THIELE CENTRE
FOR APPLIED MATHEMATICS IN NATURAL SCIENCE

Bayesian regularization of diffusion tensor images

Jesper Frandsen, Asger Hobolth,
Eva B. Vedel Jensen, Peter Vestergaard-Poulsen
and Leif Østergaard

Research Report No. 03 I January 2004

Bayesian regularization of diffusion tensor images

This Thiele Research Report is also Research Report number 441 in the Stochastics Series at Department of Mathematical Sciences, University of Aarhus, Denmark.

Bayesian regularization of diffusion tensor images

Jesper Frandsen[1], Asger Hobolth[2], Eva B. Vedel Jensen[3],
Peter Vestergaard-Poulsen[1] and Leif Østergaard[1]

[1] *Department of Neuroradiology, Centre for Functionally Integrative Neuroscience, Århus University Hospital*
[2] *Bioinformatics Research Center, Aarhus University*
[3] *The T.N. Thiele Centre, Department of Mathematical Sciences, Aarhus University*

Abstract

Diffusion tensor imaging (DTI) is currently being refined as a tool to study the course of nerve fibre bundles in the human brain. Using DTI, the local fibre orientation in each image voxel can be described by a diffusion tensor which is constructed from local measurements of diffusion coefficients along several directions. The measured diffusion coefficients and thereby the diffusion tensors are subject to noise, leading to possibly flawed representations of the three dimensional fibre bundles. Efforts to reduce noise by regularization have so far been concentrated on the analysis of the primary diffusion direction. In this paper we develop a Bayesian procedure for regularizing the full diffusion tensor field, fully utilizing the available three-dimensional information of fibre orientation. The use of the procedure is exemplified on synthetic and *in vivo* data.

Key words: Bayesian regularization, diffusion tensor imaging, Markov chain Monte Carlo.

1 Introduction

The human brain may be divided into two main components, grey matter and white matter. The grey matter contains the functional centres of the brain and has been studied for the last two decades by magnetic resonance imaging. This type of imaging technique is essential for understanding how we perform cognitive tasks, and may also be used to study a wide range of diseases. The white matter connects the functional centres of the brain. The integrity and course of white matter fibre bundles are of key importance in understanding the structural basis of the functional integration of cortical centres in cognitive tasks, of the origin of functional impairment in focal brain lesions, and finally brain plasticity.

With the development of diffusion tensor imaging, white matter microstructures can be indirectly probed by measuring the directionality of Brownian movements of free water (Basser et al., 1994, and Basser and Pierpaoli, 1996). By the hindrance of water diffusion across the cell membranes of white matter fibre tracts, the local fibre direction is believed to be inferable from the preferred free water diffusion direction, as derived from diffusion coefficients measured locally along several directions. The local fibre orientation in each image voxel are often described in terms of a diffusion tensor which is constructed from the measured diffusion coefficients. The diffusion tensor field forms the basis for determining the course of entire white matter fibre bundles. In Figure 1, an example of a diffusion tensor field is shown.

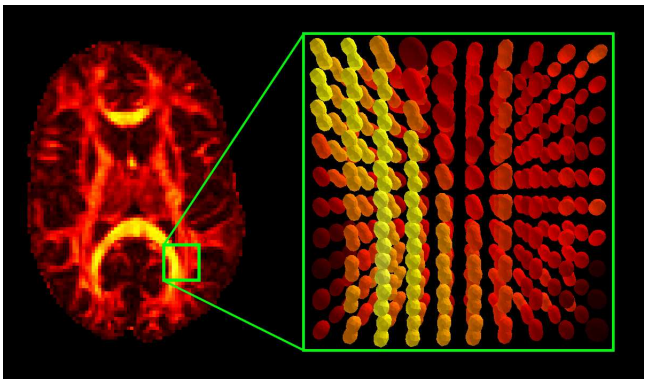

Figure 1: An example of a tensor field, obtained by diffusion tensor imaging. To the left, a slice of the brain is shown, indicating the position of the tensor field sample of size 10×10×3 shown to the right. Each tensor is represented by its diffusion function profile and its colour code by its fractional anisotropy index. For details, see Sections 2 and 6.

Due to the inherent noise of DTI measurements, the diffusion tensors are inaccurate, possibly leading to erroneous results in the derived white matter fibre paths. Voxels are likely to be part of a well organized structure (for example, a fibre direction perpendicular to that of all surrounding voxels in a well-ordered fibre bundle is very unlikely). By using Bayesian regularization techniques this *a priori* knowledge can be included in the analysis. In this paper we show that Bayesian regularization can be performed on the field of diffusion tensors, fully utilizing their

three-dimensional information on fibre orientation. The paper is a natural continuation of a paper by Poupon et al. (2000) on Bayesian regularization of the primary directions of the diffusion tensors. We use Markov chain Monte Carlo (MCMC) simulation procedures for the regularization and study the performance of the procedure on synthetic and *in vivo* data.

The paper is organized as follows. In Section 2 we describe the field of diffusion tensors, more generally the field of diffusion functions. In Sections 3 and 4, the prior and likelihood models are discussed. The Bayesian regularization procedure is described in detail in Section 5. In Section 6 we study the performance of the regularization procedure on synthetic data, and in Section 7 we consider *in vivo* data. An extension of the prior model is described in Section 8. Technical aspects of the regularization are collected in two appendices.

2 The field of diffusion functions

Let $\mathcal{W}$ be the finite set of voxels representing the white matter. For each voxel $w \in \mathcal{W}$ the true diffusion coefficient in a direction $u \in S^2$ on the unit sphere is denoted $f_w(u)$. Since the diffusion coefficient is the same in opposite directions we have

$$f_w(u) = f_w(-u), \quad u \in S^2.$$

We denote the field of diffusion functions by

$$\mathcal{F} = \{f_w : w \in \mathcal{W}\}.$$

As mentioned in the introduction, it is believed that the diffusion coefficient $f_w(u)$ is large if u is close to the main direction of the nerve fibres passing through w. One possibility for graphical representation of a diffusion function is the *diffusion function profile*

$$\{f_w(u)u : u \in S^2\} \tag{1}$$

which is the spatial surface having the distance $f_w(u)$ to the origin in the direction $u \in S^2$.

