



**HAL**  
open science

## Hierarchical microstructure informed tractography

Mario Ocampo-Pineda, Simona Schiavi, François Rheault, Gabriel Girard,  
Laurent Petit, Maxime Descoteaux, Alessandro Daducci

► **To cite this version:**

Mario Ocampo-Pineda, Simona Schiavi, François Rheault, Gabriel Girard, Laurent Petit, et al..  
Hierarchical microstructure informed tractography. *Brain connectivity*, 2021, 11 (2), pp.75-88.  
10.1089/brain.2020.0907 . hal-03225523

**HAL Id: hal-03225523**

**<https://hal.science/hal-03225523>**

Submitted on 12 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hierarchical microstructure informed tractography

Mario Ocampo-Pineda<sup>1</sup>, Simona Schiavi<sup>1</sup>, François Rheault<sup>2</sup>, Gabriel Girard<sup>3,4,5</sup>,  
Laurent Petit<sup>6</sup>, Maxime Descoteaux<sup>2</sup>, and Alessandro Daducci<sup>1</sup>

<sup>1</sup>Department of Computer Science, University of Verona, Italy

<sup>2</sup>Sherbrooke Connectivity Imaging Lab (SCIL), Université de Sherbrooke, Canada

<sup>3</sup>CIBM Center for BioMedical Imaging, Switzerland

<sup>4</sup>University Hospital Center (CHUV) and University of Lausanne (UNIL), Switzerland

<sup>5</sup>Signal Processing Lab (LTS5), École Polytechnique Fédérale de Lausanne, Switzerland

<sup>6</sup>Groupe d'Imagerie Neurofonctionnelle, Univ. Bordeaux, CNRS, CEA, IMN, UMR 5293,  
F-33000 Bordeaux, France

## Authors' contact information

Mario Ocampo-Pineda (mario.ocampo@univr.it)

Simona Schiavi (simona.schiavi@univr.it)

Alessandro Daducci (alessandro.daducci@univr.it)

Department of Computer Science, University of Verona,  
Strada le Grazie 15 - 37134 Verona, Italy.

François Rheault (francois.m.rheault@usherbrooke.ca)

Maxime Descoteaux (maxime.descoteaux@usherbrooke.ca)

Sherbrooke Connectivity Imaging Lab (SCIL), Université de Sherbrooke,  
2500 Boul. Université, Sherbrooke, Québec, J1K 2R1, Canada.

Gabriel Girard (gabriel.girard@epfl.ch)

Signal Processing Lab (LTS5), École Polytechnique Fédérale de Lausanne,  
EPFL-STI-IEL-LTS5

Station 11, CH-1015 Lausanne, Switzerland.

Laurent Petit (laurent.petit@u-bordeaux.fr)

Groupe d'Imagerie Neurofonctionnelle, Univ. Bordeaux, CNRS, CEA, IMN, UMR 5293

146 Rue Léo Saignat, 33000 Bordeaux, France.

**Short running title:** Hierarchical microstructure informed tractography

**Corresponding author**

Alessandro Daducci (alessandro.daducci@univr.it)

Department of Computer Science, University of Verona,

Strada le Grazie 15 - 37134 Verona, Italy.

+39 045 8027025

**Keywords:** Diffusion MRI, tractography, microstructure informed tractography, hierarchical organization.

## Abstract

**Background:** Tractography uses diffusion magnetic resonance imaging to noninvasively infer the macroscopic pathways of white matter fibers and it is the only available technique to probe in vivo the structural connectivity of the brain. However, despite this unique and compelling ability and its wide range of possible neurological applications, tractography is still limited, lacks anatomical precision, and suffers from a serious sensitivity/specificity trade-off. For this reason, in the past few years, tractography post-processing techniques have emerged and proved effective for improving the quality of the reconstructions. Among them, the Convex Optimization Modeling for Microstructure Informed Tractography formulation allows incorporating the anatomical prior that fibers are naturally organized in fascicles, and has obtained exceptional results in increasing the accuracy of the estimated tractograms.

**Methods:** We propose an extension to this idea and introduce a multi-level grouping of the streamlines in order to capture the white matter arrangement in fascicle and sub-fascicles. We tested our proposed formulation in synthetic and in vivo data.

**Results:** Our experiments show that using multiple levels allows considering information about the white matter organization more adequately and helps to improve further the accuracy of the resulting tractograms.

**Conclusion:** This new formulation represents a further important step towards a more accurate structural connectivity estimation.

## Impact Statement

Tractography is an invaluable tool for studying noninvasively the neuronal architecture of the brain, but recent studies have shown that the presence of a large number of false positives can significantly bias any connectivity analysis. Recently, a filtering technique called COMMIT2 has proven particularly effective in dramatically reducing their incidence by considering the prior knowledge that white matter fibers are organized in fascicles. In this work, we propose an extension to this method which allows us to increase further the anatomical accuracy of tractography reconstructions. Our new formulation represents an additional step forward towards a more veridical characterization of brain connectivity.

# 1 Introduction

Diffusion-weighted magnetic resonance imaging (DW-MRI) is a noninvasive imaging modality that provides information on the movements of water molecules within brain tissues (Basser and Pierpaoli, 1996; Le Bihan et al., 1986). This is particularly relevant in the white matter, as it is mostly composed of densely packed axons organized in fiber tracts called fascicles (Mandonnet et al., 2018) that hinder the motion of water molecules. *Tractography* algorithms are capable of inferring their macroscopic trajectories by exploiting the fact that diffusion occurs predominantly along the main direction of the fibers (Basser et al., 2000). In this context, a set of fibers is represented with a streamline, a set of streamlines with similar trajectories is commonly called bundle, and the set of all reconstructed streamlines is called tractogram. Tractography provides an invaluable tool in neuroscience research as it allows one to study *in vivo* and noninvasively the anatomy of the brain, its development as well as potential alterations due to neurodegenerative diseases along specific fiber tracts (Jeurissen et al., 2017; Yeh et al., 2020). It is possible to use the reconstructed streamlines as regions of interest (ROIs) and evaluate quantitatively all along their course additional information from other modalities, e.g., myelin content, which may provide better insight into the evolution of the pathology. Tractography makes it also possible to investigate *in vivo* the structural connectivity of the brain, i.e., the macroscopic map of the neuronal connections through the white matter, also known as connectome (Sporns et al., 2005; Hagmann, 2005). The connectome is an abstraction of the neural complexity of brain connections and can be represented as a network, where nodes correspond to gray matter regions and edges to the connections between them. With this formalism and using graph theory (Bullmore and Sporns, 2009), it is possible to investigate alterations of the connectivity due to neurological conditions (Griffa et al., 2013).

Despite these unique opportunities, several studies have recently exposed the *poor anatomical accuracy* of tractography showing, in particular, that the estimated tractograms are not truly quantitative and present problems such as partial, duplicate, and false-positive streamlines, i.e., pathways that are recovered but are not anatomically valid (Thomas et al., 2014; Drakesmith et al., 2015; Zalesky et al., 2016; Maier-Hein et al., 2017; Schilling et al., 2019; Girard et al., 2020; Rheault et al., 2020b). In particular, Zalesky et al. (2016) investigated the topological properties

of the networks estimated with tractography and concluded that specificity (true-negative rate) is more important than sensitivity (true-positive rate) for the study of brain connectivity. Maier-Hein et al. (2017) conducted a large comparison study and showed that false-positive streamlines are a major problem that affects all tractography algorithms. Therefore, it is clear that the presence of false positives in the estimated tractograms represents a crucial limitation of tractography that can heavily bias any analysis based on this technique.

A number of post-processing solutions have been recently proposed to tackle this problem and *improve the quality of the tractograms*. This class of methods is called *Microstructure Informed Tractography* (Daducci et al., 2016) and their rationale is that tractography needs more information to overcome the problem of anatomical inaccuracy. In general, these approaches combine the input tractogram with signal forward models and evaluate the contribution of each streamline to the acquired data, filtering out the most implausible or those that are not supported by the data. The Spherical-deconvolution Informed Filtering of Tractograms (SIFT) (Smith et al., 2013) considers the integral of the fiber orientation distribution to be proportional to the volume of the white matter tissue inside the voxel and uses this value to optimize the streamline density and to remove streamlines that do not fit the data. SIFT2 (Smith et al., 2015) does not remove streamlines, but it rather attempts to determine the cross-sectional area of each streamline and uses this information as a multiplier in the streamline connectivity quantification problem. The Linear Fascicle Evaluation (LiFE) (Pestilli et al., 2014) and the Convex Optimization Modeling for Microstructure Informed Tractography (COMMIT) (Daducci et al., 2013, 2015) use similar approaches, as both estimate the contributions of the streamlines from the full DW-MR signal by solving a linear system. The filtering with all these methods showed promising results for improving the quality of the tractograms (Smith et al., 2013, 2015; Pestilli et al., 2014; Daducci et al., 2013, 2015), but none of them proved effective in reducing the false positives (Schiavi et al., 2020).

To tackle this limitation, COMMIT was recently extended to take into account the fundamental knowledge that *fibers in the brain are naturally organized in fascicles*. In fact, developmental neurobiology studies have demonstrated that “axons, as they grow, remain neighbors with axons that come from neighboring neurons in the presynaptic structure, and therefore arrive at the postsynaptic structure already topographically arranged” (Udin and Fawcett, 1988). The novel formulation, called COMMIT2 (Schiavi et al., 2020), proved effective in increasing the specificity

of the tractograms without decreasing their sensitivity. However, if both true- and false-positive streamlines are reconstructed between two cortical regions, as illustrated in Figure 1 (top), then COMMIT2 would promote or penalize them as a whole and would not be able to disentangle them.

In this work, *we propose a strategy to further improve the accuracy of the tractograms* by extending the assumption made in COMMIT2 and considering that fibers in the brain are anatomically organized in fascicles and sub-fascicles, e.g., the pyramidal tract is formed by the corticospinal tract and the corticobulbar tract (Nieuwenhuys et al., 2008; Chenot et al., 2019). This biological organization of the fibers naturally induces a hierarchy in the reconstructed streamlines which, in turn, could be exploited to provide more information to the filtering procedure. To test the effectiveness of this additional anatomical prior, we created a nested arrangement of the streamlines into bundles and sub-bundles by refining the parcellation-based approach of COMMIT2 using clustering techniques. The proposed method was tested both on synthetic and in vivo data.

## 2 Materials and methods

### 2.1 Microstructure informed tractography

Given a tractogram  $\mathcal{T}$  and the corresponding DW-MR image  $\mathbf{I} \in \mathbb{R}_+^{n_v \times n_d}$  with  $n_d$  q-space samples acquired in  $n_v = n_x n_y n_z$  voxels, the acquired data can be seen as  $\mathbf{I} = \mathcal{A}(\mathcal{T}) + \eta$ , where  $\mathcal{A} : \mathcal{T} \rightarrow \mathbf{I}$  is an operator describing the signal contribution of each streamline to  $\mathbf{I}$  and  $\eta$  is the acquisition noise. The goal of tractography is to solve the inverse problem, i.e., finding the set of streamlines  $\tilde{\mathcal{T}}$  that best describe the acquired image  $\mathbf{I}$ . *Microstructure informed tractography* has been recently proposed as a new paradigm to combine the estimated streamlines with tissue microstructure models with the aim of improving the quality of the tractograms (Daducci et al., 2016). In particular, the Convex Optimization Modeling for Microstructure Informed Tractography (COMMIT) allows combining an arbitrary biophysical model for the white matter tissue (see (Alexander et al., 2019; Novikov et al., 2019) and references therein) to each streamline in order to estimate their actual contribution. The DW-MR signal in all voxels is expressed as the linear combination of the contributions from the streamlines that cross that voxel, possibly in addition to

partial volume with other tissues, e.g., cerebrospinal fluid, as follows:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \eta, \quad (1)$$

where the vector  $\mathbf{y} \in \mathbb{R}_+^{n_d n_v}$  is the concatenation of the  $n_d$  DW-MRI measurements acquired in the  $n_v$  voxels of  $\mathbf{I}$ , the matrix  $\mathbf{A} \in \mathbb{R}^{n_d n_v \times n_c}$  encodes  $n_c$  compartments with the signal associated to all the streamlines in  $\mathcal{T}$  and possibly other tissues inside each voxel according to a given multi-compartment model and  $\eta$  accounts for both acquisition noise and modeling errors. The unknowns  $\mathbf{x} \in \mathbb{R}_+^{n_c}$  represent the actual contributions of all the compartments encoded in  $\mathbf{A}$ , which are needed to explain the acquired data  $\mathbf{I}$ , and can be estimated using non-negative least squares:

$$\operatorname{argmin}_{\mathbf{x} \geq 0} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2.$$

To address the problem of false positives, COMMIT was recently extended by Schiavi et al. (2020), who added a regularization term to take into account that fibers in the brain are naturally organized in fascicles and the streamlines can be grouped in bundles:

$$\operatorname{argmin}_{\mathbf{x} \geq 0} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \sum_{g \in \mathcal{G}} \omega^{(g)} \|\mathbf{x}^{(g)}\|_2. \quad (2)$$

In this formulation, called COMMIT2,  $\mathcal{G}$  represents a partition of the streamlines into groups and  $\mathbf{x}^{(g)}$  are the coefficients of those belonging to group  $g \in \mathcal{G}$ . This additional penalty term promotes solutions that explain the measured DW-MR signal using a minimal number of groups and the parameter  $\lambda \geq 0$  controls the trade-off between the data and the regularization term. As suggested in (Wang and Leng, 2008), each component of the regularization term was scaled by the following factor to cope with the possible bias introduced by groups with different cardinality:

$$\omega^{(g)} = \frac{\sqrt{|g|}}{\|\mathbf{x}_0^{(g)}\|_2}, \quad (3)$$

where  $|g|$  is the cardinality of the group  $g$  and  $\mathbf{x}_0^{(g)}$  are the corresponding contributions estimated without regularization, i.e.,  $\lambda = 0$ .



## 2.2 Adding priors on the hierarchical organization of the fibers

In the previous formulation, the partition  $\mathcal{G}$  was created using a parcellation-based approach, i.e., grouping together all the streamlines connecting the same pair of cortical regions. However, as Eq. (2) promotes or penalizes each group of streamlines as a whole, if a group contains both true- and false-positive streamlines, as shown in Figure 1 (top), they will be promoted or penalized together. In Figure 1 (bottom) we present an example of an actual bundle of streamlines as reconstructed with tractography from in vivo data where this situation happens. This bundle consists of two distinct sub-bundles: the streamlines that directly connect the brainstem to the postcentral gyrus, which are known to be correct from anatomy, and those that incorrectly pass through the corpus callosum. Applying COMMIT2 using a gray matter parcellation as in (Schiavi et al., 2020) results in a single group for this bundle which does not allow disentangling the true-positive streamlines from the false positives. If, instead, smaller groups are used, e.g., obtained from a finer parcellation or a clustering procedure, some invalid structures are lost but at the price of losing also some of the true-positive streamlines that were connected to the inferior part of the postcentral gyrus (see below for a discussion on this topic). However, if we combine the previous groups in a hierarchical way, the false-positive streamlines can be removed and more true positives can be kept. Inspired by this example, we propose to modify COMMIT2 and make it more flexible, allowing the possibility to *group the streamlines into multiple levels* in order to capture the natural hierarchical organization of the fibers in fascicles and sub-fascicles (Chenot et al., 2019).

To take into account this organization of the streamlines, we implemented the *hierarchical sparse encoding* approach proposed by Jenatton et al. (2011), which uses sparsity-inducing norms as regularization terms to promote tree-structured sparse solutions. The set of groups  $\mathcal{G}$  is defined as a tree-structured set of groups with  $n$  levels, where each level  $d$  represents a partition of all the streamlines. The root  $\mathcal{G}_0$  contains all the streamlines of the tractogram in one group and the partitions in the lower levels called  $\mathcal{G}_1, \dots, \mathcal{G}_n$  are created by splitting each group  $g_{d-1} \in \mathcal{G}_{d-1}$ , creating a nested structure. The reconstruction problem can be written as:

$$\underset{\mathbf{x} \geq 0}{\operatorname{argmin}} \|\mathbf{Ax} - \mathbf{y}\|_2^2 + \lambda \sum_{d=1}^n \sum_{g_d \in \mathcal{G}_d} \omega^{(g_d)} \|\mathbf{x}^{(g_d)}\|_2, \quad (4)$$

where  $\mathcal{G}_d$  represents a partition into groups of level  $d$ ,  $\mathbf{x}^{(g_d)}$  are the coefficients of the streamlines in group  $g_d \in \mathcal{G}_d$  and  $\lambda \geq 0$  controls the regularization strength. Note that if  $n = 1$  then Eq. (4) reduces to Eq. (2). The scaling factor to balance groups with different cardinality is modified as follows:

$$\omega^{(g_d)} = \frac{\sqrt{|g_d|}}{\|\mathbf{x}_0^{(g_d)}\|_2}, \quad (5)$$

where  $|g_d|$  is the cardinality of the group  $g_d$  and  $\mathbf{x}_0^{(g_d)}$  are the contributions of the streamlines that belong to the group  $g_d$  obtained as described above. The regularization in Eq. (4) is also called *Hierarchical Group Lasso* and was introduced by Zhao et al. (2009). Different implementations exist of this formulation, and for our application we implemented the one proposed by Jenatton et al. (2011). The main advantage of this formulation is that its complexity is linear in the number of streamlines and levels, so increasing the number of levels does not significantly affect the computation time (for further details we refer the reader to the mathematical proofs contained in (Jenatton et al., 2011)). In the remaining of the manuscript, we will refer to this approach as COMMIT2<sub>tree</sub>. Figure 2 visually compares the differences between COMMIT, COMMIT2, and COMMIT2<sub>tree</sub> formulations.

**Illustrative toy example.** Figure 3 shows a simple synthetic example to illustrate how the introduction of the hierarchical regularization can improve the accuracy of the tractograms. The ground truth consists of four regions of interest (A, B, C, and D) connected by three streamlines (S1, S2, and S3). From this dataset, a generic tractography algorithm (top row, center) reconstructed the three true-positive streamlines as well as two false positives (S4 and S5). This tractogram was then filtered with COMMIT, COMMIT2 and COMMIT2<sub>tree</sub>. Without any prior information, COMMIT keeps all five streamlines (top row, right), as they are all compatible with the underlying data. COMMIT2 groups streamlines based on a partition with a single level and attempts to find solutions that explain the data with the minimum number of groups. If large groups are used (bottom row, left), e.g., streamlines that connect the same pair of ROIs, the group formed by S3 and S4 is necessary to explain the data in the top-right voxel, which is covered only by S3. Therefore both S3 and S4 are kept even though S4 is a false positive. This choice, in turn, requires to keep also S5 to explain the data in the remaining voxels using the minimum number of groups.

This causes the removal of the true positives S1 and S2 because they are no longer necessary to fit the data. In general, if small groups are used (bottom row, center), e.g., dividing B in two sub-regions B1 and B2, and C into C1 and C2, we obtain more groups consisting of fewer streamlines each; in this example every group has a single streamline. This situation contains many ambiguous configurations as saw before for COMMIT (top row, right) and, as groups are treated as independent entities, these ambiguities cannot be solved. Please note that this situation was already analyzed in (Schiavi et al., 2020) (see Figure 3). On the contrary, the grouping implemented in COMMIT<sub>tree</sub> is more flexible and helps to disentangle ambiguous configurations such the one presented here. In fact, according to the hierarchical structure, the group S3-S4 can be further subdivided into two groups (bottom row, right), i.e., one containing the true positive S3 and the other one the false positive S4. This allows COMMIT<sub>tree</sub> to remove all the false positives and keep the true positives.

## 2.3 Data and experiments

For the scope of this work and to keep the presentation simple, we demonstrate the improvements of COMMIT<sub>tree</sub> over the previous formulation using a hierarchical structure with only two levels; however, our proposed formulation can easily accommodate any number of levels. In the first level, we grouped the streamlines connecting the same pair of cortical regions as in COMMIT2. Then, to capture the presence of possible sub-bundles, we constructed a second level where we subdivided these parcellation-based groups by clustering the streamlines in each with QuickBundles (Garyfallidis et al., 2012). In brief, QuickBundles is an unsupervised clustering algorithm for tractography datasets which does not require to know the number of clusters in advance, but it rather uses a threshold to group streamlines based on distance metrics. In this work we used the Average of Pointwise Euclidean Metric. The algorithm is very fast and well suited for large tractograms, as its complexity is  $O(kN)$ , where  $k$  is the number of clusters and  $N$  the number of streamlines (for technical details see (Garyfallidis et al., 2012)). For a more direct comparison with COMMIT2, we adopted the same forward model as in (Schiavi et al., 2020), which associates a contribution to each streamline proportional to its cross-sectional area using a map of the intra-axonal signal fraction (IASF).

### 2.3.1 Quantitative evaluation on synthetic data

To quantitatively evaluate the performance of our proposed formulation with respect to COMMIT2, we adopted the same synthetic phantom used in (Schiavi et al., 2020). We compared the *sensitivity and specificity of the tractograms* after filtering with both techniques by assessing the number of valid bundles (VB), i.e., true-positive connections in the corresponding connectomes, and invalid bundles (IB), i.e., false-positive ones. The phantom is shown in Figure 4A and consists of 27 ground-truth bundles between 53 ROIs that mimic challenging configurations like branching, kissing and crossing with different curvatures and sizes (Caruyer et al., 2014). The IASF map of this phantom was computed from the ground-truth geometry. The corresponding DW-MR signal was generated using the Composite Hindered And Restricted Model of Diffusion (Assaf and Basser, 2005) along 64 directions with  $b = 3000 \text{ s/mm}^2$  and Rician noise was added with a signal-to-noise ratio of 30. All tractograms and the corresponding connectomes were estimated using the MRtrix3 software (Tournier et al., 2019). First, we computed the fiber orientation distributions in each voxel using Constrained Spherical Deconvolution (Tournier et al., 2007). Then, we reconstructed 1 million streamlines with both deterministic (Tournier et al., 2012) and probabilistic (Tournier et al., 2010) algorithms, using a white matter mask as seeding region and default parameters. Finally, we assigned each endpoint of a streamline to a node if that point fell within 2 mm from one of the 53 gray matter ROIs. A streamline was considered as connecting two nodes if both endpoints were assigned, otherwise it was discarded from the analysis.

### 2.3.2 Qualitative evaluation on in vivo data

To appreciate the improved quality of the tractograms, we visually inspected *known true- and false-positive bundles* in 10 subjects (5 males and 5 females of ages 26–30) of the Human Connectome Project (HCP) (Van Essen et al., 2013). These bundles were extracted with RecobundlesX (Garyfallidis et al., 2018; Rheault, 2020), a multi-atlas, multi-parameter segmentation tool that uses bundles manually segmented by an expert neuroanatomist as reference (description in Appendix A). We performed RecobundlesX using nine different parameter configurations and a threshold for the voting system of 75%. If a streamline had been selected as true positive, it was not taken into account in the segmentation of the false-positive bundles.

The intersections of streamlines in different bundles were eliminated, that is, each streamline could only belong to one bundle. We also reported the voxel coverage of each of them, i.e., the number of voxels traversed by the streamlines associated with each bundle. We segmented the T1-weighted image using the FMRIB's Automated Segmentation Tool (FAST) (Zhang et al., 2001) to derive the multi-tissue image and perform the multi-shell, multi-tissue constrained spherical deconvolution (Jeurissen et al., 2014) on the preprocessed DW-MRI data (Glasser et al., 2013). We generated 5 million streamlines of length between 20 mm and 200 mm with the probabilistic algorithm (Tournier et al., 2010), using Anatomically Constrained Tractography (Smith et al., 2012) and default parameters. As in the synthetic experiments, the white matter mask was used as seeding region. The IASF map needed by the forward model used in COMMIT2 and COMMIT2<sub>tree</sub> was computed using the Spherical Mean Technique (Kaden et al., 2016). Besides inspecting individual bundles, we also compared the resulting connectomes, which were created using the 85 gray matter ROIs defined in the Desikan-Killiany atlas (Desikan et al., 2006) derived from the T1-weighted images using FreeSurfer (<http://surfer.nmr.mgh.harvard.edu/>), and replacing the brainstem with its last part, i.e., medulla oblongata (Iglesias et al., 2015). Finally, as edge weights we used the streamline count for the raw tractograms and the sum of the estimated streamline contributions for COMMIT2 and COMMIT2<sub>tree</sub>. For each connectome, we report its density, i.e., the ratio between the actual and the possible connections.

### 3 Results

Figure 4 compares the *sensitivity and specificity of the tractograms* estimated from the synthetic data with both deterministic and probabilistic tracking, before and after filtering with COMMIT2 and COMMIT2<sub>tree</sub>. As a reference, results with COMMIT are also reported. All reconstructions included the 27 true-positive bundles and, as expected, the raw tractograms were contaminated by a large number of false-positive bundles, respectively IB=441 (probabilistic) and IB=235 (deterministic). Only few IB were removed with COMMIT, whereas COMMIT2 was able to drastically reduce them from 441 to 20 (probabilistic) and from 235 to 17 (deterministic). Using the proposed multi-level regularization, COMMIT2<sub>tree</sub> was able to further improve these results, although not substantially. Using different distance metrics in the clustering did not produce particular effects on the results

(see Supplementary Figure S1), probably due to the simple geometry of this phantom.

Figure 5 evaluates the effectiveness of the multi-level formulation in real brain data, where the fiber configurations are expected to be more complex. We *qualitatively compared known true- and false-positive bundles* obtained in one HCP subject after filtering the tractograms with COMMIT2 and COMMIT2<sub>tree</sub>. The volume coverage of the bundles is reported in terms of number of voxels and, in the last two cases, we also report the estimated contributions (weight). The filtering with COMMIT2 does not negatively affect the *true-positive bundles*, i.e., their volume is comparable with the raw tractogram, and, as expected, the *false-positive bundles* are sensibly thinned, i.e., reduced volume coverage. The introduction of the additional level in the regularization term of COMMIT2<sub>tree</sub> *further improved these performances*. On the one hand, the voxel coverage of the true-positive bundles is comparable but the estimated contributions are larger, which translates into a higher likelihood to be true positives. On the other hand, the contributions of the false positives are further reduced and all of them are almost completely removed. These results were consistent across all the ten subjects (Table 1).

To better appreciate the improved anatomical accuracy of the tractograms, in Figure 6 we *visually inspect two representative bundles* as reconstructed in each of the ten subjects. They correspond to the streamlines connecting the medulla oblongata to, respectively, the right precentral gyrus (Bundle 1) and the left precentral gyrus (Bundle 2). As can be seen, these bundles have streamlines that are coherent and well organized in two sub-bundles, and one of these describes an implausible pathway passing through the corpus callosum. The streamlines in the raw tractogram are colored based on their local orientation, whereas for COMMIT2 and COMMIT2<sub>tree</sub> they are colored according to the estimated sum of the streamline contributions in each sub-bundle. To make this visual comparison more effective, we use two different colormaps, *cool* for the true-positive streamlines and *hot* for the false positives. As expected, since COMMIT2 treats all the streamlines in each bundle as a whole, it is not able to completely remove any false-positive sub-bundles that are included in such bundles. However, if sub-groups of streamlines are considered in the filtering as done in COMMIT2<sub>tree</sub>, it is possible to assign low contributions to these false positives, or even remove them completely. For all the subjects, there is an evident improvement using COMMIT2<sub>tree</sub>: lower values are assigned to all the false positives as compared to COMMIT2, e.g., yellow and green arrows, and in some subjects it was able to completely

remove them. Moreover, COMMIT2<sub>tree</sub> assigned higher contributions than COMMIT2 to all the true positives. We can also appreciate that the patterns of the streamlines of these two bundles are consistent across subjects. Table 2 reports the mean and standard deviation among the ten subjects of the sum of the streamline contributions and the voxel coverage for each sub-bundle.

Figure 7 compares the estimated connectomes in one HCP subject before and after filtering the tractogram reconstructed with probabilistic tractography with COMMIT2 and COMMIT2<sub>tree</sub>. To assess how well each filtered tractogram explains the input data, we report the input IASF map used for the filtering and the one predicted from the filtered tractograms, using Eq. (1), as well as the root mean square error (RMSE) between the input and the fitted maps. We observe that the connectome of the raw tractogram is very dense (density=0.896), which is an expected result for probabilistic tracking, whereas after the filtering with COMMIT2 and COMMIT2<sub>tree</sub>, the connectomes are sparser and have comparable density between them (0.245 and 0.307, respectively). Despite sparser, the corresponding tractograms seem to explain well the input IASF map; the RMSE map is comparable and shows very small errors in the white matter, with higher values located in voxels at the interface with cortical and subcortical gray matter.

## 4 Discussion

Tractography is a unique technique that is able to describe the major neural pathways in the white matter. In the last decades, it has been used to study the structural network organization of the brain, but its anatomical inaccuracy and limitations have been exposed in different studies (Thomas et al., 2014; Drakesmith et al., 2015; Zalesky et al., 2016; Maier-Hein et al., 2017; Schilling et al., 2019; Girard et al., 2020; Rheault et al., 2020b). One of the major limitations still unresolved is the intrinsic trade-off between specificity and sensitivity. In a recent international challenge (Maier-Hein et al., 2017), it was suggested that a possible way to overcome this issue would be to inject microstructural and anatomical prior during tractography reconstruction. A step forward in this direction was achieved by the recent work of Schiavi et al. (2020), where microstructure informed tractography was combined with the anatomical prior that white matter fibers (axons) are organized in fascicles giving origin to COMMIT2. Using the prior knowledge that bundles are organized in sub-bundles (Chenot et al., 2019; Mandonnet et al., 2018), in this work

we extended COMMIT2 proposing the use of a multilevel organization of the streamlines in order to identify streamlines following implausible trajectories inside valid bundles. This new formulation, called COMMIT2<sub>tree</sub>, considers partitions of the streamlines with different granularity and organizes them in a hierarchical structure which provides more degrees of freedom than COMMIT2 but in a constrained clever way. The sub-groups used in COMMIT2<sub>tree</sub> share anatomical constraints like the starting and ending region in the gray matter, but they have different geometrical properties, i.e., the trajectory. Figure 6 clearly shows that by considering this way of grouping in the optimization, more false-positive streamlines inside true-positive bundles are removed, boosting further the accuracy of the final tractogram.

In synthetic data, the connectivity graphs obtained after filtering the tractograms with COMMIT2 and COMMIT2<sub>tree</sub> showed comparable results. However, the simple geometry of this phantom does not allow us to fully exploit the advantages of our proposed multi-level regularization, as the bundles consist of regular and smooth tubes. Nonetheless, our findings indicate that this enhanced formulation results in a *stable and robust* regularization of the original problem. In particular, Figure 4 shows that adding more levels to the structure does not affect the stability of the method, but can slightly improve the quality of the estimated tractograms.

To evaluate the quality of the in vivo reconstructions, we analyzed ten HCP subjects and we looked both at the anatomical accuracy of individual bundles and at the overall structural connectivity estimated after filtering the tractograms with COMMIT2 and COMMIT2<sub>tree</sub>. The density of the filtered connectomes was about one third of the unfiltered one, showing again similar performances of both methods. These results confirm and extend the results presented in (Schiavi et al., 2020), which were in agreement with the theory of the economy of brain networks (Bullmore and Sporns, 2012). However, when looking more closely to the individual bundles, before and after the filtering, it is evident that a significant improvement was achieved by COMMIT2<sub>tree</sub>. In fact, in Figure 5 and Table 1 we can appreciate how COMMIT2<sub>tree</sub> was able to filter more the false-positive bundles than COMMIT2 without negatively affecting the true positives, in particular because of its capability to remove invalid streamlines inside valid bundles (Figure 6 and Table 2). From a quantitative point of view, we observe that both methodologies assigned lower contributions to the false positives than the true positives; notably, it is important to note that COMMIT2<sub>tree</sub> assigned *higher* contributions to the true positives and *smaller* values to the false positives as compared to



its predecessor. This is an important property when one is interested in network properties relying on connections strengths.

Although COMMIT2<sub>tree</sub> was able to further improve the anatomical accuracy of the tractograms, it did not eliminate completely all known false positives in every analyzed subject. This may be partly due to the regularization parameter  $\lambda$  used and to the actual groups created with clustering in each subject. First, in our experiments we used the same regularization strength for all subjects, which was empirically determined as the value that overall produced visually satisfactory results. However, optimizing the value of  $\lambda$  on a subject-specific basis might considerably improve the performance of the filtering. Second, the results presented here show that the quality of the anatomical prior, i.e, structure definition and how streamlines are grouped, plays a fundamental role in the optimization procedure. The original COMMIT was only data driven and the filtering did not improve significantly the anatomical accuracy of the tractograms; COMMIT2 added the prior that white matter fibers are organized in fascicles which, instead, had a big impact on the performance. In COMMIT2<sub>tree</sub> we added geometrical priors based on the trajectories described by the streamlines and we built a hierarchical structure to combine both *anatomical* (based on which gray matter regions the streamlines connect) and *geometrical* (clustering the streamlines based on their trajectory) information. However, different definitions of structures can be implemented in our formulation. For example we could use more than two levels to create the hierarchy as well as explore different strategies to group the streamlines and inject additional prior information, e.g., see (Côté et al., 2015; Siless et al., 2018) and references therein. All these aspects will be the subject of future research.

It is worth noting that using a finer partitioning of the streamlines may create groups that might capture better the plausible and implausible sub-bundles; for instance, in the extreme case where all groups have a cardinality equal to 1, i.e., they consist of a single streamline, true- and false-positive streamlines are clearly separated in distinct groups. However, as the granularity of this partitioning becomes smaller, the number of generated groups increases accordingly. But since the groups are treated as independent entities in the penalty term in Eq. (4), this process has the side effect of gradually reintroducing the ambiguous configurations that, instead, COMMIT2 was able to disentangle using a coarser partition. Supplementary Figure S2 shows that, indeed, the performances gradually deteriorate as the groups become smaller. As it is not possible to predict a

priori the behavior of the algorithm using partitions with different granularity, our proposed approach represents an effective way to consider and take advantage of more than one granularity in the filtering procedure. Clearly, future research will be required to identify the optimal configuration of groups in this multi-level structure.

Finally, it is important to note that the proposed approach is still a filtering procedure, so even if it can be applied on any type of input tractograms (coming from either deterministic or probabilistic tractography approaches), it is still limited on the quality of the input reconstruction. Clearly, if a bundle is not present in the original tractogram it will not be recovered by COMMIT2<sub>tree</sub>. So, particular attention should be given to this aspect. Another important factor that affects the results is the selection of the regularization parameter  $\lambda$ . As it was analyzed in (Schiavi et al., 2020), a high value of this parameter would result in a very strong filtering and would end up in removing too many connections. On the contrary, a too low value would end up in removing only a few. Up until now, and for both COMMIT2 and COMMIT2<sub>tree</sub> formulations, we do not know a close formula to define the best value of this parameter according to the input tractogram, signal and structure for grouping the streamlines. Thresholds for upper and lower bounds when using the adaptive Group Lasso approach like ours have been suggested for the case of independent and identically distributed columns of the matrix  $\mathbf{A}$ , but they still need to be generalized to our case. In the experiments presented in this work the choice was made based on the knowledge we have on known bundles: i.e., we choose the largest value that was able to filter out the known invalid bundles from the input tractogram while keeping all the known valid ones.

## 5 Conclusion

The anatomical accuracy of tractography has been heavily challenged in the past few years and seriously questioned the use of this technique for mapping reliably the structural connectivity of the brain. Our group recently developed a post-processing filtering procedure called COMMIT2 that allowed a dramatic increase in the accuracy of the reconstructions. However, this formulation did not consider the fact that fibers in the brain are anatomically organized in fascicles and sub-fascicles. In this work, we showed that taking into account this multilevel organization of the fibers allows considering information about the white matter structure in a more specific

way. This improved formulation, i.e., COMMIT2<sub>tree</sub>, not only reduces the number of false-positive connections in the estimated tractograms as COMMIT2 but it is also able to filter out implausible streamlines within true-positive bundles, boosting even further the quality of the reconstructions. Our new formulation represents an additional step forward to improve the anatomical accuracy of the tractograms and our understanding of how different brain regions are interconnected.

**Code and data availability.** The numerical phantom used as validation is publicly available and can be downloaded from <https://github.com/ecaruyer/phantomas>. The in vivo MRI data used are from 5 males and 5 females subjects of ages 26–30 years from the Open Access Data from the HCP and are available at <https://www.humanconnectome.org>. The code is open source and freely available at <https://github.com/daducci/COMMIT>.

**Authorship Confirmation.** All the authors, Mario Ocampo-Pineda, Simona Schiavi, François Rheault, Gabriel Girard, Laurent Petit, Maxime Descoteaux, and Alessandro Daducci contributed significantly to the work. M.O.-P., S.S., and A.D. conceptualized the problem. M.O.-P., S.S., G.G., and A.D. developed, implemented, and tested the technical framework. L.P. manually segmented the in vivo VB and IB toward which we compared our method. M.O.-P., F.R., and M.D. conceptualized and developed the in vivo validation framework. M.O.-P., S.S., F.R., G.G., L.P., M.D., and A.D. wrote the manuscript.

**Authors' Disclosure.** The authors have nothing to disclose.

**Funding.** This work was supported by the Rita Levi Montalcini Programme for young researchers of the Italian Ministry of Education, University and Research.

**Acknowledgments.** Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

## Appendix A RecobundlesX

To study and evaluate tractography reconstructions, a common practice is to divide them into bundles that represent white matter fascicles. This is a complicated task, called virtual dissection, that requires great knowledge of Neuroanatomy and is not an exact procedure, there is variability among segmentations performed by experts (Rheault et al., 2020a), and up to now, there is not a ground truth for in vivo tractography to validate the reconstructions.

RecobundlesX (Rheault, 2020) uses the method Recobundles (Garyfallidis et al., 2018) that is an algorithm that uses bundle models as shape priors for detecting and segment similar streamlines in a tractogram, that is computationally demanding for datasets of millions of streamlines. To reduce the dimension of the problem of comparing all the streamlines in the models (hundreds) with all the streamlines in a tractogram (millions) to select the streamlines, Recobundles uses the algorithm QuickBundles (Garyfallidis et al., 2012) to cluster the streamlines and uses the Streamline-based Linear Registration (SLR) (Garyfallidis et al., 2015) to performs the comparison among centroids at global and local level. This procedure has a big variability in the results, generated by the clustering configurations used, in addition to the registration of the models and the tractogram.

To reduce this variability, RecobundlesX (Rheault, 2020) takes into account multi-atlas models, multi-parameters clustering, and a voting scheme that generates more stable solutions. The segmentation method supports to use multiple models of the same bundle as input. These models can be generated from different subjects, with different tractography algorithms or segmented by different experts, to cover more variability in the definition of the bundles. It also allows us to use different configurations of the segmentation parameters, e.g., clustering threshold, number of iterations for SLR, type of transformation, pruning, etc, to cover more variability in the segmentation procedure. RecobundlesX combines the results of the multi-atlas and multi-parameters segmentation using a voting system, that merges the multiple segmentation performed for each bundle definition.

**True-positive bundle models.** Five subjects from the Open Access Data from the Human Connectome Project (Van Essen et al., 2013) were used. Tractography reconstruction was

performed and 37 major bundles were manually segmented. Then, these bundles were smoothed and refined to obtain well defined bundle models (Rheault, 2020).

**False-positive bundle models.** The submitted tractograms to the ISMRM 2015 Tractography Challenge were used. In the study performed by Maier-Hein et al. (2017) about the challenge, several bundles from the submitted tractograms were labeled as false positives because they were not in the ground-truth tractogram. A quality control was performed on these bundles to select those with the most anatomical implausible trajectory. From these bundles, 44 were selected and refined to obtain smooth and well defined bundle models (Rheault, 2020).

## References

- Alexander, D. C., Dyrby, T. B., Nilsson, M., et al. (2019). Imaging brain microstructure with diffusion MRI: practicality and applications. *NMR in Biomedicine*, 32(4):e3841.
- Assaf, Y. and Basser, P. J. (2005). Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage*, 27(1):48 – 58.
- Basser, P. J., Pajevic, S., Pierpaoli, C., et al. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44(4):625–632.
- Basser, P. J. and Pierpaoli, C. (1996). Microstructural and Physiological Features of Tissues Elucidated by Quantitative-Diffusion-Tensor MRI. *Journal of Magnetic Resonance, Series B*, 111(3):209 – 219.
- Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198.
- Bullmore, E. and Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, 13(5):336–349.
- Caruyer, E., Daducci, A., Descoteaux, M., et al. (2014). Phantoms: a flexible software library to simulate diffusion MR phantoms. In *ISMRM 22nd Annual Meeting*, volume 2666.

- Chenot, Q., Tzourio-Mazoyer, N., Rheault, F., et al. (2019). A population-based atlas of the human pyramidal tract in 410 healthy participants. Brain Structure and Function, 224(2):599–612.
- Côté, M.-A., Garyfallidis, E., Laroche, H., et al. (2015). Cleaning up the mess: tractography outlier removal using hierarchical QuickBundles clustering. In ISMRM 2015.
- Daducci, A., Dal Palú, A., Descoteaux, M., et al. (2016). Microstructure Informed Tractography: Pitfalls and Open Challenges. Frontiers in Neuroscience, 10:247.
- Daducci, A., Dal Palù, A., Lemkaddem, A., et al. (2015). COMMIT: Convex Optimization Modeling for Microstructure Informed Tractography. IEEE Transactions on Medical Imaging, 34(1):246–257.
- Daducci, A., Palù, A. D., Lemkaddem, A., et al. (2013). A convex optimization framework for global tractography. In 2013 IEEE 10th International Symposium on Biomedical Imaging, pages 524–527.
- Desikan, R. S., Ségonne, F., Fischl, B., 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.
- Drakesmith, M., Caeyenberghs, K., Dutt, A., et al. (2015). Overcoming the effects of false positives and threshold bias in graph theoretical analyses of neuroimaging data. NeuroImage, 118:313 – 333.
- Garyfallidis, E., Brett, M., Correia, M. M., et al. (2012). QuickBundles, a Method for Tractography Simplification. Frontiers in Neuroscience, 6:175.
- Garyfallidis, E., Côté, M.-A., Rheault, F., et al. (2018). Recognition of white matter bundles using local and global streamline-based registration and clustering. NeuroImage, 170:283 – 295. Segmenting the Brain.
- Garyfallidis, E., Ocegueda, O., Wassermann, D., et al. (2015). Robust and efficient linear registration of white-matter fascicles in the space of streamlines. NeuroImage, 117:124 – 140.

- Girard, G., Caminiti, R., Battaglia-Mayer, A., et al. (2020). On the cortical connectivity in the macaque brain: A comparison of diffusion tractography and histological tracing data. NeuroImage, 221:117201.
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., et al. (2013). The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage, 80:105 – 124. Mapping the Connectome.
- Griffa, A., Baumann, P. S., Thiran, J.-P., et al. (2013). Structural connectomics in brain diseases. NeuroImage, 80:515 – 526. Mapping the Connectome.
- Hagmann, P. (2005). From diffusion mri to brain connectomics. page 141.
- Iglesias, J. E., Leemput, K. V., Bhatt, P., et al. (2015). Bayesian segmentation of brainstem structures in MRI. NeuroImage, 113:184 – 195.
- Jenatton, R., Mairal, J., Obozinski, G., et al. (2011). Proximal Methods for Hierarchical Sparse Coding. Journal of Machine Learning Research, 12:2297–2334.
- Jeurissen, B., Descoteaux, M., Mori, S., et al. (2017). Diffusion MRI fiber tractography of the brain. NMR in Biomedicine. e3785 NBM-17-0045.R2.
- Jeurissen, B., Tournier, J.-D., Dhollander, T., et al. (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., et al. (2016). Multi-compartment microscopic diffusion imaging. NeuroImage, 139:346 – 359.
- Le Bihan, D., Breton, E., Lallemand, D., et al. (1986). MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology, 161(2):401–407. PMID: 3763909.
- Maier-Hein, K. H., Neher, P. F., and Houde, J.-C. o. (2017). The challenge of mapping the human connectome based on diffusion tractography. Nature communications, 8(1):1349.

- Mandonnet, E., Sarubbo, S., and Petit, L. (2018). The Nomenclature of Human White Matter Association Pathways: Proposal for a Systematic Taxonomic Anatomical Classification. Frontiers in Neuroanatomy, 12:94.
- Nieuwenhuys, R., Voogd, J., and van Huijzen, C. (2008). The Human Central Nervous System. Springer, Berlin, Heidelberg.
- Novikov, D. S., Fieremans, E., Jespersen, S. N., et al. (2019). Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. NMR in Biomedicine, 32(4):e3998.
- Pestilli, F., Yeatman, J. D., Rokem, A., et al. (2014). Evaluation and statistical inference for human connectomes. Nature Methods, 11.
- Rheault, F. (2020). Analyse et reconstruction de faisceaux de la matière blanche. PhD thesis.
- Rheault, F., De Benedictis, A., Daducci, A., et al. (2020a). Tractostorm: The what, why, and how of tractography dissection reproducibility. Human Brain Mapping, 41(7):1859–1874.
- Rheault, F., Poulin, P., Caron, A. V., et al. (2020b). Common misconceptions, hidden biases and modern challenges of dMRI tractography. Journal of Neural Engineering, 17(1):011001.
- Schiavi, S., Ocampo-Pineda, M., Barakovic, M., et al. (2020). A new method for accurate in vivo mapping of human brain connections using microstructural and anatomical information. Science Advances, 6(31).
- Schilling, K. G., Nath, V., Hansen, C., et al. (2019). Limits to anatomical accuracy of diffusion tractography using modern approaches. NeuroImage, 185:1 – 11.
- Siless, V., Chang, K., Fischl, B., et al. (2018). AnatomicCuts: Hierarchical clustering of tractography streamlines based on anatomical similarity. NeuroImage, 166:32 – 45.
- Smith, R. E., Tournier, J.-D., Calamante, F., et al. (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., et al. (2013). SIFT: Spherical-deconvolution informed filtering of tractograms. NeuroImage, 67:298 – 312.



- Smith, R. E., Tournier, J.-D., Calamante, F., et al. (2015). SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. NeuroImage, 119:338 – 351.
- Sporns, O., Tononi, G., and Kötter, R. (2005). The Human Connectome: A Structural Description of the Human Brain. PLOS Computational Biology, 1(4).
- Thomas, C., Ye, F. Q., Irfanoglu, M. O., et al. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proceedings of the National Academy of Sciences, 111(46):16574–16579.
- Tournier, J.-D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. NeuroImage, 35(4):1459–1472.
- Tournier, J.-D., Calamante, F., and Connelly, A. (2010). Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. Proceedings of the International Society for Magnetic Resonance in Medicine, 88(2003):2010.
- Tournier, J.-D., Calamante, F., and Connelly, A. (2012). Mrtrix: Diffusion tractography in crossing fiber regions. International Journal of Imaging Systems and Technology, 22(1):53–66.
- Tournier, J.-D., Smith, R., Raffelt, D., et al. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. NeuroImage, 202:116137.
- Udin, S. B. and Fawcett, J. W. (1988). Formation of topographic maps. Annual Review of Neuroscience, 11(1):289–327. PMID: 3284443.
- Van Essen, D. C., Smith, S. M., Barch, D. M., et al. (2013). The WU-Minn Human Connectome Project: An overview. NeuroImage, 80:62–79.
- Wang, H. and Leng, C. (2008). A note on adaptive group lasso. Computational Statistics & Data Analysis, 52(12):5277–86.
- Yeh, C.-H., Jones, D. K., Liang, X., et al. (2020). Mapping Structural Connectivity Using Diffusion MRI: Challenges and Opportunities. Journal of Magnetic Resonance Imaging, n/a(n/a).

Zalesky, A., Fornito, A., Cocchi, L., et al. (2016). Connectome sensitivity or specificity: which is more important? NeuroImage, 142:407 – 420.

Zhang, Y., Brady, M., and 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.

Zhao, P., Rocha, G., and Yu, B. (2009). The composite absolute penalties family for grouped and hierarchical variable selection. Annals of Statistics, 37(6A):3468–3497.

## Figure legends

Figure 1. **Possible improvement over the COMMIT2 formulation.** *Top:* This simplistic example shows some representative streamlines between two cortical regions obtained with a generic tractography algorithm. This bundle consists of three true-positive sub-bundles (blue, red and green) but there are also some streamlines that seem to describe implausible trajectories (purple), which are probably false positives. However, using the parcellation-based approach introduced in COMMIT2, all these streamlines are grouped together as one unique bundle and considered as a whole, so they cannot be differentiated in true and false trajectories. *Bottom:* Example of an actual bundle of streamlines from in vivo data as reconstructed in the input tractogram and after filtering with COMMIT2, using either large or small groups, and with COMMIT2<sub>tree</sub> using both large and small groups in a hierarchical structure. This bundle consists of plausible streamlines that directly connect the brainstem to the postcentral gyrus (cyan arc), as well as of implausible streamlines that incorrectly pass through the corpus callosum (orange arrow).

Figure 2. **Visual comparison between the COMMIT, COMMIT2 and COMMIT2<sub>tree</sub> formulations.** COMMIT considered the streamlines as independent entities and COMMIT2 added a regularization term to take into account that streamlines are instead naturally organized in bundles. With COMMIT2<sub>tree</sub> we propose to improve this regularization by allowing a multi-level hierarchical organization of the streamlines, where every bundle may be composed of several sub-bundles and each can be selectively promoted or penalized.

Figure 3. **Illustrative toy example.** *First row:* ground-truth streamlines configuration as well as representative streamlines reconstructed with a generic tractography algorithm, i.e., raw tractogram, and after filtering with COMMIT. *Second row:* results after filtering with COMMIT2 (single level, once using large groups and another with small groups) and COMMIT2<sub>tree</sub> (multiple levels, combining parcellation and clustering). S1, S2, and S3, in green shades, represent true-positive streamlines while S4 and S5, in red shades, are false positives.

Figure 4. **Quantitative comparison on synthetic data between COMMIT2 and COMMIT2<sub>tree</sub>.** We compared the valid (VB, reported in green) and invalid (IB, reported in red) bundles in the connectomes generated on a synthetic phantom (A, top) for which we know the ground-truth connectivity (A, bottom). Panel B compares the quality of the connectomes estimated

from the probabilistic (top row) and deterministic (bottom row) tractograms before (raw tractogram) and after filtering with COMMIT2 and COMMIT2<sub>tree</sub>; performances of COMMIT are reported for reference.

Figure 5. **True- and false-positive bundles estimated in one HCP subject.** We report the raw tractogram and after filtering it with COMMIT2 and COMMIT2<sub>tree</sub>. The voxel coverage is reported for all the bundles, and for the filtered ones, it is also reported the sum of the streamline contributions estimated (weight). The performance of both methods is similar, the true positives keep comparable coverage to the raw tractogram, and the false positives are drastically reduced. Green arrows highlight the cases where COMMIT2<sub>tree</sub> assigned larger contributions to the true positives and further reduced the false-positives bundles than COMMIT2.

Figure 6. **Visual inspection of two representative bundles across the ten HCP subjects.** Reconstructed streamlines belonging to two representative bundles which connect, respectively: medulla oblongata with the precentral gyrus left (Bundle 1), and medulla oblongata with the precentral gyrus right (Bundle 2), as obtained across the ten subjects before (raw tractogram) and after filtering the tractograms with COMMIT2 and COMMIT2<sub>tree</sub>. For display and visual inspection purposes, the filtered bundles of streamlines are divided in two sub-bundles, true positives (TP) and false positives (FP), these definitinos were not used in the filtering procedure. The colormaps display the sum of the streamline contribution per sub-bundle, *cool* for TP and *hot* for FP. For all the bundles, the contributions estimated with COMMIT2<sub>tree</sub> are better that the ones estimated with COMMIT2, the FP have smaller contributions and the TP are more significant. The yellow and green arrows point to the region in the raw tractogram with false-positive sub-bundles after filtering with COMMIT2 and COMMIT2<sub>tree</sub>, respectively, highlighting their voxel coverage and the sum of the streamline contributions.

Figure 7. **Connectomes estimated in one HCP subject.** In the first row we show the estimated connectomes; in the case of the raw tractogram we used the number of streamlines, whereas for COMMIT2 and COMMIT2<sub>tree</sub> the sum of the streamline contributions was used. In the second row we present the intra-axonal signal fraction (IASF) maps used as input and estimated to perform a visual inspection. The third row presents root mean square error (RMSE) maps of the estimations, along with their mean and standard deviation for a quantitative comparison.

## Figures

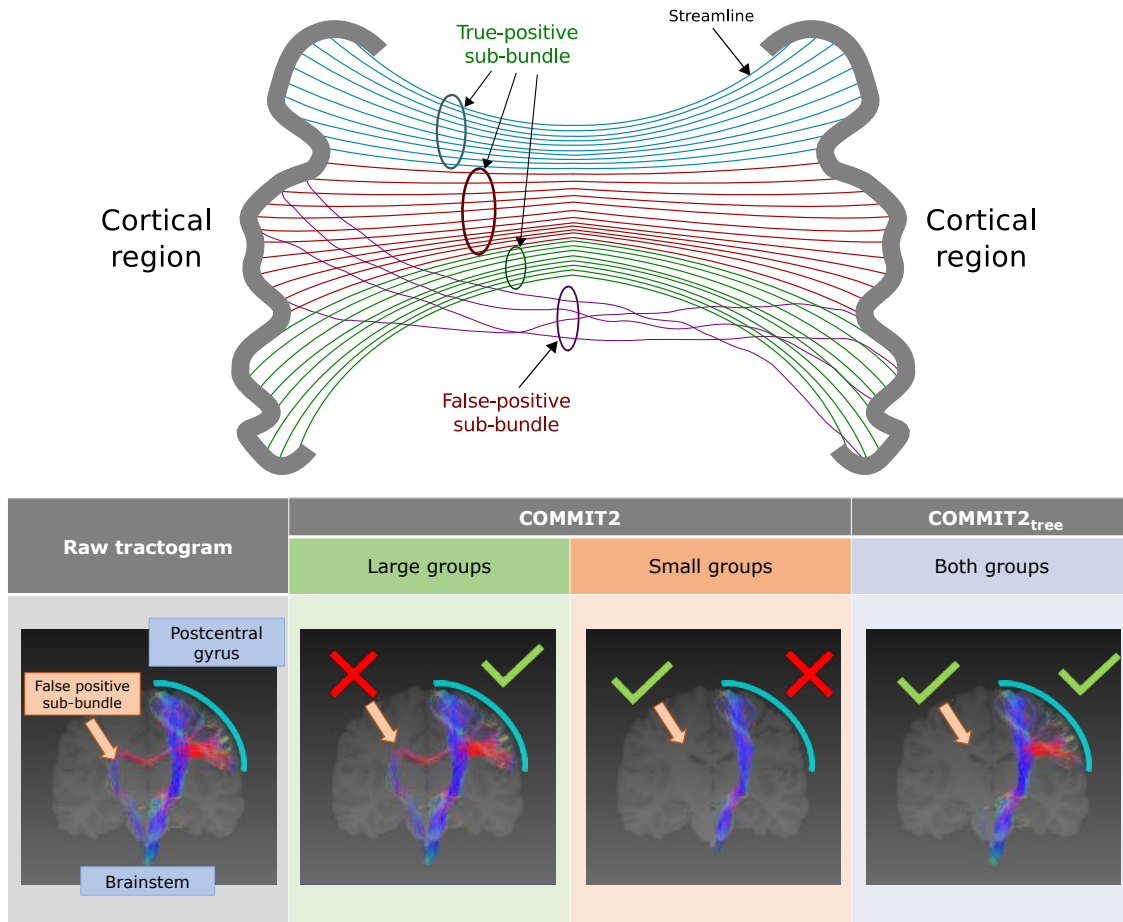


Figure 1: **Possible improvement over the COMMIT2 formulation.** *Top:* This simplistic example shows some representative streamlines between two cortical regions obtained with a generic tractography algorithm. This bundle consists of three true-positive sub-bundles (blue, red and green) but there are also some streamlines that seem to describe implausible trajectories (purple), which are probably false positives. However, using the parcellation-based approach introduced in COMMIT2, all these streamlines are grouped together as one unique bundle and considered as a whole, so they cannot be differentiated in true and false trajectories. *Bottom:* Example of an actual bundle of streamlines from in vivo data as reconstructed in the input tractogram and after filtering with COMMIT2, using either large or small groups, and with COMMIT2<sub>tree</sub> using both large and small groups in a hierarchical structure. This bundle consists of plausible streamlines that directly connect the brainstem to the postcentral gyrus (cyan arc), as well as of implausible streamlines that incorrectly pass through the corpus callosum (orange arrow).

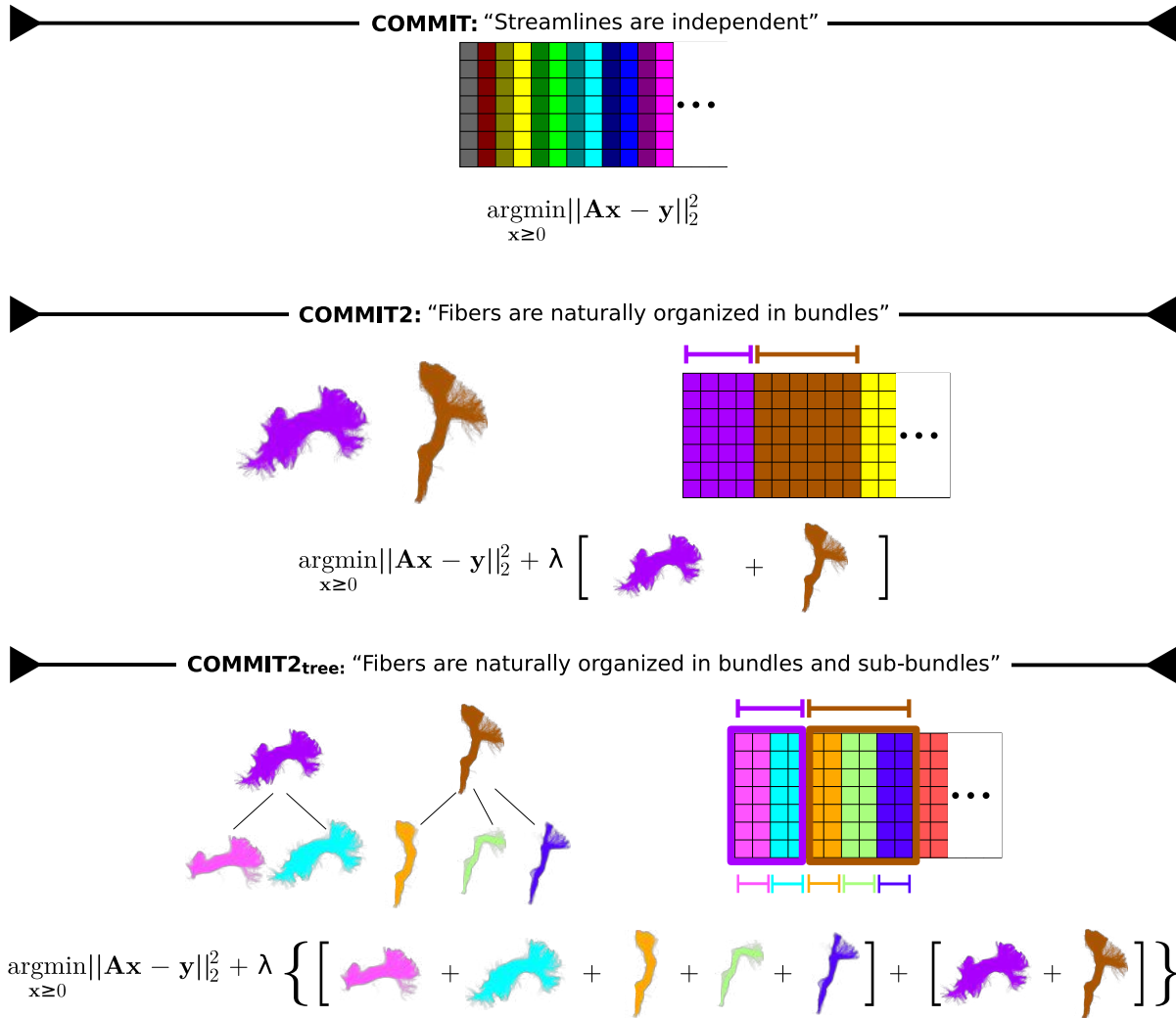


Figure 2: **Visual comparison between the COMMIT, COMMIT2 and COMMIT2<sub>tree</sub> formulations.** COMMIT considered the streamlines as independent entities and COMMIT2 added a regularization term to take into account that streamlines are instead naturally organized in bundles. With COMMIT2<sub>tree</sub> we propose to improve this regularization by allowing a multi-level hierarchical organization of the streamlines, where every bundle may be composed of several sub-bundles and each can be selectively promoted or penalized.

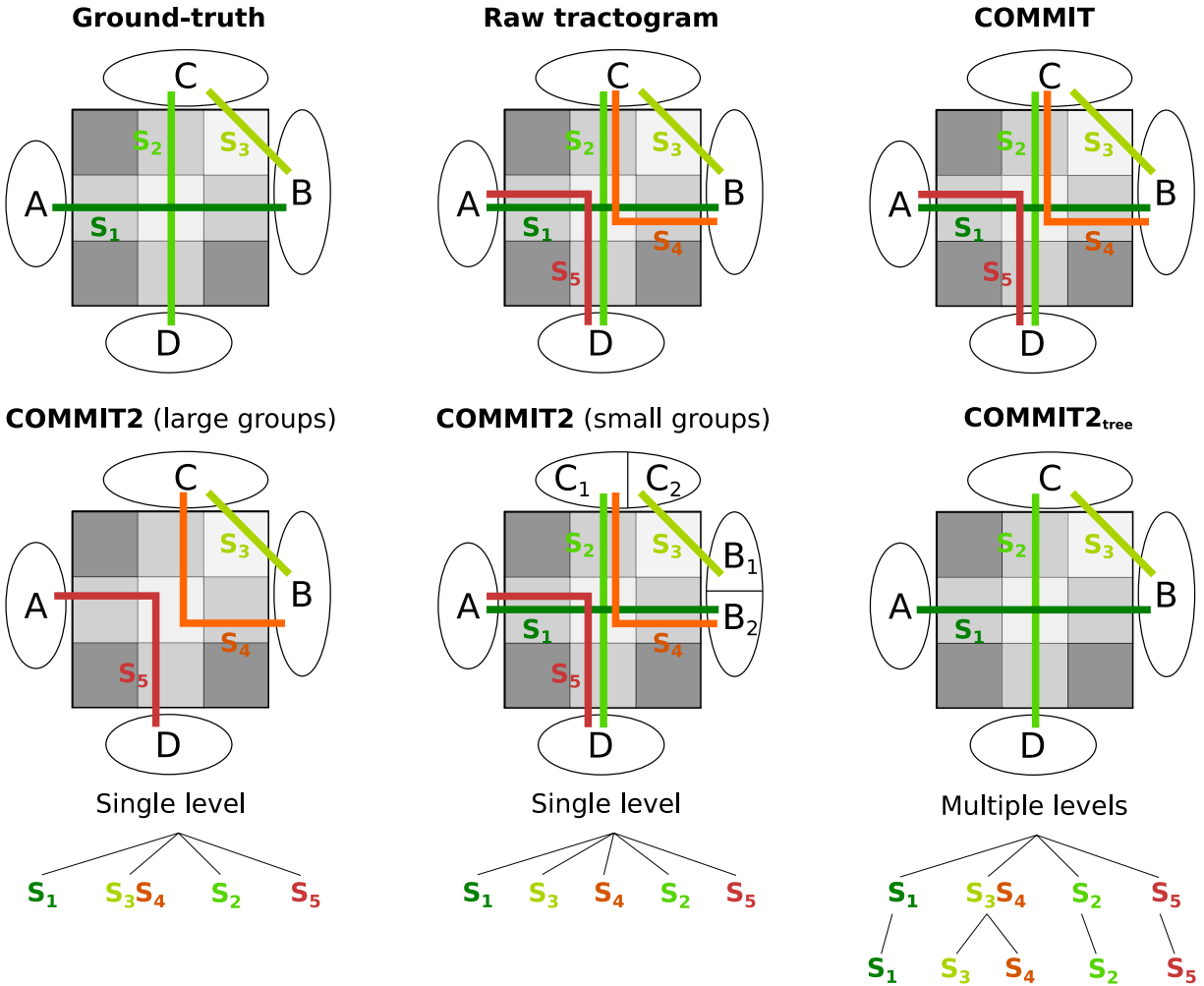


Figure 3: **Illustrative toy example.** *First row:* ground-truth streamlines configuration as well as representative streamlines reconstructed with a generic tractography algorithm, i.e., raw tractogram, and after filtering with COMMIT. *Second row:* results after filtering with COMMIT2 (single level, once using large groups and another with small groups) and COMMIT2<sub>tree</sub> (multiple levels, combining parcellation and clustering).  $S_1$ ,  $S_2$ , and  $S_3$ , in green shades, represent true-positive streamlines while  $S_4$  and  $S_5$ , in red shades, are false positives.

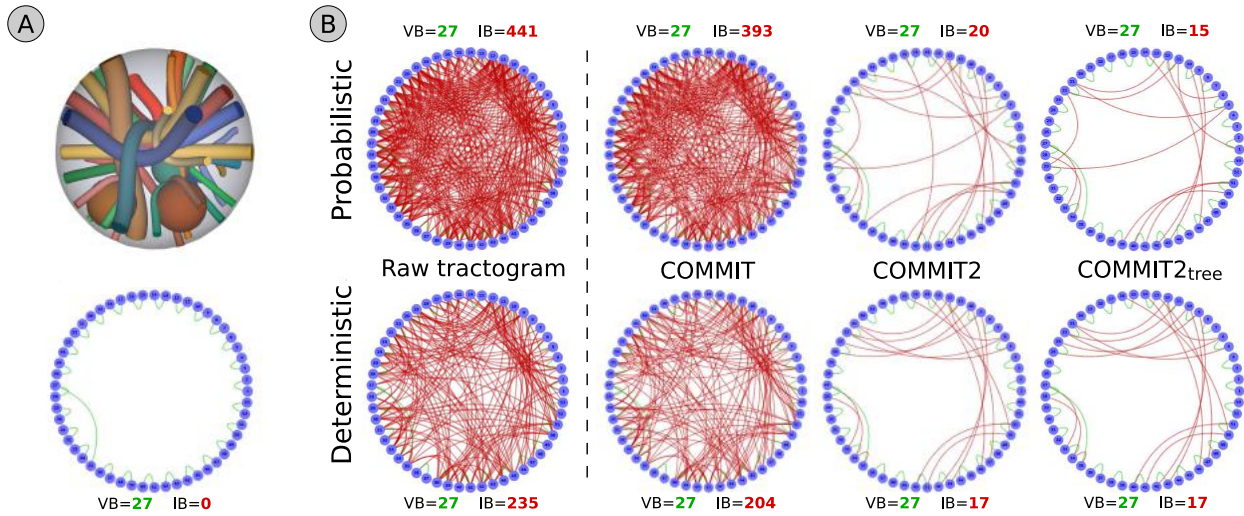


Figure 4: **Quantitative comparison on synthetic data between COMMIT2 and COMMIT2<sub>tree</sub>.** We compared the valid (VB, reported in green) and invalid (IB, reported in red) bundles in the connectomes generated on a synthetic phantom (A, top) for which we know the ground-truth connectivity (A, bottom). Panel B compares the quality of the connectomes estimated from the probabilistic (top row) and deterministic (bottom row) tractograms before (raw tractogram) and after filtering with COMMIT2 and COMMIT2<sub>tree</sub>; performances of COMMIT are reported for reference.



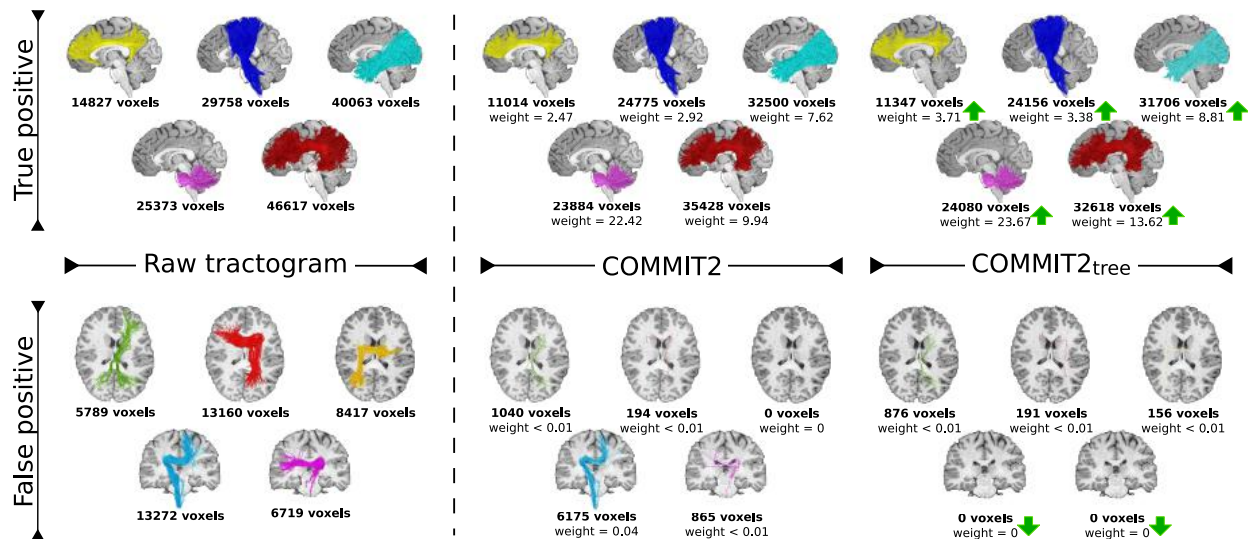


Figure 5: **True- and false-positive bundles estimated in one HCP subject.** We report the raw tractogram and after filtering it with COMMIT2 and COMMIT2<sub>tree</sub>. The voxel coverage is reported for all the bundles, and for the filtered ones, it is also reported the sum of the streamline contributions estimated (weight). The performance of both methods is similar, the true positives keep comparable coverage to the raw tractogram, and the false positives are drastically reduced. Green arrows highlight the cases where COMMIT2<sub>tree</sub> assigned larger contributions to the true positives and further reduced the false-positives bundles than COMMIT2.

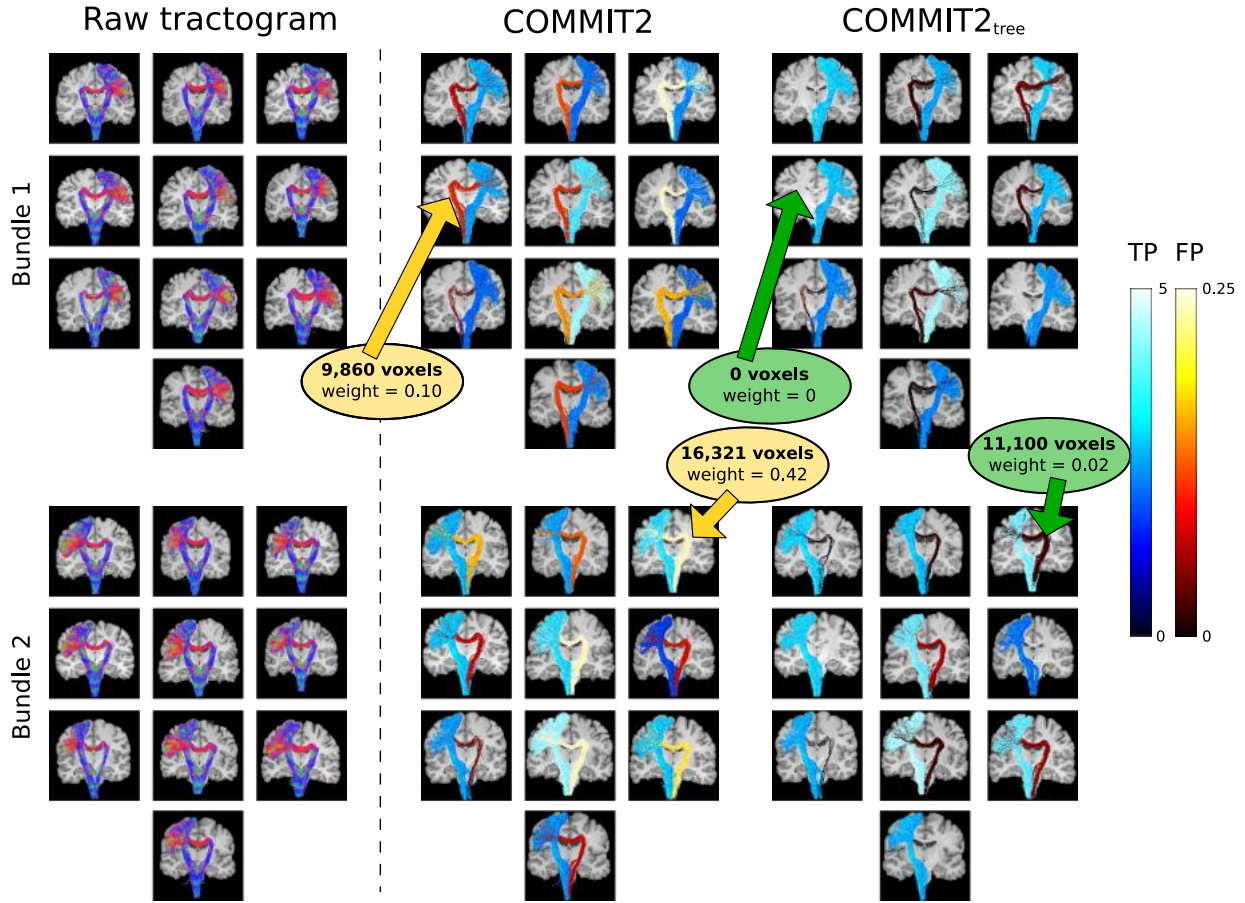


Figure 6: **Visual inspection of two representative bundles across the ten HCP subjects.** Reconstructed streamlines belonging to two representative bundles which connect, respectively: medulla oblongata with the precentral gyrus left (Bundle 1), and medulla oblongata with the precentral gyrus right (Bundle 2), as obtained across the ten subjects before (raw tractogram) and after filtering the tractograms with COMMIT2 and COMMIT2<sub>tree</sub>. For display and visual inspection purposes, the filtered bundles of streamlines are divided in two sub-bundles, true positives (TP) and false positives (FP), these definitions were not used in the filtering procedure. The colormaps display the sum of the streamline contribution per sub-bundle, *cool* for TP and *hot* for FP. For all the bundles, the contributions estimated with COMMIT2<sub>tree</sub> are better than the ones estimated with COMMIT2, the FP have smaller contributions and the TP are more significant. The yellow and green arrows point to the region in the raw tractogram with false-positive sub-bundles after filtering with COMMIT2 and COMMIT2<sub>tree</sub>, respectively, highlighting their voxel coverage and the sum of the streamline contributions.

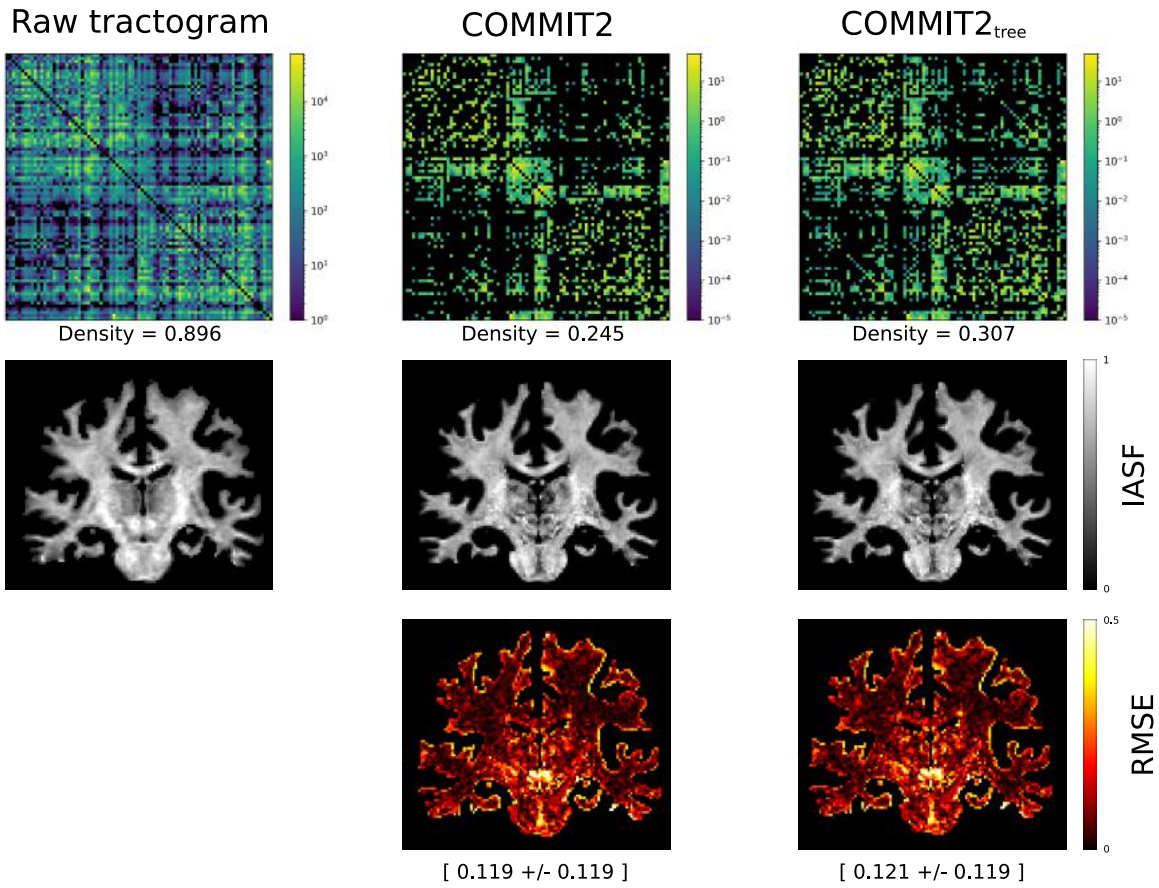


Figure 7: **Connectomes estimated in one HCP subject.** In the first row we show the estimated connectomes; in the case of the raw tractogram we used the number of streamlines, whereas for COMMIT2 and COMMIT2<sub>tree</sub> the sum of the streamline contributions was used. In the second row we present the intra-axonal signal fraction (IASF) maps used as input and estimated to perform a visual inspection. The third row presents root mean square error (RMSE) maps of the estimations, along with their mean and standard deviation for a quantitative comparison.

## Tables

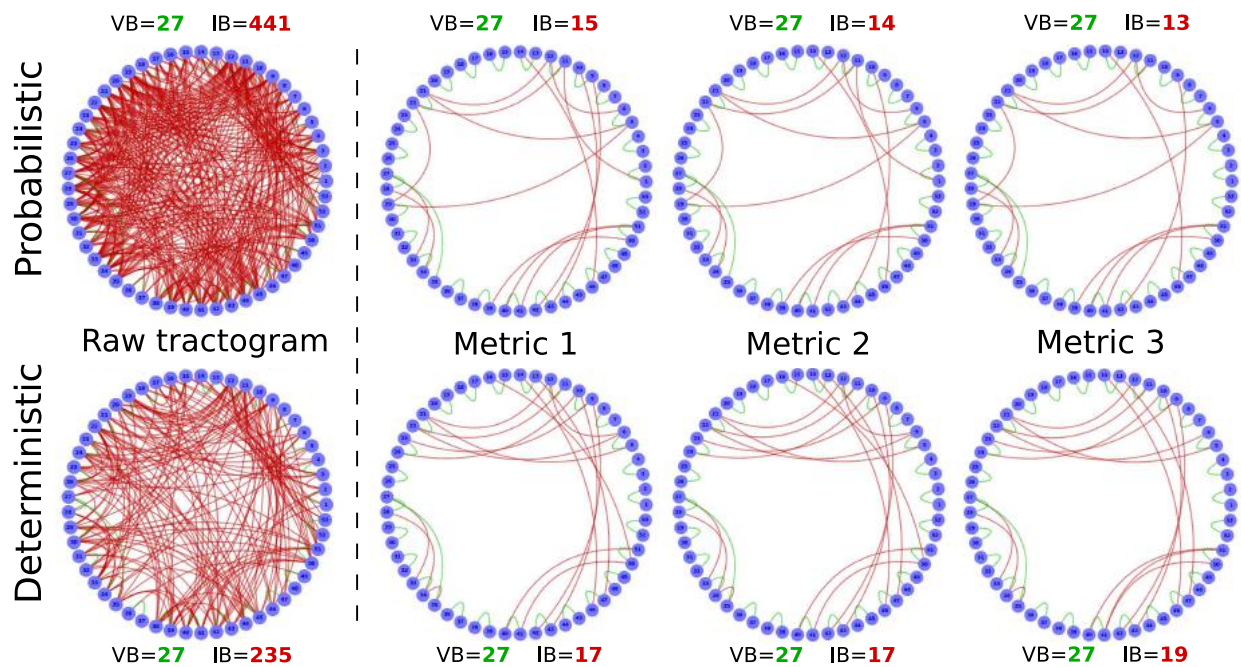
		INPUT	COMMIT2		COMMIT2 <sub>tree</sub>	
		<i>voxels</i>	<i>voxels</i>	<i>weight</i>	<i>voxels</i>	<i>weight</i>
<b>True positives</b>	#1	18,968 ± 5,229	12,983 ± 4,997	2.92 ± 1.80	10,678 ± 4,421	2.82 ± 2.12
	#2	39,920 ± 3,637	30,892 ± 4,108	3.88 ± 1.82	21,160 ± 5,285	3.90 ± 2.15
	#3	44,899 ± 2,995	37,840 ± 3,143	12.73 ± 3.24	31,203 ± 3,780	14.76 ± 4.15
	#4	28,403 ± 3,681	26,101 ± 3,372	18.95 ± 4.16	24,243 ± 3,214	17.99 ± 4.13
	#5	57,300 ± 6,645	43,324 ± 7,729	10.03 ± 4.39	34,406 ± 6,835	14.26 ± 5.52
<b>False positives</b>	#1	19,362 ± 4,924	7,847 ± 3,134	0.09 ± 0.04	1,211 ± 1,256	<0.01 ± 0.01
	#2	22,037 ± 5,221	4,110 ± 6,605	0.04 ± 0.08	229 ± 686	<0.01 ± 0.01
	#3	10,066 ± 6,567	3,617 ± 3,077	0.05 ± 0.05	1,333 ± 1,511	0.02 ± 0.03
	#4	20,908 ± 7,060	10,805 ± 3,720	0.13 ± 0.08	897 ± 2,036	0.01 ± 0.02
	#5	11,243 ± 5,714	3,876 ± 3,013	0.03 ± 0.03	474 ± 634	0.01 ± 0.02

Table 1: **Stability across the ten HCP subjects.** Voxel coverage (*voxels*) and estimated sum of streamline contributions (*weight*) of the 5 true-positive bundles and the 5 false-positives bundles shown in Figure 5. *True positives*: #1 (yellow), #2 (dark blue), #3 (cyan), #4 (fuchsia), #5 (dark red). *False positives*: #1 (green), #2 (light red), #3 (orange), #4 (light blue), #5 (magenta).

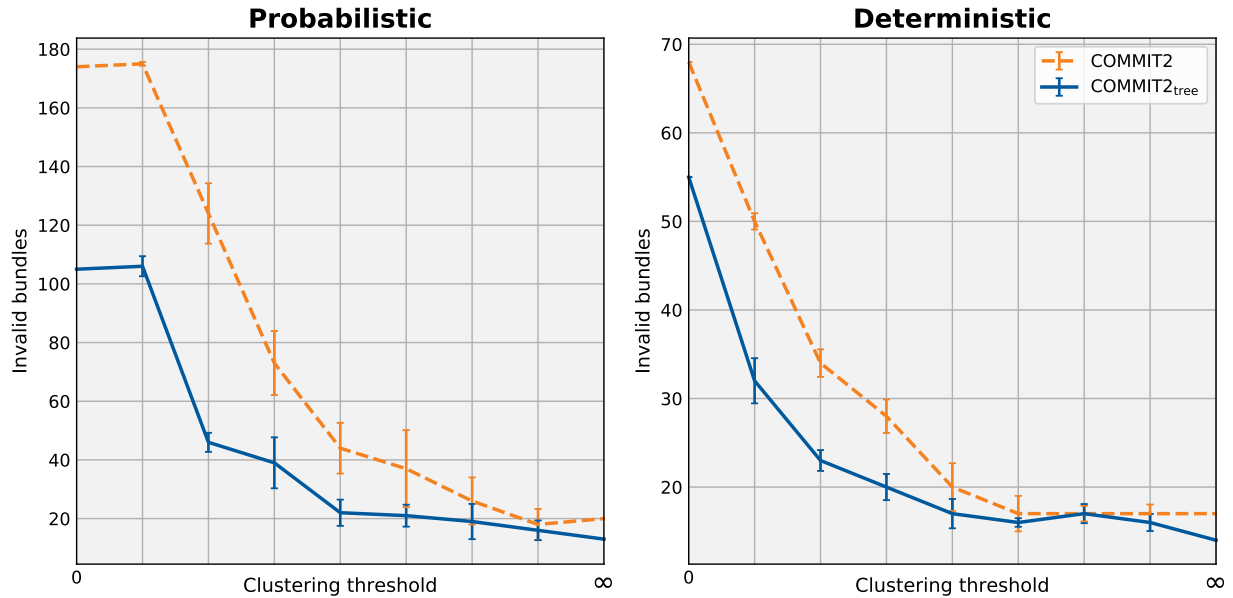
		COMMIT2		COMMIT2 <sub>tree</sub>	
		<i>voxels</i>	<i>weight</i>	<i>voxels</i>	<i>weight</i>
<b>Bundle 1</b>	TP	20,280 ± 2,159	2.75 ± 0.81	17,733 ± 2,181	3.18 ± 0.76
	FP	9,325 ± 3,471	0.14 ± 0.08	2,842 ± 2,689	0.01 ± 0.01
<b>Bundle 2</b>	TP	20,252 ± 2,577	2.97 ± 0.71	17,819 ± 2,166	3.41 ± 0.76
	FP	10,195 ± 3,876	0.18 ± 0.12	4,122 ± 4,300	0.02 ± 0.02

Table 2: **Quantitative comparison across the ten HCP subjects of the sub-bundles presented in Figure 6.** Mean and standard deviation of the sum of the streamline contributions (weight) and voxel coverage (voxels) for the true-positive (TP) and false-positive (FP) sub-bundles. *Bundle 1*: Medulla oblongata and Precentral gyrus left. *Bundle 2*: Medulla oblongata and Precentral gyrus right.

## Supplementary material



Supplementary Figure S1: **Impact of the distance metric on the quality of the tractograms filtered with  $\text{COMMIT2}_{\text{tree}}$ .** Following the same notation as Figure 4, we evaluated in the synthetic phantom the effect of changing the clustering metric to build the second level of the hierarchy. We tested the following metrics (see <https://www.dipy.org/> for more details): *Metric 1* is the Average of Pointwise Euclidean Metric between streamlines (same as Figure 4), *Metric 2* is the Euclidean Metric between the centers of mass of the streamlines, and *Metric 3* is the Cosine Similarity between the vectors of the streamline endpoints.



Supplementary Figure S2: **Impact of the group size on the quality of the tractograms filtered with COMMIT2<sub>tree</sub>.** In the synthetic phantom, we investigated the effect of changing the size of the groups to build the second level of the hierarchy, by changing the threshold of the Average of Pointwise Euclidean Metric between streamlines from 0 (one streamline per cluster) to “infinity” (bundles are not clustered). We repeated the clustering 10 times and, here, we report the mean and standard deviation of invalid bundles kept after filtering with COMMIT2<sub>tree</sub> (continuous blue line) and, for comparison, COMMIT2 (orange dashed line).