



Published in final edited form as:

*Neuroimage*. 2020 November 01; 221: 117147. doi:10.1016/j.neuroimage.2020.117147.

## Groupwise track filtering via iterative message passing and pruning

Yihao Xia<sup>a,b</sup>, Yonggang Shi<sup>a,b,\*</sup>,<sup>1</sup>

<sup>a</sup>USC Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

<sup>b</sup>Department of Electrical and Computer Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA, USA

### Abstract

Tractography is an important tool for the *in vivo* analysis of brain connectivity based on diffusion MRI data, but it also has well-known limitations in false positives and negatives for the faithful reconstruction of neuroanatomy. These problems persist even in the presence of strong anatomical priors in the form of multiple region of interests (ROIs) to constrain the trajectories of fiber tractography. In this work, we propose a novel track filtering method by leveraging the groupwise consistency of fiber bundles that naturally exists across subjects. We first formalize our groupwise concept with a flexible definition that characterizes the consistency of a track with respect to other group members based on three important aspects: *degree*, *affinity*, and *proximity*. An iterative algorithm is then developed to dynamically update the *localized* consistency measure of all streamlines via message passing from a reference set, which then informs the pruning of outlier points from each streamline. In our experiments, we successfully applied our method to diffusion imaging data of varying resolutions from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and Human Connectome Project (HCP) for the consistent reconstruction of three important fiber bundles in human brain: the fornix, locus coeruleus pathways, and corticospinal tract. Both qualitative evaluations and quantitative comparisons showed that our method achieved significant improvement in enhancing the anatomical fidelity of fiber bundles.

### Keywords

Tractography; Fiber bundle; Groupwise; Filtering; Connectivity

---

<sup>1</sup>Data used in preparation of this article were partly obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\*Corresponding author. 2025 Zonal Ave, Los Angeles, CA, 90033, USA., [yshi@loni.usc.edu](mailto:yshi@loni.usc.edu) (Y. Shi).  
CRediT authorship contribution statement

**Yihao Xia:** Methodology, Software, Formal analysis, Investigation, Visualization, Writing - original draft. **Yonggang Shi:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition.

## 1. Introduction

The advent of diffusion magnetic resonance imaging (dMRI) (Basser et al., 1994) allows the study of structural connectivity in the human brain *in vivo*. To noninvasively reveal and study the trajectories of white matter pathways of human brain based on dMRI, tractography is a central approach (Basser et al., 2000; Mori et al., 1999) and has been successfully applied in neuroimaging studies of various brain disorders. On the other hand, recent validation studies (Aydogan et al., 2018; Maier-Hein et al., 2017) showed tractography techniques had critical limitations in the reliable reconstruction of neuroanatomy. To this end, we propose in this paper a novel track filtering algorithm for the robust reconstruction of fiber bundles with groupwise consistency. We will demonstrate that the groupwise consistency is able to compensate for limited anatomical knowledge in tractography-based fiber bundle reconstruction.

Fiber bundles can be generated by either deterministic (Basser et al., 2000; Mori et al., 1999) or probabilistic tractography techniques (Behrens et al., 2007; Descoteaux et al., 2009; Tournier et al., 2019). To remove artifacts from tractography results, various approaches have been proposed for the filtering of fiber tracks with the inclusion of different degree of anatomical priors. For the filtering of whole brain tractograms, several methods have been developed that examine how well the tractogram fits the dMRI data or the fiber orientation models computed from the dMRI data and remove streamlines with low data fidelity (Daducci et al., 2015; Pestilli et al., 2014; Smith et al., 2015). Clustering techniques (Guevara et al., 2011; O'Donnell and Westin, 2007) for the reconstruction of major fiber bundles are most related to our current work. By taking advantages of the geometric similarity of pathways, clustering algorithms can be applied to fiber tracks from individual subjects or multiple subjects warped into a common space. Recently, this approach was applied to the whole brain tractography of 100 HCP subjects for the extraction of common clusters and construction of white matter atlases, which were then applied to the whole brain tractogram of individual subjects for bundle reconstruction (Zhang et al., 2018). This approach is still largely data-driven and anatomical labels were typically assigned after the generation of clusters (O'Donnell and Westin, 2007). For fiber bundle reconstruction, the most conventional type of approach uses strong anatomical priors in the form of region-of-interests (ROIs) to increase the validity of tractography-based solutions (Catani et al., 2002; Wakana et al., 2004). This method is most suitable for the reconstruction of fiber bundles with well characterized anatomy (Tang et al., 2018). Based on this approach, a white matter query language (WMQL) was developed as an automated framework for ROI-based bundle segmentation (Wassermann et al., 2016).

Even with the use of strong ROI-based anatomical priors, however, residual artifacts still frequently occur in the reconstructed fiber bundles (Rheault et al., 2020). This is because we usually can only provide a small number of ROIs to constrain the fiber trajectories due to either incomplete anatomical knowledge or the high cost of generating extensive anatomical labels. To remove outliers from ROI-based reconstruction of fiber bundles, track filtering methods based on geometric distances (Garyfallidis et al., 2012) or topological analysis (Aydogan and Shi, 2015) were proposed. More recently, topographic regularity was proposed as a novel criterion for the removal of outlier streamlines (Wang et al., 2018). A

cluster confidence index (CCI) was introduced to model the geometric similarity of neighboring tracks and remove outliers (Jordan et al., 2018). One common theme of these methods is that they assume a certain level of geometric or topographic regularity at the *individual* level to compensate for the insufficiency of anatomical constraints, but consistency across subjects is not considered. In addition, the filtering process typically operates at the level of whole streamlines.

In this paper, we develop a novel track filtering method for fiber bundles by incorporating regularity from the perspective of groupwise consistency. In essence, our method will filter the fiber bundles from a group of subjects simultaneously and take advantage of the regularity that naturally exists at the group level to remove randomly occurring errors in each subject, thus compensating for the gaps in anatomical priors. Compared to previous track filtering methods for fiber bundles, there are several unique aspects in our work. First, we develop an intuitive definition of groupwise consistency that provides flexible controls over the level of desired consistency across the group from three different aspects: *degree*, *affinity*, and *proximity*. Second, we measure the groupwise consistency at each point on a streamline via an iterative message passing mechanism from a set of carefully constructed reference set of fibers. Third, we perform the filtering at the local level by iteratively pruning away inconsistent portion of each streamline instead of making a binary decision as in many previous works. This is one of the unique aspects of our method that not only allows the elimination of whole streamlines with defects but also the extraction of streamline segments reproducible across subjects, which can be highly valuable when precise end ROIs cannot be properly defined for a fiber bundle. In our experiments, we will demonstrate this property enables the reconstruction of anatomically meaningful and consistent sub-bundle structures from fiber bundles dominated by highly spurious outliers. We will also show the proposed method can be easily applied to perform conventional filtering tasks that either accept or reject a track from a bundle.

The rest of the paper is organized as follows. In section 2, we propose our definition of groupwise consistency and develop the numerical algorithm to implement this concept to achieve track filtering at the group level. In section 3, we present experimental results on the reconstruction of three important bundles: fornix, locus coeruleus (LC) pathways, and the corticospinal tract (CST) to demonstrate the efficacy of our method over conventional filtering methods. Finally, discussions and conclusions are made in section 4.

## 2. Method

In this section, we develop the proposed track filtering algorithm based on a novel and flexible definition of groupwise consistency across fiber bundles. The main steps of our method are illustrated in Fig. 1. After nonlinearly warping all fiber bundles into a common space, we iteratively estimate the level of groupwise consistency at each point on each streamline and conduct the pruning of outlier points to enhance the overall consistency across subjects. Streamlines will be rejected or refined during the filtering process for the generation of the final outputs.

## 2.1. Definition

We denote a set of input fiber bundles from  $N$  subjects as  $\hat{\mathcal{F}} = \{\hat{F}_1, \hat{F}_2, \dots, \hat{F}_N\}$ , where  $\hat{F}_n$  is the input bundle from the  $n$ -th subject. To perform the proposed groupwise filtering, these fiber bundles are first coregistered into a common coordinate space as illustrated in Fig. 1 and denoted as  $\mathcal{F} = \{F_1, F_2, \dots, F_N\}$ . For practical implementation, we typically warp all fiber bundles into the common MNI152 space (Fonov et al., 2011) using the nonlinear registration computed by the ANTS software (Avants et al., 2008). Each streamline in  $F$  is represented as a polyline  $f = \{x(l) \mid l \in [1, S]\}$  with  $S$  points in  $\mathbb{R}^3$ . For groupwise filtering, our goal is to estimate a subset of the streamline  $f$ , which we denote as  $\hat{f} = \{x(l) \mid l \in [l_a, l_b]\}$ , that are consistent with streamlines from other subjects. More specifically, we consider the streamline segment,  $\hat{f} \subseteq f$ , as groupwise consistent if all points in  $\hat{f}$  are *close* to streamlines from a certain number of other subjects in  $F$ . Formally, we define the conditions  $G(L; K; \xi)$  for  $\hat{f}$ , a portion of a streamline  $f$ , to have groupwise consistency as follows:

- Degree ( $L$ ): the minimum length requirement of the sub-streamline structure  $\hat{f}$  that are consistent to streamlines from other subjects.
- Affinity ( $K$ ): the number of subjects in the group that contain streamlines to which  $f$  shares consistency.
- Proximity ( $\xi$ ): a distance parameter reflecting the extent of closeness between  $\hat{f}$  and streamlines from other subjects.

This definition is flexible in several aspects. The first parameter  $L$  specifies the minimum length or proportion of a streamline that needs to be consistent with tracks of other subjects. This will avoid the inclusion of overly short segments that do not reflect the connectivity of the fiber bundle. For fiber bundles with well-defined end ROIs, setting a relatively high degree parameter  $L$  will ensure a valid representation of the overall bundle similar to conventional filtering approaches can be obtained. The second parameter  $K$  can be considered as an affinity measure that controls the trade-off between inter-subject consistency and individual variability. The distance parameter  $\xi$  determines the closeness when evaluating the consistency among fiber trajectories. Taken together, these conditions characterize groupwise consistency locally at the sub-streamline level. On the contrary, previous filtering methods often first perform fiber clustering and make a decision about each cluster. Thus, the proposed groupwise definition allows the development of filtering algorithms that can measure groupwise consistency and perform the pruning/filtering at a higher resolution than previous methods.

## 2.2. Algorithmic details

To computationally realize this flexible definition of groupwise consistency for fiber bundles, we essentially need to estimate a fraction  $\hat{f}$  of each streamline  $f$  that is consistent with streamlines from  $K$  other subjects in the group. Since this portion  $\hat{f}$  is unknown, the  $K$  subjects contributing to its groupwise consistency cannot be determined *a priori*. To tackle this challenge, we develop an iterative algorithm that is composed of consistency estimation and pruning of inconsistent points. At each iteration, a reference set is constructed and dynamically updated for each streamline and a message passing mechanism is developed to

estimate the level of consistency at each point, which then guides the pruning process to remove inconsistent portion of the streamline.

**1) Reference Set:** For any pruned or unpruned streamline  $f^t$  at the  $t$ -th ( $t = 0, 1, \dots$ ) iteration, which equals the input streamline  $f$  at  $t = 0$ , we construct a subject specific reference set  $R_k^t = \{r_k^m | m = 1, \dots, M\}$  composed of the  $M$  most similar streamlines from the fiber bundle  $F_k$  of the  $k$ -th subject. In most scenarios, the fiber bundle  $F_k$  is saturated with streamlines sharing similar trajectory and a subsampled one  $\tilde{F}_k^t$  would be enough to be used for reference researching. Practically, we offer an optional parameter, subsampling rate  $r$ , to control the percentage of streamlines stochastically selected from  $F_k$  for reference set construction. The parameter  $r$  provides a trade-off between accuracy and efficiency. Low subsampling rate could effectively reduce the computational burden, while the overly subsampled fiber bundle may underrepresent the original trajectories. We used the fast fiber  $k$ -NN algorithm proposed in (Wang and Shi, 2019) to efficiently construct the reference set. Given a streamline  $f^t$  and a searching scope, e.g.  $F_k$  or  $\tilde{F}_k^t$ , the fast fiber  $k$ -NN algorithm ranks each streamline in the scope according to its similarity to  $f^t$ . Then the top- $M$  streamlines are extracted to form the subject specific reference set.