The diffusion function is commonly modelled as a quadratic form, cf. Basser et al. (1994). In this case

$$f_w(u) = u^* \Sigma_w u, \quad u \in S^2, \tag{2}$$

where Σ_w is a 3×3 positive semi-definite matrix, referred to as a tensor. The field of tensors will also be denoted by $\mathcal{F}$, for convenience. Let $\lambda_{w1} \geq \lambda_{w2} \geq \lambda_{w3} \geq 0$ be the eigenvalues of Σ_w with corresponding orthonormal eigenvectors u_{w1}, u_{w2} and u_{w3}. In Figure 2, diffusion function profiles of quadratic diffusion functions (tensors) are shown. The same type of graphical representaion of tensors is used in Figure 1. If $\lambda_{w1} > \lambda_{w2}$ a diffusion function of the form (2) attains its maximal value λ_{w1} in the unique unorientated direction $u = \pm u_{w1}$. This direction is called the *primary diffusion direction.*

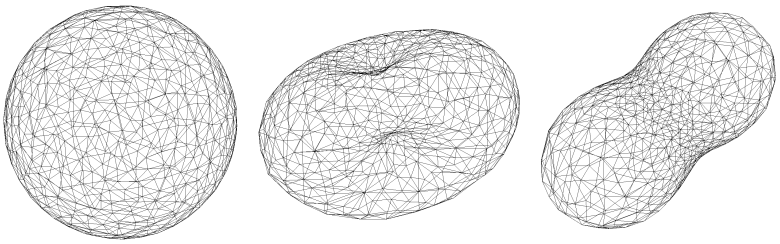

Figure 2: Diffusion function profiles for quadratic diffusion functions (tensors) with eigenvalues $\lambda_1 = \lambda_2 = \lambda_3$ (left), $\lambda_1 = \lambda_2 > \lambda_3$ (middle) and $\lambda_1 > \lambda_2 = \lambda_3$ (right).

3 Prior model

In order to perform Bayesian regularization, a prior model must be defined on the field of diffusion functions. In Bayesian analysis, Gibbs type distributions (also sometimes referred to as Markov random fields) have been successfully applied as prior models for low level tasks such as image restoration, see e.g. Geman & Geman (1984), Besag (1986), Guttorp (1995) and Hurn et al. (2003). In the present paper, the same type of approach is suggested for the field of diffusion functions. Under the prior model, diffusion functions in neighbour voxels tend to be similar.

Let $\sim$ be a neighbourhood relation on $\mathcal{W}$, indicating the voxels from which local information is taken into account during regularization. In 3D, the common neighbourhood structures are the 6, 26 or 32 nearest neighbours. In our application we use the 26 nearest neighbours. A general Gibbs type prior density of $\mathcal{F}$ is

$$p(\mathcal{F}) = \frac{1}{Z_\alpha} \exp\left(-\alpha \sum_{w \sim w'} \frac{d(f_w, f_{w'})}{||w - w'||} \right), \tag{3}$$

where Z_α is a normalizing constant and $\alpha > 0$. The summation involves all pairs of neighbour voxels. The function $d(\cdot, \cdot)$ measures the distance between diffusion functions and $||w - w'||$ is the Euclidean distance between voxel w and w'. In principle any distance between diffusion functions can be used, with the constraint that (3) should specify a proper prior, i.e. the integral of the exponential term should be finite. Note that under the prior model high probability fields $\mathcal{F}$ have the property that diffusion functions in neighbour voxels are similar.

As the diffusion function contains directional as well as size information the field of diffusion functions is commonly normalized before any further analysis (Basser and Pierpaoli, 1996). The function f_w is hence replaced by

$$\bar{f}_w(u) = \frac{f_w(u)}{\int_{S^2} f_w(v) \frac{dv}{4\pi}}, \quad u \in S^2. \tag{4}$$

Under the tensor model (2) we have

$$\int_{S^2} f_w(v) \frac{dv}{4\pi} = \frac{1}{3}(\lambda_{w1} + \lambda_{w2} + \lambda_{w3}) = \bar{\lambda}_w, \tag{5}$$

cf. Appendix A, and (4) reduces to

$$\bar{f}_w(u) = u^* \bar{\Sigma}_w u, \;\; \bar{\Sigma}_w = \Sigma_w / \bar{\lambda}_w, \;\; u \in S^2.$$

In the following we consider Gibbs type prior densities on the normalized diffusion function field

$$p(\mathcal{F}) = \frac{1}{Z_\alpha} \exp\Big(-\alpha \sum_{w \sim w'} \frac{d(\bar{f}_w, \bar{f}_{w'})}{||w - w'||} \Big). \tag{6}$$

There are many possible choices of distances between diffusion functions. Without specific assumptions on the form of the diffusion functions we may choose

$$d(\bar{f}_w, \bar{f}_{w'}) = g\Big(\int_{S^2} \big[\bar{f}_w(u) - \bar{f}_{w'}(u) \big]^2 \, du \Big), \tag{7}$$

where $g : [0, \infty) \to [0, \infty)$ is an increasing function with $g(0) = 0$. A simple choice would be $g(d) = \sqrt{d}$.

If the tensor model (2) holds, the integral of (7) reduces to

$$\int_{S^2} \big[\bar{f}_w(u) - \bar{f}_{w'}(u) \big]^2 \, du = \int_{S^2} \big[u^* (\bar{\Sigma}_w - \bar{\Sigma}_{w'}) u \big]^2 \, du = \frac{8\pi}{15} \big\| \bar{\Sigma}_w - \bar{\Sigma}_{w'} \big\|^2,$$

cf. Appendix A. Here the Frobenius norm

$$\|\Sigma\| = \big[\mathrm{tr}(\Sigma \Sigma^*) \big]^{1/2} = \Big[\sum_{i=1}^{3} \sum_{j=1}^{3} \sigma_{ij}^2 \Big]^{1/2},$$

cf. Basser and Pierpaoli (1996), is used on the space of 3×3 matrices $\Sigma = \{\sigma_{ij}\}$. Thus (7) reduces to

$$d(\bar{f}_w, \bar{f}_{w'}) = g(||\bar{\Sigma}_w - \bar{\Sigma}_{w'}||^2), \tag{8}$$

where the function g in (8) has the properties as in (7).

We now address the special case where only one eigenvalue is positive. Then

$$\int_{S^2} \big[\bar{f}_w(u) - \bar{f}_{w'}(u) \big]^2 \, du = \frac{48\pi}{5} (1 - [u_{w1} \cdot u_{w'1}]^2), \tag{9}$$

where $\cdot$ indicates inner product, cf. Appendix A. Here, $u_{w1} \cdot u_{w'1}$ is cosine of the angle between the two primary diffusion directions u_{w1} and $u_{w'1}$. In this case,

$$d(\bar{f}_w, \bar{f}_{w'}) = g(1 - [u_{w1} \cdot u_{w'1}]^2), \tag{10}$$

where g in (10) has the same properties as in (7). Poupon et al. (2000) use a model similar to (10) as prior for the primary diffusion direction.

4 Likelihood model

Let us suppose that the DTI data set consists of the following measured diffusion coefficients

$$F = \{F_w(u_i) : i = 1, \ldots, k, w \in \mathcal{W}\}.$$

Here, $u_1, \ldots, u_k$ are the directions in which the diffusion coefficients are measured. The minimum value of the number of directions k is 6.

