

# Gray matter parcellation constrained full brain fiber bundling with diffusion tensor imaging

Qing Xu<sup>a)</sup>

Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, Tennessee 37232-2310 and Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, Tennessee 37235-1826

Adam W. Anderson and John C. Gore

Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, Tennessee 37232-2310 and Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee 37235-1631

Zhaohua Ding

Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, Tennessee 37232-2310; Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, Tennessee 37235-1826; and Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee 37235-1631

(Received 9 June 2012; revised 29 May 2013; accepted for publication 31 May 2013; published 1 July 2013)

**Purpose:** Studying white matter fibers from diffusion tensor imaging (DTI) often requires them to be grouped into bundles that correspond to coherent anatomic structures, particularly bundles that connect cortical/subcortical basic units. However, traditional fiber clustering algorithms usually generate bundles with poor anatomic correspondence as they do not incorporate brain anatomic information into the clustering process. On the other hand, image registration-based bundling methods segment fiber bundles by referring to a coregistered atlas or template with pre-labeled anatomic information, but these approaches suffer from the uncertainties introduced from misregistration and fiber tracking errors and thus the resulting bundles usually have poor coherence. In this work, a bundling algorithm is proposed to overcome the above issues.

**Methods:** The proposed algorithm combines clustering- and registration-based approaches so that the bundle coherence and the consistency with brain anatomy are simultaneously achieved. Moreover, based on this framework, a groupwise fiber bundling method is further proposed to leverage a group of DTI data for reducing the effect of the uncertainties in a single DTI data set and improving cross-subject bundle consistency.

**Results:** Using the Montreal Neurological Institute template, the proposed methods are applied to building a full brain bundle network that connects cortical/subcortical basic units. Based on several proposed metrics, the resulting bundles show promising bundle coherence and anatomic consistency as well as improved cross-subject consistency for the groupwise bundling.

**Conclusions:** A fiber bundling algorithm has been proposed in this paper to cluster a set of whole brain fibers into coherent bundles that are consistent to the brain anatomy. © 2013 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [<http://dx.doi.org/10.1118/1.4811155>]

Key words: bundling, white matter fibers, diffusion tensor imaging

## I. INTRODUCTION

Magnetic resonance diffusion tensor imaging (DTI) has become a primary neuroimaging technique for noninvasive characterization of the structure and architecture of the human brain *in vivo*.<sup>1,2</sup> Since its first introduction in the early 1990s, this technique has been widely used to elucidate the structural basis for brain function both in healthy<sup>3,4</sup> and disease conditions.<sup>5-11</sup> As alterations of brain function are necessarily accompanied by structural changes in the underlying neural circuits, characterizing these changes may offer valuable insights into the pathogenesis, extent or progression of the disease, thus holding the potential of guiding therapeutic interventions.

A large body of DTI-based structural studies rests on fiber tractography,<sup>12</sup> in which local tissue orientations depicted by DTI are combined and used to delineate and visualize the courses of fiber tracts at a macroscopic scale. Typically, individual fibers are first reconstructed, then grouped into coherent fiber bundles and assigned labels that have anatomic interpretations.<sup>13</sup> Anatomically labeled fiber bundles provide the basis for quantitative analysis of fiber structures and comparisons of structural properties across different populations in a physiologically meaningful manner.

To group individual fibers into coherent and anatomically meaningful bundles, a plethora of fiber bundling methods have been proposed to date. These methods come in three different flavors: manual,<sup>14-16</sup> clustering based,<sup>17-25</sup>

and knowledge-based bundling.<sup>26–29</sup> Manual fiber bundling groups fibers using one or more regions of interest (ROIs) placed by an experienced operator as constraints to the fiber courses.<sup>14–16</sup> Performance of the manual bundling hinges significantly on the accuracy of identifying anatomic landmarks or structural boundaries. Evidently, such a bundling method suffers from the drawback of low time efficiency, high operator biases, and poor intraoperator reproducibility. The situation is made worse by fuzzy or missing structural boundaries when inspected visually, which renders accurate and precise definitions of ROIs nontrivial.

The difficulties and limitations with manual bundling have been greatly ameliorated by automated fiber bundling, for which a rich literature of clustering algorithms has been developed in the recent past.<sup>17–25</sup> With no need for prior ROI definitions, these methods group fibers automatically according to their similarities in geometrical properties such as shapes, locations, or other attributes. While the clustering based methods are able to generate coherent fiber bundles efficiently by means of fully automated algorithms, they are purely data driven, and hence bundle fibers without due regard to their anatomical plausibility.

A logical compromise between the manual- and clustering-based bundling methods is knowledge-based bundling.<sup>13,26–29</sup> A common knowledge-based method uses a template of prototypical fiber bundles that are constructed from a subset of data.<sup>13,28</sup> To bundle reconstructed fiber tracts for a given subject, the template is first transformed into the subject's native space via a registration procedure. All the fibers are then classified in reference to the prototypical bundles in the template. Because the prototypical bundles are usually defined manually, it would be pretty awkward to perform full brain fiber bundling, and quite problematic as well given the fact that our knowledge on distributions of fiber connection routes in the human brain is far from complete. A ramification of knowledge-based methods is atlas-based fiber bundling.<sup>26,27</sup> Instead of defining the courses of prototypical fiber bundles as above, the atlas based method typically bundles fibers with parcellated gray matter as constraints to fiber terminals. Of particular note, due to the availability of complete gray matter parcellations,<sup>30</sup> the atlas-based method has the inherent capability of bundling fibers of the entire brain automatically. A salient advantage of both the template- and atlas-based methods is that expert knowledge is incorporated into the bundling process, and thus all resulting bundles in principle bear anatomical interpretations naturally. Meanwhile, these methods have largely alleviated the problems of low time efficiency and poor inter- and intraoperator reproducibilities encountered in manual bundling. It should be pointed out, however, that essential to the accuracy of the knowledge-based methods is the performance of the registration algorithm used to transform the template or atlas into the subject's native space.

In our earlier work,<sup>29</sup> we developed a unified fiber bundling and registration (UFIBER) algorithm using a template based approach. A set of well-established fiber bundles were chosen to debut the process and performance of the algorithm. In the present study, we extend the UFIBER algo-

rithm so that fibers that connect cortical/subcortical units in the entire brain can be bundled in a fully automatic manner and consistently across subjects. Similar to the work in Refs. 26 and 27, an atlas of gray matter parcellation defined in the standard MNI space is used.<sup>30</sup> Obviously, using parcellated gray matter in the form of individual functional regions to constrain fiber bundling has direct benefits to integrated structural-functional studies of the human brain.<sup>31</sup> But unlike Refs. 26 and 27, which bundle fibers with some heuristic proximity rules, we pose the problem of full brain bundling as an optimization process cast with a more rigorous mathematical framework, in which constraints from gray matter region parcellation and coherence of white matter fiber bundling are simultaneously considered and optimized. This relaxes the requirements for accurate registration, which is often hard to satisfy if at all owing to practical difficulties such as poor image contrast between gray matter regions and fairly complex structures therein. Furthermore, consistent bundling across subjects offers great advantages in subsequent statistical analysis of fiber bundle properties as clinical studies are almost exclusively population based.

In the remainder of this paper, a gray matter constrained bundle model that integrates constraints from anatomical parcellation into a statistical fiber bundle model is first proposed in Sec. II. This is followed by a fiber bundling algorithm that generates both geometrically coherent and anatomically consistent bundles for a single subject in Sec. III. In Sec. IV, a groupwise bundling approach that improves cross-subject bundling consistency is described. Evaluation of the proposed algorithm is presented in Sec. V. Finally, the main contributions of this work and potential future directions are discussed in Sec. VI.

## II. GRAY MATTER CONSTRAINED BUNDLE MODEL

Given a brain atlas, in which gray matter is fully partitioned into  $L$  disjoint ROIs and each ROI is associated with a label  $l = 1, 2, \dots, L$ , we assume a registration procedure transforms all the labeled ROIs into a subject's space using homomorphic mapping. Let  $\mathbf{r}$  denote the coordinates of a terminal point of a fiber  $\mathbf{x}$  in the subject's space, and  $l(\mathbf{r})$  denote the ROI label of  $\mathbf{r}$ . Therefore, a fiber  $\mathbf{x}$  can be represented by a space curve connecting  $l(\mathbf{r1})$  and  $l(\mathbf{r2})$ , where  $\mathbf{r1}$  and  $\mathbf{r2}$  are the two terminals of fiber  $\mathbf{x}$ , respectively.