To meet the affinity criterion that  $K$  subjects will be needed to define the groupwise consistency, we build the groupwise reference set  $\mathcal{R}^t$  of each track  $f^t$  at the  $t$ -th iteration as follows:

$$\mathcal{R}^t(f^t) = \underset{\mathcal{R}' \subseteq \mathcal{R}^t, |\mathcal{R}'| = K}{\operatorname{argmin}} \sum_{R_k^t \in \mathcal{R}'} \sum_{r_k^m \in R_k^t} d_{mc}(f^t, r_k^m),$$

where  $\mathcal{R}^t = \{R_k^t | k = 1, \dots, N - 1\}$  is the collection of all subject specific reference sets and  $d_{mc}(\cdot, \cdot)$  is the mean closest point (Corouge et al., 2004) distance depicting the streamline-wise distance. Specifically, this distance measure between two streamlines  $f^t$  and  $r_k^m \in F_k$  is defined as:

$$d_{mc}(f^t, r_k^m) = \frac{1}{|f^t|} \sum_{x \in f^t} \min_{y \in r_k^m} |x - y|,$$

where  $|f^t|$  is the number of points on  $f^t$ . The groupwise reference set  $\mathcal{R}^t(f^t)$  consists of the top  $K$  subject specific reference sets in terms of the total distance between  $f^t$  and the reference tracks.

**2) Message Passing:** Using the groupwise reference set  $\mathcal{R}^t(f^t)$  at the current iteration, we will define a consistency measure at each point on  $f^t$  to enable localized track filtering. Inspired by the message passing mechanism in graph-based optimization (Wainwright et al., 2005), we will quantify the consistency level of each point by measuring the *messages* that it received from the reference set.

Given any point  $x \in \mathcal{I}^t$ , we denote a neighborhood point set  $\mathcal{N} = \{n_i | i = 1, \dots, |\mathcal{N}|\}$ , where  $n_j$  is the closest point to the  $i$ -th streamline in the reference set  $\mathcal{F}^t(f^t)$ . Note that the total number of points in  $\mathcal{N}$  equals the number of reference streamlines in  $\mathcal{F}^t(f^t)$ , i.e.,  $M \times K$ . The message

the  $i$ -th neighborhood point sends to  $x$  is defined as  $e^{-\frac{x - n_i^2}{\sigma^2}}$ , which decays exponentially as the distance between these two points increases. Note the distance  $x - n_j$  is normalized with respect to the distance scale parameter  $\sigma$  which controls the quantitative conversion from point-wise distance to point-wise affinity. By summing up the messages from all neighboring points, we obtain the consistency measure at  $x$  as:

$$p(x) = \sum_{i=1}^{|\mathcal{N}|} e^{-\frac{x - n_i^2}{\sigma^2}}.$$

This message passing process is carried out at all points on  $\mathcal{I}^t$  to define the localized consistency measure on the streamline. An illustration of the consistency measure calculation based on message passing is shown in Fig. 2. From the final consistency measure plotted in Fig. 2 (b), we can see that the local variation of the consistency level has been successfully captured, which will then inform the pruning process for the removal of outlier in fiber bundles.

We consider any point on a streamline with its consistency measure below **THD** as an outlier. To preserve the continuity of each streamline during the filtering process, we filter them at each iteration by pruning the outlier points at both ends. For a streamline  $f^t$ , we obtain the filtered

**3) Track Pruning:** After the estimation of the groupwise consistency measure for all the tracks from all subjects, we filter the fiber tracks via a pruning process. We denote the set of filtered bundles from all subjects as  $\mathcal{F}^t$  at the  $t$ -th iteration and  $p(\mathcal{F}^t)$  as the distribution of the consistency measure of the points on all the fiber bundles in  $\mathcal{F}^t$ . At the  $t$ -th iteration, we first calculate the mean  $\mu_p^t$  and standard deviation  $\sigma_p^t$  of  $p(\mathcal{F}^t)$  to determine the filtering threshold *at the group level*:

$$\mathbf{THD} = \mu_p^t - 2\sigma_p^t.$$

one  $f^{t+1}$  for the next iteration after removing its outlier points.

A filtered track will be rejected if it fails to meet the minimum degree of consistency in  $\mathcal{G}(L; K; \xi)$ , i.e., a sufficient amount of points after the pruning. We thus employ a length constraint  $L_{min}$  to realize this condition on the minimum degree of consistency. In our algorithm, the  $L_{min}$  is a ratio parameter, indicating that minimum ratio of the length of the pruned track to the overall mean streamline length. The mean streamline length, the average number of points on streamlines in input fiber bundles, characteristics the streamline length able to reflect meaningful connectome. A higher  $L_{min}$  requires the filtering result to preserve

more end-to-end interconnectivity. Besides, note that outlier points may exist in the interior of the track, which could result in local inconsistency. We constrain such interior local inconsistency by using another ratio parameter  $L_{max}$  controlling the maximum number of outlier points that are tolerable in the filtered track. The filtered track will be rejected if the ratio of outlier point number to the overall mean streamline length exceed  $L_{max}$ . With a small  $L_{max}$ , the filtering process would be sensitive to the interior local outliers. Together these two parameters will control the filtered track to have enough degree of consistency while keeping the number of residual outlier points small.

**4) Proximity Estimation and Termination Criterion:** Consider that the proximity ( $\xi$ ) constraint locally requires that the filtering result  $\hat{f}$  and its counterparts, white matter trajectories from other subjects, have certain extent of closeness. We use the average distance between the pruned streamline and its references to quantitatively reflect the group-wise closeness as follow:

$$d_{mean}(f^t) = \frac{1}{K \times M} \sum_{r_k^m \in R_k^t, R_k^t \in \mathcal{F}^t} d_{mc}(f^t, r_k^m).$$

The iterative process would not terminate until all pruned streamlines reach the proximity requirement. Thus, the overall groupwise inconsistency,  $\xi^t = \max_{f \in F_i, F_i \in \mathcal{F}^t} d_{mean}(f^t)$ , is used

as the stopping indicator. The pruning process would terminate once  $\xi^t$  is below a certain threshold.  $\delta$ .

The overall implementation of our groupwise filtering algorithm is summarized in Algorithm 1. The operations *ReferenceSet*, *MessagePassing*, *TrackPruning*, and *ProximityEstimation* implement the main steps described above. Once the filtering process stops, the same pruning operations are applied to corresponding points in the original fiber bundles in the last step of the algorithm, which produces the filtered fiber bundles  $\hat{F}$  for all the subjects in the original space. If filtering at the whole streamline level is desired to have complete end-to-end connections, our method can also recover the pruned portion and generate whole-streamline filtering results for tracks that are retained after the group-wise filtering process, i.e., tracks including a portion meeting the groupwise consistency criteria.

While there are multiple parameters in the proposed algorithm, many of the parameters can be set *a priori* and perform robustly across different track filtering tasks. For all our experiments, we set  $r = 0.2$ ,  $M = 3$ ,  $\sigma = 8$  mm, and  $\delta = 3$  mm. The rest of the parameters about groupwise consistency such as affinity  $K$ , anatomical length constraint  $L_{min}$ , and local inconsistency tolerance  $L_{max}$  can be adjusted in different filtering scenarios. We will demonstrate the intuitive ways of parameter setting and corresponding filtering results next in the experiments.

**Algorithm 1.****Groupwise Track Filtering Algorithm**


---

**Input:** Original fiber bundles of  $N$  subjects  $\dot{\mathcal{F}} = \{\dot{F}_k | k = 1, \dots, N\}$   
The corresponding warped bundles in a common space:  $\mathcal{F} = \{F_k | k = 1, \dots, N\}$ .

**Parameters:**  $K$ : affinity parameter,  $L_{min}$ : consistency parameter,  $L_{max}$ : local inconsistency tolerance parameter,  $\delta$ : proximity parameter for termination,  $r$ : subsampling rate,  $M$ : reference set size,  $\sigma$ : distance scale parameter

**Output:** Filtered fiber bundles in original space  $\widehat{\mathcal{F}} = \{\widehat{F}_k | k = 1, \dots, N\}$

**Main Steps:**

$\mathcal{F}^0 \leftarrow \mathcal{F};$   
 $t \leftarrow 0;$   
 $\xi^t \leftarrow \infty;$   
**while**  $\xi^t \leftarrow \delta$  **do**  
  **for**  $\forall f^t \in \mathcal{F}^t$  **do**  
     $\mathcal{G}^t(f^t) \leftarrow \text{ReferenceSet}(f^t, \mathcal{F}, K, M, r);$   
     $p(f^t) \leftarrow \text{MessagePassing}(f^t, \mathcal{G}^t(f^t), \sigma);$   
  **end**  
   $\mathcal{F}^{t+1} \leftarrow \text{TrackPruning}(p(\mathcal{F}^t), L_{min}, L_{max});$   
   $\xi^t \leftarrow \text{ProximityEstimation}(\mathcal{F}^{t+1});$   
   $t \leftarrow t + 1;$   
**end**  
 $\widehat{\mathcal{F}} \leftarrow \mathcal{F}^t;$

---

**3. Experimental results**

In this section, we present experimental results to demonstrate the proposed algorithm on the groupwise filtering and reconstruction of important fiber bundles in human brain. We applied our method to diffusion MRI (dMRI) data from both the second phase of Alzheimer's Disease Neuroimaging Initiative (ADNI2) (Mueller et al., 2005) and Human Connectome Project (HCP) (Van Essen et al., 2012). With the aim of developing biomarkers for the early detection of Alzheimer's disease (AD), the ADNI enrolls subjects ranging from 55 to 90 years old. The study cohort consists of varying disease stages: cognitively normal (CN), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI), and AD. The ADNI2 dMRI data used in our experiment were acquired on 3-T GE Medical Systems scanners. Each diffusion MRI scan contains 59 axial slices reconstructed to  $256 \times 256$  matrix with voxel size  $2.7 \times 2.7 \times 2.7$  mm<sup>3</sup>. Each scan includes 46 separate image volumes: 5 T2-weighted b0 images and 41 diffusion-weighted images ( $b = 1000$  s/mm<sup>2</sup>). The HCP enrolls healthy young adults in the age range of 22–35 years. The advanced multi-shell diffusion MRI data of HCP was acquired on a 3T Siemens Connectome Skyra scanner. The dMRI data of HCP has an isotropic spatial resolution of  $1.25 \times 1.25 \times 1.25$  mm<sup>3</sup> from 270



gradient directions over three b-values ( $b = 1000, 2000, 3000 \text{ s/mm}^2$ ) (Sotiropoulos et al., 2013). In our experiments, we used the preprocessed dMRI data from the 500-Subject release of HCP. For both HCP and ADNI2 dMRI data, we first reconstructed the fiber orientation distribution (FOD) (Tran and Shi, 2015) and then ran FOD-based tractography in MRtrix (Tournier et al., 2019) for ROI-based bundle reconstruction. We used the *iFODI* algorithm in MRtrix for FOD-based probabilistic tractography. As shown in previous validation studies (Aydogan et al., 2018), key parameters including *step\_size*, *angle*, and *cutoff\_threshold* of the FOD at each step of the tractography algorithm all contribute to the regularity of the fiber streamlines. For each fiber bundle, we picked these tractography parameters to ensure a sufficiently complete representation of these bundles are reconstructed according to our experience.

### 3.1. Fornix bundle reconstruction from ADNI2 data

In the first experiment, we applied our method for the groupwise reconstruction of the fornix bundle of 40 subjects from ADNI2 including 10 subjects each from groups with Alzheimer's disease (AD), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI), and cognitively normal (CN). As an important white matter tract of the limbic system, the fornix bundle was shown to be sensitive to the early neurodegeneration in the hippocampus (Mielke et al., 2012). While the fornix anatomy is relatively well described in neuroanatomy (Nieuwenhuys et al., 2008), the limited resolution in clinical dMRI data does not provide sufficient information to accurately identify small ROIs such as the mammillary body that receives fornix projection.

