Double diffeomorphism: combining morphometry and structural connectivity analysis

Pietro Gori, Olivier Colliot, Linda Marrakchi Kacem, Yulia Worbe, Alexandre Routier, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman

Double diffeomorphism: combining morphometry and structural connectivity analysis

Pietro Gori*, Olivier Colliot, Linda Marrakchi Kacem, Yulia Worbe, Alexandre Routier, Cyril Poupon, Andreas Hartmann, Nicholas Ayache and Stanley Durrleman

Abstract—The brain is composed of several neural circuits which may be seen as anatomical complexes composed of grey matter structures interconnected by white matter tracts. Grey and white matter components may be modelled as 3D surfaces and curves respectively. Neurodevelopmental disorders involve morphological and organizational alterations which can not be *jointly* captured by usual shape analysis techniques based on single diffeomorphisms. We propose a new deformation scheme, called double diffeomorphism, which is a combination of two diffeomorphisms. The first one captures changes in structural connectivity, whereas the second one recovers the global morphological variations of both grey and white matter structures. This deformation model is integrated into a Bayesian framework for atlas construction. We evaluate it on a data-set of 3D structures representing the neural circuits of patients with Gilles de la Tourette syndrome (GTS). We show that this approach makes it possible to *localise*, *quantify* and *easily visualise* the pathological anomalies altering the morphology and organization of the neural circuits. Furthermore, results also indicate that the proposed deformation model better discriminates between controls and GTS patients than a single diffeomorphism.

Index Terms—shape, morphometry , complex , multi-object , atlas , structural connectivity , Tourette , neural circuits

I. INTRODUCTION

THE pathophysiology of neurodegenerative and neurode-velopmental disorders, such as Parkinson's disease and Gilles de la Tourette syndrome (GTS), often involves mor-phological alterations of the cortico-basal ganglia and cortico-thalamus circuits [1], [2]. These networks are composed of neural projections connecting specific areas of the cortical surface and sub-cortical nuclei. Abnormalities can affect: i) the shape of every component of the circuits from both grey and white matter, ii) the relative position between grey matter structures and iii) the structural connectivity, namely the areas where white matter tracts integrate grey matter structures. Most of the studies present in the literature focus either on the first or on the last point [2], [3]. Few of them analyse the first two points together [4], [5]. In this paper, we propose a new method to tackle all points at the same time.

Every component of the neural circuits may be segmented as a 3D object. Grey matter structures, such as cortical surface and basal ganglia, are represented as surface meshes

P. Gori, O. Colliot, L. Marrakchi-Kacem, A. Routier 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@telecom-paristech.fr

segmented from Magnetic Resonance (MR) T1-w images. Neural projections of the white matter are instead modelled as bundles of 3D streamlines, called fiber bundles, which result from tractography algorithms applied on diffusion MR images. Every streamline is an estimate of the trajectories of large groups of neural axons. The geometrical representation of the neural circuits combines thus both surface and curve meshes into a single multi-object complex, called *shape complex*.

Neural circuits or, more often, parts of them (e.g. only grey or white matter structures) can be analysed using either images, or 3D objects or by combining them together in an iconic-geometric setting [6]–[8]. In this paper, we will focus on the combined analysis of 3D streamlines and 3D surfaces.

A. Related Work

The statistical shape analysis of 3D meshes has been the subject of several works. One of the most popular strategy relies on the selection of consistent correspondences between the structures of the subjects [9]. Correspondences can be, for instance, manually chosen by an expert (i.e. landmarks) [10], estimated using shape descriptors [11]–[13] or found with the iterative closest point (ICP) algorithm (or variants of it [14]). All structures are then aligned to a common reference frame where both mean and covariance matrix can be estimated. Principal Component Analysis (PCA) or Principal Geodesic Analysis (PGA) [15] can be employed to analyse the main morphological variations[1]. This strategy has been successfully employed with several brain structures. However, most of the shape descriptors are conceived for only a particular kind of mesh, i.e. genus-zero surfaces or streamlines. Thus, they can not be used to handle both grey and white matter structures into a single framework. Moreover, this approach may not preserve the anatomical organization of the neural circuits, which means that separated structures may intersect when computing the average or the main morphological variations.

Another class of statistical shape models, which naturally allows the combination of different mesh types, is based on the Grenander's pattern theory [16]. Every shape complex is modelled as a deformation of a reference shape complex called *template complex*. Deformations put into correspondence the components of the template complex with the homologous ones of the subjects. The "amount" of deformation needed to warp the template complex to the subject complex quantifies their morphological differences. The joint estimate of template

---

[1]Note that in [14] authors use a MAP approach based on EM-ICP

complex and deformations is called *atlas construction* [17]–[20]. They capture the common anatomical characteristics and the morphological variability of the population respectively.

