qlike}^{locPOAS}$. Fig. 7 b), d) and e) show color coded FA obtained from the quasi-likelihood estimates $\mathcal{D}_{qlike}^{orig}$, $\mathcal{D}_{qlike}^{glPOAS}$ and $\mathcal{D}_{qlike}^{locPOAS}$, respectively. We observe a clearly improved color coded FA in e) compared to b) (original data) and d) where the inadequate global σ in msPOAS leads to a spatially varying quality of the result. Fig. 7 c) and f) reflect the change in FA when employing quasi-likelihood instead of nonlinear regression for the original and the smoothed (local σ) data, respectively. The FA-differences are provided on a color scale with green referring to a zero difference. Both plots indicate a FA-dependent bias of the regression estimate, with the effect being much more prominent and stable for the smoothed data.

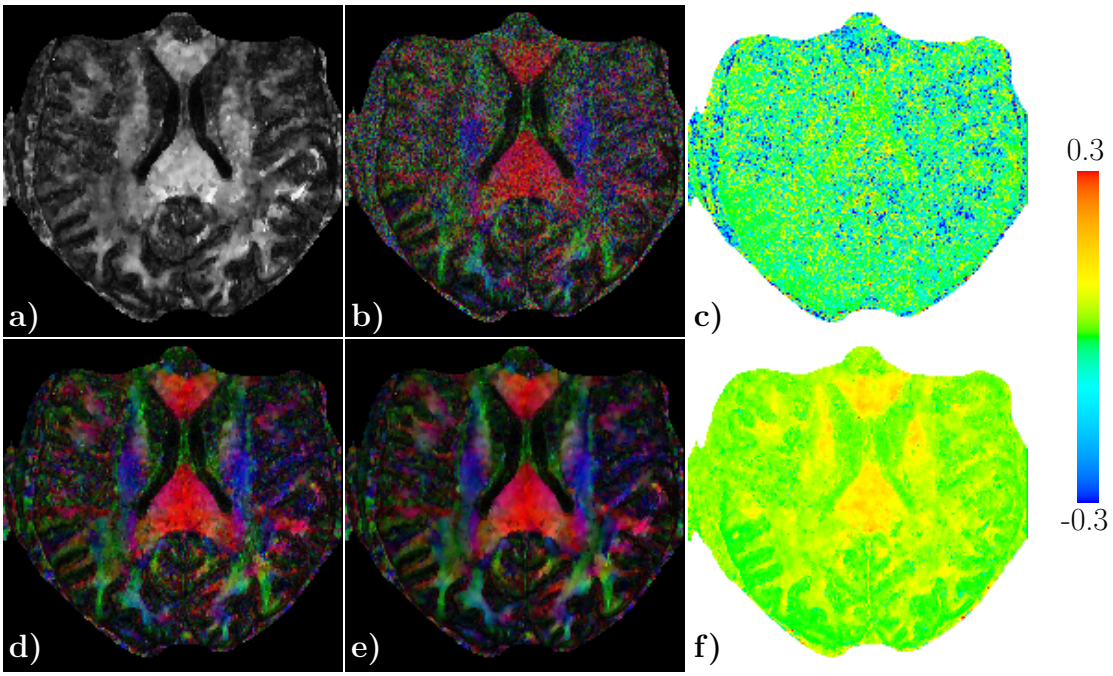

Figure 7 – *Fractional anisotropy (FA) maps for the first repetition of the diffusion weighted dataset. a) FA grayscale image, and e) corresponding color-coded FA image from data smoothed by msPOAS with local variance estimates. d) Color-coded FA estimates from data smoothed by msPOAS with a single global estimate of* σ*. b) Result for original (unsmoothed) data. a), b), d), and e) provide results obtained using the quasi-likelihood* (2.13)*. c) and f) refer to the local differences for the FA estimates when the "correct" quasi-likelihood is used compared to the "biased" nonlinear regression estimate Eq.* (2.11)*.*

5 Discussion and Conclusion

We developed a novel method for the *local* estimation of the noise parameter in magnetic resonance imaging *in the presence of an MR signal*. The method can thus be applied to estimate the noise level in regions with tissue, which is not accessible by methods that rely on the background distribution, see e.g. the comprehensive list of methods in Aja-Fernández et al. [2009a].

The method assumes the standardized MR signal to follow a non-central χ distribution with a locally constant non-centrality parameter, and spatially slowly varying noise parameter and number of degrees of freedom. The procedure is based on the propagation-separation approach [Polzehl and Spokoiny, 2006] to infer on local homogeneity regions of the non-centrality parameter of the signal distribution. Then a local weighted Maximum Likelihood estimator (over these regions) is used to obtain local estimates for the noise parameter. The parameters of the method influence the smoothness of the resulting estimate (e.g. h_{med}) or reduce the potential bias due to mis-specification of the homogeneity regions (e.g. λ).

