A Bayesian Approach for Stochastic White Matter Tractography

Ola Friman*, Gunnar Farnebäck, and Carl-Fredrik Westin

***Abstract*—White matter fiber bundles in the human brain can be located by tracing the local water diffusion in diffusion weighted magnetic resonance imaging (MRI) images. In this paper, a novel Bayesian modeling approach for white matter tractography is presented. The uncertainty associated with estimated white matter fiber paths is investigated, and a method for calculating the probability of a connection between two areas in the brain is introduced. The main merits of the presented methodology are its simple implementation and its ability to handle noise in a theoretically justified way. Theory for estimating global connectivity is also presented, as well as a theorem that facilitates the estimation of the parameters in a constrained tensor model of the local water diffusion profile.**

***Index Terms*—Bayesian modeling, diffusion tensor-magnetic resonance imaging (DT-MRI), fiber tracking, magnetic resonance imaging (MRI), probabilistic tracking, uncertainty.**

I. Introduction

In WHITE matter tractography, white matter fiber paths are estimated by tracing the direction of maximal water diffusion in diffusion weighted magnetic resonance imaging (MRI) images. Such estimated fiber paths can subsequently be used, for example, to investigate connectivity in mental and neurological disorders. To infer changes in connectivity in such disorders, the inherent variability of the estimated paths should be taken into account. Another application of white matter tractography is surgical planning, where it is important to consider the uncertainty associated with the estimated fiber bundles. In general, when estimating a quantity based on some available data, it is natural to ask what uncertainty is associated with the ensuing estimate. Indeed, uncertainty in white matter tractography, which arises due to noise, motion, imaging artifacts and partial volume effects (i.e., crossing, merging and splitting fiber tracts), has recently been investigated by several authors [1]–[10]. The question at issue is how image noise and partial volume effects are translated into uncertainty in the local fiber orientation, and how this uncertainty induces variability in global connectivity measures. Uncertainty is naturally quantified by means of probability, and in this paper, we are interested in the probability of a fiber connecting two areas in the brain, given the measured diffusion data. There are two main statistical approaches that can deal with this problem. A nonparametric approach does not make any assumptions regarding the noise or the underlying water diffusion profile. For example, Lazar *et al.* [8], and Jones *et al.* [9], recently investigated the bootstrap method for dealing with uncertainty stemming from noise. Another example is the nonparametric descriptions of the water diffusion profile that are obtained using q ball imaging [3], the persistent angular structure MRI [11], and related high angular resolution methods [12]. The drawback with nonparametric methods is that they require large amounts of data, and thereby also scanning times which under many circumstances are unacceptable. The alternative is to adopt a parametric approach, in which the estimation is regularized by prior knowledge in form of models. The widely used tensor model for the local water diffusion profile is an example [13]. However, when a parametric method is applied, care must be taken to introduce our prior knowledge and models at points where it can be theoretically or experimentally justified. To this end, Behrens *et al.* [7] pioneered a method where the probability density function of the local fiber orientation is derived in a Bayesian framework. Here we present an alternative Bayesian approach, which is efficient and straightforward to implement. The outline is as follows. First, we describe the global modeling of fibers and how to estimate the probability of a fiber going from a point or area A to an area B in the brain. Then a Bayesian approach for deriving probability density functions of the local fiber orientation is described, including a novel model for the water diffusion profile and an associated theorem that facilitates the estimation of the parameters in this model. Finally, results that justify our method are presented. Related theory and preliminary results have been presented by the authors before in [14].

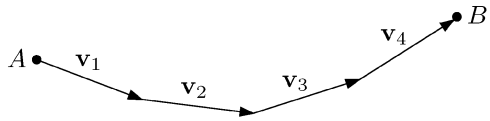

Fig. 1. A fiber path is represented by a train of vectors.

II. Global Modeling and Estimation

A white matter fiber can be modeled as a finite length path described by a train of vectors, see Fig. 1. We assume that all vectors have the same predetermined length, referred to as the step length. A fiber path can then be parameterized by a train of unit length vectors, and we use the following notation for such a path: $\mathbf{v}_{1:n} = \{\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_n\}$. Denote by Ω_A^n the set of all possible paths of length n that originate in a point or area A, and

Manuscript received January 18, 2006; revised April 4, 2006. This work was supported in part by the National Institutes of Health (NIH) under Grant P41-RR13218 and Grant 2R01-MH050740, and is a part of the National Alliance for Medical Image Computing (NAMIC), funded by the NIH through the NIH Roadmap for Medical Research, Grant U54 EB005149. *Asterisk indicates corresponding author*.

*O. Friman is with the Laboratory of Mathematics in Imaging, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115 USA (e-mail: friman@bwh.harvard.edu).

G. Farnebäck and C.-F. Westin are with the Laboratory of Mathematics in Imaging, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115 USA.

Digital Object Identifier 10.1109/TMI.2006.877093

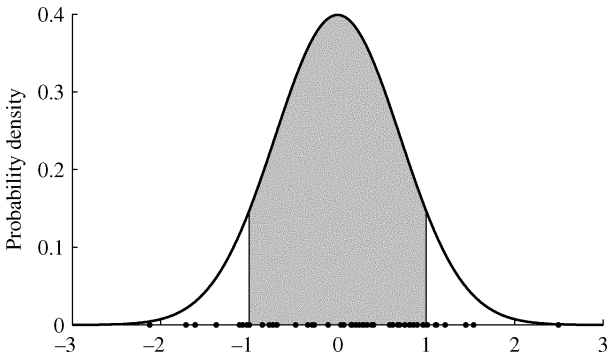

Fig. 2. To integrate the standard normal distribution between -1 and 1 using Rejection sampling, first a large number of samples are drawn from this distribution. An estimate of the area is then given by the fraction of samples that fall between -1 and 1. Using only the 50 random samples plotted along the x axis yields in an estimated area of 0.70, the correct area is 0.683.

assume that we can assign a probability $p(\mathbf{v}_{1:n})$ to each path in this space (or more correctly, define a probability function on the path space). We also introduce a discrete probability function $p(n)$ for the path length. Hence, we have

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

If we have any prior information about the expected path lengths, this can be encoded in $p(n)$. Without prior knowledge, a noninformative uniform distribution over a range of reasonable lengths is used. Integrals over path spaces, like the ones in (1), also arise in for example quantum mechanics [15], and computer graphics [16]. Now, let Ω_{AB}^n be the set of all possible paths of length n between A and another area B. We can find the probability $p(A \rightarrow B|\boldsymbol{\mathcal{D}})$ of a fiber going from A to B, given the diffusion data $\boldsymbol{\mathcal{D}}$, by summing the probabilities for all paths of all lengths between these areas

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

In the above equation, we have suppressed the dependence of the path length n on $\boldsymbol{\mathcal{D}}$. The integrals in (2) are defined over complicated path spaces Ω_{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 it is possible to estimate such high-dimensional integrals. A technique known as Rejection sampling is well suited for this problem. To illustrate this technique, consider the problem of integrating a standard normal distribution between -1 and 1, see Fig. 2. An estimate of the shaded area can be obtained by drawing random samples from the normal distribution and calculating the fraction of samples between -1 and 1. Similarly, since $\Omega_{AB}^n \subset \Omega_A^n$, we can approximate the integrals in (2) by drawing a large number of random sample paths $\mathbf{v}_{1:n}^k$, $k = 1, \ldots N_n$, from $p(\mathbf{v}_{1:n}|\boldsymbol{\mathcal{D}})$ over the domain Ω_A^n, i.e., with starting point in A, and calculate the fraction of paths that reach B. Formally, by defining an indicator function

$$\mathbb{I}(\mathbf{v}_{1:n}^k) = \begin{cases} 1, & \mathbf{v}_{1:n}^k \in \Omega_{AB}^n \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

we have

$$p(A \rightarrow B|\boldsymbol{\mathcal{D}}) \approx \sum_{n=1}^{\infty} \sum_{k=1}^{N_n} p(n) \frac{\mathbb{I}\left(\mathbf{v}_{1:n}^k\right)}{N_n}. \tag{4}$$

Expressed in words, a path that reaches B after n steps gives a contribution to the sum in (4) that depends on the total number of paths N_n of length n, and the probability of a path of length n, $p(n)$. In practice, the outer sum in (4) is finite because the probability function $p(n)$ of the path lengths is assumed zero above some finite length.

Instead of the probability $p(A \rightarrow B|\boldsymbol{\mathcal{D}})$, Tuch [13] considers 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 $p(A \rightarrow B|\boldsymbol{\mathcal{D}})$; the probability will depend on whether we start sampling paths in A or in B. To illustrate why, imagine that a fiber bundle starts in a point A and subsequently splits into two equal parts, which reach two areas B and C, respectively. $p(A \rightarrow B|\boldsymbol{\mathcal{D}})$ will then be estimated to around 0.5. But if we start in B, and assume that no other tracts reach B, almost all traced paths will reach A, and $p(B \rightarrow A|\boldsymbol{\mathcal{D}})$ will be close to 1. Furthermore, note that in order to calculate $p(A \rightarrow B|\boldsymbol{\mathcal{D}})$, there is no need to explicitly evaluate the path probabilities $p(\mathbf{v}_{1:n}|\boldsymbol{\mathcal{D}})$; it is sufficient if we can draw sample paths. In contrast, to find the maximal probability path, the path probabilities must be computed. Even though we will not examine this further, it should be mentioned that an estimate of the maximal probability path is obtained as a byproduct in the methodology presented later.

A. Sampling Paths

To implement the above scheme, we need a method for drawing random paths $\mathbf{v}_{1:n} = \{\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_n\}$ from a high-dimensional probability density function (pdf) $p(\mathbf{v}_{1:n})$. Remember that the step length is assumed fixed, and that we, therefore, can work with unit length vectors. 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 $\hat{\mathbf{v}}_i$ only depends only on the previous vector $\hat{\mathbf{v}}_{i-1}$, and not on $\hat{\mathbf{v}}_{i-2}, \ldots, \hat{\mathbf{v}}_1$. This assumption is trivially relaxed. The probability for a path of a given length n then factorizes into

$$p(\mathbf{v}_{1:n}|\boldsymbol{\mathcal{D}}) = p_1(\hat{\mathbf{v}}_1|\boldsymbol{\mathcal{D}}) \prod_{i=2}^{n} p_i(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1}, \boldsymbol{\mathcal{D}}). \tag{5}$$

Note that the probability function of the step direction is different for every vector in the path, hence, the indexes of the probability functions. Henceforth, this index will be dropped where it is not important. Equation (5) tells us that random paths

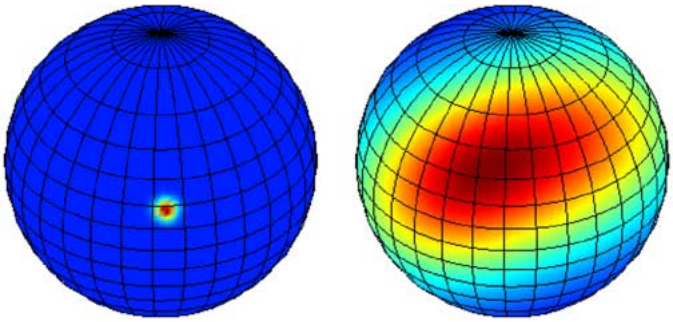

Fig. 3. The posterior distributions of the fiber orientation for a voxel in the Corpus callosum (left) and for a gray matter voxel (right). Posterior distribution of the Corpus callosum voxel is very focused, whereas the fiber orientation in the gray matter voxel is very uncertain.