Deformations are usually defined as single diffeomorphisms of the entire ambient space which are smooth invertible transformations with smooth inverse. This kind of deformation preserves the anatomical organization of the components of the template complex, namely they can not intersect, fold or shear. Moreover, deformations are defined locally and they can vary across different areas of the ambient space. This makes it possible to capture the variations in relative position between *separate* structures. However, using a single diffeomorphism, one implicitly assumes that the relative position between structures *in contact* with each other or, in practice, close to each other, does not change across subjects. This implies that a particular fiber bundle of the neural circuits should integrate the same areas of the cortical surface and basal ganglia across the whole population. This assumption precludes the study of changes in structural connectivity which could be caused by an abnormal brain development. In Fig.1, we present a toy example composed of a template complex and a subject shape complex characterized by a different structural connectivity. A single diffeomorphism could not put into correspondence all structures and capture the differences in structural connectivity.

Structural connectivity analysis is usually based on the partition of the cortical surface and sub-cortical nuclei in consistent parcels across subjects [21]. Every parcel is considered as a node of a graph and the number of streamlines connecting two nodes (or other quantities such as the projected Fractional Anisotropy) represents the weighted edge. Variability in structural connectivity across subjects can be analysed in each parcel independently or with indexes and methods from the complex network theory [2], [22]. In both cases, the analysis highly depends on the chosen parcellation scheme and it does not take into consideration the morphological variability of grey and white matter structures.
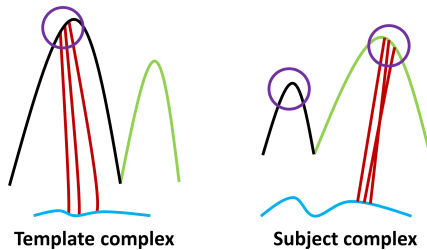

Fig. 1. Two complexes composed of a pseudo cortex, divided into black and green gyrus, a blue sub-cortical nucleus and a red fiber bundle. A single diffeomorphism could not put into correspondence all structures and capture the differences in structural connectivity. The points within the violet circle in the template complex would be matched either to the black gyrus of the subject complex or to the red fiber bundle. A double diffeomorphism would first move the fiber bundle from the left to the right gyrus and then it would change the shape of all structures, producing an accurate matching and capturing also the dissimilarities in structural connectivity.

B. Our contribution

In this paper, extending [23], we propose to join together *shape* and *structural connectivity* analysis into a unified framework based on a *double diffeomorphic* atlas construction. The

template complex is warped towards every shape complex of the population using a composition of two diffeomorphisms. The first diffeomorphism acts only on the white matter of the template complex, keeping fixed the grey matter. During this transformation, the fiber bundles are repositioned with respect to the grey matter structures, capturing the variations in structural connectivity. The second diffeomorphism acts on the whole template complex, namely on both the resulting white matter and grey matter, bringing all structures of the template complex into the subject's space. White matter tracts are rearranged by the first diffeomorphism so that the second one can correctly put into correspondence all the components of the template complex. The two diffeomorphisms are optimised together minimising a single cost function. The data-term only depends on the deformed template complex resulting from the second diffeomorphism. Using again the example in Fig.1, the first diffeomorphism would move the fiber bundle from the left gyrus to the right one. The second diffeomorphism would then modify the shape of all structures producing an accurate matching. The first diffeomorphism would capture the changes in structural connectivity, whereas the second one would recover the global morphological differences. Both diffeomorphisms are parametrized using control points as proposed in [24]. The number of control points is fixed by the user and their position is automatically adjusted during the atlas construction. To note that, we estimate two distinct deformation fields (no composition is performed) and that smoothness across white and grey matter is guaranteed by the fact that they are jointly deformed only by the second diffeomorphism.

Our approach is different from other multi-diffeomorphic methods with sliding conditions such as [25]–[27]. These methods aim to correctly register longitudinal scans or anatomical complexes characterized by sliding regions. Every region is smoothly and *independently* deformed. Contrary to that, we are interested in studying the *relative variation* of one region, white matter, with respect to another one, grey matter. The aforementioned sliding registrations, if applied to the example shown in Fig.1, would result into two independent deformations, one for the white matter and one for the grey matter. It would be thus impossible to understand whether the deformation of the white matter is due to a difference in grey matter or to a variation in structural connectivity. Furthermore, the proposed method differs from multi-scale diffeomorphisms, such as [28], [29], which combine multiple kernels at different scales to create one single diffeomorphic deformation. In this case, the goal is mainly to improve the registration accuracy and remove the scale tuning.

In order to deal with the considerable amount of streamlines resulting from tractography algorithms, we rely on the parsimonious representation, based on weighted prototypes, introduced in [30]. Both prototypes and streamlines are modelled as weighted currents [30]. This model is well suited for any kind of fiber bundle, both sheet-like [31] and tubular [12]. Furthermore, we propose to model grey matter structures as varifolds [32], [33], the non-oriented extension of the framework of currents [7], [34], or landmarks, if correspondences across subjects are available. The atlas is estimated within a Bayesian

framework based on a generative model similar to the one proposed in [35]–[37] and adapted to double diffeomorphisms.