### II.A. Uncertainties

The variables  $\mathbf{r}$  and  $l(\bullet)$  are observations of their true values, and thus contain uncertainty due to observational errors. Uncertainties of  $\mathbf{r}$  arise from the fiber tracking procedure, which has been widely recognized as being highly susceptible to image noise.<sup>32,33</sup> Fundamentally, the noise susceptibility is attributable to the integrative nature of fiber tracking, which leads to greater errors and uncertainties toward the terminals of fiber tracts.<sup>34</sup> Moreover, the commonly used echo-planar imaging (EPI) sequences for DTI data acquisitions typically generate images with poor signal-to-noise ratio (SNR), creating a fair amount of uncertainty in the local fiber orientations

estimated. Although measures can be taken to partly suppress the effect of noise prior to or during fiber tracking,<sup>35–37</sup> the uncertainty from noise by nature always exists, albeit to a lesser extent after noise suppression. In addition, complex structures of white matter fibers at certain locations, along with inadequacies of the tensor model to capture them and the available fiber tracking algorithms, add another source of uncertainties to fiber tracts.<sup>38</sup> Of note, complex fiber structures may, in principle, be resolved by using high angular resolution diffusion imaging (HARDi) (Refs. 39 and 40) or handled by probabilistic fiber tracking,<sup>41–43</sup> but these techniques have not become routine utilities owing to other complications they involve.

Uncertainties of  $l(\bullet)$  are caused by the fact that registration of brain magnetic resonance (MR) images is particularly error-prone. As briefly alluded to before, there are two primary difficulties in the registration: (1) there is a virtual lack of intensity contrast between neighboring functional regions in the gray matter, which provides no intensity cues for registering structures therein; (2) structures in the gray matter are pretty complex and have considerable anatomical variations among individuals, rendering precise matching of structural details a daunting task if possible at all.<sup>44</sup> Image noise and geometric distortions may contribute to the complexity of gray matter registration as well.

## II.B. Gray matter projection model

Given a subject, whose gray matter is parcellated into  $L$  disjoint, labeled regions via atlas mapping, the likelihood of a fiber terminal  $\mathbf{r}$  belonging to ROI  $l = 1, 2, \dots, L$  is modeled with a Gaussian distribution as follows:

$$p(\mathbf{r}|l) = (2\pi\sigma_{\text{ROI}}^2)^{-3/2} \times \exp\left(-\frac{(\mathbf{r} - \mathbf{v}(l, \mathbf{r}))^T(\mathbf{r} - \mathbf{v}(l, \mathbf{r}))}{2\sigma_{\text{ROI}}^2}\right), \quad (1)$$

where  $\mathbf{v}(l, \mathbf{r})$  is a point in ROI  $l$  closest to  $\mathbf{r}$  and  $T$  denotes the transpose operation on a vector.

In Eq. (1), distance is used in modeling of terminal-to-ROI probabilities, and the terminals of a fiber are projected into their closest points in cortical or subcortical gray matter regions. Variance  $\sigma_{\text{ROI}}$ , which is isotropic in this model, is a parameter related to uncertainties in image registration and fiber tracking. It is set to a smaller value for more precise registration and tracking, reflecting the fact that terminal points are expected to lie closer to their true ROIs, and vice versa.

Brute-force evaluation of the probability  $p(\mathbf{r}|l)$  in Eq. (1) is computationally expensive. To improve time efficiency, a distance transform of ROI  $l$  is precomputed and stored in a function  $\phi_l(\bullet)$ . This avoids the time consuming process of searching in ROI  $l$  for the closest point to  $\mathbf{r}$ . Therefore,  $p(\mathbf{r}|l)$  can be evaluated as

$$p(\mathbf{r}|l) = (2\pi\sigma_{\text{ROI}}^2)^{-3/2} \exp\left(-\frac{(\phi_l(\mathbf{r}))^2}{2\sigma_{\text{ROI}}^2}\right). \quad (2)$$

## II.C. Gray matter constrained bundle model

For efficient parameter estimation, fibers in a bundle are modeled with a Gaussian distribution, as in previous works.<sup>22,29</sup> With a further assumption of point independence in a fiber, the probability of a fiber  $\mathbf{x}$  belonging to the bundle connecting  $l1$  and  $l2$  can be expressed as

$$p(\mathbf{x}|\mu_{l1,l2}, \sigma_{l1,l2}) = \prod_{i=1}^m (2\pi\sigma_{l1,l2}^2)^{-3/2} \times \exp\left(-\frac{(\mathbf{x}_i - \mu_{l1,l2,i})^T(\mathbf{x}_i - \mu_{l1,l2,i})}{2\sigma_{l1,l2}^2}\right), \quad (3)$$

where  $i$  indexes points along the fiber  $\mathbf{x}$  and  $\mu_{l1,l2}$  is the medial axis of the bundle connecting  $l1$  and  $l2$ . Typically,  $\mu_{l1,l2}$  is an unknown parameter that needs to be estimated by an optimization scheme. Here, an isotropic variance  $\sigma_{l1,l2}$  is used for all points along the fibers in all bundles (denoted as  $\sigma_{\text{bundle}}$ ), as some bundles may contain a small number of fibers, which makes estimation of point-specific covariance and its inverse computationally unstable.  $\mathbf{x}$  and  $\mu_{l1,l2}$  are resampled to an equal number of points and  $\mathbf{x}_i$  and  $\mu_{l1,l2,i}$  denote their  $i$ th point.

When clustering fibers solely based on this Gaussian model, fibers with similar spatial courses would be grouped into a bundle without respect to their terminal locations. To use gray matter parcellation as constraints to fiber terminals, the projection model of fiber terminal points can be included. Thus, for a fiber  $\mathbf{x}$  with terminal points  $\mathbf{r1}$  and  $\mathbf{r2}$ , the likelihood of  $\mathbf{x}$  being in the bundle connecting  $l1$  and  $l2$  is defined as the joint probability of fiber distribution in a bundle and gray matter parcellation distribution, i.e.,

$$p(\mathbf{x}|l1, l2) = p(\mathbf{x}|\mu_{l1,l2}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}) = p(\mathbf{x}|\mu_{l1,l2}, \sigma_{\text{bundle}})p(\mathbf{r1}, \mathbf{r2}|l1, l2), \quad (4)$$

where

$$p(\mathbf{r1}, \mathbf{r2}|l1, l2) = \max(p(\mathbf{r1}|l1)p(\mathbf{r2}|l2), p(\mathbf{r1}|l2)p(\mathbf{r2}|l1)).$$

Note that the correspondence between  $\mathbf{r1}$ ,  $\mathbf{r2}$ , and  $l1$ ,  $l2$  is chosen to be the one yielding a greater overall probability of terminal assignments evaluated using Eq. (2).

According to the model in Eq. (4), fibers are bundled considering both the coherence of fiber tracts and proximity of their terminals to designated ROIs. Trade-offs between the fiber bundle coherence and terminal proximity are regulated by relative magnitudes of the values of bundle variance  $\sigma_{\text{bundle}}$  and ROI variance  $\sigma_{\text{ROI}}$ .

## III. GRAY MATTER PARCELLATION CONSTRAINED FIBER BUNDLING FOR A SINGLE SUBJECT

In this section, an algorithm is proposed to bundle the entire brain of a single subject based on the gray matter constrained bundle model.

### III.A. Objective function

With a complete gray matter parcellation scheme, fibers of the entire brain can be modeled as a mixture of gray matter constrained bundle models

$$p(\mathbf{x}_{\text{all}}|\boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}) = \prod_{j=1}^J \sum_{l2=l1+1}^L \sum_{l1=1}^L p(\mathbf{x}_j|\boldsymbol{\mu}_{l1,l2}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}), \quad (5)$$

where  $j$  indexes a fiber in the set of  $J$  fibers over the entire brain and  $\sigma_{\text{bundle}}, \sigma_{\text{ROI}}$  are, respectively, the bundle and ROI variance, which in this work are fixed based on several investigation studies (see Sec. V.D below). Mixture proportions are set to one for all the bundles. This is because it is observed that the number of fibers in some bundles, such as those passing the corpus callosum, is significantly greater than that of other bundles and it is not desirable that fibers prefer to be assigned to bundles with big mixture proportions. Therefore, the only variable that needs to be estimated is the bundle medial axis  $\boldsymbol{\mu}_{l1,l2}$ . Note that the second summation is taken from  $l1 + 1$  to  $L$  for the index  $l2$  since we assume  $\boldsymbol{\mu}_{l1,l1}$  does not exist and  $\boldsymbol{\mu}_{l2,l1}$  and  $\boldsymbol{\mu}_{l1,l2}$  are identical.