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], [17], and it is similar to what in the white matter tractography literature is known as streamlining. The sequential sampling is terminated when the path reaches a forbidden area, e.g., outside the brain, or if the uncertainty regarding the next step is too high.

An interpolation problem arises in the sampling process because we only have the diffusion data $\boldsymbol{\mathcal{D}}$ on a discrete grid, whereas the sampled paths do not have such a restriction. The probabilistic framework allows us to employ the probabilistic interpolation method suggested by Behrens *et al.* [7], which utilizes the diffusion data at a grid point chosen at random based on the distance to the current sample point.

III. Local Modeling and Estimation

To carry out the sequential sampling of fiber paths, in every step we need to find a pdf of the local fiber orientation $p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1},\boldsymbol{\mathcal{D}})$, and to draw a random sample from this pdf. Examples of such pdf's are shown in Fig. 3. Assume that we have a model that relates the diffusion measurements $\boldsymbol{\mathcal{D}}$ with the underlying tissue properties and fiber architecture. Such a model necessarily contains at least one fiber direction $\hat{\mathbf{v}}_i$, which here is the parameter of interest, and a set of nuisance parameters collectively denoted by $\boldsymbol{\theta}$. By applying Bayes' theorem, we have

$$p(\hat{\mathbf{v}}_i,\boldsymbol{\theta}|\hat{\mathbf{v}}_{i-1},\boldsymbol{\mathcal{D}}) = \frac{p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_{i-1},\hat{\mathbf{v}}_i,\boldsymbol{\theta})p(\hat{\mathbf{v}}_i,\boldsymbol{\theta}|\hat{\mathbf{v}}_{i-1})}{p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_{i-1})}. \quad (6)$$

This expression can be simplified if we introduce a few plausible assumptions. First, we assume that the prior distribution can be factorized $p(\hat{\mathbf{v}}_i,\boldsymbol{\theta}|\hat{\mathbf{v}}_{i-1}) = p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1})p(\boldsymbol{\theta})$, meaning that our prior knowledge about the nuisance parameters in the current point is independent of both the previous step direction and our prior knowledge about the next step direction. Likewise, we assume that the diffusion measurements in the current point do not depend on the previous step direction, giving that $p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_i,\hat{\mathbf{v}}_{i-1},\boldsymbol{\theta}) = p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_i,\boldsymbol{\theta})$ and $p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_{i-1}) = p(\boldsymbol{\mathcal{D}})$. The equation can then be written

$$p(\hat{\mathbf{v}}_i,\boldsymbol{\theta}|\hat{\mathbf{v}}_{i-1},\boldsymbol{\mathcal{D}}) = \frac{p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_i,\boldsymbol{\theta})p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1})p(\boldsymbol{\theta})}{p(\boldsymbol{\mathcal{D}})}. \quad (7)$$

TABLE I
Three Models of How the Voxel Intensity μ_j Depends on the Underlying Water Diffusion Profile

Tensor model	$\mu_j = \mu_0\, e^{-b_j \hat{\mathbf{g}}_j^T \mathbf{D} \hat{\mathbf{g}}_j}$
Constrained model	$\mu_j = \mu_0\, e^{-\alpha b_j}\, e^{-\beta b_j \left(\hat{\mathbf{g}}_j^T \hat{\mathbf{v}}\right)^2}$
Compartment model	$\mu_j = \mu_0 [(1-f)e^{-b_j d} + f e^{-b_j d \left(\hat{\mathbf{g}}_j^T \hat{\mathbf{v}}\right)^2}]$

The factor $p(\boldsymbol{\mathcal{D}})$ normalizes the posterior probability distribution to have unit volume and can thus, be written as the integral of the numerator

$$p(\boldsymbol{\mathcal{D}}) = \int_{\hat{\mathbf{v}}_i,\boldsymbol{\theta}} p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_i,\boldsymbol{\theta})p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1})p(\boldsymbol{\theta}). \quad (8)$$

Finally, to find the probability function of interest $p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1},\boldsymbol{\mathcal{D}})$ used in (5), we need to marginalize, i.e., integrate, (7) over the nuisance parameters $\boldsymbol{\theta}$. This marginalization and the integral in (8) have to be calculated in every step in the sequential sampling of the fiber paths. Since these integrals are high-dimensional and generally intractable, the cost of the Bayesian approach will be prohibitive unless approximations for these integrals are found. As will be discussed in further detail later, Behrens *et al.* [7] avoid evaluating the integrals by applying Markov Chain Monte Carlo (MCMC) methods for drawing samples from a joint posterior distribution similar to the one in (7). In the following sections, we describe an alternative method, and discuss different observation models $p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_i,\boldsymbol{\theta})$ and priors, $p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1})$ and $p(\boldsymbol{\theta})$, that allow fast and simple computation of the posterior distribution.

A. Observation Model

The true intensities μ_j, $j = 1,\ldots,N$, in a voxel in the diffusion sensitized images depend on the local water diffusion profile. In Table I, three models of how the voxel intensity is modulated by the underlying water diffusivity and fiber orientation are listed. The gradient directions $\hat{\mathbf{g}}_j$ and the b-values b_j are known experimental parameters. In white matter tractography, the diffusion tensor $\mathbf{D}$ or the orientation of maximal water diffusivity $\hat{\mathbf{v}}$ are the parameters of interest. The orientation of maximal water diffusivity is generally assumed to coincide with the orientation of an underlying fiber tract. The remaining parameters in the models in Table I are of little interest in the current application, but they are required for modeling the connection between water diffusivity and voxel intensity. They are for this reason referred to as nuisance parameters.

The tensor model is widely used in diffusion imaging [13]. It assumes a Gaussian shaped water diffusion profile described by the diffusion tensor $\mathbf{D}$ with eigen-decomposition

$$\mathbf{D} = \lambda_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \lambda_2 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T + \lambda_3 \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3^T \quad (9)$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$. The compartment model was proposed by Behrens *et al.* [7], though in Table I it is reproduced in a different parametrization. This model assumes that a voxel can be divided into two compartments: one compartment with isotropic water diffusion and one compartment where the water diffuses anisotropically due to a potential fiber bundle going through the

voxel. In this work we use the constrained model, which is obtained by assuming that the two smallest eigenvalues of the diffusion tensor $\mathbf{D}$ are equal, i.e., $\lambda_2 = \lambda_3 = \alpha$. This gives

$$\begin{aligned}\mathbf{D} &= \lambda_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \alpha\left(\hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T + \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3^T\right) \\ &= (\lambda_1 - \alpha)\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \alpha \mathbf{I} = \beta \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \alpha \mathbf{I} \end{aligned} \tag{10}$$

which after substitution into the expression for the tensor model yields the constrained model. The constrained model describes cigar-shaped diffusion profiles, with a sphere and a line as extremes. This model has recently been suggested independently by Alexander [18], and by Friman and Westin [14]. Using the theorem presented in the Appendix, here an efficient estimation procedure for this model is provided.

While all three models in Table I describe diffusion profiles, the constrained model and the compartment model can also be viewed as modeling the effect of a single underlying fiber tract. Due to the relatively large voxel size, voxels may contain several different fiber tracts, and models that handle multiple fiber orientations within a voxel have been proposed [11], [12], [19]. In this paper, however, we assume that there is only one fiber orientation in each voxel, and any deviations from this model will be captured as uncertainty in this orientation via the posterior distribution in (7). This assumption will be discussed further in the results and discussion sections. The reason we prefer the constrained model over the compartment model is because of its mathematical tractability, as will become clear later.

The voxel intensity y_j in a diffusion weighted image is a noisy observation of μ_j. Moreover, the intensity decays exponentially (or bi-exponentially) with the water diffusivity, as implied by the models in Table I. Hence, by taking the logarithm of the observations, $z_j = \ln y_j$, we obtain a more linear relationship between observations and model parameters. The noise in magnitude MRI images is Rician distributed [20], [21], which should be taken into account when estimating parameters using the logarithmed data. Basser *et al.* [13] and Salvador *et al.* [22] show that $z_j = \ln \mu_j + \epsilon$, where $\epsilon \in N(0, \sigma^2/\mu_j^2)$, is an accurate 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, or the likelihood, of the observed (log-)data $\boldsymbol{\mathcal{D}} = [z_1, \ldots, z_N]$ in a voxel can then be written

$$p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{\mu_j}{\sqrt{2\pi\sigma^2}} e^{-\frac{\mu_j^2}{2\sigma^2}(z_j - \ln \mu_j)^2} \tag{11}$$

where μ_j is to be replaced by the expression for any of the models in Table I, and $\boldsymbol{\theta}$ denotes the nuisance parameters in this model. This likelihood function is substituted into the expressions for the posterior distribution in (7) and (8).

B. Priors

Via the probability functions $p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1})$ and $p(\boldsymbol{\theta})$, we encode our prior knowledge about fiber regularity and nuisance parameters. The prior distribution of the fiber regularity can be inferred by invasive examination of real white matter fibers. However, here we use a simple family of distributions given by

$$p(\hat{\mathbf{v}}_i|\hat{\mathbf{v}}_{i-1}) \propto \begin{cases} \left(\hat{\mathbf{v}}_i^T \hat{\mathbf{v}}_{i-1}\right)^\gamma, & \hat{\mathbf{v}}_i^T \hat{\mathbf{v}}_{i-1} \geq 0 \\ 0, & \hat{\mathbf{v}}_i^T \hat{\mathbf{v}}_{i-1} < 0 \end{cases} \tag{12}$$

where $\gamma \geq 0$. 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° and above. The γ-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 I, for example the constrained model in which $\boldsymbol{\theta} = \{\mu_0, \alpha, \beta, \sigma^2\}$, we generally do not have any detailed prior information, except that these parameters should be nonnegative. Considering the computational effort required to evaluate the integral in (8), 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, in the calculation of the posterior pdf in (7), the nuisance parameters are fixed to some values $\hat{\mu}_0$, $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\sigma}^2$. Note, however, that the nuisance parameters will vary from voxel to voxel. With this simplification, the integral in (8) and posterior pdf in (7) collapses to be defined only over $\hat{\mathbf{v}}_i$, i.e., over the unit sphere. Hence, they are easily calculated numerically and significant computational savings are made.

