

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/262840989>

# Fiber Direction Estimation in Diffusion MRI

Article · June 2014

Source: arXiv

---

CITATIONS

0

READS

645

5 authors, including:



Debashis Paul

University of California, Davis

58 PUBLICATIONS 3,057 CITATIONS

[SEE PROFILE](#)



Jie Peng

University of Science and Technology of China

97 PUBLICATIONS 2,812 CITATIONS

[SEE PROFILE](#)

# Fiber Direction Estimation in Diffusion MRI

Raymond K. W. Wong\* Thomas C. M. Lee\* Debashis Paul\* Jie Peng\*

for the Alzheimer's Disease Neuroimaging Initiative†

*\*Department of Statistics, University of California, Davis*

June 1, 2014

## Abstract

Diffusion magnetic resonance imaging is a medical imaging technology to probe anatomical architectures of biological samples in an in vivo and non-invasive manner through measuring water diffusion. It is widely used to reconstruct white matter fiber tracts in brains. This can be done in several steps. Typically, the first step is to estimate the diffusion direction(s) for each voxel of the biological sample under study by extracting the leading eigenvector from the estimated diffusion tensor at each voxel. As it is reasonable to assume that the diffusion directions from neighboring voxels are similar, a local smoothing may be applied to the estimated tensors or directions to improve the estimation of diffusion directions. Finally, a tracking algorithm is used to reconstruct fiber tracts based on (estimated) diffusion directions.

Most commonly used tensor estimation methods assume a single tensor and do not work well when there are multiple principal diffusion directions within a single voxel. The first contribution of this paper is the proposal of a new method which is able to identify and estimate

---

\*Department of Statistics, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA. Email: {rkwwong, tcmllee, debpaul, jiepeng}@ucdavis.edu

†Data used in preparation of this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([www.loni.ucla.edu/ADNI](http://www.loni.ucla.edu/ADNI)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://loni.ucla.edu//ADNI//Collaboration//ADNI\\_Authorship\\_list.pdf](http://loni.ucla.edu//ADNI//Collaboration//ADNI_Authorship_list.pdf)

multiple diffusion directions within a voxel. This method is based on a new parametrization of the multi-tensor model and it produces reliable results even when there are multiple principal diffusion directions within the voxels. As a second contribution, this paper proposes a novel direction smoothing method which greatly improves diffusion direction estimation in regions with crossing fibers. This smoothing method is shown to have excellent theoretical and empirical properties. Lastly this paper develops a novel fiber tracking algorithm which takes (estimated) diffusion directions as input and accommodates multiple directions within a voxel. The overall methodology is illustrated with data sets collected for the study of Alzheimer’s disease.

**Keywords:** DTI, direction smoothing, multi-tensor model, fiber tracking, tractography

## 1 Introduction

Diffusion magnetic resonance imaging (dMRI) is a medical imaging technology that uses magnetic field gradients to measure water diffusion on a three-dimensional (3D) grid of biological tissue along a set of predetermined directions (Bammer *et al.*, 2009; Beaulieu, 2002; Chanraud *et al.*, 2010; Mukherjee *et al.*, 2008). In biological tissues, water diffusion is anisotropic due to the presence of fiber bundles with coherent orientations and thus anatomical structures can be deduced from the diffusion characteristics of water. Due to its in vivo and non-invasive nature, dMRI has been widely applied to delineate the white matter fiber tracts in human brain. Mapping white matter fiber tracts is of great importance in the study of neuronal connectivity and understanding of brain functionality (Mori, 2007; Sporns, 2011).

Water diffusion in any location in the brain is often modeled as a 3D Gaussian process. At each voxel, diffusion is described by a  $3 \times 3$  positive definite matrix, which is referred to as a diffusion tensor; see Mori (2007) for an introduction to diffusion tensor imaging (DTI) techniques. One then extracts the direction information from the estimated diffusion tensor (e.g., the principal eigenvector) at each voxel and reconstructs the white matter fiber tracts by computer aided tracking algorithms via a process named tractography (Basser *et al.*, 2000).

However, DTI cannot resolve multiple fiber populations with distinct orientations (i.e., crossing

fibers) within a voxel since a tensor only has one principal direction. In crossing fiber regions, estimated diffusion tensors may lead to low anisotropy estimation or oblate tensor estimation. Poor tensor estimation results in poor direction estimation which affects fiber reconstruction, e.g., early termination of the fiber tracking or biased fiber tracking.

In order to resolve intravoxel orientational heterogeneity, several approaches have been proposed. Tuch *et al.* (2002) propose a multi-tensor model which assumes a finite number of homogeneous fiber directions with a voxel and Gaussian diffusion along each direction. However, it has been shown that the parameters in the multi-tensor model are not identifiable (Scherrer and Warfield, 2010). Nonparametric methods such as Q-ball and Q-space imaging have been proposed (Tuch, 2004; Descoteaux *et al.*, 2007). However such methods rely on high angular resolution diffusion imaging (HARDI) (Tuch *et al.*, 2002; Hosey *et al.*, 2005) where a large number of gradients is sampled (e.g., a few hundreds). Most currently available data sets, and particularly those obtained under clinical settings, have much less number of gradient directions (a few tens at most), rendering such methods not applicable.

The primary goal of this paper is to develop a new method for fiber detection and tracking that works exceptionally well in the presence of crossing fibers. Our method is completely automatic and improves existing methods in several aspects. Loosely, the method can be divided into the following three major steps.

In the first step, we estimate the tensor directions within each voxel under a multi-tensor model. We propose a new parametrization which makes the tensor directions identifiable. We develop an efficient and numerically stable computational procedure to obtain the global MLE of the tensor directions.

Once the tensor direction estimates are obtained for all individual voxels, in the second step, a direction smoothing procedure is applied to further improve the diffusion direction estimates by borrowing information from neighboring voxels. A distinctive and unique feature of this new procedure is that it handles crossing fibers through the clustering of directions into homogeneous

groups. We note that, although various tensor smoothing methods have been proposed (e.g., Pennec *et al.*, 2006; Arsigny *et al.*, 2006; Fillard *et al.*, 2007; Fletcher and Joshi, 2007; Yuan *et al.*, 2012; Carmichael *et al.*, 2013), to the best of our knowledge, little work in the literature on direct diffusion direction smoothing. Since diffusion directions rather than tensors are used as input for tracking algorithms, methods on direction estimation and smoothing should be more efficient in terms of fiber tracking.

In the last step, a fiber tracking algorithm is applied to reconstruct fiber tracts through (smoothed) diffusion direction estimates. Our tracking algorithm is designed to explicitly allow for multiple directions within a voxel.

It is shown by extensive numerical studies that the proposed procedure is effective in direction estimation as well as fiber tracts reconstruction.

The rest of the paper is organized as follows. Section 2 provides background material for some common tensor models. The proposed methods for tensor direction estimation, smoothing of estimated directions, and fiber tracking are presented in, respectively, Sections 3, 4 and 5. Theoretical support for the direction smoothing method are presented in Section 6. The empirical performance of the overall methodology is illustrated with numerical experiments in Section 7 and with a real data set in Section 8. Section 9 provides some concluding remarks, while additional results and technical details are collected in an online supplementary.

## 2 Tensor models

Suppose dMRI measurements are made on  $N$  voxels on a three dimensional grid representing a brain. For each voxel, we have measurements of diffusion weighted signals (complex numbers) along a fixed set (i.e., the same for all voxels) of unit-norm gradient vectors  $\mathcal{U} = \{\mathbf{u}_i : i = 1, \dots, m\}$ .

By assuming Gaussian additive noise on both real and imaginary parts of the signal, the observed signal intensity can be modeled as

$$S(\mathbf{s}, \mathbf{u}) = \|\bar{S}(\mathbf{s}, \mathbf{u})\phi(\mathbf{s}, \mathbf{u}) + \sigma\epsilon(\mathbf{s}, \mathbf{u})\|,$$

where  $\bar{S}(\mathbf{s}, \mathbf{u})$  is the intensity of the noiseless signal,  $\phi(\mathbf{s}, \mathbf{u})$  is a unit vector in  $\mathbb{R}^2$  representing the phase of the signal,  $\epsilon(\mathbf{s}, \mathbf{u})$  is the noise random variable following  $\mathcal{N}_2(\mathbf{0}, \mathbf{I}_2)$  and  $\sigma > 0$  denotes the noise level. The observed signal intensity then follows a Rician distribution (Gudbjartsson and Patz, 1995):

$$S(\mathbf{s}, \mathbf{u}) \sim \text{Rician}(\bar{S}(\mathbf{s}, \mathbf{u}), \sigma).$$

Moreover, we assume the noise  $\epsilon(\mathbf{s}, \mathbf{u})$ 's are independent across different voxels and gradient directions. We write the set of measurements as  $\{S(\mathbf{s}, \mathbf{u}) : \mathbf{u} \in \mathcal{U}\}$ , where  $\mathbf{s}$  is the three dimensional coordinate of the center of this voxel.

Assuming Gaussian diffusion, the noiseless signal intensity is given by (e.g., Mori, 2007)

$$\bar{S}(\mathbf{s}, \mathbf{u}) = S_0(\mathbf{s}) \exp \{-b\mathbf{u}^\top \mathbf{D}(\mathbf{s})\mathbf{u}\},$$

where  $S_0(\mathbf{s})$  is the non-diffusion-weighted intensity,  $b > 0$  is an experimental constant referred to as the  $b$ -value and  $\mathbf{D}(\mathbf{s})$  is a  $3 \times 3$  covariance matrix referred to as the diffusion tensor. This model is called the single tensor model and suits for the case of at most one dominant diffusion direction within a voxel. To indicate the degree of anisotropy of the diffusion, one commonly used measure is the fractional anisotropy (FA),

$$FA = \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}, \quad (1)$$

where  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  are the eigenvalues of  $\mathbf{D}$ . FA value lies between zero and one and the larger it is, the more anisotropic the water diffusion is at the corresponding voxel.

Although the single tensor model is probably the most widely used tensor model in practice (implemented by most softwares for DTI), it is not suitable for crossing fiber regions. To deal with crossing fibers, this model has been extended to a multi-tensor model (e.g., Tuch, 2002; Behrens *et al.*, 2003, 2007; Tabelow *et al.*, 2012):

$$\bar{S}(\mathbf{s}, \mathbf{u}) = S_0(\mathbf{s}) \sum_{j=1}^{J(\mathbf{s})} p_j(\mathbf{s}) \exp \{-b\mathbf{u}^\top \mathbf{D}_j(\mathbf{s})\mathbf{u}\}, \quad (2)$$

where  $\sum_{j=1}^{J(\mathbf{s})} p_j(\mathbf{s}) = 1$  and  $p_j(\mathbf{s}) > 0$  for  $j = 1, \dots, J(\mathbf{s})$ . Here  $J(\mathbf{s})$  represents the number of fiber populations and  $p_j(\mathbf{s})$ 's denote weights of the corresponding fibers.

### 3 Voxel-wise estimation of diffusion directions

One important goal of DTI studies is to estimate principal diffusion directions, referred to as diffusion directions hereafter, at each voxel. They may be interpreted as tangent directions along fiber bundles at the corresponding voxel. The estimated diffusion directions are then used as an input for tractography algorithms to reconstruct fiber tracts. This section explores the diffusion direction estimation within a single voxel. For notational simplicity, dependence on voxel index  $\mathbf{s}$  is temporarily dropped. Moreover, for ease of exposition, we assume that  $\sigma$  and  $S_0(\mathbf{s})$  are known and delay the discussion of their estimation to Section 8.