The paper is organized as follows. We first describe the double diffeomorphic generative model and the Bayesian framework for atlas construction. We initially model all shapes with landmarks and then we describe how to integrate the computational models of varifolds and weighted currents. We conclude Sec.II showing how to compute the diffeomorphic deformations and with a description of the optimization procedure. Eventually, we evaluate the discriminative power of the proposed method to distinguish between a population of controls and one of GTS patients. We also compare the resulting classification scores with the ones obtained using a single diffeomorphism.

II. METHODS

A. Double Diffeomorphic Generative Model

The proposed atlas construction is based on a generative statistical model. We assume that the population under study is composed of N subjects. Every subject i is characterised of $M = M^G + M^W$ 3D discrete geometric representations (points, polylines or polygon meshes) from both grey (M^G) and white matter (M^W). We define the representation of structure j belonging to subject i as S_{ij}. Every subject shape complex $\boldsymbol{S}_i$, defined for the moment in a generic way as the ensemble of all meshes S_{ij}, is modelled as a *double deformation* of a common template complex $\boldsymbol{T}$ plus a residual noise $\boldsymbol{\epsilon}_i$. Both $\boldsymbol{T}$ and $\boldsymbol{\epsilon}_i$ are also defined as the ensembles of the templates T_j and residuals ϵ_{ij}. The first deformation ϕ^W acts only on the white matter structures of the template complex: $\boldsymbol{T}^W$. The grey matter of the template complex $\boldsymbol{T}^G$ is not modified. The second deformation ϕ^{All} deforms both the resulting white matter $\phi^W(\boldsymbol{T}^W)$ and $\boldsymbol{T}^G$. This formulation derives from the forward model [35], [38], [39] where we assume that all elements belong to an algebraic structure where addition is defined. It results:

$$\boldsymbol{S}_i = \phi_i^{All}\left(\phi_i^W(\boldsymbol{T}^W) \cup \boldsymbol{T}^G\right) + \boldsymbol{\epsilon}_i \tag{1}$$

The two deformations, ϕ_i^W and ϕ_i^{All}, are two diffeomorphisms of the entire ambient space. They follow one another creating a cascade of diffeomorphisms. White matter streamlines of $\boldsymbol{T}^W$ are re-positioned by ϕ_i^W within the grey matter $\boldsymbol{T}^G$, which is kept fixed. This can be seen as a relative change of coordinates with respect to $\boldsymbol{T}^G$, which is considered as a fixed reference frame common to all subjects i. The entire template complex, both $\boldsymbol{T}^G$ and $\phi_i^W(\boldsymbol{T}^W)$, are then registered to the subject shape complex $\boldsymbol{S}_i$ by ϕ_i^{All}. This is instead a global change of coordinates which brings the template complex to the subject space. The two deformations, ϕ_i^W and ϕ_i^{All}, capture the differences in structural connectivity and the global morphological changes, common to both white and grey matter, respectively. A diagram based on the toy example of Fig. 1 can be found in the Appendix.

B. Bayesian Atlas Construction

The goal of the atlas construction is to estimate the template complex $\boldsymbol{T} = \boldsymbol{T}^W \cup \boldsymbol{T}^G$, the variations in structural connectivity within the population described by the ensemble of first diffeomorphisms $\{\phi_i^W\}$ and the global morphological variations captured by the second diffeomorphisms $\{\phi_i^{All}\}$. Both diffeomorphisms are parametrized by a set of parameters, $\boldsymbol{\alpha}_i^W$ and $\boldsymbol{\alpha}_i^{All}$ respectively, specific to every subject i. We assume that these parameters follow a Gaussian distribution with zero mean and covariance matrix equal to Γ_α^W and Γ_α^{All} respectively: $\boldsymbol{\alpha}_i^W \sim N(0, \Gamma_\alpha^W)$ and $\boldsymbol{\alpha}_i^{All} \sim N(0, \Gamma_\alpha^{All})$. Moreover, as usual in statistical learning, we assume that the residuals follow a Gaussian distribution centered at 0 and with a scalar matrix as covariance matrix ($\epsilon_{ij} \sim N(0, \sigma_j^2 \mathbb{1}_{\Lambda_j})$). For now, we model all structures with landmarks, which are 3D points reproducible among subjects that establish a point-correspondence. For every subject, structure j is modelled using Λ_j landmarks. Thus, $\mathbb{1}_{\Lambda_j}$ is the identity matrix of size Λ_j. The norm of the difference between two meshes is the L^2-norm ($|| \cdot ||_2$, i.e. the sum of squared differences between corresponding landmark pairs). The likelihoods of the residuals of white (W) and grey (G) matter structures modelled as landmarks are:

$$p(\epsilon_{ij}^W | \sigma_j^W) \propto \frac{1}{|(\sigma_j^W)^2|^{\Lambda_j/2}} \exp\left[-\frac{1}{2(\sigma_j^W)^2}||S_{ij} - \phi_i^{All}\left(\phi_i^W(T_j^W)\right)||_2^2\right]$$

$$p(\epsilon_{ij}^G | \sigma_j^G) \propto \frac{1}{|(\sigma_j^G)^2|^{\Lambda_j/2}} \exp\left[-\frac{1}{2(\sigma_j^G)^2}||S_{ij} - \phi_i^{All}\left(T_j^G\right)||_2^2\right] \tag{2}$$

where σ_j^G and σ_j^W refer to grey and white matter structures respectively. In the following, we refer to σ_j when we make no distinction between grey and white matter structures. In Sec.II-C, we will make clear how to adapt these equations when modeling a structure as weighted current or varifold. Whatever the model employed, the variance only depends on the structure-dependent parameter σ_j^2. Moreover, from Eq.1 and Eq.2, it follows that all shapes S_{ij} follow a Gaussian distribution: $S_{ij}^W \sim N(\phi_i^{All}\left(\phi_i^W(T_j^W)\right), (\sigma_j^W)^2 \mathbb{1}_{\Lambda_j})$, $S_{ij}^G \sim N(\phi_i^{All}\left(T_j^G\right), (\sigma_j^G)^2 \mathbb{1}_{\Lambda_j})$. The two covariance matrices of the deformation parameters, Γ_α^W and Γ_α^{All}, are also considered as parameters of the model. We can thus reformulate the goal of the atlas construction as estimating $\boldsymbol{T}$, Γ_α^W and Γ_α^{All}, knowing the shape complexes $\{S_{ij}\}$ and assuming they follow a Gaussian distribution. This can be achieved by maximizing the joint posterior distribution of $\boldsymbol{T}$, σ_j^2, Γ_α^W and Γ_α^{All}. Assuming independence between all random variables and considering $\{\boldsymbol{\alpha}_i^W\}$ and $\{\boldsymbol{\alpha}_i^{All}\}$ as hidden variables, it results:

$$\{\boldsymbol{T}^*, \Gamma_\alpha^{W*}, \Gamma_\alpha^{All*}, \sigma_j^{2*}\} = \underset{\boldsymbol{T}, \Gamma_\alpha^W, \Gamma_\alpha^{All}, \sigma_j^2}{\arg\max} \tag{3}$$

$$\prod_i^N \prod_j^M \int \int p(T_j, \Gamma_\alpha^W, \Gamma_\alpha^{All}, \sigma_j^2, \boldsymbol{\alpha}_i^W, \boldsymbol{\alpha}_i^{All}, S_{ij}) d\boldsymbol{\alpha}_i^{All} d\boldsymbol{\alpha}_i^W$$

Not using priors for Γ_α^W, Γ_α^{All} and σ_j^2 can produce degenerate estimates with small training data-sets, as demonstrated in [35]. A possible solution is to regularize the estimates using adapted versions of the inverse Wishart distributions as priors $\sigma_j^2 \sim \mathcal{W}^{-1}(P_j, w_j)$, $\Gamma_\alpha^W \sim \mathcal{W}^{-1}(P_\alpha^W, w_\alpha^W)$, $\Gamma_\alpha^{All} \sim \mathcal{W}^{-1}(P_\alpha^{All}, w_\alpha^{All})$:

$$p(\sigma_j^2) \propto (\sigma_j^2)^{-\frac{w_j}{2}} \exp\left[-\frac{1}{2}\frac{w_j P_j}{\sigma_j^2}\right] \qquad (4)$$

$$p(\Gamma_\alpha^W) \propto |\Gamma_\alpha^W|^{-\frac{w_\alpha^W}{2}} \exp\left[-\frac{1}{2}w_\alpha^W Tr((P_\alpha^W)^T(\Gamma_\alpha^W)^{-1})\right]$$

$$p(\Gamma_\alpha^{All}) \propto |\Gamma_\alpha^{All}|^{-\frac{w_\alpha^{All}}{2}} \exp\left[-\frac{1}{2}w_\alpha^{All} Tr((P_\alpha^{All})^T(\Gamma_\alpha^{All})^{-1})\right]$$

The scalars w_j, P_j, w_α^W and w_α^{All} are strictly positive and P_α^W and P_α^{All} are positive symmetric matrices. They are hyper-parameters fixed by the user (see Sec.II-E to get more insight). Since the maximization of Eq.3 is not tractable analytically, we use the EM (Expectation Maximization) algorithm where we approximate the conditional distribution of the E step with a Dirac distribution at its mode. See [35], [37] for more information about the E and M step. Assuming that the template T has a non-informative prior distribution, it results:

$$\boxed{\sum_{j=1}^{M^W}\sum_{i=1}^{N}\frac{1}{2(\sigma_j^W)^2}\left(||S_{ij}^W - \phi_i^{All}(\phi_i^W(T_j^W))||_2^2 + \frac{P_j w_j}{N}\right)} +$$