C. Point Estimates of Model Parameters

To define the priors above, or to initialize an MCMC process for drawing samples from the posterior distribution, it is important to have access to good point estimators for the parameters in the observation models in Table I. For the tensor model, point estimates are readily obtained by means of linear least squares estimation [13], [22]. The parameters in the constrained model can be found through the following theorem.

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

$$\mathbf{S} = \lambda_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \frac{\lambda_2 + \lambda_3}{2}\left(\hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T + \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3^T\right). \tag{13}$$

A proof of this theorem can be found in the Appendix. Recalling (10), we can find the parameters of the constrained model by first solving for the full diffusion tensor $\mathbf{D}$ in the tensor model, and then setting

$$\alpha = \frac{\lambda_2 + \lambda_3}{2} \quad \beta = \lambda_1 - \alpha \quad \hat{\mathbf{v}} = \hat{\mathbf{e}}_1. \tag{14}$$

The compartment model is more difficult to handle mathematically, and we need to apply computationally demanding nonlinear optimization techniques to find accurate estimates of its parameters. Finally, once the model parameters have been estimated, an estimate of the noise variance $\hat{\sigma}^2$ is obtained from the residuals, see [22].

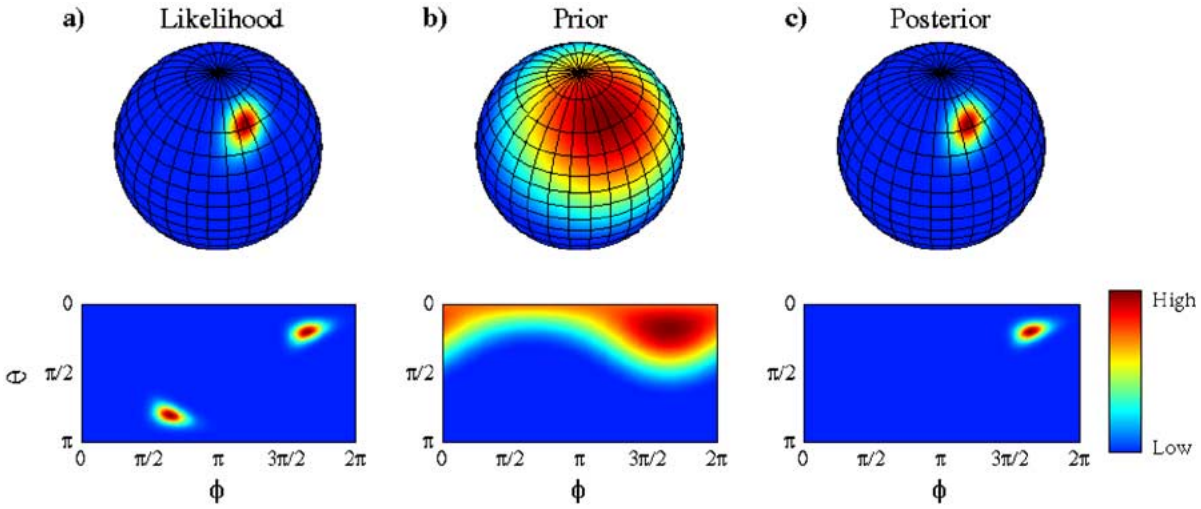

Fig. 4. Example of how the posterior pdf of the next step direction is computed in an example voxel in the External capsule. In the top panel, distributions for the next step direction are plotted on the unit sphere. The bottom panel shows same distributions using spherical coordinates, where θ is the angle from the positive z axis and ϕ is the azimuth, i.e., the angle from the x axis in the xy plane. (a) The likelihood given the diffusion data. Note that the likelihood function is symmetric because the diffusion data points out an orientation and not a direction. (b) Hypothetical prior distribution of the direction of the next step. This prior was generated from an assumed previous step direction. (c) Posterior pdf for the next step direction, obtained as the product of the likelihood and prior, and then normalized so as to become a probability density.

D. Drawing Samples From the Posterior Distribution

Finally, to perform the sequential sampling of fiber paths, we need to draw random samples of the fiber direction from the posterior distribution in (7). For drawing samples from complicated and high dimensional pdfs, one can always resort to MCMC techniques. However, while MCMC sampling is conceptually simple, there are implementation issues, such as convergence and so-called mixing of the Markov chain, that complicate practical use. In contrast, with the dirac priors proposed 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 denote such a set of unit vectors with $\mathcal{S}$. Care must be taken to sample the continuous pdf densely enough, so as to cover high-probability areas well. In our experiments, we evaluate the posterior distribution at 2562 predefined unit length vectors obtained by a fourfold tessellation of an icosahedron. This number was found sufficient by examining voxels with the most anisotropic diffusion in the Corpus callosum. In case a larger number of directions is desired or required, such directions can for example be found by an additional tessellation, or by an electrostatic repulsion algorithm used previously in diffusion tensor-magnetic resonance imaging (DT-MRI) to find suitable gradient directions [23].

IV. Implementation

In this section, the implementation of the methodology presented in the previous sections is discussed. The procedure for calculating the probability of a connection between two areas in the brain consists of two separate steps. First, a large number of fiber paths must be generated. Subsequently, a map of connection probabilities can be calculated. These two steps are detailed in the sections below. MATLAB code is available on request.

A. Sampling Fiber Paths

A central issue in probabilistic tractography is how the probability distribution of the underlying fiber orientation is calculated, and how random samples are drawn from this distribution. In Section III, a Bayesian derivation leading to a posterior distribution of the next step direction is presented. An example of the posterior distribution in a single voxel is visualized in Fig. 4, and two more examples of posterior distributions are shown in Fig. 3. The distribution of the next step direction is found by marginalizing, i.e., integrating, (7) over the nuisance parameters $\boldsymbol{\theta}$. The integrations required for finding the posterior distribution are demanding operations, and in the sections above two simplifications are suggested to facilitate the calculation. First, it is suggested to use dirac priors for the nuisance parameters. This has the effect of collapsing the integrals over the nuisance parameters. Second, using the set of unit vectors densely distributed over the unit sphere $\mathcal{S}$, a discretization of the possible step directions is suggested, which converts the integration over the step direction $\hat{\mathbf{v}}_i$ in (8) to a summation. Thus, combining (7) and (8), and taking the simplifications into account, gives the distribution

$$p\left(\hat{\mathbf{v}}_i = \hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{i-1}, \boldsymbol{\mathcal{D}}\right) = \frac{p\left(\boldsymbol{\mathcal{D}} | \hat{\mathbf{v}}_k, \boldsymbol{\theta}\right) p\left(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{i-1}\right)}{\sum\limits_{\hat{\mathbf{v}}_k \in \mathcal{S}} p\left(\boldsymbol{\mathcal{D}} | \hat{\mathbf{v}}_k, \boldsymbol{\theta}\right) p\left(\hat{\mathbf{v}}_k | \hat{\mathbf{v}}_{i-1}\right)}. \quad (15)$$

In this equation, $\hat{\mathbf{v}}_i$ is a random variable which equals one of the predefined step directions $\hat{\mathbf{v}}_k \in \mathcal{S}$ with a certain probability. The likelihood function $p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_k, \boldsymbol{\theta})$ is given by (11), and the prior distribution $p(\hat{\mathbf{v}}_k|\hat{\mathbf{v}}_{i-1})$ by (12).

The steps below outline the procedure for generating one fiber. The parameters that must be chosen beforehand are the set of vectors $\mathcal{S}$ spread over the unit sphere, a seed point, a step length, and termination criteria.

1) *Estimate Tensor Model for the Current Point*:
The tensor can be estimated using a conventional linear least squares approach, or some more advanced approach where the eigenvalues are constrained to be nonnegative [24], [25].
2) *Termination Criteria*:
The tracking should be terminated if the diffusion anisotropy is too low, according to for example the fractional anisotropy or relative anisotropy [26], or some other measure of anisotropy.
3) *Estimate Nuisance Parameters in the Constrained Model*:
The nuisance parameters in the constrained model α, β are estimated according to (14), and $\hat{\sigma}^2$ as the residual noise variance.
4) *Calculate Likelihood Function*:
The likelihood $p(\boldsymbol{\mathcal{D}}|\hat{\mathbf{v}}_k, \boldsymbol{\theta})$ of each of the predefined directions $\hat{\mathbf{v}}_k \in \mathcal{S}$ is obtained according to (11).
5) *Calculate Prior Distribution*:
The prior probability $p(\hat{\mathbf{v}}_k|\hat{\mathbf{v}}_{i-1})$ for each of the predefined directions $\hat{\mathbf{v}}_k \in \mathcal{S}$ is obtained according to (12).
6) *Calculate Posterior Distribution*:
The posterior probability $p(\hat{\mathbf{v}}_i = \hat{\mathbf{v}}_k|\hat{\mathbf{v}}_{i-1}, \boldsymbol{\mathcal{D}})$ for each of the directions $\hat{\mathbf{v}}_k \in \mathcal{S}$ is obtained by multiplying the likelihood with the prior and normalizing to unity, according to (15).
7) *Draw Random Direction*:
One of the directions in $\mathcal{S}$ is drawn at random according to the posterior distribution $p(\hat{\mathbf{v}}_i = \hat{\mathbf{v}}_k|\hat{\mathbf{v}}_{i-1}, \boldsymbol{\mathcal{D}})$. Since the posterior distribution is discrete, this is simply done by letting each of the directions in $\mathcal{S}$ occupy an interval between 0 and 1 of length equal to its probability, and then choosing the interval in which a uniform random number between 0 and 1 falls. A step of the predefined step length is taken in the drawn direction.
8) *Stochastic Interpolation*:
As described in [27], choose a neighboring voxel at random according to the distance to the current point. Use the diffusion data in this voxel for the next step.
9) *Go to Step 1)*

When generating a large number of fibers according to the above steps, many voxels will be visited many times. The computationally most demanding step is the calculation of the likelihood function in Step 4. Great computational savings are therefore made if the likelihood function is stored once it has been calculated for a voxel. The tracking algorithm can then start from Step 5 when a previously visited voxel is encountered.

B. Calculating Probability Maps

Section II describes how the probability $p(A \rightarrow B|\boldsymbol{\mathcal{D}})$ is calculated. By letting A be a user-defined seed voxel, and calculating $p(A \rightarrow B|\boldsymbol{\mathcal{D}})$ for B equal to all other voxels, a map of connection probabilities is obtained. An important question is how many fiber samples are required for an acceptable approximation of the probability of a connection, i.e., for (4) to approximate (2). The accuracy of the rejection sampling procedure underlying (4) depends on the number of sampled paths. A sampled path that was terminated after 10 steps can be seen as 10 paths of different lengths. Hence, the number of sampled paths N_n of a particular length n decreases monotonically as the length increases. For example, there will be many paths of length 1 but fewer of length 100, because some paths will be terminated before they reach this length. At some length $n_{\max}$ there are not enough paths to give an acceptable approximation of the integral in (2), that is, N_n in (4) becomes too small for lengths $n > n_{\max}$. As a rule of thumb, we generate between 5000 and 10 000 paths from each seed point, and require N_n to be at least 1000. This seems to be sufficient in the sense that the spatial connectivity maps do not change much if more samples are generated. Another way of viewing this is that the distribution of the path length $p(n)$ is set to zero for $n > n_{\max}$. In this work, we set $p(n)$ to a uniform distribution over the lengths 1 to $n_{\max}$. The probability of a connection for each voxel is obtained via (4), with the outer sum going from 1 to $n_{\max}$ (or to between limits dictated by $p(n)$ in case another distribution is chosen). The time it takes to calculate a map of connection probabilities is negligible compared to the time it takes to sample the fiber paths. Finally, each generated fiber path has a probability associated with it according to (5). An estimate of the maximum probability path between two regions is thus given by the fiber with the maximum probability among the generated paths between these regions.

V. Results

For the results, a diffusion weighted data set consisting of 31 volumes acquired with different gradient directions ($b = 1000$ s/mm^2), and one baseline volume acquired with no diffusion weighting ($b = 0$ s/mm^2), is used. The data were acquired using a General Electric 3 T whole-body scanner and a dual spin echo EPI imaging sequence, with $\mathrm{TR} = 11.5$ ms and $\mathrm{TE} = 88.9$ ms. The images cover a field of view of 24 cm using a 128 × 128 grid, which subsequently was interpolated to a 256 × 256 grid. 31 image slices were acquired, with a slice thickness of 3.5 mm. The signal-to-noise ratio (SNR) in white matter in the $b = 0$ volume, calculated as image intensity divided by standard error of the noise, is $\mathrm{SNR} = 4.6$ in this data set. A slight smoothing of the images, using a Gaussian kernel of a width of 1.2 voxel full-width at half-maximum (FWHM), was applied to the original data volumes to avoid problems with negative eigenvalues in the tensor estimation. A set of T1-weighted anatomical images were also acquired, and used as overlay images for some of the results below.

Before showing tracking results, we first compare the behavior of the diffusion models in Table I; investigate how the dirac priors for the nuisance parameters affect the uncertainty in fiber orientation; and examine how the single-fiber model behaves when there are multiple fiber orientations present in a voxel. We then move on to show examples of tracked fibers and maps of connection probabilities. Unless otherwise stated, the constrained observation model in Table I is used.

A. Diffusion Model

We begin with a comparison of the water diffusion models in Table I. The focus is on any potential differences between these models. The widely used tensor model is used as reference with which the compartment model used by Behrens *et al.* [7] and the constrained model used here are compared. To this end, these

TABLE II
COMPARISON OF THE THREE WATER DIFFUSION MODELS LISTED TO THE LEFT. THE TENSOR MODEL IS USED AS REFERENCE. THE ANGULAR DIFFERENCE IS THE AVERAGE DIFFERENCE IN PRINCIPAL DIFFUSION ORIENTATION IN WHITE MATTER. THE RESIDUAL VARIANCE IS THE VARIANCE OF THE NOISE REMAINING AFTER THE MODEL HAS BEEN FITTED, AVERAGED ACROSS ALL WHITE MATTER VOXELS

	Angular difference	Residual variance
Tensor model	-	1
Constrained model	0°	1.20
Compartment model	3.2°	1.20

models were first fitted to the diffusion measurements in white matter voxels. The tensor model was fitted using a conventional linear least squares approach, the constrained model via (14), and the compartment model was fitted using an iterative nonlinear least squares method using the MATLAB Optimization Toolbox. Subsequently, using the tensor model as reference, the differences in the principal water diffusion orientation and the variance of the residual data were investigated. The angular difference in diffusion orientation was calculated for each voxel as $(360/2\pi)\arccos(|\hat{\mathbf{v}}_C^T\hat{\mathbf{v}}_T|)$, where $\hat{\mathbf{v}}_T$ is the first eigenvector of the diffusion tensor in the tensor model, and $\hat{\mathbf{v}}_C$ is the diffusion direction in the constrained or compartment model. The residual variance in a voxel using the tensor model was calculated as

$$\sigma_T^2 = \frac{1}{N-7}\sum_{j=1}^{N}\left(y_j - \mu_j^*\right)^2 \tag{16}$$

where N again is the number of diffusion weighted image volumes (32 in our case), y_j is the voxel intensity in volume number j, and μ_j^* is the intensity in this voxel predicted by the tensor model with the optimal estimated parameters inserted. The variances for the compartment and the constrained models are calculated similarly, but with the normalizing factor $N-5$ instead, because there are only five parameters in these models, compared to the seven parameters in the tensor model. The angular difference and the residual variance, averaged over all white matter voxels, are reported in Table II. First, the principal water diffusion orientation is the same in the tensor model and the constrained model, leading to the 0° angular difference. The angular difference between the main diffusion orientation in the tensor model and the compartment model in white matter averages on 3.2°. The practical consequence of this difference is illustrated in Fig. 5, where a conventional tracking following the principal diffusion orientation in the constrained model and the compartment model is shown. Turning to the residual noise variance comparison, we see that the residual noise variances for the constrained model and the compartment model are 20% higher compared to the tensor model. This is explained by the extra free parameters the tensor model has to improve the model fit (seven versus five parameters). The important observation here is that there is no difference in the goodness of fit between the constrained model and the compartment model.

The aforementioned results are valid only for white matter voxels, where a certain diffusion anisotropy can be expected. In voxels with more isotropic water diffusion, such as in gray matter voxels, larger differences in, for example, the estimated main diffusion orientation can be expected. However, since the models are supposed to operate in white matter voxels, and the fiber tracking is terminated when isotropic voxels are encountered, any differences outside white matter have little effect on the final results. As stated previously, the advantage of the constrained model lies in its mathematical tractability; fitting the constrained model in every voxel in a data set is done in a couple of seconds while fitting the compartment model takes a couple of hours, i.e., a difference of three orders of magnitude.

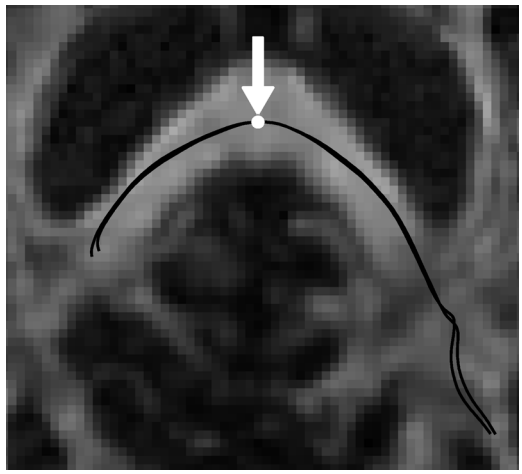

Fig. 5. Tracking example using the constrained model and the compartment model. Tracking was carried out by tracing the principal water diffusion direction, as pointed out by the constrained model and compartment model. The tracking was initiated in the marked point in Corpus callosum, and then the diffusion orientation was followed in both directions and for both diffusion models.

B. Prior Distributions

In this section, we investigate how the use of dirac priors influences the posterior distribution of the fiber orientation. To this end, we compare the posterior distribution obtained using dirac priors with the posterior distribution obtained using uniform distributions as priors. Predictably, saying that we are absolutely certain about the values of the nuisance parameters leads to a reduced subjective uncertainty regarding the fiber orientation. The uniform prior distributions are chosen to cover a $\pm 20\%$ interval around the point estimates of the nuisance parameters, indicating an uncertainty regarding their exact values. To calculate the posterior distribution in this case, we need to evaluate a six-dimensional integral for the normalization factor in (8). This is done using trapezoidal integration. Because of the dimensionality, the computational time is a few minutes. This should be compared to the dirac prior case, where the integral in (8) is defined over the unit sphere (i.e., two-dimensional) and the numerical evaluation is extremely fast. Resulting posterior pdf's for three voxels with different degrees of diffusion anisotropy are shown in Fig. 6. Clearly, the distributions obtained using uniform priors and dirac priors are nearly indistinguishable. A very close examination reveals that the posterior pdfs obtained using the uniform priors indeed are somewhat broader, but this small difference has no practical relevance. This result indicates that taking uncertainty in nuisance parameters into account has a small effect on the uncertainty in fiber orientation, and it thereby justifies the use of dirac priors for the nuisance parameters and eliminates the need to apply MCMC sampling.

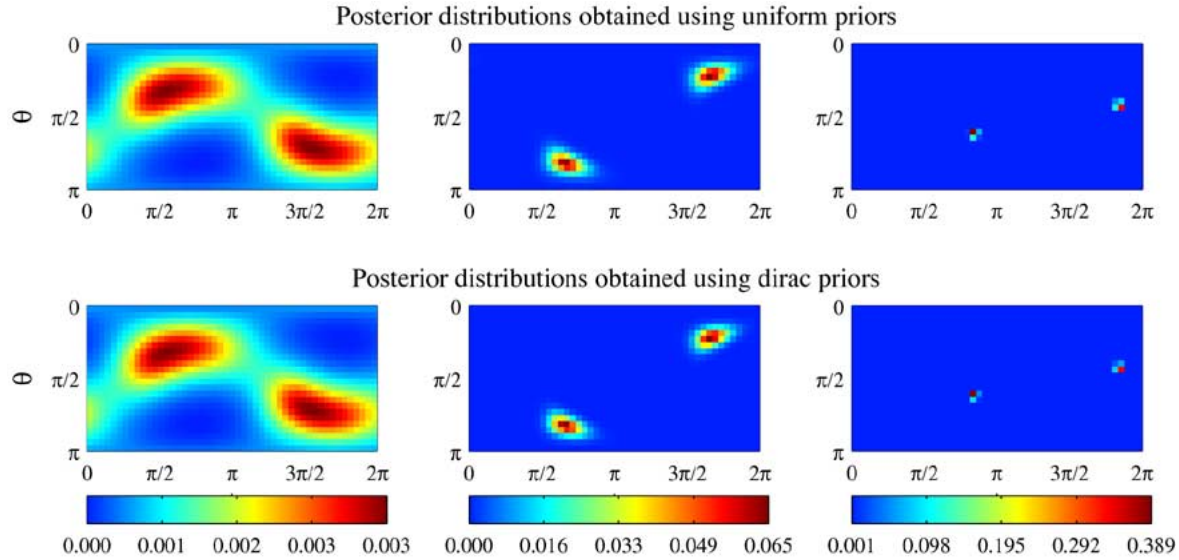

Fig. 6. Comparison of the posterior pdfs obtained using noninformative uniform priors (upper row) and dirac priors (lower row) for the nuisance parameters. Results for three voxels, one in each column, with different degrees of diffusion anisotropy are shown. The leftmost column shows a voxel from gray matter with low anisotropy, the middle column shows a voxel from the External capsule, and the rightmost column shows a voxel with high anisotropy from the Corpus callosum.

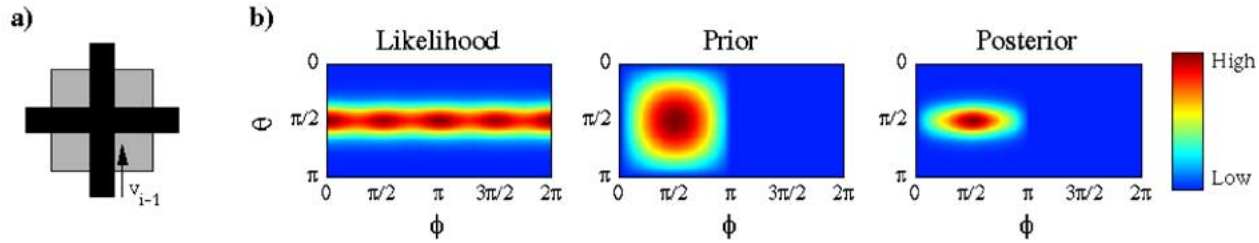

Fig. 7. (a) Schematic image of the fiber crossing used as basis for simulating diffusion measurements. It is assumed that we are tracking the vertical fiber bundle, with the last step direction $\mathbf{v}_{i-1}$ as indicated in the figure. (b) Likelihood of the next step direction given the simulated diffusion measurements, the prior distribution of the next step direction given the previous step, and the posterior obtained as the product of the likelihood and the prior.

C. Crossing Fibers

Next, the behavior in a voxel where two fibers cross is investigated. For this purpose, simulated diffusion measurements for a 90° crossing in the x–y plane were generated, see Fig. 7(a) for a schematic image. The compartment model with three compartments, one compartment with isotropic water diffusion and two compartments with anisotropic diffusion along the fiber directions, was used to generate simulated measurements (31 gradient directions and a b value of 1000 s/mm^2). The compartment model was used to deliberatively create a mismatch between the simulated data and the constrained model used for analyzing the data. We assume that we are tracking the fiber running along the y axis, with the previous step $\mathbf{v}_{i-1}$ indicated in Fig. 7(a). Fig. 7(b) shows how the likelihood (11) and prior (12) are combined to a posterior distribution of the next step direction. The likelihood function indicates that the single-fiber model fits better when aligned with either of the two fiber bundles, but the uncertainty as to where to go next is high in the x–y plane. When the information about the previous step direction is added, the uncertainty becomes significantly lower and the probability of continuing through the crossing is high, as indicated by the posterior distribution. Hence, when repeated a large number of times, some fiber paths will meander off to the right or left, but the majority of the fibers will find the correct path. This illustrates that the single-fiber model has some ability to handle complex neighborhoods if used in a stochastic tracking framework, and that the lack of model fit will be manifested as an increased uncertainty captured by the posterior distribution of the fiber orientation.

D. Sampling Fiber Paths

In this section, we show examples of real data fiber paths. A step length of 1 mm was used for generating these paths. The tracking was terminated when the ratio $\beta/(\alpha+\beta)$, where α and β are parameters indicating isotropy and anisotropy, respectively, in the constrained model, became too low (in this case, lower than 0.2). This ratio will attain values between 0 and 1, where 0 means that the water diffusion is isotropic and values close to 1 indicate highly anisotropic diffusion. This anisotropy measure is related to the linear tensor shape measure proposed by Westin *et al.* [27]. However, one can use for example the fractional anisotropy or relative anisotropy as terminating criteria as well; this choice has no important impact on the final result. In the first example, one point in the right and one point in the left Cingulum bundle were chosen as seed points. 1000 paths initiated in each of these points were then sampled, see

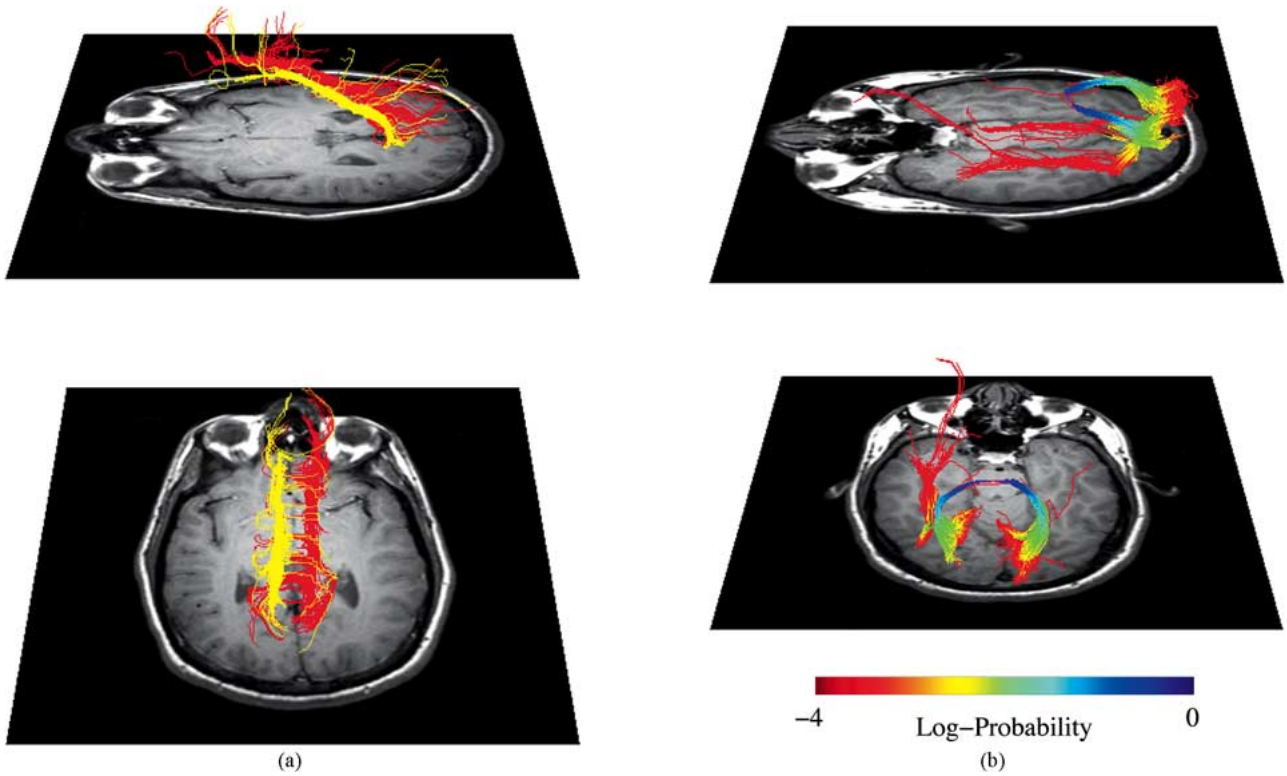

Fig. 8. (a) Fiber samples tracing the left and right Cingulum bundles. All the red fibers originate from the same seed point, as do all the yellow paths. The seed points are located where the bundles cross the axial slice. (b) 3000 fiber samples initiated in the splenium of Corpus callosum, see Fig. 9. The coloring indicates how the probability evolves along the fiber paths according to (5).

Fig. 8(a). Note how some paths slip into the underlying Corpus callosum, but that the massive sampling approach yields a robust delineation of the bundles. Fig. 8(b) shows another example with 3000 paths seeded in a point in the splenium of Corpus callosum (marked in Fig. 9). The coloring of these paths indicates how the probability for each fiber evolves along the paths according to (5).

These examples show how a stochastic tracking framework is able to handle splitting fiber bundles and ambiguous neighborhoods in a way that is not possible with conventional tracking of the principal diffusion direction. Computationally, the stochastic tracking of the fiber paths is by far the most demanding part in the presented methodology. Using a high-end PC of today, our MATLAB implementation typically generates a few paths per second. Still, since a few thousand fibers are required, the computational time ranges from a few minutes to an hour.

E. Probability of a Connection

Finally, from the sample paths shown in the previous section, the probability that a fiber seeded in a user defined point reaches a specific voxel can be estimated by applying the theory presented in Section II. As described in Section IV, to get an overview of the uncertainty associated with a fiber path seeded in a specific point, a map of connection probabilities can be produced. Here, a few examples of such probability maps are demonstrated. Each of these maps has been generated from 10 000 fibers initiated in the points indicated by the arrows in the figures. In the first example, shown in Fig. 9, tracking was initiated in the splenium of Corpus callosum. The maximum length $n_{\max}$, discussed in the Implementation section, in this example was 84 steps, i.e., 8.4 centimeters using a step length

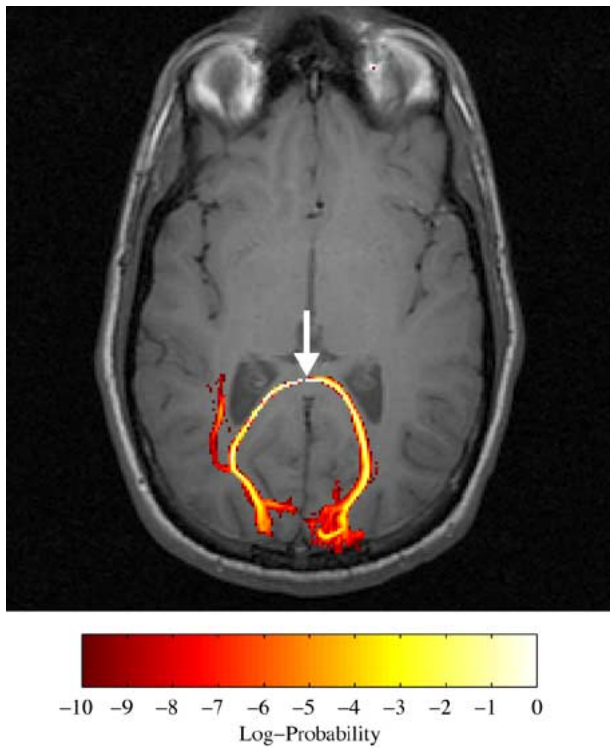

Fig. 9. Maximum intensity projection probability map. The coloring indicates the probability that a fiber seeded in the splenium of Corpus callosum, marked by the arrow, reaches respective voxel.

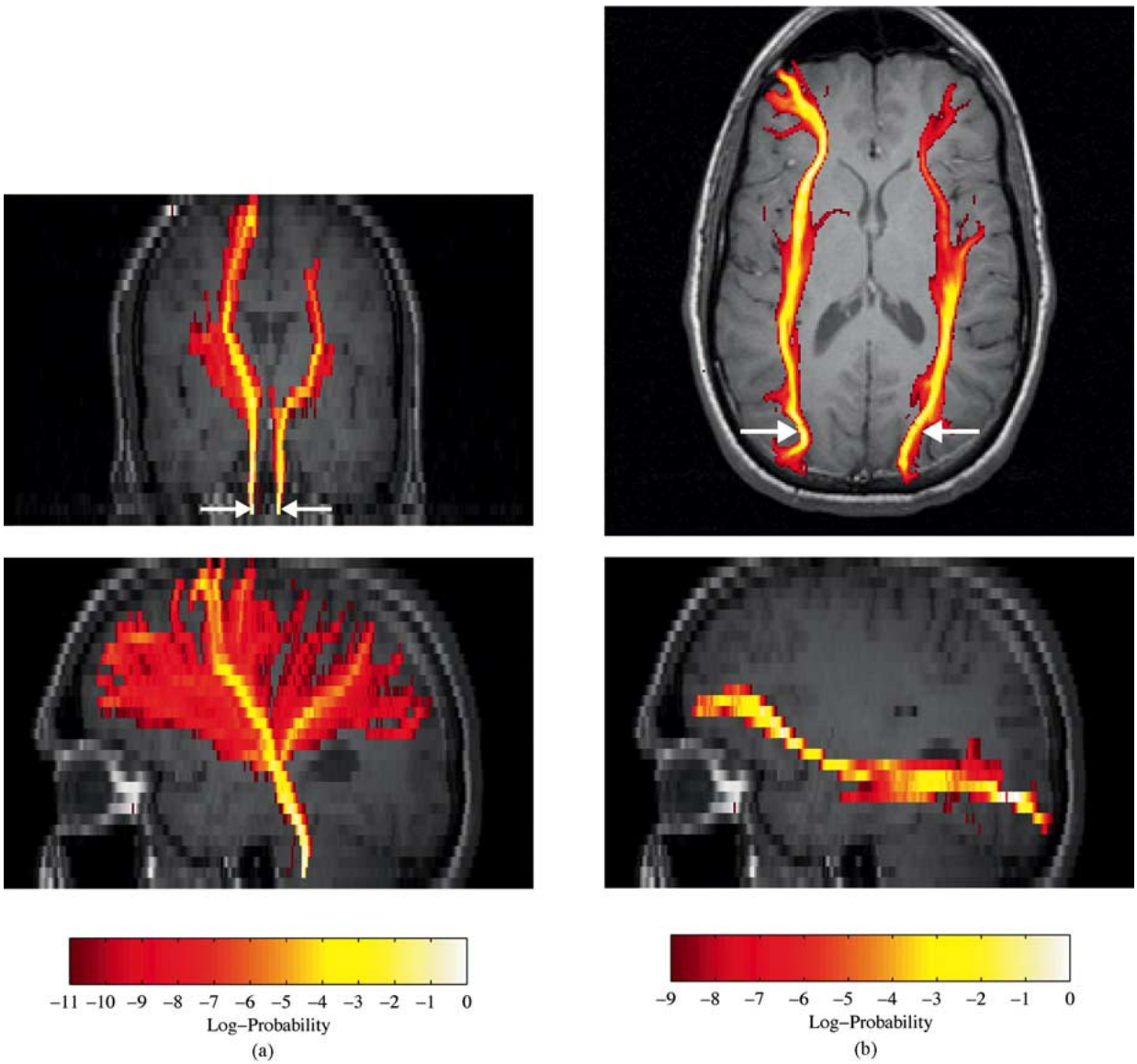

Fig. 10. (a) Coronal and sagittal maximum intensity projection probability maps seeded in the Corticospinal tract. (b) Axial and sagittal maximum intensity projection probability maps of fibers seeded in the points in the Inferior occipitofrontal fasciculi marked by the arrows.

of 1 mm. It is important to stress that the coloring should be interpreted as our belief that a single fiber initiated at the seed point reaches respective voxel, and not that fibers going through the seed point in Corpus callosum diverge and spread into the Occipital cortex. It is known that fibers in specific parts of the Corpus callosum connect to very specific locations of the cerebral cortex [28], [29], but Fig. 9 indicates that the diffusion data does not have the precision to resolve such a mapping.

More examples of connection probability maps of longer fiber tracts are shown in Fig. 10. In Fig. 10(a), the Corticospinal tract is tracked by seeding two points at the level of the Pons. The image shows how the Corticospinal tract is followed until the Cerebral peduncles, where the tracking fans out in a v-shape into the Internal capsule. Note also the high probability of the paths heading towards the Motor cortex. The maximum length considered in this case is $n_{\max} = 144$ or 14.4 cm. In Fig. 10(b), the left and right Inferior occipitofrontal fasciculi are delineated by seeding fibers at the points marked by the arrows. To achieve this delineation, prior knowledge about the approximate extent of the Inferior occipitofrontal fasciculi were implemented by restricting the space of allowable fiber paths, i.e., fibers going outside a region of interest were removed. Without this restriction, there would be some fibers entering the Corpus callosum due partial voluming effects (cf. Fig. 9). The maximum considered length in this case was $n_{\max} = 168$ steps or 16.8 cm.

VI. Discussion

White matter tractography methods are currently under intensive development. The first methods used for tracking fibers in DT-MRI data are commonly referred to as streamline methods, because they draw paths in vector fields given by the principal diffusion orientations [30]–[35]. A major drawback with the streamline tracking methods is that they do not provide or visualize the confidence in the estimated fiber paths; and if not interpreted carefully, they might even give an impression of false

certainty. Several authors have, therefore, recently developed so-called probabilistic or stochastic tracking methods, which attempt to take uncertainty stemming from for example noise and diffusion model imperfections into account [1]–[10]. Apart from giving a more complete picture of the evidence of connections in the brain, probabilistic methods have also been noted to behave more robustly in the presence of fiber crossings and to be able to operate under looser constraints, for example in terms of stopping criteria, allowing them to pass low-anisotropy areas and to penetrate deeper into gray matter where the fibers start and end [7].

In this work, a novel probabilistic method for assessing uncertainty in white matter tractography is introduced. The main merit of this method is that it rests on a solid Bayesian theoretical foundation, while yet being simple to implement. A top-down presentation of the method is given, in which all the modeling and approximation steps are highlighted. Modeling and estimation are carried out at two levels: a global level and a local level. At the global level, the question is how the probability that a fiber seeded in a point or area A reaches and area B is estimated. The theoretical foundation of this estimation is presented in Section II. A similar procedure has been used heuristically by several authors previously [2], [4], [6], [7], [9]. In this work, a general and comprehensive description of the theoretical background of this procedure is presented.

Algorithmically, the main difference between the approach presented in this work and previously presented probabilistic methods lies in the local modeling of the fiber orientation. The local fiber orientation is estimated from the diffusion data, and uncertainty enters in this process due to image noise and complex fiber architectures which the observation model is unable to resolve. The goal is to find a probability density function of the local fiber orientation. Since this can be computationally demanding, several authors have employed heuristic distributions, based for example on parameters derived from the diffusion tensor [1], [2], [4], [6]. However, such distributions have little theoretical justification, and they should be contrasted with a Bayesian approach, in which our prior knowledge about parameters and noise enter naturally.

To apply Bayes' theorem, we need to specify an observation model for the water diffusion and provide prior probability distributions of the parameters in this model. It is preferable to use an observation model where the underlying fiber architecture is expressed explicitly, because it is on this architecture we wish to make inference. Herein lies the difference between the tensor model and the constrained and compartment models. The latter models cannot attain an oblate diffusion profile, i.e., the case when the eigenvalues in the tensor model have the relation $\lambda_1 \approx \lambda_2 \gg \lambda_3$. The tensor adopts an oblate shape in voxels where fibers cross or split. As was exemplified in Section V-C, in a stochastic framework using the constrained model or the compartment model, these cases will instead be captured as uncertainty regarding the fiber orientation. This is a key mechanism in the probabilistic tracking that cannot be stressed enough. Compared to the tensor model, which due to its two additional parameters can describe oblate diffusion shapes, the constrained model and the compartment model will in some cases have poorer model fits, as is shown in Section V-A. However, information is not lost in such voxels; it is captured in the posterior distribution of the next step direction and subsequently reflected in the maps of connection probabilities. This is in contrast to traditional streamline tracking using the tensor model, where the information in oblate tensor shapes is ignored because only the principal diffusion orientation is utilized. In the comparison between the compartment model and the constrained model in Section V-A, it was found that these models perform similarly in terms of describing the diffusion data. The simple estimation of the parameters via the theorem in the Appendix makes the constrained model an attractive choice. Note that the popularity enjoyed by the tensor model is much due to the same reason; its widespread use can largely be attributed to the simple linear least squares procedure for estimating the diffusion tensor.