Under the single tensor model, various methods for tensor estimation have been proposed including linear regression, nonlinear regression and ML estimation; e.g., see Carmichael *et al.* (2013) for a comprehensive review. Then diffusion directions are derived as principal eigenvectors of (estimated) diffusion tensors. However, for the multi-tensor models, severe computational issues have been observed and additional prior information and assumptions are imposed to tackle these issues. For instance, Behrens *et al.* (2003, 2007) use shrinkage priors and Tabelow *et al.* (2012) assume all tensors to be axially symmetric (i.e., the two minor eigenvalues are the same) and have the same set of eigenvalues. Scherrer and Warfield (2010) show that the multi-tensor model is indeed non-identifiable and they suggest to use multiple  $b$ -values in data acquisition to make the model identifiable. However, due to practical limitations, most of the current dMRI studies are obtained under a fixed  $b$ -value and so render their suggestion inapplicable. Below we show that the identifiability issue does not prevent one from estimating the diffusion directions and so neither strong assumptions nor special experimental settings are necessary if one is only interested in diffusion directions rather than the diffusion tensors themselves.

#### 3.1 Identifiability of multi-tensor model

From Scherrer and Warfield (2010), model (2) can be re-written as

$$\bar{S}(\mathbf{u}) = S_0 \sum_{j=1}^J p_j a_j \exp \left\{ -b \mathbf{u}^\top \left( \mathbf{D}_j + \frac{\log a_j}{b} \mathbf{I}_3 \right) \mathbf{u} \right\},$$

where  $a_j > 0$  for  $j = 1, \dots, J$  such that  $p_j a_j > 0$ ,  $\mathbf{D}_j + (\log a_j/b)\mathbf{I}_3$  is positive definite and  $\sum_{j=1}^J p_j a_j = 1$ . When  $J = 2$ , one can easily derive the explicit conditions for  $a_j$  to fulfill these criteria, and see that there are infinite sets of such  $a_j$ 's. However, note that  $\mathbf{D}_j + (\log a_j/b)\mathbf{I}_3$  shares the same set of eigenvectors with  $\mathbf{D}_j$ . Thus, one may still be able to estimate diffusion directions, which correspond to the major eigenvectors of the tensors. This motivates us to consider estimating diffusion directions directly instead of the tensors themselves.

Now we assume that  $\mathbf{D}_j$ 's are axially symmetric; that is, the two minor eigenvalues of  $\mathbf{D}_j$  are equal. This is a common assumption (Basser *et al.*, 1994) for modeling dMRI data and it implies that diffusion is symmetric around the principal diffusion direction (here, the principal eigenvector) (Tournier *et al.*, 2004, 2007). By not differentiating the two minor eigenvectors, we obtain a clear meaning of diffusion direction. In addition, this reduces the number of unknown parameters by one and thus facilities estimation. In the following, we propose a new parametrization of the multi-tensor model which is identifiable and thus can be used for direction estimation.

Write  $\mathcal{M}$  as the space of the unit principal eigenvector, i.e., the three dimensional unit sphere with equivalence relation  $\mathbf{m} \sim -\mathbf{m}$ . Let  $\alpha_j \geq 0$ ,  $\xi_j > 0$  and  $\mathbf{m}_j \in \mathcal{M}$  be the difference between the larger and smaller eigenvalue, smaller eigenvalue and the standardized principal eigenvector of  $\mathbf{D}_j$ , respectively. Since  $\mathbf{D}_j = \alpha_j \mathbf{m}_j \mathbf{m}_j^\top + \xi_j \mathbf{I}_3$ , model (2) becomes

$$\begin{aligned}\bar{S}(\mathbf{u}) &= S_0 \sum_{j=1}^J p_j \exp \left\{ -b \mathbf{u}^\top \left( \alpha_j \mathbf{m}_j \mathbf{m}_j^\top + \xi_j \mathbf{I}_3 \right) \mathbf{u} \right\} \\ &= S_0 \sum_{j=1}^J p_j \exp(-b\xi_j) \exp \left\{ -b\alpha_j (\mathbf{u}^\top \mathbf{m}_j)^2 \right\} \\ &= S_0 \sum_{j=1}^J \tau_j \exp \left\{ -b\alpha_j (\mathbf{u}^\top \mathbf{m}_j)^2 \right\},\end{aligned}\tag{3}$$

where  $\tau_j = p_j \exp(-b\xi_j) \in (0, 1)$ . From the above, one can see that  $p_j$  and  $\xi_j$  are not simultaneously identifiable, so we cannot estimate the tensors. However, the new parametrization  $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_J^\top)^\top$  is identifiable, where  $\boldsymbol{\gamma}_j = (\tau_j, \alpha_j, \mathbf{m}_j^\top)^\top$  for  $j = 1, \dots, J$ , so that we can estimate the principal diffusion directions  $\mathbf{m}_j$ 's.

### 3.2 A fast and stable algorithm for ML estimation

To estimate the parameters in model (3), we start by investigating the standard ML estimation.

Under the Rician noise assumption, the log-likelihood of  $\gamma$  is:

$$\begin{aligned} l(\gamma) &= \sum_{\mathbf{u} \in \mathcal{U}} \log \left[ \frac{S(\mathbf{u})}{\sigma^2} \exp \left\{ -\frac{S^2(\mathbf{u}) + \bar{S}^2(\mathbf{u})}{2\sigma^2} \right\} I_0 \left\{ \frac{S(\mathbf{u})\bar{S}(\mathbf{u})}{\sigma^2} \right\} \right] \\ &= \sum_{\mathbf{u} \in \mathcal{U}} \left[ \log \left\{ \frac{S(\mathbf{u})}{\sigma^2} \right\} - \frac{S^2(\mathbf{u}) + \bar{S}^2(\mathbf{u})}{2\sigma^2} + \log I_0 \left\{ \frac{S(\mathbf{u})\bar{S}(\mathbf{u})}{\sigma^2} \right\} \right], \end{aligned} \quad (4)$$

where  $I_0(x) = \int_0^\pi \exp(x \cos \phi) d\phi / \pi$  is the zeroth order modified Bessel function of the first kind (Abramowitz and Stegun, 1964). The ML estimate is obtained through maximizing the log-likelihood function (4). Although the new parametrization avoids the identifiability issue, the likelihood function usually has multiple local maxima, which makes the computation of ML estimate difficult and unstable. Next we discuss a strategy to tackle this issue.

In attempt to find the global maximizer, we develop an efficient algorithm through an approximation of model (3). This algorithm essentially performs a grid search, but it makes use of the geometry of the problem so it is quite fast. It includes three major steps: (i) lay down a grid for  $(\alpha_j, \mathbf{m}_j^\top)$ 's, (ii) evaluate the likelihood function on the grid, and (iii) return the grid point that maximizes the likelihood function. One can then use this returned grid point as a starting value in a gradient method for obtaining ML estimation of Model (3). Such a strategy results in better numerical stability and accuracy in finding ML estimates.

#### 3.2.1 An approximation of model (3)

Let  $\mathbf{c}_j = (\alpha_j, \mathbf{m}_j^\top)^\top$ ,  $\mathbf{c} = (\mathbf{c}_1^\top, \dots, \mathbf{c}_J^\top)^\top$  and  $\mathcal{C}_j$  be the set of grid points for  $\mathbf{c}_j$ . For simplicity, we take the same set of grid points,  $\mathcal{C}$ , for all  $j$ . To lay down a grid for  $\mathbf{m}_j$ 's, we apply the sphere tessellation using Icosahedron, which is depicted in Figure 1. Here, we only pick unique vertices up to a sign for the formation of the grid. In our implementation, we utilize randomly rotated versions of the tessellation with two subdivisions, which results in a grid with 321 directions. If

$\mathbf{c} \in \prod_{j=1}^J \mathcal{C}_j = \mathcal{C}^J$ , model (3) can be rewritten as

$$\bar{S}(\mathbf{u}) = \sum_{k=1}^K \tilde{\beta}_k x(\mathbf{u}, \tilde{\mathbf{m}}_k, \tilde{\alpha}_k), \quad (5)$$

where  $K = |\mathcal{C}|$ ,  $x(\mathbf{u}, \tilde{\mathbf{m}}_k, \tilde{\alpha}_k) = S_0 \exp\{-b\tilde{\alpha}_k(\mathbf{u}^\top \tilde{\mathbf{m}}_k)^2\}$ ,  $(\tilde{\alpha}_k, \tilde{\mathbf{m}}_k) \in \mathcal{C}$  and  $\tilde{\beta}_k \in [0, 1]$ . One may notice that, in this reformulation, the non-zero  $\tilde{\beta}_k$ 's have an one-to-one correspondence with  $\tau_j$ 's in model (3). If  $\mathbf{c} \notin \mathcal{C}^J$ , i.e. the set of parameters is not a grid point, then equation (5) serves as an approximation to  $\bar{S}(\mathbf{u})$  in model (3) as long as the grid is dense enough in the parameter space.

Furthermore, under the commonly used scales of  $b$ -values and tensors,  $x(\mathbf{u}, \tilde{\mathbf{m}}_k, \tilde{\alpha}_k)$  and  $x(\mathbf{u}, \tilde{\mathbf{m}}_{k'}, \tilde{\alpha}_{k'})$  are highly correlated if  $\tilde{\mathbf{m}}_k = \tilde{\mathbf{m}}_{k'}$ . Inspired by this observation, we reduce the grid size by setting  $\tilde{\alpha}_k = \tilde{\alpha}$  for all  $k$  to a common value  $\tilde{\alpha}$ . From our experience, we set  $\tilde{\alpha} = 2/b$ . With all these approximations, we consider fitting the following model:

$$\bar{S}(u) = \sum_{k=1}^K \beta_k x_k(\mathbf{u}), \quad (6)$$

where  $x_k(\mathbf{u}) = x(\mathbf{u}, \tilde{\mathbf{m}}_k, \tilde{\alpha})$  and  $\beta_k \geq 0$ . For our purpose, we want to identify nonzero  $\beta_k$ 's because those  $\tilde{\mathbf{m}}_k$ 's associated with non-zero  $\hat{\beta}_k$ 's can be regarded as selected diffusion directions. Note that model (6) converts the expensive grid search to an estimation problem of a linear model (with respect to  $\beta_k$ 's) with non-negative constraints. A fast algorithm for fitting this model with Rician noise assumption is given in Section S1 of the supplementary material (SM). As it turns out, the non-negativity constraints often result in a sparse estimate of  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)^\top$ ; i.e., only a subset of directions is selected. In particular, if the estimate of the unconstrained problem (i.e.,  $\beta_k$ 's are allowed to be negative) is not located in the first quadrant of the parameter space, the corresponding constrained solution will be sparse.

Even though the solution is often sparse, the number of selected directions is usually larger than  $J$ , the true number of tensor components. This is partly due to colinearity of  $x_k(\mathbf{u})$ 's resulting from the use of a dense grid on the directions  $\tilde{\mathbf{m}}_k$ 's.

In the following, we propose to first divide the selected directions into  $I$  groups and then generate stable estimates of  $\mathbf{m}_j$ 's via gradient methods (Section 3.2.2). Finally, Bayesian information

criterion (BIC) (Schwarz, 1978) is used to choose an appropriate  $I$  as the estimate for  $J$  (Section 3.3).

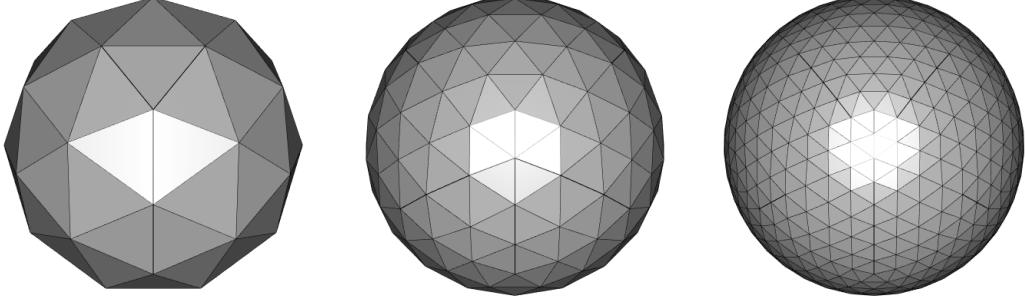


Figure 1: Sphere tessellations through triangulation using Icosahedron with level of subdivisions: 0 (Left), 1 (Middle) and 2 (Right).