$$\boxed{\sum_{j=1}^{M^G}\sum_{i=1}^{N}\frac{1}{2(\sigma_j^G)^2}\left(||S_{ij}^G - \phi_i^{All}(T_j^G)||_2^2 + \frac{P_j w_j}{N}\right)} + \qquad (5)$$

$$\boxed{\frac{1}{2}\sum_{i=1}^{N}(\boldsymbol{\alpha}_i^W)^T(\Gamma_\alpha^W)^{-1}\boldsymbol{\alpha}_i^W} + \boxed{\frac{1}{2}\sum_{i=1}^{N}(\boldsymbol{\alpha}_i^{All})^T(\Gamma_\alpha^{All})^{-1}\boldsymbol{\alpha}_i^{All}} +$$

$$\frac{w_\alpha^W}{2}tr((\Gamma_\alpha^W)^{-1}P_\alpha^W) + \frac{w_\alpha^{All}}{2}tr((\Gamma_\alpha^{All})^{-1}P_\alpha^{All}) +$$

$$\frac{(w_\alpha^W + N)}{2}\log(|\Gamma_\alpha^W|) + \frac{(w_\alpha^{All} + N)}{2}\log(|\Gamma_\alpha^{All}|) +$$

$$\sum_{j=1}^{M^W}\frac{1}{2}(w_j + \Lambda_j N)\log((\sigma_j^W)^2) + \sum_{j=1}^{M^G}\frac{1}{2}(w_j + \Lambda_j N)\log((\sigma_j^G)^2)$$

Eq. 5 represents the cost function of our algorithm. The framed terms refer to the data-terms and to the regularity terms of both diffeomorphisms respectively. The other terms are due to the use of inverse Wishart prior distributions. The proposed statistical framework is generic since it can be employed with any shape model, provided it is possible to define probability density functions, and with any parametric deformation model.

C. Similarity metrics for shape complexes

When landmarks are not available, we propose to use two correspondence-free shape metrics: varifolds [32], for grey matter surfaces, and weighted currents [30], for white matter streamlines. In both frameworks, meshes are embedded into a infinite-dimensional Hilbert space where the union of surfaces or streamlines is equal to a sum of varifolds (V) or weighted currents (C) respectively. We define Gaussian variables in these two spaces similarly to [40] for the framework of currents. Since they do not have probability density functions (pdf) in infinite dimension, we project both templates T and

shapes S, modelled as varifolds or weighted currents, to finite dimensional spaces. For each structure j, we define a regular grid composed of Λ_j points which covers the ambient space and where pdf can be computed. The reader is referred to [41, Chapter 4.2.3] for more details.

Varifolds - In the space of varifolds W^*, the inner-product between two surfaces X and Y is: $\langle V_X, V_Y \rangle_{W^*} = \sum_{l=1}^{L}\sum_{h=1}^{H}\exp\left(\frac{-||p_l-q_h||_2^2}{\lambda_W^2}\right)\left(\frac{n_l^T u_h}{|n_l|_2|u_h|_2}\right)^2 |n_l|_2|u_h|_2$ where n_l (resp. u_h) is the normal of X (resp. Y) at point p_l (resp. q_h). The only user-defined parameter is the kernel bandwidth λ_W. The distance between V_X and V_Y is: $||V_X - V_Y||_{W^*}^2 = \langle V_X - V_Y, V_X - V_Y\rangle_{W^*}$. Two important characteristics of this metric are: the absence of correspondences and the invariance to a change of orientation of some normals of the surfaces. For more information, the user is referred to [32].

Weighted currents - The inner product between two oriented 3D polygonal curves, A and B, modelled as weighted currents and composed of G and F segments respectively, is: $\langle C_A, C_B\rangle_{Q^*} = \sum_{g=1}^{G}\sum_{f=1}^{F}\exp\left(\frac{-||x_g-y_f||_2^2}{\lambda_g^2}\right)\alpha_g^T\beta_f$ $\exp\left(\frac{-||f^a-t^a||_2^2}{\lambda_a^2}\right)\exp\left(\frac{-||f^b-t^b||_2^2}{\lambda_b^2}\right)$ where Q^* indicates the space of weighted currents, x_g and α_g (resp. y_f and β_f) are the centre and tangent vector of segment g (resp. f). The two 3D vectors f^a and f^b (resp. t^a and t^b) are the coordinates of the end-points of the curve A (resp. B). Two curves are considered similar if their pathways are alike, as in usual currents [39], but also if their endpoints are close to each other. The inner product is parametrised by three user-defined bandwidths: λ_g, λ_a, λ_b. The distance between C_A and C_B is defined as: $||C_A - C_B||_{Q^*}^2 = \langle C_A - C_B, C_A - C_B\rangle_{Q^*}$. As usual currents, curves need to have a consistent orientation. This can be achieved by tracing all streamlines of a bundle from one ROI (Region Of Interest) to another one, as it is the case in this paper. For more details, please see [30].

