NeuroImage 263 (2022) 119600

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Bundle-o-graphy: improving structural connectivity estimation with adaptive microstructure-informed tractography

Matteo Battocchio [a,c,*], Simona Schiavi [a,b], Maxime Descoteaux [c], Alessandro Daducci [a]

[a] Department of Computer Science, University of Verona, Verona, Italy
[b] Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), University of Genoa, Genova, Italy
[c] Sherbrooke Connectivity Imaging Laboratory (SCIL), Departement dInformatique, Universite de Sherbrooke, Sherbrooke, Quebec, Canada

ARTICLE INFO

Keywords:
Microstructure-informed tractography
bundle-o-graphy
structural connectivity
clustering
MCMC adaptation

ABSTRACT

Tractography is a powerful tool for the investigation of the complex organization of the brain in vivo, as it allows inferring the macroscopic pathways of the major fiber bundles of the white matter based on non-invasive diffusion-weighted magnetic resonance imaging acquisitions. Despite this unique and compelling ability, some studies have exposed the *poor anatomical accuracy* of the reconstructions obtained with this technique and challenged its effectiveness for studying brain connectivity. In this work, we describe a novel method to readdress tractography reconstruction problem in a global manner by combining the strengths of so-called generative and discriminative strategies. Starting from an input tractogram, we parameterize the connections between brain regions following a bundle-based representation that allows to drastically reducing the number of parameters needed to model groups of fascicles. The parameters space is explored following an MCMC generative approach, while a discrimininative method is exploited to globally evaluate the set of connections which is updated according to Bayes' rule. Our results on both synthetic and real brain data show that the proposed solution, called *bundle-o-graphy*, allows improving the anatomical accuracy of the reconstructions while keeping the computational complexity similar to other state-of-the-art methods.

1. Introduction

First tractography algorithms for inferring the macroscopic fiber trajectories, called streamlines, of the white matter (WM) were based on "*local*" integration procedures of the fiber orientations estimated in each voxel from the measured diffusion-weighted magnetic resonance imaging data (DW-MRI). This approach is very fast but also rather sensitive to estimation errors of the local orientations (Mori et al., 1999). Maier-Hein et al. (2017) showed that such algorithms tend to follow the easiest path available in crossing regions, which represent the majority of WM voxels (Jeurissen et al., 2012), and thus fail to reconstruct some anatomical bundles, i.e. false negatives. To deal with this inadequacy of tractography to explore the whole space of brain connections, probabilistic alternatives were proposed that use probability distributions estimated in each voxel to allow uncertainty in the propagation of the trajectories. These methods have demonstrated their ability to recover hard-to-track connections and to cover the WM more adequately (Côté et al., 2013); however, this improved capability of exploring brain anatomy leads also to the reconstruction of implausible fascicles that do not anatomically exist, i.e. false positives. The effects of such false-negative and false-positive connections in tractography reconstructions has been recently investigated and, in particular, Zalesky et al. (2016) have demonstrated that these spurious connections are detrimental to the study of brain connectivity networks and they can heavily bias all analyses based on this technique.

The advent of the so called "*global*" tractography algorithms marked an important milestone on the road to significantly improve the quality of the reconstructions. These approaches introduced the use of global optimization to reconstruct the set of streamlines, i.e. tractogram, that are most consistent with the acquired DW-MRI data and, indeed, the resulting reconstructions showed improved anatomical accuracy (Christiaens et al., 2015; Close et al., 2015; Fillard et al., 2009; Kreher et al., 2008; Mangin et al., 2013; Reisert et al., 2011;). First solutions were based on Monte Carlo Markov Chain (MCMC) stochastic procedures for constructing the optimal set of streamlines, but this strategy turned out to be computationally very heavy given the large amount of parameters that need to be optimized and do not ensure the biological plausibility of the reconstructed connections (Girard and Descoteaux, 2012; Smith et al., 2012).

One major step to reduce the complexity of this *generative* strategy, while keeping a global approach, was made with the introduction of *discriminative* approaches. The idea behind these methods is to iden-

* Corresponding author.
E-mail address: matteo.battocchio@univr.it (M. Battocchio).

https://doi.org/10.1016/j.neuroimage.2022.119600.
Received 27 April 2022; Received in revised form 26 July 2022; Accepted 1 September 2022

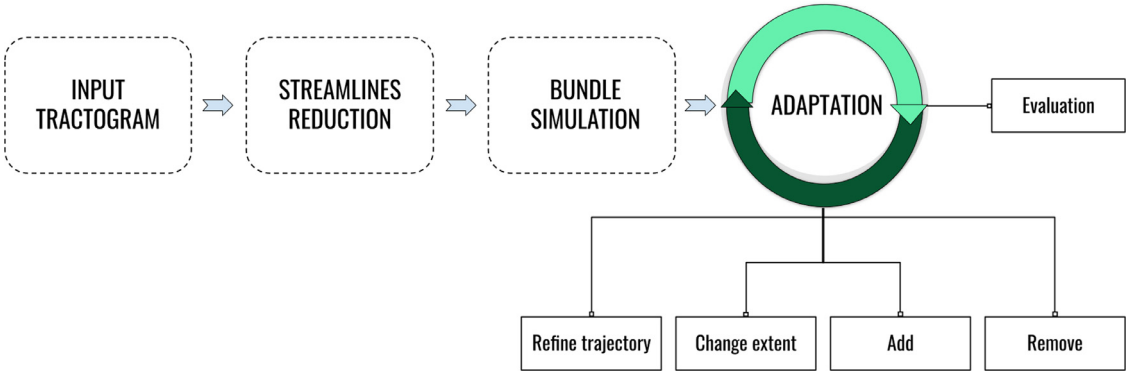

Fig. 1. Bundle-o-graphy workflow. Given an input tractogram the first step aims to reduce the number of streamlines needed to represent pathways between pairs of regions. Additionally, each streamline is parameterized using a subset of the initial points to approximate their trajectory. The second step is characterized by the shift from streamline to bundle-based representation. Here a volume is assigned to each streamline, allowing to mimic the contribution of a set of aligned fibers centered around the original trajectory. Finally the bundle-based configuration is optimized by adapting their shape and geometry following a MCMC optimization approach.

tify the optimal subset of streamlines from a pre-computed set of candidates, constructed using standard tractography algorithms, that are most compatible with the measured DW-MRI data; for this reason, they are sometimes referred to as *filtering* methods (Daducci et al., 2016). Different formulations exist; for instance, in SIFT/SIFT2 (Smith et al., 2013; 2015) streamlines are selected based on the agreement between their trajectories and the local fiber orientation distributions, whereas the full measured DW-MRI data is considered in COMMIT/COMMIT2 and LiFE (Daducci et al., 2015; Pestilli et al., 2014; Schiavi et al., 2020). These discriminative methods allowed reducing the computational cost required to perform global tractography dramatically, and showed great potential to further improve the quality of the reconstructions, notably alleviating the problem of false positives as well as improving the biological interpretability of the tractograms (Jbabdi and Johansen-Berg, 2011; Schiavi et al., 2020; Smith et al., 2020). However, unlike generative methods, they assume a static input configuration, i.e. shape / position of the candidate streamlines is fixed and cannot be modified, which means that the quality of the reconstructions remains indissolubly bounded to the quality of the algorithm used to build the candidate pathways.

In this work, we present a hybrid method which allows inheriting the strengths of both generative and discriminative approaches. Our solution tackles tractography reconstruction from a different perspective: the idea is to move away from streamline-based tracking with the aim to directly reconstruct bundles of them; for this reason we call it *bundle-o-graphy*. Thanks to a convenient parameterization, we can model groups of coherent streamlines using a minimal set of parameters which, in turn, allows us to extend a state-of-the-art *discriminative* method, i.e. COMMIT (Daducci et al., 2015), with the possibility of efficiently adapting the configuration of the bundles as in generative approaches. Our experiments conducted both on synthetic and real data clearly indicate the potential of our solution for improving the anatomical accuracy of the reconstructions.

2. Methods

The general structure of the algorithm is presented in Fig. 1 and the general workflow takes inspiration from our previous work (Battocchio et al., 2021). Bundle-o-graphy takes as input a tractogram that can be the computed using any tractography algorithm, or it can be the combination of different reconstructions. As first step, we divide the input tractogram based on a GM parcellation. Each connection is then clustered to reduce the number of streamlines by keeping only the most representative ones, which are parameterized using splines. As second step, the resulting fibers are used as prior to represent fascicles of coherent streamlines aligned along each pathway, hence introducing the concept of bundle-based representation. The set of bundles constitutes the configuration to optimize using an MCMC iterative approach, that allows to adapt their shape and extent along with the possibility to add and remove entire connections. The configuration is globally evaluated and the process is driven by a Bayesian framework to find the optimal configuration that best explains the observed signal.

In the following, we provide more details about each stage of the algorithm.

2.1. Streamline reduction and simplification

Streamline reduction is performed based on hierarchical clustering, following the approach implemented by Schiavi et al. (2020). We first divide the streamlines with respect to anatomical information, in particular based on the regions they connect given a cortical and subcortical brain parcellation (Fig. 2A). Secondly, each group of streamlines is clustered based on geometrical criterion, in our case represented by their average euclidean distance. To this aim, we exploit QuickBundle (Garyfallidis et al., 2012) (Fig. 2B) to reduce the number of streamlines needed to represent connections between regions. The clustering threshold can be manually chosen by the user and for our tests, we set it to 3 mm. Thanks to this procedure we can simplify the bundle representation and reduce the tractogram complexity. In particular, by removing unwanted redundancy we are able to downsize the input tractogram using, on average, only the 2% of the initial set of streamlines.

The resulting streamlines are then simplified to minimize the number of parameters needed to represent each trajectory, as in the work of Lemkaddem et al. (2014), using the Douglas-Peucker reduction algorithm (Douglas and Peucker, 1973), which selects the minimal subset of coordinates given an approximation threshold (Fig. 2C). We reduce the number of coordinates to represent each streamline using between 4 and 6 control points, in line with previous works (Jbabdi et al., 2007; Lemkaddem et al., 2014). The set of points is then interpolated using a particular class of cubic B-splines, named Catmull-Rom (Catmull and Rom, 1974)(Fig. 2D), characterized by several desirable properties. First of all, they preserve the starting and ending points, i.e. the first and last points remain the original ones, meaning that the original connectivity is preserved giving the fact that the reduced streamlines connect the same regions. Secondly, the interpolated trajectories intersect all the control points, which permits a better supervision of their spatial position and to avoid the reconstruction of pathways outside white matter regions.

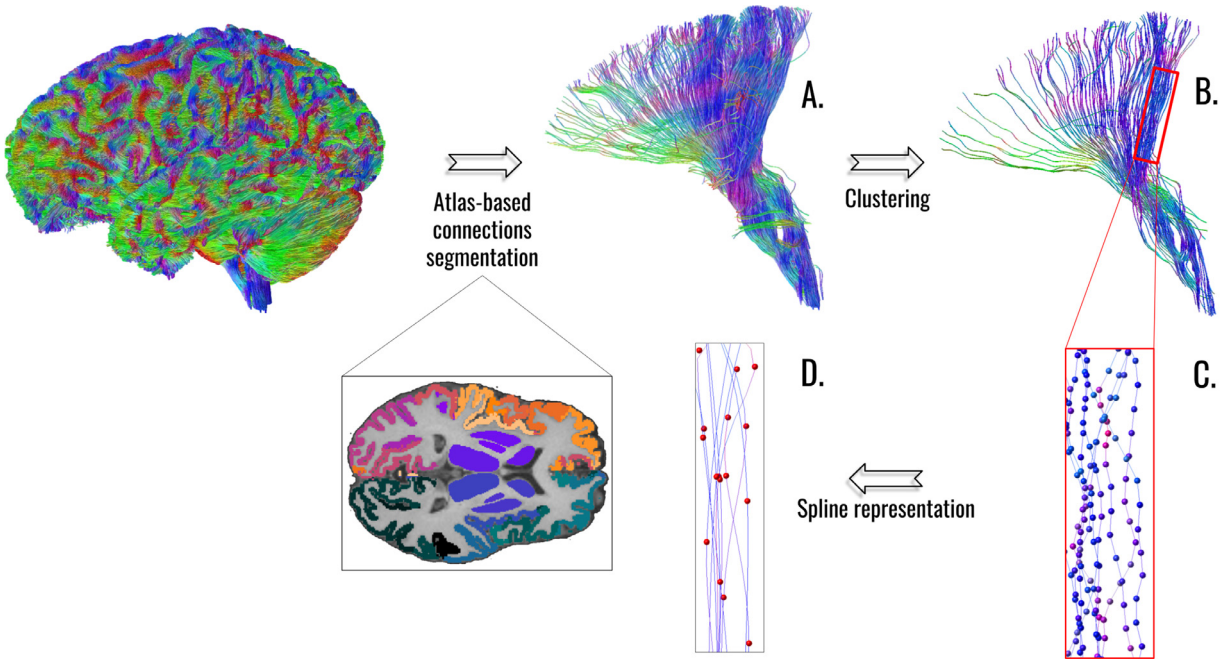

Fig. 2. Streamline reduction workflow: the first step implies separating the input tractogram into subsets based on a cortical and subcortical regions segmentation. Fig. 2A is an example showing a set of projection fibers segmented based on Freesurfer atlas. Each connection is than clustered (Fig. 2B) keeping only the representative streamlines for each cluster. These are simplified using Ramer-Douglas-Peucker algorithm (Fig. 2C) which reduces the number of points needed to approximate the streamline trajectories. Finally these coordinates are interpolated using cubic b-splines (Fig. 2D).

2.2. Bundle representation

Once the input tractogram has been reduced and simplified, we assign a volume to each streamline by exploiting a feature embedded in COMMIT. In particular, each tract is considered as the centroid of a cylinder, with constant radius, extending along the whole trajectory. Following COMMIT formulation, the signal contribution of a streamline can be computed based on the trajectory, i.e the voxel it traverses, and the response function adopted. To extend the contribution of a streamline to the neighbor voxels we create a set of replicas displaced equidistantly over concentric circles of increasing radius centered around each point of the fiber. These are computed internally, meaning that no further streamlines are added to the configuration. Starting in correspondence of the initial point, all circles lie on a plane that is always orthogonal to the streamline direction. Each point of the replicas is then computed based on Frenet-Serret frames (Frenet, 1852; Serret, 1851) which allows to compute the displacement of the replicas' following points along the streamline trajectory. The circles discretization, i.e., the number of replicas created, and the number of circles used to sample the space are empirically fixed. The signal contribution corresponding to the bundle is computed considering all the voxels traversed by the centroid and its replicas. While the signal contribution is constant along the trajectory, it can vary as we move outward from the center to take care of uncertainty at the boundaries of the bundle. To do so we implemented a *blurring function*, used to radially scale the signal contribution, shown in Fig. 3 and defined as follows:

$$P(x) = \begin{cases} 1 & \text{if } x < \sigma_C \,, \\ \exp\left(-\frac{(x-\sigma_C)^2}{2\sigma_G^2}\right) & \text{otherwise.} \end{cases} \tag{1}$$

Based on Eq. 1, the signal contribution of the bundle core, which extent is modulated by the parameter σ_C, is not scaled, while the signal corresponding to the replicas falling farther is exponentially reduced depending on the distance x from the center according to a Gaussian damping function regulated by σ_G. A more detailed description of the bundle creation process can be found in Daducci et al. (2021). Following this implementation, bundle simulation turns into modeling the space of influence of a streamline which requires only the parameter σ_C while σ_G is empirically fixed.

2.3. Optimization

The resulting set of bundles constitutes the initial configuration to adapt. More specifically, the collection of control points representing each bundle trajectory, along with their corresponding blur extent, controlled by σ_C, represent the parameters to optimize. In particular, the shape and the volume of each bundle is adapted following a generative approach, based on MCMC, which permits to exploit global information to better adapt the reconstruction with respect to the underlying WM structure. Given the observed data d and a set of competing models for the data $\{M_i | i = 1, 2, 3...\}$, each defined by a set of parameters θ^i, we can compute the posterior probability distribution function following the Bayes' theorem:

$$p(M_i|d) = \frac{p(d|M_i)p(M_i)}{p(d)}, \tag{2}$$

where $p(d|M_i)$ represents the likelihood of observing the experimental data given the model M_i, $p(M_i)$ is the prior probability of the model parameters and $p(d)$ is a normalizing constant.

In this context, M_i represents the parametric representation of a set of bundles defined by θ^i control points coordinates in the 3D space along with the corresponding blur extent σ_C, and $p(d|M_i)$ scores how well the configuration M_i explain the measured diffusion data d.

At iteration i, the probability of a given configuration is defined by the following distribution:

$$f_T(M_i) = \exp\left(-\frac{E_D(M_i, d)}{T_i}\right) \exp\left(-\frac{E_P(M_i)}{T_i}\right), \tag{3}$$

where, in our case, f_T represents the non-normalized joint distribution

$$p(d|M_i)p(M_i) \propto p(M_i|d). \tag{4}$$

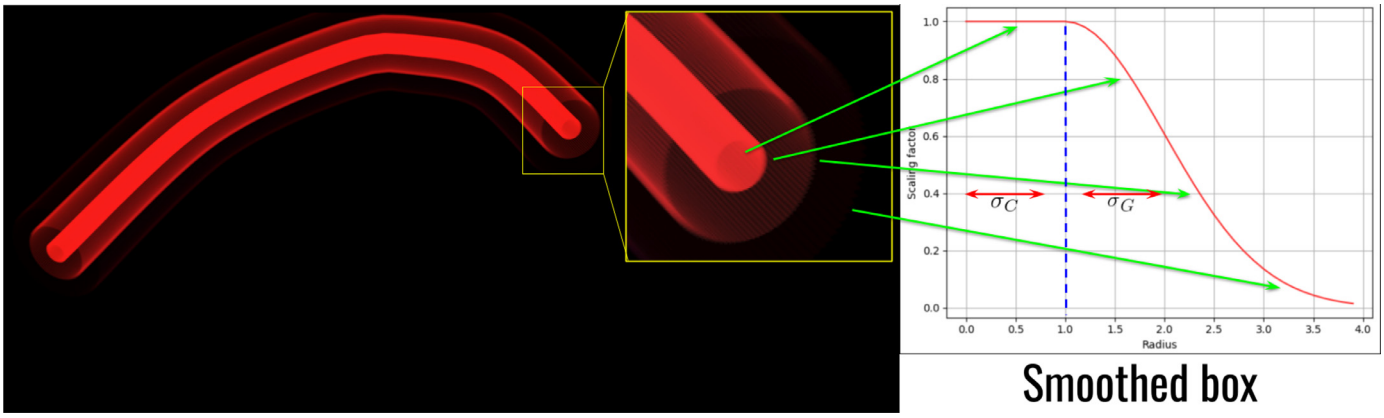

Fig. 3. Bundle simulation: the contribution of each original streamline can be modeled as a cylinder centered along the trajectory. The volume can be radially adapted using a blurring function to allow uncertainty as we move outward from the central pathway. The two parameters, σ_C and σ_G regulate the extent of the core and the Gaussian dumping respectively.

In particular, the first term, $e^{-\frac{E_D(M_i,d)}{T_i}}$, defines the likelihood function, while the second is independent from the measured data and defines the prior probability. Finally T_i is the system temperature at iteration i, a parameter used to speed up the convergence. The likelihood function scores how well the bundles configuration explains the observed data. This is computed as the differences between the original and the reconstructed signal computed based on a local forward model. In our case, instead of the full DW signal, we choose to fit the Intracellular (IC) signal fraction using a simple forward model that assigns a contribution, i.e., volume or cross-sectional area, to each bundle i of the input tractogram proportionally to its length L_i inside each voxel v. The voxel-wise signal can be expressed as:

$$S_v = \sum_{i=1}^{N} x_i L_i, \tag{5}$$

where N is the number of bundles passing through the voxel and x_i represents the actual contributions of the bundle i, estimated with COMMIT, needed to explain the acquired data d. Based on Eq. 5, the likelihood becomes

$$E_D(M_i, d) = \sqrt{\left(\frac{1}{V}\right)\sum_{v=1}^{V}(S_v - d_v)^2}, \tag{6}$$

where V is the number of voxels in the WM volume and d_v the measured voxel signal.

Our prior knowledge is represented by $E_P(M_i)$ as follows:

$$E_P(M_i) = \lambda|B_i| + \beta|M_i|, \tag{7}$$

where $|B_i|$ represents the number of connections, in the configuration M_i, between pairs of regions of the cortical and subcortical parcellation and $|M_i|$ the total number of blurred streamlines, i.e. bundles, constituting these connections. Finally, λ and β are fixed parameters to control the regularization terms to balance likelihood and priors in the cost function.

To maximize the posterior probability we adopted a MCMC Metropolis-Hastings-Green (Green, 1995; Hastings, 1970; Metropolis et al., 1953) sampling approach and Simulated Annealing (Perrin et al., 2005) optimization to explore the space of parameters (see Fig. S1 in the supplementary details for further details on the sampling efficiency and convergence).

Given the current configuration M_i, a new configuration M_{i+1} is accepted with probability:

$$p_{\text{accept}} = min(1, R) \tag{8}$$

where R is the Green's ratio

$$R = \frac{f_T(M_{i+1})}{f_T(M_i)}\frac{p(M_i|M_{i+1})}{p(M_{i+1}|M_i)}\frac{q(M_{i+1})}{q(M_i)}, \tag{9}$$

with p describing the probability density functions associated to the proposal to move from configuration M_i to configuration M_{i+1}, and q is the probability to choose one among the four proposals listed below:

1. add connection
2. remove connection
3. refine trajectory
4. change blur extent

Regarding the first proposal, a new connection is added to the configuration by drawing randomly and uniformly from the set of connections A,which is composed by all the connections removed from M during the optimization. The associated probability density takes the form $p(M_{i+1}|M_i) = |A|^{-1}$ where $|A|$ is the number of connections in A. Conversely, the removal of a connection from the configuration is equal to the probability of uniformly sampling from the current configuration M_i, meaning $p(M_i|M_{i+1}) = |B_i|^{-1}$. The corresponding Green's ratio becomes:

$$R = \frac{f_T(M_{i+1})}{f_T(M_i)}\frac{|B_i|^{-1}}{|A|^{-1}}\frac{q_r}{q_a}, \tag{10}$$

where q_r and q_a are, respectively, the probabilities for choosing the *Remove* and the *Add* transitions.

While, for the first and second proposals, the probability density functions are asymmetric, meaning that $p(M_i|M_{i+1}) \neq p(M_{i+1}|M_i)$, refining the trajectory and change the blur extent rely on normal probability density functions, which are symmetric. In particular, the trajectory can be adapted in two ways: by moving a single control point or by moving the whole bundle. The first consists in sampling the plane orthogonal to the bundle direction from a normal distribution of points centered around the coordinates of each control point. For the movement of an entire bundle we use the same sampling approach, but extended to all the control points. The last proposal randomly picks a bundle from M and adapts its blur extent, i.e the bundle cross-sectional area, by sampling the new value of σ_C from a normal distribution. Given these symmetric proposals, the acceptance probability becomes proportional to how likely each of the current state M_i and the proposed state M_{i+1} are under the full joint density. Hence, the Green's ratio for these proposals is given by:

$$R = \frac{f_T(M_{i+1})}{f_T(M_i)}. \tag{11}$$

Based on the simulated annealing approach, at the beginning, the system is characterized by a high temperature (T), which decreases as the process advances. High values of T imply that "bad" configurations are accepted, allowing the system to explore a wider range of configurations. In previous works, it has been shown how a geometric lowering

schedule of the temperature ensures the convergence (Lieshout, 1994; Perrin et al., 2005) and, at the same time, improves the chances to sample from maxima of $P(M_i|D)$. Starting from a model $M_{i=0}$, the space of its parameters is explored by altering the configuration.

To better asses convergence we run two separate optimizations. In the first case, the system is allowed only to add or remove connections while, in the second, all four proposals are adopted.

2.4. Data and Experiments

To show the effectiveness of the method, we tested it on both synthetic and real data.

Synthetic data. We used the dataset provided for the IEEE International Symposium on Biomedical Imaging (ISBI) 2013 Reconstruction Challenge (Caruyer et al., 2014), which simulates an acquisition protocol with 64 directions at b-value=3000 s/mm^2, 1 mm isotropic voxel and signal-to-noise ratio of 30. This dataset consists of 27 fascicles arranged in a configuration mimicking most of the challenging bundle configurations that can be found in the brain, like bundles with various diameters branching, kissing and crossing at different angles. From the signal reconstruction point of view, the phantom reproduces both partial volume effects, given by the presence of multiple fiber compartments within the same voxel and cerebrospinal fluid (CSF) contamination. We performed streamline reconstruction with three different tractography algorithms iFOD2 (Tournier et al., 2010), SD_Stream (Tournier et al., 2012), Trekker (Aydogan and Shi, 2021), generating 1 million streamlines using default parameters for each. We divide the tractograms into bundles based on the parcellation provided with the dataset and then performed clustering to reduce the number of streamlines used to represent each connection.

In vivo brain data. We also evaluated bundle-o-graphy on in vivo human data from the HCP test-retest dataset (Van Essen et al., 2013). We downloaded the preprocessed diffusion data corresponding to subject 172332 and the structural T1-weighted image with the corresponding standard Desikan-Killiany (Desikan et al., 2006) parcellation in 85 gray matter ROIs performed with FreeSurfer (Fischl et al., 2004). To do so, we first segment the T1-weighted image using FMRIBs automated segmentation tool (Zhang et al., 2001) to derive the multi-tissue image. This allowed performing the tissue-informed multi-shell spherical deconvolution and to recover the fiber orientation distributions (Jeurissen et al., 2014). We performed three whole brain reconstructions, using SDStream, iFOD2 and Trekker. For the deterministic and probabilistic methods we perform anatomically constrained tractography (Smith et al., 2012) with default parameters, generating 3 million streamlines, while, for Trekker, we used the white matter mask as the seed region for the tracking, generating 1 million streamlines.

2.5. Evaluation metrics

For each dataset we computed the IC signal fraction maps in each voxel. Different models can be used for the estimation, as standard models like neurite orientation dispersion and density imaging (NODDI) (Zhang et al., 2012) or spherical mean technique (SMT) (Kaden et al., 2016), implemented as open-source available at[1]. We believe that the choice of the model to use does not affect the validity of our method and, in our case, we arbitrarily decided to use the SMT.

For the synthetic dataset we processed the three input tractograms and computed the total white-matter overlap based on the corresponding IC maps, i.e., the total percentage of WM volume covered by the streamlines. For all the reconstructions, we investigated the connectivity, reporting the strength and the L1- and L2-distances between estimated and ground-truth (GT) connectivity matrices. In our case, a matrix entry reflects the connectivity strength between two regions, defined as the sum of the bundles cross-sectional area each multiplied by the corresponding weight estimated by COMMIT from the fiber density map. We assessed the number of true positive connections (TPCs), i.e., number of connections between pair of regions known to be connected and the false positives connections (FPCs), i.e., the number of those connecting regions known to be disconnected. We also reported the connectivity estimation focusing on a well known hard-to-track connection along with the WM coverage computed by segmenting the the WM mask to isolate that specific connection. Then, we carried the same analysis on the resulting configuration showing the effects of bundle-o-graphy optimization.

For the in vivo brain dataset, we compared the estimated IC maps computed by COMMIT and the signal fitting root-mean-square-error (RMSE) corresponding to the three reconstruction before and after the application of bundle-o-graphy. We also carried out analysis focusing on three major connections, Corpus Callosum (CC), Pyramidal Tract (PyT) and Arcuate Fasciculus (AF), reporting the corresponding estimated IC maps before and after the optimization along with their WM coverage.

3. Results

A visual inspection of the impact of bundle-o-graphy is shown in Fig. 4. The first row shows the IC maps estimated by COMMIT from the static reconstructions, performed with the three tractography algorithms, of a well known hard-to-track connection. We also reported the respective connectivity strength and WM overlap. On the second row the corresponding results of the optimization with bundle-o-graphy. Starting from an underrepresented connection, the method is able to better distribute the streamlines inside the WM volume and adapt their spatial extent, converging to the same connectivity value across tractograms computed with different methods. Table 1 is a summary of the results on the synthetic dataset. Quantitatively, bundle-o-graphy allows to reach full coverage of WM using less then 2% of the input streamlines. The impact is particularly notable on the hard-to-track connection between region 5 and 6, where the WM coverage increases by almost three fold for SD_STREAM after the adaptation. This, in turn, allows to better estimate its connectivity strength, which converges to similar values across the three adapted configurations. At global level, the presented method drastically reduces the number of false positives while all the true positives connections are preserved (see Fig. 5), which, in turn, is reflected by the decrease in L1 and L2 connectivity distances with respect to the GT.

Results on in vivo dataset are shown in Fig. 6 and Fig. 7. The first shows a comparison between the signal fitting error associated to the three input tractograms before (first row) and after bundle-o-graphy optimization (second row). In all three optimized tractograms, the reduction in the RMSE is coupled with an improved streamlines density estimation, as can be see in Fig. 7 (second row). Thanks to bundle-o-graphy we are able to increase the overall WM coverage while improving density homogeneity at the same time. Fig. 8 shows the results focusing on three specific tracts, comparing the reconstructions before and after the application of bundle-o-graphy. For each connection we computed the IC maps using COMMIT, reporting in the first row those corresponding to the input and in the second row the adapted configurations. The bundles adaptation allows to improve the tracts WM coverage while better representing the underlying anatomy, converging to connections that shares similar signal density patterns along their trajectories. For the deterministic tractogram, bundle-o-graphy increases the tracts volume by three folds, while, for the input tract recovered using iFOD2 and Trekker, the improvement is by two folds on average. As for the synthetic dataset, bundle-o-graphy complexity reduction allows to achieve these results using a fraction of the initial number of streamlines, in particular around 7% for SD_STREAM, 32% for iFOD2 and 16% for Trekker. Finally, Fig. 9 shows the comparison between the cortical surface projections of the CC tract corresponding to the input reconstruction (first row)

[1] https://github.com/ekaden/smt

Fig. 4. Impact of the optimization on different reconstruction algorithms on the synthetic phantom. On the first and second row are the intra-cellular signal fraction maps corresponding to the vertical hard-to-track connection of the input and optimized configurations respectively, viewed from two different perspectives.

Table 1
Summary table of the comparison between SD_STREAM, iFOD2 and Trekker reconstructions before and after the optimization on the synthetic dataset.The size corresponds to the number of input streamlines for the static case, where only those connecting were kept, while it refers to the number of bundles for the adaptive one. The metrics include the total WM overlap and the one corresponding to a hard-to-track connection along with its estimated connectivity strength. Bundle-o-graphy improves overall white matter coverage and improves connection-specific and global connectivity estimation. The optimized tractograms present a reduced number of FPCs, while keeping all the TPCs.

		size	Total WM overlap	Connection 5-6 WM overlap	Connection 5-6 WM conn. strength	TPCs/FPCs	L1 conn. distance	L2 conn. distance	Computational time hh:mm
SD_STREAM									
	Static	772'793	80.21%	32.34%	0.116	27/142	0.36	0.105	
	Adaptive	3'798	99.89%	97.25%	0.640	27/6	0.23	0.094	1:40
iFOD2									
	Static	503'371	99.81%	42.78%	0.005	27/365	0.59	0.178	
	Adaptive	3'398	99.77%	92.61%	0.474	27/19	0.35	0.134	2:27
Trekker									
	Static	453'233	98.51%	67.35%	0.197	27/490	0.48	0.120	
	Adaptive	8'848	99.95%	100%	0.709	27/0	0.01	0.003	4:31

and the respective optimized connections. The results show how sub-cortical WM volume coverage is coupled with a significantly improved cortical projection extent.

4. Discussion

Results on both synthetic and real data show the impact of our bundle-based approach. Bundle-o-graphy is able to improve the sensitivity/specificity trade-off of tractography by considering two fundamental observations, as previously done by Schiavi et al. (2020), about brain anatomy to drive the adaptation process: (i) Streamlines can be "quantified" based on the underlying microstructure, and (ii) the WM structure is organized into bundles (Rheault et al., 2019; Schilling et al., 2020; Wasserthal and Neher, 2018; Wasserthal et al., 2019). To achieve that we exploit two priors: the first, $|B_i|$, minimizing the total number of connections and the second, $|M_i|$, to specifically leverage the

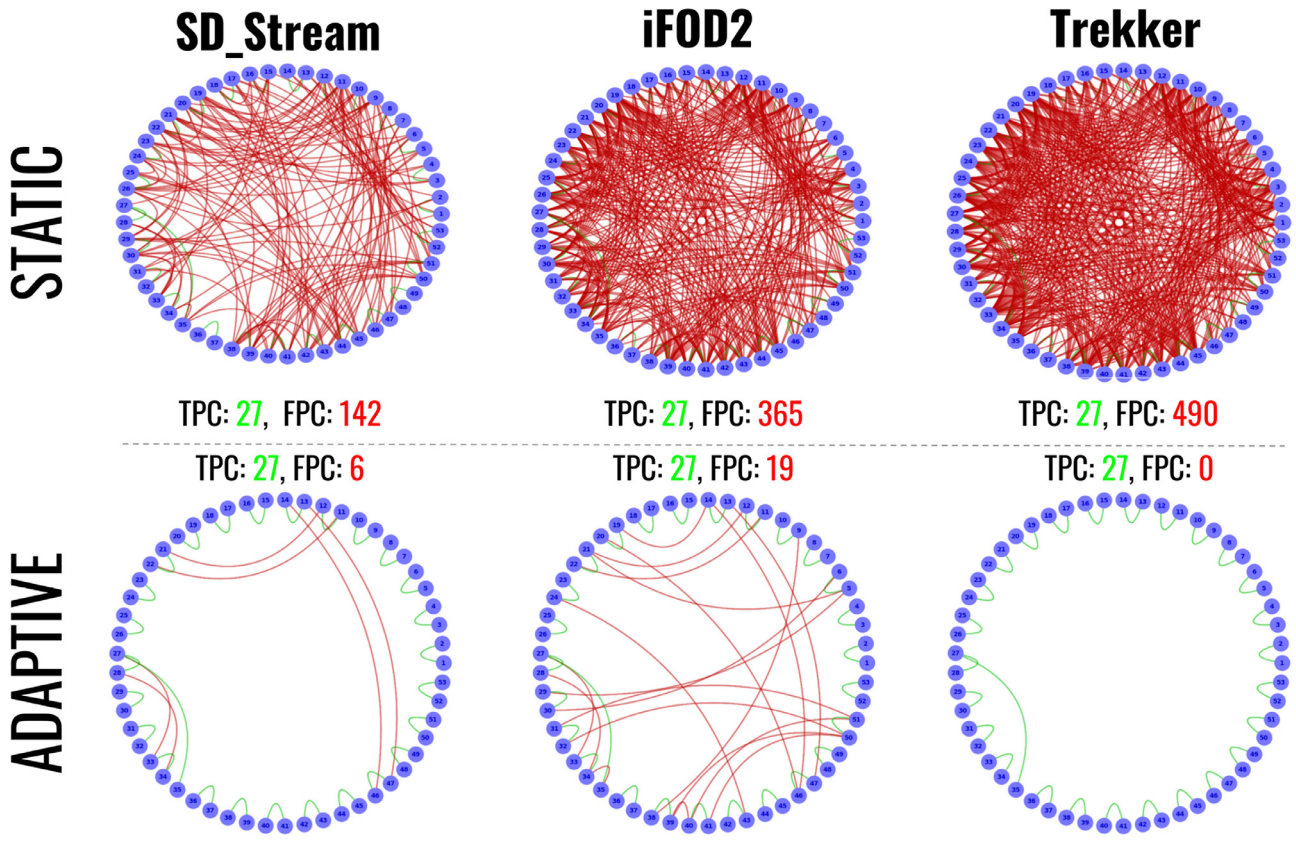

Fig. 5. Impact of the optimization on the removal of false positives connections. The connectivity graphs show in red the connections between pair of regions if they represent false positives connections or in green if they correspond to true positive ones. In the first row are the connectivity graphs of the input tractograms, while on the second row are the connectivity evaluated after bundle-o-graphy optimization.

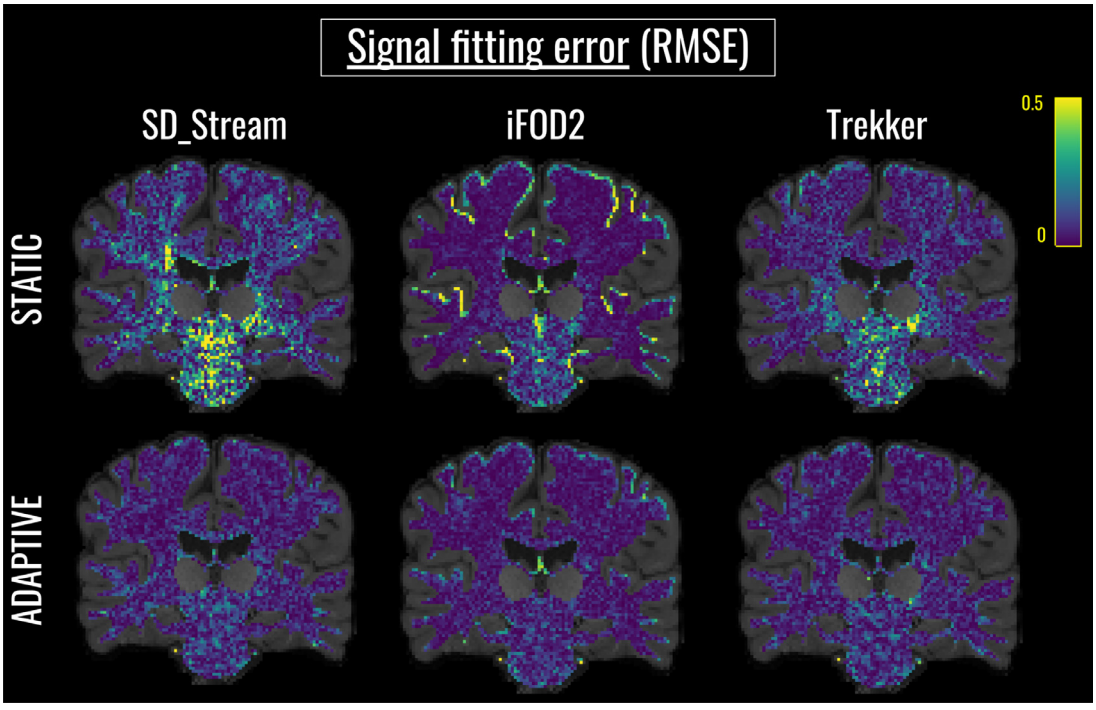

Fig. 6. Comparison between the voxel-wise RMSE maps corresponding to the input tractograms computed with the three different reconstruction algorithms (first row) and the adapted configurations (second row).

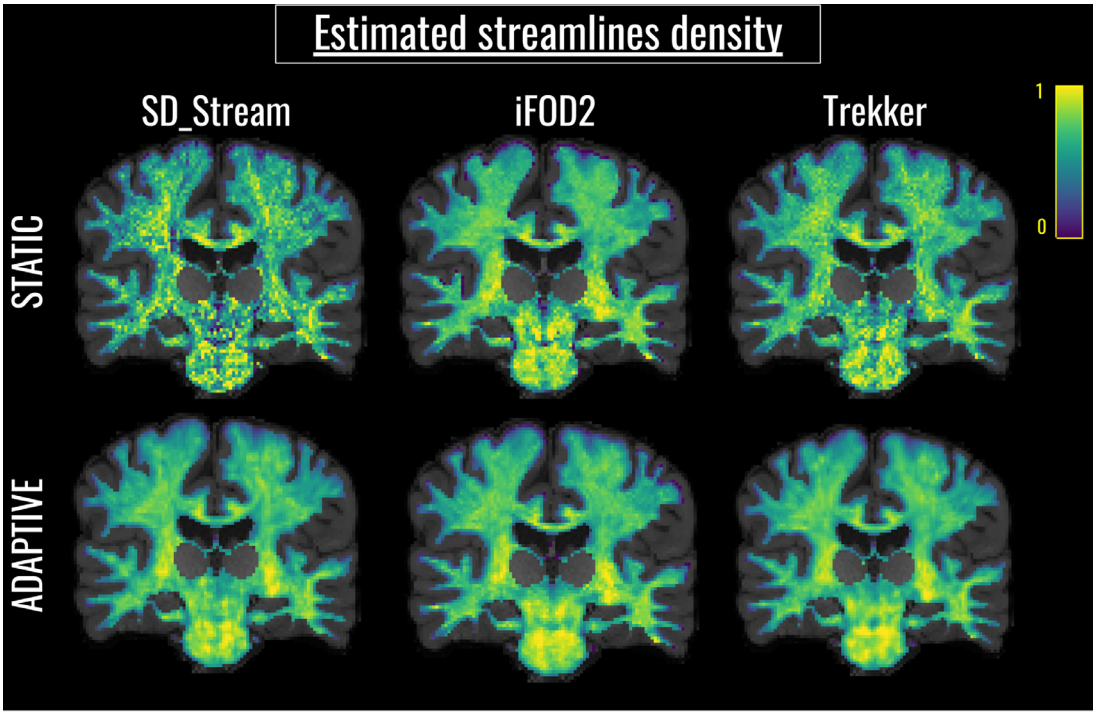

Fig. 7. Comparison between the estimated IC maps of the input tractograms (first row) and the adapted configurations (second row).

ability of the blur to reduce the complexity of the problem. In fact, the empirical image data can be equally explained using a configuration with many streamlines without blur, or using a configuration with fewer streamlines which have a blur associated to them (i.e. bundles). Clearly, this configuration requires less parameters to be estimated, and this second priors promotes this type of solutions (see Fig. S2 in the supplementary materials for further details on the impact of the priors in the optimization process). In this context, it's important to notice that adding/removing a connection and changing the σ_C value of a bundle does not have the same effect because they act at two different levels. The first allows for the addition or removal of the whole set of bundles representing a connection between two regions, while the second adapts the extent of an already existing bundle. For instance, a missing connection can be recovered only by adding a bundle, because the overlap between bundles belonging to a neighbor connection is limited and does not cover the whole extent of a connection.

With COMMIT2, Schiavi et al. (2020) achieve notable results by using a linear optimization approach but, differently from their work and discriminative methods in general, bundle-o-graphy is capable of improving tractography reconstructions by adapting the shape and the position of each bundle. The adaptation performed by the presented method has many benefits, both qualitatively and quantitatively. The first advantage is crucial in the case of WM pathways poorly reconstructed and hence underrepresented, as shown in Fig. 4, where, thanks to bundle-o-graphy, we were able to isolate and improve the reconstruction of the connection. In the second case, the removal of invalid connections and the bundles adaptation in terms of geometry and extent, allows to be remarkably independent from the tractography algorithm used to compute the initial set of streamlines, converging to similar connectivity estimates. Moreover, the possibility to represent groups of streamlines as set of packed and aligned fascicles leads to a more homogeneous and smooth tract density (see Fig. 7), which reduces the variability introduced by tracking algorithms. This is particularly evident for connections going through or passing by hard-to-track regions, where tractography algorithms are known to have challenges (see Fig. 6).

The idea of direct modeling groups of streamlines has been already explored in previous works, as in Close et al. (2015). However these methods often rely on many empirically tuned parameters to shape the geometry of the bundle and to compute the corresponding signal as well as the fact that the initial setup for each bundle need to be manually defined and they cannot reconstruct multiple bundles at once.

In our case, we don't aim to represent complex configurations with single bundles but, instead, modeling groups of streamlines by adopting a convenient parameterization. Following our approach, a connection between two regions is reconstructed using a set of bundles that, in the case of complex architecture, e.g fanning, allows to properly represent the whole extent of the tract, as shown in figure S6. One of the crucial step to reducing the number of parameters is represented by the streamlines reduction, described in section 2.1, coupled with the blur function that allows us to model groups of similar streamlines with much fewer parameters. The amount of streamlines kept, constituting the configuration we need to adapt, is influenced by three main factors: the input tractogram, the parcellation used and the clustering threshold applied. The amount of pathways reconstructed, either true or false positives, highly depends from the tracking algorithm used to compute the inital tractogram (see Fig. 5 first row). This fact, along with the adoption of atlases with different parcellation resolution, affects the initial total number of connections. Lastly, the clustering threshold used affects the number of representative streamlines extracted for each connection. In our experiments, using standard Desikan-Killiany atlas and clustering threshold of 3 mm, we reported a drop of 98% in the number of streamlines. The reduction in the number of coordinates to represent the streamlines is another fundamental step. This is combined with the interpolation using polynomial curves that closely follow the original anatomical trajectory. In Supplementary Fig. S4 we show how the fitting error changes with respect to the number of control points used to simplify the input streamlines on the synthetic phantom. In particular, we can see how the fitting error increases exponentially as we progressively reduce the number of control points. In our case, we chose an average number of 6 control points per streamline, shown by the blue vertical line, which represents a good trade-off between fitting error and parameters reduction. Control

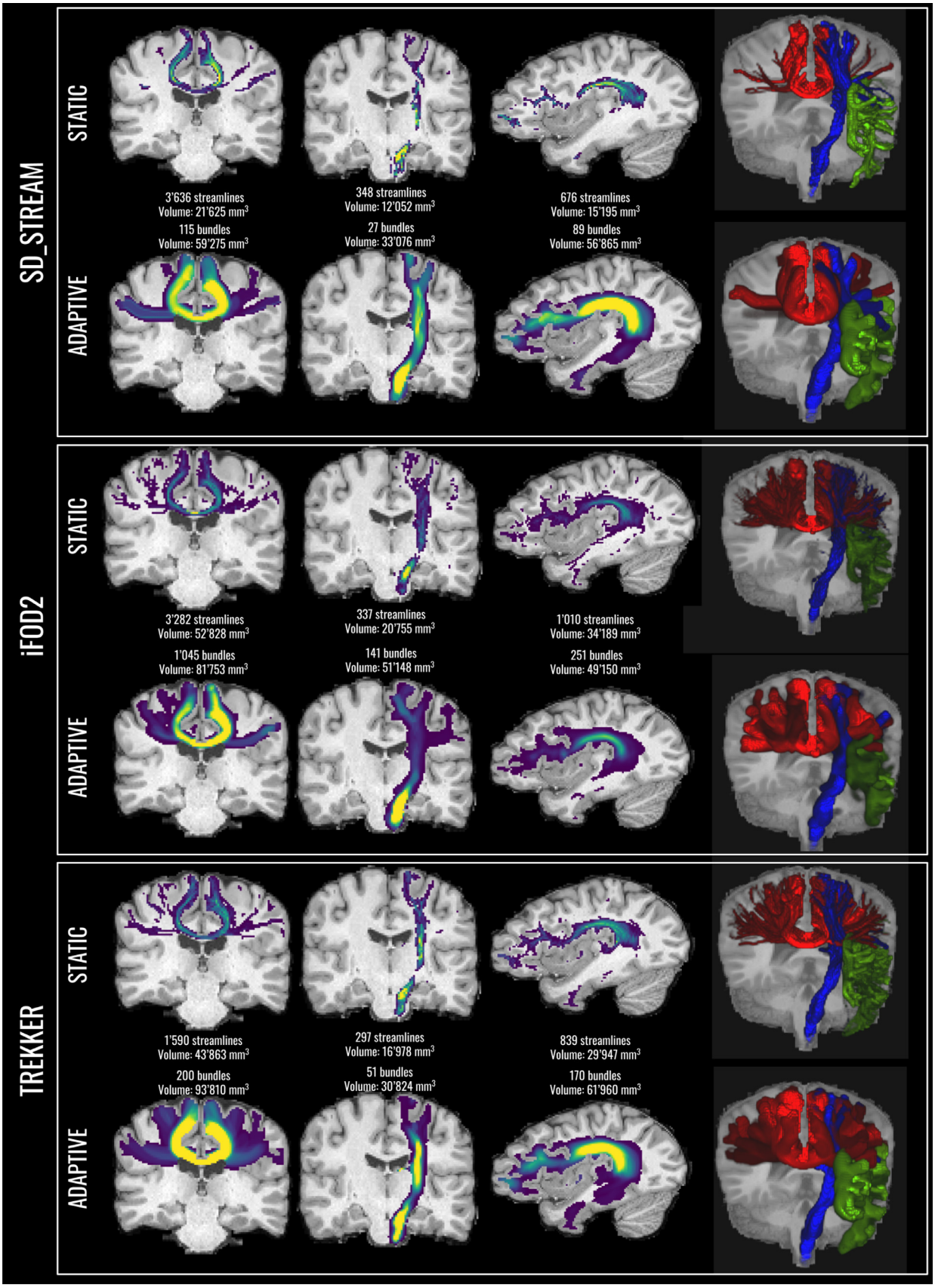

Fig. 8. Comparison between the IC maps, estimated with COMMIT, corresponding to the CC, PyT and AF reconstructed with the three different tractography algorithms. For each method, the first row shows the the connection segmented from the input tractogram while the second reports the connection segmented from the adapted one. The last column shows the corresponding visual inspection of the volume and geometry of the tracts. The visualization is performed in the same way and the thickness is related to the blur extent regulated by σ_C.

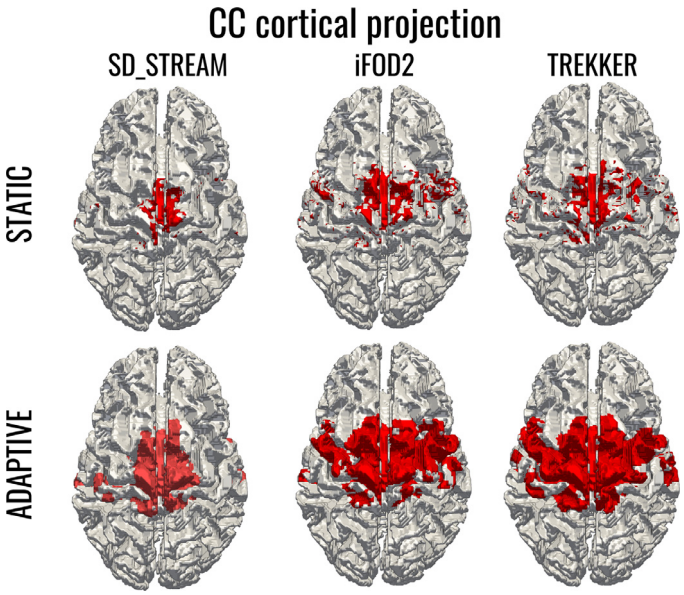

Fig. 9. Comparison between the cortical projections of the CC tracts segmented from the input tractograms computed with the three different reconstruction algorithms (first row) and the respectively adapted connections (second row).

point reduction is crucial for bundles trajectory adaptation. In fact, the possibility to model a pathway with few control points reduces notably the parameters space that need to be explored, speeding up the convergence. To furthermore reduce the computational cost, we empirically fixed some of the parameters. In particular, the Gaussian dumping, represented by σ_G, the two parameters controlling the regularization terms λ and β and the initial system temperature T. The first is used to model uncertainty on the borders of a bundle and, if set to zero it means that all the segments within the extent of the bundle contribute in the same way. The parameters λ and β are used to balance likelihood and priors in the cost function, while the initial temperature T is used to regulate the speed of convergence of the optimization process. Although moving to a simplified parametric representation of the input tractogram could affect the quality of the reconstruction, our results show how, thanks to the bundle simulation and adaptation, we were able to equivalently fit the data (see Fig. 6 and 7).

While the reconstruction method used to compute the initial set of streamlines still influences the performances, our approach allows to converge to similar connectivity and WM density patterns as shown in Fig. 4 and 8. The differences between the resulting bundles configurations can be explained by two main factors: initial number of false positive connections and reconstruction quality of the valid ones.While the first mostly impact the computational time, requiring more iterations to filter out the implausible connections, the second can affect the quality of the reconstruction. In our case, iFOD2 has a better initial coverage compared to SD_STREAM but its polluted by a greater amount of false positives, while true positives connections are still poorly reconstructed in correspondence of hard-to-track regions. In the case of valid connections particularly underrepresented, this implies that to cover the WM regions some of the invalids need to be kept to explain the signal, even if this is notably alleviated thanks to the adaptation and consequent improvement of the true positives coverage. The input tractogram computed using Trekker, although being characterized by a huge amount of false positives, provides a better support by properly recovering the true positives connections, allowing the algorithm to successfully identify and discard all the false ones. This is possible thanks to a global discriminative approach that allows to evaluate the bundle configuration as it is adapted throughout the optimization process. In our case we adopted COMMIT framework but bundle-o-graphy supports any global discriminative method that returns a fitness measure between the reconstructed and the observed signal. Moreover, thanks to its formulation, bundle-o-graphy facilitates the embedding of prior information to drive the reconstruction. For instance, these information can be exploited to introduce hierarchies between streamlines belonging to the same connection, as in COMMIT2tree Ocampo-Pineda et al. (2021), further improving our ability to filter out false positives connections. Currently, the algorithm starts from a large precomputed set of streamlines. Although efficient, this approach requires that the initial configuration contains all the true positive connections, a condition that is not always satisfied. A solution is to combine reconstructions computed with different techniques to avoid false negatives or to add a module that performs on-the-fly tracking, as the one implemented by Aydogan et al. (2021). This is part of future works, including the exploitation of priors coming from different imaging modalities, as in the work of Schiavi et al. (2022), that can be integrated in different ways.

5. Conclusions

Although tracking algorithms have shown a notable evolution, state-of-the-art streamline reconstructions are still anatomically inaccurate and difficult to reproduce, limiting their potential to study white matter connectivity which is fundamental to characterize the healthy structure of the human brain, as well as its perturbation in disease. Thanks to a convenient parameterization, bundle-o-graphy allows to combine both the potential of filtering techniques with the flexibility of generative global optimization approaches. We demonstrated the feasibility and the effectiveness of bundle-o-graphy both on synthetic and in vivo data, showing how bundle-o-graphy can improve the biological accuracy of the reconstruction regardless the input data.

We believe that our method could represent a step forward in characterizing and quantifying the structural connectivity by combining micro and macro-structure information.

Acknowledgements

We thank the University of Sherbrooke institutional research chair in Neuroinformatics and NSERC discovery grant from M. Descoteaux.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.neuroimage.2022.119600

References

Aydogan, D.B., Shi, Y., 2021. Parallel transport tractography. IEEE Transactions on Medical Imaging 40 (2), 635–647.

Aydogan, D.B., Souza, V.H., Lioumis, P., Ilmoniemi, R.J., 2021. Towards real-time tractography-based tms neuronavigation. Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation 14 (6), 1609.

Battocchio, M., Schiavi, S., Descoteaux, M., Daducci, A., 2021. Improving tractography accuracy using dynamic filtering. In: Gyori, N., Hutter, J., Nath, V., Palombo, M., Pizzolato, M., Zhang, F. (Eds.), Computational Diffusion MRI. Springer International Publishing, pp. 45–54.

Caruyer, E., Daducci, A., Descoteaux, M., Houde, J.-C., Thiran, J.-P., Verma, R., 2014. Phantomas: a flexible software library to simulate diffusion mr phantoms. Proc. Intl. Soc. Mag. Reson. Med. (ISMRM).

Catmull, E., Rom, R., 1974. A class of local interpolating splines. In: BARNHILL, R.E., RIESENFELD, R.F. (Eds.), Computer Aided Geometric Design. Academic Press, pp. 317–326.

Christiaens, D., Reisertbib2013, M., Dhollander, T., Sunaert, S., Suetens, P., Maes, F., 2015. Global tractography of multi-shell diffusion-weighted imaging data using a multi-tissue model. NeuroImage 123, 89–101.

Close, T.G., Tournier, J.-D., Johnston, L.A., Calamante, F., Mareels, I., Connelly, A., 2015. Fourier tract sampling (fouts): A framework for improved inference of white matter tracts from diffusion mri by explicitly modelling tract volume. NeuroImage 120, 412–427.

Côté, M.-A., Girard, G., Bor, A., Garyfallidis, E., Houde, J.-C., Descoteaux, M., 2013. Tractometer: Towards validation of tractography pipelines. Medical Image Analysis 17 (7), 844–857. Special Issue on the 2012 Conference on Medical Image Computing and Computer Assisted Intervention

Daducci, A., Dal Palu, A., Descoteaux, M., Thiran, J.-P., 2016. Microstructure Informed Tractography: Pitfalls and Open Challenges. Front Neurosci 10 (8), 247.
Daducci, A., Dal Palù, A., Lemkaddem, A., Thiran, J.-P., 2015. COMMIT: convex optimization modeling for microstructure informed tractography. IEEE Trans Med Imaging 34 (1), 246–257.
Daducci, A., Gobbi, F., N, F., M, B., S, S., 2021. Blurred streamlines: a new concept to improve tractography accuracy by spatially blurring signal contributions. Proc. Intl. Soc. Mag. Reson. Med. (ISMRM).
Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al., 2006. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. Neuroimage 31 (3), 968–980.
Douglas, D.H., Peucker, T.K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: The International Journal for Geographic Information and Geovisualization 10 (2), 112–122.
Fillard, P., Poupon, C., Mangin, J.-F., 2009. A Novel Global Tractography Algorithm Based on an Adaptive Spin Glass Model. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Vol. 12, pp. 927–934.
Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Sgonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex 14 (1), 11–22.
Frenet, F., 1852. Sur les courbes a double courbure. Journal de mathématiques pures et appliquées 437–447.
Garyfallidis, E., Brett, M., Correia, M., Williams, G., Nimmo-Smith, I., 2012. Quickbundles, a method for tractography simplification. Frontiers in Neuroscience 6, 175.
Girard, G., Descoteaux, M., 2012. Anatomical tissue probability priors for tractography. In: International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI12)-Computational Diffusion MRI Workshop, pp. 174–185.
Green, P.J., 1995. Reversible jump markov chain monte carlo computation and bayesian model determination. Biometrika 82 (4), 711–732.
Hastings, W.K., 1970. Monte carlo sampling methods using markov chains and their applications. Biometrika 57 (1), 97–109.
Jbabdi, S., Johansen-Berg, H., 2011. Tractography: Where Do We Go from Here? Brain Connectivity 1 (3), 169–183.
Jbabdi, S., Woolrich, M., Andersson, J., Behrens, T., 2007. A bayesian framework for global tractography. NeuroImage 37 (1), 116–129.
Jeurissen, B., Leemans, A., Tournier, J.-D., Jones, D.K., Sijbers, J., 2012. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. Human Brain Mapping 34 (11), 2747–2766.
Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., Sijbers, J., 2014. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data. NeuroImage 103, 411–426.
Kaden, E., Kelm, N.D., Carson, R.P., Does, M.D., Alexander, D.C., 2016. Multi-compartment microscopic diffusion imaging. NeuroImage 139, 346–359.
Kreher, B., Mader, I., Kiselev, V., 2008. Gibbs tracking: A novel approach for the reconstruction of neuronal pathways. Magn Reson Med 60 (4), 953–963.
Lemkaddem, A., Skildebrand, D., Dal Pal, A., Thiran, J.-P., Daducci, A., 2014. Global Tractography with Embedded Anatomical Priors for Quantitative Connectivity Analysis. Frontiers in Neurology 5, 232.
Lieshout, M.N.M.V., 1994. Stochastic annealing for nearest-neighbour point processes with application to object recognition. Advances in Applied Probability 26 (2), 281–300.
Maier-Hein, K.H., Neher, P., Stieltjes, B., Descoteaux, M., 2017. The challenge of mapping the human connectome based on diffusion tractography. Nature Communications 5.
Mangin, J.-F., Fillard, P., Cointepas, Y., Bihan, D.L., Frouin, V., Poupon, C., 2013. Toward global tractography. NeuroImage 80, 290–296.
Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. The Journal of Chemical Physics 21 (6), 1087–1092.
Mori, S., Crain, B.J., Chacko, V.P., Van Zijl, P.C.M., 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Annals of Neurology 45 (2), 265–269.
Ocampo-Pineda, M., Schiavi, S., Rheault, F., Girard, G., Petit, L., Descoteaux, M., Daducci, A., 2021. Hierarchical microstructure informed tractography. Brain Connectivity 11 (2), 75–88.
Perrin, G., Descombes, X., Zerubia, J., 2005. Adaptive simulated annealing for energy minimization problem in a marked point process application. In: Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 3–17.
Pestilli, F., Yeatman, J.D., Rokem, A., Kay, K.N., Wandell, B.A., 2014. Evaluation and statistical inference for human connectomes. Nature Methods 11, 1058.
Reisertbib2013, M., Mader, I., Anastasopoulos, C., Weigel, M., Schnell, S., Kiselev, V., 2011. Global fiber reconstruction becomes practical. NeuroImage 54 (2), 955–962.
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.
Schiavi, S., Lu, P.-J., Weigel, M., Lutti, A., Jones, D.K., Kappos, L., Granziera, C., Daducci, A., 2022. Bundle myelin fraction (bmf) mapping of different white matter connections using microstructure informed tractography. NeuroImage 249, 118922.
Schiavi, S., Ocampo-Pineda, M., Barakovic, M., Petit, L., Descoteaux, M., Thiran, J.-P., Daducci, A., 2020. A new method for accurate in vivo mapping of human brain connections using microstructural and anatomical information. Science Advances 6.
Schilling, K.G., Rheault, F., Remedios, S., Pierpaoli, C., Anderson, A.W., Anderson, A.W., Landman, B.A., Descoteaux, M., 2020. Brain connections derived from diffusion MRI tractography can be highly anatomically accurate if we know where white matter pathways start, where they end, and where they do not go. Brain Structure and Function.
Serret, J., 1851. On some formulas relating ẗo the theory of à double curvature curves. Journal of Pure and Applied Mathematics 193–207.
Smith, R., Raffelt, D., Tournier, J.-D., Connelly, A., 2020. Quantitative streamlines tractography: methods and inter-subject normalisation. OSF Preprints.
Smith, R.E., Tournier, J.-D., Calamante, F., Connelly, A., 2012. Anatomically-constrained tractography: improved diffusion mri streamlines tractography through effective use of anatomical information. Neuroimage 62 (3), 1924–1938.
Smith, R.E., Tournier, J.-D., Calamante, F., Connelly, A., 2013. SIFT: Spherical-deconvolution informed filtering of tractograms. NeuroImage 67, 298–312.
Smith, R.E., 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.
Tournier, J.-D., Calamante, F., Connelly, A., 2010. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. Proc. Intl. Soc. Mag. Reson. Med. (ISMRM) 18.
Tournier, J.-D., Calamante, F., Connelly, A., 2012. Mrtrix: Diffusion tractography in crossing fiber regions. Int J Imaging Syst Technol 22 (1), 53–66.
Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., 2013. The wu-minn human connectome project: An overview. NeuroImage 80, 62–79.
Wasserthal, J., Neher, P.F., Maier-Hein, K.H., 2018. Tract orientation mapping for bundle-specific tractography. In: Frangi, Alejandro, F., Schnabel, J.A, Davatzikos, C., Alberola-López, C., Fichtinger, G. (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer International Publishing, pp. 36–44.
Wasserthal, J., Neher, P.F., Hirjak, D., Maier-Hein, K.H., 2019. Combined tract segmentation and orientation mapping for bundle-specific tractography. Medical Image Analysis 58, 101559.
Zalesky, A., Fornito, A., Cocchi, L., Gollo, L.L., van den Heuvel, M.P., Breakspear, M., 2016. Connectome sensitivity or specificity: which is more important? NeuroImage 142, 407–420.
Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. Noddi: Practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage 61 (4), 1000–1016.
Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. IEEE transactions on medical imaging 20 (1), 45–57.