### 3.2.2 Clustering of the selected directions

Write the above ML estimate of  $\beta_k$  as  $\hat{\beta}_k$  for  $k = 1, \dots, K$ . Suppose there are  $L > 0$  nonzero  $\hat{\beta}_k$ 's, without loss of generality,  $k = 1, \dots, L$ . Thus,  $\tilde{\mathbf{m}}_1, \dots, \tilde{\mathbf{m}}_L$  are the selected directions. Now, we develop a strategy to cluster the selected directions into  $I$  groups, for a set of  $I \in \{1, \dots, L\}$ . To perform clustering, we require a metric measure on the space of directions  $\mathcal{M}$ . A natural metric is

$$d^*(\mathbf{u}, \mathbf{v}) = \arccos(|\mathbf{u}^\top \mathbf{v}|), \quad (7)$$

where  $\mathbf{u}, \mathbf{v} \in \mathcal{M}$ . Note that,  $d^*(\mathbf{u}, \mathbf{v})$  is the acute angle between  $\mathbf{u}$  and  $\mathbf{v}$ . With this distance metric, one can define dissimilarity matrix for a set of directions and make use of a generic clustering algorithm. Our choice is the Partition Around Medoids (PAM) (Kaufman and Rousseeuw, 1990) due to its simplicity. The detailed procedure is described in Algorithm S1 in the SM, where the input vectors are the selected directions. Due to the sparsity of  $\hat{\beta}_j$ 's and efficient algorithms of PAM, this clustering strategy is practically fast. Let  $\check{\mathbf{m}}_1, \dots, \check{\mathbf{m}}_I$  be the resulting group (Karcher) means. They are used as the starting value for gradient-based methods, such as L-BFGS-B algorithm (Byrd *et al.*, 1995), for obtaining  $\hat{\gamma}(I)$ , the ML estimate of  $\gamma$  under model (3) with  $I$  tensor components. More specifically, the starting value is set as  $((1/I, \tilde{\alpha}, \check{\mathbf{m}}_1^\top), \dots, (1/I, \tilde{\alpha}, \check{\mathbf{m}}_I^\top))^\top$ .

### 3.3 Selection of the number of tensor components $J$

We use BIC to select the number of components  $J$ . Under model (3), the BIC for a model with  $I$  components is

$$\text{BIC}(I) = -2l(\hat{\gamma}(I)) + 4I \log(m), \quad (8)$$

where  $m$  is the number of gradient directions. Then  $J$  is chosen as  $\hat{J} = \operatorname{argmin}_{I \in \{1, \dots, \tilde{I}\}} \text{BIC}(I)$ , where  $\tilde{I}$  is a pre-specified upper bound for the number of components. Based on our experience,  $\tilde{I} = 4$  is reasonable choice. If  $\tilde{I} > L$ , we simply compare  $\text{BIC}(1), \dots, \text{BIC}(L)$ .

In practice, there are voxels with no major diffusion directions. Under single tensor model, the corresponding diffusion tensor is isotropic, i.e., all three eigenvalues are equal.

In the case of isotropic tensor, (3) reduces to  $\bar{S}(\mathbf{u}) = \mathbf{S}_0\tau_1$ . Thus there is only one parameter  $\tau_1$ . We write the corresponding likelihood function as  $\tilde{l}$  and denote the ML estimate of  $\tau_1$  by  $\hat{\tau}_1$ , which can be obtained by a generic gradient method. The corresponding BIC criterion is

$$\text{BIC}(0) = -2\tilde{l}(\hat{\tau}_1) + \log(m),$$

where 0 represents no diffusion direction. Combined with the previous BIC formulation (8), one has a comprehensive model selection rule, which handles voxels with from zero to up to  $\tilde{I}$  (here 4) fiber populations. In practice, we follow the convention and use FA (1) (see, e.g., Mori, 2007) as a first step screening; i.e., we first remove voxels with small FA values and then apply the BIC approach over those suspected anisotropic voxels.

We summarize our voxel-wise estimation procedure in Algorithm S2 in the SM. A simulation study is conducted and the corresponding results are presented in Section S2 of the SM. These numerical results suggest that our voxel-wise estimation procedure provides extremely stable and reliable results under various model settings.

## 4 Spatial smoothing of diffusion directions

Although model (3) provides a better modeling than single tensor model for crossing fiber regions, it also leads to an increase in the number of parameters and thus the variability of the estimates. To further improve estimation, we consider borrowing information from neighboring voxels and develop a novel smoothing technique for diffusion directions.

Tensor smoothing has been widely studied in the literature (Arsigny *et al.*, 2006; Pennec *et al.*, 2006; Yuan *et al.*, 2012; Carmichael *et al.*, 2013). However, as discussed earlier, tensors are not identifiable in the multi-tensor model without additional assumptions. Moreover, if the ultimate goal is the reconstruction of fiber tracts, a good estimate of diffusion directions suffices. This motivates the proposal of the new direction smoothing method below.

We shall assume that tangent directions of fiber bundles change smoothly. This leads to the spatial smoothness of diffusion directions that belong to the same fiber bundle. In many brain regions, it is reasonable to model the fiber tracts as smooth curves at the resolution of voxels in dMRI ( $\sim 2\text{mm}$ ).

However, there is one major challenge in diffusion direction smoothing. The smoothness assumption is only reasonable along the same fiber bundle. In regions with crossing fibers, the diffusion directions may belong to several different fibers which contribute to a mixture of populations of diffusion directions. To circumvent this issue, we propose to first cluster the diffusion directions within a neighborhood into separate homogeneous populations (Section 4.2) and then apply direction smoothing within each population (Section 4.1).

### 4.1 Smoothing along a single fiber

This subsection assumes that there is only one homogeneous population of diffusion directions, which corresponds to a single fiber bundle. Let  $\{\{\hat{\mathbf{m}}_j(\mathbf{s}) : j = 1, \dots, \hat{J}(\mathbf{s})\} : \forall \mathbf{s}\}$  be the estimated diffusion directions obtained from the above voxel-wise estimation procedure in Section 3. Further, write  $T = \sum_{\mathbf{s}} \hat{J}(\mathbf{s})$  and, by re-indexing,  $\{\hat{\mathbf{m}}_k : k = 1, \dots, T\} = \{\{\hat{\mathbf{m}}(\mathbf{s}) : j = 1, \dots, \hat{J}(\mathbf{s})\} : \forall \mathbf{s}\}$ .

Also write  $\mathbf{s}_k$  as the corresponding voxel location associated with  $\hat{\mathbf{m}}_k$ . Following the idea of kernel smoothing on Euclidean space (Fan and Gijbels, 1996), the smoothing estimate at voxel  $\mathbf{s}_0$  is defined as a weighted Karcher mean of the neighboring direction vectors:

$$\arg \min_{\mathbf{v} \in \mathcal{M}} \sum_{i=1}^T w_i d^{*2}(\hat{\mathbf{m}}_i, \mathbf{v}), \quad (9)$$

where  $w_i = K_{\mathbf{H}}(\mathbf{s}_i - \mathbf{s}_0)$ 's are spatial weights and the metric  $d^*$  is defined in (7). These weights place more emphasis on spatially closer observations. Here  $K_{\mathbf{H}}(\cdot) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2} \cdot)$  with  $K(\cdot)$  as a three dimensional kernel function satisfying  $\int K(\mathbf{s}) d\mathbf{s} = 1$ , and  $\mathbf{H}$  is a  $3 \times 3$  bandwidth matrix. In our numerical work, we choose  $K$  as the standard Gaussian density, and set  $\mathbf{H} = h\mathbf{I}_3$ , where  $h$  is chosen automatically using the cross-validation approach described in Section S3 of the SM.

## 4.2 Smoothing over multiple fibers

As discussed earlier, the spatial smoothness assumption does not hold in a voxel  $\mathbf{s}_0$  with crossing fibers. To tackle this issue, we first apply clustering to estimated directions within a neighborhood of  $\mathbf{s}_0$  in an attempt to separate the direction vectors corresponding to different fiber populations into different clusters. Then we apply the smoothing procedure in the previous subsection within each direction cluster. This subsection describes this procedure in details.

First we define neighboring voxels for  $\mathbf{s}_0$ . We begin with computing the spatial weights defined in Section 4.1. We then remove those voxels with weights smaller than a threshold. By filtering out these voxels, we obtain tighter and better separated clusters of directions. Moreover, such voxels have little effects on smoothing due to their small weights. The artificial data set displayed in Figure 2 provides an illustrative example. Every black dot in the left panel represents an estimated direction (from the center of the sphere). In the middle panel, the size of each dot is proportional to its spatial weight in equation (9). Lastly, the right panel shows all dots with spatial weights larger than a threshold. Notice that such a trimming operation leads to two obvious clusters of directions, which makes the subsequent task of clustering the estimated directions much easier.

Next we utilize the same clustering strategy developed in Section 3.3 to choose the number of

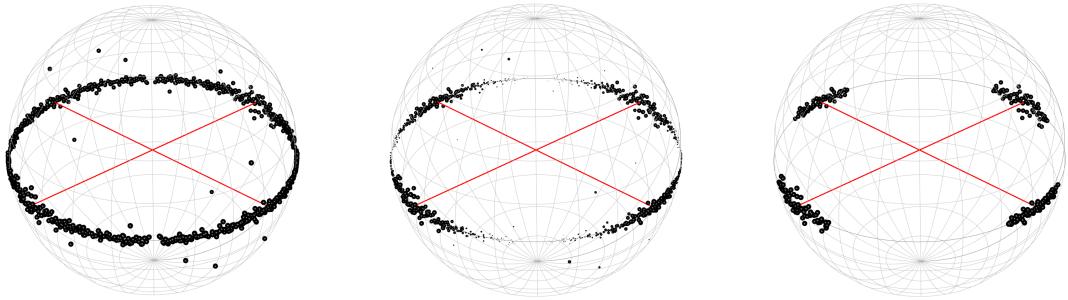


Figure 2: Finding all the neighboring voxels for separating crossing fiber directions. Left: all estimated directions. Middle: sizes of all estimated directions proportional to weights. Right: estimated directions with weights larger than a threshold. Red lines represents underlying true directions.

clusters adaptively by the average silhouette (Rousseeuw, 1987); see Algorithm S3 of the SM. The silhouette of a datum  $i$  measures the strength of its membership to its cluster, as compared to the neighboring cluster. Here, the neighboring cluster is the one, apart from cluster of datum  $i$ , that has the smallest average dissimilarity with datum  $i$ . The corresponding silhouette is defined as  $(b_i - a_i)/(\max\{a_i, b_i\})$ , where  $a_i$  and  $b_i$  represent the average dissimilarities of datum  $i$  with all other data in the same cluster and that with the neighboring cluster respectively. The average silhouette of all data gives a measure of how good the clustering is. Thus we select the number of clusters via maximizing the average silhouette.

The detailed smoothing procedure is given in Algorithm 1.