The variability of a measured diffusion coefficient $F_w(u)$ at voxel w in direction u depends on the true value of the diffusion coefficient $f_w(u)$. Expressed more formally, we have

$$\mathrm{Var}(F_w(u)|\mathcal{F}) = h(f_w(u)). \tag{11}$$

If we assume that the measured diffusion coefficients are independent and normally distributed, then the likelihood becomes

$$p(F|\mathcal{F}) = \left[\prod_{w\in\mathcal{W}} \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi h(f_w(u_i))}} \right] \exp\left(-\frac{1}{2} \sum_{w\in\mathcal{W}} \sum_{i=1}^{k} \frac{(F_w(u_i) - f_w(u_i))^2}{h(f_w(u_i))} \right). \tag{12}$$

Because of detailed insight in the DTI scanning procedure, the likelihood model can be given theoretical support for a particular choice of h. In DTI, the diffusion coefficient $F_w(u)$ in direction u is determined by the equation

$$S_w(u) = S_{w0} \exp(-bF_w(u)),$$

where S_{w0} is the signal intensity without any gradient, $S_w(u)$ is the measured signal intensity and b is the diffusion encoding strength factor. The signal intensities $S_w(u)$ and S_{w0} follow Rice distributions[1], $S_w(u) \sim R(a, \sigma^2)$ and $S_{w0} \sim R(a_0, \sigma^2)$, cf. Sijbers et al. (1998). The signal to noise ratio is defined for $S_w(u)$ and S_{w0} as $\mathrm{SNR} = a/\sigma$ and $\mathrm{SNR}_0 = a_0/\sigma$, respectively. If SNR is not too small ($\mathrm{SNR} \geq 8$), then according to Appendix B the approximation

$$F_w(u) \sim N\Big(f_w(u), \frac{\exp(2bf_w(u)) + 1}{(b\mathrm{SNR}_0)^2}\Big) \tag{13}$$

holds, i.e. noise is Gaussian. Under the tensor model, we have

$$F_w(u) \sim N\Big(\bar{\lambda}_w u^* \bar{\Sigma}_w u, \frac{\exp(2b\bar{\lambda}_w u^* \bar{\Sigma}_w u) + 1}{(b\mathrm{SNR}_0)^2}\Big). \tag{14}$$

Recall from (5) that

$$\bar{\lambda}_w = \int_{S^2} E(F_w(u)) \frac{du}{4\pi},$$

and therefore $\bar{\lambda}_w$ can be estimated by

$$\hat{\lambda}_w = \frac{1}{k} \sum_{i=1}^{k} F_w(u_i),$$

if the experimental directions $u_1, \ldots, u_k$ are approximately systematically uniformly distributed on S^2.

[1]We say that $S \sim R(a, \sigma^2)$ if S has the same distribution as $\sqrt{X^2 + Y^2}$ where X and Y are independent random variables, $X \sim N(a\cos\phi, \sigma^2)$ and $Y \sim N(a\sin\phi, \sigma^2)$.

5 The Bayesian regularization procedure

We have used the Metropolis-Hastings algorithm (see, for instance, Gilks et al. (1996)) to simulate a regularized field of diffusion functions from the posterior distribution

$$p(\mathcal{F}|F) \propto p(F|\mathcal{F})p(\mathcal{F}).$$

We have worked under the tensor model in which case the likelihood $p(F|\mathcal{F})$ given by (12) is determined by

$$f_w(u) = \hat{\lambda}_w u^* \bar{\Sigma}_w u, \qquad h(f_w(u)) = \frac{\exp(2bf_w(u)) + 1}{(b\mathrm{SNR}_0)^2}.$$

The prior $p(\mathcal{F})$ takes the form, cf. (6) and (8) with $g(d) = \sqrt{d}$,

$$p(\mathcal{F}) = \frac{1}{Z_\alpha} \exp\left(-\alpha \sum_{w \sim w'} \frac{\left\| \bar{\Sigma}_w - \bar{\Sigma}_{w'} \right\|}{||w - w'||} \right).$$

We start from some suitable initial field

$$\mathcal{F}_0 = \{\bar{\Sigma}_w(0) : w \in \mathcal{W}\}.$$

The transition from the field $\mathcal{F}_t$ at time t to the field $\mathcal{F}_{t+1}$ at time $t+1$ is performed as follows:

1. Choose a normalized diffusion tensor $\bar{\Sigma}_\omega,\ \omega \in \mathcal{W}$, for updating.

2. Sample a candidate normalized diffusion tensor $\bar{\Sigma}'_\omega$ from a proposal distribution $q(\cdot|\mathcal{F}_t)$, which may or may not depend on the current state $\mathcal{F}_t$ of the chain.

3. Accept the candidate tensor with probability β, where

$$\beta = \min\left\{ \frac{p(\mathcal{F}'|F)q(\mathcal{F}_t|\mathcal{F}')}{p(\mathcal{F}_t|F)q(\mathcal{F}'|\mathcal{F}_t)}, 1 \right\}.$$

   Here $\mathcal{F}'$ is identical to $\mathcal{F}_t$ except for the diffusion tensor at ω, which is $\bar{\Sigma}'_\omega$.

4. If $\bar{\Sigma}'_\omega$ is accepted, then $\mathcal{F}_{t+1} = \mathcal{F}'$, else $\mathcal{F}_{t+1} = \mathcal{F}_t$.

The choice of diffusion tensor to update may be done at random or according to a sequential ordering of the voxels. Another possibility is to use a random permutation scheme where each diffusion tensor is visited once during a sweep of the algorithm, but in random order, see e.g. Guttorp (1995, page 55).

In theory, the proposal distribution q may have any form (for regularity conditions, see Roberts, 1996), but in practise it is important to select q carefully so that the Markov chain $\{\mathcal{F}_t\}$ moves rapidly around the tensor space, yet has a reasonable proportion of candidate tensors accepted. This can be achieved by ensuring that the candidate tensor $\bar{\Sigma}'_\omega$ is close to the actual tensor $\bar{\Sigma}_\omega$ of the chain, yet not so close that it takes many steps to move around in tensor space.

The tensor space consists of 3×3 normalized positive definite matrices. The Wishart distribution generates positive definite matrices and a natural choice for the proposal distribution is therefore a *normalized* Wishart distribution. Let X be distributed according to the Wishart distribution $W_3(\bar{\Sigma}/n, n)$ with mean $\bar{\Sigma}$ and degrees of freedom n. We then propose to use the distribution of the normalized matrix $\bar{X}$ as proposal distribution. In order to calculate the acceptance probability we need the density of $\bar{X}$ which is derived in the following proposition.