Assuming each fiber is an independent sample from this distribution, an optimal  $\boldsymbol{\mu}$  can be estimated by maximizing the following likelihood:

$$\begin{aligned} \hat{\boldsymbol{\mu}} &= \arg \max_{\boldsymbol{\mu}} p(\mathbf{x}_{\text{all}}|\boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}) \\ &= \arg \max_{\boldsymbol{\mu}} \prod_{j=1}^J \sum_{l2=l1+1}^L \sum_{l1=1}^L p(\mathbf{x}_j|\boldsymbol{\mu}_{l1,l2}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}). \end{aligned} \quad (6)$$

### III.B. Expectation and maximization (EM) algorithm

To solve for an optimal  $\boldsymbol{\mu}$  in Eq. (6), the classic solution, EM algorithm,<sup>45</sup> is employed in this work. Given an initial  $\boldsymbol{\mu}^0$ , an expectation (E), step and maximization (M) step are alternatively performed until convergence. In the E step, based on the current estimation of  $\boldsymbol{\mu}^{n-1}$ , the fiber-to-bundle membership is computed with the formula below

$$m_{j,(l1,l2)}^n = \frac{p(\mathbf{x}_j|\boldsymbol{\mu}_{l1,l2}^{n-1}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})}{\sum_{l2=l1+1}^L \sum_{l1=1}^L p(\mathbf{x}_j|\boldsymbol{\mu}_{l1,l2}^{n-1}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})}, \quad (7)$$

where  $m_{j,(l1,l2)}^n$  represents the membership of fiber  $\mathbf{x}_j$  to bundle  $(l1, l2)$ . In essence,  $m_{j,(l1,l2)}^n$  is the likelihood of fiber  $\mathbf{x}_j$  to bundle  $(l1, l2)$  normalized across all the bundles.<sup>46,47</sup>

In the M step, using the estimated  $m_{j,(l1,l2)}^n$ , an optimal  $\boldsymbol{\mu}^n$  can be found by maximizing the likelihood below:

$$\begin{aligned} E(\boldsymbol{\mu}^n) &= \sum_{j=1}^J \sum_{l2=l1+1}^L \sum_{l1=1}^L m_{j,(l1,l2)}^n \\ &\quad \times \log(p(\mathbf{x}_j|\boldsymbol{\mu}_{l1,l2}^n, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})), \end{aligned} \quad (8a)$$

which leads to the update function for  $\boldsymbol{\mu}^n$ ,

$$\boldsymbol{\mu}_{l1,l2}^n = \frac{\sum_{j=1}^J m_{j,(l1,l2)}^n \mathbf{x}_j}{\sum_{j=1}^J m_{j,(l1,l2)}^n}. \quad (8b)$$

The above update scheme is essentially a weighted sum of all fibers using the membership. Using only the Gaussian bundle model [Eq. (3)], the membership is solely determined by the distance between individual fibers and their corresponding bundle medial axis, which leads to a minimization of the inbundle variation or the coherence of resulting bundles. Using the gray matter constrained bundle model, a decay term [Eq. (1)] is used to attenuate the membership based on the distances between fiber terminal points to the corresponding ROIs, which constrains the bundling process so that the bundle center would not deviate too much from its corresponding ROI. The inherent coherence minimization force is to correct inaccurate ROI labeling caused by image misregistration, while the parcellation constraint places a certain level of confidence on coregistered ROIs.

### III.C. Implementation issues

From Eqs. (7) and (8), it can be seen that the computational complexity is proportional to the total number of fibers, which is huge ( $\sim 40\,000$ ) for each subject. To reduce the number of fibers in the computation,  $m_{j,(l1,l2)}$  is set to zero if one of  $\mathbf{x}_j$ 's end points has a distance greater than  $3\sigma_{\text{ROI}}$  from ROI  $l1$  or  $l2$ . Then  $m_{j,(l1,l2)}$  would never be evaluated, nor would  $\mathbf{x}_j$  be involved in the update of  $\boldsymbol{\mu}_{l1,l2}$ . With this simplification, a significant amount of computation is avoided, leading to a CPU time of about 1 min per iteration (Intel Xeon 5150 2.66 GHz).

To find an initial estimate of  $\boldsymbol{\mu}^0$ , all fibers are first classified into a bundle that connects a pair of ROIs that their end points are closest to. Then  $\boldsymbol{\mu}^0$  for each pair of ROIs can be computed as the mean value of all the fibers in the bundle connecting them. Since all fibers have already been resampled to the same number of points, each point in the mean fiber is simply the average point of all the corresponding points in the bundle.

In each E step, fibers are assigned to the bundle with the maximum membership value. The algorithm automatically terminates when the total number of changes to fiber-to-bundle assignment or the number of iterations reaches a preset threshold (20 and 10, respectively, in this work).

## IV. GROUP CONSISTENT FIBER BUNDLING

In this section, an algorithm is proposed to bundle the entire brains for a group of subjects based on the gray matter constrained bundle model.

### IV.A. Objective function

The proposed groupwise bundling is formulated as estimating a common bundle model for a given set of subjects. Each fiber set is assumed to be an independent sample from

the underlying common bundle model. Such an assumption can be safely made when fiber sets are from a group of subjects with the same condition or even from the same subject. Let  $\mathbf{x}^s$  denote the  $s$ th fiber set in the subject group, where  $s$  could range from 1 to  $S$ , the total number of subjects in the group. Using the gray matter projection bundle model, the unknown parameter, bundle centroids  $\boldsymbol{\mu}$ , can be estimated using the below Bayesian rule,

$$\hat{\boldsymbol{\mu}} = \arg \max_{\boldsymbol{\mu}} \prod_{s=1}^S p(\mathbf{x}_{\text{all}}^s | \boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}),$$

where  $\sigma_{\text{bundle}}, \sigma_{\text{ROI}}$  are the bundle variance and ROI variance, respectively.

One problem with the above estimation is that each fiber set is in its own native space as subjects may be scanned in different positions in scanners. Therefore, it is necessary to transform all fiber sets into a common space, where the bundle model will be estimated. A T1 or fractional anisotropy (FA) based image registration procedure may be applied to this problem, but these approaches align fibers by only minimizing their image intensity difference without considering the underlying fiber directions. As T1 or FA intensity only indicates the type of the underlying tissue but nothing about fiber orientations, two voxels with different fiber orientations may be incorrectly considered to correspond simply due to their T1 or FA intensity similarity. To address this issue, fibers shall be utilized to make the alignment among a group of subjects. To avoid the alignment of whole fiber sets, which contain a huge number of fibers, bundle centroids can be efficiently aligned with each other through a nonrigid transformation. However, since the fiber bundles are yet to be estimated, there are no reliable bundles that can be used to make this alignment.

As the solutions to the groupwise bundling and spatial alignment could benefit each other, these two problems are coupled into a unified objective function and optimal transformations and bundling are jointly estimated. Let  $\mathbf{T}^s$  be a transformation that warps the fiber set  $\mathbf{x}^s$  into a common space. This can be cast as an optimization problem that simultaneously seeks an optimal model  $\boldsymbol{\mu}$  and optimal transformations  $\mathbf{T}^s$  from subject fiber sets' native spaces to the common space given a group of fiber sets  $\mathbf{x}^s$ . Using a Bayesian estimation framework, an optimal solution can be obtained by a maximum *a posteriori* (MAP) approach,

$$\begin{aligned} \boldsymbol{\theta} &= (\mathbf{T}^{1,2,\dots,S}, \boldsymbol{\mu}) \\ &= \arg \max_{\mathbf{T}^{1,2,\dots,S}, \boldsymbol{\mu}} \prod_{s=1}^S p(\mathbf{x}_{\text{all}}^s | \boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}, \mathbf{T}^s). \end{aligned} \quad (9)$$