In our experiment, we first manually delineated several ROIs in the T1-weighted MRI of the MNI152 atlas and then registered them to the subject space as the anatomical constraints for fornix reconstruction. As shown in Fig. 3, the seed ROI (in green) and inclusion ROI (in red) corresponding to the two ends of fornix body were depicted on the axial and coronal slices respectively. The exclusion ROIs (in blue) were drawn on the sagittal slices to avoid the tracking artifacts resulting from the entanglement of the fornix and neighboring tracts such as the anterior commissure. To further reduce false positives in bundle reconstruction, the hippocampus masks and cortical regions produced by FreeSurfer (Fischl et al., 2002) were used as inclusion and exclusion ROIs for FOD-based tractography, respectively. Other related tractography parameters are as follows: *step\_size* = 0.2 mm, *angle* =  $6^\circ$ , and *cutoff\_threshold* = 0.025. For each subject, we generated 1000 streamlines for the fornix bundle.

As shown in Fig. 5, the reconstructed fiber bundles contain a large number of outliers. With the following key parameters in our method:  $K = 24$ ,  $L_{min} = 0.6$ , and  $L_{max} = 0.05$ , our method successfully removed outlier portion of the tracks and produced a consistent reconstruction of the fornix body. Note that the affinity  $K = 24$  is 60% of the total number of subjects in the dataset. The pruning results with different termination criteria are demonstrated in Fig. 4, where the filtered bundle of a representative subject was plotted. At the convergence of the algorithm, the reconstructed fiber bundles were obtained and shown in Fig. 5. These results demonstrate that our method is capable of extracting anatomically

meaningful and consistent fiber bundles even from inputs dominated by such highly spurious outliers used in this experiment.

The subjects involved in this experiment are from multiple groups in terms of clinical diagnosis of AD. This results in the potential pathological heterogeneity in the dataset of reconstructed fornix bundle. The affinity parameter  $K$ , which controls the scope of groupwise consistency in population, would have the most significant impact on the filtering results. We examined the impact of the affinity parameter  $K$  qualitatively. By fixing other parameters, we varied the parameter  $K$  in groupwise filtering. The reconstruction results of a representative subject are shown in Fig. 6. We can see residual outliers can still be seen at relatively smaller value ( $K = 8$  or  $16$  which is 20% and 40% of total number of subjects). With the increase of  $K$ , results become more constrained and lead to a reconstruction that underrepresents the fornix. This is especially obvious when  $K$  was chosen as 39, which is the maximum value for a group of 40 subjects. This example also demonstrates that the affinity parameter  $K$  is robust. Changing  $K$  in the range from 16 to 32 results in little geometrical differences in the final filtering results. The observations conform to the expectation that varying the parameter  $K$  allows the trade-off between inter-subject consistency and individual variability. It is also detectable in Fig. 5, that the general morphological characters of each fornix bundle are preserved.

### 3.2. Locus coeruleus pathway and atlas from HCP data

In the updated Braak staging of tau pathology (Braak et al., 2011), the locus coeruleus (LC) nuclei in brainstem was considered the earliest region with tau tangles, one of the defining hallmarks of AD. There is thus increasing interests in studying the LC morphology and connectivity (Clewett et al., 2016). *In vivo* reconstruction of LC pathways in human brains, however, has been relatively under studied. In this experiment, we applied our method to obtain groupwise consistent reconstruction of LC pathways to the medial temporal lobe (MTL), which corresponds to Braak stage I after the LC (Braak stage 0). A robust and consistent reconstruction of the LC pathways could facilitate the investigation of the propagation of tau pathology along fiber pathways (Gibbons et al., 2019) and improve our understanding about the early development of AD.

We used the dMRI data from 50 HCP subjects in this experiment to demonstrate the consistent reconstruction of LC pathways on the right hemisphere of these subjects. Two ROIs were used in the tractography-based reconstruction. The first ROI was the right LC mask nonlinearly warped from an atlas in the MNI152 space (Keren et al., 2009). This ROI was used as the seed region in tractography. The second ROI was the amygdala mask produced by FreeSurfer (Fischl et al., 2002) and used as an inclusion ROI. Parameter settings for FOD-based probabilistic tractography are listed as follows:  $step\_size = 0.125$  mm,  $angle = 4.5^\circ$ , and  $cutoff\_threshold = 0.05$ . Each input LC bundle contains around 1000 streamlines.

As shown in Fig. 8, while the ROIs played an important role in constraining the trajectories of the fiber pathways, the tractography results still tend to be contaminated by erroneous outliers varying from subject to subject. Parameters used in our method were chosen as  $L_{min} = 0.8$ ,  $L_{max} = 0.05$ , and  $K = 49$ . Compared to the filtering implementation for fornix bundle,

we can choose stricter constraints on affinity and consistency in this experiment because of the high anatomical homogeneity across HCP subjects. Constrained only by the ROIs at the two ends, the original LC pathway contains many streamlines with interior false positive portions (pointed and circled out in Fig. 7 (a)). We also show how the maximum outlier length  $L_{max}$  affects the filtering result. As shown in Fig. 7, cleaner reconstruction of the LC pathway can be obtained with the decrease of the  $L_{max}$ . When the  $L_{max}$  is small enough, e.g. less than 0.1, its influence on filtering results almost vanishes. The final reconstruction results are shown in Fig. 8, where clean and consistent reconstruction of the LC pathways have been successfully obtained. These results match very well with the trajectories of the dorsal noradrenergic pathway of the LC as described in previous literature (Marien et al., 2004).

Using the reconstructed LC pathway of the 50 HCP subjects, we created a probabilistic atlas in the MNI152 space following the same approach in (Tang et al., 2018). As shown in Fig. 9, this atlas shows the support of the non-zero regions are compact and well connected, which further confirms the consistent trajectories of the reconstructed pathways.

### 3.3. Quantitative comparison of corticospinal tract reconstruction from HCP data

In the third experiment, we applied our groupwise filtering method to the reconstruction of the corticospinal tract (CST) and quantitatively compared its performance with two publicly available methods. The data of 20 HCP subjects from a previous brainstem atlas project (Tang et al., 2018) was used in this experiment. For each CST bundle, five ROIs in the brainstem region (Fig. 11 (b)) were manually delineated by an experienced neuroanatomist in (Tang et al., 2018) to guide the accurate reconstruction of the CST. To evaluate the performance of different track filtering methods, we used only one brainstem ROI (ROI 1 as shown in Fig. 11 (a)) in tractography-based reconstruction of the left CST of each subject. The other four ROIs were used as ground truth to quantitatively measure the accuracy of filtered tracks by different methods. With the brainstem ROI as an inclusion ROI, we also used the left precentral gyrus from the FreeSurfer Aseg labels (Fischl et al., 2002) as the seed region. For FOD-based tractography, the parameter setting is as follows:  $step\_size = 0.125$  mm,  $angle = 4^\circ$ , and  $cutoff\_threshold = 0.025$ . Each reconstructed CST contains around 500 streamlines.

From the input CST bundles shown in the first row of Fig. 10, we can see some frequent outliers in the brainstem area and the lateral projections to part of the precentral gyrus that do not contribute to the CST. Following the similar parameter selection strategy presented in the LC pathway filtering experiment, we used the following parameters:  $K = 19$ ,  $L_{min} = 0.8$ , and  $L_{max} = 0.01$  for our groupwise filtering algorithm. We selected a small  $L_{max}$  to remove streamlines with short false positive segments (pointed out by arrow in Fig. 10). The high consistency requirement,  $L_{min} = 0.8$ , lead the filtering process to produce result with a high degree of consistency. As shown in the second row of Fig. 10, our method successfully removed these outliers and generated consistent and clean bundles that follow the correct anatomy. We also displayed the filtering results at the whole-stream level by adding back the pruned points for tracks retained by the filtering algorithm, highlighting the end-to-end connectivity from the motor cortex to the spinal cord, in the bottom row of Fig. 10. For the

filtered bundle from all subjects, we calculated the number of tracks and listed their distribution in Table 2, which will be used to guide parameter tuning in the tools we will compare with.

The first publicly available method we compare with is the track filtering method in the QuickBundles software tool (Garyfallidis et al., 2012), which was applied to the same input bundles as our method. More specifically, we chose the threshold for distance between curves as 5 mm and cluster size as 70 streamlines in QuickBundles, which means the maximum Minimum Average Direct-flip (MDF) distance between curves within a cluster was limited to 5 mm, and all streamlines belonging to clusters with less than 70 curves were discarded. With the increase of the threshold for cluster size, more outliers will be removed but also potentially valid tracks. For all subjects we counted the number of tracks in the filtered fiber bundles and fine-tuned the threshold of cluster size such that the lower end of the track number distribution will be slightly below our method as listed in Table 2. This suggests we have filtered a comparable or more outlier tracks with the QuickBundles method as compared to our method. A comparison of the filtering results from QuickBundles and our method on an HCP subject is shown in Fig. 11. While QuickBundles successfully filtered out the outlier tracks projecting to the inferior and lateral portion of the precentral gyrus, it did not completely remove the tracks with defects in the brainstem area, which our method was able to handle consistently across the group. We also compared our method with the cluster confidence index (CCI) based streamline filtering (Jordan et al., 2018). Given a streamline, the streamlines within a certain MDF distance ( $\theta_{cci}$ ) are employed as references. The CCI qualitatively reflects the reproducibility of individual streamline according to the overall similarity with its reference, where the streamline-wise similarity is characterized by the  $K_{cci}$ -th power of the reciprocal of the MDF distance. We set  $K_{cci} = 1$ ,  $\theta_{cci} = 5$  mm and the CCI threshold was fine-tuned to be 30 based on the same criteria used above for QuickBundles, i.e., the lower end of the track number distribution will be slightly below our method to ensure a comparable number of outliers were removed (Table 2). Qualitatively, a filtering result example from CCI based method was demonstrated in Fig. 11 (d), where we can observe that inconsistent streamlines were removed properly at the cost of removing more valid tracts in comparison with our method. Both QuickBundles and CCI based filtering method have high computational efficiency by completing the processing of the 20 CST bundles in around 10 s.

Two automatic bundle reconstruction approaches, introduced in (Garyfallidis et al., 2018), and (Zhang et al., 2018) are tested for further comparisons. These methods used whole brain tractography as the input and extracted individual fiber bundles based on the precomputed tractography atlases. To apply these methods to the 20 HCP subjects used in our experiment, we first generated whole brain tractography containing 100k streamlines (top right in Fig. 11 (e)) for each subject using the FOD-based probabilistic tractography of MRtrix (Tournier et al., 2019). The following parameter setting was used for the tractography:  $step\_size = 0.125$  mm,  $angle = 4^\circ$ , and  $cutoff\_threshold = 0.025$ .

The first atlas based approach, which has been implemented as the RecoBundles tool in Dipy (Garyfallidis et al., 2014), was applied for the left CST reconstruction. In this approach, the whole brain tractography of each subject was first registered to a population-

average tractography atlas (Yeh et al., 2018) using the nonlinear registration computed by the ANTS software (Avants et al., 2008). Following the guidance of parameter setting in (Garyfallidis et al., 2018), we chose the following parameters for RecoBundles:  $cluster\_threshold = 15$  mm,  $model\_cluster\_threshold = 5$  mm, and  $reduction\_threshold = 20$  mm and fine-tuned the parameter  $pruning\_threshold$  to be 6 mm so that the lower end of track number distribution will be slightly below our method (Table 2) because this method also follows similar techniques from QuickBundles. As demonstrated in Fig. 11 (e), we can see that the extracted fiber bundle generally follows the trajectory of CST, but many of the tracks extracted from the whole brain tractography terminated prematurely before reaching the end of the medulla. The second atlas-based approach (Zhang et al., 2018) has been distributed as part of the *whitematteranalysis* tool in the SlicerDMRI project (Norton et al., 2017). Both affine and nonrigid registration (O'Donnell et al., 2012) were computed to warp the tracts to the atlas space by using tools provided in SlicerDMRI. After that, the atlas-based method in SlicerDMRI was applied to extract the left CST as shown in Fig. 11 (f). Compared to the manually delineated ROIs, we can see a large number of false positives in the brainstem area were included in this reconstruction.

