FIBER DIRECTION ESTIMATION, SMOOTHING AND TRACKING IN DIFFUSION MRI[1,2]

BY RAYMOND K. W. WONG[*], THOMAS C. M. LEE[†,3], DEBASHIS PAUL[†,4], JIE PENG[†,5] AND ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE

Iowa State University [*] *and University of California, Davis* [†]

Diffusion magnetic resonance imaging is an imaging technology designed to probe anatomical architectures of biological samples in an in vivo and noninvasive manner through measuring water diffusion. The contribution of this paper is threefold. First, it proposes a new method to identify and estimate multiple diffusion directions within a voxel through a new and identifiable parametrization of the widely used multi-tensor model. Unlike many existing methods, this method focuses on the estimation of diffusion directions rather than the diffusion tensors. Second, this paper proposes a novel direction smoothing method which greatly improves direction estimation in regions with crossing fibers. This smoothing method is shown to have excellent theoretical and empirical properties. Last, this paper develops a fiber tracking algorithm that can handle multiple directions within a voxel. The overall methodology is illustrated with simulated data and a data set collected for the study of Alzheimer's disease by the Alzheimer's Disease Neuroimaging Initiative (ADNI).

1. Introduction. Diffusion magnetic resonance imaging (dMRI) is an in vivo and noninvasive medical imaging technology that uses water diffusion as a proxy to probe anatomical structures of biological samples. The most important application of dMRI is to reconstruct white matter fiber tracts in the brain—large axonal bundles with similar destinations. In white matter, water diffusion appears to be anisotropic, as water tends to diffuse faster along the fiber bundles. Therefore, white matter fiber structures can be deduced from the diffusion characteristics of water. Mapping white matter fiber tracts is of great importance in the study of structural organization of neuronal networks and the understanding of brain functionality [Mori (2007), Sporns (2011)]. Moreover, dMRI also has many clinical

applications, including detecting brain abnormality in white matter due to axonal loss or deformation, which are thought to be related to many neuron degenerative diseases including Alzheimer's disease, and also in surgical planning by resolving complex neuronal connections between white and gray matter [Nimsky, Ganslandt and Fahlbusch (2006)].

dMRI techniques sensitize signal intensity with the amount of water diffusion by applying pulsed magnetic gradient fields on the sample. Specifically, water diffusion along the gradient field direction leads to signal loss and the amount of loss at a voxel equals the summation (across locations within the voxel) of the sinusoid waves with shifted signal phases weighted by the proton density at their respective locations. In other words, signal loss (referred to as the diffusion weighted signal) is the inverse Fourier transform of the diffusion probability density function of water molecules, and thus can be used to recover water diffusion characteristics. The amount of signal loss is also influenced by various experimental parameters, including the gradient field intensity (the stronger, the more loss), the duration of gradient fields (the longer, the more loss), etc. Their effects are aggregatively reflected by an experimental parameter called the "b-value" which is often fixed throughout the experiment (though multiple b-values are used in Q-space imaging). Since only water motion along the gradient field direction can be detected, multiple gradient directions need to be applied [Mori (2007)].

In its raw form, dMRI provides diffusion weighted signal measurements on a 3D spatial grid (of the sample) along a set of predetermined gradient directions [Bammer et al. (2009), Beaulieu (2002), Chanraud et al. (2010), Mukherjee et al. (2008)]. For example, a typical data set from the Alzheimer's Disease Neuroimaging Initiative (ADNI) has diffusion measurements along 41 gradient directions for each voxel on a $256 \times 256 \times 59$ 3D grid of the brain. The first step of dMRI analysis is to summarize these measurements into estimates of water diffusion at each voxel. A popular model for water diffusion is the so-called single tensor model where the diffusion process is modeled as a 3D Gaussian process described by a 3×3 positive definite matrix, referred to as a diffusion tensor; see Mori (2007) for an introduction to diffusion tensor imaging (DTI) techniques. Figure 1 depicts a tensor map on a 2D grid, where each diffusion tensor is represented by an ellipsoid, estimated from diffusion weighted measurements from an ADNI data set using a single tensor model. One then extracts the local diffusion direction as the principal eigenvector of the (estimated) diffusion tensor at each voxel and reconstructs the white matter fiber tracts by computer-aided tracking algorithms via a process called tractography [Basser et al. (2000)].

However, DTI cannot resolve multiple fiber populations with distinct orientations, that is, crossing fibers, within a voxel since a tensor only has one principal direction. Consequently, 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 adversely affects fiber reconstruction, for example, early termination of or biased fiber tracking.
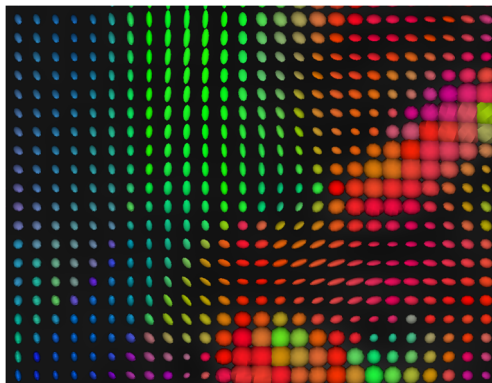
FIG. 1. *An example of a tensor map on a* 2*D grid where each diffusion tensor is represented by an elliposid.*