To derive the expression of  $p(\mathbf{x}_{\text{all}}^s | \mathbf{T}^{1,2,\dots,S}, \boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})$ , it is assumed that each transformed fiber in the common space is an independent and identically distributed sample that is drawn from the distribution of the common bundle

model, which leads to the below formula,

$$\begin{aligned} & \prod_{s=1}^S p(\mathbf{x}_{\text{all}}^s | \mathbf{T}^s, \boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}) \\ &= \prod_{s=1}^S p(\mathbf{T}^s(\mathbf{x}_{\text{all}}^s) | \boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}) \\ &= \prod_{s=1}^S \prod_{j=1}^{M^s} p(\mathbf{T}^s(\mathbf{x}_j^s) | \boldsymbol{\mu}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}) \\ &= \prod_{s=1}^S \prod_{j=1}^{M^s} \sum_{k=1}^K p(\mathbf{T}^s(\mathbf{x}_j^s) | \boldsymbol{\mu}_k, \sigma_{\text{bundle}}, \sigma_{\text{ROI}}), \end{aligned} \quad (10)$$

where  $j, k, s$  index fibers in a target fiber set, fiber bundles in the common bundle set, and subject in the group, respectively. There are totally  $K$  bundles that need to be estimated and  $M^s$  fibers for each subject fiber set  $\mathbf{x}^s$ .  $p(\mathbf{T}^s(\mathbf{x}_j^s) | \boldsymbol{\mu}_k, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})$  is evaluated using the same formula as Eq. (4), where a bundle is indexed by the ROI pair  $(l_1, l_2)$ . The index  $k$  is in essence the same as the  $(l_1, l_2)$ . The optimal parameters  $(\mathbf{T}^{1,2,\dots,S}, \boldsymbol{\mu})$  can be found by maximizing the above probability [Eq. (10)].

#### IV.B. EM algorithm

The above optimization problem can be solved with the Expectation and Maximization algorithm. Let  $n$  denote the iteration of the EM algorithm and  $(\mathbf{T}^{n,1,2,\dots,S}, \boldsymbol{\mu}^n)$  denote the resulting parameters estimated in that iteration. In the E step, the membership probability of a fiber  $\mathbf{x}_j^s$  to the  $k$ th bundle is estimated as follows:

$$m_{j,k}^{s,n} = \frac{p(\mathbf{T}^{n-1,s}(\mathbf{x}_j^s) | \boldsymbol{\mu}_k^{n-1}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})}{\sum_{k=1}^K p(\mathbf{T}^{n-1,s}(\mathbf{x}_j^s) | \boldsymbol{\mu}_k^{n-1}, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})}. \quad (11)$$

In the M step, based on the fiber membership  $m_{j,k}^{s,n}$ , the original objective likelihood is turned into

$$\begin{aligned} E_{\text{EM}}(\mathbf{T}^{n,1,2,\dots,S}, \boldsymbol{\mu}^n) &= \sum_{s=1}^S \sum_{j=1}^{M^s} \sum_{k=1}^K m_{j,k}^{s,n} \\ &\quad \times \log(p(\mathbf{T}^{n,s}(\mathbf{x}_j^s) | \boldsymbol{\mu}_k^n, \sigma_{\text{bundle}}, \sigma_{\text{ROI}})). \end{aligned} \quad (12)$$

The above objective function can be optimized by firstly fixing the transformations  $\mathbf{T}^{n,1,2,\dots,S}$  to be  $\mathbf{T}^{n-1,1,2,\dots,S}$  and then solving the differential equations,

$$\frac{dE_{\text{EM}}}{d\boldsymbol{\mu}^n} = 0,$$

which leads to the below solution

$$\boldsymbol{\mu}_{k,i}^n = \frac{\sum_{s=1}^S \sum_{j=1}^{M^s} m_{j,k}^{s,n} \mathbf{T}^{s,n-1}(\mathbf{x}_{j,i}^s)}{\sum_{k=1}^K \sum_{s=1}^S \sum_{j=1}^{M^s} m_{j,k}^{s,n}}, \quad (13)$$

where  $i$  is used to index the points on the fiber.

After  $\mu_{k,i}^n$  is computed and fixed, we minimize the below objective function to estimate transformations  $\mathbf{T}^{n,1,2,\dots,S}$ ,

$$E_{EM}(\mathbf{T}^{n,1,2,\dots,S}) = \sum_{s=1}^S \sum_{j=1}^{M^s} \sum_{k=1}^K m_{j,k}^{s,n} \sum_{i=1}^{N_k} (\mathbf{T}^{s,n}(\mathbf{x}_{j,i}^s) - \mu_{k,i}^n) \times (\mathbf{T}^{s,n}(\mathbf{x}_{j,i}^s) - \mu_{k,i}^n)^T),$$

where  $N_k$  is the number of fibers in the  $k$ th subject.

The minimization of  $E_{EM}(\mathbf{T}^{n,1,2,\dots,S})$  is actually a least-squares problem, the computational cost of which depends mainly on the number of target fibers  $M^s$  in each subject and the total number of subjects  $S$ . To improve the computation efficiency, we circumvent the direct optimization of  $E_{EM}(\mathbf{T}^{n,1,2,\dots,S})$  by minimizing a simpler form  $E_{EM}'(\mathbf{T}^{n,1,2,\dots,S})$  as follows:

$$E_{EM}'(\mathbf{T}^{n,1,2,\dots,S}) = \sum_{k=1}^K \sum_{i=1}^{N_k} (\mathbf{T}^{s,n}(\mathbf{y}_{k,i}^n) - \mu_{k,i}^n) \times (\mathbf{T}^{s,n}(\mathbf{y}_{k,i}^n) - \mu_{k,i}^n)^T), \quad (14)$$

where

$$\mathbf{y}_{k,i}^n = \sum_{s=1}^S \sum_{j=1}^{M^s} m_{j,k}^{s,n} \mathbf{x}_{j,i}^s.$$

Here,  $\mathbf{y}_{k,i}^n$  can be interpreted as the current estimation of bundle centroids for individual subject  $k$ . The optimization of Eq. (14) is essentially aligning bundle centroids of each subject to their common model centroids. This simplification would make the computational complexity proportional to the number of bundles and the number of points in each bundle, which is much smaller than the total number of fibers in the whole data set. In the case of whole brain bundling, even with reduction, the computational complexity is still unacceptable due to a large number of bundles (typically  $\sim 1000$ ). Therefore, each centroid is further downsampled (three times) to fit the computation to our hardware resources.

Theoretically, any form of transformation can be used in the above registration framework. The thin-plate spline (TPS) transformation is chosen in this work due to its smoothness in deformation fields and closed-form solution for warping and parameter estimation.<sup>42</sup>

To compute the initial values for the parameters  $(\mathbf{T}^{0,1,2,\dots,S}, \mu^0)$ , a rigid registration is first performed to give a rough alignment of all the subjects to the template  $(\mathbf{T}^{0,1,2,\dots,S})$ . Then fibers from all subjects are transformed to the template space and classified as bundles that connect a pair of ROIs.  $\mu^0$  is obtained by combining the same bundles from all the subjects and computing their mean values.

## V. EXPERIMENTS AND RESULTS

### V.A. Gray matter parcellation template

In Tzourio-Mazoyer's work,<sup>18</sup> the gray matter in the Montreal Neurological Institute (MNI) single subject MRI data is manually labeled into 90 ROIs, the so-called automated anatomical labeling (AAL) mask, including 39 cortical re-

gions on each brain hemisphere and 12 subcortical regions. Similar to the work of Gong *et al.*,<sup>2</sup> these ROIs are considered to be basic units of the brain gray matter, and fiber bundles connecting pairs of ROIs are studied in this work.

### V.B. Imaging and fiber tracking

The T1 weighted and diffusion weighted images were acquired for ten healthy human subjects using a 3T Philips Achieva MR scanner. Informed consent was given by the subject according to a protocol that was approved by the local ethics committee. Each T1 volume contains a  $170 \times 256 \times 256$  matrix with an isotropic resolution of  $1 \times 1 \times 1$  mm<sup>3</sup>. The DWI data were acquired with 32 noncollinear weighting directions and a single shot, echo-planar, pulsed gradient spin echo imaging sequence with a diffusion weighting factor (i.e., b-value) of 1000, which generated a volume of  $128 \times 128 \times 60$  voxels at an isotropic resolution of  $2 \times 2 \times 2$  mm<sup>3</sup> for each direction. Three repeated scans were performed and coregistered and corrected for motion and distortion. A linear least-squares fitting was used to estimate diffusion tensors. Then a streamline tracking algorithm was started at all voxels with FA > 0.15, and followed sequentially along the local principal diffusion direction at a step size of 2 mm. A fiber was terminated when voxels with FA below 0.15 were met or the angle between the principal diffusion directions of two consecutive points exceeded 41°. The above procedure generated a whole volume fiber set ( $\sim 40\,000$  fibers) for each subject.