To quantitatively compare the performance of different methods, we calculated the distance between the fiber bundles and the other four manually delineated ROIs (ROI 2–5) not used in bundle reconstruction. Because all the ROIs were delineated on axial slices, we denote the set of points in the  $j$ -th ROI of the  $i$ -th subject as  $U_i^j$ : Given a fiber bundle of the  $i$ -th subject, we denote its intersection with the corresponding axial slice of the  $j$ -th ROI as the point set  $V_i^j$ . For comparison, the Hausdorff distance  $d_H(U_i^j, V_i^j)$  between these two-point sets was computed for the original input bundles, filtered bundles from our method, QuickBundles, and CCI based reconstruction. These Hausdorff distances indicate the mismatches between streamlines and the underlying anatomy delineated by ROIs. Statistics of the Hausdorff distances for each method and ROI were reported in Table 1. Our method, QuickBundles, and CCI based filtering used the same input bundles (*Original CST* as listed in Table 1), and the results show that our algorithm achieved the best performance while all three methods were able to enhance the fidelity to manually delineated ROIs. We also estimated the Hausdorff distance between these four ROIs (ROI 2–5) and CSTs generated by the two atlas-based tools in RecoBundles and SlicerDMRI. The results in Table 1 suggests much larger errors as compared to the manually delineated labels in the brainstem. This is consistent with the large number of outliers as illustrated in Fig. 11 (e) and (f). Because the results from SlicerDMRI contain much larger number of tracks than other method, we applied manually delineated ROI 1 as an inclusion region to the bundle reconstructed by SlicerDMRI for a fair comparison to other methods. This removed many of the outliers and improved distance measure to other ROIs are listed in Table 1 as *SlicerDMRI + ROI 1*, which are comparable to the performance of the input bundles listed as *Original CST*. Such anatomical constraints were not imposed for CSTs generated by RecoBundles because it tends to extract tracks from the whole brain tractography that terminate early and rarely reaches ROI 1.

With the manually delineated ground truth labels for brainstem ROIs, we can also demonstrate the impact of parameters used in our method. Because the CST bundle has clear

end ROIs on both the cortical and brainstem area, we have selected very high *degree* related parameters (large  $L_{min} = 0.8$  and smaller  $L_{max} = 0.01$ ). In addition, we have selected a high affinity parameter ( $K = 19$ ) because the prior knowledge that HCP subjects are young and healthy. Thus, we focused on the fine tuning of parameters related to the proximity condition:  $\sigma$  and  $\delta$ , and examined their impact on the Hausdorff distances of the reconstructed bundle with respect to the ground truth labels. As shown in Fig. 12, the colored curves and associated shaded regions demonstrate the mean and standard deviations of the Hausdorff distances from the reconstructed bundle to each ROI with respect to the change of these two parameters. The black dots show the number of fiber tracks in the reconstructed bundle under each parameter value. With the increase of  $\sigma$ , the consistent measure computed from message passing becomes more insensitive to proximity conditions and could result in the pruning of more tracks. From Fig. 12 (a), we can see that  $\sigma = 6 \sim 8$  mm would be good trade-offs. For the  $\delta$  parameter, its decrease will lead to stricter stopping criteria and the elimination of more tracks. We thus set it as  $\delta = 3$  mm to ensure a sufficient number of tracks can be retained (around or above 100 for the CST bundle) in the final reconstruction. Overall, the fine tuning of the parameters strikes a balance between applying the proximity condition and ensuring enough number of tracks to effectively represent the geometry of the fiber bundle.

### 3.4. Computational cost

In all experiments, our method was implemented in python on a desktop computer with 3.60 GHz Intel i7-6850K CPUs and 64 GB RAM. We conducted each experiment ten times to obtain a robust estimation of running times. The overall processing times of all subjects in the experiments are summarized in Table 3.

## 4. Discussion and conclusion

In this work we developed a novel groupwise track filtering algorithm for the consistent reconstruction of fiber bundles from diffusion imaging data. Our method is based on a flexible definition of groupwise consistency that controls the *degree*, *affinity*, and *proximity* of each track with respect to other group members. A key element of our algorithm is the dynamic construction and update of a reference set for each track that allows the efficient implementation of *localized* consistency evaluation based on message passing and outlier pruning. In summary, the main contributions of our work are a) Proposed a general conceptual framework for characterizing groupwise consistency of fiber tracks; b) Developed a novel numerical algorithm that iteratively and locally prunes away inconsistent portion of each track; c) Demonstrate the general applicability of the proposed algorithm on fiber bundles with varying level of artifacts and complexity; d) Performed quantitative comparisons based on ground truth from manually delineated labels and showed that groupwise filtering can compensate for the gap in anatomical knowledge and achieve more faithful reconstruction of fiber bundles.

Accurate bundle reconstruction relies on intensive anatomical priors. Multiple automated bundle reconstruction methods such as TRACULA (Yendiki et al., 2011), *whitematteranalysis* (O'Donnell and Westin, 2007; Zhang et al., 2018), Recobundles

(Garyfallidis et al., 2018), TractSeg (Wasserthal et al., 2018), WMQL (Wassermann et al., 2016) taking advantages of the anatomical priors in the form of segmentations of reference tracts, tractography atlas, and brain parcellation atlases to effectively and efficiently reconstruct various fiber bundles. With these methods, the analysis of white matter at the fiber bundle level becomes very conveniently. While limited attention was paid for considering whether each individual streamline in the fiber bundles is reliable and reproducible. From (Rheault et al., 2020), we know that the reproducibility of streamlines is much more sensitive than which of whole bundle volume. If we would like to leverage the information at the streamline level, it is necessary to determine how reliable each streamline is. In (Jordan et al., 2018), the cluster confidence index is proposed to quantitatively indicate the reliability of individual streamline according to its similarity with neighborhoods. In this work, we further generalize the reliability concept to be the reproducibility across subjects (groupwise consistency) and construct a framework to extract the most reproducible sub-bundle structures.

The underlying assumption for the success of the proposed method is the existence of certain level of commonness in the fiber bundles across subjects. This is commonly adopted in brain mapping research, where image or surface registration was first applied to factor out variability across population before group level analysis. In our experiments, we warp all fiber trajectories into a common space using nonlinear image registration (Avants et al., 2008) before the filtering process. This ensures that individual variability in white matter anatomy is taken into account. Tractography registration methods proposed in (Garyfallidis et al., 2015; O'Donnell et al., 2012) could also be useful to align the fiber bundles before the application of our groupwise filtering algorithm. Without referring to the images, these registration methods have the potential advantages of handling the bundle alignment task from patients with severe white matter atrophy, lesion, and tumor.

In the algorithmic implementation of the proposed groupwise filtering framework, the parameters used for controlling the *degree* ( $L_{min}$  and  $L_{max}$ ) and *affinity* measures ( $K$ ) are typically chosen according to the prior knowledge with respect to individual fiber bundles. For fiber bundles without stringent ROIs that determine the end points of the tracks, we demonstrate a proper choice of the degree parameter can help prune away spurious portions on both ends of the fiber bundle and produce a consistent representation of the fornix bundle. This strategy can be generally applicable to various fiber bundles that project diffusively to broad cortical areas and therefore have high variability in their tractography based reconstruction. For example, the anterior commissure projects to the broad areas including the middle and inferior temporal gyrus and would benefit from our method to achieve a consistent reconstruction for comparison across subjects. For fiber bundles that can be defined with relatively precise end ROIs, we can increase the *degree* requirement (larger  $L_{min}$  and smaller  $L_{max}$ ). Similar to existing track filtering methods, this will achieve essentially a binary decision (accept/reject) on each streamline with the assistance of groupwise consistency. For the affinity measure, our method allows its adjustment for fiber bundles or cohorts with varying levels of heterogeneity across the group. For the fornix bundle reconstruction from the ADNI data, we demonstrated the robustness of the results with respect to the change of the affinity ( $K$ ) parameter and the preservation of subject level variability in the reconstructed fiber bundles. For young and healthy subjects from the HCP,

a higher affinity parameter is selected in our experiments that reflects the prior knowledge about the higher degree of similarity in this cohort. For the *proximity* condition, the strengthening of its requirement (smaller  $\delta$ ) will result in the filtering of more artifacts and potentially the removal of valid tracks from the reconstructed fiber bundles. As demonstrated in Fig. 12, we chose proximity parameter by balancing the removal of artifacts and the preservation of a sufficient number of valid tracks in the CST bundle. Overall, the parameters for the degree, affinity and proximity conditions in our method can be selected intuitively because the clear expectation as explained above about their effects on the filtered filter bundles.

As a fiber bundle filtering tool, our method depends on the original inputs and could be affected by the bias issue of tractography algorithms (Maier-Hein et al., 2017). In our CST experiment, the fiber pathways emanating from the lateral portion of the precentral gyrus are harder to reconstruct because the need of crossing regions with complicated fiber geometry and the dramatic turning angles as they join the descending portion of the CST at the internal capsule. Fiber tracks from the lateral portion of the motor cortex thus tend to be severely under-represented and highly variable across subjects. As a result, our method will filter out these tracks due to their lack of groupwise consistency. Improvements in tractography algorithms will help provide more balanced representation of input fiber bundles to our algorithm and hence generate better reconstruction results. For example, the Anatomically Constrained Tractography (ACT) (Smith et al., 2012) can improve tractography quality and hence the downstream filtering algorithms. Recently (Rheault et al., 2019), proposed a tractography algorithm to mitigate the bias in the fiber bundle reconstruction by introducing anatomical and orientational prior knowledge for tractography. For future work, we will investigate such tractography tools and examine their impact on our groupwise filtering method. In addition, it will be highly valuable to perform validations against ground truth provided by tracer injection data in other anatomical regions such as the internal capsule for the CST bundle (Innocenti et al., 2018).

In summary, we developed an iterative algorithm to prune away inconsistent artifacts of fiber tracks for the reconstruction of fiber bundles with groupwise consistency. Results from our method can improve the comparability of bundle-based representations of white matter connectivity and hence potentially provide increased power in the detection of group differences. In future work, we will apply our method to perform more extensive validations and study the impact of groupwise filtering on the detection of connectivity changes of critical fiber bundles in brain disorder such as the Alzheimer's disease.

## Acknowledgements

This work was in part supported by the National Institute of Health (NIH) under grants RF1AG056573, R01EB022744, R21AG064776, P41EB015922, P50AG05142. The brainstem region masks for the corticospinal tracts used in this work were from a previously published brainstem pathway atlas (Tang et al., 2018), where all manually delineated masks were generated by an experienced neuroanatomist Dr. Yuchun Tang when he was a postdoctoral fellow in Dr. Yonggang Shi's group. Data used in this paper were partly provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Data used in this paper were partly provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012).



ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

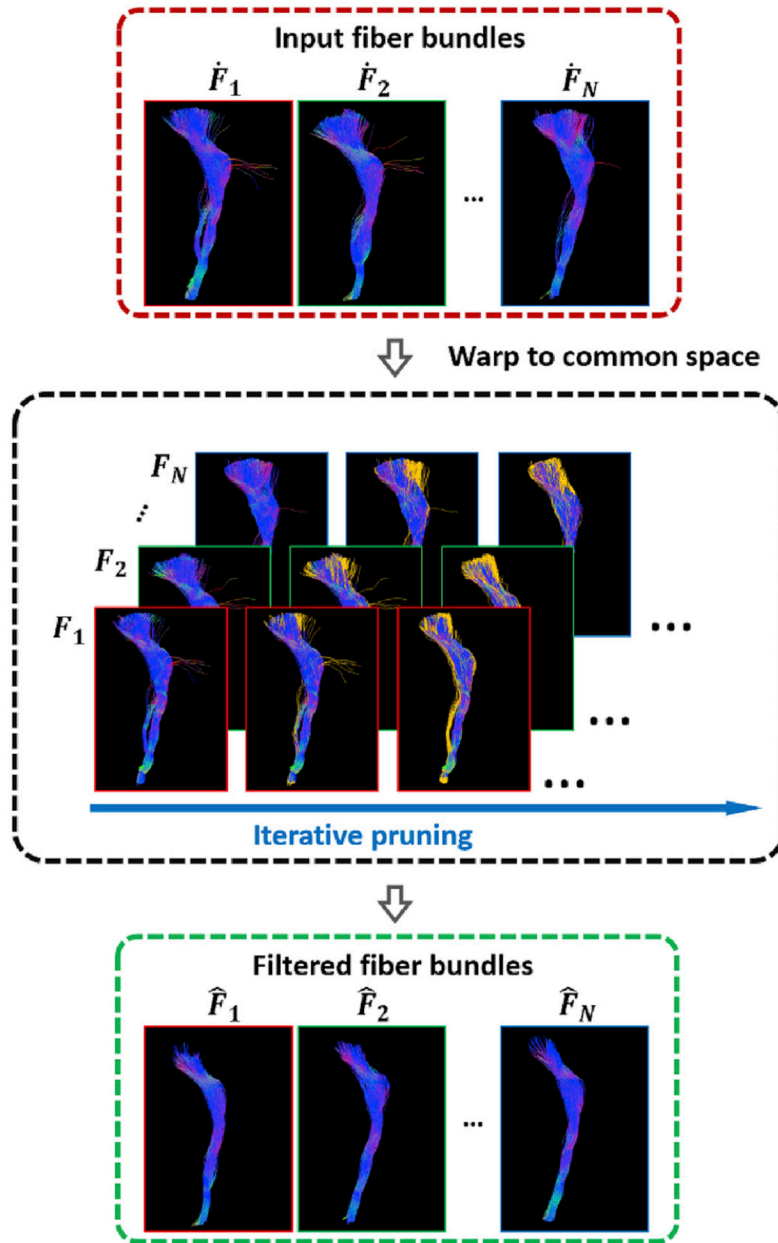
## References

- Avants BB, Epstein CL, Grossman M, Gee JC, 2008 Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal* 12, 26–41. [PubMed: 17659998]
- Aydogan DB, Shi Y, 2015 Track filtering via iterative correction of TDI topology. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 20–27.
- Aydogan DB, Jacobs R, Dulawa S, Thompson SL, Francois MC, Toga AW, Dong H, Knowles JA, Shi Y, 2018 When tractography meets tracer injections: a systematic study of trends and variation sources of diffusion-based connectivity. *Brain Struct. Funct* 223, 2841–2858. [PubMed: 29663135]
- Basser PJ, Mattiello J, LeBihan D, 1994 MR diffusion tensor spectroscopy and imaging. *Biophys. J* 66, 259–267. [PubMed: 8130344]
- Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A, 2000 In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med* 44, 625–632. [PubMed: 11025519]
- Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW, 2007 Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *Neuroimage* 34, 144–155. [PubMed: 17070705]
- Braak H, Thal DR, Ghebremedhin E, Del Tredici K, 2011 Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. *J. Neuropathol. Exp. Neurol* 70, 960–969. [PubMed: 22002422]
- Catani M, Howard RJ, Pajevic S, Jones DK, 2002 Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *Neuroimage* 17, 77–94. [PubMed: 12482069]
- Clewett DV, Lee T-H, Greening S, Ponzio A, Margalit E, Mather M, 2016 Neuromelanin marks the spot: identifying a locus coeruleus biomarker of cognitive reserve in healthy aging. *Neurobiol. Aging* 37, 117–126. [PubMed: 26521135]
- Corouge I, Gouttard S, Gerig G, 2004 Towards a shape model of white matter fiber bundles using diffusion tensor MRI In: *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium On*, pp. 344–347.
- Daducci A, Dal Palù A, Lemkaddem A, Thiran J, 2015 COMMIT: convex optimization modeling for microstructure informed tractography. *IEEE Trans. Med. Imag* 34, 246–257.
- Descoteaux M, Deriche R, Knosche TR, Anwander A, 2009 Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Trans. Med. Imag* 28, 269–286.
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM, 2002 Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355. [PubMed: 11832223]
- Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL, 2011 Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* 54, 313–327. 10.1016/j.neuroimage.2010.07.033. [PubMed: 20656036]

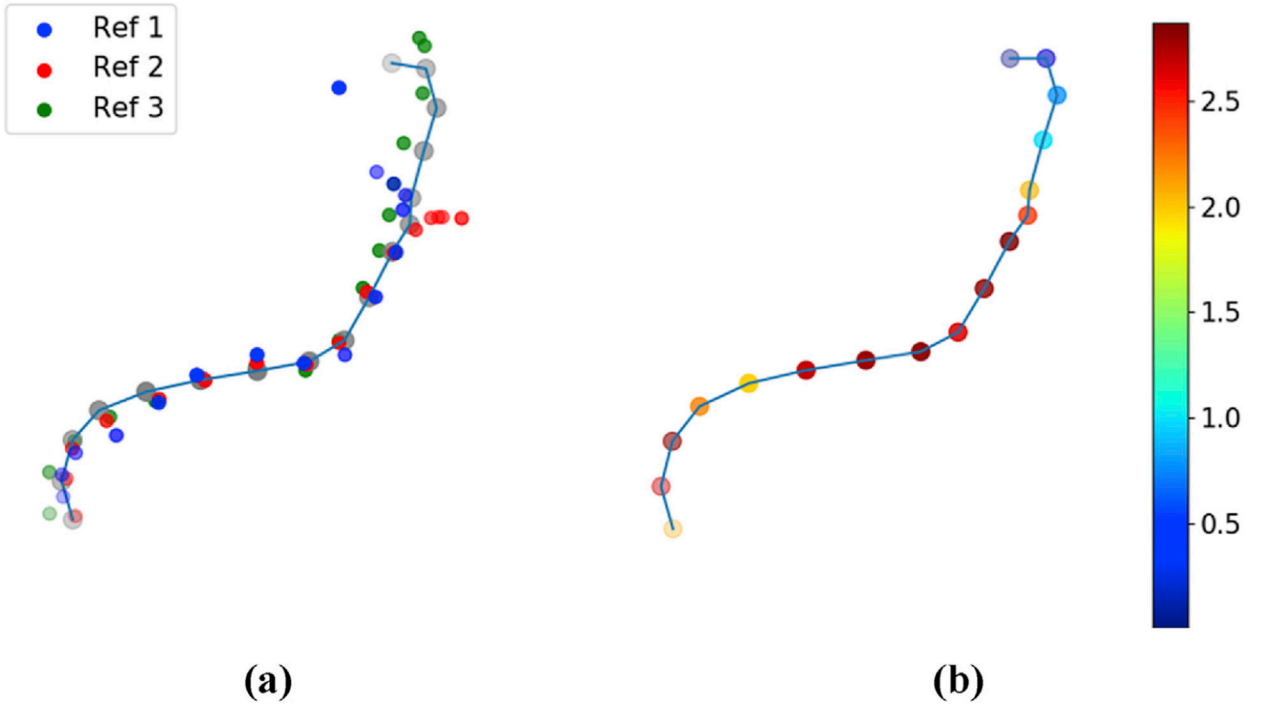
- Garyfallidis E, Brett M, Correia MM, Williams GB, Nimmo-Smith I, 2012 QuickBundles, a method for tractography simplification. *Front. Neurosci* 6, 1–13. 10.3389/fnins.2012.00175. [PubMed: 22294978]
- Garyfallidis E, Brett M, Amirbekian B, Rokem A, van der Walt S, Descoteaux M, Nimmo-Smith I, 2014 Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinf* 8, 1–17. 10.3389/fninf.2014.00008.
- Garyfallidis E, Ocegueda O, Wassermann D, Descoteaux M, 2015 Robust and efficient linear registration of white-matter fascicles in the space of streamlines. *Neuroimage* 117, 124–140. 10.1016/j.neuroimage.2015.05.016. [PubMed: 25987367]
- Garyfallidis E, Côté MA, Rheault F, Sidhu J, Hau J, Petit L, Fortin D, Cunanne S, Descoteaux M, 2018 Recognition of white matter bundles using local and global streamline-based registration and clustering. *Neuroimage* 170, 283–295. 10.1016/j.neuroimage.2017.07.015. [PubMed: 28712994]
- Gibbons GS, Lee VMY, Trojanowski JQ, 2019 Mechanisms of cell-to-cell transmission of pathological tau: a review. *JAMA Neurol* 76, 101–108. [PubMed: 30193298]
- Guevara P, Poupon C, Riviere D, Cointepas Y, Descoteaux M, Thirion B, Mangin J-F, 2011 Robust clustering of massive tractography datasets. *Neuroimage* 54, 1975–1993. [PubMed: 20965259]
- Innocenti GM, Caminiti R, Rouiller EM, Knott G, Dyrby TB, Descoteaux M, Thiran J-P, 2018 Diversity of cortico-descending projections: histological and diffusion MRI characterization in the monkey. *Cerebr. Cortex* 29, 788–801.
- Jordan KM, Amirbekian B, Keshavan A, Henry RG, 2018 Cluster confidence index: a streamline-wise pathway reproducibility metric for diffusion-weighted MRI tractography. *J. Neuroimaging* 28, 64–69. 10.1111/jon.12467. [PubMed: 28940825]
- Keren NI, Lozar CT, Harris KC, Morgan PS, Eckert MA, 2009 In vivo mapping of the human locus coeruleus. *Neuroimage* 47, 1261–1267. [PubMed: 19524044]
- Maier-Hein KH, Neher PF, Houde J-C, Côté M-A, Garyfallidis E, Zhong J, Chamberland M, Yeh F, Lin Y, Ji Q, Reddick WE, Glass JO, Chen DQ, Feng Y, Gao C, Wu Y, Ma J, He R, Li Q, Westin C, Deslauriers-Gauthier S, Omar Ocegueda Gonz alez J, Paquette M, St-Jean S, Girard G, Rheault F, Sidhu J, Tax CMW, Guo F, Mesri H, Y, D avid S, Froeling M, Heemskerk AM, Leemans A, Bor e A, Pinsard B, Bedetti C, Desrosiers M, Brambati S, Doyon J, Sarica A, Vasta R, Cerasa A, Quattrone A, Yeatman J, Khan AR, Hodges W, Alexander S, Romascano D, Barakovic M, Auría A, Esteban O, Lemkaddem A, Thiran J, Cetingul HE, Odry BL, Boris Mailhe B, Mariappan S, Nadar MS, Pizzagalli F, Prasad G, Julio E, Villalon-Reina JE, Galvis J, Thompson PM, Requejo FDS, Laguna RL, Lacerda LM, Barrett R, Dell'Acqua F, Catani M, Petit L, Caruyer E, Daducci A, Dyrby TB, Holland-Letz T, Hilgetag CC, Stieltjes B, Descoteaux M, 2017 The challenge of mapping the human connectome based on diffusion tractography. *Nat. Commun* 8, 1349. [PubMed: 29116093]
- Marien MR, Colpaert FC, Rosenquist AC, 2004 Noradrenergic mechanisms in neurodegenerative diseases: a theory. *Brain Res. Rev* 45, 38–78. [PubMed: 15063099]
- Mielke MM, Okonkwo OC, Oishi K, Mori S, Tighe S, Miller MI, Ceritoglu C, Brown T, Albert M, Lyketsos CG, 2012 Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease. *Alzheimer's Dementia* 8, 105–113.
- Mori S, Crain BJ, Chacko VP, Van Zijl PCM, 1999 Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann. Neurol* 45, 265–269. [PubMed: 9989633]
- Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, Trojanowski JQ, Toga AW, Beckett L, 2005 The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin* 15, 869–877.
- Nieuwenhuys R, Voogd J, van Huijzen C, 2008 *The Human Central Nervous System*. Steinkopff-Verlag Heidelberg
- Norton I, Essayed WI, Zhang F, Pujol S, Yarmarkovich A, Golby AJ, Kindlmann G, Wassermann D, Estepar RSJ, Rathi Y, Pieper S, Kikinis R, Johnson HJ, Westin C, O'Donnell LJ, 2017 SlicerDMRI: open source diffusion MRI software for brain cancer research. *Canc. Res* 77, e101–e103.
- O'Donnell LJ, Westin C-F, 2007 Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Trans. Med. Imag* 26, 1562–1575.