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 within a voxel. However, it has been shown that the parameters in the multi-tensor model are not identifiable [Scherrer and Warfield (2010)]. Imaging techniques such as Q-ball and Q-space and the corresponding nonparametric methods have also been proposed [Descoteaux et al. (2007), Tuch (2004)]. However, such methods often require high angular resolution diffusion imaging (HARDI) [Hosey, Williams and Ansorge (2005), Tuch et al. (2002)] where a large number of gradients is sampled. In light of these facts, the goal of this paper is to develop a new fiber direction estimation and tracking method that can handle crossing fibers without requiring any high resolution techniques. The proposed method, named DiST, short for Diffusion Direction Smoothing and Tracking, is completely automated and improves existing methods in several aspects. Particularly, it is applicable either when there is a large number of gradient directions (as in the HARDI setting) or when only a relatively small number of gradient directions are available (as in most clinical settings).

The DiST method can be divided into three major steps.

Step 1: Estimate the tensor directions within each voxel under a multi-tensor model. A new parametrization is proposed which makes the tensor directions identifiable. An efficient and numerically stable computational procedure is developed to obtain the maximum likelihood (ML) estimate of the tensor directions. Here we highlight that this method focuses on the estimation of the tensor directions rather than the actual tensors themselves.

Step 2: Using the voxel-wise tensor direction estimates from step 1 as input, a new 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 smoothing 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., Arsigny et al. (2006), Carmichael et al. (2013), Fillard et al. (2007), Fletcher and Joshi (2007), Pennec, Fillard and Ayache (2006), Yuan et al. (2012)], little work has been done on direct diffusion direction smoothing. One notable exception is the work of Schwartzman, Dougherty and Taylor (2008), which harnesses diffusion directions directly to construct a map of test statistics for detecting differences between diffusion direction maps from two groups of subjects while the spatial smoothness of the test statistics is being considered. Also, note that approaches to averaging unsigned directions in the real projective space are known in the directional statistics literature.

Step 3: Last, a fiber tracking algorithm is applied to reconstruct fiber tracts using the smoothed diffusion direction estimates obtained in step 2. This tracking algorithm is designed to explicitly allow for multiple directions within a voxel.

We apply DiST to an ADNI data set measured on a healthy elderly person with a 41-direction dMRI scan on a 3 Tesla GE Medical Systems MRI scanner. ADNI is a longitudinal study (since 2005) that collects serial MRI, cognitive assessments and numerous additional measurements approximately twice per year from hundreds of elderly individuals spanning a range from cognitive health to clinically diagnosed Alzheimer's disease. We also examine DiST using simulated data sets which mimic the most commonly encountered experimental situations in terms of number of gradient directions and signal-to-noise ratio. DiST is shown to lead to superior results than those based on the single tensor model in the simulation study, as well as more biologically sensible results in the real data application.

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. Section 6 summarizes simulation results. The application to an ADNI data set is presented in Section 7. Section 8 provides some concluding remarks, while additional simulation results and technical details are collected in the online Supplemental Material [Wong et al. (2016)].

2. Tensor models. Suppose dMRI measurements are made on N voxels on a 3D grid representing a brain. For each voxel, we have measurements of diffusion weighted signals along a fixed set (i.e., the same for all voxels) of unit-norm gradient vectors $\mathcal{U} = \{\mathbf{u}_i : i = 1, \ldots, m\}$. We write the set of measurements as $\{S(\mathbf{s}, \mathbf{u}) : \mathbf{u} \in \mathcal{U}\}$, where $\mathbf{s}$ is the 3D coordinate of the center of this voxel.

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

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

where $S_0(\mathbf{s})$ is the nondiffusion-weighted intensity, $b > 0$ is an experimental constant referred to as the b-value and $\mathbf{D}(\mathbf{s})$ is a 3×3 covariance matrix referred to as the diffusion tensor. This model is called the single tensor model and is suited for the case of at most one dominant diffusion direction within a voxel.

Although the single tensor model is the most widely used tensor model in practice, 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., Behrens et al. (2003, 2007), Tabelow, Voss and Polzehl (2012), Tuch (2002)]:

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

where $\sum_{j=1}^{J(\mathbf{s})} p_j(\mathbf{s}) = 1$ and $p_j(\mathbf{s}) > 0$ for $j = 1, \ldots, J(\mathbf{s})$. 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 dMRI 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 σ and $S_0(\mathbf{s})$ are known and delay the discussion of their estimation to Section 7.

