Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles

Pietro Gori, Olivier Colliot, Linda Marrakchi-Kacem, Yulia Worbe, Fabrizio
de Vico Fallani, Mario Chavez, Cyril Poupon, Andreas Hartmann, Nicholas
Ayache, Stanley Durrleman

HAL Id: hal-01346067
https://hal.science/hal-01346067

Submitted on 18 Jul 2016

Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles

Pietro Gori*, Olivier Colliot, Linda Marrakchi-Kacem, Yulia Worbe, Fabrizio De Vico Fallani, Mario Chavez, Cyril Poupon, Andreas Hartmann, Nicholas Ayache and Stanley Durrleman

Abstract—**Fiber bundles stemming from tractography algorithms contain many streamlines. They require therefore a great amount of computer memory and computational resources to be stored, visualised and processed. We propose an approximation scheme for fiber bundles which results in a parsimonious representation of weighted prototypes. Prototypes are chosen among the streamlines and they represent groups of similar streamlines. Their weight is related to the number of approximated streamlines. Both streamlines and prototypes are modelled as weighted currents. This computational model does not need point-to-point correspondences and two streamlines are considered similar if their endpoints are close to each other and if their pathways follow similar trajectories. Moreover, the space of weighted currents is a vector space with a closed-form metric. This permits easy computation of the approximation error and the selection of the prototypes is based on the minimisation of this error. We propose an iterative algorithm which approximates independently and simultaneously all the fascicles of the bundle in a fast and accurate way. We show that the resulting representation preserves the shape of the bundle and it can be used to accurately reconstruct the original structural connectivity. We evaluate our algorithm on bundles obtained from both deterministic and probabilistic tractography algorithms. The resulting approximations use on average only 2% of the original streamlines as prototypes. This drastically reduces the computational burden of the processes where the geometry of the streamlines is considered. We demonstrate its effectiveness using as example the registration between two fiber bundles.**

Index Terms—**Diffusion weighted imaging, Brain, Connectivity analysis, Dimensionality reduction, Registration, Tractography, Visualization**

I. INTRODUCTION

TRACTOGRAPHY [1], [2] from diffusion-weighted magnetic resonance imaging (DW-MRI) [3] is the only non-invasive technique capable to trace in vivo the wiring architecture of the human brain white matter. It is widely employed for both clinical (i.e. stroke [4], surgical procedures [5]) and research purposes (i.e. Alzheimer's disease, schizophrenia [6]). The 3D polylines stemming from tractography algorithms, called streamlines or fibers, are only estimates of the trajectories of large groups of neural axons. Streamlines are traced from points inside a starting voxel, called seeds, and they are constituted of segments connecting neighbouring

P. Gori, O. Colliot, L. Marrakchi-Kacem, F. De Vico Fallani, M. Chavez and S. Durrleman are with Aramis project-team, Inria, UPMC Univ Paris 06, Inserm U1127, CNRS UMR 7225, ICM, Paris, France.

O. Colliot, Y. Worbe and A. Hartmann are with AP-HP, Pitié-Salpêtrière hospital, F-75013, Paris, France.

C. Poupon is with NeuroSpin, CEA, Gif-Sur-Yvette, France.

N. Ayache is with Asclepios project-team, Inria, Sophia Antipolis, France.

* Corresponding author: pietro.gori@inria.fr

voxels. The direction of these segments is defined by a local diffusion model (i.e. tensor, Q-ball [7]) computed at each voxel and by a tractography method: deterministic or probabilistic [7]. Deterministic algorithms produce segments which follow the principal direction of the local diffusion model whereas probabilistic ones use randomly perturbed versions of the main direction. Depending on the tractography step-size parameter (i.e. the distance between two consecutive points), segments may be collinear. A piece-wise linearization step, as proposed in [8], may be applied to reduce the number of segments. Seeds are usually placed in every voxel of the white matter (whole-brain tractography) and the resulting streamlines can be divided into different fiber bundles based on clustering algorithms or starting/ending Regions of Interest (ROI). Fiber bundles may then be decomposed into *fascicles* which are groups of fibers with a similar pathway and whose extremities are close to each other, connecting therefore the same functional territories.

Fiber bundles are difficult to analyse both qualitatively and quantitatively due to their considerable number of streamlines. The size of a bundle can make computationally intractable processes such as clustering [9], registration [10], atlas construction [11] or shape analysis [12], where the trajectory of the streamlines is considered. Moreover, the great quantity of streamlines might also complicate the rendering, the visualisation and the interpretation of a bundle, thus limiting possible clinical applications.

In this paper, we propose to approximate a fiber bundle with a parsimonious representation of weighted streamline prototypes. We exploit the fact that many streamlines, starting from seeds in the same voxel or in neighbour voxels, share the same pathway and ending area. We approximate these streamlines with one of them, called prototype. We use a computational model for both streamlines and prototypes characterised by an explicit and easily computable metric. This allows us to control the approximation error and to select the streamlines which minimize it as prototypes. Furthermore, the resulting representation of weighted prototypes preserve both the shape and the fiber density on grey matter structures (i.e. structural connectivity) with a controlled error.

The proposed approximation scheme is conceived for fiber bundles resulting from both deterministic and probabilistic streamline tractography algorithms. The definitions of streamline, fascicle and bundle, as employed throughout this paper, are as follows. A streamline is a curve composed of a finite, ordered and connected sequence of 3D points. The distance between connected points is not assumed to be constant. The

number of points may vary between two different streamlines. A fascicle is an ensemble of streamlines with similar pathway and whose extremities are close to each other. A bundle is a group of streamlines with a consistent orientation and connecting two specific ROIs defined by the user. Every bundle is composed of one or more fascicles.

II. RELATED WORK

In the last years, there has been a great effort to compactly represent a fiber bundle. A pragmatic strategy is to randomly choose a smaller subset of streamlines. The sampling is not driven by the minimisation of an approximation error and it is not possible to control the selection of the streamlines. This can cause the loss of the smallest fascicles of the bundle which might be important for the purpose of the study. More sophisticated solutions have been proposed and they can be divided into two categories. The first group gathers computational models conceived to compactly represent single streamlines. The second category assembles instead strategies focused on simplifying the representation of the entire fiber bundle.

A. Compact representation of streamlines

Streamlines are composed of contiguous variable-length segments whose number might also vary among fibers. Different computational models have been proposed whose goal is to concisely parametrise a streamline. A widely employed method consists of defining point-to-point correspondence among streamlines parametrising them as sets of points [13] or with cubic B-splines [14] for instance. This technique eases the computations but it can be applied only if streamlines have a similar length and the definition of corresponding points can be very challenging. Other authors proposed to characterise a streamline using only its extremities [15] or its connectivity signature [16], namely the probabilities to be linked to a defined set of ROIs. These methods have been used for clustering, visualisation and interpretation purposes but they do not take into consideration the shape of the streamlines which is important for registration and morphometry. Conversely, different authors proposed to evaluate only the geometry of the fibers, without taking into consideration their extremities. A first example is given by the methods based on Fourier descriptors [17], [18], which result in a concise parametrisation useful for clustering and shape analysis. In these models the number of descriptors needs to be fixed though and the optimal number depends on the length and shape of the streamlines, which might vary even within a single bundle. Lately, other authors proposed to represent a streamline as a blurred indicator function modelled as a Gaussian process [19], [20]. This representation can be easily employed to compare and average streamlines. However, it is not a geometric primitive and it is therefore difficult to use in multi-object registrations together with other geometric primitives such as 3D surface meshes.

B. Compact representation of the whole bundle

The second category is composed of methods which approximate the entire fiber bundle. The most common strategy is to divide the fiber bundle into subsets, usually called clusters, which are then characterised by representative fibers (i.e. prototypes.) These fibers represent the average trajectory of the streamlines of the clusters. They can be computed as the mode [21] or mean [22] of the streamlines, if there is a point-correspondence, or according to a fiber dissimilarity measure [23], [24]. Representative fibers are mainly used to ease the interpretation and visualisation of a bundle and to reduce the memory footprint and computational time for shape analysis and registration. A first approach to speed up affine registration with prototype fiber representations was the one proposed in [25]. Other authors have also employed isosurfaces to represent the spatial variation of the clusters [26]. This representation can be used only for tubular-shaped bundles that can be modelled as convex envelopes. Other bundles, such as the corpus callosum and the rostral part of the corticospinal tract, have a different topology and they are defined as sheet-like bundles. In [27] the authors proposed to represent those bundles as 3D surface meshes whereas in [28] it was suggested to use deformable medial models (cm-reps). In both cases, the medial surface representations are employed only for visualisation and clustering and to provide statistics about diffusion coefficients. A different representation, which can be employed for any kind of bundle, is the tract probability map [19], [29], [30]. It indicates the probability of a voxel to belong to a given bundle. This method is very concise but it is not based on a geometrical primitive and it has been used for visualisation, interpretation and clustering. A last example is the sparse representation based on the matching pursuit algorithm for currents presented in [11]. In the framework of currents [31], a bundle is considered as a single mathematical object composed of disconnected oriented points which model the local orientation of the streamlines. The approximation presented in [11] represents a bundle with a sparse set of oriented points. This representation is very concise but it has the drawback to accurately approximate only the areas of the bundle characterised by a high density of streamlines, like the central mass of the bundle. Thus, the small fascicles may not be well approximated. Moreover, the framework of currents does not take into account the extremities of the streamlines. This prevents the analysis of the structural connectivity, namely the areas of the gray matter connected by the bundle.