Proposition 1. Let $X \sim W_3(\bar{\Sigma}/n, n)$. Then the density of the normalized matrix $\bar{X}$ is given by

$$\frac{3}{\pi^{3/2}} \frac{\Gamma(\frac{3n}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{n-1}{2})\Gamma(\frac{n-2}{2})} \frac{|\bar{X}|^{(n-4)/2}}{\operatorname{tr}(\bar{\Sigma}^{-1}\bar{X})^{3n/2}|\bar{\Sigma}|^{n/2}}. \tag{15}$$

Proof. According to Anderson (1958, Theorem 7.2.2), the density of X with respect to the Lebesgue measure in $\mathbf{R}^6$ is

$$\frac{(\frac{n}{2})^{3n/2}}{\pi^{3/2}} \frac{1}{\Gamma(\frac{n}{2})\Gamma(\frac{n-1}{2})\Gamma(\frac{n-2}{2})} \frac{|X|^{(n-4)/2} e^{-n \operatorname{tr}(\bar{\Sigma}^{-1}X)/2}}{|\bar{\Sigma}|^{n/2}}, \quad n \geq 3.$$

Now consider the transformation

$$g : X = (x_{11}, x_{21}, x_{22}, x_{31}, x_{32}, x_{33}) \rightarrow$$
$$\left(\frac{x_{11}}{\bar{t}_X}, \frac{x_{21}}{\bar{t}_X}, \frac{x_{22}}{\bar{t}_X}, \frac{x_{31}}{\bar{t}_X}, \frac{x_{32}}{\bar{t}_X}, \bar{t}_X\right) = (\bar{X}, \bar{t}_X) = g(X),$$

where $\bar{t}_X = (x_{11} + x_{22} + x_{33})/3$. Note that the first five components $\bar{X}$ of $g(X)$ correspond to the normalized matrix, while the last component $\bar{t}_X$ is the scaling factor. The Jacobian of g^{-1} is $3\,\bar{t}_X^5$ and the density of $g(X)$ becomes

$$\frac{3(\frac{n}{2})^{3n/2}}{\pi^{3/2}} \frac{1}{\Gamma(\frac{n}{2})\Gamma(\frac{n-1}{2})\Gamma(\frac{n-2}{2})} \frac{\bar{t}_X^{\frac{3n}{2}-1} |\bar{X}|^{(n-4)/2} e^{-n\bar{t}_X \operatorname{tr}(\bar{\Sigma}^{-1}\bar{X})/2}}{|\bar{\Sigma}|^{n/2}}.$$

Integrating out with respect to $\bar{t}_X$ the density of the normalized matrix $\bar{X}$ is given by expression (15) with respect to the Lebesgue measure in $\mathbf{R}^5$. □

Let $\lambda_i,\ i = 1, 2, 3$, denote the eigenvalues of the product $\bar{\Sigma}'^{-1}_{\omega}\bar{\Sigma}_{\omega}$. Then we get from Proposition 1

$$\frac{q(\mathcal{F}_t|\mathcal{F}')}{q(\mathcal{F}'|\mathcal{F}_t)} = (\lambda_1\lambda_2\lambda_3)^{n-2} \left[\frac{\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3}}{\lambda_1 + \lambda_2 + \lambda_3}\right]^{3n/2}.$$

For the calculation of the posterior ratio $p(\mathcal{F}'|F)/p(\mathcal{F}_t|F)$ it is important to notice that the chain only updates the diffusion tensor in one voxel at a time and therefore only terms in (6) and (12) involving this particular voxel contribute to the ratio.

Using that the density of $X = \sum_{i=1}^{n} Z_i Z_i^T$ is a $W_3(\bar{\Sigma}/n, n)$ if $Z_i, i = 1, \ldots, n$ $(n \geq 3)$, are independent, each according to $N_3(0, \bar{\Sigma}/n)$, it is easy to simulate from the Wishart distribution. The value of the degrees of freedom n determines how close the candidate tensor is to the current tensor; as n increases the density of the proposal becomes more concentrated around the mean $\bar{\Sigma}$.

6 Synthetic data

To study the performance of the regularization procedure we analyzed the torus model presented in Tournier et al. (2002). The fibre bundle is modelled as a torus of radius R and circular cross-section of radius r. The side length of a voxel is taken to be one unit length and each voxel is represented by its centre point.

At a voxel w inside the torus, a tensor Σ_w is positioned with primary diffusion direction horizontal and perpendicular to the ray from O to w. The eigenvalues of Σ_w satisfy

$$\lambda_{w1} > \lambda_{w2} = \lambda_{w3} > 0$$

and are normalized such that $\bar{\lambda}_w = 1$ and $\Sigma_w = \bar{\Sigma}_w$. The degree of anisotropy is the same for all tensors and is determined by $\delta = \lambda_{w2}/\lambda_{w1}$. Since $\bar{\lambda}_w = 1$, we furthermore have

$$\lambda_{w1} = 3/(1 + 2\delta).$$

In the DTI literature, the anisotropy is often specified by the fractional anisotropy index

$$\text{FA} = \sqrt{\frac{\frac{1}{2}\sum_{i=1}^{3}(\lambda_{wi} - \bar{\lambda}_w)^2}{\frac{1}{3}\sum_{i=1}^{3}\lambda_{wi}^2}},$$

cf. e.g. Basser and Pierpaoli (1996). In terms of δ, FA can be expressed as

$$\text{FA} = \sqrt{\frac{(1-\delta)^2}{1+2\delta^2}}.$$

If a voxel w is outside the torus we let $\Sigma_w = \bar{\Sigma}_w = I_3$. There is thereby no preferred direction, $\delta = 1$ and FA $= 0$. If the voxel is on the boundary of the torus, the diffusion tensors are determined as a volume weighted average of the tensors outside and inside the torus. Finally, noise is added in all k directions for each voxel, according to the observation model (14) with $\bar{\lambda}_w = 1$.

The effect of regularization was assessed as follows. The normalized tensor field

$$\{\bar{\Sigma}_w : w \in \mathcal{W}\}$$

was calculated and the noisy data

$$F_w(u_i), \quad i = 1, \ldots, k, \quad w \in \mathcal{W},$$

were simulated. Next, a least squares method was used to fit a 3×3 positive definite tensor matrix $\bar{\Sigma}_w(0)$ to the data in each voxel, cf. Fraleigh and Beauregard (1990, page 315). The corresponding normalized tensor field is denoted

$$\{\bar{\Sigma}_w(0) : w \in \mathcal{W}\}.$$

This tensor field was used as starting value for the Markov chain Monte Carlo (MCMC) regularization algorithm described in Section 5.