Under the single tensor model, various methods for tensor estimation have been proposed, including linear regression, nonlinear regression and ML estimation; for example, see Carmichael et al. (2013) for a comprehensive review. Then diffusion directions are derived as principal eigenvectors of (estimated) diffusion tensors. However, for the estimation of multi-tensor models, severe computational issues have been observed and additional prior information and additional assumptions are usually imposed to tackle these issues. For instance, Behrens et al. (2003, 2007) use shrinkage priors and Tabelow, Voss and Polzehl (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 nonidentifiable in the sense that there exist multiple parameterizations that are observationally equivalent. These authors 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.*   Model (1) can be rewritten as

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

where $a_j > 0$ for $j = 1, \ldots, J$ 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 for modeling dMRI data and it implies that diffusion is symmetric around the principal diffusion direction [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 for each tensor in the multiple tensor model and thus facilitates 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, that is, the 3D 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 $\mathbf{D}_j = \alpha_j \mathbf{m}_j \mathbf{m}_j^\mathsf{T} + \xi_j \mathbf{I}_3$, model (1) becomes

$$\bar{S}(\mathbf{u}) = S_0 \sum_{j=1}^{J} p_j \exp\{-b\mathbf{u}^\mathsf{T}(\alpha_j \mathbf{m}_j \mathbf{m}_j^\mathsf{T} + \xi_j \mathbf{I}_3)\mathbf{u}\}$$

(2)

$$= S_0 \sum_{j=1}^{J} \tau_j \exp\{-b\alpha_j (\mathbf{u}^\mathsf{T}\mathbf{m}_j)^2\},$$

where $\tau_j = p_j \exp(-b\xi_j) \in (0, 1)$. From the above, one can see that p_j and ξ_j are not simultaneously identifiable, so we cannot estimate the tensors. However, as stated in the following theorem, $\tau_j, \alpha_j, \mathbf{m}_j^\mathsf{T}$ are identifiable, and hence we can estimate the principal diffusion directions $\mathbf{m}_j$'s.

THEOREM 1.   *Under model* (2), *for any arbitrary J, the parameters $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\mathsf{T}, \ldots, \boldsymbol{\gamma}_J^\mathsf{T})^\mathsf{T}$ are identifiable, where $\boldsymbol{\gamma}_j = (\tau_j, \alpha_j, \mathbf{m}_j^\mathsf{T})^\mathsf{T}$ for $j = 1, \ldots, J$.*

The proof of this theorem can be found in Section S5.1 of the Supplemental Material. Note that, compared to the model in Tabelow, Voss and Polzehl (2012), model (2) allows for different eigenvalues and shapes of the tensors within a voxel, and thus is much more flexible.

3.2. *Parameter estimation using maximum likelihood* (*ML*). We first consider the case when J is known and delay the selection of J to Section 3.3. By assuming Gaussian additive noise on both real and imaginary parts of the complex signal, the observed signal intensity can be modeled as [see, e.g., Zhu et al. (2007)]

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

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

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

Moreover, we assume the noise $\boldsymbol{\epsilon}(\mathbf{u})$'s are independent across different voxels and gradient directions.

Under the Rician noise assumption, the log-likelihood of $\boldsymbol{\gamma}$ in model (2) is

(3)
$$\begin{aligned}
l(\boldsymbol{\gamma}) &= \sum_{\mathbf{u}\in\mathcal{U}} \log\bigg[\frac{S(\mathbf{u})}{\sigma^2} \exp\bigg\{ -\frac{S^2(\mathbf{u}) + \bar{S}^2(\mathbf{u})}{2\sigma^2} \bigg\} I_0\bigg\{ \frac{S(\mathbf{u})\bar{S}(\mathbf{u})}{\sigma^2} \bigg\} \bigg] \\
&= \sum_{\mathbf{u}\in\mathcal{U}} \bigg[\log\bigg\{ \frac{S(\mathbf{u})}{\sigma^2} \bigg\} - \frac{S^2(\mathbf{u}) + \bar{S}^2(\mathbf{u})}{2\sigma^2} + \log I_0\bigg\{ \frac{S(\mathbf{u})\bar{S}(\mathbf{u})}{\sigma^2} \bigg\} \bigg],
\end{aligned}$$

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. The ML estimate is obtained through maximizing (3). Although the above new parametrization avoids the identifiability issue, the likelihood function usually has multiple local maxima, which makes the computation of the ML estimate difficult and unstable.

The method that we used to overcome this issue can be briefly described as follows. We first develop an approximation of model (2) whose likelihood can be globally maximized via a grid search. We utilize the geometry of the problem so that the grid search can be done efficiently. Then we use the ML estimate of this approximated model as the initial value in a gradient method to obtain the ML estimate of model (2). This method provides very reliable estimates. To speed up the pace of this article, its full description is given in Section S1 of the Supplemental Material.

3.3. *Selecting the number of tensor components* J. Common model selection methods can be applied to select the number of components J. Results from extensive numerical experiments suggest that the Bayesian information criterion (BIC) [Schwarz (1978)] is a good choice; see Section S2 of the Supplemental Material.

Under model (2), each tensor corresponds to four free scalar parameters since $\mathbf{m}_j$ is characterized by two free scalar parameters. The BIC for a model with I tensors is

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

where m is the number of gradient directions and $\hat{\boldsymbol{\gamma}}(I)$ is the ML estimate of $\boldsymbol{\gamma}$ under I tensors. Then J is chosen as $\hat{J} = \text{argmin}_{I \in \{1, \ldots, \tilde{I}\}} \text{BIC}(I)$, where $\tilde{I}$ is a prespecified upper bound for the number of components. Based on our experience, $\tilde{I} = 4$ is a reasonable choice.