## 5 Fiber tracking

For dMRI, fiber tractography can be divided into deterministic and probabilistic methods. Deterministic methods (e.g. Mori *et al.*, 1999; Weinstein *et al.*, 1999; Mori and van Zijl, 2002) track fiber bundles by utilizing the principal eigenvectors of tensors. Probabilistic methods (e.g. Koch *et al.*, 2002; Parker and Alexander, 2003; Friman *et al.*, 2006) use the probability density of diffusion ori-

---

**Algorithm 1** Algorithm for direction smoothing

---

**Input:** Target voxel  $\mathbf{s}^*$ , voxel-wise estimate  $\{(\mathbf{s}_k, \hat{\mathbf{m}}_k), k = 1, \dots, T\}$ , estimated number of fibers  $\{\hat{J}(\mathbf{s}) : \mathbf{s} \in \mathcal{S}\}$ , kernel function  $K$ , bandwidth matrix  $\mathbf{H}$ , threshold  $c$ , maximum number of cluster (Algorithm S3 of the SM)  $K$ , angular threshold (Algorithm S3 of the SM)  $\xi$

**Output:** Updated number of directions and updated directions at  $\mathbf{s}^*$

**Description:** To perform smoothing for diffusion directions at  $\mathbf{s}^*$

- 1: **for**  $k = 1$  to  $T$  **do** Compute spatial weight:  $w_k \leftarrow K_{\mathbf{H}}(\mathbf{s}_k - \mathbf{s}^*)$
  - 2: **for**  $k = 1$  to  $T$  **do** Standardize spatial weights:  $w_k \leftarrow w_k / \sum_{j=1}^T w_j$
  - 3: Sort  $w_k$ 's in decreasing order such that  $w_{l_1} \geq \dots \geq w_{l_T}$
  - 4: Identify neighborhood for clustering (Section 4.2):  
Compute  $L \leftarrow \min_{M \in \{1, \dots, T\}} \mathbb{1}\{\sum_{m=M+1}^T w_{l_m} \leq c\}$  (The summation  $\sum_{m=T+1}^T w_{l_m}$  is defined as 0.)
  - 5: Clustering via Algorithm S3 (SM):  $(\{\mathbf{u}_1, \dots, \mathbf{u}_C\}, C) \leftarrow \text{CLUSTDIRN}(\{\hat{\mathbf{m}}_{l_1}, \dots, \hat{\mathbf{m}}_{l_L}\}, K, \xi)$
  - 6: **if**  $C \geq \hat{J}(\mathbf{s}^*)$  **then**
  - 7: Match the smoothed directions,  $\{u_1, \dots, u_C\}$ , to the voxel-wise estimates at  $\mathbf{s}^*$ ,  $\{\hat{\mathbf{m}}_1(s^*), \dots, \hat{\mathbf{m}}_{\hat{J}(s^*)}(s^*)\}$ :
$$\left( \hat{k}_1, \dots, \hat{k}_{\hat{J}(\mathbf{s}^*)} \right) \leftarrow \arg \min_{\{k_1, \dots, k_{\hat{J}(\mathbf{s}^*)} \in \{1, \dots, C\} : k_i \neq k_j\}} \sum_{j=1}^{\hat{J}(\mathbf{s}^*)} d^*(\hat{\mathbf{m}}_j(\mathbf{s}^*), \mathbf{u}_{k_j})$$
  - 8: **for**  $j = 1$  to  $\hat{J}(\mathbf{s}^*)$  **do**  $\hat{\mathbf{m}}_j(\mathbf{s}^*) \leftarrow \mathbf{u}_{\hat{k}_j}$
  - 9: **else**
  - 10: Match the voxelwise estimates at  $\mathbf{s}^*$  to the smoothed directions: :
$$\left( \hat{k}_1, \dots, \hat{k}_C \right) \leftarrow \arg \min_{\{k_1, \dots, k_C \in \{1, \dots, \hat{J}(\mathbf{s}^*)\} : k_i \neq k_j\}} \sum_{j=1}^C d^*(\hat{\mathbf{m}}_{k_j}(\mathbf{s}^*), \mathbf{u}_j)$$
  - 11: **for**  $j = 1$  to  $C$  **do**  $\hat{\mathbf{m}}_{\hat{k}_j}(\mathbf{s}^*) \leftarrow \mathbf{u}_j$
  - 12:  $\hat{J}(\mathbf{s}^*) \leftarrow C$  and remove non-updated  $\hat{\mathbf{m}}_j(\mathbf{s}^*)$ 's
- 

entations. Deterministic methods, including the popular Fiber Assignment by Continuous Tracking (FACT) (Mori *et al.*, 1999) and Tensorlines (Weinstein *et al.*, 1999) algorithms, typically require a diffusion tensor field, where there is a single diffusion tensor (either isotropic or anisotropic) associated with each voxel, as an input. In below, we propose a deterministic tracking algorithm which takes diffusion directions (associated with the location information) as input. This algorithm allows for multiple or no principal diffusion directions at a voxel. One advantage of the proposed

algorithm is that it makes use of the directional information from individual fibers at voxel level.

To construct our procedure, we adopt similar tracking ideas from FACT, as depicted in Figure 3 (Left). Tracking starts at the center of a voxel (Voxel 1 in Figure 3) and continues in the direction of the estimated diffusion direction. When it enters the next voxel (Voxel 2 in Figure 3), the track changes its direction to align with the new diffusion direction and so on. The above tracking rule may produce many short and fragmented fiber tracts due to either a wrongfully identified isotropic voxel or spurious directions which go nowhere. In addition, it does not tell us which direction to follow in case there are multiple directions in a voxel, which happens in crossing fiber regions. To address these issues, we modify the above procedure in the following.

Given a current diffusion direction (we refer to the corresponding voxel as the current voxel), the voxel that it points to (we refer to this voxel as the destination voxel) may have (i) at least one direction; (ii) no direction (i.e., isotropic). In case (i), we will first identify the direction with the smallest angular difference with the current direction. If its separation angle is smaller than a pre-specified threshold (e.g.,  $\pi/6$ ), we enter the destination voxel and tracking will go on along this direction. See Figure 3 (Middle). On the other hand, if the separation angle is greater than the threshold, or case (ii) happens, we deem that the destination voxel does not have a viable direction. In this case, tracking will go along the current direction if it finds a viable direction within a pre-specified number of voxels. The number of voxels that are allowed to be skipped is set to be 1 in our numerical illustrations. See Figure 3 (Right). On the other hand, the tracking stops at the current voxel if no viable directions within a pre-specified number of voxels can be found. The detailed tracking algorithm is described in Algorithm S4 in the SM.

As for the choice of starting voxels (known as seeds), there are two common strategies. One can choose seeds based on tracts of interest and starts the tracking from a region of interest (ROI). This approach is based on knowledge on ROI and may not give a full picture of the tracts of interest if there are diverging branches. The other approach is called brute-force approach, where tracking starts from every voxel. It usually leads to a more comprehensive picture of tracts at a higher

computational cost.

The proposed algorithm can be coupled with either strategy. In the brute-force approach, we apply Algorithm S4 of the SM for every pair of  $(\mathbf{s}_k, \hat{\mathbf{v}}_k)$  twice, i.e.,  $(\mathbf{s}_k, \hat{\mathbf{v}}_k)$  and  $(\mathbf{s}_k, -\hat{\mathbf{v}}_k)$ . Due to the continuity of fiber, one would not expect a fiber going to and from nowhere, and only exists within a single voxel. Therefore, if that happens, we remove the corresponding fiber.

The simplicity of the proposed algorithm makes various extensions possible. For instance, we may use weighted average of neighboring directions to produce smoother tracts, similar to Mori and van Zijl (2002).

Combining with the aforementioned smoothing procedure, we call the resulting technique **Diffusion direction Smoothing and Tracking** (DiST).

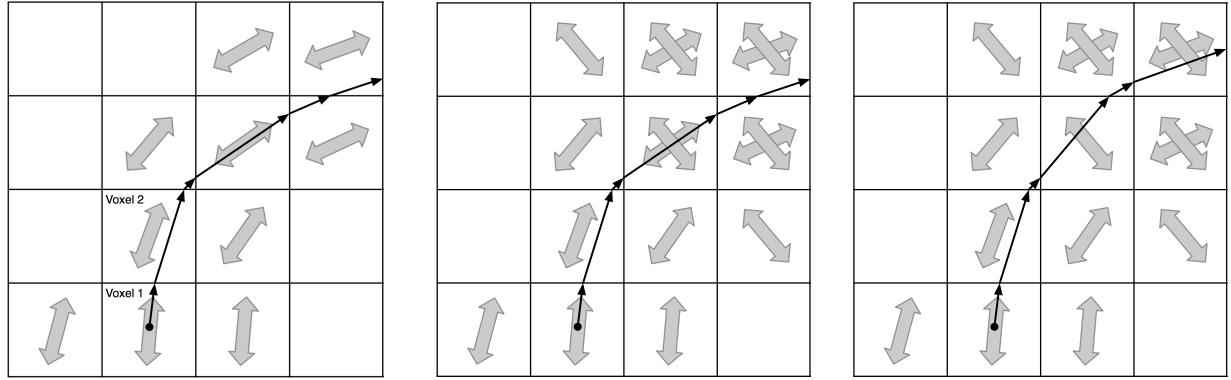


Figure 3: Left: Demonstration of the proposed algorithm in single fiber region. Middle: Demonstration of the proposed algorithm in crossing fiber region. Right: Demonstration of the proposed algorithm in case of absence of viable directions.

## 6 Theoretical results

This section derives some asymptotic properties of the proposed direction smoothing estimator. Note that, since the space of direction vectors has a non-Euclidean geometry and so the theoretical framework is different from that of classical smoothing estimators. Without loss of generality,

suppose we observe  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathcal{M}$  at spatial locations  $\mathbf{s}_1, \dots, \mathbf{s}_n$  respectively. Let  $\mathcal{V}$  be the three dimensional unit sphere. Then  $\mathcal{M}$  is the quotient space of  $\mathcal{V}$  with equivalence relation  $\mathbf{v} \sim -\mathbf{v}$  for any  $\mathbf{v} \in \mathcal{V}$ . This space is also identified with the so-called real projective space  $\mathbb{R}P^2$ .

In the following, we derive our theoretical results under random design where  $\mathbf{s}_i$ 's are independently and identically sampled from a distribution with density  $f_S$ , but our results also apply to regular voxels. Given a spatial location  $\mathbf{s}_0$ , our target is to estimate  $\mathbf{v}_0$ , namely the diffusion direction at  $\mathbf{s}_0$ , in the sense that it minimizes  $\mathbb{E} \{ d^{*2}(\mathbf{V}, \mathbf{v}) | \mathbf{S} = \mathbf{s}_0 \}$ , where  $d^*(\mathbf{u}, \mathbf{v}) = \arccos(|\mathbf{u}^\top \mathbf{v}|)$ . For simplicity, we assume  $\mathbf{s}_i \in \mathbb{R}$  and write it as  $s_i$  thereafter. Thus, our estimator (9) at  $s_0$  can be written as

$$\hat{\mathbf{v}}(s_0) = \arg \min_{\mathbf{v} \in \mathcal{M}} \sum_{i=1}^n K_h(s_i - s_0) d^{*2}(\mathbf{v}_i, \mathbf{v}),$$

where  $K_h(\cdot) = K(\cdot/h)/h$ . Here, with slight notation abuse,  $K(\cdot)$  represents a one dimensional kernel function throughout the theoretical developments.