- O'Donnell LJ, Wells WM, Golby AJ, Westin CF, 2012 Unbiased groupwise registration of white matter tractography. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 7512 LNCS 123–130. 10.1007/978-3-642-33454-2\_16.
- Pestilli F, Yeatman JD, Rokem A, Kay KN, Wandell BA, 2014 Evaluation and statistical inference for living connectomes. *Nat. Methods* 11, 1058–1063. [PubMed: 25194848]
- Rheault F, St-Onge E, Sidhu J, Maier-Hein K, Tzourio-Mazoyer N, Petit L, Descoteaux M, 2019 Bundle-specific tractography with incorporated anatomical and orientational priors. *Neuroimage* 186, 382–398. 10.1016/j.neuroimage.2018.11.018. [PubMed: 30453031]
- Rheault F, De Benedictis A, Daducci A, Maffei C, Tax CMW, Romascano D, Caverzasi E, Morency FC, Corrivetti F, Pestilli F, Girard G, Theaud G, Zemmoura I, Hau J, Glavin K, Jordan KM, Pomiecko K, Chamberland M, Barakovic M, Goyette N, Poulin P, Chenot Q, Panesar SS, Sarubbo S, Petit L, Descoteaux M, 2020 Tractostorm: the what, why, and how of tractography dissection reproducibility. *Hum. Brain Mapp* 41, 1859–1874. 10.1002/hbm.24917. [PubMed: 31925871]
- Smith RE, Tournier J-D, Calamante F, Connelly A, 2012 Anatomically-constrained tractography: improved diffusion MRI streamlines tractography through effective use of anatomical information. *Neuroimage* 62, 1924–1938. [PubMed: 22705374]
- Smith RE, Tournier J-D, Calamante F, Connelly A, 2015 SIFT2: enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. *Neuroimage* 119, 338–351. [PubMed: 26163802]
- Sotiropoulos SN, Jbabdi S, Xu J, Andersson JL, Moeller S, Auerbach EJ, Glasser MF, Hernandez M, Sapiro G, Jenkinson M, Feinberg DA, Yacoub E, Lenglet C, Van Essen DC, Ugurbil K, Behrens TEJ, WU-Minn HCP Consortium, 2013 Advances in diffusion MRI acquisition and processing in the human connectome project. *Neuroimage* 80, 125–143. [PubMed: 23702418]
- Tang Y, Sun W, Toga AW, Ringman JM, Shi Y, 2018 A probabilistic atlas of human brainstem pathways based on connectome imaging data. *Neuroimage* 169, 227–239. [PubMed: 29253653]
- Tournier JD, Smith R, Raffelt D, Tabbara R, Dhollander T, Pietsch M, Christiaens D, Jeurissen B, Yeh CH, Connelly A, 2019 MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage* 202, 116137 10.1016/j.neuroimage.2019.116137.
- Tran G, Shi Y, 2015 Fiber orientation and compartment parameter estimation from multi-shell diffusion imaging. *IEEE Trans. Med. Imag* 34, 2320–2332.
- Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, Della Penna S, Feinberg D, Glasser MF, Harel N, Heath AC, Larson-Prior L, Marcus D, Michalareas G, Moeller S, Oostenveld R, Petersen SE, Prior F, Schlaggar BL, Smith SM, Snyder AZ, Xu J, Yacoub E, WU-Minn HCP Consortium, 2012 The Human Connectome Project: a data acquisition perspective. *Neuroimage* 62, 2222–2231. [PubMed: 22366334]
- Wainwright MJ, Jaakkola TS, Willsky AS, 2005 MAP estimation via agreement on trees: message-passing and linear programming. *IEEE Trans. Inf. Theor* 51, 3697–3717.
- Wakana S, Jiang H, Nagae-Poetscher LM, van Zijl PCM, Mori S, 2004 Fiber tract-based atlas of human white matter anatomy. *Radiology* 230, 77–87. [PubMed: 14645885]
- Wang J, Shi Y, 2019 A fast fiber k-Nearest-Neighbor algorithm with application to group-wise white matter topography analysis In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* Springer International Publishing 10.1007/978-3-030-20351-1\_25.
- Wang J, Aydogan DB, Varma R, Toga AW, Shi Y, 2018 Modeling topographic regularity in structural brain connectivity with application to tractogram filtering. *Neuroimage* 183, 87–98. [PubMed: 30081193]
- Wassermann D, Makris N, Rathi Y, Shenton M, Kikinis R, Kubicki M, Westin CF, 2016 The white matter query language: a novel approach for describing human white matter anatomy. *Brain Struct. Funct* 221, 4705–4721. 10.1007/s00429-015-1179-4. [PubMed: 26754839]
- Wasserthal J, Neher P, Maier-Hein KH, 2018 TractSeg - fast and accurate white matter tract segmentation. *Neuroimage* 183, 239–253. 10.1016/j.neuroimage.2018.07.070. [PubMed: 30086412]

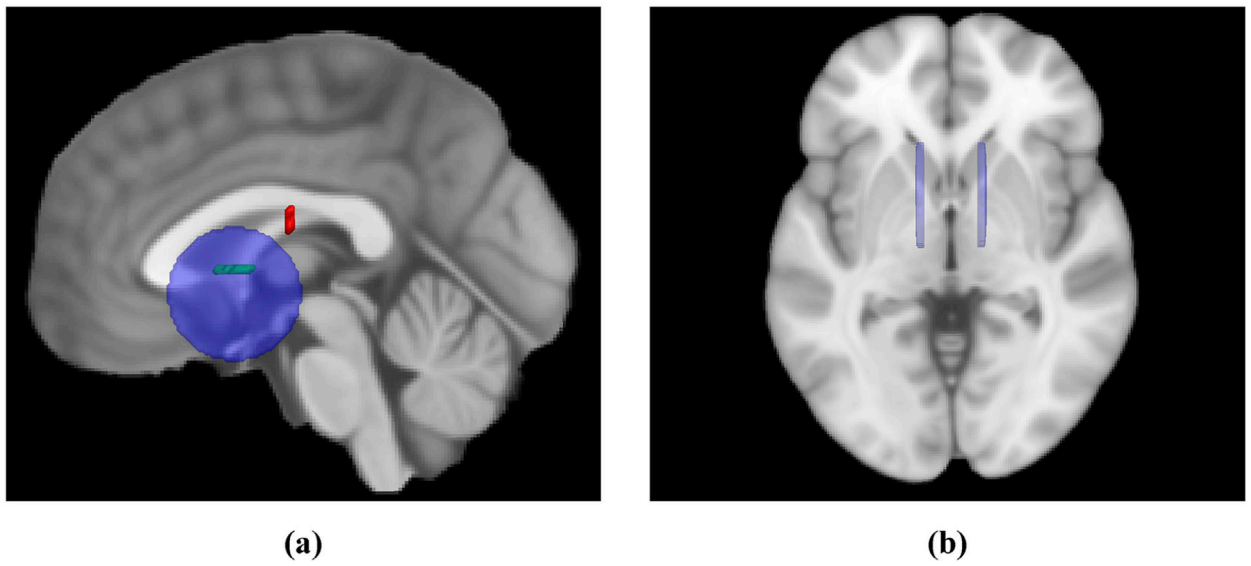
- Yeh FC, Panesar S, Fernandes D, Meola A, Yoshino M, Fernandez-Miranda JC, Vettel JM, Verstynen T, 2018 Population-averaged atlas of the macroscale human structural connectome and its network topology. *Neuroimage* 178, 57–68. 10.1016/j.neuroimage.2018.05.027. [PubMed: 29758339]
- Yendiki A, Panneck P, Srinivasan P, Stevens A, Zöllei L, Augustinack J, Wang R, Salat D, Ehrlich S, Behrens T, Jbabdi S, Gollub R, Fischl B, 2011 Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Front. Neuroinf* 5, 1–12. 10.3389/fninf.2011.00023.
- Zhang F, Wu Y, Norton I, Rigolo L, Rathi Y, Makris N, et al., 2018 An anatomically curated fiber clustering white matter atlas for consistent white matter tract parcellation across the lifespan. *Neuroimage* 179, 429–447. [PubMed: 29920375]

**Fig. 1.**

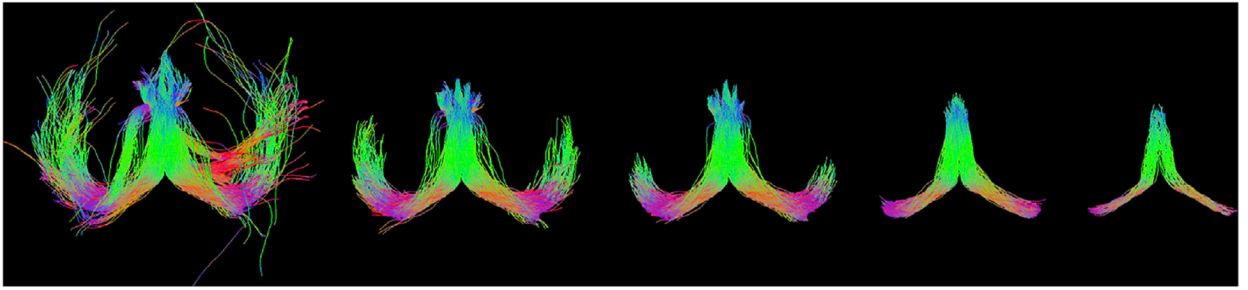
An illustration of the proposed groupwise track filtering framework. All fiber bundles from  $N$  subjects will first be nonlinearly warped to a common space such as the MNI152 atlas. Outlier points (colored in yellow) will be iteratively pruned based on the consistency measure obtained via messages passed from tracks in a reference set, which is updated dynamically. Once the pruning process is completed, the same pruning operations are applied to points on the original tracks to obtain the filtered fiber bundles.



**Fig. 2.** Calculation of localized consistency measures on a track based on message passing. (a) The points on the track under consideration (cyan) are plotted as gray dots. The points on the three reference tracks are plotted in red, green, and blue, respectively. (b) Each point on the track under consideration is color-coded according to the consistency measure from messages they receive from neighboring tracks. The consistency measures range from 0 to 3 as 3 reference tracks are used in this example.