Weighted prototypes - We approximate white matter fiber bundles with a parsimonious representation of weighted streamlines prototypes. Prototypes are chosen among the streamlines by minimizing an approximation error based on the metric of weighted currents. Every prototype represents an ensemble of streamlines which share similar endpoints and pathway. The weight of the prototype is related to the number of streamlines approximated. An outlier detection and removal step is also performed during the algorithm. This approximation preserves the global shape of the bundle and its structural connectivity, which is fundamental for the scope of this paper. For more information the user is referred to [30].

D. Diffeomorphic deformations

We define here how to compute the diffeomorphic deformations of the template complex. Our approach relies on the Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework based on the control point formulation presented in [24]. For every subject i, both ϕ_i^W and ϕ_i^{All} are defined as the last deformations of two flows of diffeomorphisms $\{\phi_{it}^W\}_{t\in[0,1]}$ and $\{\phi_{it}^{All}\}_{t\in[0,1]}$. Calling $\phi_i(\boldsymbol{x},t) = \phi_{it}(\boldsymbol{x}) = \boldsymbol{x}_i(t)$ the position of a point at time t which was located in $\boldsymbol{x}$ at time $t = 0$, each flow is built by integrating:

$\frac{\partial \phi_i(\boldsymbol{x},t)}{\partial t} = v_i(\phi_i(\boldsymbol{x},t),t) = v_i(\boldsymbol{x}_i(t),t)$ over $t \in [0,1]$ where $v_i(\boldsymbol{x}_i(t),t)$ is a time-varying vector field representing the instantaneous velocity of a point located in $\boldsymbol{x}_i(t)$ at time t. Both vector fields v_i^{All} and v_i^W belong to the same RKHS D with Gaussian kernel $\boldsymbol{K}_D$. They are defined by two different sets of 3D control points, $\boldsymbol{c}^{All}$ and $\boldsymbol{c}^W$, shared among all subjects, and by two distinct sets of 3D vectors, called momenta, $\boldsymbol{\alpha}_i^{All}$ and $\boldsymbol{\alpha}_i^W$ linked to the control points and specific to each subject i: $v_i^{All}(\boldsymbol{x}_i(t),t) = \boldsymbol{K}_D(\boldsymbol{x}_i(t),\boldsymbol{c}^{All}(t))\boldsymbol{\alpha}_i^{All}(t)$ and $v_{it}^W(\boldsymbol{x}_i(t)) = \boldsymbol{K}_D(\boldsymbol{x}_i(t),\boldsymbol{c}^W(t))\boldsymbol{\alpha}_i^W(t)$, where $\boldsymbol{x}_i(0) = \boldsymbol{x}$ and $\boldsymbol{K}_D(\boldsymbol{x}_i(t),\boldsymbol{c}(t))$ represents a block matrix of Gaussian kernels with an equal fixed bandwidth for both v_i^{All} and v_i^W. The deformation of every point $\boldsymbol{x}$ in the ambient space depends on its initial position at $t = 0$ and on the evolution of the system $\boldsymbol{L}_i^{All}(t) = \{\boldsymbol{c}^{All}(t),\boldsymbol{\alpha}_i^{All}(t)\}$ if the point belongs to the grey matter, and on both systems $\boldsymbol{L}_i^W(t) = \{\boldsymbol{c}^W(t),\boldsymbol{\alpha}_i^W(t)\}$ and $\boldsymbol{L}_i^{All}(t)$ if the point belongs to the white matter. At $t = 0$ the deformations ϕ_{i0}^W and ϕ_{i0}^{All} are equal to the identity transformations. For both systems, the path from ϕ_{i0}^{All} (resp. ϕ_{i0}^W) to ϕ_{i1}^{All} (resp. ϕ_{i1}^W), the latter being the deformation of interest, is chosen as the geodesic one, which means the one that minimizes the total kinetic energy along the path: $\int_0^1 ||v_{it}^{All}||_D^2$ (resp. $\int_0^1 ||v_{it}^W||_D^2$). It has been shown in [20] that the extremal paths are such that both systems $\boldsymbol{L}_i^W(t)$ and $\boldsymbol{L}_i^{All}(t)$ satisfy:

$$\dot{\boldsymbol{c}}_i(t) = \boldsymbol{K}_D(\boldsymbol{c}_i(t),\boldsymbol{c}_i(t))\boldsymbol{\alpha}_i(t) = F^c(\boldsymbol{c}_i(t),\boldsymbol{\alpha}_i(t))$$
$$\dot{\boldsymbol{\alpha}}_i(t) = -\boldsymbol{\alpha}_i(t)^T\boldsymbol{\alpha}_i(t)\nabla_1\boldsymbol{K}_D(\boldsymbol{c}_i(t),\boldsymbol{c}_i(t)) = F^\alpha(\boldsymbol{c}_i(t),\boldsymbol{\alpha}_i(t))$$
$$\text{s.t. } \boldsymbol{c}_i(0) = \boldsymbol{c}(0) = \boldsymbol{c}_0 , \ \boldsymbol{\alpha}_i(0) = \boldsymbol{\alpha}_{i0} \qquad (6)$$