A single-fiber model is clearly insufficient in voxels containing two or more fiber tracts, and it is well known that such voxels exist [19], [36]–[39]. Therefore, we investigated how the single-fiber model behaves when there are two fibers crossing in a voxel, and demonstrated how the lack of model fit is translated into uncertainty captured by the posterior distribution of the fiber orientation. Conversely, a two-fiber model will over-fit in a voxel with only one fiber, leading to an artificial reduction of uncertainty. Hence, the single-fiber model will always give a more conservative approximation of the uncertainty than a higher order model. Note that both a single-fiber and two-fiber model will over-fit in voxels with no dominant fiber structure, e.g., in gray matter voxels. To obtain an accurate approximation of the uncertainty, which ideally should stem from noise only, we must choose the correct model order for every voxel. Work in this direction has recently been presented by Parker and Alexander [10]. When discussing models, it should also be mentioned that some very recent work has been presented where tracking is performed using nonparametric (i.e., model free) descriptions of the water diffusion [10], [40]. As mentioned previously, nonparametric descriptions of the water diffusion, obtained for example via q-ball imaging [3] or persistent angular structure MRI [11], has the potential to resolve complex fiber neighborhoods and may prove a very successful strategy in the future. The current drawbacks of these methods are the longer scanning times and a sensitivity to noise (in the absence of a regularizing model) that has been observed to generate spurious structure in voxels with isotropic diffusion [11], [18], [40].

A critical part of a Bayesian framework is the prior distributions of the model parameters. Behrens *et al.* [7] use noninformative prior distributions for the nuisance parameters, leading to a full Bayesian modeling. Since the expression for the posterior distribution of the local fiber orientation then becomes complex, involving high-dimensional integrals, they use MCMC methods for generating random samples from the posterior distribution. In the current work, dirac priors are used instead. This is an approximation that implies an overconfidence in the estimation of the nuisance parameters, which in turn may lead to an underestimate of the uncertainty in fiber orientation. To investigate this potential problem, we compared the posterior distributions of the fiber orientation obtained using noninformative priors and the dirac priors, and found that the difference is negligible; the uncertainty in nuisance parameters has a very small effect on the uncertainty in fiber direction. In our Bayesian formula-

tion, a prior distribution of the next step direction is also used. This prior distribution replaces the momentum term commonly used in traditional streamline tracking to prevent the fibers from having implausible curvature. The prior on the next step direction used in this work is quite heuristic, and in our future research we plan to replace it with an empirically derived distribution obtained by studying real fibers using histology techniques.

One of the crucial differences between the work by Behrens *et al.* [7], and the work presented here, is how random step directions are drawn from a posterior distribution. For this Behrens *et al.* use MCMC techniques. MCMC methods simulate Markov chains which are known theoretically to converge and give samples from the target posterior distribution [41]. This is a general approach which conceptually is quite simple. In practice, however, it is not straightforward to implement a well working MCMC method. MCMC methods are by nature iterative, and issues such as convergence, i.e., how many iterations that are required for the Markov chain to produce samples from the desired distribution, must be checked. The MCMC process is also governed by a number of additional parameters that must be adjusted to give a desired so-called mixing rate of the Markov chain. These issues become especially cumbersome in white matter tractography, where a large number of chains must be simulated, different parameters are required for each voxel, and convergence and mixing of the Markov chains must be monitored automatically. With the dirac priors used in this work, we circumvent computationally demanding numerical integration and marginalization procedures in (7) and (8), and there is no need to apply MCMC methods. Instead the posterior distribution is approximated by a discrete probability function using 2562 predefined directions evenly spread over the unit sphere, and one of these directions is chosen as the next step direction when sampling fiber paths. This makes the calculation of the posterior and the generation of random samples from it very simple. The validity of the approach requires that small errors in the estimation of the nuisance parameters have no, or at least only weak, influence on the uncertainty of the fiber orientation. This is in fact required in the method by Behrens *et al.* too, because the MCMC method proposed in [7] draws samples from a joint distribution of the step direction *and* the nuisance parameters $p(\hat{\mathbf{v}}_i, \boldsymbol{\theta})$, and not from the marginalized distribution $p(\hat{\mathbf{v}}_i)$. This means the random step direction will have a dependence on the nuisance parameters. Rewriting $p(\hat{\mathbf{v}}_i, \boldsymbol{\theta}) = p(\hat{\mathbf{v}}_i|\boldsymbol{\theta})p(\boldsymbol{\theta})$ we see that $\hat{\mathbf{v}}_i$ can be seen drawn from the distribution $p(\hat{\mathbf{v}}_i|\boldsymbol{\theta})$ and not from $p(\hat{\mathbf{v}}_i)$; these two distributions are the same only when the nuisance parameters $\boldsymbol{\theta}$ are independent of the step direction $\hat{\mathbf{v}}_i$. This is clearly not true in general, but as is shown in Section V-B, the dependence is quite weak close to the actual values of the nuisance parameters. Hence, the results in Section V-B validate not only the work presented here, but also the work presented by Behrens *et al.* [7]. It can also be noted that the maximum probability path is not accessible if we use MCMC methods, because MCMC methods do not provide the actual probabilities of the random directions they generate, which is required for evaluating the probability of an entire path according to (5).

The end product in this work is the probability $p(A \rightarrow B|\mathcal{D})$, i.e., the probability that a single fiber seeded in a point or area A reaches an area B, given the diffusion measurements. Maps of this probability can be useful, for example, in surgical planning and for studying connectivity in mental and neurological disorders. As we are working in a Bayesian paradigm, probability should be interpreted as a subjective belief which is dependent on the supplied models. Hence, to validate the proposed approach, we need to investigate the models on which it is based, i.e., the water diffusion model, noise model and prior distributions. The constrained diffusion model is based on the tensor model, which is extensively used and generally assumed to be a reasonable model. Note that when applying a conventional principal diffusion direction tracking there is no difference between the tensor model and the constrained model, and issues such the alignment of the fibers and water diffusion and error analyzes have been discussed elsewhere [42], [43]. The noise model is based on a Gaussian additive noise in the frequency domain, in which MRI measurements are acquired. This is a theoretically and experimentally well-founded model [20]. Using dirac priors for the nuisance parameters is clearly a simplification, but again, it is shown in the Results section that this has a very small influence on the uncertainty in fiber orientation.

VII. Conclusion

A method for stochastic white matter tractography has been presented. The necessary modeling and estimation have been described at both a global level and a local level. At the global level, a theoretical foundation for estimating the probability of a connection between two areas in the brain has been detailed. At the local level, probability density functions of the fiber orientation are derived in a theoretically justified way via Bayes' theorem. In addition, a theorem that facilitates estimation of the parameters in a constrained version of the popular tensor model of the water diffusion has also been presented. To arrive in a computationally feasible solution, the use of dirac prior distributions for the nuisance parameters was investigated, and it was found that this approximation has a very weak influence on the uncertainty in fiber orientation in a voxel. As a consequence there is no need to apply MCMC sampling for drawing samples in the stochastic tracking of fiber paths, which makes the method simple to implement.

Appendix

Theorem: Let $\mathbf{D}$ be a symmetric 3×3 matrix with eigenvalue factorization

$$\mathbf{D} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^T + \lambda_3 \mathbf{e}_3 \mathbf{e}_3^T, \qquad \lambda_1 \geq \lambda_2 \geq \lambda_3.$$

Let $\mathfrak{S}$ be the space of symmetric 3×3 matrices where the two smallest eigenvalues are equal. The matrix $\mathbf{S} \in \mathfrak{S}$ minimizing $||\mathbf{D} - \mathbf{S}||_{\mathrm{F}}$ is given by

$$\mathbf{S} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \frac{\lambda_2 + \lambda_3}{2} \left(\mathbf{e}_2 \mathbf{e}_2^T + \mathbf{e}_3 \mathbf{e}_3^T \right). \tag{17}$$

If $\lambda_1 > \lambda_2$, the solution is unique. If $\lambda_1 = \lambda_2 > \lambda_3$, there are multiple solutions which are all given by (17) for different choices of $\mathbf{e}_1$ in the eigen-space corresponding to the largest

eigenvalue. If all eigenvalues are equal, the solution is trivially unique.

Proof: Any matrix $\mathbf{S} \in \mathfrak{S}$ can be written

$$\mathbf{S} = \gamma_1 \hat{\mathbf{u}}\hat{\mathbf{u}}^T + \gamma_2(\mathbf{I} - \hat{\mathbf{u}}\hat{\mathbf{u}}^T)$$

where $\hat{\mathbf{u}}$ is a unit vector and $\gamma_1 \geq \gamma_2$. The squared Frobenius distance to $\mathbf{D}$ is

$$\begin{aligned}
||\mathbf{D}-\mathbf{S}||_F^2 &= \mathrm{tr}\left[\left(\mathbf{D} - \gamma_1\hat{\mathbf{u}}\hat{\mathbf{u}}^T - \gamma_2(\mathbf{I}-\hat{\mathbf{u}}\hat{\mathbf{u}}^T)\right)^2\right] \\
&= ||\mathbf{D}||_F^2 + \gamma_1^2 + 2\gamma_2^2 - 2\gamma_1\hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}} \\
&\quad - 2\gamma_2(\mathrm{tr}\mathbf{D} - \hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}}) \\
&= ||\mathbf{D}||_F^2 + (\gamma_1 - \hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}})^2 - (\hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}})^2 \\
&\quad + 2\left(\gamma_2 - \frac{1}{2}(\mathrm{tr}\mathbf{D} - \hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}})\right)^2 \\
&\quad - \frac{1}{2}(\mathrm{tr}\mathbf{D} - \hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}})^2 \\
&= ||\mathbf{D}||_F^2 + (\gamma_1 - \hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}})^2 \\
&\quad + 2\left(\gamma_2 - \frac{1}{2}(\mathrm{tr}\mathbf{D} - \hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}})\right)^2 \\
&\quad - \frac{3}{2}\left(\frac{1}{3}\mathrm{tr}\mathbf{D} - \hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}}\right)^2 - \frac{1}{3}(\mathrm{tr}\mathbf{D})^2. \quad (18)
\end{aligned}$$

The expression $\hat{\mathbf{u}}^T\mathbf{D}\hat{\mathbf{u}}$ can take any value a in the interval $\lambda_1 \geq a \geq \lambda_3$. Let $b = (1/2)(\mathrm{tr}\mathbf{D} - a)$. With this notation, the squared distance can be written

$$||\mathbf{D}||_F^2 + (\gamma_1 - a)^2 + 2(\gamma_2 - b)^2 - \frac{2}{3}(a-b)^2 - \frac{1}{3}(\mathrm{tr}\mathbf{D})^2. \quad (19)$$

For the moment, let a in (19) be fixed and study the two terms depending on γ_1 and γ_2, $(\gamma_1 - a)^2 + 2(\gamma_2 - b)^2$. If $a \geq b$, this is minimized by $\gamma_1 = a, \gamma_2 = b$, and the minimum value is 0. If instead $a < b$, the minimum is achieved for $\gamma_1 = \gamma_2 = (a + 2b)/3$ and the minimum value is $(2/3)(a-b)^2$. This gives

$$\begin{aligned}
&\min_{\gamma_1,\gamma_2} ||\mathbf{D}-\mathbf{S}||_F^2 \\
&\quad = \begin{cases} ||\mathbf{D}||_F^2 - \frac{1}{3}(\mathrm{tr}\mathbf{D})^2 - \frac{2}{3}(a-b)^2, & a \geq b \\ ||\mathbf{D}||_F^2 - \frac{1}{3}(\mathrm{tr}\mathbf{D})^2, & a < b \end{cases} \\
&\quad = \begin{cases} ||\mathbf{D}||_F^2 - \frac{1}{3}(\mathrm{tr}\mathbf{D})^2 - \frac{3}{2}\left(a - \frac{1}{3}\mathrm{tr}\mathbf{D}\right)^2, & a \geq \frac{1}{3}\mathrm{tr}\mathbf{D} \\ ||\mathbf{D}||_F^2 - \frac{1}{3}(\mathrm{tr}\mathbf{D})^2, & a < \frac{1}{3}\mathrm{tr}\mathbf{D} \end{cases}. \quad (20)
\end{aligned}$$

It is clear that this is minimized by choosing a as large as possible, i.e., $a = \lambda_1$. It follows that $\hat{\mathbf{u}}$ must be an eigenvector of $\mathbf{D}$ corresponding to the largest eigenvalue, $\gamma_1 = a = \lambda_1$, and $\gamma_2 = b = (\lambda_2 + \lambda_3)/2$. Equation (17) follows.