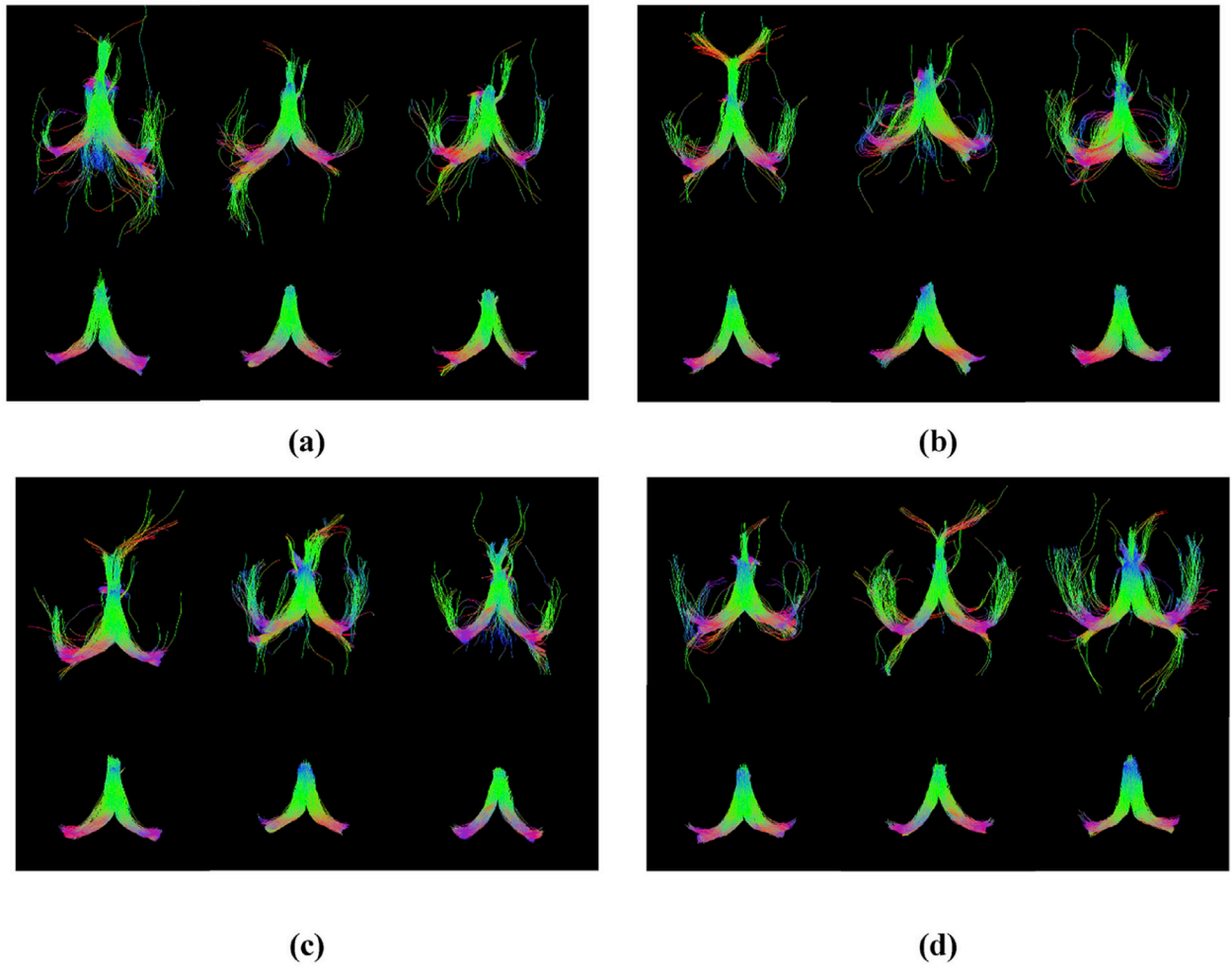


**Fig. 3.** ROIs used for fornix bundle reconstruction are plotted against the MNI152 T1 image. (a) A sagittal view of the manually drawn seed (green), inclusion (red), and exclusion (blue) ROIs. (b) An axial view of the manually draw exclusion ROIs. Note they do not overlap with the seed and inclusion ROIs.

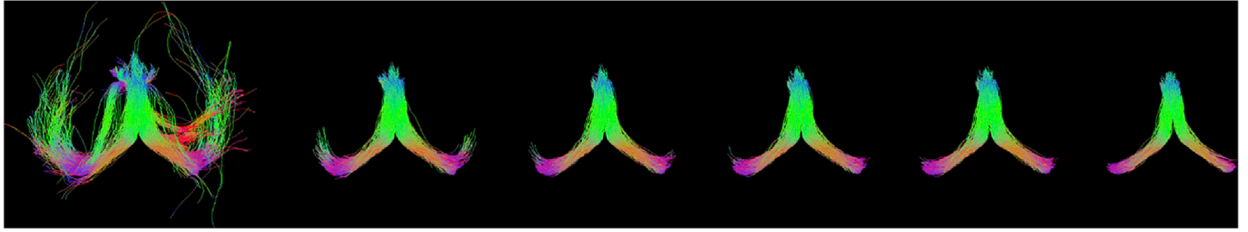


**Fig. 4.** The iterative pruning results with different proximity requirement of the fornix bundle of an EMCI subject. The original bundle, and the filtering outcomes with proximity threshold  $\delta = 20$  mm, 7 mm, 3 mm, and 2 mm, corresponding to the results after 1, 2, 5, and 10 iterations of filtering, are displayed from left to right, respectively.

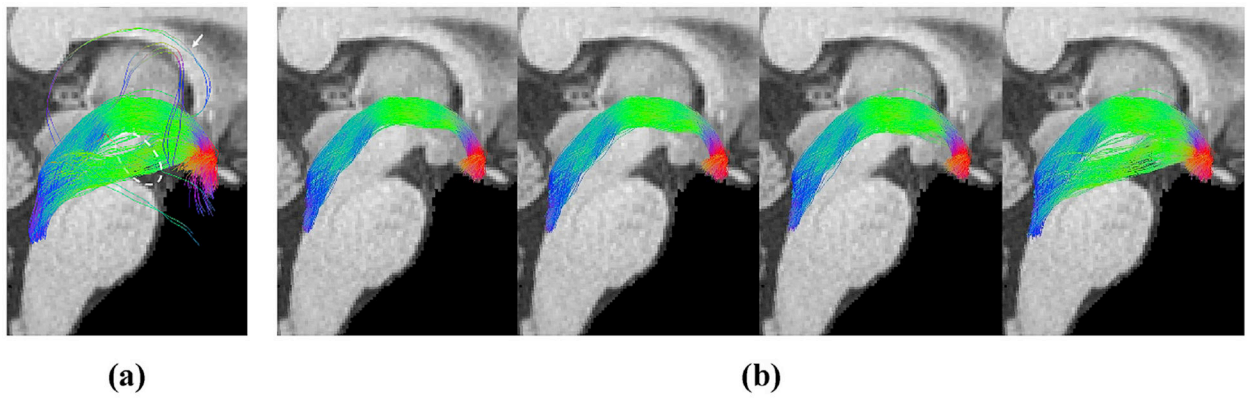




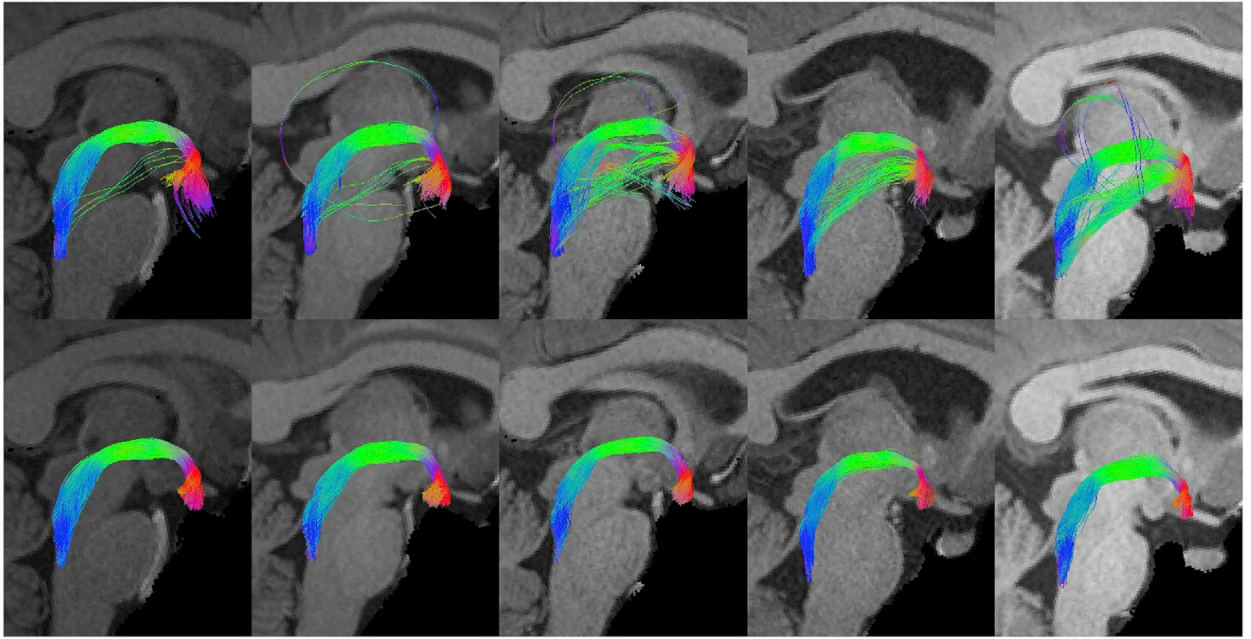
**Fig. 5.** Input and reconstructed fornix bundles from ADNI2 subjects. Each sub-figure from (a) to (d) shows three input (top row) and filtered (bottom row) bundles from the AD, LMCI, EMCI, and CN group, respectively.



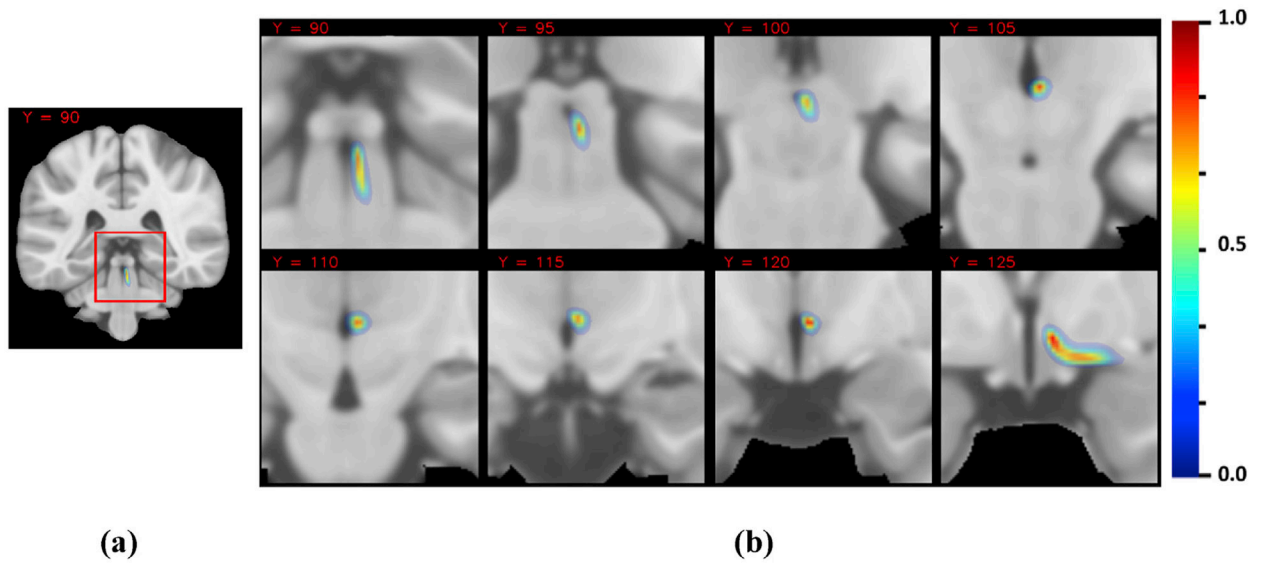
**Fig. 6.** The filtering results of the same EMCI subject in Fig. 4 with different choices of the affinity parameter  $K$ . The input fiber bundle, and the filtered bundle with  $K = 8, 16, 24, 32$  and  $39$ , are shown from left to right, respectively.