which can be summarized as $\dot{\boldsymbol{L}}_i^{All}(t) = F(\boldsymbol{L}_i^{All}(t))$ (resp. $\dot{\boldsymbol{L}}_i^W(t) = F(\boldsymbol{L}_i^W(t))$). The last diffeomorphisms ϕ_{i1}^{All} and ϕ_{i1}^W are completely parametrized by the initial conditions of the systems: $\boldsymbol{L}_i^{All}(0) = \boldsymbol{L}_{i0}^{All} = \{\boldsymbol{c}_0^{All},\boldsymbol{\alpha}_{i0}^{All}\}$ (resp. $\boldsymbol{L}_{i0}^W = \{\boldsymbol{c}_0^W,\boldsymbol{\alpha}_{i0}^W\}$). Thus, in order to put into correspondence the template $\boldsymbol{T}$ with the subject complex $\boldsymbol{S}_i$, we deform only the white matter of the template $\boldsymbol{T}^W$ integrating forward in time first $\dot{\boldsymbol{L}}_i^W(t)$ and then also:

$$\dot{\boldsymbol{T}}_i^W(t) = \boldsymbol{K}_D(\boldsymbol{T}_i^W(t),\boldsymbol{c}_i^W(t))\boldsymbol{\alpha}_i^W(t) = Z[\boldsymbol{T}_i^W(t),\boldsymbol{L}_i^W(t)]$$
$$\text{s.t. } \boldsymbol{T}_i^W(0) = \boldsymbol{T}_{i0}^W = \boldsymbol{T}^W \qquad (7)$$

The deformed white matter of the template $\phi_{i1}^W(\boldsymbol{T}^W) = \boldsymbol{T}_{i1}^W$, together with the un-deformed grey matter of the template $\boldsymbol{T}^G = \boldsymbol{T}_{i0}^G$, constitute $\boldsymbol{T}_{i0}^{All} = \boldsymbol{T}_{i1}^W \cup \boldsymbol{T}_{i0}^G$. They are deformed by the second diffeomorphism ϕ_{i1}^{All} computed integrating forward in time first $\dot{\boldsymbol{L}}_i^{All}(t)$ and then:

$$\dot{\boldsymbol{T}}_i^{All}(t) = \boldsymbol{K}_D(\boldsymbol{T}_i^{All}(t),\boldsymbol{c}_i^{All}(t))\boldsymbol{\alpha}_i^{All}(t) = Z[\boldsymbol{T}_i^{All}(t),\boldsymbol{L}_i^{All}(t)]$$
$$\text{s.t. } \boldsymbol{T}_i^{All}(0) = \boldsymbol{T}_{i0}^{All} = \boldsymbol{T}_{i1}^W \cup \boldsymbol{T}_{i0}^G \qquad (8)$$

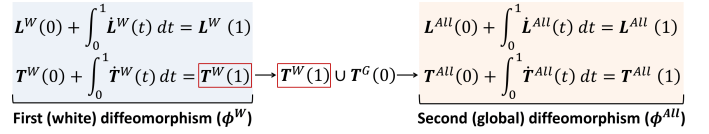A diagram of the double diffeomorphism is shown in Fig.2.

Fig. 2. Diagram of the template deformation based on the proposed double diffeomorphism. We omit the subject index i for clarity purpose.

E. Optimization procedure

The double diffeomorphism can be integrated in the previously presented Bayesian setting for atlas construction. The two sets of initial control points and momenta, $\{\boldsymbol{c}_0^{All},\{\boldsymbol{\alpha}_{i0}^{All}\}\}$ and $\{\boldsymbol{c}_0^W,\{\boldsymbol{\alpha}_{i0}^W\}\}$, represent the deformation parameters which warp the template $\boldsymbol{T}$ towards the subject complex $\boldsymbol{S}_i$. Initial control points, $\boldsymbol{c}_0^{All}$ and $\boldsymbol{c}_0^W$, are considered as parameters of the model since they are fixed effect common to the entire population. Initial momenta, $\boldsymbol{\alpha}_{i0}^{All}$ and $\boldsymbol{\alpha}_{i0}^W$, are instead the subject-specific deformation parameters and, as previously $\boldsymbol{\alpha}_i^{All}$ and $\boldsymbol{\alpha}_i^W$, they follow a Gaussian distribution. Assuming that all random variables are independent, the cost function in Eq.5 does not change except for the framed terms where we exchange the L2-norm with the one of varifolds and weighted currents, for grey and white matter structures respectively, and where we substitute $\boldsymbol{\alpha}_i$ with $\boldsymbol{\alpha}_{i0}$.

