

# Uncertainty in White Matter Fiber Tractography

Ola Friman and Carl-Fredrik Westin

Laboratory of Mathematics in Imaging,  
Department of Radiology Brigham and Women's Hospital, Harvard Medical School

**Abstract.** In this work we address the uncertainty associated with fiber paths obtained in white matter fiber tractography. This uncertainty, which arises for example from noise and partial volume effects, is quantified using a Bayesian modeling framework. The theory for estimating the probability of a connection between two areas in the brain is presented, and a new model of the local water diffusion profile is introduced. We also provide a theorem that facilitates the estimation of the parameters in this diffusion model, making the presented method simple to implement.

## 1 Introduction

When estimating a quantity based on some available data, it is natural to ask what uncertainty is associated with this estimate. In white matter fiber tractography, connections in the brain are estimated by tracing the direction of maximal water diffusion in diffusion weighted MRI images. In this process, uncertainty arises due to noise and partial volume effects, i.e., crossing, merging and splitting fiber tracts. Uncertainty is naturally quantified by means of probability, and in this work we are interested in the probability of a fiber connection between two points  $A$  and  $B$  in the brain, given the measured diffusion data  $\mathcal{D}$ . We write this probability  $p(A \rightarrow B|\mathcal{D})$ . There is some previous work in the literature on this subject [1–6], but most of these approaches account for uncertainty in heuristic ways. Exceptions are Tuch [3] and Behrens et al. [6], who were the first to present approaches with stronger theoretical support. In this paper we build on the work by Tuch and Behrens et al., and present a more detailed theory as well as a computationally tractable solution.

## 2 Global Modeling and Estimation

A fiber can be modeled as a finite length path, described by a train of vectors  $\mathbf{v}_{1:n} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ . Denote by  $\Omega_A^n$  the set of all possible paths of length  $n$  that originate in a point  $A$ , and assume that we can assign a probability  $p(\mathbf{v}_{1:n})$  to each path in this space. Hence,

$$\int_{\Omega_A^n} p(\mathbf{v}_{1:n}) = 1 \quad \text{and} \quad \sum_{n=1}^{\infty} \int_{\Omega_A^n} p(n)p(\mathbf{v}_{1:n}) = 1, \quad (1)$$

where  $p(n)$  is the probability of length  $n$  path. If we have any prior information about the expected path lengths this can be encoded in  $p(n)$ . Without prior knowledge,  $p(n)$

is chosen to be a uniform distribution over a range of reasonable lengths. Integrals over path spaces, like the ones in Eq. 1, also arise for example in quantum mechanics [7] and computer graphics [8]. Now, let  $\Omega_{AB}^n$  be the set of all paths of length  $n$  between  $A$  and another point  $B$ . We can then find the probability of a fiber going from  $A$  to  $B$ , given the diffusion data  $\mathcal{D}$ , by summing the probabilities for all paths of all lengths between these points:

$$p(A \rightarrow B|\mathcal{D}) = \sum_{n=1}^{\infty} \int_{\Omega_{AB}^n} p(n)p(\mathbf{v}_{1:n}|\mathcal{D}). \tag{2}$$

In the above equation we have suppressed the dependence of the path length  $n$  on  $\mathcal{D}$ . The integrations in Eq. 2 are performed over  $2n$ -dimensional path spaces  $\Omega_{AB}^n$ , and we cannot hope to find analytical solutions. Hence, we must resort to numerical integration, and it is only by applying Monte Carlo methods we can hope to estimate such high-dimensional integrals. Since  $\Omega_{AB}^n \subset \Omega_A^n$ , we can approximate the integrals by drawing a large number of random paths  $\mathbf{v}_{1:n}^k$ ,  $k = 1, \dots, N_n$ , from  $p(\mathbf{v}_{1:n}|\mathcal{D})$  over the domain  $\Omega_A^n$ , i.e., with starting point in  $A$ , and calculate the fraction of paths that reach  $B$ . This technique is known as Rejection sampling. Formally, by defining an indicator function

$$\mathbb{I}(\mathbf{v}_{1:n}^k) = \begin{cases} 1 & \mathbf{v}_{1:n}^k \in \Omega_{AB}^n \\ 0 & \text{Otherwise} \end{cases}, \text{ we have } p(A \rightarrow B|\mathcal{D}) \approx \sum_{n=1}^{\infty} \sum_{k=1}^{N_n} p(n) \frac{\mathbb{I}(\mathbf{v}_{1:n}^k)}{N_n}.$$