## 6.1 Working coordinate system

For each  $\mathbf{p} \in \mathcal{V}$ , one can endow a tangent space  $T_{\mathbf{p}}\mathcal{V} = \{\mathbf{v} \in \mathbb{R}^3 : \mathbf{v}^\top \mathbf{p} = 0\}$  with the metric tensor  $g_{\mathbf{p}} : T_{\mathbf{p}}\mathcal{V} \times T_{\mathbf{p}}\mathcal{V} \rightarrow \mathbb{R}$  defined as  $g_{\mathbf{p}}(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_1^\top \mathbf{u}_2$ . Note that the tangent space is identified with  $\mathbb{R}^2$ . The geodesics are great circles and the geodesic distance is  $\arccos(\mathbf{p}_1^\top \mathbf{p}_2)$ , for any  $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{V}$ . The corresponding exponential map at  $\mathbf{p} \in \mathcal{V}$ ,  $\text{Exp}_{\mathbf{p}} : T_{\mathbf{p}}\mathcal{V} \rightarrow \mathcal{V}$ , is given by

$$\text{Exp}_{\mathbf{p}}(\mathbf{0}) = \mathbf{p} \quad \text{and} \quad \text{Exp}_{\mathbf{p}}(\mathbf{u}) = \cos(\|\mathbf{u}\|)\mathbf{p} + \frac{\sin(\|\mathbf{u}\|)}{\|\mathbf{u}\|} \mathbf{u} \quad \text{when } \mathbf{u} \neq \mathbf{0},$$

while the corresponding logarithm map at  $\mathbf{p} \in \mathcal{V}$ ,  $\text{Log}_{\mathbf{p}} : \mathcal{V} \setminus \{-\mathbf{p}\} \rightarrow T_{\mathbf{p}}\mathcal{V}$ , is given by

$$\text{Log}_{\mathbf{p}}(\mathbf{p}) = \mathbf{0} \quad \text{and} \quad \text{Log}_{\mathbf{p}}(\mathbf{v}) = \frac{\arccos(\mathbf{v}^\top \mathbf{p})}{\sqrt{1 - (\mathbf{v}^\top \mathbf{p})^2}} [\mathbf{v} - (\mathbf{v}^\top \mathbf{p})\mathbf{p}] \quad \text{when } \mathbf{v} \neq \mathbf{p}.$$

One can use the exponential map and the logarithm map to define a coordinate system for the  $\mathcal{V} \setminus \{-\mathbf{v}_0\}$  in the following way. Given  $\mathbf{v} \in \mathcal{V}$ , we define the logarithmic coordinate as

$$\omega_1 = \mathbf{e}_1^\top \text{Log}_{\mathbf{v}_0}(\mathbf{v}) \quad \text{and} \quad \omega_2 = \mathbf{e}_2^\top \text{Log}_{\mathbf{v}_0}(\mathbf{v}),$$

where  $\mathbf{e}_1, \mathbf{e}_2 \in T_{\mathbf{v}_0}\mathcal{V}$  and  $\{\mathbf{e}_1, \mathbf{e}_2\}$  forms an orthonormal basis for  $T_{\mathbf{v}_0}\mathcal{V}$ . Write  $\phi(\mathbf{v}) = (\omega_1, \omega_2)^\top$ . In addition, we define

$$\rho_{\mathbf{v}_0}(\mathbf{v}) = \begin{cases} \text{sign}(\mathbf{v}_0^\top \mathbf{v}) \mathbf{v} & \mathbf{v}_0^\top \mathbf{v} \neq 0 \\ \mathbf{v} & \mathbf{v}_0^\top \mathbf{v} = 0 \end{cases},$$

and

$$d(\boldsymbol{\omega}, \boldsymbol{\theta}) = d^*(\tilde{\phi}^{-1}(\boldsymbol{\omega}), \tilde{\phi}^{-1}(\boldsymbol{\theta})), \quad \boldsymbol{\omega}, \boldsymbol{\theta} \in \mathbb{R}^2,$$

where  $\tilde{\phi} = \phi \circ \rho_{\mathbf{v}_0}$ . Here, we define  $\rho_{\mathbf{v}_0}^{-1}$  as an identity map.

## 6.2 Asymptotic results

Now, write  $\boldsymbol{\theta}_i = \tilde{\phi}(\mathbf{v}_i)$  for  $i = 1, \dots, n$ , and  $\psi(\boldsymbol{\omega}, \boldsymbol{\theta}) = d^2(\boldsymbol{\omega}, \boldsymbol{\theta})$ . We have  $\boldsymbol{\theta}_0 = \tilde{\phi}(\mathbf{v}_0) = \mathbf{0}$ . Also, let  $\boldsymbol{\psi}_1(\boldsymbol{\omega}, \boldsymbol{\theta})$  and  $\boldsymbol{\psi}_2(\boldsymbol{\omega}, \boldsymbol{\theta})$  be the first and second order derivative of  $\psi$  with respect to  $\boldsymbol{\theta}$  respectively. Let  $\mathbf{m}(s) = (m_1(s), m_2(s))^\top = \mathbb{E}(\boldsymbol{\theta}_1 | S_1 = s)$  and  $\boldsymbol{\Sigma}(s) = [\Sigma_{jk}(s)]_{1 \leq j, k \leq 2} = \text{Var}(\boldsymbol{\theta}_1 | S_1 = s)$ . Also, denote  $\boldsymbol{\Psi}(s) = [\Psi_{jk}(s)]_{1 \leq j, k \leq 2} = \mathbb{E}[\boldsymbol{\psi}_2(\boldsymbol{\theta}_1, \boldsymbol{\theta}_0) | S_1 = s]$ . Write  $\mathcal{B}_\delta(\boldsymbol{\theta}_0) = \{\boldsymbol{\theta} \in \mathbb{R}^2 : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| < \delta\}$ , for  $\delta > 0$ . Throughout our discussion, we use the  $L_2$ -norm for matrix. We need the following assumptions to proceed.

**Assumption 1.** *There exists  $\epsilon > 0$  such that  $\text{supp}(\mathbf{V}_1 | S_1 = s) \subseteq \{\mathbf{v} \in \mathbb{R}^3 : d^*(\mathbf{v}, \mathbf{v}_0) \leq \pi/2 - \epsilon\}$ , in a neighborhood of  $s_0$ .*

**Assumption 2.**  *$h \rightarrow 0$  and  $nh \rightarrow \infty$ .*

**Assumption 3.**  *$K(\cdot)$  is bounded, compactly supported kernel function satisfying (i)  $\int K(x)dx = 1$  and (ii)  $\int xK(x)dx = 0$ .*

**Assumption 4.** *The density of  $S$ ,  $f_S(\cdot)$ , is twice continuously differentiable in a neighborhood of  $s_0$  and  $f_S(s_0) > 0$ .*

**Assumption 5.**  *$m_j(\cdot)$  is twice continuously differentiable in a neighborhood of  $s_0$ , for  $j = 1, 2$ .*

**Assumption 6.**  *$\Sigma_{jk}(\cdot)$  is continuous in a neighborhood of  $s_0$ , for  $j, k = 1, 2$ .*

**Assumption 7.**  $\Psi_{jk}(\cdot)$  is continuous in a neighborhood of  $s_0$ , for  $j, k = 1, 2$ .

**Assumption 8.**  $E\{[\psi_2(\boldsymbol{\theta}_1, \boldsymbol{\theta}_0)]_{j,k}^2 | S_1 = s\} \leq C_{jk}$  for all  $s$ , for  $j, k = 1, 2$ .

**Assumption 9.** Let  $\gamma(\delta, s) = \mathbb{E}[\sup_{\tilde{\boldsymbol{\theta}} \in \mathcal{B}_\delta(\boldsymbol{\theta}_0)} \|\psi_2(\boldsymbol{\theta}_1, \tilde{\boldsymbol{\theta}}) - \psi_2(\boldsymbol{\theta}_1, \boldsymbol{\theta}_0)\| | S_1 = s]$ . There exists a neighborhood of  $s_0$ ,  $\mathcal{W}(s_0)$ , such that

$$\tilde{\gamma}(\delta) = \sup_{s \in \mathcal{W}(s_0)} \gamma(\delta, s) = o(1),$$

as  $\delta \rightarrow 0$ .

**Assumption 10.**  $\boldsymbol{\Psi}(s_0)$  is positive definite.

Assumption 1 is a technical assumption for avoiding the unnecessary complication arising from the representation of geodesic distance as a function of the working coordinate system. As a result of Assumption 1, one can use a representation of  $\pm \mathbf{v}$ , which aligns with  $\mathbf{v}_0$ , and reduces the geodesic distance of  $\mathcal{M}$  to the geodesic distance of  $\mathcal{V}$ . This assumption is usually satisfied by our procedure, as a results of thresholding and clustering. Assumptions 2-10 are standard conditions for consistency and distributional limits for smoothing estimators.

**Theorem 1.** Assume that Assumptions 1-10 hold. Let  $M_n(\boldsymbol{\theta}) = \sum_{i=1}^n h K_h(S_i - s_0) d^2(\boldsymbol{\theta}_i, \boldsymbol{\theta})$ .

(a) There exists a sequence of solutions,  $\hat{\boldsymbol{\theta}}_n(s_0)$ , to  $M_n^{(1)}(\boldsymbol{\theta}) = 0$ , such that  $\hat{\boldsymbol{\theta}}_n(s_0)$  converges in probability to  $\boldsymbol{\theta}_0$ .

(b) And  $\hat{\boldsymbol{\theta}}_n$  is asymptotically normal:

$$\sqrt{nh} \left\{ (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) - h^2 \boldsymbol{\eta} \right\} \implies \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Omega}),$$

where

$$\boldsymbol{\eta} = 2 \int x^2 K(x) dx \boldsymbol{\Psi}^{-1}(s_0) \left\{ \frac{f_S^{(1)}(s_0)}{f_S(s_0)} m^{(1)}(s_0) + \frac{1}{2} m^{(2)}(s_0) \right\}$$

and

$$\boldsymbol{\Omega} = 4 \int K^2(x) dx \boldsymbol{\Psi}^{-1}(s_0) \boldsymbol{\Sigma}(s_0).$$

The proof of the Theorem 1 can be found in Section S5 of the SM.

## 7 Simulation results

This section presents simulation results of the proposed DiST procedure. For simulation results of the voxel-wise estimation procedure proposed in Section 3, see Section S2 of the SM.

We simulate 200 diffusion tensor data sets from the tensor field given in Figure 4 (Top). The tensors all have the principal eigenvalues being  $4 \times 10^{-3}$  and FA (1) being 0.9. The  $b$ -value is set to be 1000 across all voxels. This mimics the  $b$ -value and diffusivity (reflected by the numerical scale of the tensor) in a real dMRI study.

At each voxel there is either one tensor or there are two tensors. For crossing fiber regions,  $p_1$  and  $p_2$  are set to 0.7 and 0.3 respectively, and the separation angles between the two tensors range from 66.3 to 86.6 degree. In crossing fiber regions of Figure 4 (Top), the more transparent the tensor is, the less weight it takes.

In addition,  $S_0(\mathbf{s})$ 's have the same value which is set to 1000. Two choices of the noise standard deviation  $\sigma$  are used, namely 50 and 100, which corresponds to signal-to-noise ratio ( $S_0/\sigma$ ) of 20 and 10, respectively. The case that  $\text{SNR} = 20$  is typical for dMRI studies while that  $\text{SNR} = 10$  corresponds to a high noise setting. The set of gradient directions  $\mathcal{U}$  is obtained from the sphere tessellation with 3 subdivision using octahedron and  $|\mathcal{U}| = 33$ , which is in a typical range for dMRI studies nowadays. With these gradient directions, the observed signal intensities  $\mathbf{S}(\mathbf{s})$ 's are simulated according to the multi-tensor model (2) with the Rician noise. A total of four different procedures are compared:

- raw: voxel-wise estimation without any smoothing;
- DiST-cv: DiST using ordinary cross-validation score for choosing  $h$ ;
- DiST-tcv: DiST using 5% trimmed cross-validation score for choosing  $h$ ;
- DiST-mcv: DiST using median cross-validation score for choosing  $h$ .

See Section S3 of the SM for definitions of the various cross-validation variants.

Table 1 shows numerical summaries of the simulation results. In addition to the proportion of correctly estimated number of diffusion directions, we also report the mean MSE (MMSE) and the mean root MSE (MRMSE), defined as follows. Conditional on the correct estimation of  $J$ , the squared error of  $\mathbf{m}$  is defined as

$$\min_{\{k_1, \dots, k_J \in \{1, \dots, J\}: k_i \neq k_j\}} \sum_{j=1}^J d^{*2}(\mathbf{m}_j, \hat{\mathbf{u}}_{k_j}), \quad (10)$$

where  $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_J$  are the estimated diffusion directions. Here, the MSE is the mean of squared errors (10) over voxels with  $\hat{J} = J$  in one simulated data set and root MSE (RMSE) is the square root of MSE. Then MMSE and MRMSE are defined, respectively, as the means of MSEs and RMSEs over the 200 simulated data sets.

The voxel-wise estimation works reasonably well in estimating both the number of diffusion directions  $J$  and the diffusion directions. Even for the low **SNR** setting, the correctness of estimation of  $J$  is around 75% and the angular error is no more than 11 degree. For the single tensor region ( $J = 1$ ), smoothing improves upon estimation of both  $J$  and diffusion directions. For regions with two tensors ( $J = 2$ ), smoothing only improves direction estimation. Among the three smoothing procedures, DiST-mcv works the best.

Table 2 shows the five-number summary of the maximum angular error with  $\hat{J} = J = 2$  across the 200 simulated data sets. Again smoothing procedures have smaller errors than the raw procedure and DiST-mcv is the best among all methods. For DiST-mcv, the mean and median of angular errors are around 2.5 degree and 1 degree for **SNR** = 10 and **SNR** = 20, respectively. Such magnitude of errors has little impact on tracking.

We then apply the proposed tracking algorithm in Section 5 (Algorithm S4, SM) to the estimated diffusion directions based on the above procedures. The tracking results of a simulation with **SNR** = 10 are shown in Figures 4 (Bottom) and 5. As can be seen in Figure 5, the lines produced by DiST are much more aligned when compared to the tracking result based on voxel-wise estimation without smoothing (raw).

Table 1: Diffusion direction estimation results. **Correct-select**: proportion of  $\hat{J} = J$ . **MMSE**: mean of MSEs (Each MSE is computed over voxels with  $\hat{J} = J$  in one simulated data set.), in squared degree, of the estimated diffusion direction, with the corresponding standard error stated in brackets. **MRMSE**: mean of RMSEs (Each RMSE is computed over voxels with  $\hat{J} = J$  in one simulated data set.), in degree, of the estimated diffusion direction, with the corresponding standard error stated in brackets.

SNR	$J$		raw	DiST-cv	DiST-tcv	DiST-mcv
10	1	Correct-select	97.12%	99.09%	99.15%	99.45%
		MMSE	9.84 (3.84e-02)	4.95 (2.94e-01)	2.70 (1.09e-01)	3.06 (1.40e-01)
		MRMSE	3.14 (6.12e-03)	2.09 (5.46e-02)	1.60 (2.60e-02)	1.69 (3.13e-02)
	2	Correct-select	75.18%	74.38%	75.37%	75.44%
		MMSE	114 (2.42)	50.9 (3.45)	40.0 (3.11)	9.81 (1.40)
		MRMSE	10.6 (1.07e-01)	6.05 (2.68e-01)	5.26 (2.49e-01)	2.49 (1.35e-01)
20	1	Correct-select	98.59%	99.46%	99.69%	99.75%
		MMSE	2.30 (8.50e-03)	1.25 (1.23e-01)	7.97e-01 (3.02e-02)	1.15 (5.47e-02)
		MRMSE	1.52 (2.80e-03)	1.02 (3.28e-02)	8.79e-01 (1.10e-02)	1.03 (2.04e-02)
	2	Correct-select	99.38%	99.94%	99.99%	99.99%
		MMSE	19.8 (2.12e-01)	6.43 (5.18e-01)	2.00 (2.84e-01)	1.48 (2.13e-01)
		MRMSE	4.43 (2.34e-02)	2.13 (9.75e-02)	1.13 (6.02e-02)	9.93e-01 (4.98e-02)

Table 2: Summary statistics of the maximum absolute error across the voxels with  $\hat{J} = J = 2$ .

SNR	Method	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
10	raw	0.530	6.63	9.86	11.8	14.6	89.3
	DiST-cv	0.132	2.32	4.99	6.97	9.59	89.3
	DiST-tcv	0.0933	2.08	4.01	6.00	8.07	89.3
	DiST-mcv	0.135	1.35	2.11	2.91	3.35	65.1
20	raw	0.350	3.20	4.67	5.20	6.65	29.5
	DiST-cv	0.0803	0.931	1.73	2.48	3.28	26.1
	DiST-tcv	0.0494	0.613	0.965	1.33	1.53	15.7
	DiST-mcv	0.0473	0.531	0.841	1.16	1.40	15.9

## 8 Real data application

In this section, we apply the proposed methodology to a real dMRI data set, which was obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database ([www.loni.ucla.edu/ADNI](http://www.loni.ucla.edu/ADNI)). The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and onset of Alzheimer’s disease

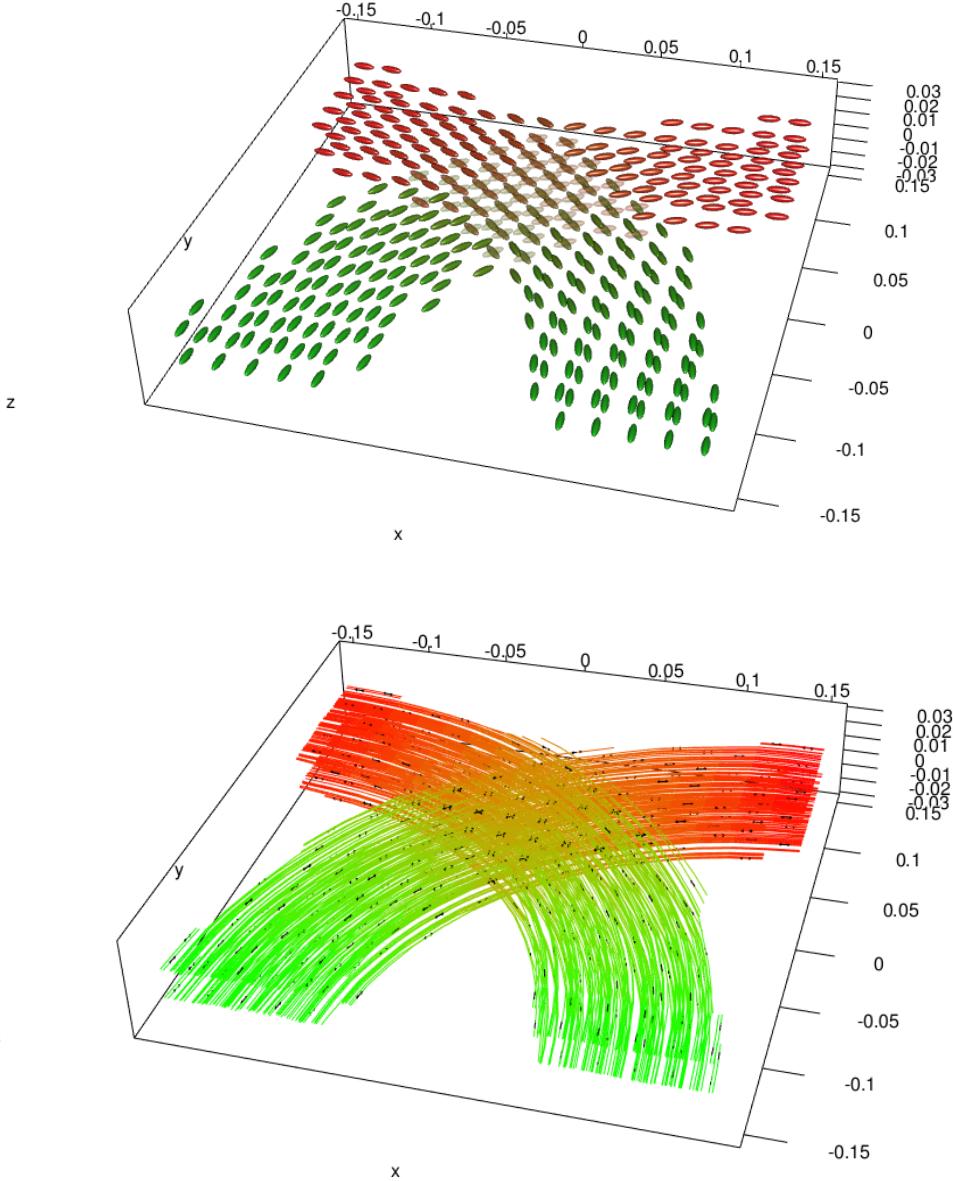


Figure 4: Top: The true tensor field used in the simulation study (Section 7). Bottom: Illustration of fiber tracking using DiST-mcv.

(AD). In the following, we use an eddy-current-corrected ADNI data set of a normal subject for illustration of our technique.

This data set contains 41 distinct gradient directions with  $b$ -value set as  $1000s/mm^2$ . In addition, there are 5  $b_0$  images (corresponding to  $b = 0$ ), forming in total 46 measurements for each of

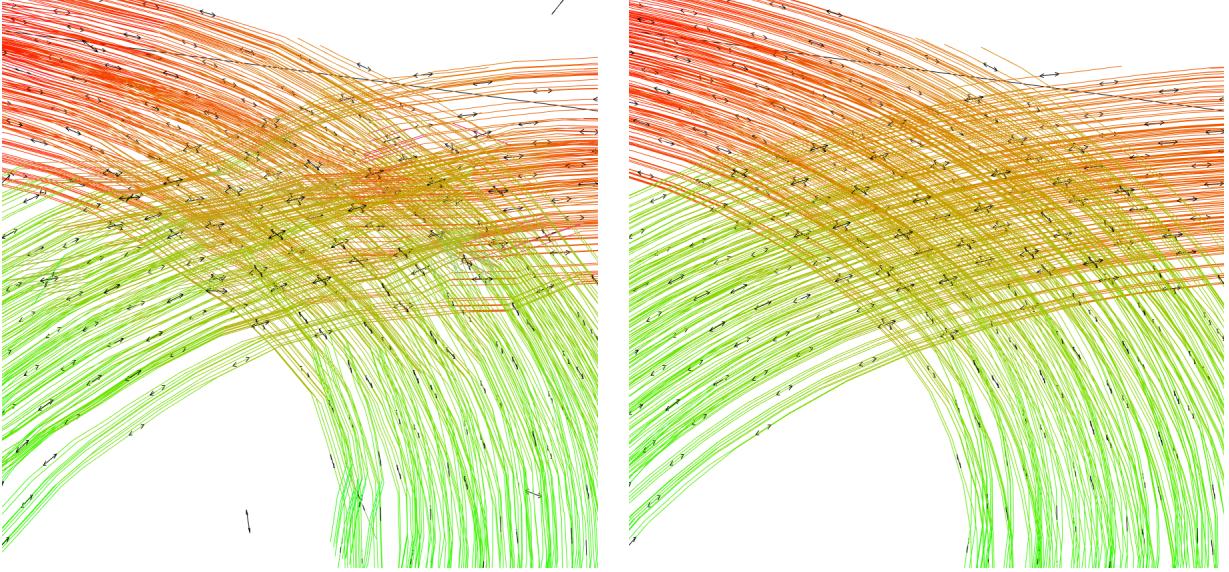


Figure 5: Illustration of fiber tracking over the crossing fiber region by `raw` (left) and DiST-mcv (right) respectively.

the  $256 \times 256 \times 59$  voxels. To implement our technique, we require estimates of  $S_0(\mathbf{s})$ 's and  $\sigma$ . We first estimate  $S_0(\mathbf{s})$  and  $\sigma(\mathbf{s})$  for each voxel by ML estimation based on the 5  $b_0$  images. Then we fix  $\sigma$  as the median of estimated  $\sigma(\mathbf{s})$ 's for voxel-wise estimation of the diffusion directions. Since the original  $256 \times 256 \times 59$  voxels contain volume outside the brain, we only take median over a human-chosen set of  $81 \times 81 \times 20$  voxels. The estimated  $\sigma$  is 56.9.

In this analysis, we focus on a subset of voxels ( $15 \times 15 \times 5$ ), which contains the intersection of corpus callosum (CC) and corona radiata (CR). This region is known to contain significant fiber crossing (Wiegell *et al.*, 2000). Figure 6 shows the fiber orientation color map (derived from the single tensor model). The aforementioned region is indicated by a white rectangular box. Within this region,  $S_0(\mathbf{s})$ 's have mean 1860.1 and standard deviation 522.7.

We then apply voxel-wise estimation to individual voxels followed by the DiST-mcv procedure. Distributions of the estimated number of diffusion directions are summarized in Table 3. For comparison purposes, we also fit the single tensor model with the commonly used regression estimator (e.g., Mori, 2007).