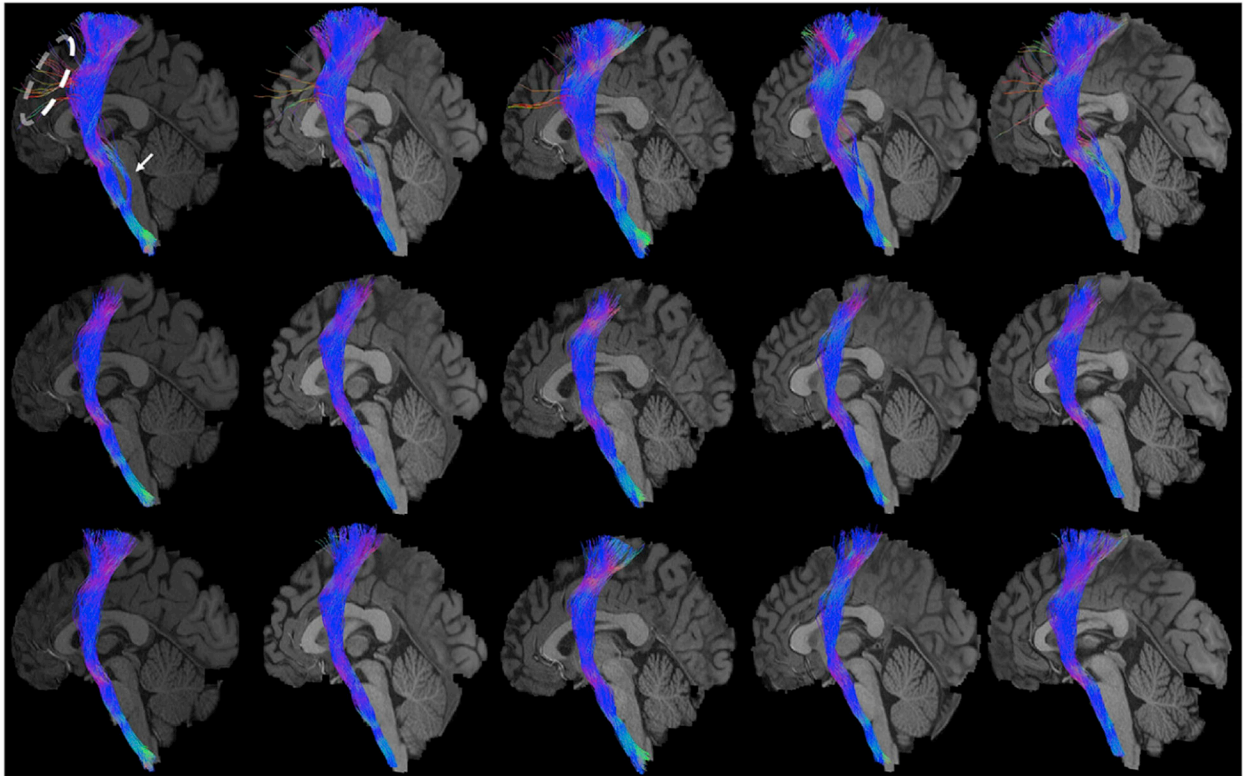
**Fig. 7.** The impact of the parameter  $L_{max}$  on track filtering performance. (a) Input LC pathway of an HCP subject. Spurious tracks were highlighted by the ellipsoid and arrow. (b) Filtered LC pathways with varying  $L_{max} = 0.05, 0.1, 0.3,$  and  $0.5,$  from left to right, respectively.



**Fig. 8.** Results of filtering the right LC bundle from five HCP subjects. Top row: input LC pathways; Bottom row: filtered bundles.

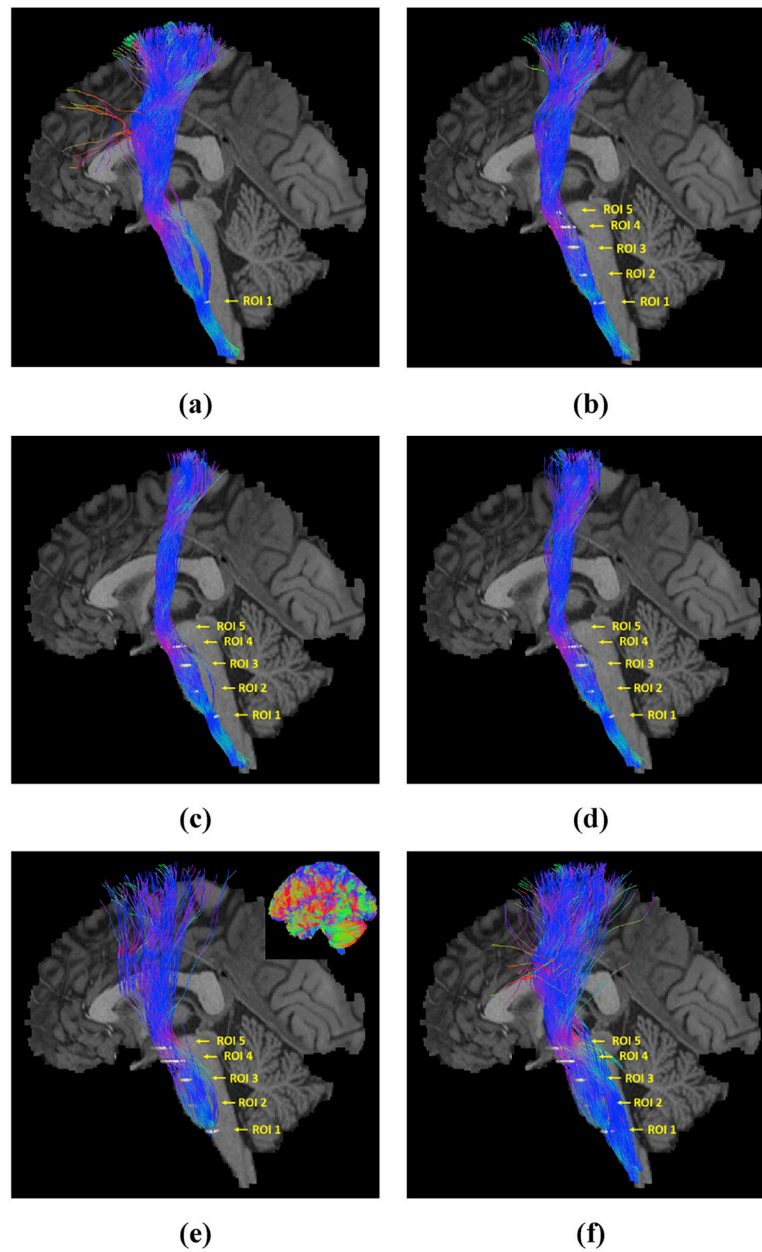


**Fig. 9.** A probabilistic atlas of the right LC bundle is plotted on 8 coronal slices in the MNI152 space. (a) indicates the zoom-in region (within the red box) on one coronal slice. The corresponding regions on different coronal slices are magnified and displayed in (b).



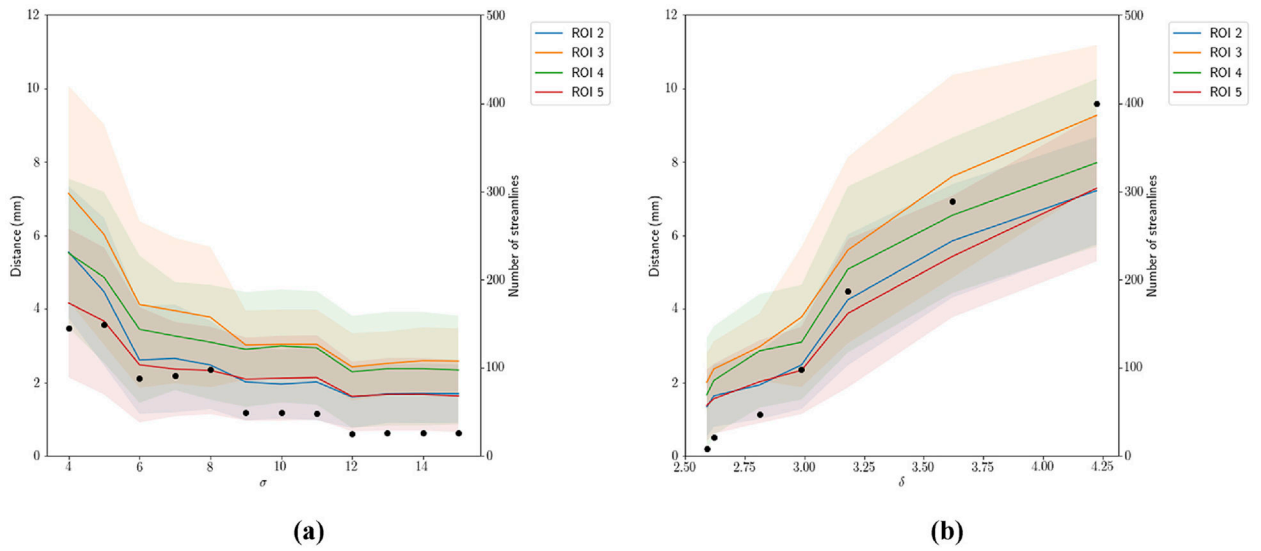
**Fig. 10.**

Groupwise track filtering results of the left CST from five representative HCP subjects. Top row: input fiber bundles from FOD-based tractography. The white arrow and ellipse highlight main outliers removed during the filtering process. Middle row: filtered fiber bundles generated by our groupwise filtering algorithm. Bottom row: filtering results at the whole-streamline level by adding back pruned points for tracks retained after the groupwise filtering process, i.e., tracks shown in the middle row.



**Fig. 11.**

A comparison of the filtering results from our method and other track filtering methods and two tract reconstruction methods for the left CST of an HCP subject. (a) The input bundle reconstructed with one inclusion ROI (the white disk). (b) The overlay of the filtered bundle from our method and the five manually delineated ROIs (white disks). (c) The overlay of the filtered bundle from QuickBundles and the five manually delineated ROIs (white disks). (d) CCI-based result. The left CST bundle reconstructed by the atlas-based method in RecoBundles and SlicerDMRI are displayed in panel (e) and (f), respectively. The whole brain tractography, used as the input for both atlas-based methods, is displayed in the top-right in (e).

**Fig. 12.**

The impact of proximity parameters on CST reconstruction with our groupwise method. The mean and standard deviation of Hausdorff distances (y-axis on the left side) from the reconstructed bundle to each ROI (2–5) are plotted in (a) and (b) as colored curves and shaded regions with respect to the change of the parameters  $\sigma$  and  $\delta$ , respectively. In addition, the number of tracks (y-axis on the right side) in the reconstructed bundle are plotted as black dots with respect to the change of parameters.



**Table 1**

Hausdorff Distance of Fiber Bundles to Manually Delineated ROIs. The unit of the distance is in *mm*.

ROIs	Approaches	Hausdorff Distance (mean $\pm$ std <i>mm</i> )
ROI 5	Original CST	9.49 $\pm$ 2.20
	Proposed method	<b>2.33 <math>\pm</math> 1.19</b>
	QuickBundles	4.76 $\pm$ 1.92
	CCI	3.10 $\pm$ 1.76
	RecoBundles	11.81 $\pm$ 3.63
	SlicerDMRI	16.25 $\pm$ 3.10
	SlicerDMRI + ROI 1	9.34 $\pm$ 2.18
ROI 4	Original CST	9.74 $\pm$ 2.50
	Proposed method	<b>3.10 <math>\pm</math> 1.56</b>
	QuickBundles	5.89 $\pm$ 2.21
	CCI	4.26 $\pm$ 2.10
	RecoBundles	13.19 $\pm$ 4.02
	SlicerDMRI	18.15 $\pm$ 2.79
	SlicerDMRI + ROI 1	9.32 $\pm$ 2.20
ROI 3	Original CST	9.86 $\pm$ 1.86
	Proposed method	<b>3.78 <math>\pm</math> 1.91</b>
	QuickBundles	6.68 $\pm$ 2.65
	CCI	5.41 $\pm$ 2.07
	RecoBundles	13.14 $\pm$ 3.22
	SlicerDMRI	18.75 $\pm$ 3.30
	SlicerDMRI + ROI 1	9.76 $\pm$ 1.90
ROI 2	Original CST	7.46 $\pm$ 1.49
	Proposed method	<b>2.48 <math>\pm</math> 1.20</b>
	QuickBundles	4.92 $\pm$ 1.99
	CCI	3.65 $\pm$ 1.76
	RecoBundles	12.27 $\pm$ 4.23
	SlicerDMRI	16.43 $\pm$ 3.00
	SlicerDMRI + ROI 1	7.23 $\pm$ 1.23

**Table 2**

Size (num of tracks) of filtered fiber bundles.

<b>Approaches</b>	<b>Num of Tracks (mean <math>\pm</math> std; [min, max])</b>
Proposed method	168.90 $\pm$ 45.37; [98, 253]
QuickBundles	197.70 $\pm$ 80.89; [84, 381]
CCI	229.45 $\pm$ 75.64; [80, 342]
RecoBundles	146.90 $\pm$ 51.08; [78, 269]
SlicerDMRI	381.10 $\pm$ 101.26; [220, 611]
SlicerDMRI + ROI 1	81.75 $\pm$ 43.17; [25, 210]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**

Summary of the computing cost in each experiment.

<b>Filtering Task</b>	<b># Subjects</b>	<b>Time (mean <math>\pm</math> std <i>seconds</i>)</b>
Fornix	40	1087.99 $\pm$ 20.74
LC Pathway	50	991.24 $\pm$ 17.92
CST	20	535.19 $\pm$ 9.07

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript