



Embarrassingly Parallel Acceleration of Global Tractography via Dynamic Domain Partitioning

Haiyong Wu^{1,2}, Geng Chen², Yan Jin², Dinggang Shen^{2,3*} and Pew-Thian Yap^{2*}

¹ School of Information Engineering, Xiaozhuang University, Nanjing, China, ² Department of Radiology and Biomedical Research Imaging Center (BRIC), University of North Carolina, Chapel Hill, NC, USA, ³ Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea

OPEN ACCESS

Edited by:

Tianzi Jiang,
The Chinese Academy of Sciences,
China

Reviewed by:

Tjeerd Olde Scheper,
Oxford Brookes University, UK
Lijuan Zhang,
Shenzhen Institutes of Advanced
Technology (SIAT) of the Chinese
Academy of Science, China
Dajiang Zhu,
University of Southern California, USA

*Correspondence:

Pew-Thian Yap
ptyap@med.unc.edu;
Dinggang Shen
dgshen@med.unc.edu

Received: 08 March 2016

Accepted: 23 June 2016

Published: 13 July 2016

Citation:

Wu H, Chen G, Jin Y, Shen D and Yap P-T (2016) Embarrassingly Parallel Acceleration of Global Tractography via Dynamic Domain Partitioning. *Front. Neuroinform.* 10:25. doi: 10.3389/fninf.2016.00025

Global tractography estimates brain connectivity by organizing signal-generating fiber segments in an optimal configuration that best describes the measured diffusion-weighted data, promising better stability than local greedy methods with respect to imaging noise. However, global tractography is computationally very demanding and requires computation times that are often prohibitive for clinical applications. We present here a reformulation of the global tractography algorithm for fast parallel implementation amenable to acceleration using multi-core CPUs and general-purpose GPUs. Our method is motivated by the key observation that each fiber segment is affected by a limited spatial neighborhood. In other words, a fiber segment is influenced only by the fiber segments that are (or can potentially be) connected to its two ends and also by the diffusion-weighted signal in its proximity. This observation makes it possible to parallelize the Markov chain Monte Carlo (MCMC) algorithm used in the global tractography algorithm so that concurrent updating of independent fiber segments can be carried out. Experiments show that the proposed algorithm can significantly speed up global tractography, while at the same time maintain or even improve tractography performance.

Keywords: diffusion magnetic resonance imaging, global tractography, Markov chain Monte Carlo, brain connectivity, parallel computing

1. INTRODUCTION

Diffusion magnetic resonance imaging (DMRI) (Basser et al., 1994) relies on the fundamental observation that the diffusion of water molecules in white matter (WM) is much less restricted along the direction of axonal bundles than perpendicular to them. DMRI is widely used as a non-invasive imaging modality for studying WM changes in association with development, growth, and disorders, such as Alzheimer's disease (Li et al., 2013; Termenon et al., 2013; Jin et al., 2015a,b), Parkinson's disease (Martínez-Murcia et al., 2014), multiple sclerosis (Goldberg-Zimring et al., 2005), autism (Thomas et al., 2011), and traumatic brain injury (Dennis et al., 2015a,b). DMRI is also used to map out the comprehensive wiring of the brain, generating what is commonly known as the brain connectome (Van Essen et al., 2013; Nossenson et al., 2016).

Diffusion tractography algorithms estimate the WM pathways by constructing streamlines that trace through the directional information at each voxel given by DMRI data (Conturo et al., 1999; Mori et al., 1999; Basser et al., 2000). They can be generally divided into two categories: local tractography (LT) and global tractography (GT). LT starts from a random or predetermined region and traces local voxel-wise fiber orientations in small successive steps. At each voxel, there can be a single direction, e.g., based on the principal direction of the tensor model (Mori et al., 1999; Basser et al., 2000), or multiple directions as given by the peaks of the orientation distribution function (ODF)

(Tuch, 2004). The tracing of these directions can be done deterministically (Mori et al., 1999; Basser et al., 2000) or probabilistically (Behrens et al., 2003; Parker et al., 2003). The main strength of LT lies in its speed. For example, whole-brain LT takes only a few minutes with the FACT algorithm (Mori et al., 1999). However, LT is susceptible to error accumulation owing to the estimation uncertainty of the directions at each voxel (Behrens et al., 2003). This causes the estimated tracts to deviate from the true WM trajectories.

On the other hand, GT (Reisert et al., 2009, 2011) uses global optimization techniques to decrease sensitivity to local estimation errors. GT reconstructs all fiber trajectories simultaneously by considering their agreement with the underlying diffusion data. The reconstructed fiber trajectories are the result of the interaction between signal-generating fiber segments and their matching with the diffusion-weighted measurements. The inter-segment interaction is reflected by the *internal energy* and the matching of the signals of the fibers with the data is reflected by the *external energy*. The configuration of these fiber segments is optimized globally by minimizing the global energy, computed as the sum of the internal energy and the external energy (Kreher et al., 2008; Fillard et al., 2009; Reisert et al., 2009, 2011). GT algorithms can better resolve ambiguous WM pathways and afford better robustness to imaging noise (Mangin et al., 2013). Utilizing a Monte Carlo optimization framework, GT randomly perturbs the fiber segments using a creation/deletion, connection/disconnection, and shifting mechanism to determine a configuration of the fiber segments that best fits the data. Although studies have shown that GT generally outperforms LT (Fillard et al., 2011), the main drawback of GT methods is their computational costs. Whole-brain GT takes from a few hours up to a day on a standard machine (Reisert et al., 2011; Neher et al., 2012). Therefore, reducing the computational cost of GT is critical to improving its usability in clinical settings.

The key idea of recent works on accelerating LT algorithms is to generate fiber trajectories from the seed voxels in parallel instead of sequentially (Lee and Kim, 2013). However, this approach cannot be applied to GT because of the interdependence of the fiber tracts in the global optimization framework. The GT algorithm has been recently modified in different forms with multithreading capability (Reisert et al., 2014; Christiaens et al., 2015). It is however unsure whether in these implementations the statistical independence structure of the problem has been taken into account to allow a mathematically valid parallelization of the associated MCMC algorithm.

In this work, we will focus on accelerating the GT algorithm proposed by Reisert et al. (2011). This algorithm utilizes a Markov chain Monte Carlo (MCMC) technique, called the Metropolis–Hastings algorithm (Neal, 1993; Van Lieshout, 2000), to determine the configuration of fiber segments with the minimum global energy. In theory, the Markov process asymptotically reaches a unique stationary distribution that equals the posterior distribution of the fiber configuration given the data. However, MCMC methods can be prohibitively slow and require a large number of “burn-in” steps before producing representative samples.

An *embarrassingly parallel* approach was recently proposed to parallelize burn-in and sampling in MCMC (Neiswanger et al., 2014). The key idea is to apply any classical MCMC method

independently to subsets of data without requiring much communication between them. First, the data are partitioned into multiple subsets. Next, an MCMC method is used to draw samples from a *posteriori* distribution given by each subset. Finally, the samples from all of the subsets are combined to form samples from the full posterior. This method is termed *embarrassingly parallel* because the processing of each subset is performed independently without communication with other subsets until the final combination stage.

Building on this concept, we show in this work that the GT algorithm can be improved significantly in terms of speed by MCMC parallelization. The key observation that drives our algorithm, called the parallel global tractography algorithm (PGT), is that the spatial extent of the influence of each fiber segment is limited. That is, the influence of each fiber segment depends only on the fiber segments that are connected (or can potentially be connected) to its both ends and also on the diffusion-weighted signal that is in its proximity. In other words, despite the fact that we try to decrease the total fitting energy in a global sense, the influence of each fiber segment on the variation of the energy is in fact *local*. Based on this observation, significant parallelism can be harnessed for improving the speed of the GT algorithm. The data can be partitioned into subsets similar to Neiswanger et al. (2014) and processed separately before combining the results to form samples for the original problem. Experimental results confirm the effectiveness of the proposed method and demonstrate that comparable tractography performance can be achieved in a reduced amount of time.

Part of this work has been reported in our recent workshop paper (Wu et al., 2015). Herein, we provide additional examples, results, derivations, and insights that are not part of the workshop publication.

2. ACCELERATING GLOBAL TRACTOGRAPHY

2.1. Background

The GT algorithm assumes that a fiber streamline is composed of fiber *segments*, each of which can be represented by a cylinder, as illustrated in **Figure 1**. The i -th segment can be defined as a tuple $h_i = (\mathbf{x}_i, \mathbf{v}_i, l, d)$, where $\mathbf{x}_i \in \mathbb{R}^3$ denotes the spatial location of the center of the segment, $\mathbf{v}_i \in S_2$ denotes the orientation, l is the half length, and d is the diameter. Both the length and the diameter are identical for all segments. The extremities of the fiber segment are \mathbf{e}_i^α , where $\alpha \in \{+, -\}$ indicate the positive or negative endpoint.

The goal of GT (Reisert et al., 2011) is to determine the optimal configuration \mathcal{M} of a set of signal-generating fiber segments given the measured diffusion-weighted signals \mathcal{D} , including their existence, spatial positions, orientations, and connections at both ends to other fiber segments. More formally, one is interested

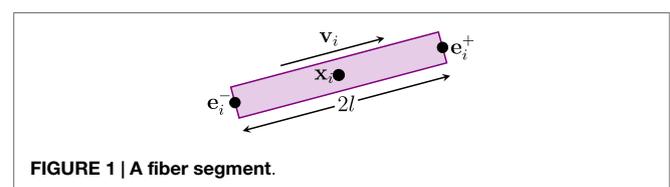


FIGURE 1 | A fiber segment.

in determining the \mathcal{M} that maximizes the posterior probability $P(\mathcal{M}|\mathcal{D})$ defined as

$$P(\mathcal{M}|\mathcal{D}) \propto P(\mathcal{M})P(\mathcal{D}|\mathcal{M}) = \exp\left(\frac{-E_{\text{int}}(\mathcal{M}) - E_{\text{ext}}(\mathcal{M}, \mathcal{D})}{T}\right), \quad (1)$$

where $E_{\text{int}}(\mathcal{M})$ and $E_{\text{ext}}(\mathcal{M}, \mathcal{D})$ are the internal energy and the external energy, respectively, and T is the temperature associated with simulated annealing (Aarts and Korst, 1989).

The internal energy $E_{\text{int}}(\mathcal{M})$ characterizes the smoothness of the fibers and is defined as the sum of all the interaction potentials between connected segments:

$$E_{\text{int}}(\mathcal{M}) = \lambda_{\text{int}} \sum_{(\mathbf{e}_i^{\alpha_i \rightarrow j}, \mathbf{e}_j^{\alpha_j \rightarrow i})} \frac{1}{l^2} (|\mathbf{e}_i^{\alpha_i \rightarrow j} - \bar{\mathbf{x}}_{ij}|^2 + |\mathbf{e}_j^{\alpha_j \rightarrow i} - \bar{\mathbf{x}}_{ij}|^2) - L, \quad (2)$$

where $\bar{\mathbf{x}}_{ij}$ represents the midpoint of the line that connects the centers of these two segments, and $\mathbf{e}_i^{\alpha_i \rightarrow j}$ denotes that the endpoint of the i -th segment that is connected to the j -th fiber segment. The bias L affects the probability of connections between segments.

The signal predicted by the fiber segments at location \mathbf{x} and orientation \mathbf{v} is defined as

$$F_{\mathcal{M}}(\mathbf{x}, \mathbf{v}) = w \sum_{i=1}^N \exp(-c(\mathbf{v}^T \mathbf{v}_i)^2) \exp\left(\frac{-|\mathbf{x} - \mathbf{x}_i|^2}{\sigma^2}\right). \quad (3)$$

where N is the total number of fiber segments. The constant w controls the amount of signal contribution from each fiber segment. Parameter $\sigma > 0$ controls the spatial extent of the influence of each fiber segment. Parameter $c > 0$ controls the shape of the signal profile generated by each fiber segment (Reisert et al., 2011). In practical implementation, the second exponential term is truncated and hence the spatial extent is limited.

The external energy $E_{\text{ext}}(\mathcal{M}, \mathcal{D})$ measures the difference between the observed data \mathcal{D} and the predicted signal $F_{\mathcal{M}}(\mathbf{x}, \mathbf{v})$ given by the configuration \mathcal{M} :

$$\begin{aligned} E_{\text{ext}}(\mathcal{M}, \mathcal{D}) &= \lambda_{\text{ext}} \|F_{\mathcal{M}} - \mathcal{D}\|^2 \\ &= \lambda_{\text{ext}} \int_{\mathbb{R}^3 \times S_2} |F_{\mathcal{M}}(\mathbf{x}, \mathbf{v}) - \mathcal{D}(\mathbf{x}, \mathbf{v})|^2 d^3 \mathbf{x} d^2 \mathbf{v}, \end{aligned} \quad (4)$$

where λ_{ext} is a tuning parameter.

2.2. Parallel Global Tractography

To maximize the posterior probability $P(\mathcal{M}|\mathcal{D})$, an MCMC method called the Metropolis–Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970) is employed in Reisert et al. (2011). However, MCMC methods may take a prohibitively long time, depending on the number of data points. Furthermore, MCMC methods might require a large number of “burn-in” steps before beginning to generate representative samples (Liu, 2001). Here, we propose a parallelized version of the GT algorithm, called parallel global tractography (PGT), to improve the speed.

PGT utilizes the structure of the problem to allow embarrassingly parallel speed up of the GT algorithm. Similar to Neiswanger et al. (2014), this is achieved by partitioning the data into subsets,

on which an MCMC algorithm can act independently without communication between them until the final combination stage. More formally, this is done by partitioning the data \mathcal{D} into K subsets $\{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_K\}$ and associate with each subset a *subposterior* probability $P(\mathcal{M})^{\frac{1}{K}} P(\mathcal{D}_k|\mathcal{M})$. MCMC sampling is then performed independently for each *subposterior* probability before eventually combining their samples to produce samples from an estimate of the subposterior density product, which is proportional to the full-data posterior. If the density product estimator is consistent, it can then be shown that one is drawing asymptotically exact samples from the full posterior distribution. Parametric, non-parametric, and semiparametric estimation techniques are described in Neiswanger et al. (2014).

While straightforward, the method described above (Neiswanger et al., 2014) is not directly applicable to the parallelization of GT. The problem lies in the difficulty in designing an appropriate consistent estimator of the subposterior density estimator that will allow us to draw asymptotically exact samples from the posterior. Combination of the samples drawn from the subposteriors is further complicated by the fact that the dimensionality of the parameters of the configuration \mathcal{M} is not fixed due to the creation/deletion proposals (Kreher et al., 2008). In fact, for this reason, GT requires a reversible jump version of MCMC (Green, 1995). In PGT, we overcome this problem by leveraging the structure of the GT problem to further improve parallelism. We first rewrite the posterior probability as

$$P(\mathcal{M}|\mathcal{D}) = P(\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_K|\mathcal{D}) \quad (5)$$

$$= P(\mathcal{M}_1, \dots, \mathcal{M}_K|\mathcal{D}, \mathcal{M}_0)P(\mathcal{M}_0|\mathcal{D}). \quad (6)$$

Note that here \mathcal{M}_0 denotes the configuration of the fiber segments between the other K regions, in which the fiber configurations are denoted as $\{\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_K\}$ (see **Figure 2**). If the region covered by \mathcal{M}_0 gives sufficient separation between the K regions, and noting that each fiber segment is affected by its neighboring and not distant fiber segments (as required by the internal energy), we have

$$P(\mathcal{M}|\mathcal{D}) = P(\mathcal{M}_0|\mathcal{D}) \prod_{k=1}^K P(\mathcal{M}_k|\mathcal{D}, \mathcal{M}_0). \quad (7)$$

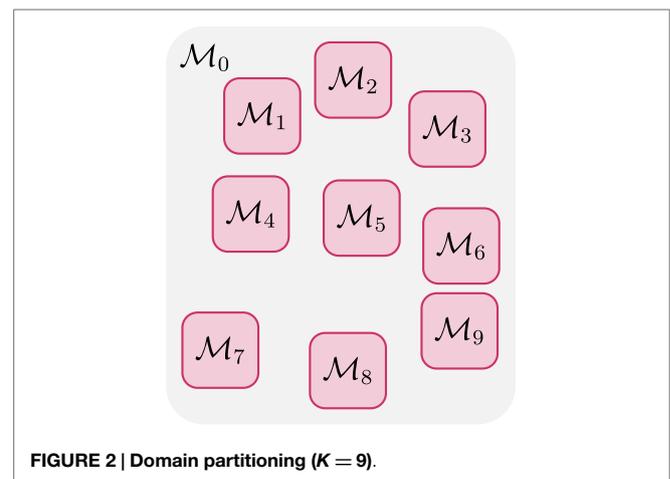
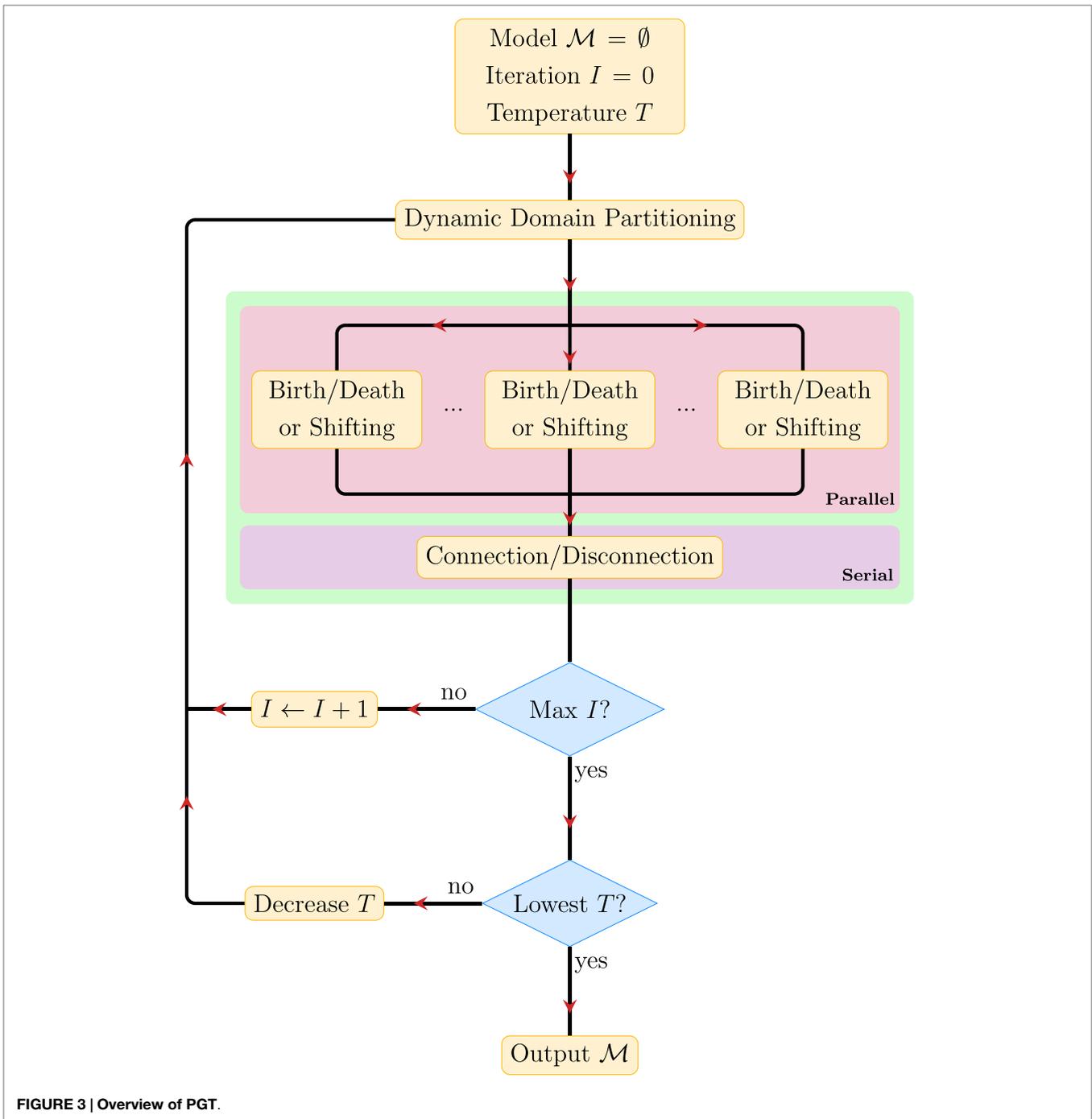


FIGURE 2 | Domain partitioning ($K = 9$).



This indicates that the configurations $\{\mathcal{M}_1, \dots, \mathcal{M}_K\}$ are conditionally independent given \mathcal{D} and \mathcal{M}_0 . Noting the fact that the configuration of each fiber segment is only dependent on the data in its vicinity, we further have

$$P(\mathcal{M}|\mathcal{D}) = P(\mathcal{M}_0|\mathcal{D}_0) \prod_{k=1}^K P(\mathcal{M}_k|\mathcal{D}_k, \mathcal{M}_0). \quad (8)$$

This implies that once the samples for \mathcal{M}_0 have been drawn, the samples for $\{\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_K\}$ can be drawn independently and

hence in parallel. In contrast to Neiswanger et al. (2014), instead of partitioning the data, we partition the parameters according to the spatial location and then partition the data accordingly. This formulation allows us to make local changes (creation/deletion, connection/disconnection, and shifting) in parallel. We can switch to $K=0$ to accommodate for more global proposals like the connection/disconnection proposal (Reisert et al., 2009).

The MH algorithm is applied in parallel by proposing changes to the fiber segments in the subregions associated with $\{\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_K\}$. The proposals in these subregions are

accepted or rejected based on their individual acceptance ratios. The independence condition guarantees that the proposals for each of the K subregions can be accepted and rejected separately but in parallel.

Then, for $k = 1, \dots, K$, samples are drawn from the subposterior densities $P(\mathcal{M}_k | \mathcal{D}_k, \mathcal{M}_0)$. With \mathcal{D}_k and \mathcal{M}_0 fixed and hence $P(\mathcal{D}_k, \mathcal{M}_0)$ being a constant, we have

$$P(\mathcal{M}_k | \mathcal{D}_k, \mathcal{M}_0) \propto P(\mathcal{M}_k, \mathcal{M}_0) P(\mathcal{D}_k | \mathcal{M}_k, \mathcal{M}_0). \quad (9)$$

Proposals for modification of configuration are made for the fiber segments in each region according to its subposterior density by randomly selecting at each time a fiber segment, perturbing it using a creation/deletion, connection/disconnection, and shifting mechanism, and examining if the regional energy can be decreased. In this process, \mathcal{M}_0 remains fixed and $\{\mathcal{M}_k\}$ are updated. After $\{\mathcal{M}_k\}$ are sufficiently updated, they are combined to form \mathcal{M} . The random partitioning of the image space into subregions is performed iteratively so that each time the fiber configurations of a different set of K random subregions can be updated. The decision of whether to accept a proposal is based on the individual Green's ratio of the i -th region

$$R_k = \frac{P(\mathcal{M}'_k | \mathcal{D}_k, \mathcal{M}_0) Q(\mathcal{M}_k | \mathcal{M}'_k)}{P(\mathcal{M}_k | \mathcal{D}_k, \mathcal{M}_0) Q(\mathcal{M}'_k | \mathcal{M}_k)}, \quad (10)$$

where $Q(\mathcal{M}'_k | \mathcal{M}_k)$ is the transition probability associated with the MH algorithm. From Equation (9), the internal energy contributed by the fiber segments in the k -th region alone is

$$E_{\text{int}}(\mathcal{M}_k) = \lambda_{\text{int}} \sum_{(\mathbf{e}_i^{\alpha_i \rightarrow j}, \mathbf{e}_j^{\alpha_j \rightarrow i}) \in \mathcal{N}_k \times \mathcal{N}_k} \frac{1}{j^2} (||\mathbf{e}_i^{\alpha_i \rightarrow j} - \bar{\mathbf{x}}_{ij}||^2 + ||\mathbf{e}_j^{\alpha_j \rightarrow i} - \bar{\mathbf{x}}_{ij}||^2) - L \quad (11)$$

and the external energy is

$$E_{\text{ext}}(\mathcal{M}_k, \mathcal{D}_k) = \lambda_{\text{ext}} \int_{\mathcal{N}_k \times \mathcal{S}_2} |F_{\mathcal{M}_k}(\mathbf{x}, \mathbf{n}) - \mathcal{D}_k(\mathbf{x}, \mathbf{n})|^2 d^3 \mathbf{x} d^2 \mathbf{n}, \quad (12)$$

where \mathcal{N}_k is the region containing all fiber segments associated with \mathcal{M}_k .

Note that some proposals are parallelizable and some are not. For each fiber segment, the change in internal energy associated with creation/deletion and shifting proposals is affected only by the fiber segments it is (or will be) connected to. The change in external energy involves only the diffusion signals in a localized neighborhood surrounding the fiber segment. Hence, creation/deletion and shifting proposals can be performed independently and simultaneously in different subregions. However, the connection/disconnection proposals, which attempt to determine new fibers with lower energy, involve a larger spatial extent and are hence more difficult to parallelize. To overcome this problem, we alternate between parallel proposals (i.e., creation/deletion and shifting) and serial proposals (i.e., connection/disconnection) according to MH transition probabilities assigned to them.

2.3. Implementation

Figure 3 summarizes the key steps in PGT. The PGT algorithm involves repeating the following steps until convergence.

1. **Dynamic Domain Partitioning:** Partition the image space into K subregions, between which the configurations of the fiber segments are independent.
2. **Parallel Proposals:** Make creation/deletion and shifting proposals in parallel for the fiber segments in these regions according to the corresponding transition probabilities, accept/reject the proposals based on their acceptance ratios, and repeat this step for a sufficient number of times.
3. **Serial Proposals:** Make connection/disconnection proposals and determine fiber tracts that better explain the data.

We dynamically partition the image data into K subregions that are mutually non-influential and statistically independent. We randomly choose K points, and the K subregions are defined as cubic blocks with these K points as centers. These subregions are chosen to be far enough to avoid overlap so as to maintain statistical independence. The region excluding these K subregions is the region containing \mathcal{M}_0 . After sufficient proposals have been proposed in parallel and in serial, the partitioning is repeated to generate a new set of subregions. **Figure 4** shows an example of random subregions with a block size that is typical in our implementation.

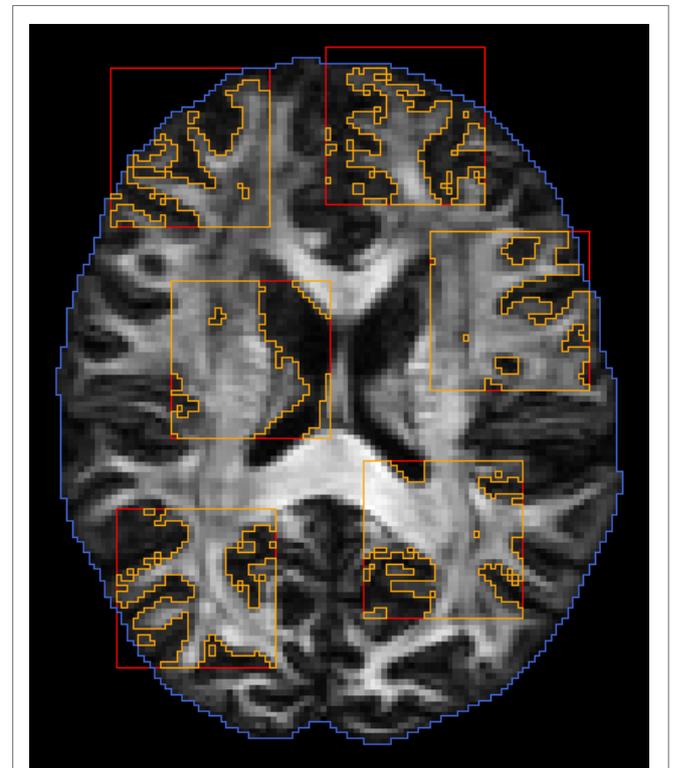


FIGURE 4 | Dynamic domain partitioning. Blue: brain region; red: random block regions; orange: regions after applying white matter mask.

3. EXPERIMENT RESULTS

3.1. Datasets

Two synthetic datasets and a real human dataset were utilized to evaluate the performance of PGT.

- *Spiral data*: A 128×128 field of diffusion-weighted signals was generated to simulate curved fiber bundles in white matter of human brain, forming a spiral as shown in **Figure 8**. Each voxel within the spiral was simulated by a tensor model with principal diffusivities $\lambda_1 = 1.7 \times 10^{-3} \text{ mm}^2/\text{s}$, $\lambda_2 = \lambda_3 = 3.0 \times 10^{-4} \text{ mm}^2/\text{s}$, and diffusion weighting $b = 2000 \text{ s/mm}^2$. The gradient directions and the voxel resolution were identical to those of the real data described below.
- *Cross data*: Another synthetic dataset was generated to simulate two fiber bundles crossing at 90° . The image dimension is 60×60 , and the signal at each voxel was simulated using a tensor model or its mixture with principal diffusivities $\lambda_1 = 1.5 \times 10^{-3} \text{ mm}^2/\text{s}$, $\lambda_2 = \lambda_3 = 1.0 \times 10^{-3} \text{ mm}^2/\text{s}$, and diffusion weighting $b = 2000 \text{ s/mm}^2$. The gradient directions and the voxel resolution were the same as those of the real data described below.
- *In vivo data*: Diffusion-weighted images were acquired from an adult with a Siemens 3T Tim Trio MR scanner using an EPI sequence. Diffusion gradients were applied in 120 non-collinear directions with diffusion weighting $b = 2000 \text{ s/mm}^2$. The acquisition parameters were: repetition

time (TR) = 12,400 ms, echo time (TE) = 116 ms, volume cropped dimensions = $83 \times 97 \times 76$, and voxel resolution = $2 \text{ mm} \times 2 \text{ mm} \times 2 \text{ mm}$.

3.2. Results

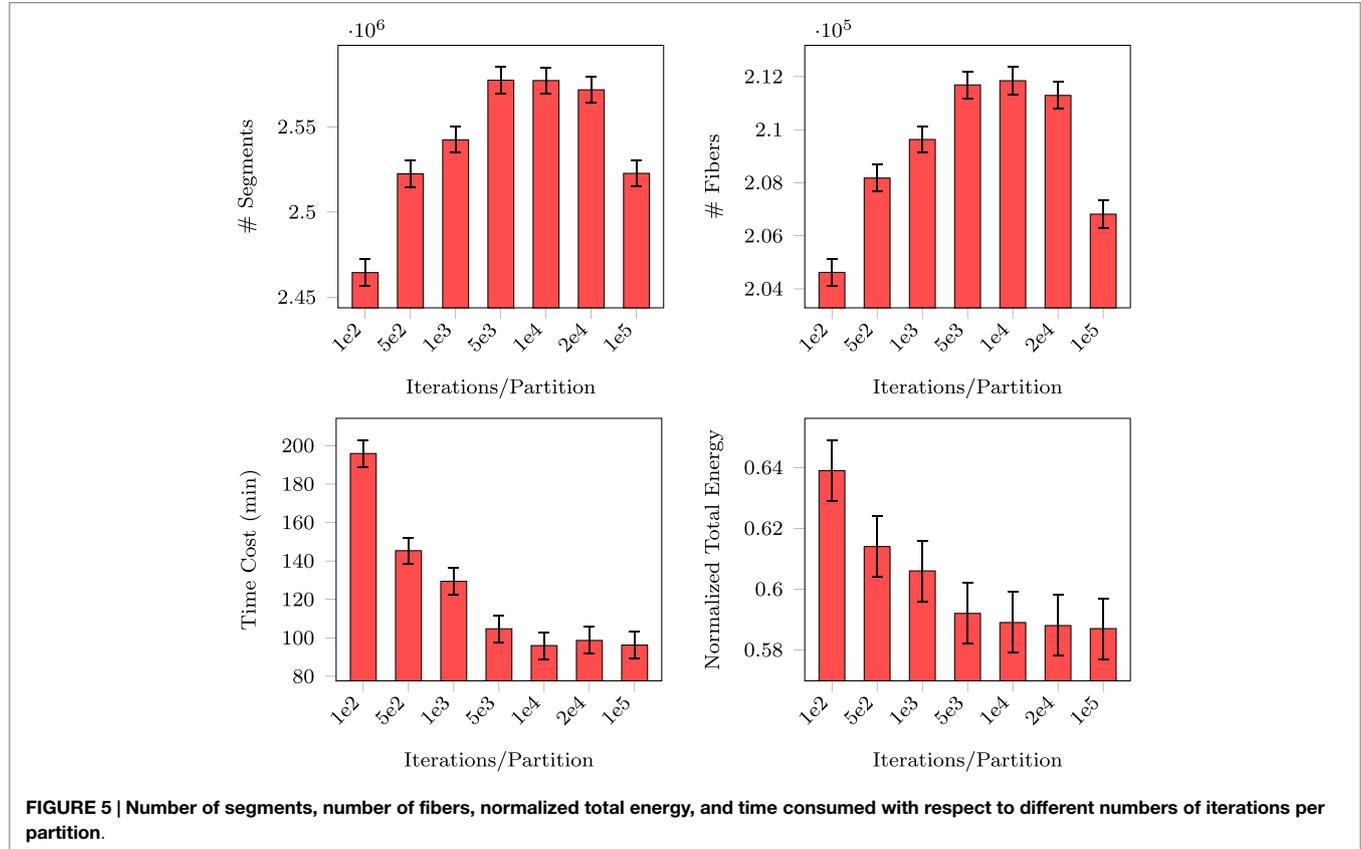
3.2.1. Iterations Per Partition and Number of Random Partitions

For fair comparison of PGT with GT, we fixed the total number of proposals that were eventually generated (i.e., 10^8). For PGT, this means balancing between the number of iterations per partition and the number of random partitions. Since the total number of proposals is proportional to the product of the two quantities, we only need to report the performance of PGT with respect to the former. The number of generated segments, the number of connected fibers, the normalized total energy, and the total computing time are reported in **Figure 5**. PGT was repeated 10 times using the *in vivo* data with 10 parallel threads.

The figure shows that the numbers of generated segments and connected fibers stabilize at 10^4 iterations per partition. The figure also indicates that with this amount of per partition iterations the total energy and the time consumed became stable, implying that this is sufficient for MCMC “burn-in.” In the following experiments, we iterated MCMC 10^4 times for each set of partitioned regions before randomly partitioning the data again.

3.2.2. Energy

As mentioned in Section 2.1, the internal energy characterizes the smoothness of the fibers while the external energy characterizes



the consistency between the predicted signal and the diffusion-weighted signal. We compared both the internal energy and external energy of GT and PGT (with 10 parallel threads). The parameters used for GT and PGT are set as recommended in Reisert et al. (2011). **Figure 6** shows the results with respect to the number of generated proposals for the spiral, the crossing and the *in vivo* data. It can be observed that PGT outperforms GT with lower internal energy and external energy. It is interesting to note that when the energy curves of GT flatten out, those of PGT continue to decrease. This indicates that, when the adjustments of the fiber segments are done concurrently in multiple regions, a configuration with lower energy can be reached with greater ease.

3.2.3. Computational Costs

The speed improvement given by PGT over GT was evaluated based on a computing cluster with 2.9 GHz Intel Xeon CPUs and 48 GB RAM. **Figure 7** shows that, for both synthetic and *in vivo* data, PGT requires only less than approximately 1/3 of the time required by GT. Note that it is not possible to achieve the ideal $10\times$ speed increase because the GT algorithm is only partially parallelized, as discussed in Section 2. Moreover, part of the time

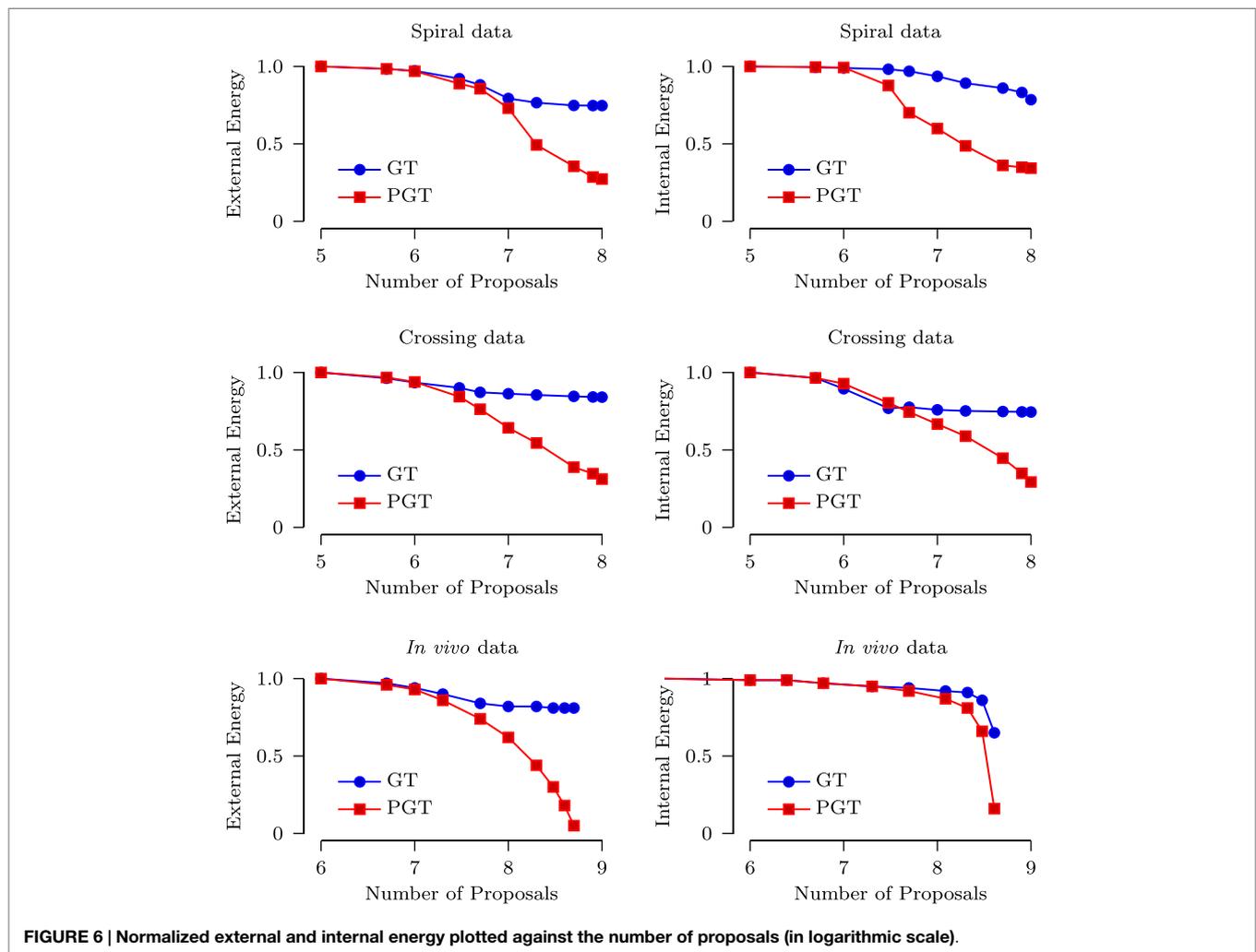
cost is associated with the computational overhead involved in the parallelized implementation.

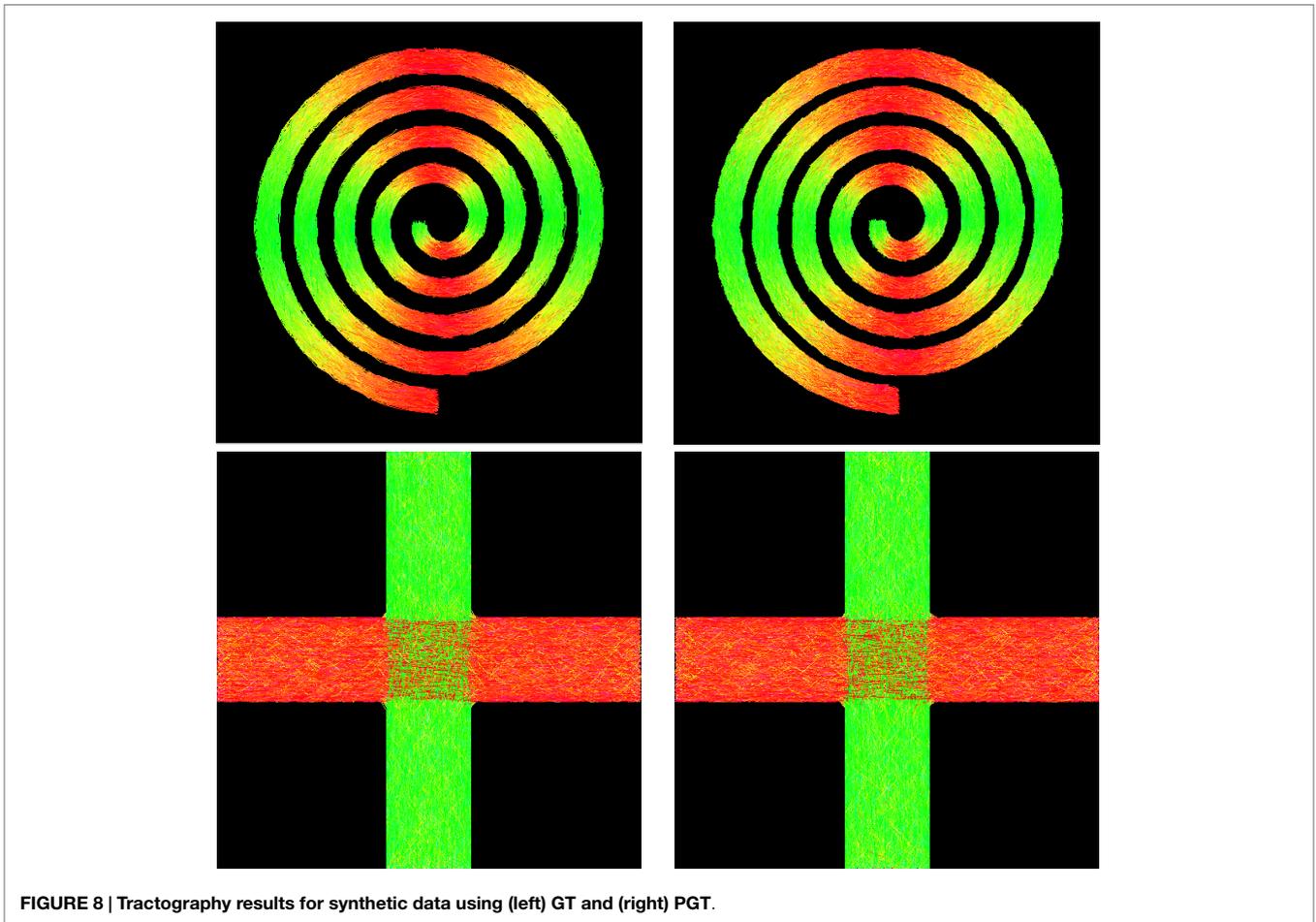
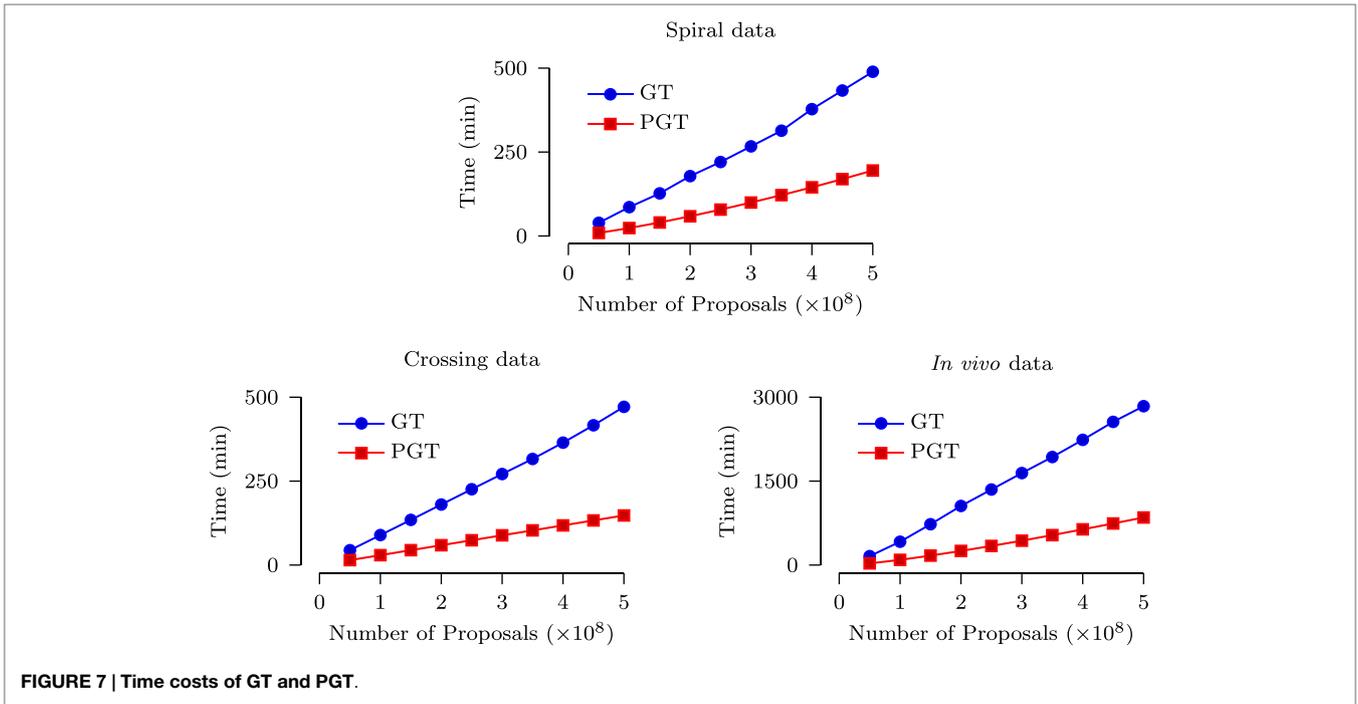
3.2.4. Tractography Results

Figure 8 shows the tractography results of GT and PGT for the synthetic data. Both GT and PGT create reasonable and consistent fiber tracts that are in agreement with the data. For numerical evaluation, we computed the distance of fiber bundles given by GT and PGT using the distance defined in Yap et al. (2011a). The distance between two fiber bundles \mathcal{F} and \mathcal{G} is defined as

$$\frac{1}{|\mathcal{F}| + |\mathcal{G}|} \left[\sum_{F_i \in \mathcal{F}} \min_{G_j \in \mathcal{G}} d(F_i, G_j) + \sum_{G_i \in \mathcal{G}} \min_{F_j \in \mathcal{F}} d(G_i, F_j) \right] \quad (13)$$

where $d(F_i, G_j)$ is a pairwise distance between fibers $F_i \in \mathcal{F}$ and $G_j \in \mathcal{G}$, which in our case, is defined as the mean of the closest distances calculated for all points on fiber F_i to fiber G_j . When two fiber tracts are identical, the value returned is zero. The respective distances for the spiral and the cross are 1.46 mm and 0.52 mm, indicating high similarity between the tractography results of GT and PGT. Despite the visual





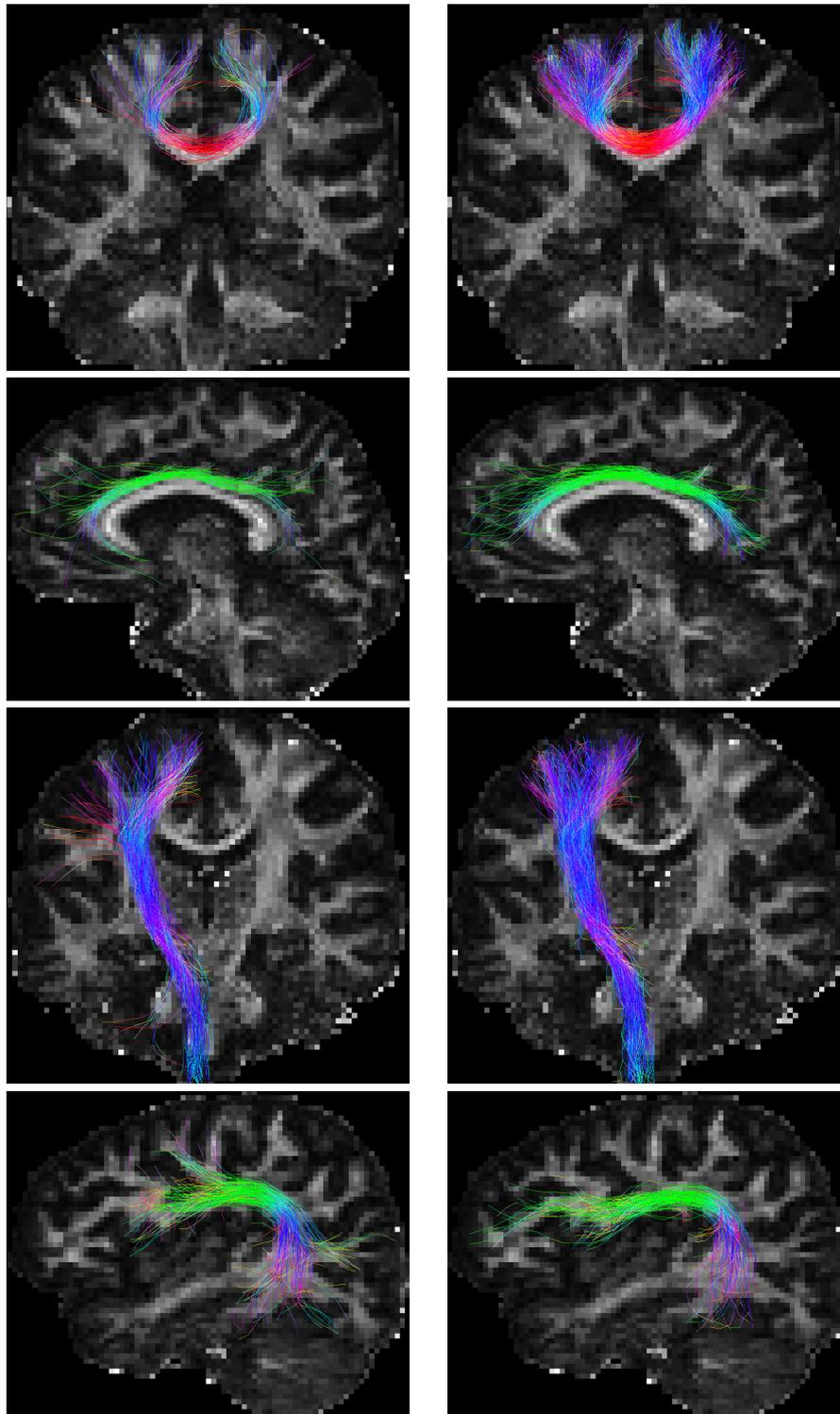


FIGURE 9 | Fiber bundles given by (left) GT and (right) PGT. From top to bottom are the CC-PRC, CGC, CST, and ARC tracts, respectively.

similarity between the GT and PGT, PGT yields in general a higher number of fiber tracts with greater lengths. Therefore, some dissimilarities exist between the results given by the two algorithms.

For the *in vivo* data, fiber bundles extracted with multiple ROIs (Wakana et al., 2007) are shown in **Figure 9**. The extracted fiber tracts include (Jin et al., 2013, 2014): (1) association tracts such as the cingulum tract (CGC); (2) the arcuate fasciculus (ARC), a part

GC, P-TY, and YJ. Contributed reagents/materials/analysis tools: HW, P-TY, GC. Wrote the paper: HW, YJ, P-TY, and DS.

ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their comments, which have helped improve the manuscript.

REFERENCES

- Aarts, E., and Korst, J. (1989). *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*. New York, NY: John Wiley & Sons, Inc.
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* 54, 2033–44. doi:10.1016/j.neuroimage.2010.09.025
- Basser, P. J., Mattiello, J., and LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66, 259–267. doi:10.1016/S0006-3495(94)80775-1
- Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., and Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med.* 44, 625–632. doi:10.1002/1522-2594(200010)44:4<625::AID-MRM17>3.0.CO;2-O
- Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., et al. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.* 50, 1077–1088. doi:10.1002/mrm.10609
- Christians, D., Reisert, M., Dhollander, T., Sunaert, S., Suetens, P., and Maes, F. (2015). Global tractography of multi-shell diffusion-weighted imaging data using a multi-tissue model. *Neuroimage* 123, 89–101. doi:10.1016/j.neuroimage.2015.08.008
- Conturo, T. E., Lori, N. F., Cull, T. S., Akbudak, E., Snyder, A. Z., Shimony, J. S., et al. (1999). Tracking neuronal fiber pathways in the living human brain. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10422–10427. doi:10.1073/pnas.96.18.10422
- Dennis, E. L., Ellis, M. U., Marion, S. D., Jin, Y., Moran, L., Olsen, A., et al. (2015a). Callosal function in pediatric traumatic brain injury linked to disrupted white matter integrity. *J. Neurosci.* 35, 10202–10211. doi:10.1523/JNEUROSCI.1595-15.2015
- Dennis, E. L., Jin, Y., Villalon-Reina, J. E., Zhan, L., Kernan, C. L., Babikian, T., et al. (2015b). White matter disruption in moderate/severe pediatric traumatic brain injury: advanced tract-based analyses. *Neuroimage Clin.* 7, 493–505. doi:10.1016/j.nicl.2015.02.002
- Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., et al. (2011). Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. *Neuroimage* 56, 220–234. doi:10.1016/j.neuroimage.2011.01.032
- Fillard, P., Poupon, C., and Mangin, J.-F. (2009). “A novel global tractography algorithm based on an adaptive spin glass model,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, eds G.-Z. Yang, D. J. Hawkes, D. Rueckert, A. Noble, and C. Taylor (London: Springer), 927–934.
- Goldberg-Zimring, D., Mewes, A. U., Maddah, M., and Warfield, S. K. (2005). Diffusion tensor magnetic resonance imaging in multiple sclerosis. *J. Neuroimaging* 15, 68S–81S. doi:10.1177/1051228405283363
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82, 711–732. doi:10.1093/biomet/82.4.711
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97–109. doi:10.1093/biomet/57.1.97
- Jin, Y., Shi, Y., Zhan, L., de Zubicaray, G. I., McMahon, K. L., Martin, N. G., et al. (2013). “Labeling white matter tracts in HARDI by fusing multiple tract atlases with applications to genetics,” in *International Symposium on Biomedical Imaging (ISBI)* (San Francisco: IEEE), 512–515.
- Jin, Y., Shi, Y., Zhan, L., Gutman, B. A., de Zubicaray, G. I., McMahon, K. L., et al. (2014). Automatic clustering of white matter fibers in brain diffusion MRI with an application to genetics. *Neuroimage* 100, 75–90. doi:10.1016/j.neuroimage.2014.04.048
- Jin, Y., Shi, Y., Zhan, L., and Thompson, P. M. (2015a). “Automated multi-atlas labeling of the fornix and its integrity in Alzheimer’s disease,” in *International Symposium on Biomedical Imaging (ISBI)* (New York: IEEE), 140–143.
- Jin, Y., Wee, C. Y., Shi, F., Thung, K. H., Ni, D., Yap, P. T., et al. (2015b). Identification of infants at high-risk for autism spectrum disorder using multiparameter multiscale white matter connectivity networks. *Hum. Brain Mapp.* (2015) 36, 4880–4896. doi:10.1002/hbm.22957
- Kreher, B., Mader, I., and Kiselev, V. (2008). Gibbs tracking: a novel approach for the reconstruction of neuronal pathways. *Magn. Reson. Med.* 60, 953–963. doi:10.1002/mrm.21749
- Lee, J., and Kim, D.-S. (2013). Divide et impera: acceleration of DTI tractography using multi-GPU parallel processing. *Int. J. Imaging Syst. Technol.* 23, 256–264. doi:10.1002/ima.22059
- Li, J., Jin, Y., Shi, Y., Dinov, I. D., Wang, D. J., Toga, A. W., et al. (2013). “Voxelwise spectral diffusional connectivity and its applications to Alzheimer’s disease and intelligence prediction,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, eds K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab (Nagoya: Springer), 655–662.
- Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer Series in Statistics, Springer.
- Mangin, J.-F., Fillard, P., Cointepas, Y., Le Bihan, D., Frouin, V., and Poupon, C. (2013). Toward global tractography. *Neuroimage* 80, 290–296. doi:10.1016/j.neuroimage.2013.04.009
- Martínez-Murcia, F. J., Górriz, J. M., Ramírez, J., Illán, I., Ortiz, A., Initiative, P. P. M., et al. (2014). Automatic detection of parkinsonism using significance measures and component analysis in datscan imaging. *Neurocomputing* 126, 58–70. doi:10.1016/j.neucom.2013.01.054
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953). Equations of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087–1092. doi:10.1063/1.1699114
- Mori, S., Crain, B. J., Chacko, V., and Van Zijl, P. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann. Neurol.* 45, 265–269. doi:10.1002/1531-8249(199902)45:2<265::AID-ANA21>3.0.CO;2-3
- Neal, R. M. (1993). *Probabilistic Inference Using Markov Chain Monte Carlo Methods*. Technical Report. University of Toronto.
- Neher, P. F., Stieltjes, B., Reisert, M., Reicht, I., Meinzer, H.-P., and Fritzsche, K. H. (2012). “MITK global tractography,” in *SPIE Medical Imaging*, eds D. R. Haynor and S. Ourselin (Orlando, FL: International Society for Optics and Photonics), 83144–83149.
- Neiswanger, W., Wang, C., and Xing, E. (2014). “Asymptotically exact, embarrassingly parallel MCMC,” in *The Annual Conference on Uncertainty in Artificial Intelligence (UAI)* (Quebec City: AUAI Press), 623–632.
- Nossenson, N., Magal, A., and Messer, H. (2016). Detection of stimuli from multi-neuron activity: empirical study and theoretical implications. *Neurocomputing* 174, 822–837. doi:10.1016/j.neucom.2015.10.007
- Parker, G. J., Haroon, H. A., and Wheeler-Kingshott, C. A. (2003). A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements. *J. Magn. Reson. Imaging* 18, 242–254. doi:10.1002/jmri.10350
- Reisert, M., Kiselev, V. G., Dihtal, B., Kellner, E., and Novikov, D. (2014). “MesoFT: unifying diffusion modelling and fiber tracking,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, eds P. Golland, N. Hata, C. Barillot, J. Hornegger, and R. Howe (Boston, MA: Springer), 201–208.
- Reisert, M., Mader, I., Anastasopoulos, C., Weigel, M., Schnell, S., and Kiselev, V. (2011). Global fiber reconstruction becomes practical. *Neuroimage* 54, 955–962. doi:10.1016/j.neuroimage.2010.09.016

FUNDING

This work was supported in part by Educational Science in Jiangsu Provincial Grants (D201501125), Guangdong Provincial Key Laboratory of Medical Image Processing (2014B030301042), a UNC BRIC-Radiology start-up fund, and NIH grants (NS093842, EB006733, EB008374, EB009634, MH088520, AG041721, and MH100217).

- Reisert, M., Mader, I., and Kiselev, V. (2009). "Global reconstruction of neuronal fibres," in *MICCAI Workshop on Diffusion Modelling*. London.
- Termonon, M., Grana, M., Besga, A., Echeveste, J., and Gonzalez-Pinto, A. (2013). Lattice independent component analysis feature selection on diffusion weighted imaging for Alzheimer's disease classification. *Neurocomputing* 114, 132–141. doi:10.1016/j.neucom.2012.08.044
- Thomas, C., Humphreys, K., Jung, K.-J., Minshew, N., and Behrmann, M. (2011). The anatomy of the callosal and visual-association pathways in high-functioning autism: a DTI tractography study. *Cortex* 47, 863–873. doi:10.1016/j.cortex.2010.07.006
- Tuch, D. S. (2004). Q-ball imaging. *Magn. Reson. Med.* 52, 1358–1372. doi:10.1002/mrm.20279
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi:10.1006/nimg.2001.0978
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., et al. (2013). The WU-Minn human connectome project: an overview. *Neuroimage* 80, 62–79. doi:10.1016/j.neuroimage.2013.05.041
- Van Lieshout, M. (2000). *Markov Point Processes and Their Applications*. World Scientific.
- Wakana, S., Caprihan, A., Panzenboeck, M. M., Fallon, J. H., Perry, M., Gollub, R. L., et al. (2007). Reproducibility of quantitative tractography methods applied to cerebral white matter. *Neuroimage* 36, 630–644. doi:10.1016/j.neuroimage.2007.02.049
- Wee, C.-Y., Yap, P.-T., Zhang, D., Denny, K., Browndyke, J. N., Potter, G. G., et al. (2012). Identification of MCI individuals using structural and functional connectivity networks. *Neuroimage* 59, 2045–2056. doi:10.1016/j.neuroimage.2011.10.015
- Wu, H., Chen, G., Jin, Y., Shen, D., and Yap, P.-T. (2015) "Accelerating global tractography using parallel markov chain monte carlo," in *MICCAI Workshop on Computational Diffusion MRI (CDMRI)*, Munich.
- Yap, P. T., An, H., Chen, Y., and Shen, D. (2014). Fiber-driven resolution enhancement of diffusion-weighted images. *Neuroimage* 84, 939–950. doi:10.1016/j.neuroimage.2013.09.016
- Yap, P. T., Chen, Y., An, H., Yang, Y., and Gilmore, J. H. (2011a). SPHERE: spherical harmonic elastic registration of HARDI data. *Neuroimage* 55, 545–556. doi:10.1016/j.neuroimage.2010.12.015
- Yap, P.-T., Gilmore, J. H., Lin, W., and Shen, D. (2011b). PopTract: population-based tractography. *IEEE Trans. Med. Imaging* 30, 1829–1840. doi:10.1109/TMI.2011.2154385
- Yap, P.-T., Wu, G., Zhu, H., Lin, W., and Shen, D. (2010). F-TIMER: fast tensor image morphing for elastic registration. *IEEE Trans. Med. Imaging* 29, 1192–1203. doi:10.1109/TMI.2010.2043680

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Wu, Chen, Jin, Shen and Yap. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.