The tracking results are produced by applying the proposed tracking algorithm to the estimated diffusion directions from DiST-mcv and those from the single tensor model estimation. Figure 7 shows the corresponding tracking results. For visualization purposes, we also present the longest 300 tracts in Figure 8. From anatomy, the CC has a mediolateral direction while the CR has a superoinferior orientation. They are clearly shown in both tracking results. In these figures, reconstructed fiber tracts are colored by a RGB color model with red for left-right, green for anteroposterior, and blue for superior-inferior. Thus, one can easily locate the CC and the CR as the red fiber bundle and the blue fiber bundle respectively. Tracking result based on DiST-mcv shows clear crossing between mediolateral fiber and the superoinferior fiber (In the figure, the crossing of red and blue fiber tracts). From neuroanatomic atlases and previous studies, Wiegell *et al.* (2000) conclude that there are several fiber populations with crossing structure in this conjunction region of CC and CR, which matches with the tracking based on DiST-mcv. However, the single tensor model estimation can only reconstruct one major diffusion direction in each voxel and thus the corresponding tracking result does not show crossing structure. Instead, the CC (red fiber bundle) is blocked by the CR (blue fiber bundle) and this leads to either termination of the CC fiber tracts or significant merging of the CC and the CR fiber tracts instead of the known crossing structure. To give further illustration, Figure 9 shows the locations of the CC, the CR and the region of crossing fibers (Cross). One can see that tracking based on DiST-mcv reproduces the crossing fiber structures between the CC and the CR, while the result based on single tensor model tends to connect the CC and the CR fibers.

Moreover, the green fiber on top of the CC represents the cingulum bundle. Both fiber tracking based DiST and single fiber model produce clear and sensible reconstruction of cingulum bundle. All these features match with neuroanatomic atlases and provide a good demonstration of our proposed method.

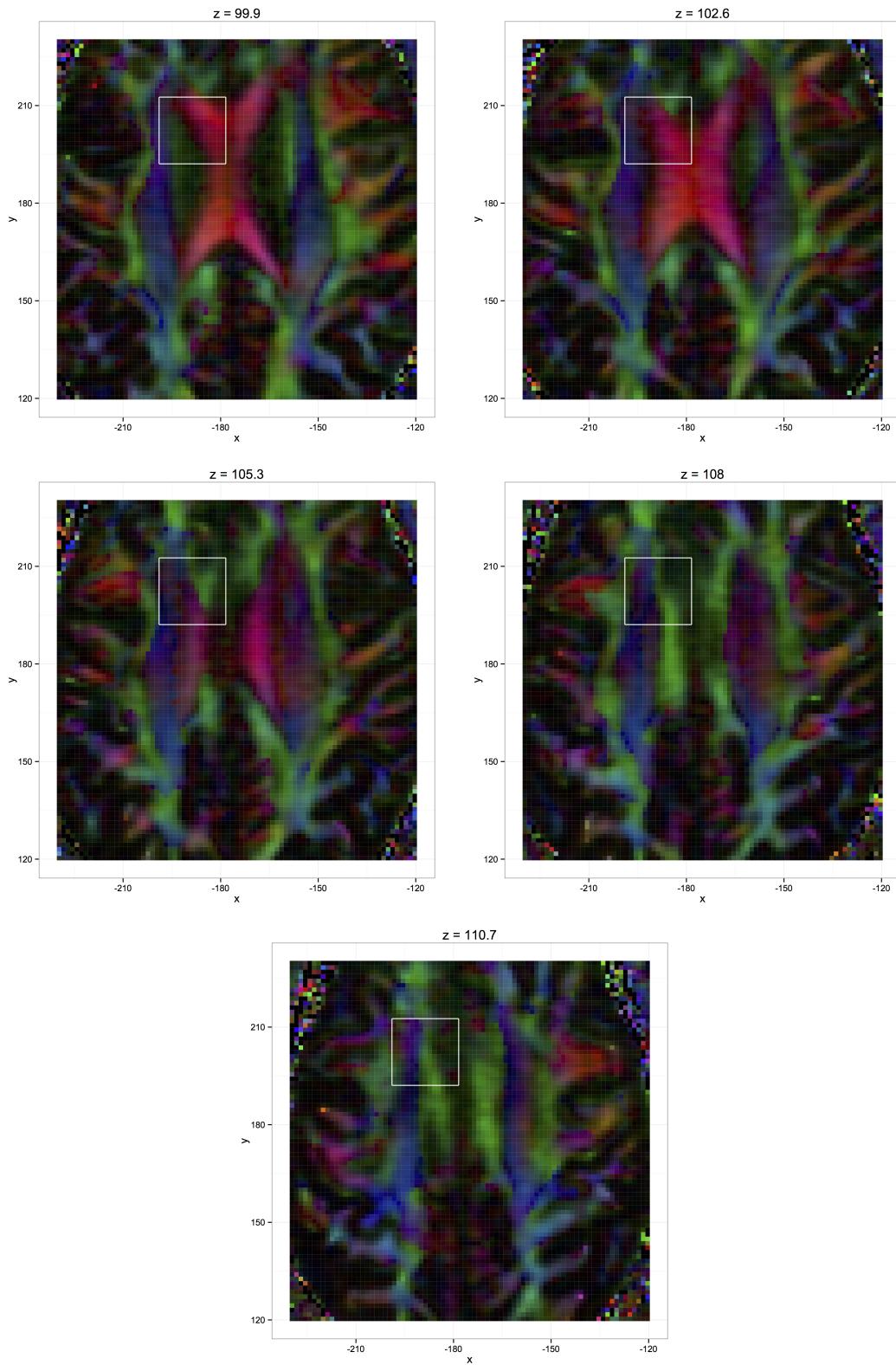


Figure 6: The fiber orientation color map (based on the single tensor model). The focused region is indicated by white rectangular boxes.

Table 3: Number of voxels with different estimated number of diffusion directions.

	Number of diffusion directions				
	0	1	2	3	4
Voxel-wise estimation	37	476	589	23	0
Smoothing	37	476	593	19	0

## 9 Discussion

Using tensor estimation to resolve cross-fiber can be problematic, due to the non-identifiability issue in multi-tensor model. In this paper, we take a different route by focusing on the estimation of diffusion directions rather than the non-identifiable diffusion tensors. We develop the corresponding direction smoothing procedure and fiber tracking strategy, together called DiST, along this route. Our technique gives promising empirical results in both simulation study and real data analysis.

The procedure we presented works well even with moderate number of gradient directions (a few tens), as long as the number of distinct crossing fibers within a voxel is not large. With HARDI data, which can have up to a couple of hundreds gradient directions, rather than modeling the direction distribution within a tensor framework, we can estimate the fiber orientation distribution nonparametrically (Tuch, 2004; Descoteaux *et al.*, 2007). In that case, we can potentially extend the fiber tracking procedure presented here by adopting a probabilistic approach in which the directions for moving from one voxel to another are sampled from the fiber orientation distribution. Such a probabilistic fiber tracking has the additional advantage of giving a measure of uncertainty of the fiber tracts extracted from the data. This is a topic of future research.

## Acknowledgement

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering

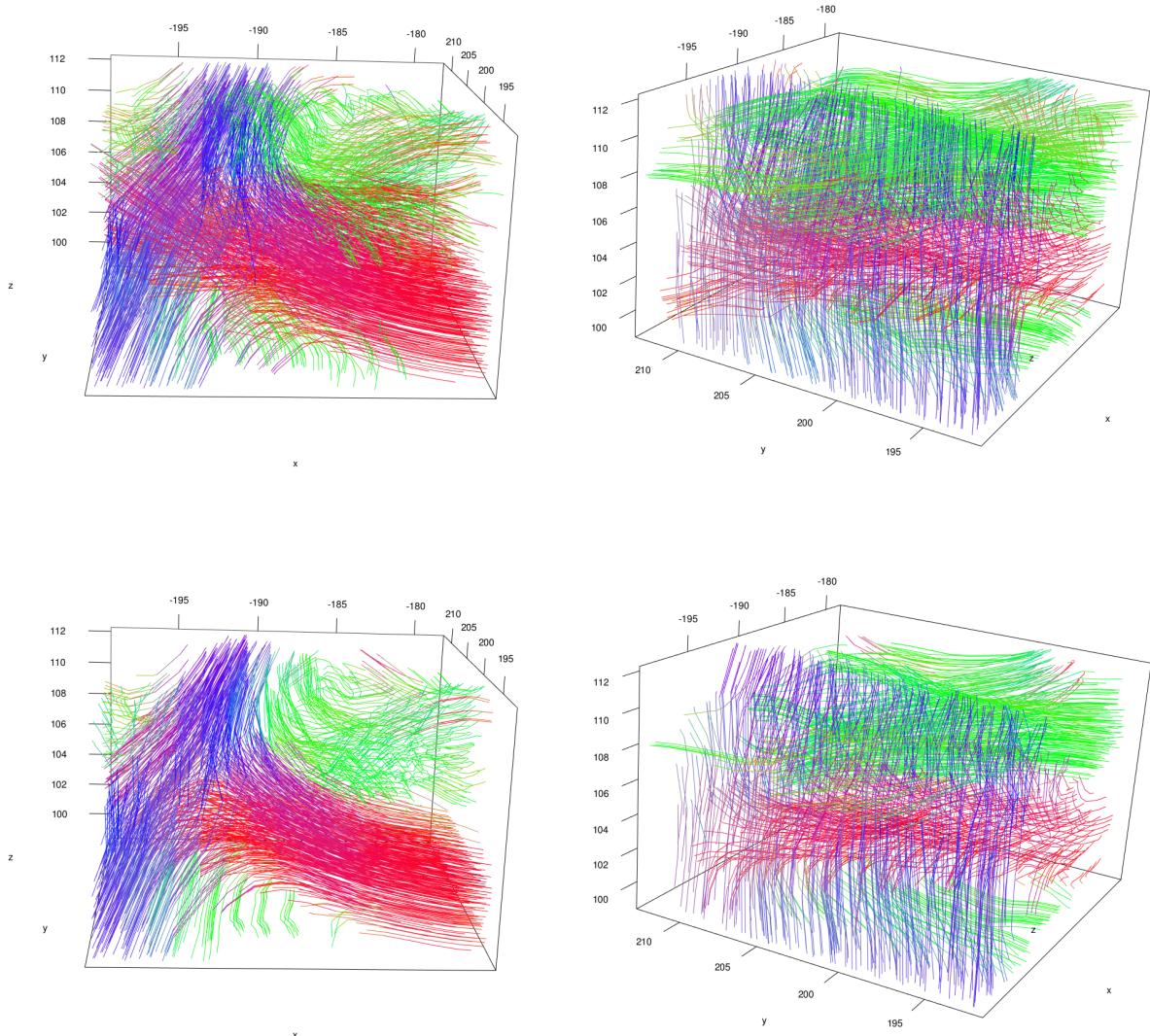


Figure 7: Top: Tracking using DiST-mcv. Bottom: Tracking using single tensor model.

Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research

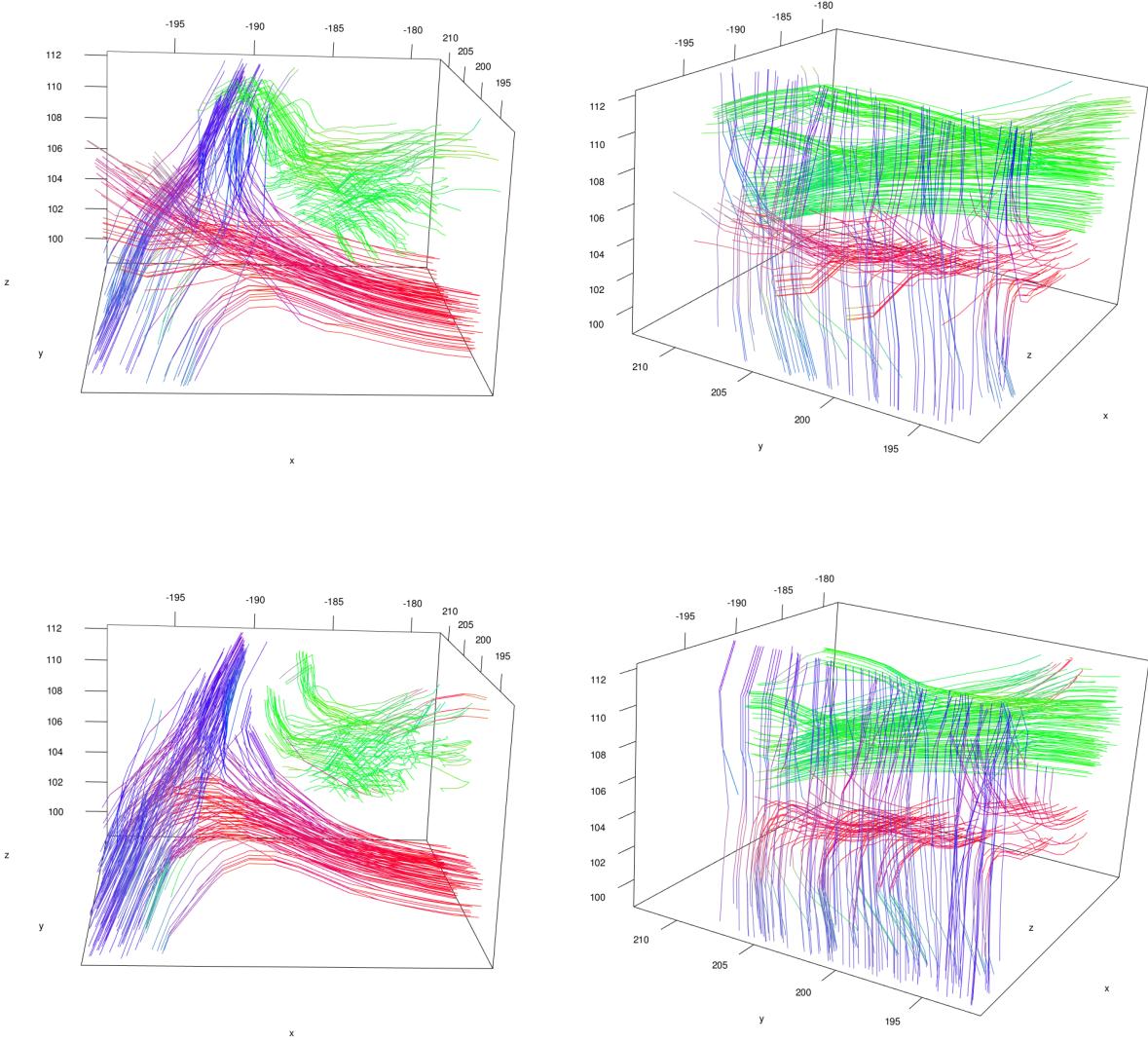


Figure 8: Top: The longest 300 tracks using DiST-mcv (The left and right figures correspond to different view angles). Bottom: Similarly for single tensor model.

and Education, and the study is coordinated by the Alzheimer’s Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation. The authors would like to thank Professor Owen Carmichael for making available the data and his valuable comments. The work of Lee was supported in part by the National Science Foundation under Grants DMS-10-07520, DMS-

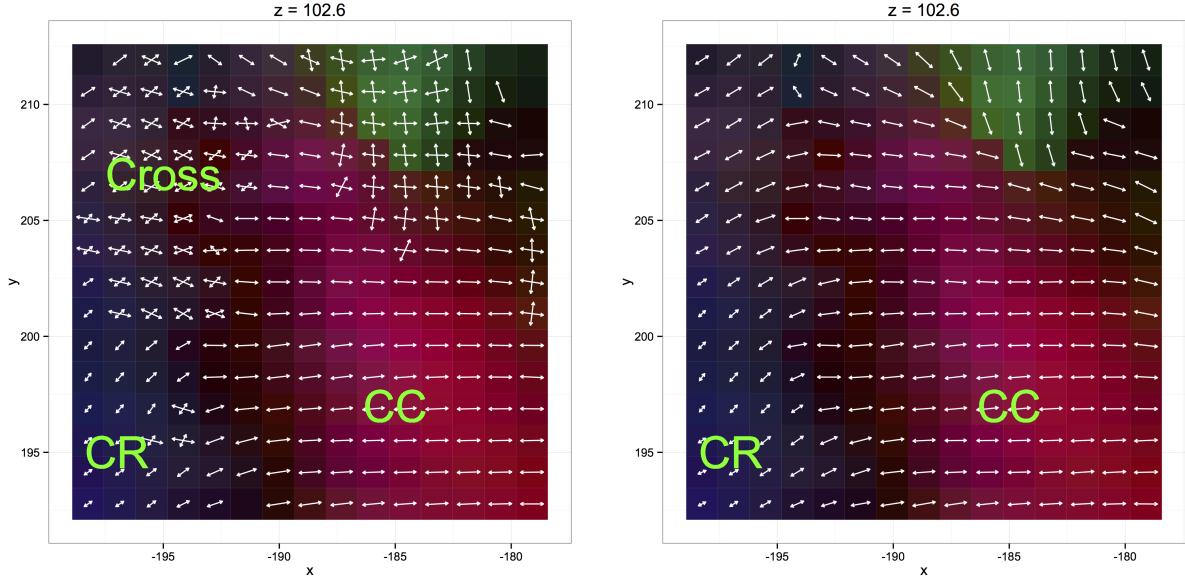


Figure 9: The projection of fiber directions to the  $xy$ -plane at  $z = 102.6$  for illustration of crossing fibers. The plot also shows the location of corpus callosum (CC), corona radiata (CR) and crossing region (Cross). The fiber orientation color map is overlaid as the background. Left: for DiST-mcv. Right: for single tensor model.

12-09226 and DMS-12-09232. The work of Paul was supported in part by the National Science Foundation under Grant DMS-11-06690. The work of Peng was supported in part by the National Science Foundation under Grant DMS-10-07583.

## References

- Abramowitz, M. and Stegun, I. A. (1964) *Handbook of Mathematical Functions: With Formulars, Graphs, and Mathematical Tables*, vol. 55. National Bureau of Standards.
- Arsigny, V., Fillard, P., Pennec, X. and Ayache, N. (2006) Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic resonance in medicine*, **56**, 411–421.
- Bammer, R., Holdsworth, S. J., Veldhuis, W. B. and Skare, S. T. (2009) New methods in diffusion-

- weighted and diffusion tensor imaging. *Magnetic resonance imaging clinics of North America*, **17**, 175–204.
- Basser, P. J., Mattiello, J. and LeBihan, D. (1994) Mr diffusion tensor spectroscopy and imaging. *Biophysical journal*, **66**, 259–267.
- Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J. and Aldroubi, A. (2000) In vivo fiber tractography using dt-mri data. *Magnetic Resonance in Medicine*, **44**, 625–632.
- Beaulieu, C. (2002) The basis of anisotropic water diffusion in the nervous system – a technical review. *NMR in Biomedicine*, **15**, 435–455.
- Behrens, T., Berg, H. J., Jbabdi, S., Rushworth, M. and Woolrich, M. (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage*, **34**, 144–155.
- Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, J. and Smith, S. (2003) Characterization and propagation of uncertainty in diffusion-weighted mr imaging. *Magnetic Resonance in Medicine*, **50**, 1077–1088.
- Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995) A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, **16**, 1190–1208.
- Carmichael, O., Chen, J., Paul, D. and Peng, J. (2013) Diffusion tensor smoothing through weighted karcher means. *Electronic Journal of Statistics*, **7**, 1913–1956.
- Chanraud, S., Zahr, N., Sullivan, E. V. and Pfefferbaum, A. (2010) Mr diffusion tensor imaging: a window into white matter integrity of the working brain. *Neuropsychology review*, **20**, 209–225.
- Descoteaux, M., Angelino, E., Fitzgibbons, S. and Deriche, R. (2007) Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, **58**, 497–510.
- Fan, J. and Gijbels, I. (1996) *Local Polynomial Modelling and Its Applications*. London: Chapman and Hall.

- Fillard, P., Pennec, X., Arsigny, V. and Ayache, N. (2007) Clinical dt-mri estimation, smoothing, and fiber tracking with log-euclidean metrics. *Medical Imaging, IEEE Transactions on*, **26**, 1472–1482.
- Fletcher, P. T. and Joshi, S. (2007) Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing*, **87**, 250–262.
- Friman, O., Farneback, G. and Westin, C.-F. (2006) A bayesian approach for stochastic white matter tractography. *Medical Imaging, IEEE Transactions on*, **25**, 965–978.
- Gudbjartsson, H. and Patz, S. (1995) The rician distribution of noisy mri data. *Magnetic Resonance in Medicine*, **34**, 910–914.
- Hosey, T., Williams, G. and Ansorge, R. (2005) Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, **54**, 1480–1489.
- Kaufman, L. and Rousseeuw, P. J. (1990) *Finding groups in data: an introduction to cluster analysis*, vol. 344. New Jersey: John Wiley & Sons.
- Koch, M. A., Norris, D. G. and Hund-Georgiadis, M. (2002) An investigation of functional and anatomical connectivity using magnetic resonance imaging. *Neuroimage*, **16**, 241–250.
- Mori, S. (2007) *Introduction to diffusion tensor imaging*. Amsterdam: Elsevier.
- Mori, S., Crain, B. J., Chacko, V. and Van Zijl, P. (1999) Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of neurology*, **45**, 265–269.
- Mori, S. and van Zijl, P. (2002) Fiber tracking: principles and strategies—a technical review. *NMR in Biomedicine*, **15**, 468–480.
- Mukherjee, P., Berman, J., Chung, S., Hess, C. and Henry, R. (2008) Diffusion tensor mr imaging and fiber tractography: theoretic underpinnings. *American journal of neuroradiology*, **29**, 632–641.

- Parker, G. J. M. and Alexander, D. C. (2003) Probabilistic monte carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *Information Processing in Medical Imaging*, 684–695. Springer.
- Pennec, X., Fillard, P. and Ayache, N. (2006) A riemannian framework for tensor computing. *International Journal of Computer Vision*, **66**, 41–66.
- Rousseeuw, P. J. (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, **20**, 53–65.
- Scherrer, B. and Warfield, S. K. (2010) Why multiple b-values are required for multi-tensor models. evaluation with a constrained log-euclidean model. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 1389–1392.
- Schwarz, G. (1978) Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.
- Sporns, O. (2011) *Networks of the Brain*. The MIT Press.
- Tabelow, K., Voss, H. and Polzehl, J. (2012) Modeling the orientation distribution function by mixtures of angular central gaussian distributions. *Journal of neuroscience methods*, **203**, 200–211.
- Tournier, J., Calamante, F., Connelly, A. *et al.* (2007) Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, **35**, 1459–1472.
- Tournier, J., Calamante, F., Gadian, D. G., Connelly, A. *et al.* (2004) Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage*, **23**, 1176–1185.
- Tuch, D. S. (2002) *Diffusion MRI of complex tissue structure*. Ph.D. thesis, Massachusetts Institute of Technology.

- Tuch, D. S. (2004) Q-ball imaging. *Magnetic Resonance in Medicine*, **52**, 1358–1372.
- Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W. and Wedeen, V. J. (2002) High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, **48**, 577–582.
- Weinstein, D., Kindlmann, G. and Lundberg, E. (1999) Tensorlines: Advection-diffusion based propagation through diffusion tensor fields. In *Proceedings of the conference on Visualization*, 249–253.
- Wiegell, M. R., Larsson, H. B. and Wedeen, V. J. (2000) Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology*, **217**, 897–903.
- Yuan, Y., Zhu, H., Lin, W. and Marron, J. S. (2012) Local polynomial regression for symmetric positive definite matrices. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **74**, 697–719.