The regularized tensor field at iteration t of the algorithm is denoted

$$\{\bar{\Sigma}_w(t) : w \in \mathcal{W}\}.$$

The average gain obtained by regularization was assessed by comparing the regularized tensor field at iteration t to the true tensor field, i.e. by considering the average tensor difference

$$\frac{1}{|\mathcal{W}|}\sum_{w\in\mathcal{W}}||\bar{\Sigma}_w(t)-\bar{\Sigma}_w||$$

as a function of iteration number t.

Figure 3 shows the results obtained in an example where the parameters were chosen as $R = 7.5$, $r = 3$ (in voxel units), FA $= 0.6$, $\alpha = 7.5$, $k = 17$, $b = 1000$ and $\mathrm{SNR}_0 = 25$. These parameters are typical in a DTI setting. Tensors positioned outside the torus were not regularized. In the upper panel of Figure 3, a central section of the torus is shown. At each voxel, the projection onto the section of the primary diffusion direction is shown before (left) and after (right) regularization. An example of fibre tracking in the raw and regularized fields of tensors is also shown in Figure 3, using an algorithm presented in Mori et al. (1999). Regularization seemingly results in a more well-defined fibre orientation distribution. In the lower part of Figure 3, the average tensor difference is shown as a function of iteration number.

7 *In vivo* data

All the scanning experiments were performed on a normal adult male volunteer in a 1.5 T GE Signa system (GE Medical Systems, Milwaukee, WI, USA). The diffusion weighted scanning was performed in k=14 directions isotropically distributed in space. The diffusion encoding strength was characterized by a b-factor of 1000 s/mm^2. The scanning resulted in 55 consecutive sections of the brain, each of thickness 2.5 mm. Each section consisted of 128 by 128 pixels, covering an area of 22 by 22 cm. The dimensions of a voxel is therefore 1.7×1.7×2.5 mm. The scanning time was 306 seconds. In addition, two scannings without diffusion encoding ($b =$ 0 s/mm^2) were acquired. This design was repeated 10 times providing 10 individual data sets. Finally a T1 weighted sequence provided the anatomical images shown in Figure 4 below. The observed diffusion coefficients will be denoted

$$F_{w,i,j},\quad w\in\mathcal{W}, i=1,\ldots,k, j=1,\ldots,10,$$

where $F_{w,i,j}$ is the observed coefficient at voxel $w\in\mathcal{W}$, direction i and replication j. In order to get an impression of the individual data available we show in Figure 4 the central 10×10 slice of a representative 10×10×9 cube $\mathcal{W}'$ and its position in the brain.

In order to find an appropriate value of the prior parameter α, we fitted a normalized tensor $\bar{\Sigma}_{wj}$ to each individual data set j at each voxel w. The averages

$$\bar{\Sigma}_w=\frac{1}{10}\sum_{j=1}^{10}\bar{\Sigma}_{wj}$$

were regarded as the 'true' tensors in an subsequent estimation of α. The total prior

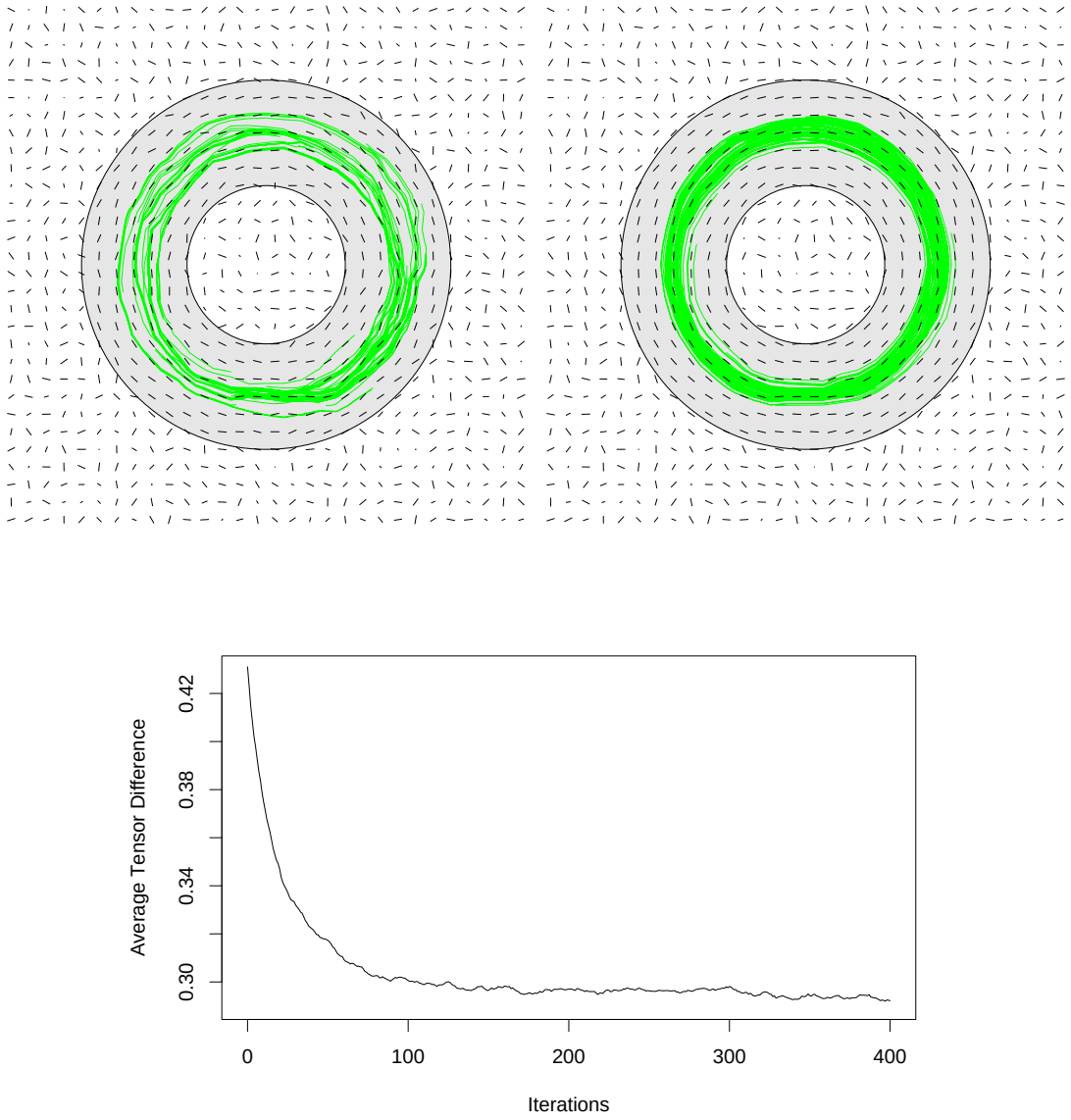

Figure 3: Upper left: Image of the projection of the primary diffusion direction and the corresponding fibre tracking before regularization. Upper right: Image of the projection of the primary diffusion direction and the corresponding fibre tracking after regularization. Lower: Plot of the average difference between the true diffusion tensors and the regularized diffusion tensors as a function of iteration number.