In practice, there are voxels with no major diffusion directions. This corresponds to the case where there is only one isotropic tensor. In the case of an isotropic tensor, (2) reduces to $\bar{S}(\mathbf{u}) = \mathbf{S}_0 \tau_1$. Thus, there is only one parameter τ_1. We write the corresponding likelihood function as $\tilde{l}$ and denote the ML estimate of τ_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 (4), 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 fractional anisotropy (FA) [see, e.g., Mori (2007)],

$$\text{(5)} \qquad \qquad \text{FA} = \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}},$$

where λ_1, λ_2 and λ_3 are the eigenvalues of the corresponding tensor in the single tensor model, to conduct an initial screening to speed up the whole procedure. The FA value lies between zero and one and the larger it is, the more anisotropic the water diffusion is at the corresponding voxel. Thus, we first remove voxels with very small FA values and then apply the BIC approach over those suspected anisotropic voxels. Note that such removal is mainly for reducing computational cost, as a typical dMRI data set consists of hundreds of thousands of voxels. From our experiences, this has little effect on the final tracking results. We also note that the proposed framework including selection of J can be applied without such removal if enough computational resources are available.

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

4. Spatial smoothing of diffusion directions.   Although model (2) provides a better modeling than the 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.

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

4.1. *Smoothing along a single fiber.*   This subsection considers the simpler situation where there is only one homogeneous population of diffusion directions; that is, there is only one single fiber bundle without crossing. Write T as the total number of estimated diffusion directions from all voxels and $\{\hat{\mathbf{m}}_k : k = 1, \ldots, T\}$ as the set of all estimated diffusion directions. Also, write $\mathbf{s}_k$ as the corresponding voxel location associated with $\hat{\mathbf{m}}_k$. Note that some $\mathbf{s}_k$'s share the same value, as some voxels contain multiple estimated directions. Following the idea of kernel smoothing on Euclidean space [e.g., Fan and Gijbels (1996)], the smoothing estimate at voxel $\mathbf{s}_0$ is defined as a weighted Karcher mean of the neighboring direction vectors:

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

where $w_i = K_{\mathbf{H}}(\mathbf{s}_i - \mathbf{s}_0)$'s are spatial weights and the metric d^* is defined as

$$
(7) \qquad\qquad d^*(\mathbf{u}, \mathbf{v}) = \arccos(|\mathbf{u}^{\mathsf{T}}\mathbf{v}|), \qquad \mathbf{u}, \mathbf{v} \in \mathcal{M};
$$

that is, $d^*(\mathbf{u}, \mathbf{v})$ is the acute angle between $\mathbf{u}$ and $\mathbf{v}$. The weights w_i's 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 3D kernel function satisfying $\int K(\mathbf{s}) \, d\mathbf{s} = 1$, and $\mathbf{H}$ is a 3×3 bandwidth matrix. In our numerical work, we choose $K(\cdot)$ as the standard Gaussian density, and set $\mathbf{H} = h\mathbf{I}_3$, where h is chosen using the cross-validation (CV) approach described in Section S3 of the Supplemental Material. We adopt the leave-one-out CV idea to develop an ordinary CV score and two robust CV scores. Their practical performances are reported in Section S6 of the Supplemental Material.

4.2. *Smoothing over multiple fibers.*   When there are crossing fibers in a voxel $\mathbf{s}_0$, the above smoothing procedure will not work well. To address this issue, we first cluster the neighboring estimated directions of $\mathbf{s}_0$ into groups that correspond to different fiber populations. Then we apply the above smoothing procedure to each individual 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
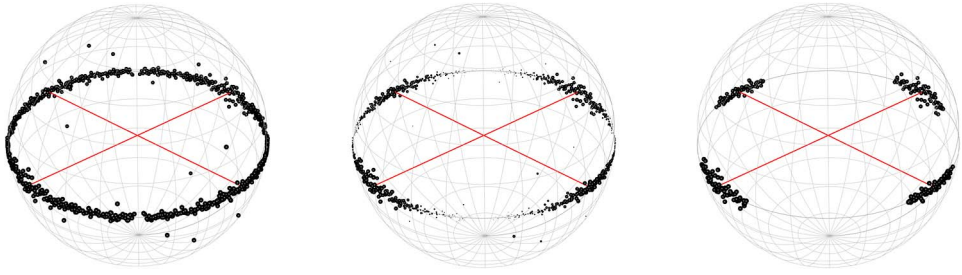
FIG. 2.   *Direction clustering. Left: all estimated directions. Middle: sizes of all estimated directions proportional to weights. Right: estimated directions with weights larger than a threshold. Red lines represent underlying true directions.*

than a threshold. By filtering out these voxels, we obtain tighter and better sepa-rated clusters of directions. Moreover, such voxels have little effect on smoothing due to their small weights. The artificial data set displayed in Figure 2 provides an illustrative example. Each 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 (6). Last, 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 directions much easier.

Next we need a clustering strategy to choose the number of clusters adaptively. With the distance metric (7), one can define a dissimilarity matrix for a set of di-rections and make use of a generic clustering algorithm. Our choice is the Partition Around Medoids (PAM) [Kaufman and Rousseeuw (1990)] due to its simplicity. Also, we apply the average silhouette [Rousseeuw (1987)] to choose the number of clusters; see Algorithm S3 of the Supplemental Material. 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 correspond-ing 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 S4.

4.3. *Theoretical results.*   This subsection derives asymptotic properties of the proposed direction smoothing estimator. Note that since the space of direction vec-tors has a non-Euclidean geometry, the theoretical framework is different from that of classical smoothing estimators. Without loss of generality, suppose we observe $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{M}$ at spatial locations $\mathbf{s}_1, \ldots, \mathbf{s}_n$, respectively. Let $\mathcal{V}$ be the 3D 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 $\mathbb{R}P^2$.

The theoretical results below were derived under the more convenient random design where $\mathbf{s}_i$'s are independently and identically sampled from a distribution with density f_S. The below theorem (Theorem 2) remains valid even under a fixed, regular design setting, with the number of grid points increasing to infinity. In this case, in the statement of the asymptotic formulae and their proofs, the density function f_S is replaced with a constant-valued function, representing a regular grid, with corresponding changes wherever derivatives of f_S appear.

Given a spatial location $\mathbf{s}_0$, our target is to estimate $\mathbf{v}_0$, namely, the diffusion direction at $\mathbf{s}_0$, which is defined as the minimizer of $\mathbb{E}\{d^{*2}(\mathbf{V}, \mathbf{v}) | \mathbf{S} = \mathbf{s}_0\}$ over $\mathbf{v}$, where $d^*(\mathbf{u}, \mathbf{v}) = \arccos(|\mathbf{u}^\mathsf{T} \mathbf{v}|)$. Here $\mathbf{V}$ is a random unit vector representing a random diffusion direction and the expectation is taken over $\mathbf{V}$ conditional on $\mathbf{S} = \mathbf{s}_0$, where $\mathbf{S}$ represents the location of where $\mathbf{V}$ is observed. For simplicity, we assume $\mathbf{s}_i \in \mathbb{R}$ and write it as s_i thereafter. Thus, our estimator (6) 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 n is the number of diffusion direction vectors and $K_h(\cdot) = K(\cdot/h)/h$. Here, with slight notation abuse, $K(\cdot)$ represents a one-dimensional kernel function throughout the theoretical developments.

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

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

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

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

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

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

WONG ET AL.

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 $\phi(\mathbf{v}) = (\omega_1, \omega_2)^\mathsf{T}$. In addition, we define

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

which aligns $\mathbf{v}$ with $\mathbf{v}_0$, and $d(\boldsymbol{\omega}, \boldsymbol{\theta}) = d^*(\phi^{-1}(\boldsymbol{\omega}), \phi^{-1}(\boldsymbol{\theta}))$ for $\boldsymbol{\omega}, \boldsymbol{\theta} \in \mathbb{R}^2$. Note that, for any $\mathbf{v}, \mathbf{p} \in \mathcal{V}$, we have $d(\tilde{\phi}(\mathbf{v}), \tilde{\phi}(\mathbf{p})) = d^*(\mathbf{v}, \mathbf{p})$, where $\tilde{\phi} = \phi \circ \rho_{\mathbf{v}_0}$. Here $\tilde{\phi}(\mathbf{v})$ first aligns a direction $\mathbf{v}$ with the true diffusion direction $\mathbf{v}_0$ and then represents it by its logarithmic coordinate.

We now present the asymptotic results. Now, write $\boldsymbol{\theta}_i = \tilde{\phi}(\mathbf{v}_i)$ for $i = 1, \ldots, n$ 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 $\boldsymbol{\psi}_j(\boldsymbol{\omega}, \boldsymbol{\theta})$ be the jth order derivative of ψ with respect to $\boldsymbol{\theta}$ for $j = 1, 2$. Let $\mathbf{m}(s) = (m_1(s), m_2(s))^\mathsf{T} = \mathbb{E}(\boldsymbol{\theta}_1 | S_1 = s)$ and $\boldsymbol{\Sigma}(s) = [\Sigma_{jk}(s)]_{1 \leq j,k \leq 2} = \text{Var}(\boldsymbol{\theta}_1 | S_1 = s)$. Also, denote $\boldsymbol{\Psi}(s) = [\Psi_{jk}(s)]_{1 \leq j,k \leq 2} = \mathbb{E}[\boldsymbol{\psi}_2(\boldsymbol{\theta}_1, \boldsymbol{\theta}_0) | S_1 = s]$.

Under the assumptions 1–10 laid out in Section S5.2 of the Supplemental Material which are all standard technical conditions (except for assumption 1 which is to ensure the representation of the geodesic distance as a function of the working coordinate system), we have the following theorem.

THEOREM 2. *Let $M_n(\boldsymbol{\theta}) = \sum_{i=1}^n h K_h(S_i - s_0) d^2(\boldsymbol{\theta}_i, \boldsymbol{\theta})$, and assume assumptions 1–10 hold.*

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

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

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 Theorem 2 can be found in Section S5.2 of the Supplemental Material.

5. Fiber tracking. For dMRI, fiber tractography can be classified as a deterministic and probabilistic method. Deterministic methods [e.g., Mori and van Zijl (2002), Mori et al. (1999), Weinstein, Kindlmann and Lundberg (1999)] track fiber bundles by utilizing the principal eigenvectors of tensors, while probabilistic
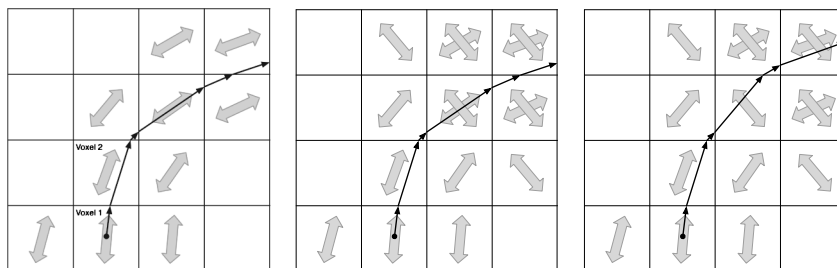
FIG. 3. *Left*: *Demonstration of the proposed algorithm in the single fiber region. Middle*: *Demonstration of the proposed algorithm in the crossing fiber region. Right*: *Demonstration of the proposed algorithm in the case of absence of viable directions.*

methods [e.g., Friman, Farneback and Westin (2006), Koch, Norris and Hund-Georgiadis (2002), Parker and Alexander (2003)] use the probability density of diffusion orientations. Most deterministic methods assume one single diffusion tensor in each voxel, and hence are unable to handle voxels with crossing fibers. In view of this, this section develops a deterministic tracking algorithm that allows for multiple or no principal diffusion directions in a voxel.

The proposed algorithm can be seen as a generalization of the popular Fiber Assignment by Continuous Tracking (FACT) [Mori et al. (1999)] algorithm. A brief description of FACT is as follows. Tracking starts at the center of a voxel (Voxel 1 in Figure 3, left panel) and continues in the direction of the estimated diffusion direction. When it enters the next voxel (Voxel 2 in Figure 3, left panel), the track changes its direction to align with the new diffusion direction and so on. This 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 cannot determine 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 manner. 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 prespecified 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 prespecified 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 prespecified number of voxels can be found.

The detailed tracking algorithm is described in Algorithm S5 in the Supplemental Material.

As for the choice of starting voxels, also known as seeds, there are two common strategies. One can choose seeds based on tracts of interest and start 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 the 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.

Combining the voxel-wise estimation method in Section 3 and the direction smoothing procedure in Section 4 gives the proposed DiST method.

6. Simulation study. Extensive simulation experiments have been conducted to evaluate the practical performances of DiST. They are reported in Section S6 of the Supplemental Material. Overall, the DiST method provided highly promising results.

7. 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). 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 (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 the b-value set as 1000 s/mm^2. In addition, there are 5 $b0$ images (corresponding to $b = 0$), forming in total 46 measurements for each of the $256 \times 256 \times 59$ voxels. To implement our technique, we require estimates of $S_0(\mathbf{s})$'s and σ. We first estimate $S_0(\mathbf{s})$ and $\sigma(\mathbf{s})$ for each voxel by ML estimation based on the 5 $b0$ images. Then we fix σ 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 σ 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, Larsson and Wedeen (2000)]. See Figure 4 (Left) for a fiber orientation color map of one of the five xy-planes. Within the whole focused 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
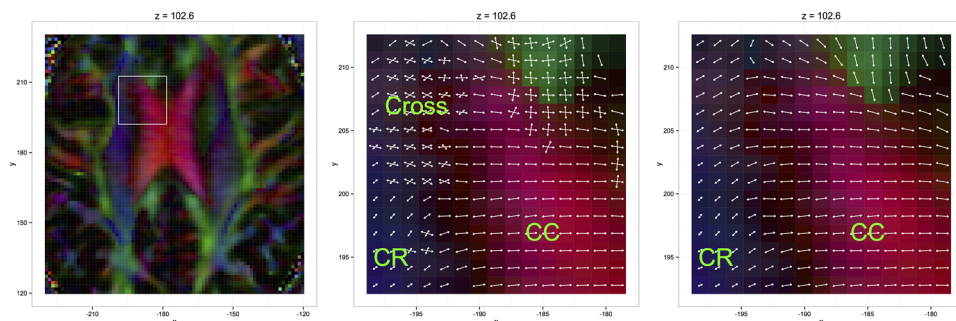
FIG. 4. *Left*: *the fiber orientation color map* (*based on the single tensor model*). *The focused region is indicated by a white rectangular box. Middle* (*from* DiST-mcv) *and right* (*from single tensor model*): *The projection of fiber directions to the* xy-*plane at* $z = 102.6$ *for illustration of crossing fibers.* (*The five* xy-*planes that we focus on have reference values* $z = 99.9, 102.6, 105.3, 108, 110.7$ *from bottom to top.*) *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.*

summarized in Table 1. 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, which represents the DiST procedure with h chosen by the median cross-validation score (see Sections S3 and S6 of the Supplemental Material), and those from the single tensor model estimation. For visualization purposes, we present the longest 900 tracts in Figure 5. 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. The tracking result based on DiST-mcv shows clear crossing between the mediolateral fiber and the superoinferior fiber (in the figure, the crossing of the red and blue fiber tracts). From neuroanatomic atlases and previous studies, Wiegell, Larsson and Wedeen (2000) conclude that there are several fiber populations with crossing structure in this conjunction region of CC and CR, which matches with

TABLE 1
Number of voxels with different estimated number of diffusion directions

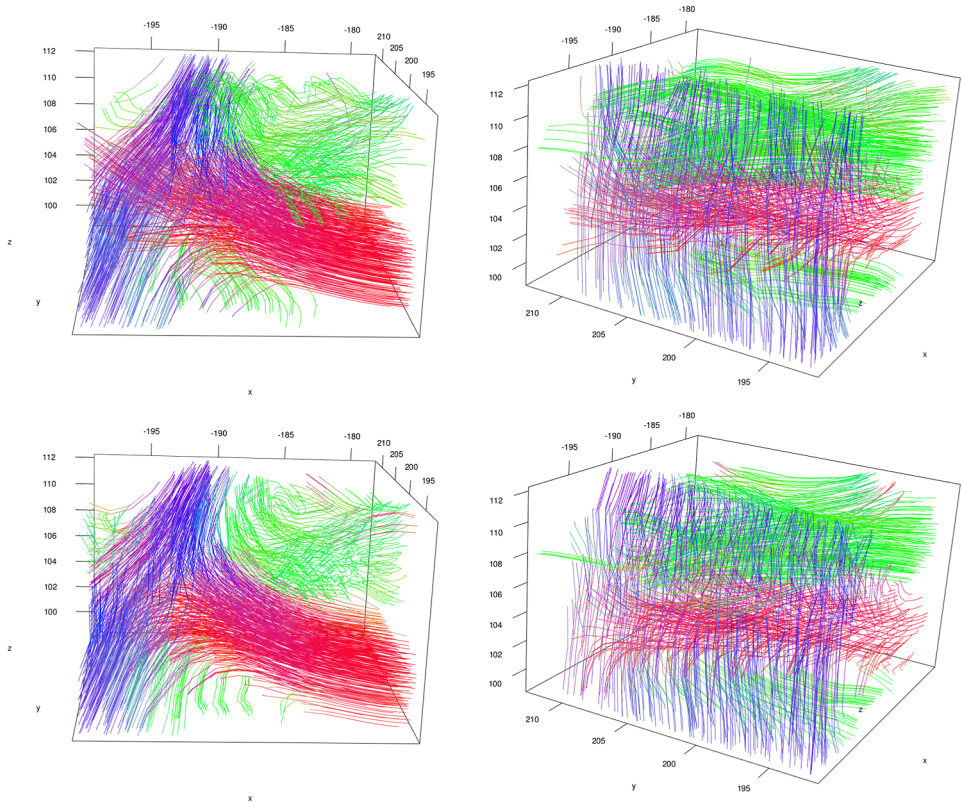|	**Number of diffusion directions**					
	0	**1**	**2**	**3**	**4**	**Total**
Voxel-wise estimation	37	476	589	23	0	1125
Smoothing	37	476	593	19	0	1125

FIG. 5. *Top*: *The longest* 900 *tracks using* DiST-mcv. *Bottom*: *The longest* 900 *tracks using the single tensor model. The left and right figures correspond to different viewing angles.*

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 4 shows the locations of the CC, the CR and the region of crossing fibers (Cross). One can see that estimated directions based on DiST-mcv reproduces the crossing fiber structures between the CC and the CR, while the result based on the 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 on DiST and the single fiber model produce clear and sensible reconstruction of the cingulum bundle. All these features match with neuroanatomic atlases and provide a good demonstration of our proposed method.

As shown by Figures 4 and 5, when comparing with the results obtained by the single tensor model, DiST produces more biologically sensible and interpretable

tracking results. This provides more reliable information on brain connectivity and in turn could lead to better understanding of neuro-degenerative diseases such as Alzheimer's disease and autism, as well as better detection of brain abnormality, such as deformation and neuron loss in white matter regions.

8. Discussion. Using tensor estimation to resolve crossing fiber can be problematic due to the inability of estimating multiple diffusion directions by the single tensor model and the nonidentifiability issue in the multi-tensor model. In this paper, we take a different route by focusing on the estimation of diffusion directions rather than the 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 (see Section S6 of the Supplemental Material) and real data analysis.

The procedure we presented works well even with a moderate number of gradient directions (a few tens), as long as the number of distinct crossing fibers within a voxel is not larger than three. With HARDI data, which can have up to a couple of hundred gradient directions, rather than modeling the direction distribution within a tensor framework, we can estimate the fiber orientation distribution nonparametrically [Descoteaux et al. (2007), Tuch (2004)].

Applying DiST to multiple images from ADNI (either from the same subject over time or from multiple subjects) and then relating the tracking results with clinical outcomes such as cognitive measures would provide valuable information about the role of white matter connectivity in initiation and progression of Alzheimer's disease and dementia. Although this is an important direction of research, it is beyond the scope of this paper which focuses on developing a statistical procedure to denoise dMRI data in order to provide better tracking results. We plan to explore more sophisticated applications of the proposed procedure in our future research.

SUPPLEMENTARY MATERIAL

Supplemental document for "Fiber direction estimation, smoothing and tracking in diffusion MRI" (DOI: 10.1214/15-AOAS880SUPP; .pdf). This supplement provides additional material.

REFERENCES

ARSIGNY, V., FILLARD, P., PENNEC, X. and AYACHE, N. (2006). Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.* **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. *Magn. Reson. Imaging Clin. N. Am.* **17** 175–204.

BASSER, P. J., PAJEVIC, S., PIERPAOLI, C., DUDA, J. and ALDROUBI, A. (2000). In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med.* **44** 625–632.

BEAULIEU, C. (2002). The basis of anisotropic water diffusion in the nervous system—A technical review. *NMR Biomed.* **15** 435–455.

BEHRENS, T. E. J., WOOLRICH, M. W., JENKINSON, M., JOHANSEN-BERG, H., NUNES, R. G., CLARE, S., MATTHEWS, P. M., BRADY, J. M. and SMITH, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.* **50** 1077–1088.

BEHRENS, T. E. J., BERG, H. J., JBABDI, S., RUSHWORTH, M. F. S. and WOOLRICH, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34** 144–155.

CARMICHAEL, O., CHEN, J., PAUL, D. and PENG, J. (2013). Diffusion tensor smoothing through weighted Karcher means. *Electron. J. Stat.* **7** 1913–1956. MR3084676

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. *Neuropsychol. Rev.* **20** 209–225.

DESCOTEAUX, M., ANGELINO, E., FITZGIBBONS, S. and DERICHE, R. (2007). Regularized, fast, and robust analytical Q-ball imaging. *Magn. Reson. Med.* **58** 497–510.

FAN, J. and GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications. Monographs on Statistics and Applied Probability* **66**. Chapman & Hall, London. MR1383587

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. *Magn. Reson. Med.* **34** 910–914.

HOSEY, T., WILLIAMS, G. and ANSORGE, R. (2005). Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magn. Reson. Med.* **54** 1480–1489.

KAUFMAN, L. and ROUSSEEUW, P. J. (1990). *Finding Groups in Data*: *An Introduction to Cluster Analysis*. Wiley, New York. MR1044997

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*. Elsevier, Amsterdam.

MORI, S. and VAN ZIJL, P. C. M. (2002). Fiber tracking: Principles and strategies—A technical review. *NMR Biomed*. **15** 468–480.

MORI, S., CRAIN, B. J., CHACKO, V. P. and VAN ZIJL, P. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology* **45** 265–269.

MUKHERJEE, P., BERMAN, J. I., CHUNG, S. W., HESS, C. P. and HENRY, R. G. (2008). Diffusion tensor MR imaging and fiber tractography: Theoretic underpinnings. *AJNR Am. J. Neuroradiol*. **29** 632–641.

NIMSKY, C., GANSLANDT, O. and FAHLBUSCH, R. (2006). Implementation of fiber tract navigation. *Neurosurgery* **58** ONS–292–303; discussion ONS–303–4.

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, Berlin.

PENNEC, X., FILLARD, P. and AYACHE, N. (2006). A Riemannian framework for tensor computing. *Int. J. Comput. Vis*. **66** 41–66.

ROUSSEEUW, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math*. **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. IEEE, Rotterdam.

SCHWARTZMAN, A., DOUGHERTY, R. F. and TAYLOR, J. E. (2008). False discovery rate analysis of brain diffusion direction maps. *Ann. Appl. Stat*. **2** 153–175. MR2415598

SCHWARZ, G. (1978). Estimating the dimension of a model. *Ann. Statist*. **6** 461–464. MR0468014

SPORNS, O. (2011). *Networks of the Brain*. MIT Press, Cambridge, MA.

TABELOW, K., VOSS, H. U. and POLZEHL, J. (2012). Modeling the orientation distribution function by mixtures of angular central Gaussian distributions. *J. Neurosci. Methods* **203** 200–211.

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.

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.

TUCH, D. S. (2002). Diffusion MRI of complex tissue structure. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

TUCH, D. S. (2004). Q-ball imaging. *Magn. Reson. Med*. **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. *Magn. Reson. Med*. **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 imaging1. *Radiology* **217** 897–903.

WONG, R. K. W., LEE, T. C. M., PAUL, D. and PENG, J. (2016). Supplement to "Fiber direction estimation, smoothing and tracking in diffusion MRI." DOI:10.1214/15-AOAS880SUPP.

YUAN, Y., ZHU, H., LIN, W. and MARRON, J. S. (2012). Local polynomial regression for symmetric positive definite matrices. *J. R. Stat. Soc. Ser. B. Stat. Methodol*. **74** 697–719. MR2965956

ZHU, H., ZHANG, H., IBRAHIM, J. G. and PETERSON, B. S. (2007). Statistical analysis of diffusion
    tensors in diffusion-weighted magnetic resonance imaging data. *J. Amer. Statist. Assoc.* **102** 1085–
    1102. MR2412530

R. K. W. WONG   T. C. M. LEE
DEPARTMENT OF STATISTICS   D. PAUL
IOWA STATE UNIVERSITY   J. PENG
2218 SNEDECOR HALL   DEPARTMENT OF STATISTICS
AMES, IOWA 50011   UNIVERSITY OF CALIFORNIA, DAVIS
USA   4118 MATHEMATICAL SCIENCES BUILDING
E-MAIL: raywong@iastate.edu   ONE SHIELDS AVENUE
                                                DAVIS, CALIFORNIA 95616
                                                USA
                                                E-MAIL: tcmlee@ucdavis.edu
                                                        debpaul@ucdavis.edu
                                                        jiepeng@ucdavis.edu