Instead of calculating  $p(A \rightarrow B|\mathcal{D})$ , Tuch [3] consider the path of maximal probability between  $A$  and  $B$ . As a measure of connectivity between  $A$  and  $B$ , the maximal probability path has the interesting property of being symmetric. In general, this does not hold for  $p(A \rightarrow B|\mathcal{D})$ ; the probability will depend on whether we start sampling paths in  $A$  or in  $B$ . The reason is that  $p(A \rightarrow B|\mathcal{D})$  pertains to the probability for a *single* fiber, which cannot split or merge. Furthermore, note that in order to calculate  $p(A \rightarrow B|\mathcal{D})$ , there is no need to explicitly evaluate the path probabilities  $p(\mathbf{v}_{1:n}|\mathcal{D})$ ; it is sufficient if we can draw sample paths. In contrast, to find the maximal probability path, the path probabilities must be computed. In the approach described in the following sections, both alternatives are possible.

### 2.1 Sampling Paths

To implement the above scheme we need a method for drawing random paths  $\mathbf{v}_{1:n} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  from a probability density function (pdf)  $p(\mathbf{v}_{1:n})$ . Since the sampled paths are models of actual fibers, a certain regularity must be imposed upon them. For simplicity and notational convenience we assume that the vector  $\mathbf{v}_k$  only depends on the previous vector  $\mathbf{v}_{k-1}$ , and not on  $\mathbf{v}_{k-2}, \dots, \mathbf{v}_1$ . This assumption is trivially relaxed. We also assume a pre-determined vector length, i.e., a step length, so that we can work with normalized vectors  $\hat{\mathbf{v}}_k$ . The probability for a path of length  $n$  then factorizes into

$$p(\mathbf{v}_{1:n}|\mathcal{D}) = p_1(\hat{\mathbf{v}}_1|\mathcal{D}) \prod_{k=2}^n p_k(\hat{\mathbf{v}}_k|\hat{\mathbf{v}}_{k-1}, \mathcal{D}). \tag{3}$$

Hence, random paths can be built sequentially by first drawing a random direction  $\hat{\mathbf{v}}_1$ , then a random direction  $\hat{\mathbf{v}}_2$  given  $\hat{\mathbf{v}}_1$ , and so on. This procedure is known as Sequential

Importance Sampling [1]. An interpolation problem arises because we only have the diffusion data  $\mathcal{D}$  on a discrete grid, whereas the sampled paths do not have such a restriction. Due to the probabilistic framework we can use the method suggested in [6], which uses the diffusion data at a grid point chosen at random based on the distance to the current sample point. The sequential sampling is terminated when the path reaches an area where the uncertainty regarding the next step is too high.

### 3 Local Modeling and Estimation

The goals in this section are to find a pdf of the local fiber orientation  $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}, \mathcal{D})$  and to draw random samples from this pdf. Examples of pdf's are shown in Fig. 1a. The local pdf's should take the uncertainty arising from partial volume effects and measurement noise into account. There are a few different ways to approach this problem. For example, with  $q$ -ball imaging [3], it is possible to resolve complex fiber neighborhoods and to obtain a detailed description of the local diffusion profile. The drawback with this technique is that it does not account for the uncertainty stemming from noise. Another approach is to acquire several diffusion data sets and generate a pdf by bootstrapping [9], with the obvious disadvantage that several data set are required. The most commonly used approach in diffusion weighted imaging is to use an explicit model for the water diffusion profile. As pioneered by Behrens et al. [6], the uncertainty in fiber direction can then be inferred via Bayes' theorem, as described below.

#### 3.1 Local Bayesian Modeling

Assume that we have a model that relates the diffusion measurements  $\mathcal{D}$  with the underlying tissue properties and architecture. Such a model necessarily contains at least one fiber direction  $\hat{\mathbf{v}}_k$  and a set of nuisance parameters collectively denoted by  $\theta$ . By applying Bayes' theorem we have

$$p(\hat{\mathbf{v}}_k, \theta | \hat{\mathbf{v}}_{k-1}, \mathcal{D}) = \frac{p(\mathcal{D} | \hat{\mathbf{v}}_k, \theta) p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) p(\theta)}{p(\mathcal{D})}, \quad (4)$$

where we have assumed that the prior distribution can be factorized  $p(\hat{\mathbf{v}}_k, \theta | \hat{\mathbf{v}}_{k-1}) = p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) p(\theta)$ . To find  $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}, \mathcal{D})$  we need to marginalize Eq. 4 over the nuisance parameters  $\theta$ . Furthermore, the normalizing factor

$$p(\mathcal{D}) = \int_{\hat{\mathbf{v}}_k, \theta} p(\mathcal{D} | \hat{\mathbf{v}}_k, \theta) p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) p(\theta) \quad (5)$$

is in general difficult to evaluate because of the high-dimensional integral and an intractable integrand. The posterior in Eq. 4 has to be calculated in every step in the sequential sampling of the fiber paths, and unless an approximation for the integral in Eq. 5 is found the cost of the Bayesian approach will be prohibitive. Behrens et al. [6] avoid evaluating the integral by applying Markov Chain Monte Carlo (MCMC) methods for drawing samples from the posterior. In the following sections we describe an alternative method, and discuss different observation models  $p(\mathcal{D} | \hat{\mathbf{v}}_k, \theta)$  and priors,  $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1})$  and  $p(\theta)$ , that allow fast computation of the posterior.

**Observation Model.** The true underlying voxel intensities  $\mu_i$  in the diffusion weighted images depend on the local water diffusion profile, and the fiber orientation is generally assumed to coincide with the direction of largest diffusivity. In Table 1, three models of how the voxel intensity is modulated by the water diffusion are listed. The gradient directions  $\hat{\mathbf{g}}_i$  and the  $b$ -values  $b_i$  are known experimental parameters. In these models,

**Table 1.** Three models of how the voxel intensity  $\mu_i$  depends on the water diffusion profile

Gaussian model	$\mu_i = \mu_0 e^{-b_i \hat{\mathbf{g}}_i^T \mathbf{D} \hat{\mathbf{g}}_i}$
Constrained model	$\mu_i = \mu_0 e^{-\alpha b_i} e^{-\beta b_i (\hat{\mathbf{g}}_i^T \hat{\mathbf{v}})^2}$
Compartment model	$\mu_i = \mu_0 [(1-f)e^{-b_i d} + f e^{-b_i d (\hat{\mathbf{g}}_i^T \hat{\mathbf{v}})^2}]$

the diffusion tensor  $\mathbf{D}$  or the local fiber direction  $\hat{\mathbf{v}}$  are parameters of interest. The remaining parameters are nuisance parameters. The Gaussian model is widely used in diffusion imaging [10], and the Compartment model was proposed by Behrens et al. [6] (although here it is given in a reparameterized version). We propose the Constrained model, which is obtained if we assume that the two smallest eigenvalues of  $\mathbf{D}$  are equal, i.e.,  $\lambda_2 = \lambda_3 = \alpha$ . This gives

$$\mathbf{D} = \lambda_1 \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \alpha (\hat{\mathbf{v}}_2 \hat{\mathbf{v}}_2^T + \hat{\mathbf{v}}_3 \hat{\mathbf{v}}_3^T) = (\lambda_1 - \alpha) \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \alpha \mathbf{I} = \beta \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \alpha \mathbf{I}, \quad (6)$$

which after substitution into the expression for the Gaussian model yields the Constrained model. The Constrained model describes cigar-shaped diffusion profiles, with a sphere and a line as extremes. While all three models in Table 1 describe diffusion profiles, the Constrained model and the Compartment model can also be viewed as modeling the effect of an underlying fiber. They can also be extended to handle multiple fiber directions. In this work, however, we assume that there is only one fiber direction in each voxel, and any deviations from this model will be captured as uncertainty in this direction via the posterior in Eq. 4. The reason we prefer the Constrained model over the Compartment model is out of mathematical tractability, as will become clear later.