In this paper we demonstrated for real and simulated data the robustness and validity of the method comparing it to the ground truth. Furthermore, we compared its performance with the intriguingly simple and fast method proposed by Aja-Fernández et al. [2013].

Noise estimation for magnetic resonance imaging data is often essential to evaluate image quality and to successfully run image enhancing methods like noise reduction. We showed in this paper, that local estimation of the noise estimation outperforms the use of a global parameter in the noise reduction method msPOAS [Becker et al., 2014] as an example. However, this improvement will also apply to other denoising methods, see e.g., Aja-Fernández et al. [2013].

An often underestimated fact is the bias introduced in parameter estimates of diffusion models for dMRI data like DTI or DKI at low SNR, see e.g. Jones and Basser [2004]. We also demonstrated, that the estimation of the model parameters (for DTI) will benefit from the local determination of the noise as proposed, see e.g. also Landman et al. [2009], Veraart et al. [2011a], Veraart et al. [2011b], or Ghosh et al. [2013] for related ideas. While the diffusion model parameter estimation using maximum likelihood based on the non-central χ distribution is valid for the original unsmoothed data, data processed by non-local noise reduction methods like msPOAS follows a different signal distribution, such that a quasi-likelihood formulation of the estimation problem is preferable.

The package primarily used for the analysis in this paper is freely available: **dti** [Tabelow and Polzehl, 2014].

6 Acknowledgment

We thank A. Anwander and R. Heidemannn for the permission to re-use the diffusion weighted dataset for comparison. This work is supported by the DFG Research Center MATHEON.

References

S. Aja-Fernández and A. Tristán-Vega. Influence of noise correlation in multiple-coil statistical models with sum of squares reconstruction. *Magnetic Resonance in Medicine*, 67(2):580–585, 2012.

S. Aja-Fernández, M. Niethammer, M. Kubicki, M.E. Shenton, and C.-F. Westin. Restoration of DWI data using a Rician LMMSE estimator. *IEEE Transactions on Medical Imaging*, 27: 1389–1403, 2008.

S. Aja-Fernández, A. Tristán-Vega, and C. Alberola-López. Noise estimation in single- and multiple-coil magnetic resonance data based on statistical models. *Magnetic Resonance Imaging*, 27(10):1397–1409, 2009a.

S. Aja-Fernández, G. Vegas-Sánchez-Ferrero, M. Martín-Fernández, and C. Alberola-López. Automatic noise estimation in images using local statistics. Additive and multiplicative cases. *Image and Vision Computing*, 27(6):756–770, 2009b.

S. Aja-Fernández, A. Tristán-Vega, and W.S. Hoge. Statistical noise analysis in grappa using a parametrized noncentral chi approximation model. *Magnetic Resonance in Medicine*, 65(4): 1195–1206, 2011.

S. Aja-Fernández, V. Brion, and A. Tristán-Vega. Effective noise estimation and filtering from correlated multiple-coil MR data. *Magnetic Resonance Imaging*, 31(2):272–285, 2013.

E.D. André, F. Grinberg, E. Farrher, I.I. Maximov, N.J. Shah, C. Meyer, M. Jaspar, V. Muto, C. Phillips, and E. Balteau. Influence of noise correction on intra- and inter-subject variability of quantitative metrics in diffusion kurtosis imaging. *PLoS One*, 9(4):e94531, 2014.

P.J. Basser and S. Pajevic. Statistical artefacts in diffusion tensor MRI (DT-MRI) caused by background noise. *Magnetic Resonance in Medicine*, 44(1):41–50, 2000.

P.J. Basser, J. Mattiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994a.

P.J. Basser, J. Mattiello, and D. Le Bihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance B*, 103:247–254, 1994b.

S.M.A. Becker, K. Tabelow, H.U. Voss, A. Anwander, R.M. Heidemann, and J. Polzehl. Position-orientation adaptive smoothing of diffusion weighted magnetic resonance data (POAS). *Medical Image Analysis*, 16(6):1142–1155, 2012.

S.M.A. Becker, K. Tabelow, S. Mohammadi, N. Weiskopf, and J. Polzehl. Adaptive smoothing of multi-shell diffusion-weighted magnetic resonance data by msPOAS. *NeuroImage*, 95: 90–105, 2014.

P.T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford Science Publications, 1991.

D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998.

C.D. Constantinides, E. Atalar, and E.R. McVeigh. Signal-to-noise measurements in magnitude images from NMR phased arrays. *Magnetic Resonance in Medicine*, 38(5):852–857, 1997.

P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Transactions on Medical Imaging*, 27(4):425–441, 2008.

A. Ghosh, T. Milne, and R. Deriche. Constrained diffusion kurtosis imaging using ternary quartics & MLE. *Magnetic Resonance in Medicine*, 71(4):1581–1591, 2013.

H. Gudbjartsson and S. Patz. The rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34:910–914, 1995.

J. P. Haldar, V. J. Wedeen, M. Nezamzadeh, G. Dai, M. W. Weiner, N. Schuff, and Z.-P. Liang. Improved diffusion imaging through SNR-enhancing joint reconstruction. *Magnetic Resonance in Medicine*, 69(1):277–289, 2013.

R.M. Heidemann, A. Anwander, T. Feiweier, T.R. Knösche, and R. Turner. k-space and q-space: Combining ultra-high spatial and angular resolution in diffusion imaging using ZOOPPA at 7T. *NeuroImage*, 60(2):967–978, 2012.

D.K. Jones, editor. *Diffusion MRI: Theory, Methods, and Applications*. Oxford University Press, 2010.

D.K. Jones and P.J. Basser. "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine*, 52:979–993, 2004.

C.G. Koay, J.D. Carew, A.L. Alexander, P.J. Basser, and M.E. Meyerand. Investigation of anomalous estimates of tensor-derived quantities in diffusion tensor imaging. *Magnetic Resonance in Medicine*, 55:930–936, 2006.

B. Landman, P.-L. Bazin, and J. Prince. Estimation and application of spatially variable noise fields in diffusion tensor imaging. *Magnetic Resonance Imaging*, 27:741–751, 2009.

J.N. Morelli, V.M. Runge, T. Feiweier, J.E. Kirsch, K.W. Williams, and U.I. Attenberger. Evaluation of a modified Stejskal-Tanner diffusion encoding scheme, permitting a marked reduction in TE, in diffusion-weighted imaging of stroke patients at 3 T. *Investigative Radiology*, 45(1): 29–35, 2010.

C. Pierpaoli and P.J. Basser. Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance in Medicine*, 36(6):893–906, 1996.

J. Polzehl and V. Spokoiny. Propagation-separation approach for local likelihood estimation. *Probability Theory and Related Fields*, 135:335–362, 2006.

P. Pullens, A. Roebroeck, and R. Goebel. Ground truth hardware phantoms for validation of diffusion-weighted MRI applications. *Journal of Magnetic Resonance Imaging*, 32(2):482–488, 2010.

J. Rajan, J. Veraart, J. Van Audekerke, M. Verhoye, and J. Sijbers. Nonlocal maximum likelihood estimation method for denoising multiple-coil magnetic resonance images. *Magnetic Resonance Imaging*, 30(10):1512–1518, 2012.

P.B. Roemer, W.A. Edelstein, C.E. Hayes, S.P. Souza, and O.M. Mueller. The NMR phased

array. *Magnetic Resonance in Medicine*, 16(2):192–225, 1990.

D.W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley, 1992.

J. Sijbers, A.J. den Dekker, J. Van Audekerke, M. Verhoye, and D. Van Dyck. Estimation of the noise in magnitude MR images. *Magnetic Resonance Imaging*, 16(1):87–90, 1998.

S.N. Sotiropoulos, S. Jbabdi, J. Xu, J.L. Andersson, S. Moeller, E.J. Auerbach, M.F. Glasser, M. Hernandez, G. Sapiro, M. Jenkinson, D.A. Feinberg, E. Yacoub, C. Lenglet, D.C. Van Essen, K. Ugurbil, and T.E.J. Behrens. Advances in diffusion MRI acquisition and processing in the human connectome project. *NeuroImage*, 80:125–143, 2013a.

S.N. Sotiropoulos, S. Moeller, S. Jbabdi, J. Xu, J.L. Andersson, E.J. Auerbach, E. Yacoub, D. Feinberg, K. Setsompop, L.L. Wald, T.E.J. Behrens, K. Ugurbil, and C. Lenglet. Effects of image reconstruction on fibre orientation mapping from multichannel diffusion MRI: Reducing the noise floor using SENSE. *Magnetic Resonance in Medicine*, 70(6):1682–1689, 2013b.

K. Tabelow and J. Polzehl. *dti: DTI/DWI Analysis*, 2014. URL `http://CRAN.R-project.org/package=dti`. R package version 1.2-0.

J. Veraart, D.H.J. Poot, W. Van Hecke, I. Blockx, A. Van der Linden, M. Verhoye, and J. Sijbers. More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging. *Magnetic Resonance in Medicine*, 65(1):138–145, 2011a.

J. Veraart, W. Van Hecke, and J. Sijbers. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. *Magnetic Resonance in Medicine*, 66 (3):678–686, 2011b.

K.J. Worsley, C. Liao, J.A.D. Aston, V. Petre, G.H. Duncan, F. Morales, and A.C. Evans. A general statistical analysis for fMRI data. *NeuroImage*, 15:1–15, 2002.