tensor difference

$$\sum_{w \sim w'} \frac{||\bar{\Sigma}_w - \bar{\Sigma}_{w'}||}{||w - w'||} \tag{16}$$

is a sufficient statistic of the prior model which belongs to the family of exponential models. According to standard exponential family theory, the maximum likelihood estimate of α is the value at which the mean of the sufficient statistic equals the observed value. Figure 5 shows the mean of the sufficient statistic as a function

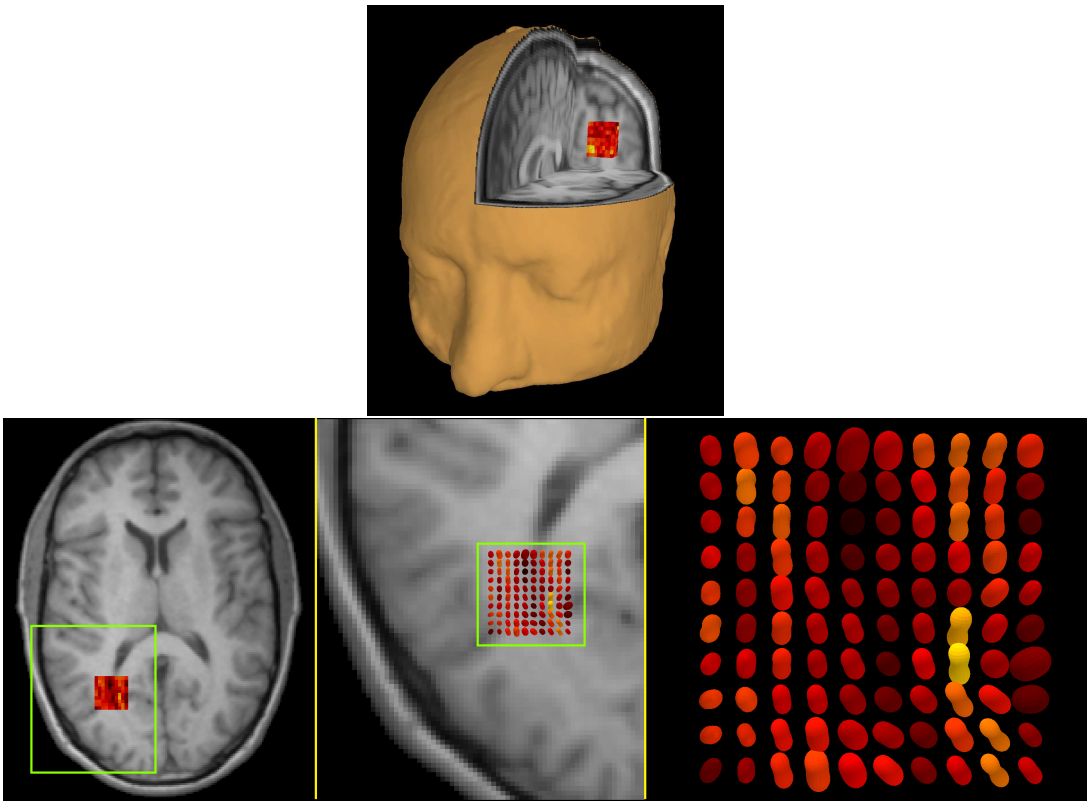

Figure 4: A slice of tensors determined from an individual data set and its position in the human brain.

of α, determined by MCMC simulation, for voxels in the $10\times10\times9$ cube $\mathcal{W}'$. The observed value of the sufficient statistic is also indicated in Figure 5. The value of α was estimated to be 7.5.

We applied our regularization procedure to each of the ten individual data sets. All individual data sets gave similar results. In Figure 6 the average tensor difference

$$\frac{1}{|\mathcal{W}|}\sum_{w\in\mathcal{W}}||\bar{\Sigma}_{w1}(t)-\bar{\Sigma}_w||$$

is shown for one of the data sets, as a function of iteration number t together with the corresponding plot for the voxels in $\mathcal{W}'$.

In Figure 7, a central slice of tensors are shown before and after regularization together with the average tensor field. The overall impression is that the tensors after regularization resemble more closely the average tensors. The changes are moderate but important from a practical point of view. Analyzing averages of two data sets we found that the precision of the regularized tensors based on a single scan corresponds to that of unregularized tensors based on two scans. The scanning time may therefore be reduced by a factor of 2 if the regularization procedure is used.

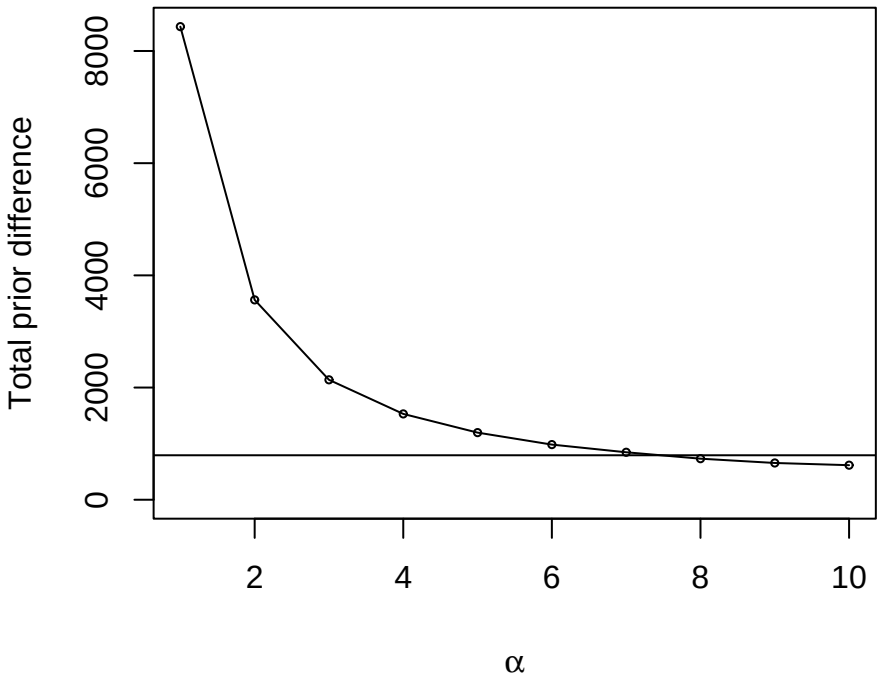

Figure 5: Total prior tensor difference as a function of α for the representative $10\times10\times9$ cube. The observed total tensor difference is 793 resulting in an α-estimate of 7.5.