The voxel intensity  $y_i$  in a diffusion weighted image is a noisy observation of  $\mu_i$ . Moreover, the intensity decays exponentially with the water diffusivity, as implied by the models in Table 1. Hence, by taking the logarithm of the observations,  $z_i = \ln y_i$ , we obtain a more linear relationship between observations and model parameters. In references [10, 11] it is shown that  $z_i = \ln \mu_i + \epsilon$ , where  $\epsilon \in N(0, \sigma^2 / \mu_i^2)$ , is a good model for the observation noise. That is, after taking the logarithm the noise can be modeled as additive Gaussian with a variance that depends on the voxel intensity. The joint distribution for the observed (log-)data  $\mathcal{D} = [z_1, \dots, z_N]$  is then

$$p(\mathcal{D} | \hat{\mathbf{v}}, \boldsymbol{\theta}) = \prod_{i=1}^N \frac{\mu_i}{\sqrt{2\pi\sigma^2}} e^{-\frac{\mu_i^2}{2\sigma^2} (z_i - \ln \mu_i)^2}, \quad (7)$$

where  $\mu_i$  is to be replaced by the expression for any of the models in Table 1, and  $\boldsymbol{\theta}$  denotes the nuisance parameters in this model. This expression is used in the equations for the posterior, Eq. 4 and Eq. 5.

**Priors.** Via the probability functions  $p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1})$  and  $p(\boldsymbol{\theta})$  we encode our prior knowledge about fiber regularity and nuisance parameters. Starting with the fiber regularity, a simple family of distributions is given by

$$p(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{k-1}) \propto \begin{cases} (\hat{\mathbf{v}}_k^T \hat{\mathbf{v}}_{k-1})^\gamma & \hat{\mathbf{v}}_k^T \hat{\mathbf{v}}_{k-1} \geq 0 \\ 0 & \hat{\mathbf{v}}_k^T \hat{\mathbf{v}}_{k-1} < 0 \end{cases} \quad \text{with } \gamma \geq 0. \quad (8)$$

This prior gives preference to continue in the previous step-direction, with a decreasing probability for sharper turns until it reaches a zero probability for turns of 90 degrees and above. The  $\gamma$ -parameter controls the sharpness of the distribution, i.e., the regularity of the path. In our experiments we use  $\gamma = 1$ . Turning to the nuisance parameters in the models in Table 1, e.g. the Constrained model for which  $\boldsymbol{\theta} = \{\mu_0, \alpha, \beta, \sigma^2\}$ , we generally do not have any detailed prior information, except that these parameters should be non-negative. Considering the computational effort required to evaluate the integral in Eq. 5 we choose dirac impulses as priors for the nuisance parameters, for example  $p(\boldsymbol{\theta}) = \delta(\mu_0 - \hat{\mu}_0) \delta(\alpha - \hat{\alpha}) \delta(\beta - \hat{\beta}) \delta(\sigma^2 - \hat{\sigma}^2)$ . That is, we fix the nuisance parameters to some values  $\hat{\mu}_0, \hat{\alpha}, \hat{\beta}$  and  $\hat{\sigma}^2$ . The integral in Eq. 5, as well as the posterior pdf in Eq. 4, is then evaluated over the unit sphere. Hence, they are easily calculated numerically and significant computational savings are made. Also, note that the Constrained model and the Compartment model in Table 1 differ only in the way the nuisance parameters enter, and once these parameters are fixed the models are essentially equivalent.

### 3.2 Point Estimates of Local Model Parameters

To define the priors above, or to initialize an MCMC process for drawing samples from the posterior, it is important to have access to good point estimators of the parameters in the observation models in Table 1. For the Gaussian model, point estimates are readily obtained by means of linear least squares estimation [10, 11]. The parameters in the Constrained model can be found through the following non-trivial theorem:

**Theorem.** *Let  $\mathbf{D}$  be a symmetric  $3 \times 3$  matrix with eigenvalue factorization  $\mathbf{D} = \lambda_1 \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \lambda_2 \hat{\mathbf{v}}_2 \hat{\mathbf{v}}_2^T + \lambda_3 \hat{\mathbf{v}}_3 \hat{\mathbf{v}}_3^T$ ,  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ . The closest, in terms of the Frobenius norm, symmetric matrix  $\mathbf{C}$  with the two smallest eigenvalues equal is given by*

$$\mathbf{C} = \lambda_1 \hat{\mathbf{v}}_1 \hat{\mathbf{v}}_1^T + \frac{\lambda_2 + \lambda_3}{2} (\hat{\mathbf{v}}_2 \hat{\mathbf{v}}_2^T + \hat{\mathbf{v}}_3 \hat{\mathbf{v}}_3^T). \quad (9)$$

A proof of this theorem is left out due to space restrictions. Hence, we can find the parameters of the Constrained model by first solving for the full diffusion tensor  $\mathbf{D}$  in the Gaussian model and then setting  $\alpha = (\lambda_2 + \lambda_3)/2$ ,  $\beta = \lambda_1 - \alpha$  and  $\hat{\mathbf{v}} = \hat{\mathbf{e}}_1$ . The Compartment model is more difficult to handle mathematically and we need to apply computationally demanding non-linear optimization techniques to find accurate estimates of its parameters.