ACKNOWLEDGMENT

The authors would like to thank Dr. R. S.-J. Estepar, Dr. G. Kindlmann, and Dr. S. Peled for valuable input.

REFERENCES

[1] M. Björnemo, A. Brun, R. Kikinis, and C.-F. Westin, "Regularized stochastic white matter tractography using diffusion tensor MRI," presented at the 5th Int. Conf. Med. Image Computing Computer Assist. Intervention Tokyo, Japan, Sep. 25–28, 2002.

[2] M. Koch, D. Norris, and M. Hund-Georgiadis, "An investigation of functional and anatomical connectivity using magnetic resonance imaging," in *NeuroImage*, 2002, vol. 16, no. 1, pp. 241–250.

[3] D. Tuch, "Diffusion MRI of complex tissue structure," Ph.D. dissertation, Division of Health Sciences and Technology, Harvard–MIT, Cambridge, MA, 2002.

[4] G. Parker, H. Haroon, and C. Wheeler-Kingshott, "A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements," *J. Magn. Reson. Imag.*, vol. 18, no. 2, pp. 242–254, 2003.

[5] D. Jones, "Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI," *Magn. Reson. Med.*, vol. 49, no. 1, pp. 7–12, 2003.

[6] P. Hagmann, J. Thiran, L. Jonasson, P. Vandergheynst, S. Clarke, P. Maeder, and R. Meuli, "DTI mapping of human brain connectivity: Statistical fibre tracking and virtual dissection," *NeuroImage*, vol. 19, no. 3, pp. 545–554, 2003.

[7] T. Behrens, M. Woolrich, M. Jenkinson, H. Johansen-Berg, R. Nunes, S. Clare, P. Matthews, J. Brady, and S. Smith, "Characterization and propagation of uncertainty in diffusion-weighted MR imaging," *Magn. Reson. Med.*, vol. 50, no. 5, pp. 1077–1088, 2003.

[8] M. Lazar and A. Alexander, "Bootstrap white matter tractography (BOOT-TRAC)," *NeuroImage*, vol. 24, no. 2, pp. 524–532, 2005.

[9] D. Jones and C. Pierpaoli, "Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach," *Magn. Reson. Med.*, vol. 53, no. 5, pp. 1143–1149, 2005.

[10] G. Parker and D. Alexander, "Probabilistic anatomical connectivity using persistent angular structure obtained from diffusion weighted imaging," *Philos. Trans. Roy. Soc.*, ser. B, vol. 360, no. 1457, pp. 893–902, 2005.

[11] K. Jansons and D. Alexander, "Persistent angular structure: New insights from diffusion magnetic resonance imaging data," *Inverse Problems*, vol. 19, no. 5, pp. 1031–1046, 2003.

[12] J. Tournier, F. Calamante, D. Gadian, and A. Connelly, "Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution," *NeuroImage*, vol. 23, no. 2, pp. 1176–1185, 2004.

[13] P. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *J. Magn. Reson.*, ser. B, vol. 103, no. 3, pp. 247–254, 1994.

[14] O. Friman and C.-F. Westin, "Uncertainty in white matter tractography," in *Proc. 8th Int. Conf. Med. Image Comput. Computer Assist. Intervention (MICCAI'05)*, Palm Springs, CA, Oct. 2005, pp. 107–114.

[15] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*. Singapore: World Scientific, 1995.

[16] E. Veach, "Robust Monte Carlo methods for light transport simulation," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1997.

[17] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, 2001.

[18] D. Alexander, "An introduction to computational diffusion MRI: The diffusion tensor and beyond," in *Visualization and Image Processing of Tensor Fields*. New York: Springer, 2005.

[19] D. Tuch, T. Reese, M. Wiegell, N. Makris, J. Belliveau, and V. Wedeen, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magn. Reson. Med.*, vol. 48, no. 4, pp. 577–582, 2002.

[20] A. Macovski, "Noise in MRI," *Magn. Reson. Med.*, vol. 36, pp. 494–497, 1996.

[21] J. Sijbers, J. den Dekker, J. van Audekerke, M. Verhoye, and D. van Dyck, "Estimation of the noise in magnitude MR images," *Magn. Reson. Imag.*, vol. 16, no. 1, pp. 87–90, 1998.

[22] R. Salvador, A. Pena, D. Menon, T. Carpenter, J. Pickard, and E. Bullmore, "Formal characterization and extension of the linearized diffusion tensor model," *Hum. Brain Mapp.*, vol. 24, no. 2, pp. 144–155, 2005.

[23] D. Jones, M. Horsfield, and A. Simmons, "Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging," *Magn. Reson. Med.*, vol. 42, no. 3, pp. 515–525, 1999.

[24] Z. Wang, B. Vemuri, Y. Chen, and T. Mareci, "A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI," *IEEE Trans. Med. Imag.*, vol. 23, no. 8, pp. 930–939, Aug. 2004.

[25] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Fast and simple calculus on tensors in the log-Euclidean framework," in *Proc. 8th Int. Conf. Med. Image Comput. Computer Assist. Intervention (MICCAI'05)*, Palm Springs, CA, Oct. 2005, pp. 115–122.

[26] P. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *J. Magn. Reson.*, ser. B, vol. 111, no. 3, pp. 209–219, 1996.

[27] C.-F. Westin, S. Maier, H. Mamata, A. Nabavi, F. Jolesz, and R. Kikinis, "Processing and visualization for diffusion tensor MRI," *Med. Image Analy.*, vol. 6, no. 2, pp. 93–108, 2002.

[28] S. Lomber, B. Payne, and A. Rosenquist, "The spatial relationship between the cerebral cortex and fiber trajectory through the corpus callosum of the cat," *Behav. Brain Res.*, vol. 64, no. 1–2, pp. 25–35, 1994.

[29] K. Matsunaml, T. Kawashlma, S. Uekl, M. Fujita, and T. Konishi, "Topography of commissural fibers in the corpus callosum of the cat: A study using WGA-HRP method," *Neurosci. Res.*, vol. 20, no. 2, pp. 137–148, 1994.

[30] T. Conturo, N. Lori, T. Cull, E. Akbudak, A. Snyder, J. Shimony, R. McKinstry, H. Burton, and M. Raichle, "Tracking neuronal fiber pathways in the living human brain," *Proc. Nat. Acad. Sci.*, vol. 96, no. 18, pp. 10422–10427, 1999.

[31] S. Mori, B. Crain, V. Chacko, and P. van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann. Neurol.*, vol. 45, no. 2, pp. 265–269, 1999.

[32] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "In vivo fiber tractography using DT-MRI data," *Magn. Reson. Med.*, vol. 44, no. 4, pp. 625–632, 2000.

[33] B. Stieltjes, W. Kaufmann, P. van Zijl, K. Fredericksen, G. Pearlson, M. Solaiyappan, and S. Mori, "Diffusion tensor imaging and axonal tracking in the human brainstem," *NeuroImage*, vol. 14, no. 3, pp. 723–735, 2001.

[34] C. Poupon, J. Mangin, C. Clark, V. Frouin, J. Rgis, D. L. Bihan, and I. Bloch, "Towards inference of human brain connectivity from MR diffusion tensor data," *Med. Image Anal.*, vol. 5, no. 1, pp. 1–15, 2001.

[35] S. Mori and P. van Zijl, "Fiber tracking: Principles and strategies—A technical review," *NMR Biomed.*, vol. 15, no. 7–8, pp. 468–480, 2002.

[36] V. Wedeen, T. Reese, D. Tuch, M. Weigel, J. Dou, R. Weiskoff, and D. Chessler, "Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI," in *Proc. 8th Sci. Meeting Int. Soc. Magnetic Resonance Medicine*, Denver, CO, April 2000, p. 82.

[37] M. Wiegell, H. Larsson, and V. Wedeen, "Fiber crossing in human brain depicted with diffusion tensor MR imaging," *Radiology*, vol. 217, no. 3, pp. 897–903, 2000.

[38] C. Pierpaoli, A. Barnett, S. Pajevic, R. Chen, L. Penix, A. Virta, and P. Basser, "Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture," *NeuroImage*, vol. 13, no. 6, pp. 1174–1185, 2001.

[39] D. Alexander, G. Barker, and S. Arridge, "Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data," *Magn. Reson. Med.*, vol. 48, no. 2, pp. 331–340, 2002.

[40] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Rivire, A. Constantinesco, D. L. Bihan, and J.-F. Mangin, "Fiber tracking in q-ball fields using regularized particle trajectories," in *Proc. 19th Int. Conf. Inf. Process. Med. Imag.*, Glenwood Springs, CO, Jul. 2005, pp. 52–63.

[41] M. Chen, Q. Shao, and J. Ibrahim, *Monte Carlo Methods in Bayesian Computation*. New York: Springer Verlag, 2000.

[42] C. Beaulieu, "The basis of anisotropic water diffusion in the nervous system—A technical review," *NMR Biomed.*, vol. 15, no. 7–8, pp. 435–455, 2002.

[43] M. Lazar and A. Alexander, "An error analysis of white matter tractography methods: Synthetic diffusion tensor field simulations," *NeuroImage*, vol. 20, no. 2, pp. 1140–1153, 2003.