8 An extension of the prior model

It has been stressed recently that the diffusion in voxels with e.g. crossing fibres can be quite complicated and not necessarily well characterized by the standard diffusion tensor model (2), cf. Frank (2001).

In Frank (2002) and Alexander et al. (2002), a method of extending the tensor model to more complex configurations of fibres is presented. The diffusion function profile is modelled by truncating a spherical harmonic expansion at several orders. If the spherical harmonic expansion is truncated at order 0 an isotropic tensor is obtained, truncating at order 2 gives the tensor model, while higher order truncations result in more complex diffusion function profiles. In order to describe multi directional fibres and improve the regularization such extensions are needed. The prior model described in the present paper can be generalized to accommodate various complex fibre configurations. One possibility is to use the notion of a marked field of diffusion functions

$$\mathcal{F}_m = \{(f_w, m_w) : w \in \mathcal{W},\ m_w \in \{1, 2\}\}.$$

Here the mark $m_w = 1$ refers to a voxel with one predominant fibre direction while

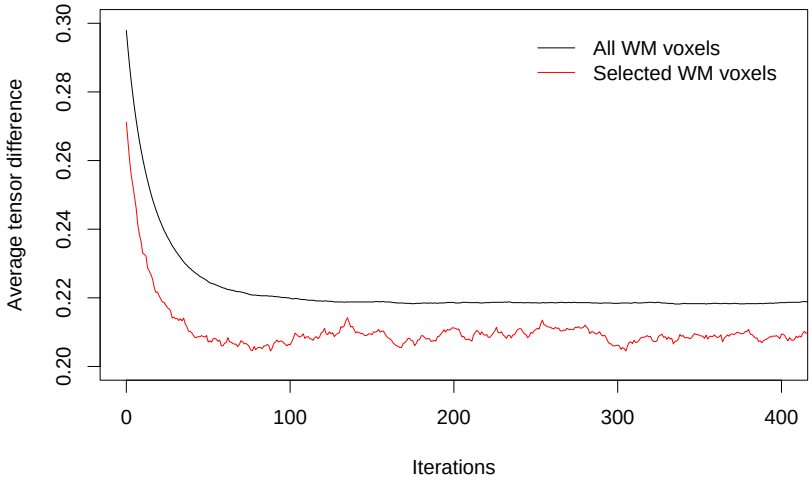

Figure 6: Average tensor difference between tensors based on an individual data set and the pooled data set. The upper curve is for all voxels in white matter while the lower curve is for voxels in the 10×10×9 subsample.

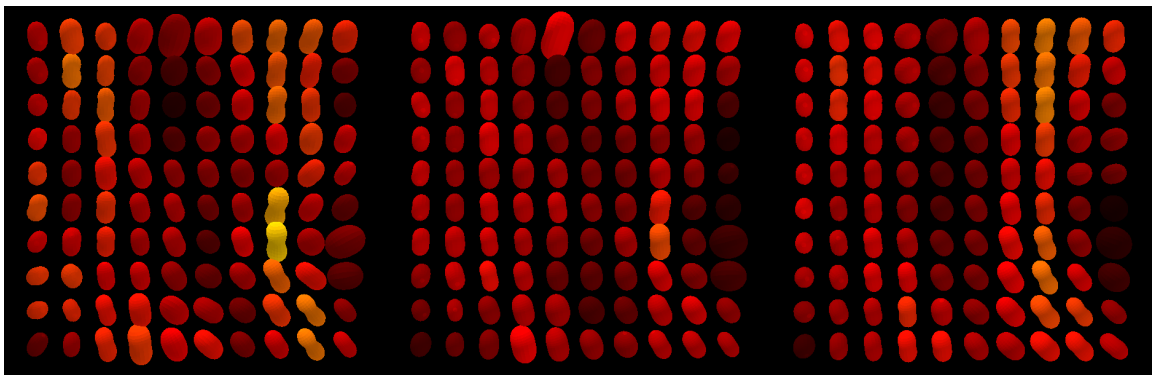

Figure 7: Left: Tensors based on an individual data set before regularization. Middle: Tensors after regularization (400 iterations). Right: Average tensor field.

$m_w = 2$ refers to a fibre crossing. If we let $n(\mathcal{F}_1)$ be the number of voxels with mark 1, then a generalized prior model is

$$p(\mathcal{F}_m) = \frac{1}{Z_{\alpha,\beta}} \beta^{n(\mathcal{F}_1)} \exp\left(-\alpha \sum_{w \sim w'} \frac{d((f_w, m_w), (f_{w'}, m_{w'}))}{||w - w'||} \right).$$

The parameter $\beta > 0$ determines the fraction of voxels with crossing fibres. The distance function should depend on whether the voxels w and w' both represent unidirectional fibre bundles or one of them is a fibre crossing location.

Appendix A

In this appendix, we derive some results concerning integrals of diffusion functions that can be described by the tensor model.

The first result concerns the normalization of a diffusion function of the form

$$f_w(u) = u^* \Sigma_w u.$$

Let us write Σ_w as

$$\Sigma_w = A^* \Lambda A,$$

where Λ is a diagonal 3×3 matrix with diagonal elements equal to the eigenvalues of Σ_w and A is an orthogonal 3×3 matrix with row vectors equal to the eigenvectors of Σ_w. Then,

$$\int_{S^2} (u^* \Sigma_w u) \, \frac{du}{4\pi} = \frac{1}{4\pi} \sum_{i=1}^{3} \lambda_{wi} \int_{S^2} (u_{wi}^* u)^2 \, du.$$

The integrals on the right-hand side of this equation do not depend on i. Letting $u_0 = (0, 0, 1)^*$, the common value of the integrals becomes

$$\int_{S^2} (u_0^* u)^2 \, du = \int_0^{\pi} \int_0^{2\pi} (\cos^2 \theta) \, \sin\theta \, d\varphi \, d\theta = \frac{4\pi}{3},$$

where we have used polar coordinates. It follows that

$$\int_{S^2} (u^* \Sigma_w u) \, \frac{du}{4\pi} = \frac{1}{3}(\lambda_{w1} + \lambda_{w2} + \lambda_{w3}).$$

Next, we derive the explicit expression for

$$\int_{S^2} \left[\bar{f}_w(u) - \bar{f}_{w'}(u) \right]^2 du \tag{17}$$

under the tensor model. Let us write

$$\bar{\Sigma}_w - \bar{\Sigma}_{w'} = A^* \Lambda A,$$

where Λ and A now contains the eigenvalues and eigenvectors of $\bar{\Sigma}_w - \bar{\Sigma}_{w'}$. If the eigenvalues of $\bar{\Sigma}_w - \bar{\Sigma}_{w'}$ are denoted μ_i, $i = 1, 2, 3$, then

$$\begin{aligned}
&\int_{S^2} \left[u^* (\bar{\Sigma}_w - \bar{\Sigma}_{w'}) u \right]^2 du \\
&\qquad = \int_{S^2} \left[\mu_1 u_1^2 + \mu_2 u_2^2 + \mu_3 u_3^2 \right]^2 du \\
&\qquad = \int_0^{\pi} \int_0^{2\pi} \left[\mu_1 \sin^2\theta \cos^2\varphi + \mu_2 \sin^2\theta \sin^2\varphi + \mu_3 \cos^2\theta \right]^2 \, \sin\theta \, d\varphi \, d\theta \\
&\qquad = \frac{8\pi}{15} \left\| \bar{\Sigma}_w - \bar{\Sigma}_{w'} \right\|^2.
\end{aligned}$$

The last equality sign represents elementary, but somewhat lengthy calculations. It is used that $\mu_1 + \mu_2 + \mu_3 = 0$ since $\bar{\Sigma}_w$ and $\bar{\Sigma}_{w'}$ are normalized tensors.

Finally, we want to find the explicit expression for (17) in the case where the tensors have exactly one positive eigenvalue. Because of the normalization, the common eigenvalue of the tensors is $\lambda = 3$. Furthermore,

$$u^*\bar{\Sigma}_w u = \lambda(u^*u_{w1})^2, \quad u^*\bar{\Sigma}_{w'} u = \lambda(u^*u_{w'1})^2.$$

We find

$$\begin{aligned}\int_{S^2}\left[\bar{f}_w(u) - \bar{f}_{w'}(u)\right]^2 du &= \lambda^2 \int_{S^2}\left[(u^*u_{w1})^2 - (u^*u_{w'1})^2\right]^2 du \\ &= \lambda^2\Big[2\int_{S^2}(u^*u_{w1})^4\, du - 2\int_{S^2}(u^*u_{w1})^2(u^*u_{w'1})^2\, du\Big].\end{aligned}$$

The two integrals can be evaluated, using polar coordinates. Without loss of generality, it can be assumed that $u_{w1} = (0,0,1)^*$, and we obtain the result (9).

Appendix B

The Rice distribution $R(a,\sigma^2)$ with parameters $a, \sigma^2 > 0$ is the distribution of $\sqrt{X^2+Y^2}$, where X and Y are independent random variables, and

$$X \sim N(a\cos\phi, \sigma^2), \;\; Y \sim N(a\sin\phi, \sigma^2).$$

If $\mathrm{SNR} = a/\sigma$ is large enough, then

$$-\ln\sqrt{X^2+Y^2}$$

can be regarded as normally distributed. The parameters in the approximating normal distribution can be found by linearizing the function

$$f(x,y) = -\ln\sqrt{x^2+y^2}$$

around $f(a\cos\phi, a\sin\phi) = -\ln a$. We find

$$f(x,y) = -\ln a - \frac{\cos\phi}{a}(x - a\cos\phi) - \frac{\sin\phi}{a}(x - a\sin\phi).$$

Therefore,

$$-\ln\sqrt{X^2+Y^2} \approx N(-\ln a, \frac{\sigma^2}{a^2}).$$

It follows that the distribution of $F_w(u)$ is approximately

$$F_w(u) \approx N\left(\frac{1}{b}\ln\frac{a_0}{a}, \frac{\sigma^2}{b^2}(\frac{1}{a^2} + \frac{1}{a_0^2})\right).$$

Writing $f_w(u) = \frac{1}{b}\ln\frac{a_0}{a}$ and $\mathrm{SNR}_0 = \frac{a_0}{\sigma}$, we get (13).

References

Alexander, D.C., Barker, G.J. and Arridge, S.R. 2002. Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. Magn Reson Med. 48(2): 331–340.

Anderson, T.W. 1958. An introduction to multivariate statistical analysis. Wiley, New York.

Basser, P.J., Mattiello, J. and LeBihan, D. 1994. MR diffusion tensor spectroscopy and imaging. Biophys J. 66(1): 259–267.

Basser, P.J. and Pierpaoli, C. 1996. Microstructural and physiological features of tissues elucidated by quantitive-diffusion-tensor MRI. J Magn Reson B. 111(3): 209–219.

Besag, J.E. 1986. Statistical analysis of dirty pictures. J. Roy. Statist. Soc. Ser. B; 48: 259–302.

Fraleigh and Beauregard 1990. Linear Algebra. Addison-Wesley, New York.

Frank, L.R. 2001. Anisotropy in high angular resolution diffusion-weighted MRI. Magn Reson Med. 45(6): 935–939.

Frank, L.R. 2002. Characterization of Anisotropy in High Angular Resolution Diffusion-Weighted MRI. Magn Reson Med. 47:1083–1099.

Geman, S. and Geman, D. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern analysis and Machine Intelligence; 6: 721–741.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. 1996. Introducing Markov chain Monte Carlo. In: W.R. Gilks, S. Richardson, and D.J. Spiegelhalter (Eds.), Markov Chain Monte Carlo in Practice, pp. 1–19.

Guttorp, P. 1995. Stochastic modeling of scientific data. Chapman and Hall, London.

Hurn, M., Husby, O. and Rue, H. 2003. Advances in Bayesian image analysis. In: P. Green, N. L. Hjorth, and S. Richardson (Eds.), Highly Structured Stochastic Systems, Oxford University Press, Oxford, pp. 302–322.

Mori S., Crain B.J., Chacko V.P., van Zijl P.C.M. 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann Neurol. 45(2): 265–269.

Poupon, C., Clark, C.A., Frouin, V., Régis, J., Bloch, I., Le Bihan, D., and Mangin, J.-F. 2000. Regularization of diffusion-based directional maps for the tracking of brain white matter fascicles. Neuroimage 12(2): 184–195.

Roberts, G.O. 1996. Markov chain concepts related to sampling algorithms. In: W.R. Gilks, S. Richardson, and D.J. Spiegelhalter (Eds), Markov Chain Monte Carlo in Practice, pp. 45–54.

Sijbers J., den Dekker A. J., Verhoye M., Van Audekerke J., and Van Dyck D. 1998. Estimation of noise from magnitude MR images. Magn Reson Imaging. 16(1): 87–90.

Tournier, J.D., Calamante, F., King, M.D., Gadian, D.G. and Connelly, A. 2002. Limitations and requirements of diffusion tensor fiber tracking: an assessment using simulations. Magn Reson Med. 47(4): 701–8.

Acknowledgements

We would like to thank Dr. Cyril Poupon, Prof. Carsten Gyldensted and Dr. Fabien Schneider for valuable discussions. This work is supported by grants from the Danish National Research Foundation and the Danish Natural Science Research Council.