### 3.3 Drawing Samples from the Posterior

Finally, to perform the sequential sampling of fiber paths, we need to draw random samples of the fiber direction from the posterior in Eq. 4. For drawing samples from

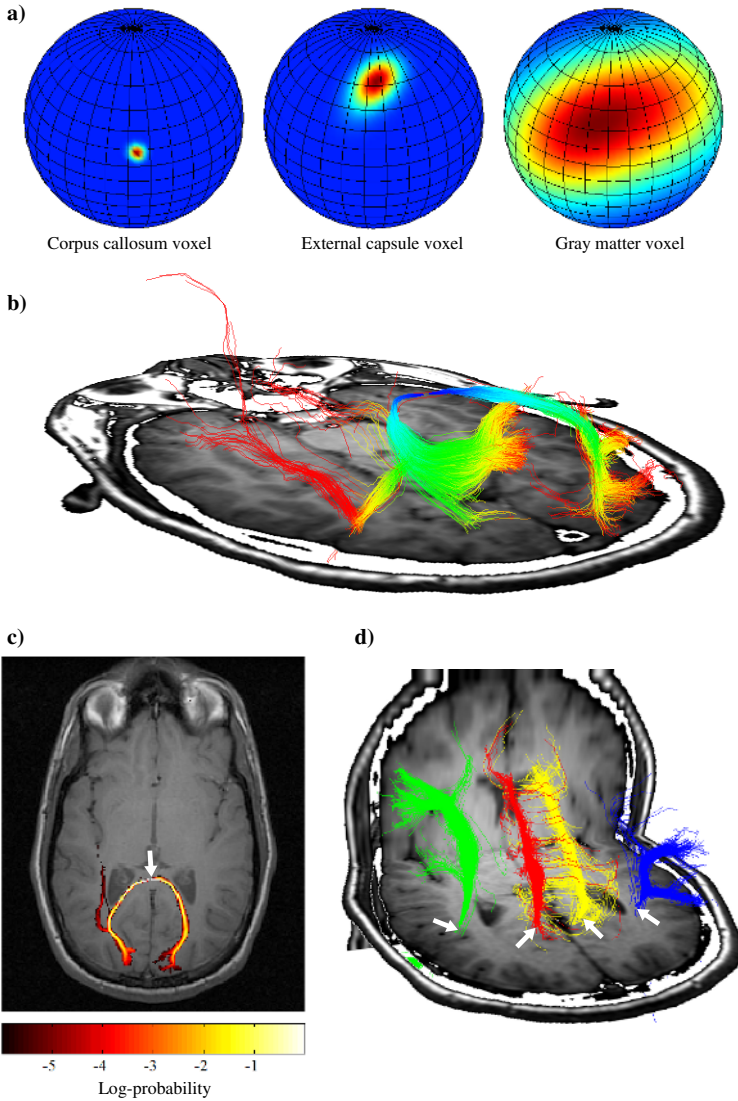
complicated and high dimensional pdf's one can always resort to MCMC techniques. However, while MCMC sampling is conceptually simple there are implementational issues, such as convergence and so-called mixing of the chain, that complicate practical use. In contrast, with the dirac priors suggested above we need only draw samples from a pdf defined on the unit sphere. This can be accomplished by evaluating the pdf at a sufficiently large number of points evenly spread over the unit sphere, effectively approximating the continuous pdf with a discrete pdf, from which it is straightforward to draw random samples. We must make take care to sample the continuous pdf densely enough so that high probability areas are covered well. In our experiments we evaluate the posterior at 2,562 predefined unit length vectors obtained by a fourfold tessellation of an icosahedron.

## 4 Results

Here we show results obtained by applying the methodology described in the above sections to a diffusion weighted data set with 31 different gradient directions and a  $b$ -value of 1000 s/mm<sup>2</sup>. Figure 1a shows example pdf's of the fiber orientation in three different voxels in this data set. To produce sample fiber paths we sequentially draw fiber orientations from such pdf's. A Matlab implementation typically produces a few sample paths per second on a high-end PC, and the sampling process is trivially parallelized. Figure 1b shows 3,000 sample fibers originating in one point in the splenium of Corpus callosum (marked in Fig. 1c). The fibers have been colored according to how the probability evolves along the paths (Eq. 3). From the sampled fibers in Fig. 1b, the probability map in Fig. 1c was produced as described in Section 2. This map indicates the probability, or our belief, of a single fiber going from the starting point to any other point in the image slice. Finally, Fig. 1d shows more examples of sampled fiber paths, started in four different points in the axial slice. Note how the stochastic framework handles splitting fiber bundles and partial volume effects. For example, a single sample fiber in the Cingulum bundles (red and yellow fibers) easily slips into the underlying Corpus callosum bundle, but the massive sampling approach gives a robust delineation.

## 5 Discussion

To conclude this paper, we have presented a method for assessing uncertainty in white matter fiber tractography. The main contribution is a novel way of obtaining pdf's of the local fiber direction, including a new model of the water diffusion suitable for a stochastic tracking framework, and a theorem that allows fast computation. Also, while the procedure for calculating the probability of a connection has been used before [2, 4, 5, 6], the theoretical justification in Section 2 has not been described previously. Among the cited works, the approach presented by Behrens et al. is closest to the one presented here. The essential difference lies in how the posterior pdf of the fiber orientation is approximated in order to obtain a computationally feasible solution. We use simplified priors for the nuisance parameters, which result in an algorithm that is computationally fast and easy to implement. Since some of the priors are estimated from the data, the approach can be classified as empirical Bayesian. In contrast, the approach presented



**Fig. 1.** **a)** Example pdf's of the local fiber orientation taken at three different locations. A uniform prior for the step direction was used when generating these pdf's, i.e, the previous step direction has no influence. **b)** 3,000 fiber samples originating in a point located at the splenium of Corpus callosum (marked in c)). Each fiber is colored according to how the probability evolves along its path. **c)** Probability map of the existence of a fiber going from the marked point to all other voxels in the image slice. The map was generated using the fiber samples in b). **d)** More examples of fiber samples. The sampling was started in the four different points marked by the arrows, and the fibers are colored accordingly.

by Behrens et al. employs full Bayesian modeling of the priors combined with MCMC sampling, which is slower but has the theoretical advantage of taking uncertainty in nuisance parameter estimates into account. However, to exploit this advantage we need to assess convergence and adaptively control the mixing of a large number of Markov chains automatically, which we have found difficult. Further work is required to investigate if there are any interesting differences in result between the faster empirical Bayesian and the full Bayesian approaches.

## Acknowledgments

This work was funded by NIH grant P41-RR13218. The authors thank Dr. Raul San-Jose Estepar, Dr. Gunnar Farnebäck and Dr. Gordon Kindlmann for valuable input.

## References

1. Björnemo, M., Brun, A., Kikinis, R., Westin, C.F.: Regularized stochastic white matter tractography using diffusion tensor MRI. In: Proceedings of the 5th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'02), Tokyo, Japan (2002)
2. Koch, M., Norris, D., Hund-Georgiadis, M.: An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage* **16** (2002) 241–250
3. Tuch, D.: Diffusion MRI of complex tissue structure. PhD thesis, Harvard-MIT Division of Health Sciences and Technology (2002)
4. Parker, G., Haroon, H., Wheeler-Kingshott, C.: A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements. *Journal of Magnetic Resonance Imaging* **18** (2003) 242–254
5. Hagmann, P., Thiran, J., Jonasson, L., Vandergheynst, P., Clarke, S., Maeder, P., Meuli, R.: DTI mapping of human brain connectivity: Statistical fibre tracking and virtual dissection. *NeuroImage* **19** (2003) 545–554
6. Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, J., Smith, S.: Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine* **50** (2003) 1077–1088
7. Kleinert, H.: Path integrals in quantum mechanics, statistics, and polymer physics. World Scientific, Singapore (1995)
8. Veach, E.: Robust Monte Carlo methods for light transport simulation. PhD thesis, Stanford University (1997)
9. Lazar, M., Alexander, A.: Bootstrap white matter tractography (BOOT-TRAC). *NeuroImage* **24** (2005) 524–532
10. Basser, P., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B* **103** (1994) 247–254
11. Salvador, R., Pena, A., Menon, D., Carpenter, T., Pickard, J., Bullmore, E.: Formal characterization and extension of the linearized diffusion tensor model. *Human Brain Mapping* **24** (2005) 144–155