III. OUR CONTRIBUTION

In this paper, extending [32], we propose to approximate *any* fiber bundle with a set of weighted prototypes. Prototypes are chosen among the streamlines and they represent ensembles of similar fibers. Their weights are related to the number of streamlines approximated. Both prototypes and streamlines are modelled as weighted currents, an extension of the framework of currents. This computational model takes into consideration both the *pathway* of the fiber/prototype and the *anatomical location* of its extremities. Two fibers/prototypes are considered similar if their endpoints are close to each other and if their trajectories are similar. The space of weighted currents is a *vector space* with an explicit and easily computable metric.

This implies that the approximation error of the proposed scheme can be easily controlled. The resulting parsimonious representation, up to a reasonable approximation level, preserves both the shape and the structural connectivity (computed as streamlines density) of the original bundle. Moreover, the framework of weighted currents inherits from the one of currents [33] the fact that it does *not* need either point-to-point or streamline-to-streamline correspondences. Nevertheless, it requires the definition of the starting and ending point of each streamline. The uncertainty of tractography algorithms near the grey/white matter boundaries is taken into account by a Gaussian smoothing kernel. Prototypes are visualised as tubes whose constant radii are proportional to their weights. This concise representation can be easily combined in multi-object studies with other geometric primitives such as surface meshes modelled as landmarks, currents or varifolds [34], [35].

The paper is organised as follows. In Section IV, we first present the framework of weighted currents and then the different steps of the proposed approximation scheme. After that, we demonstrate the effectiveness of our algorithm on deterministic and probabilistic fiber bundles from both a qualitative and quantitative point of view.

IV. METHOD

A. Weighted currents

The framework of weighted currents is an extension of the one of currents [31] [36, Chapter 2]. In this framework, a streamline is considered as a set of disconnected oriented points which are *weighted* by the spatial coordinates of the streamline extremities. In this way, every oriented point encodes not only the local orientation of a streamline, as with currents, but also its connectivity. It is an adaptation of the framework of functional currents [37].

A streamline X is a polygonal line of N segments. The coordinates of the two extremities f^a and f^b are two 3D vectors defined in the space $Q = \mathbf{R}^3 \mathrm{x} \mathbf{R}^3$. The fiber X is modelled as a 1-weighted current C_X via a line integral of a vector field ω:

$$C_X(\omega) = \int_X w(x, f^a, f^b)^T \alpha(x) dx \approx \sum_{i=1}^{N} \omega_{(x_i, f^a, f^b)}^T (\alpha_i)$$
$$\approx \sum_{i=1}^{N} \delta_{(x_i, f^a, f^b)}^{\alpha_i} (\omega) \quad (1)$$

where x_i and α_i are respectively the center and the tangent vector of segment i which is approximated by a Dirac delta weighted current δ. It can be shown that the approximation error tends to zero as the sampling becomes more accurate i.e. the length of the segments decreases [38]. The vector field ω belongs to a reproducing kernel Hilbert space (RKHS) W defined on the product space $\mathbf{R}^3 \mathrm{x} Q$. The space of weighted currents is a continuous linear form on W and every weighted current C_X belongs to its dual space W^*. As a space of mappings, the space of weighted currents is a *vector* space. A natural way to build a kernel K associated to the product
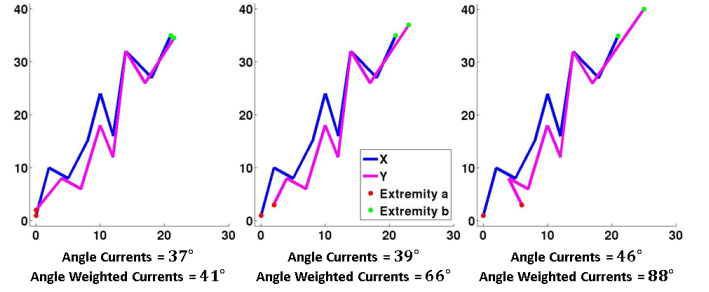

Fig. 1. Two 2D streamlines X and Y are compared using both the framework of currents and the one of weighted currents. In the three figures only the position of the extremities change, the overall pathway remains almost unchanged. The framework of weighted currents is more sensitive to the distance between the extremities which explains why the two streamlines X and Y are almost orthogonal in the last figure on the right. The bandwidths of all kernels employed in both frameworks are equal to 5.

space W is as tensor product of two kernels defined separately in $\mathbf{R}^3$ (K_g) and in Q (K_f): $K = K_g \otimes K_f$. Since even Q is a product space, K_f is also defined as a tensor product between two kernels K_a and K_b. Thus, the kernel K results: $K((x, f^a, f^b), (y, t^a, t^b)) = K_g(x, y) K_a(f^a, t^a) K_b(f^b, t^b)$. All kernels K_g, K_a and K_b are defined as Gaussian and they are parametrised by their bandwidths λ_g, λ_a and λ_b. Using these kernels, the inner product in the framework of weighted currents between two Diracs is defined as: $\langle \delta_{(x,f^a,f^b)}^{\alpha}, \delta_{(y,t^a,t^b)}^{\beta} \rangle_{W*} = K_a(f^a, t^a) \ K_b(f^b, t^b) (\alpha^T K_g(x, y) \beta)$. By linearity, the inner product between two streamlines X and Y ($C_Y(\omega) \approx \sum_{j=1}^{M} \omega_{(y_j, t^a, t^b)}(\beta_j)$) is:

$$\langle C_X, C_Y \rangle_{W*} = K_a(f^a, t^a) K_b(f^b, t^b) \boxed{\sum_{i=1}^{N} \sum_{j=1}^{M} \alpha_i^T K_g(x_i, y_j) \beta_j}$$

$$= \exp\left(\frac{-||f^a - t^a||_2^2}{\lambda_a^2}\right) \exp\left(\frac{-||f^b - t^b||_2^2}{\lambda_b^2}\right) \quad (2)$$

$$\boxed{\sum_{i=1}^{N} \sum_{j=1}^{M} \exp\left(\frac{-||x_i - y_j||_2^2}{\lambda_g^2}\right) \alpha_i^T \beta_j}$$

The framed part would be the inner product between X and Y if modelled as currents. It measures overall differences between the geometry of their trajectories where the degree of detail is determined by the scale λ_g, which is in distance unit. If the average distance between the segments of X and Y is definitely greater than 3 times λ_g, the two streamlines are considered orthogonal ($\langle C_X, C_Y \rangle_{W*} \sim 0$). The two other terms evaluate how far the streamline extremities are from each other and they are parametrised by λ_a and λ_b respectively. If f^a and t^a (resp. f^b and t^b) are spaced more than 3 times λ_a (resp. λ_b) apart, the two streamlines are considered orthogonal. This means that two streamlines are considered orthogonal if they do not share either the pathway or the ending areas. As shown in Fig.1, the framework of currents is almost "blind" to a change of the positions of the end-points. Even if the extremities of the two streamlines are far from each other, with respect to the kernel bandwidth λ_g, the angle

$(cos^{-1} \frac{\langle C_X, C_Y \rangle_{W^*}}{|C_X|_{W^*} |C_Y|_{W^*}})$ between the two streamlines varies by only few degrees. On the contrary, the streamlines become almost orthogonal in the framework of weighted currents when their extremities are spaced out. This makes the definition of similarity twofold in the framework of weighted currents. Two streamlines are considered similar if their pathways are alike, as in currents, but also if their endpoints are close to each other. In Fig.2 we show the most similar fibers to the red streamline in the framework of currents (green) and in the one of weighted currents (blue). The green fibers share a similar pathway with the red one but some of them connect different anatomical areas. On the contrary, the blue fibers are similar to the red streamline both in terms of geometry and connectivity.

As currents, the framework of weighted currents does not need point-to-point correspondence, except for the extremities. This can be obtained, for instance, by tracing all the streamlines of a bundle from one ROI to another one, as it is done for the bundles considered in this paper. Moreover, every streamline S_i is considered as a vector in a Hilbert space. Thus, a fiber bundle, which is the union of many fibers $B = \cup_i^n S_i$, is represented as a sum in this framework: $C_B = \sum_i^n C_{S_i}$. The difference between two streamlines, modelled as weighted currents, is defined as their sum with the orientation of the second fiber inverted. If two fibers are equal, their difference cancels out. Furthermore, it is also possible to compute the average weighted current $\bar{S}$ of a fiber bundle as: $C_{\bar{S}} = \frac{1}{n} \sum_i^n C_{S_i}$. Given the inner product defined in Eq.2, the squared norm of the difference between two bundles $C_B = \sum_{i=1}^n C_{S_i}$ and $C_{B'} = \sum_{p=1}^m C_{S'_p}$ is equal to: $\|C_B - C_{B'}\|_{W^*}^2 = \sum_{i=1}^n \sum_{j=1}^n \langle C_{S_i}, C_{S_j} \rangle_{W^*} + \sum_{p=1}^m \sum_{q=1}^m \langle C_{S'_p}, C_{S'_q} \rangle_{W^*} - 2 \sum_{i=1}^n \sum_{p=1}^m \langle C_{S_i}, C_{S'_p} \rangle_{W^*}$.

A bundle B composed of two streamlines, X and Y, is modelled as $C_B = C_X + C_Y$ in the framework of weighted currents. If the two streamlines X and Y are similar in this framework, their sum can be well approximated by $C_B = 2C_X$ or $C_B = 2C_Y$. This is crucial for the scope of this paper since an ensemble of streamlines can be represented with a single *weighted prototype* where the weight is related to the number of streamlines approximated. In the previous example both X and Y could be chosen as prototype and the weight would be 2. A weighted prototype can be visualised as a tube where the streamline chosen as prototype is the central axis and the constant radius is proportional to the weight (see Fig.3). In the following, we will describe how to use this idea to approximate a complex bundle stemming from a tractography algorithm. We will also assume that both streamlines and bundles are modelled as weighted currents writing simply S (resp. B) instead than C_S (resp. C_B).

B. Approximation scheme

The goal of the proposed approximation scheme is to represent a fiber bundle B with a set of weighted prototypes $\{\tau_k P_k\}$. The resulting parsimonious representation should preserve both the shape and the structural connectivity (streamlines density on the grey matter surfaces) of the original bundle. The proposed algorithm is based on a *greedy approach* where we first subdivide the bundle into fascicles and then select the prototypes in each fascicle independently.
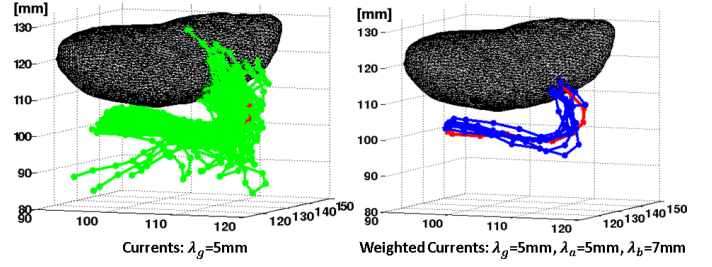

Fig. 2. Streamlines that have an angle smaller than 45 degrees with the red one using currents (green, #118) and weighted currents (blue, #8). Green streamlines are more spread than the blue ones, connecting anatomical locations far from the ones of the red fiber. The concept of similarity in the framework of weighted currents is more stringent than using currents.
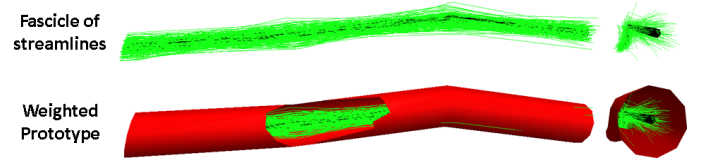

Fig. 3. Weighted prototype visualised as a red tube. It approximates the fascicle of streamlines coloured in green. The streamline chosen as prototype is the central axis of the tube and it is coloured in black. The radius of the tube is proportional to the weight of the prototype and it does *not* represent the spatial coverage of the prototype. The spatial coverage is the same for all prototypes and it depends on the three bandwidths λ_g, λ_a and λ_b.

a) Fascicles detection: A fascicle is a group of streamlines which are considered similar in the framework of weighted currents, namely they have a similar pathway and end-points close to each other. The subdivision of a bundle into fascicles is based on the maximization of a quality function called *modularity* [39]:

$$Q = \sum_{F=1}^{N_F} \left\{ \|\sum_{i \in F} S_i\|_{W^*}^2 \|\sum_{j \notin F} S_j\|_{W^*}^2 - (\sum_{i \in F} \sum_{j \notin F} \langle S_i, S_j \rangle_{W^*})^2 \right\}$$
$$(3)$$

where F is a fascicle, N_F is the number of fascicles and it is constrained by $\sum_{F=1}^{N_F} \sum_{i \in F} S_i = B$. In the simple case of $N_F=2$, Eq.3 can be rewritten as: $Q = \|\bar{S}_1\|_{W^*}^2 \|\bar{S}_2\|_{W^*}^2 - \langle \bar{S}_1, \bar{S}_2 \rangle_{W^*}^2$ where $\bar{S}_1$ and $\bar{S}_2$ are the averages of the two fascicles. Maximizing Q means therefore dividing the bundle into two fascicles such that their averages have a similar norm and tend to be orthogonal to each other. In the general form of Eq.3, one looks for N_F fascicles with balanced norms and which tend to have streamlines orthogonal to the streamlines of the other fascicles and parallel to the streamlines of their own fascicle.

Modularity is often employed in the field of complex networks to detect densely connected communities of nodes within a network [39]. It has been demonstrated that exact modularity optimization is strongly NP-complete [40]. Several approximation schemes exist in the literature and one of the state-of-the-art methods is the "Louvain" algorithm [39]. It is a greedy solution where every fascicle is considered as a vertex of a graph. Two vertices F_1 and F_2 have a weighted

edge equal to the sum of the inner products between the streamlines of the fascicles $\sum_{i \in F_1} \sum_{j \in F_2} \langle S_i, S_j \rangle_{W^*}$. At the beginning, every streamline is considered as an independent fascicle. The algorithm is divided into two parts which are repeated iteratively. The first part consists of moving all the streamlines of a vertex to its neighbour vertices finding the relocation that leads to the greatest increase in modularity. If none movement produces a positive gain in modularity, the streamlines remain in their initial vertex. This part is repeated until no change would produce an increase in modularity. In the second part, one redefines the graph by discarding the empty vertices and recomputing the weighted edges between the new vertices. The two steps are repeated until no change would produce an increase in modularity. At the end, the fiber bundle is separated into different fascicles without fixing in advance neither the number of fascicles nor the number of streamlines per fascicle.

Remark: This step could be seen as a clustering. However, the word "clustering" is often employed in the literature with a connotation of "segmentation", namely the subdivision of a white matter tractogram into tracts reproducible across subjects. Here, the goal is not to create anatomically relevant clusters but to subdivide the multi-modal distribution of fibers into different fascicles (modes). In order not to create ambiguity about this step, we avoided the word "clustering".

b) Prototypes Streamline Selection (PSS): Once the fascicles are defined, a PSS is performed in each fascicle independently. We propose an iterative algorithm in the spirit of orthogonal matching pursuit [41]. Let F be a fascicle with L streamlines modelled as weighted currents, the first prototype P_1 is chosen as the streamline minimising the residual squared error, namely: $P_1 = \operatorname{argmin}_{S_i} ||F - \tau_1 S_i||^2_{W^*}$. Since the space of weighted currents is a vector space, we can easily minimize it. The optimal weight is: $\tau_1 = \frac{\langle F, P_1 \rangle_{W^*}}{||P_1||^2_{W^*}}$ and the prototype is: $P_1 = \operatorname{argmax}_{S_i} \langle F, \frac{S_i}{||S_i||_{W^*}} \rangle^2_{W^*} = \operatorname{argmax}_{S_i} L^2 \langle \bar{S}, \frac{S_i}{||S_i||_{W^*}} \rangle^2_{W^*}$ with $i=1,...,L$. The prototype is therefore the most parallel streamline to the average $\bar{S}$ of the fascicle.

Once the first prototype is selected, we remove from each streamline S_i its orthogonal projection onto the prototype, resulting in the residual: $r(S_i) = S_i - \pi(S_i) = S_i - \frac{\langle S_i, P_1 \rangle_{W^*} P_1}{||P_1||^2_{W^*}}$. We keep therefore only the components of the streamlines orthogonal to the prototype P_1. In this new space, we select the second prototype as: $P_2 = \operatorname{argmax}_{r(S_i)} \langle r(F), \frac{r(S_i)}{||r(S_i)||_{W^*}} \rangle^2_{W^*}$. In this way, P_1 and P_2 should be very different as well as the streamlines they approximate. We iterate this process until: $||F - \sum_{k=1}^{K} \tau_k P_k||_{W^*} \leq \gamma ||F||_{W^*}$ where $||F||_{W^*}$ is the norm of the fascicle, K is the number of prototypes and γ indicates the required approximation level. At each iteration t, the set of weights $\{\tau_k\}_{k=1,...,t}$ is computed as the orthogonal projection of all the streamlines of F to the space spanned by the selected set of prototypes $\{P_k\}_{k=1,...,t}$. The final number of prototypes depends on the chosen approximation level γ and on the kernel bandwidths of the framework of weighted currents.

It is important to notice that the computations are based on the Gram matrix Γ of the fascicle F which has size $[LxL]$. Thus, instead of computing directly $r(S_i)$, we simply update

Γ as: $\langle r(S_i), r(S_j) \rangle_{W^*} = \langle S_i, S_j \rangle_{W^*} - \frac{\langle S_i, P \rangle_{W^*} \langle S_j, P \rangle_{W^*}}{||P||^2_{W^*}} = \Gamma_{(i,j)} - \frac{\Gamma_{(i,P)} \Gamma_{(j,P)}}{||\Gamma_{(P,P)}||^2_2}$. A sketch of the algorithm can be found in *Algorithm* 1 where $\Gamma_{(i,j)}$ indicates the value of the matrix Γ at row i and column j, $\Gamma_{(K,L)}$ refers to the submatrix of Γ containing the K rows of the prototypes and all the L columns, $\Gamma_{(K,K)}$ is the square submatrix with the rows and columns of the K prototypes and $\mathbf{1}_{(L,1)}$ is a L-dimensional vector of ones.

Algorithm 1 Prototype Fiber Selection

Input: Fascicle F with L streamlines S_i: $F = \sum_{i=1}^{L} S_i$; threshold γ ; pre-computed matrix Γ ; K=1

1: $P_1 \leftarrow \operatorname{argmax}_{S_i} \langle F, \frac{S_i}{||S_i||_{W^*}} \rangle^2_{W^*}$
2: $\tau_1 \leftarrow \frac{\langle F, P_1 \rangle_{W^*}}{||P_1||^2_{W^*}}$
3: **for** $i = 1$ **to** L **do**
4:     **for** $j = 1$ **to** L **do**
5:         $\langle S_i, S_j \rangle_{W^*} \leftarrow \langle S_i, S_j \rangle_{W^*} - \frac{\langle S_i, P_1 \rangle_{W^*} \langle S_j, P_1 \rangle_{W^*}}{||P_1||^2_{W^*}}$
6:     **end for**
7: **end for**
8: **while** $||F - \sum_{k=1}^{K} \tau_k P_k||_{W^*} \leq \gamma ||F||_{W^*}$ **do**
9:     $K \leftarrow K + 1$
10:    $P_K \leftarrow \operatorname{argmax}_i \frac{(\Gamma_{(i,L)} \mathbf{1}_{(L,1)})^2}{\Gamma_{(i,i)}}$
11:    $\{\tau_k\}_{k=1,...,K} \leftarrow \frac{\Gamma_{(K,L)} \mathbf{1}_{(L,1)}}{\Gamma_{(K,K)}}$
12:    **for** $i = 1$ **to** L **do**
13:       **for** $j = 1$ **to** L **do**
14:          $\Gamma_{(i,j)} \leftarrow \Gamma_{(i,j)} - \frac{\Gamma_{(i,P_K)} \Gamma_{(j,P_K)}}{||\Gamma_{(P_K,P_K)}||^2_2}$
15:       **end for**
16:    **end for**
17: **end while**
Output: $\{\tau_k\}$, $\{P_k\}$ k=1,...,K

After selecting the prototypes of each fascicle independently, they are all gathered into a single bundle of prototypes B_P. The weights are then recomputed as the orthogonal projection of the whole bundle B to the entire set of prototypes B_P in order to retrieve the correct values also for the weights of the prototypes close to the boundary between two different fascicles. Moreover, before the PSS algorithm, we also perform an outlier detection step in every fascicle. The streamlines characterised by an average angle with the other streamlines between $88°$ and $90°$ are considered as outliers and discarded from the analysis.

In Fig.4, we show a sketch of the entire approximation scheme. A fiber bundle is first divided into fascicles. The subdivision is not meant to be related to anatomy. Its goal is purely algorithmic. It permits the detection and the approximation of all fascicles, even the small ones. This is crucial, for instance, when registering two bundles in order to correctly align all the fascicles and not only the central and bigger ones. The size and number of fascicles is determined by the bandwidths λ_g, λ_a and λ_b. The smaller the bandwidths, the greater the number of fascicles. In fact, by decreasing these values we make the definition of similarity between two streamlines more stringent. For instance, if their extremities are more distant than 3 times λ_b, they will be considered orthogonal (i.e. far from each other). At the end of the PSS step, if the approximation level is sufficiently small, the resulting weighted prototypes
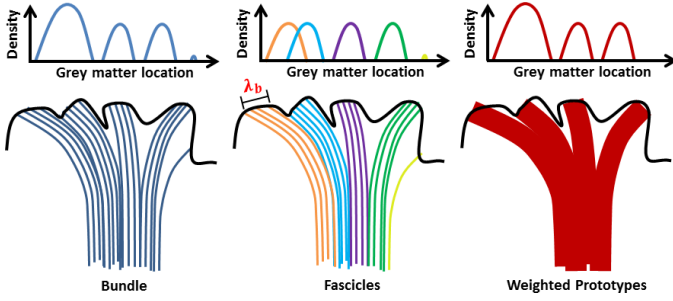
Fig. 4. Sketch of the approximation scheme on a toy bundle composed of blue streamlines and a black cortex. The first row shows the structural connectivity computed as the density of the streamlines extremities onto the grey matter. The second row presents the results of the two steps. At the end of the first step, the bundle is divided into different fascicles. Streamlines considered as outliers (i.e. the yellow one on the right) are discarded before the PSS step. The resulting weighted prototypes preserve the original structural connectivity.
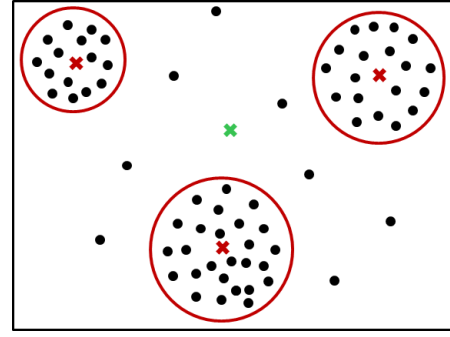

Fig. 5. Visual explanation of the bundle subdivision into fascicles (modes). Dots and crosses represent the streamlines of a fiber bundle in the space of weighted currents. The green cross is the most similar streamline to the average of the bundle. It is far from almost all the other streamlines and it could be considered as an outlier. The three circles represent the fascicles (modes) composed by similar fibers. The red crosses are the prototypes of the fascicles considered independently. These fibers are more representative than the green cross and they better approximate the fiber bundle.

should preserve both the shape and the structural connectivity of the original bundle. Structural connectivity is computed as the density of the streamlines/prototypes extremities onto the grey matter surface. When using the prototypes, we also take into account their weights. We use kernel density estimation with a Gaussian kernel.

Remark: Performing the PSS in each fascicle independently allows us to distribute the computations to different processors, reducing therefore the computational time. Moreover, it also decreases the chance to select as prototype a streamline which might be considered an outlier. This is explained in Fig.5 where every dot represents a streamline of a bundle modelled as weighted current. We oversimplify this space assuming it is simply $\mathbb{R}^2$. In this space, it is likely that a tractography bundle has a multi-modal distribution, where every mode is a fascicle. If we wanted to approximate the whole bundle with a single prototype, it would be the most parallel streamline to the average of the bundle. In Fig.5, we would choose the streamline represented by the green cross. This fiber is far from almost all the other streamlines and it could be considered as an outlier. Instead, if we applied the same selection process in each fascicle independently, we would obtain the three prototypes highlighted in red. These streamlines are more representative than the green fiber and they better approximate the bundle. Obviously, an actual fascicle does not lie in a 2D space and therefore we may need more than one prototype to explain its variability.

V. EXPERIMENTS AND RESULTS

In this section, we first describe the dataset used in the following experiments and some technical details about the implementation of the proposed algorithm. Then, we present the approximation of two probabilistic bundles and we show that their structural connectivity is similar to the one of the original bundles. Furthermore, we evaluate from a qualitative point of view how a variation in the parameters of the algorithm affects the approximation of a deterministic bundle. Eventually, we assess the effectiveness of our algorithm showing that the registration between two approximated fiber

bundles is definitely faster than using the original bundles for the same registration accuracy.

A. Materials

We test the proposed approximation scheme on 25 subjects. Diffusion-weighted (DW) scans are acquired with sequences of 50 directions with a B-factor of 1000 and a voxel size of 2x2x2 mm^3. We use the Spherical Deconvolution Transform (SDT) model [42] to estimate the local underlying orientation distribution function (ODF). Whole brain connectivity is then inferred within an anatomy-based tractography mask [24] using both a deterministic (1 seed per voxel) and a probabilistic (8 seeds per voxel) tractography algorithm available in BrainVISA/Connectomist-2.0 [43]. In this paper, we consider three distinct fiber bundles connecting the left hemisphere of the cortical surface to the left thalamus, putamen and caudate respectively. We extract them from both the deterministic and probabilistic whole brain tractography as explained in [44]. All bundles also include the commissural fibers which are truncated at the inter-hemispheric plane. The other streamlines are cut at the intersection with their respective sub-cortical structure and at the border between white and gray matter of the cortex. Sub-cortical structures are segmented with FSL [45] from 3D T_1-weighted images (voxel size: 1x1x1 mm^3) and we merge the segmentations of nucleus accumbens and caudate in order to consider them as a single structure. The 3D meshes are created using the marching cubes algorithm of BrainVISA v4.4.0. The cortical surface is segmented using FreeSurfer v5.3 [46]. More information about acquisition and preprocessing of both T1-w and DW images can be found in [44].

B. Numerical aspects

The parameters needed to be fixed by the user are the bandwidths of the three kernels of weighted currents λ_g, λ_a and λ_b and the approximation level γ. In the following, λ_g refers to the kernel of currents, λ_a to the end-point on the nuclei and λ_b to the end-point on the cortical surface.

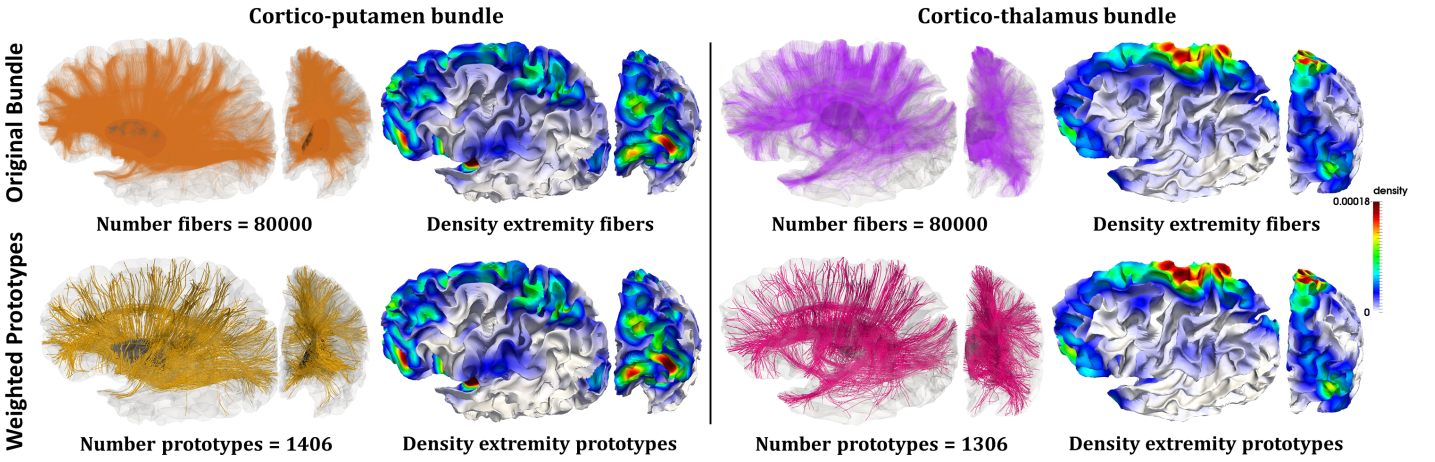**Cortico-putamen bundle**   **Cortico-thalamus bundle**

Fig. 6. Weighted prototype approximations of two probabilistic bundles: a cortico-putamen and a cortico-thalamus. As it is possible to notice, our approximation alters neither the global shape of the bundle nor the densities of the endpoints onto the cortical surface. We use: γ=0.13, λ_g=7mm, λ_a=5mm and λ_b=10mm.

All experiments shown in this paper are computed on a Intel Xeon, 8 cores, CPU E5-1620, 3.60 GHz with a memory of 16Gb and a graphic card NVIDIA Quadro K4000.

The computational times for the approximations of the fiber bundles shown in Fig.6 and in Fig.8, composed of 80.000 and 35.674 streamlines, were of 150 and 19 minutes respectively.

C. Weighted prototypes representation

We present in Fig.6 the weighted prototype approximations of two probabilistic bundles using γ= 0.13, λ_g=7mm, λ_a=5mm and λ_b=10mm. The proposed representation preserves the global shape of the bundle and it approximates thoroughly all the fascicles. We also show the densities of the endpoints (i.e. structural connectivity) on the cortical surface computed using either the streamlines of the original bundles or the weighted prototypes. The two densities are very similar from a qualitative point of view and the Kolmogorov-Smirnov test fails to show statistically significant differences between them at the 5% level. Probability densities are computed using Gaussian kernels, taking into account the weights of the prototypes for the proposed approximation. Furthermore, we use the same parameters to approximate all the bundles of our data-set (75 deterministic and 75 probabilistic) and in no case the density of the endpoints is statistically significantly different from the one of the original bundle at the 5% level. The average compression ratios 100(1- K/N), where N is the number of streamlines of the original bundles and K is the number of prototypes, are shown in Table I and II for the deterministic and probabilistic bundles respectively. In Fig.7, we also show the evolution of the average compression ratio and number of prototypes for different approximation levels. These results show that our algorithm leads to a much more compact representation than the original bundle while preserving the overall structural connectivity.

D. Qualitative evaluation of the parameters influence

In Fig.8 we evaluate the influence of the parameters of our algorithm on the approximation of a deterministic

Deterministic			
Bundle	Cortex-caudate	Cortex-putamen	Cortex-thalamus
N (Streamlines)	17079±4881	28056±5247	28371±6806
K (Prototypes)	344±58	409±55	341±87
Compression	97.85 %	98.49 %	98,77 %

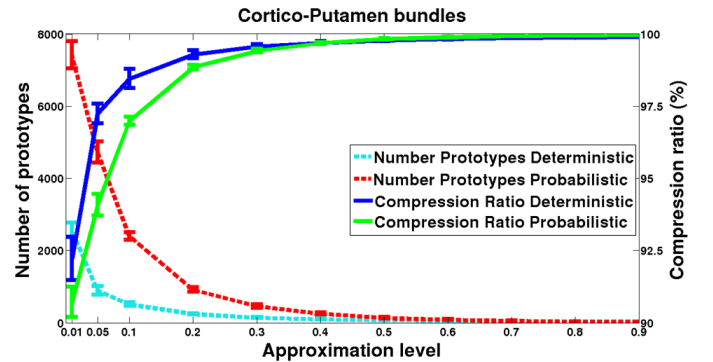| Probabilistic			
Bundle	Cortex-caudate	Cortex-putamen	Cortex-thalamus
N (Streamlines)	75389±4646	78125±2223	68640±8568
K (Prototypes)	1182±358	1411±393	1000±319
Compression	98,41 %	98,19 %	98,54 %

Fig. 7. Evolution of the average number of prototypes and compression ratio at different approximation levels for 5 deterministic and 5 probabilistic cortico-putamen bundles. Bars represent one standard deviation.

cortico-putamen bundle. In the first row we employ λ_g=5mm, λ_a=4mm, λ_b=6mm and in the second row λ_g=7mm, λ_a=5mm, λ_b=10mm. Every column corresponds to a different approx-
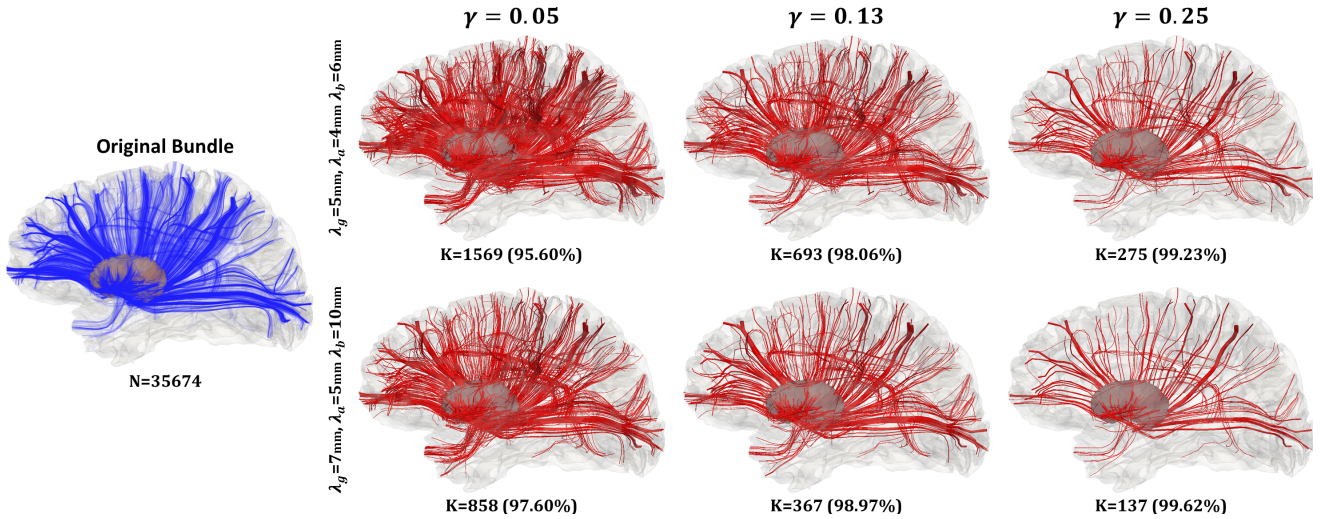
Fig. 8. Weighted prototype representations (in red) at different approximation levels of a deterministic cortico-putamen bundle (in blue). Each row is based on a different set of kernel bandwidths. The letters N and K refer respectively to the number of streamlines of the bundle and to the number of prototypes. The compression ratios are indicated in brackets.
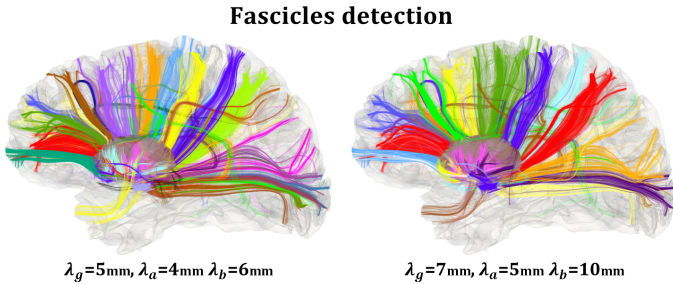

Fig. 9. Results of the fascicles detection step applied to the bundle shown in Fig.8 using the two sets of kernel bandwidths. The number of fascicles are respectively: 65 (left) and 35 (right). Colours are chosen randomly.

imation level. It can be noticed that at γ=0.05 and γ=0.13 all fascicles are well approximated, whatever the set of parameters. At γ=0.25 and higher values of γ (not shown here) only the denser parts of the fascicles are well approximated. Moreover, the results based on the first row use almost twice the prototypes than in the second row. This is expected since the bandwidths of the kernels are smaller and therefore the definition of similarity between two streamlines is more stringent. Thus, for a given γ, one needs more prototypes to approximate the same number of streamlines. Furthermore, this also influences the number/size of the fascicles, as shown in Fig.9. The fascicles obtained with the first set of parameters are smaller in size and greater in number than the ones obtained using the second set of parameters.

E. Registration-based evaluation of the algorithm

Here we evaluate the impact of the proposed approximation scheme on the quality of a registration between two deterministic cortico-putamen bundles of different subjects. We use the diffeomorphic transformation implemented in the software *Deformetrica* (www.deformetrica.org). It is based on a control points formulation [35] of the Large Deformation

Diffeomorphic Metric Mapping (LDDMM) framework [47]. In the following experiments, we use a diffeomorphic kernel bandwidth equal to 10mm and 1309 control points. The source bundle B_S is the one shown in Fig.8. First, we approximate both B_S and the target bundle B_T, composed of 35674 and 25916 streamlines respectively, at different approximation levels using λ_g=7mm, λ_a=5mm, λ_b=10mm. Then, for each level, we register the approximation of B_S onto the one of B_T. After that, we apply the obtained deformation to the original bundle B_S and we compute the residual error between the transformed original source bundle $\phi(B_S)$ and the original target bundle B_T: $\|\phi(B_S) - B_T\|^2_{W_*}$ in the framework of weighted currents. Ideally, we would compare this residual error with the one resulting from the registration of the original fiber bundles. Unfortunately, the computational time would be too long (see Table III). Thus, we decide to register smaller sub-samples of the original fiber bundles and then to use the resulting residual errors for comparison. In Fig.10, we apply two different deformations to the original source bundle. The first one results from the registration of the prototype approximations with γ=0.13. The second one is instead obtained by matching the sub-samples of 5000 streamlines of the original bundles. It is possible to notice that the results look very similar. This is confirmed in Table III where we show that the difference between their residual errors is very small. We can conclude that the registration based on the approximation at γ=0.13 is as accurate as using the sub-sample of 5000 streamlines but 93 times faster! We also present the results for other approximation levels. Compared to γ=0.13, the other registrations are either less accurate or slower and with a similar accuracy.

Remark: It is important to highlight that the proposed approximation scheme is general and that it could have been tested with other registration approaches (i.e. affine transformation [13], [21], [48]). We use the LDDMM framework as one possible example. We expect similar results with other registration methods.
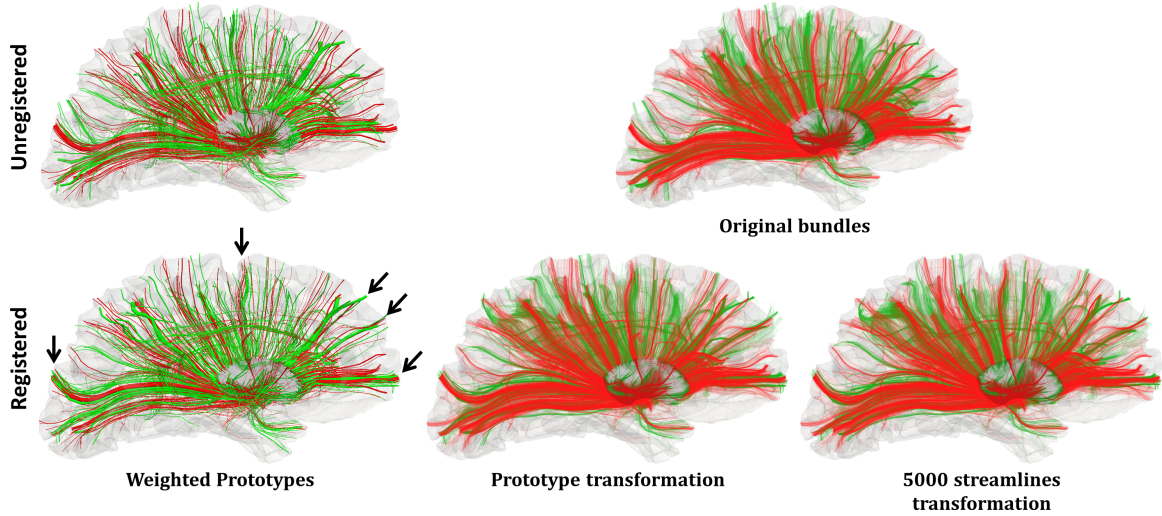
Fig. 10. On the left: registration between the green source bundle and the red target bundle approximated with weighted prototypes. On the right top: original fiber bundles. On the right bottom: deformed original fiber bundle. The deformation applied on the left results from the registration of the prototype approximations with γ=0.13. The deformation applied on the right results from the matching of the sub-samples of 5000 streamlines of the original bundles. Black arrows highlight the areas where the alignment is more noticeable.

TABLE III
REGISTRATION ERROR AND COMPUTATIONAL TIME USING DIFFERENT
APPROXIMATION LEVELS AND SAMPLES OF THE ORIGINAL BUNDLES.

	Registration Error	Computational Time
γ=0.40	1.75e+09	2h 39min
γ=0.25	1.42e+09	4h 34min
γ=0.13	9.88e+08	5h 51min
γ=0.05	9.79e+08	18h 35min
1000 streamlines	9.96e+08	23h 12min
5000 streamlines	9.94e+08	547h 32min
8000 streamlines	-	$\sim$ 1120h
15000 streamlines	-	$\sim$ 4484h
30000 streamlines	-	∞

VI. DISCUSSION AND CONCLUSIONS

We presented an algorithm to approximate a fiber bundle with a small and scattered set of weighted streamlines prototypes. We tested it on 150 bundles resulting from both deterministic and probabilistic tractography algorithms. The number of prototypes was on average 2% of the number of streamlines of the original bundles. We showed that such a parsimonious representation preserves the shape of the bundles and it can be used to reconstruct their original structural connectivity, for sufficiently small approximation levels (γ). It is interesting to notice that the density of streamlines on grey matter surfaces could be also stored in a connectivity matrix. This would provide an efficient encoding of the structural connectivity but it would discard the trajectory (i.e. geometry) of the streamlines. Moreover, it would also require the definition of a parcellation scheme for grey matter surfaces. This might be a tricky task since, to date, there is still not a globally accepted parcellation of the grey matter in the scientific community. On the contrary, our approach could be used to drive an adapted parcellation in future works.

The streamlines considered in this paper have been truncated at the intersection between gray and white matter which is an area usually characterised by a low Signal to Noise Ratio. This means that the estimates of the end-points of the streamlines are not very robust. To account for this uncertainty, we use Gaussian kernels for measuring the dissimilarity between two streamlines. Theoretically, the greater the uncertainty, the greater the bandwidths of the kernels. In the proposed algorithm, these bandwidths (i.e. λ_a and λ_b) are considered as parameters fixed by the user. Their values are chosen by looking at how much the streamlines fan out when approaching to the boundary between white and grey matter. Streamlines deviate more when they are close to the cortex than to the sub-cortical nuclei, which explains why λ_b is always greater than λ_a in our experiments. It would be of interest to automatically estimate these parameters taking into consideration the type of bundle, the SNR of the diffusion image, the tractography algorithm and the diffusion model.

Another parameter fixed at the beginning of the algorithm is the approximation level γ. It defines the stopping criteria and the value $(1 - \gamma)$ is the minimal percentage of the norm of the fascicle explained by the prototypes. In fact, thanks to the triangle inequality, we can rewrite the stopping criteria obtaining $(1-\gamma)\|F\|_{W_*} \leq \|\sum_{k=1}^{K} \tau_k P_k\|_{W_*} \leq (1+\gamma)\|F\|_{W_*}$. This means that, using γ=0.13, the norm of the prototypes will be at least 87% of the norm of the fascicle at the end of the algorithm. Furthermore, we noticed that between γ=0.5 and γ=0.01 the number of prototypes grows exponentially (see Fig.7). The first prototypes to be estimated approximate the parts of the fascicle with an higher density of streamlines (i.e. greater redundancy). Thus, their weights have a great value and few prototypes can explain a considerable percentage of the norm of the fascicle. Instead, between γ=0.1 and γ=0.01, most of the streamlines have already been approximated and every new prototype can explain only a few of the remaining fibers. We found that a value of γ=0.13 results in a parsimonious representation which exhaustively approximates all the fasci-

cles. An interesting improvement would be to automatically estimate the value of γ taking into consideration both the approximation error and the number of prototypes.

The proposed algorithm does not take into account any microstructure measurement such as FA (Fractional Anisotropy). This means that there is no certainty as to whether the approximation would not change the microstructural properties of a bundle. A possible solution would be to augment the framework of weighted currents with a "functional signal" describing the microstructure architecture of the bundle in the same spirit of [49]. Based on this new computational model, we could use analyses such as "tractometry" [50] to investigate the quality of the approximation with respect to microstructure measurements such as FA. In this work, we focused on the ability of the proposed method to preserve the fiber density in specific ROIs, which encodes the structural connectivity.

We showed that our representation can approximate not only the central and denser mass of a bundle, as with currents, but also its smaller fascicles and extremities. We demonstrated its usefulness by registering two approximated bundles where we correctly matched the entire pathway of the weighted prototypes, including their extremities (see Fig.10). Furthermore, bundle registrations based on our approximation scheme present a lower computational time and memory footprint than using the original fiber bundles. This makes thus possible population studies, like the atlas construction [51], based on multiple template-to-subject non-linear registrations, which would be very time-consuming - or even unfeasible - with the original fiber bundles.

APPENDIX A
INTERPRETATION OF τ

We show here that the weight τ of a prototype P is related to the number of fibers approximated by P. Given a bundle B composed of 3 fibers: $B = \sum_{i=1}^{3} S_i$ we want to approximate it with one prototype P. Let assume that S_1 is the prototype P (the reasoning does not change modifying the prototype), the value of its weight τ is $1 + \frac{\|S_2\|_{W^*}^2}{\|P\|_{W^*}^2}\cos(P, S_2)_{W^*} + \frac{\|S_3\|_{W^*}^2}{\|P\|_{W^*}^2}\cos(P, S_3)_{W^*}$. This means that if S_1 is parallel to the other fibers and their norms are similar, the value of τ will be about 3. Instead, if either S_2 or S_3 is orthogonal to S_1, the prototype will not approximate that fiber and the weight will be smaller than 3. This shows that τ is related to the number of fibers approximated by the prototype or, more precisely, to the "amount of similarity" between the prototype and the other fibers of the bundle. When dealing with more prototypes, every τ also depends on the inner product between prototypes. Since we project at each iteration all the streamlines onto the orthogonal space of the last estimated prototype, the inner product between prototypes should be small.

APPENDIX B
MODULARITY BASED ON WEIGHTED CURRENTS

The definition of modularity in [39] is:

$$Q = \sum_{c=1}^{N_C} \left[\frac{W_c}{m} - \left(\frac{S_c}{2m} \right)^2 \right] \qquad (4)$$

where N_C is the number of modules (fascicles), $W_c = 1/2\|\sum_{i \in c} S_i\|_{W^*}^2$ is the sum of the weights (inner products) of all the edges joining only the vertices (fibers) of module c, $S_c = \sum_{i \in c} \sum_{j=1}^{N} \langle S_i, S_j \rangle_{W^*}$ is the sum of the weights of the edges between the vertices in c and the N vertices in the graph (bundle) and $m = 1/2\|\sum_{i=1}^{N} S_i\|_{W^*}^2$ is the sum of the weights of all edges in the graph. Substituting these equations in Eq.4 and noting that m is a constant term and that it can be rewritten as $2m = \|\sum_{i \in c} S_i\|_{W^*}^2 + \|\sum_{j \notin c} S_j\|_{W^*}^2 + 2\sum_{i \in c}\sum_{j \notin c}\langle S_i, S_j\rangle_{W^*}$, one obtains Eq.3.

REFERENCES

[1] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle, "Tracking neuronal fiber pathways in the living human brain," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 18, pp. 10 422–10 427, Aug. 1999. 1

[2] S. Mori and P. C. M. van Zijl, "Fiber tracking: principles and strategies – a technical review," *NMR in Biomedicine*, vol. 15, no. 7-8, pp. 468–480, Nov. 2002. 1

[3] P. J. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *Journal of Magnetic Resonance. Series B*, vol. 103, no. 3, pp. 247–254, Mar. 1994. 1

[4] P. Mukherjee, "Diffusion Tensor Imaging and Fiber Tractography in Acute Stroke," *Neuroimaging Clinics of North America*, vol. 15, no. 3, pp. 655–665, Aug. 2005. 1

[5] O. Ciccarelli, M. Catani, H. Johansen-Berg, C. Clark, and A. Thompson, "Diffusion-based tractography in neurological disorders: concepts, applications, and future developments," *The Lancet Neurology*, vol. 7, no. 8, pp. 715–727, Aug. 2008. 1

[6] M. Kubicki, R. McCarley, C.-F. Westin, H.-J. Park, S. Maier, R. Kikinis, F. A. Jolesz, and M. E. Shenton, "A review of diffusion tensor imaging studies in schizophrenia," *Journal of Psychiatric Research*, vol. 41, no. 1–2, pp. 15–30, Jan. 2007. 1

[7] P. Fillard, M. Descoteaux, A. Goh, S. Gouttard, B. Jeurissen, J. Malcolm, A. Ramirez-Manzanares, M. Reisert, K. Sakaie, F. Tensaouti, T. Yo, J.-F. Mangin, and C. Poupon, "Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom," *NeuroImage*, vol. 56, no. 1, pp. 220–234, May 2011. 1

[8] C. Presseau, P.-M. Jodoin, J.-C. Houde, and M. Descoteaux, "A new compression format for fiber tracking datasets," *NeuroImage*, vol. 109, pp. 73–83, Apr. 2015. 1

[9] L. O'Donnell and C.-F. Westin, "Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas," *IEEE Transactions on Medical Imaging*, vol. 26, no. 11, pp. 1562–1575, Nov. 2007. 1

[10] V. Siless, J. Glaunès, P. Guevara, J.-F. Mangin, C. Poupon, D. L. Bihan, B. Thirion, and P. Fillard, "Joint T1 and Brain Fiber Log-Demons Registration Using Currents to Model Geometry," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, N. Ayache, H. Delingette, P. Golland, and K. Mori, Eds. Springer Berlin Heidelberg, 2012, no. 7511, pp. 57–65. 1

[11] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache, "Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents," *NeuroImage*, vol. 55, no. 3, pp. 1073–1090, Apr. 2011. 1, 2

[12] I. Corouge, S. Gouttard, and G. Gerig, "Towards a shape model of white matter fiber bundles using diffusion tensor MRI," in *IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004*, Apr. 2004, pp. 344–347 Vol. 1. 1

[13] L. J. O'Donnell, W. M. W. Iii, A. J. Golby, and C.-F. Westin, "Unbiased Groupwise Registration of White Matter Tractography," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, N. Ayache, H. Delingette, P. Golland, and K. Mori, Eds. Springer Berlin Heidelberg, 2012, no. 7512, pp. 123–130. 2, 8

[14] I. Corouge, P. T. Fletcher, S. Joshi, S. Gouttard, and G. Gerig, "Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis," *Medical Image Analysis*, vol. 10, no. 5, pp. 786–798, Oct. 2006. 2

[15] A. Brun, H.-J. Park, H. Knutsson, and C.-F. Westin, "Coloring of DT-MRI Fiber Traces Using Laplacian Eigenmaps," in *Computer Aided Systems Theory - EUROCAST 2003*, R. Moreno-Díaz and F. Pichler, Eds. Springer Berlin Heidelberg, 2003, no. 2809, pp. 518–529. 2

[16] B. Tunç, A. R. Smith, D. Wasserman, X. Pennec, W. M. Wells, R. Verma, and K. M. Pohl, "Multinomial Probabilistic Fiber Representation for Connectivity Driven Clustering," in *Information Processing in Medical Imaging*, J. C. Gee, S. Joshi, K. M. Pohl, W. M. Wells, and L. Zöllei, Eds. Springer Berlin Heidelberg, 2013, no. 7917, pp. 730–741. 2

[17] P. G. Batchelor, F. Calamante, J.-D. Tournier, D. Atkinson, D. L. G. Hill, and A. Connelly, "Quantification of the shape of fiber tracts," *Magnetic Resonance in Medicine*, vol. 55, no. 4, pp. 894–903, Apr. 2006. 2

[18] M. K. Chung, N. Adluru, J. E. Lee, M. Lazar, J. E. Lainhart, and A. L. Alexander, "Cosine series representation of 3d curves and its application to white matter fiber bundles in diffusion tensor imaging," *Statistics and its interface*, vol. 3, no. 1, pp. 69–80, 2010. 2

[19] D. Wassermann, L. Bloy, E. Kanterakis, R. Verma, and R. Deriche, "Unsupervised white matter fiber clustering and tract probability map generation: Applications of a Gaussian process framework for white matter fibers," *NeuroImage*, vol. 51, no. 1, pp. 228–241, May 2010. 2

[20] M. Liu, B. C. Vemuri, and R. Deriche, "Unsupervised automatic white matter fiber clustering using a Gaussian mixture model," vol. 2012, no. 9, pp. 522–525, Jul. 2012. 2

[21] O. Zvitia, A. Mayer, R. Shadmi, S. Miron, and H. K. Greenspan, "Co-registration of white matter tractographies by adaptive-mean-shift and Gaussian mixture modeling," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 1, pp. 132–145, 2010. 2, 8

[22] E. Garyfallidis, M. Brett, M. M. Correia, G. B. Williams, and I. Nimmo-Smith, "QuickBundles, a Method for Tractography Simplification," *Frontiers in Neuroscience*, vol. 6, 2012. 2

[23] L. J. O'Donnell, C.-F. Westin, and A. J. Golby, "Tract-based morphometry for white matter group analysis," *NeuroImage*, vol. 45, no. 3, pp. 832–844, Apr. 2009. 2

[24] P. Guevara, C. Poupon, D. Rivière, Y. Cointepas, M. Descoteaux, B. Thirion, and J. F. Mangin, "Robust clustering of massive tractography datasets," *NeuroImage*, vol. 54, no. 3, pp. 1975–1993, Feb. 2011. 2, 6

[25] O. Zvitia, A. Mayer, and H. Greenspan, "Adaptive mean-shift registration of white matter tractographies," in *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008*, May 2008, pp. 692–695. 2

[26] M. Maddah, W. M. Wells, S. K. Warfield, C.-F. Westin, and W. E. L. Grimson, "Probabilistic Clustering and Quantitative Analysis of White Matter Fiber Tracts," *Information Processing in Medical Imaging*, vol. 20, pp. 372–383, 2007. 2

[27] M. Maddah, J. V. Miller, E. V. Sullivan, A. Pfefferbaum, and T. Rohlfing, "Sheet-Like White Matter Fiber Tracts: Representation, Clustering, and Quantitative Analysis," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011*, G. Fichtinger, A. Martel, and T. Peters, Eds. Springer Berlin Heidelberg, 2011, no. 6892, pp. 191–199. 2

[28] P. A. Yushkevich, H. Zhang, T. J. Simon, and J. C. Gee, "Structure-specific statistical mapping of white matter tracts," *NeuroImage*, vol. 41, no. 2, pp. 448–461, Jun. 2008. 2

[29] K. Hua, J. Zhang, S. Wakana, H. Jiang, X. Li, D. S. Reich, P. A. Calabresi, J. J. Pekar, P. C. M. van Zijl, and S. Mori, "Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification," *NeuroImage*, vol. 39, no. 1, pp. 336–347, Jan. 2008. 2

[30] U. Bürgel, K. Amunts, L. Hoemke, H. Mohlberg, J. M. Gilsbach, and K. Zilles, "White matter fiber tracts of the human brain: Three-dimensional mapping at microscopic resolution, topography and inter-subject variability," *NeuroImage*, vol. 29, no. 4, pp. 1092–1105, Feb. 2006. 2

[31] M. Vaillant and J. Glaunès, "Surface Matching via Currents," in *Information Processing in Medical Imaging*, G. E. Christensen and M. Sonka, Eds. Springer Berlin Heidelberg, 2005, no. 3565, pp. 381–392. 2, 3

[32] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, F. D. V. Fallani, M. Chavez, S. Lecomte, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman, "A Prototype Representation to Approximate White Matter Bundles with Weighted Currents," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*, P. Golland, N. Hata, C. Barillot, J. Hornegger, and R. Howe, Eds. Springer International Publishing, Sep. 2014, no. 8675, pp. 289–296. 2

[33] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache, "A statistical model of white matter fiber bundles based on currents," in *Information Processing in Medical Imaging*. Springer Berlin Heidelberg, 2009, pp. 114–125. 3

[34] P. Gori, O. Colliot, Y. Worbe, L. Marrakchi-Kacem, S. Lecomte, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman, "Bayesian Atlas Estimation for the Variability Analysis of Shape Complexes," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds. Springer Berlin Heidelberg, 2013, no. 8149, pp. 267–274. 3

[35] S. Durrleman, M. Prastawa, N. Charon, J. R. Korenberg, S. Joshi, G. Gerig, and A. Trouvé, "Morphometry of anatomical shape complexes with dense deformations and sparse parameters," *NeuroImage*, vol. 101, pp. 35–49, Nov. 2014. 3, 8

[36] S. Durrleman, "Statistical models of currents for measuring the variability of anatomical curves, surfaces and their evolution," Ph.D. dissertation, University of Nice-Sophia Antipolis, 2010. 3

[37] N. Charon and A. Trouvé, "Functional Currents: A New Mathematical Tool to Model and Analyse Functional Shapes," *Journal of Mathematical Imaging and Vision*, vol. 48, no. 3, pp. 413–431, Jan. 2013. 3

[38] J. Glaunès, "Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique," Ph.D. dissertation, Université Paris 13, 2005. 3

[39] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, Oct. 2008. 4, 10

[40] U. Brandes, D. Delling, M. Gaertler, R. Goerke, M. Hoefer, Z. Nikoloski, and D. Wagner, "Maximizing Modularity is hard," *arXiv:physics/0608255*, Aug. 2006, arXiv: physics/0608255. 4

[41] J. A. Tropp and A. C. Gilbert, "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007. 5

[42] M. Descoteaux, R. Deriche, T. Knosche, and A. Anwander, "Deterministic and Probabilistic Tractography Based on Complex Fibre Orientation Distributions," *IEEE Transactions on Medical Imaging*, vol. 28, no. 2, pp. 269–286, Feb. 2009. 6

[43] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Rivière, A. Constantinesco, D. L. Bihan, and J.-F. Mangin, "Fiber Tracking in q-Ball Fields Using Regularized Particle Trajectories," in *Information Processing in Medical Imaging*, G. E. Christensen and M. Sonka, Eds. Springer Berlin Heidelberg, 2005, no. 3565, pp. 52–63. 6

[44] Y. Worbe, L. Marrakchi-Kacem, S. Lecomte, R. Valabregue, F. Poupon, P. Guevara, A. Tucholka, J.-F. Mangin, M. Vidailhet, S. Lehericy, A. Hartmann, and C. Poupon, "Altered structural connectivity of cortico-striato-pallido-thalamic networks in Gilles de la Tourette syndrome," *Brain*, vol. 138, no. 2, pp. 472–482, Feb. 2015. 6

[45] B. Patenaude, S. M. Smith, D. N. Kennedy, and M. Jenkinson, "A Bayesian model of shape and appearance for subcortical brain segmentation," *NeuroImage*, vol. 56, no. 3, pp. 907–922, Jun. 2011. 6

[46] B. Fischl, A. v. d. Kouwe, C. Destrieux, E. Halgren, F. Ségonne, D. H. Salat, E. Busa, L. J. Seidman, J. Goldstein, D. Kennedy, V. Caviness, N. Makris, B. Rosen, and A. M. Dale, "Automatically Parcellating the Human," *Cerebral Cortex*, vol. 14, no. 1, pp. 11–22, Jan. 2004. 6

[47] M. I. Miller, A. Trouve, and L. Younes, "On the Metrics and Euler-Lagrange Equations of Computational Anatomy," *Annual Review of Biomedical Engineering*, vol. 4, no. 1, p. 375, Aug. 2002. 8

[48] E. Garyfallidis, O. Oceguedo, D. Wassermann, and M. Descoteaux, "Robust and efficient linear registration of white-matter fascicles in the space of streamlines," *NeuroImage*, vol. 117, pp. 124–140, Aug. 2015. 8

[49] B. Charlier, N. Charon, and A. Trouvé, "The fshape framework for the variability analysis of functional shapes," *arXiv:1404.6039 [cs, math]*, Apr. 2014. 10

[50] J. Derek and N. Markus, "Tractometry and the hunt for the missing link: a physicist perspective," *Frontiers in Neuroscience*, vol. 9, 2015. 10

[51] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, A. Routier, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman, "Joint morphometry of fiber tracts and gray matter structures using double diffeomorphisms," in *Information Processing in Medical Imaging*. Springer International Publishing, 2015, pp. 275–287. 10