### V.C. Registration of AAL mask with fibers

To transform a subject's fibers to the MNI space, the MNI T1 template is registered with the subject's T1 images using intensity difference as a metric. As the subject's T1 images are not in the exact space as the DTI data and fibers, the T1 images need to be first aligned with the associated b0 diffusion weighted images. As the two types of images have different intensity distributions, normalized mutual information<sup>42</sup> is used as the registration metric. Moreover, a rigid transformation is estimated since both image sets are from the same subject and thus a rigid transformation should be sufficient to characterize the transformation. The T1 images are warped and then registered with the T1 MNI template. The resulting transformation maps fibers from DTI space to MNI space, where the AAL mask is used as the parcellation constraint.

### V.D. Evaluations of single-subject based bundling

Although an anatomical mapping is already provided by the above registration process, the inaccuracy of registration likely results in bundles with poor coherence. On the other hand, a coherent bundle from a clustering algorithm may have poor consistency with this anatomical mapping.

#### V.D.1. Metrics

To quantitatively characterize the bundling results, two metrics are proposed in this work: (1) mean inbundle

variation and (2) mean end-to-ROI distance. Mean inbundle variation (MIV) generally measures the coherence of a bundle, which is expressed as the mean distance of fibers to their corresponding bundle centroids,

$$\text{MIV} = \frac{1}{J} \sum_{j=1}^J |\mathbf{x}_j - \boldsymbol{\mu}_{\varphi(j)}|, \quad (15)$$

where  $|\bullet|$  denotes the distance between two fibers and  $\varphi(\bullet)$  is the fiber assignment function that maps fiber  $\mathbf{x}_j$  to the bundle with maximum membership. The total distances are normalized by the number of fibers  $J$ . The mean end-to-ROI distances (MED) characterize the deviation of bundles from their assigned ROI pair. This metric can be expressed as follows:

$$\text{MED} = \frac{1}{J} \sum_{j=1}^J \phi_{l_{1\varphi(j)}}(\mathbf{r}_{1j}) + \phi_{l_{2\varphi(j)}}(\mathbf{r}_{2j}), \quad (16)$$

where  $\mathbf{r}_{1j}$  and  $\mathbf{r}_{2j}$  are start and end points of fiber  $\mathbf{x}_j$ , and  $l_{1\varphi(j)}$  and  $l_{2\varphi(j)}$  correspond to the two ROIs, which define the bundle that  $\mathbf{x}_j$  is assigned to. A smaller MED usually suggests resulting bundles are consistent with the original ROI definition, while a bigger MED shows resulting bundles may not connect their corresponding ROIs correctly.

### V.D.2. Convergence

The single-subject bundling algorithm ( $\sigma_{\text{bundle}} = 2$  voxels,  $\sigma_{\text{ROI}} = 2$  voxels) is applied to all ten subjects' DTI data. The number of fiber assignment changes is recorded at each iteration, and the algorithm achieves convergence (below 20 changes) in nine iterations for all ten cases (see Fig. 1).

### V.D.3. Effect of the ROI variance

For one of the subjects, the proposed bundling algorithm is applied several times with variance  $\sigma_{\text{ROI}}$  changing from 0.5 to 5. The plot in Fig. 2(a) shows that the number of iterations that is needed to reach convergence increases with the increase of  $\sigma_{\text{ROI}}$ , which indicates that a smaller  $\sigma_{\text{ROI}}$  would make the algorithm converge faster. The increase of variance puts less confidence on the ROI mapping and thus the effect of the ROI constraint is weakened so that the clustering pro-

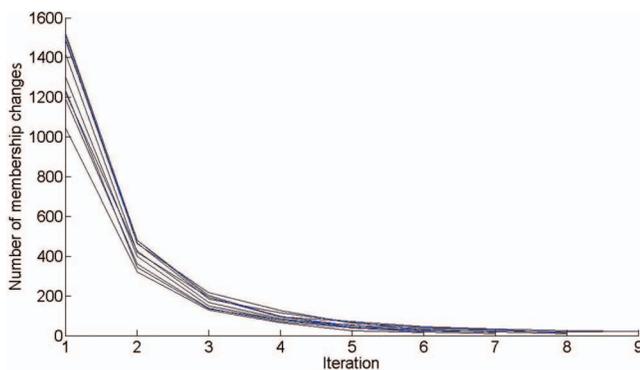


FIG. 1. The variation of the number of fiber-to-bundle assignment changes with respect to iteration index for all ten subjects.

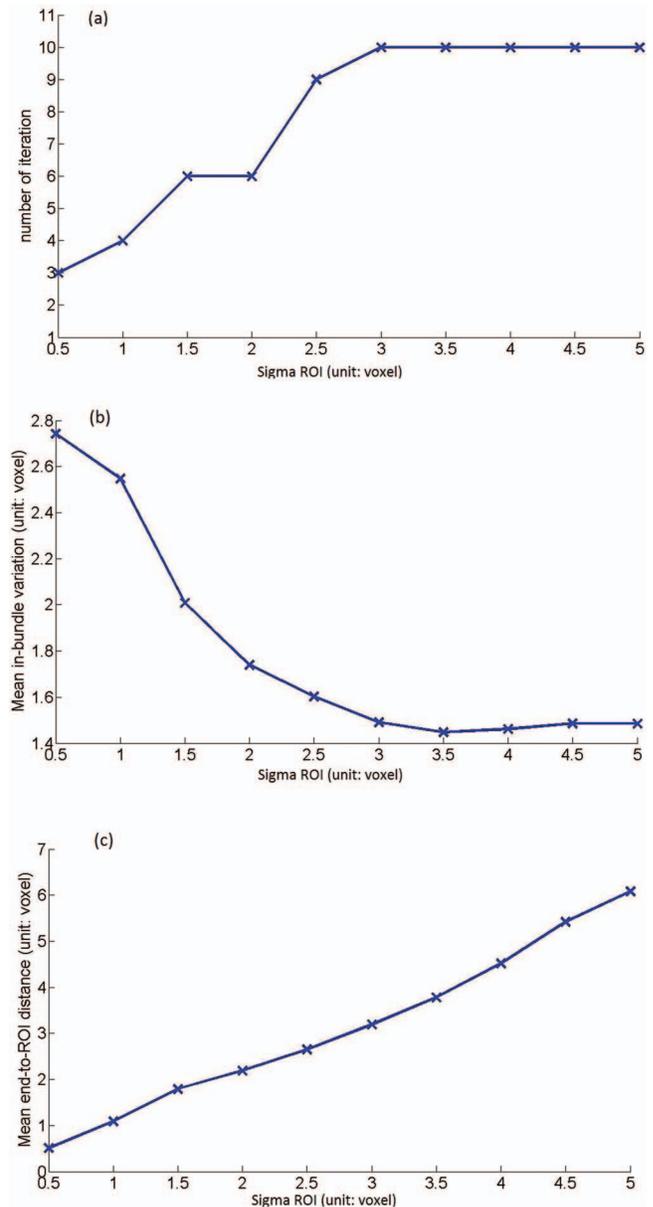


FIG. 2. The variations of (a) number of iterations to convergence, (b) mean inbundle variation, and (c) mean end-to-ROI distance with respect to different ROI variance.

cess takes longer. In Fig. 2(b), the mean inbundle variation decreases with the increase of  $\sigma_{\text{ROI}}$ , which leads to a decreasing contribution of end point positions to the bundle model. It also can be seen from Fig. 2(c) that the mean end-to-ROI distances are increasing with the increase of  $\sigma_{\text{ROI}}$ , as a bigger  $\sigma_{\text{ROI}}$  would reduce the “force” of dragging a bundle to its corresponding ROI pair so that the bundle is more free to move away from ROIs. From these results, we can see that a zero  $\sigma_{\text{ROI}}$  would turn the algorithm into a simple one that just labels fibers based on their closest ROI pairs, while a positive infinite  $\sigma_{\text{ROI}}$  would turn the algorithm into a pure clustering algorithm without consideration of the anatomical information. Based on this result,  $\sigma_{\text{ROI}}$  is set to 2 for the following experiments as this setting can generate relatively smaller

values for both the mean inbundle variance and end-to-ROI distances.

#### V.D.4. Comparisons with baseline methods

In this experiment, the proposed algorithm is compared with two other baseline methods: (1) a clustering without gray matter projection model and (2) directly assigning fibers based on their closest ROI label.

Two examples of the resulting bundles from these methods are illustrated in Fig. 3. In the first row, the bundle connecting the triangular part of the left inferior frontal gyrus to the left caudate is displayed for both the proposed method Fig. 3(a) and the ROI constraint only method Fig. 3(b). It can be seen that in Fig. 3(b) several outlier fibers, which deviate significantly from the majority of the bundle, are also grouped into this bundle as their end points fall into the corresponding ROIs. These fibers are discarded in the proposed method as their existence in this bundle would increase the inbundle variation or reduce the coherence of the bundle. On the other hand, a clustering algorithm that only aims at minimizing such coherence will also produce some erroneous results as in Fig. 3(d). Although initialized as a bundle connecting the left hippocampus to the left supplementary motor area, the bundle still finally converges to a very small bundle that does not even connect these two specific ROIs. The proposed algorithm is capable of constraining the bundle to the two ROIs so that the resulting bundle will not move too far away.

To further quantify this performance, the mean inbundle variation and end-to-ROI distances are summarized in Fig. 4 for all ten subjects. Although the clustering only and ROI constrained only methods achieved minimum inbundle distance and end-to-ROI distances, respectively, they also generate large values for the other metric. On the other hand,

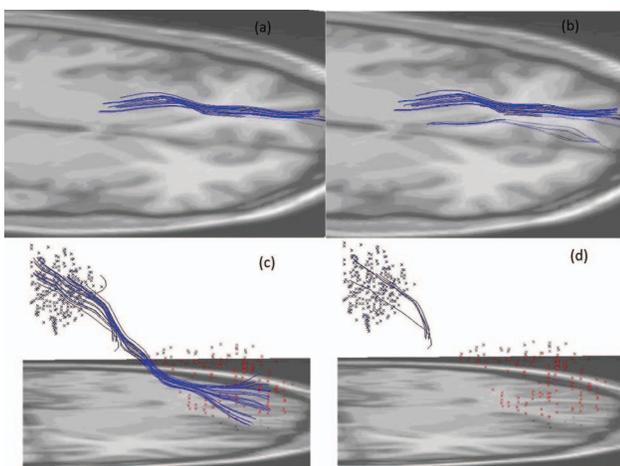


FIG. 3. An illustrative example of differences of the parcellation constrained bundling with the clustering only and the parcellation constraint only method. (a) The bundle connecting the triangular part of the left inferior frontal gyrus to the left caudate (parcellation constrained clustering). (b) The same bundle as (a) (parcellation constraint only). (c) The bundle connecting the left hippocampus to the left supplementary motor area (parcellation constrained clustering). (d) The same bundle as (c) (clustering only).

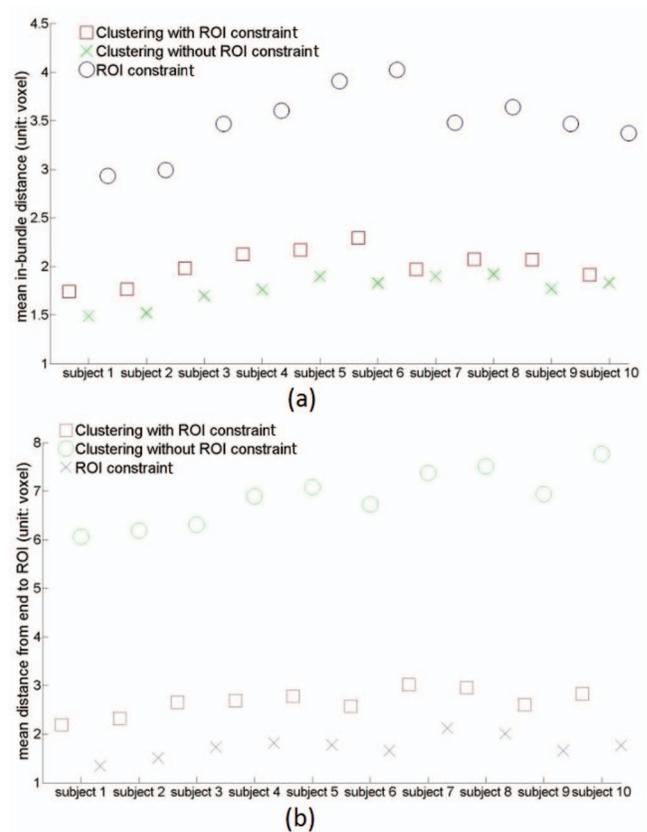


FIG. 4. The plot of mean inbundle variation (a) and end-to-ROI distances (b) for the parcellation constrained bundling and the two baseline methods.

the proposed algorithm yields close-to-minimum values for both metrics without deteriorating the other metric.

#### V.D.5. Demonstrations of resulting bundles

Due to the image misregistration and the uncertainty in the DTI fibers, it is possible that some bundles are generated for nonexistent connections between two cortical/subcortical regions. To eliminate these outlier bundles, only bundles consistent across the whole group are identified and kept as valid connections. To measure the consistency, we computed the mean differences of each individual bundle's mean and its group mean (discussed further in Sec. V.E). Bundles with above 2.5 voxels difference are discarded, resulting in a total of 36 bundles which are rendered in Fig. 5 for two example subjects.

#### V.E. Evaluations of groupwise bundling

In the case of the groupwise bundling, the initial registration is not considered to be sufficient to provide bundlewise alignment. Therefore, in addition to bundling, bundles from different subjects are aligned simultaneously, which contributes to the preservation of cross-subject consistency.



FIG. 5. Saggittal (left), axial (middle), and coronal (right) view of whole brain bundling results for two example subjects (subjects 3 and 6) using the parcellation constrained bundling.

### V.E.1. Metrics

To measure the cross-subject bundle consistency, a mean bundle centroid difference (MBCD) is proposed and computed as follows:

$$\text{MBCD} = \frac{1}{\text{KS}} \sum_{k=1}^K \sum_{s=1}^S |\mu_k^s - \mu_k|, \quad (17)$$

where  $\mu^s$  is the subject-specific bundle centroid,  $\mu$  is the bundle centroid of the group, and  $k$  indexes the bundle. This metric essentially measures the deviation of subject-specific models from the group mean. The groupwise mean inbundle variation (GMIV) is also measured by computing the average distance of fibers to their corresponding group bundle centroids

$$\text{GMIV} = \frac{1}{\sum_{s=1}^S M^s} \sum_{s=1}^S \sum_{j=1}^{M^s} |\mathbf{x}_j^s - \mu_{\varphi(\mathbf{x}_j^s)}|, \quad (18)$$

where the assignment function  $\varphi(\mathbf{x}_j^s)$  assigns the fiber  $\mathbf{x}_j^s$  to a group common bundle. The coherence metric GMIV is related to consistency, as consistent bundling tends to yield coherent bundles.

### V.E.2. Baseline methods

To demonstrate the main advantages of the proposed algorithm, we compared it with three different baseline methods. In method I each fiber set from the subject group is individually bundled using the gray matter parcellation constrained bundling algorithm with no attempt to correct for image mis-

TABLE I. The consistency and coherence metrics (in voxels) for all four types of groupwise bundling methods.

|      | Group consistent fiber bundling | Method I | Method II | Method III |
|------|---------------------------------|----------|-----------|------------|
| GMIV | 3.74                            | 4.98     | 4.54      | 4.41       |
| MBCD | 1.31                            | 2.78     | 2.20      | 1.90       |

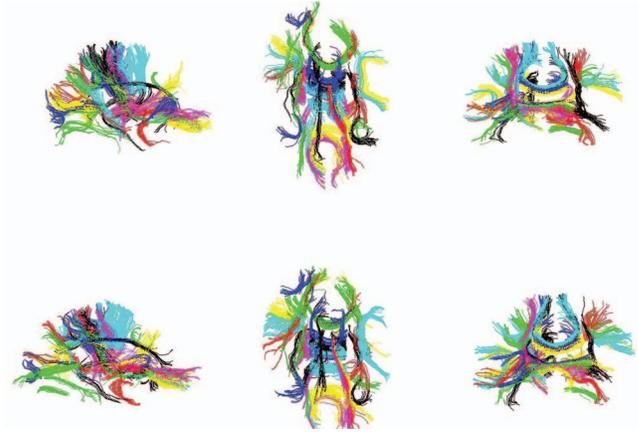


FIG. 6. Saggittal (left), axial (middle), and coronal (right) view of whole brain bundling results for two example subjects (subjects 3 and 6) using the groupwise bundling.

registration. Method II bundles the fibers in a way identical to method I except that a TPS transformation is applied afterwards to transform each individual subject's bundle centroids to the corresponding common centroids. Method III uses a joint clustering scheme that treats fibers from all subjects as a single fiber set, and performs gray matter parcellation constrained bundling on this combined data set. Note that all fibers are transformed to the MNI space by the preliminary image coregistration for all the three methods.

The above metrics are computed and summarized in Table I for all four methods, including the proposed method and the other three baseline methods. It can be seen that the proposed consistent groupwise bundling algorithm has the smallest values for both metrics, which indicates its superiority in preserving consistency. Although a nonrigid transformation is also used in method II to align bundles with their group mean, the resulting metric values still cannot compete with the proposed algorithm due to the absence of this transformation in the bundling process. In each iteration of the proposed algorithm, the clustering would favor the direction that could generate more consistent bundling. On the other hand, the joint clustering without transformation is less likely to reduce consistency very much due to the misalignment caused by scalar-image coregistration. Therefore, from this example, one can see that it is important to integrate the nonrigid transformation into the clustering process.

### V.E.3. Demonstrations of resulting bundles

To demonstrate some identified connections, the same elimination procedure as used for the single subject bundling experiments (see Sec. V.D.5) is applied to the resulting fiber bundles, generating 45 bundles for each subject (see Fig. 6 for two representative subjects).

## VI. CONCLUSION AND DISCUSSION

To obtain a set of DTI fiber bundles, one must start with DTI imaging, track fibers by connecting DTI tensor directions, and then cluster these fibers based on fiber similarities.

There is a significant amount of noise and uncertainty in each step. Noise introduced in the DTI imaging step makes the estimated fiber direction at each voxel unreliable. By following these unreliable directions a fiber is generated. Therefore, any uncertainty in these directions causes the fiber to vary from its true shape and position. This issue is more serious on the ends of the fiber as these points are farther away from the starting point and errors are accumulated in the tracking process. To cluster these fibers into bundles, fiber similarity needs to be measured. However, this step is more challenging, as it is still an open question on how to measure curve similarity and there are lots of issues such as point correspondence and metric and so on. Therefore, this kind of noise must be reduced to generate accurate fiber bundles for a study.

Using prior knowledge and smoothing is the major way of reducing noise. For example, on DTI images it could be assumed that intensities of close voxels should be similar. This prior knowledge can be utilized to smooth images by averaging intensities of neighboring voxels to get rid of imaging noise. If it is only assumed that only voxels in homogenous regions have similar intensities, one should only smooth voxels on nonboundary regions, which leads to an anisotropic denoising method. Different prior knowledge could lead to different denoising algorithms. Similarly in fiber tracking, the prior knowledge that two neighboring points along the curve should have similar directions has been frequently used to reduce the noise in the fiber tracking, although this prior can be integrated into fiber tracking in different ways, such as simply smoothing a curve to reduce curvature or using it as a Bayesian prior.

The basic motivation of this paper is to introduce useful priors to reduce the noise in the fiber bundling process. In addition to the traditional prior that similar fibers should stay in the same bundle, we proposed a novel prior based on the anatomical information of human brain, i.e., bundles should have meaningful correspondence to the real brain anatomy. The anatomical prior is much more reliable than the imaging data obtained through a single DTI scan, as it is built from numerous imaging data sets and validated by human experts. Therefore, fiber bundles deviating from the anatomy structure are corrected by enforcing this prior. We also developed a Bayesian algorithm to combine the traditional and the anatomical prior to make them work together. Another useful prior introduced into fiber bundling is that bundles from a group of similar subjects should have similar structures. With this prior, bundles in a subject are smoothed by bundles from his/her "neighboring" subjects. A groupwise bundling algorithm is proposed in this paper to make this prior integrated into the Bayesian clustering framework. In summary, the major contribution of this paper is the exploring of different priors to reduce the uncertainty of DTI fiber bundling.

Similar to other smoothing techniques, oversmoothing could be a potential issue to the use of priors in reducing noise. In other words, true individual variations could also be falsely treated as noise and removed due to its difference from the prior, such as the brain anatomical template or the mean bundle of the subject group. This could be the focus of future

work, identifying the right procedure of tuning the algorithm parameters.

Two important aspects ignored in this paper are the underlying diffusion models and fiber tracking algorithms that are used to generate the input to the presented method. Although they are not the focus of this paper, we do realize that the resulting fiber bundles are heavily affected by these factors. The proposed method can only yield the fiber bundles for the entire brain only if the input fibers contains all the fibers in the entire brain. We use a tensor model and streamline fiber tracking algorithm for demonstration purposes. However, more accurate models with higher angular resolutions and more accurate fiber tracking algorithms (e.g., probabilistic tracking algorithms) could be potentially used to improve the coverage and accuracy of the resulting fiber bundles.

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: qingxu2009@gmail.com

<sup>1</sup> P. J. Basser, J. Mattiello, and D. Le Bihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.* **66**(1), 259–267 (1994).

<sup>2</sup> D. Le Bihan, J. F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion tensor imaging: Concepts and applications," *J. Magn. Reson. Imaging* **13**, 534–546 (2001).

<sup>3</sup> V. J. Schmithorst, M. Wilke, B. J. Dardzinski, and S. K. Holland, "Cognitive functions correlate with white matter architecture in a normal pediatric population: A diffusion tensor MRI study," *Hum Brain Mapp* **26**(2), 139–47 (2005).

<sup>4</sup> V. L. Morgan, A. Mishra, A. T. Newton, J. C. Gore, and Z. Ding, "Integrating functional and diffusion magnetic resonance imaging for analysis of structure-function relationship in the human language network," *PLoS ONE* **4**(8), e6660 (2009).

<sup>5</sup> M. Filippi, M. Cercignani, M. Inglese, M. A. Horsfield, and G. Comi, "Diffusion tensor magnetic resonance imaging in multiple sclerosis," *Neurology* **56**(3), 304–311 (2001).

<sup>6</sup> C. Pierpaoli, A. Barnett, S. Pajevic, R. Chen, L. R. Penix, A. Virts, and P. Basser, "Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture," *NeuroImage* **13**, 1174–1185 (2001).

<sup>7</sup> B. A. Assaf, F. B. Mohamed, K. J. Abou-Khaled, J. M. Williams, M. S. Yazeji, J. Haselgrove, and S. H. Faro, "Diffusion tensor imaging of the hippocampal formation in temporal lobe epilepsy," *AJNR Am. J. Neuroradiol.* **24**(9), 1857–1862 (2003).

<sup>8</sup> J. R. Wozniak, B. A. Mueller, P. N. Chang, R. L. Muetzel, L. Caros, and K. O. Lim, "Diffusion tensor imaging in children with fetal alcohol spectrum disorders," *Alcohol Clin. Exp. Res.* **30**(10), 1799–806 (2006).

<sup>9</sup> J. E. Lee, E. D. Bigler, A. L. Alexander, M. Lazar, M. B. DuBray, M. K. Chung, M. Johnson, J. Morgan, J. N. Miller, W. M. McMahon, J. Lu, E. K. Jeong, and J. E. Lainhart, "Diffusion tensor imaging of white matter in the superior temporal gyrus and temporal stem in autism," *Neurosci. Lett.* **424**(2), 127–132 (2007).

<sup>10</sup> M. Kyriakopoulos, T. Bargiotas, G. J. Barker, and S. Frangou, "Diffusion tensor imaging in schizophrenia," *Eur. Psychiatry* **23**(4), 255–273 (2008).

<sup>11</sup> K. Kantarci, R. Avula, M. L. Senjem, A. R. Samikoglu, B. Zhang, S. D. Weigand, S. A. Przybelski, H. A. Edmonson, P. Vemuri, D. S. Knopman, T. J. Ferman, B. F. Boeve, R. C. Petersen, and C. R. Jack, Jr., "Dementia with Lewy bodies and Alzheimer disease: Neurodegenerative patterns characterized by DTI," *Neurology* **74**(22), 1814–1821 (2010).

<sup>12</sup> S. Mori and P. C. van Zijl, "Fiber tracking: Principles and strategies—a technical review," *NMR Biomed.* **15**(7–8), 468–480 (2002).

<sup>13</sup> S. Zhang, S. Correia, and D. H. Laidlaw, "Identifying white-matter fiber bundles in DTI data using an automated proximity-based fiber-clustering method," *IEEE Trans. Vis. Comput. Graph.* **14**(5), 1044–1053 (2008).

<sup>14</sup> M. Catani, R. J. Howard, S. Pajevic, and D. K. Jones, "Virtual *in vivo* interactive dissection of white matter fasciculi in the human brain," *NeuroImage* **17**(1), 77–94 (2002).

<sup>15</sup> S. Wakana, H. Jiang, L. M. Nagae-Poetscher, P. C. van Zijl, and S. Mori, "Fiber tract-based atlas of human white matter anatomy," *Radiology* **230**(1), 77–87 (2004).

- <sup>16</sup>L. Concha, D. W. Gross, and C. Beaulieu, "Diffusion tensor tractography of the limbic system," *AJNR Am. J. Neuroradiol.* **26**, 2267–2274 (2005).
- <sup>17</sup>Z. Ding, J. C. Gore, and A. W. Anderson, "Classification and quantification of neuronal fiber pathways using diffusion tensor MRI," *Magn. Reson. Med.* **49**(4), 716–721 (2003).
- <sup>18</sup>A. Brun, H. Knutsson, H. J. Park, M. E. Shenton, and C. F. Westin, "Clustering fiber traces using normalized cuts," *Lect. Notes Comput. Sci.* **3216**, 368–375 (2004).
- <sup>19</sup>L. O'Donnell and C. F. Westin, "White matter tract clustering and correspondence in populations," *Medical Image Computing Computer-Assisted Intervention (MICCAI)* (Springer, Heidelberg, 2005), pp. 140–147.
- <sup>20</sup>I. Corouge, P. T. Fletcher, S. Joshi, S. Gouttard, and G. Gerig, "Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis," *Med. Image Anal.* **10**(5), 786–798 (2006).
- <sup>21</sup>M. Maddah, W. E. L. Grimson, S. K. Warfield, and W. M. Wells, "A unified framework for clustering and quantitative analysis of white matter fiber tracts," *Med. Image Anal.* **12**(2), 191–202 (2008).
- <sup>22</sup>U. Ziyang, M. R. Sabuncu, W. E. Grimson, and C. F. Westin, "Consistency clustering: A robust algorithm for group-wise registration, segmentation and automatic atlas construction in diffusion MRI," *Int. J. Comput. Vis.* **85**(3), 279–290 (2009).
- <sup>23</sup>E. Visser, E. H. Nijhuis, J. K. Buitelaar, and M. P. Zwiers, "Partition-based mass clustering of tractography streamlines," *Neuroimage* **54**(1), 303–312 (2011).
- <sup>24</sup>P. Guevara, C. Poupon, D. Rivière, Y. Cointepas, M. Descoteaux, B. Thirion, and J. F. Mangin, "Robust clustering of massive tractography datasets," *Neuroimage* **54**(3), 1975–1993 (2011).
- <sup>25</sup>D. Wassermann, L. Bloy, E. Kanterakis, R. Verma, and R. Deriche, "Unsupervised white matter fiber clustering and tract probability map generation: Applications of a Gaussian process framework for white matter fibers," *Neuroimage* **51**(1), 228–241 (2010).
- <sup>26</sup>H. J. Park, M. Kubicki, C. F. Westin, I. F. Talos, A. Brun, S. Peiper, R. Kikinis, F. A. Jolesz, R. W. McCarley, and M. E. Shenton, "Method for combining information from white matter fiber tracking and grey matter parcellation," *AJNR Am. J. Neuroradiol.* **25**(8), 1318–1324 (2004).
- <sup>27</sup>Y. Xia, U. Turken, S. L. Whitfield-Gabrieli, and J. D. Gabrieli, "Knowledge-based classification of neuronal fibers in entire brain," *Medical Image Computing Computer-Assisted Intervention (MICCAI)* (Springer, Heidelberg, 2005), pp. 205–212.
- <sup>28</sup>M. Maddah, A. U. J. Mewes, S. Haker, W. E. L. Grimson, and S. K. Warfield, "Automated atlas-based clustering of white matter fiber tracts from DTMRI," *Medical Image Computing Computer-Assisted Intervention (MICCAI)* (Springer, Heidelberg, 2005), pp. 188–195.
- <sup>29</sup>Q. Xu, A. W. Anderson, J. C. Gore, and Z. Ding, "Unified bundling and registration of brain white matter fibers," *IEEE Trans. Med. Imaging* **28**(9), 1399–1411 (2009).
- <sup>30</sup>S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *J. Neurosci.* **26**(1), 63–72 (2006).
- <sup>31</sup>M. Guye, F. Bartolomei, and J. P. Ranjeva, "Imaging structural and functional connectivity: Towards a unified definition of human brain organization?" *Curr. Opin. Neurol.* **21**(4), 393–403 (2008).
- <sup>32</sup>D. K. Jones and C. Pierpaoli, "Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach," *Magn. Reson. Med.* **53**(5), 1143–1149 (2005).
- <sup>33</sup>M. Lazar and A. L. Alexander, "Bootstrap white matter tractography (BOOT-TRAC)," *Neuroimage* **24**(2), 524–532 (2005).
- <sup>34</sup>A. W. Anderson, "Theoretical analysis of the effects of noise on diffusion tensor imaging," *Magn. Reson. Med.* **46**(6), 1174–1188 (2001).
- <sup>35</sup>C. Poupon, C. A. Clark, V. Frouin, J. Régis, I. Bloch, D. Le Bihan, and J. Mangin, "Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles," *Neuroimage* **12**(2), 184–195 (2000).
- <sup>36</sup>Z. Ding, J. C. Gore, and A. W. Anderson, "Reduction of noise in diffusion tensor images using anisotropic smoothing," *Magn. Reson. Med.* **53**(2), 485–490 (2005).
- <sup>37</sup>Y. Lu, A. Aldroubi, J. C. Gore, A. W. Anderson, and Z. Ding, "Improved fiber tractography with Bayesian tensor regularization," *Neuroimage* **31**(3), 1061–1074 (2006).
- <sup>38</sup>A. L. Alexander, K. M. Hasan, M. Lazar, J. S. Tsuruda, and D. L. Parker, "Analysis of partial volume effects in diffusion-tensor MRI," *Magn. Reson. Med.* **45**(5), 770–780 (2001).
- <sup>39</sup>D. S. Tuch, "Q-ball imaging," *Magn. Reson. Med.* **52**(6), 1358–1372 (2004).
- <sup>40</sup>A. W. Anderson, "Measurement of fiber orientation distributions using high angular resolution diffusion imaging," *Magn. Reson. Med.* **54**(5), 1194–1206 (2005).
- <sup>41</sup>O. Friman, G. Farneback, and C. F. Westin, "A Bayesian approach for stochastic white matter tractography," *IEEE Trans. Med. Imaging* **25**(8), 965–978 (2006).
- <sup>42</sup>T. E. Behrens, H. J. Berg, S. Jbabdi, M. F. Rushworth, and M. W. Woolrich, "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?" *Neuroimage* **34**(1), 144–155 (2007).
- <sup>43</sup>A. Mishra, A. W. Anderson, X. Wu, J. C. Gore, and Z. Ding, "An improved Bayesian tensor regularization and sampling algorithm to track neuronal fiber pathways in the language circuit," *Med. Phys.* **37**(8), 4274–4287 (2010).
- <sup>44</sup>P. M. Thompson, R. P. Woods, M. S. Mega, and A. W. Toga, "Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain," *Hum. Brain Mapp.* **9**(2), 81–92 (2000).
- <sup>45</sup>A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **39**(1), 1–38 (1977).
- <sup>46</sup>K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. H. Kuhn, "Landmark-based elastic registration using approximating thin-plate splines," *IEEE Trans. Med. Imaging* **20**(6), 526–534 (2001).
- <sup>47</sup>P. Hagmann, L. Cammoun, X. Gigandet, S. Gerhard, P. Ellen Grant, V. Wedeen, R. Meuli, J. P. Thiran, C. J. Honey, and O. Sporns, "MR connectomics: Principles and challenges," *J. Neurosci. Methods* **194**(1), 34–45 (2010).