The variables $\boldsymbol{T}$, $\{\boldsymbol{\alpha}_{i0}^{All}\}$, $\{\boldsymbol{\alpha}_{i0}^W\}$, $\boldsymbol{c}_0^{All}$, $\boldsymbol{c}_0^W$ are minimised using a gradient descent scheme. Instead, Γ_α^{All}, Γ_α^W and $\{\sigma_j^2\}$ have closed form solutions due to the use of conjugate priors:

$$\hat{\Gamma}_\alpha^W = \frac{\sum_{i=1}^N \left[(\boldsymbol{\alpha}_{i0}^W)(\boldsymbol{\alpha}_{i0}^W)^T\right] + w_\alpha^W(P_\alpha^W)^T}{(w_\alpha^W + N)} \qquad (9)$$

$$\hat{\Gamma}_\alpha^{All} = \frac{\sum_{i=1}^N \left[(\boldsymbol{\alpha}_{i0}^{All})(\boldsymbol{\alpha}_{i0}^{All})^T\right] + w_\alpha^{All}(P_\alpha^{All})^T}{(w_\alpha^{All} + N)}$$

$$(\hat{\sigma}_j^G)^2 = \frac{\sum_{i=1}^N ||\Pi(S_{ij} - \phi_{i1}^{All}(T_j^G))||_{W_{\Lambda_j}^*}^2 + w_j P_j}{(w_j + N\Lambda_j)}$$

$$(\hat{\sigma}_j^W)^2 = \frac{\sum_{i=1}^N ||\Pi(S_{ij} - \phi_{i1}^{All}(\phi_{i1}^W(T_j^W)))||_{Q_{\Lambda_j}^*}^2 + w_j P_j}{(w_j + N\Lambda_j)}$$

where $W_{\Lambda_j}^*$ and $Q_{\Lambda_j}^*$ are the finite dimensional spaces where varifolds and weighted currents are projected to. For more details about the projection Π, the user is referred to [41, Chapter 4.2.3.3].

The two covariance matrices $\hat{\Gamma}_\alpha^W$ and $\hat{\Gamma}_\alpha^{All}$ are equal to a weighted sum between the sample covariance matrix of the initial momenta and a prior. We choose $P_\alpha^W = \boldsymbol{K}_D^{-1}(\boldsymbol{c}_0^W,\boldsymbol{c}_0^W)$ and $P_\alpha^{All} = \boldsymbol{K}_D^{-1}(\boldsymbol{c}_0^{All},\boldsymbol{c}_0^{All})$ which are block matrices of Gaussian kernels between the initial control points. Note that $\boldsymbol{K}_D$ is the kernel of the RKHS to which belong both vector fields v_i^{All} and v_i^W. This choice is motivated by the fact that, when $N << w_\alpha^W$, it results $\hat{\Gamma}_\alpha^W \propto \boldsymbol{K}_D^{-1}(\boldsymbol{c}_0^W,\boldsymbol{c}_0^W)$ and, consequently, the regularity term in Eq.5 becomes $\sum_{i=1}^N (\boldsymbol{\alpha}_{i0}^W)^T \boldsymbol{K}_D(\boldsymbol{c}_0^W,\boldsymbol{c}_0^W)\boldsymbol{\alpha}_{i0}^W = \sum_{i=1}^N ||v_{i0}^W||_D^2$ which is the sum of the lengths of the geodesic paths over all subjects. This kind of regularity term has been often employed in previous

atlas construction methods not based on a statistical setting [18]. The same reasoning is also valid for $\hat{\Gamma}_\alpha^{All}$.

The two other parameters, $\hat{\sigma}_j^G$ and $\hat{\sigma}_j^W$, are equal to a weighted sum between the data-term of the j-th structure and the prior P_j. Each parameter balances the importance of structure j with respect to the other structures and with respect to the regularity terms of both diffeomorphisms. The prior P_j imposes a minimum value to $\hat{\sigma}_j$ which is useful to avoid overfitting. In fact, without a prior, the minimisation process might focus only on a structure k, reducing its residuals almost to zero and ignoring the other structures. This would result in $\hat{\sigma}_j^2 \rightarrow 0$ and therefore also to $\log(\hat{\sigma}_j^2) \rightarrow -\infty$.

The gradients of the cost function E in Eq.5 with respect to $T_k^G, T_k^W, \boldsymbol{\alpha}_{s0}^{All}, \boldsymbol{\alpha}_{s0}^W, \boldsymbol{c}_0^{All}, \boldsymbol{c}_0^W$, where k and s are the indexes for the structure and subject respectively, are equal to:

$$\nabla_{T_k^G} E = \sum_{i=1}^N \frac{1}{2\sigma_k^2} \nabla_{T_k^G} D_{ik}^G \quad \nabla_{\boldsymbol{c}_0^{All}} E = \sum_{i=1}^N \sum_{j=1}^M \frac{1}{2\sigma_j^2} \nabla_{\boldsymbol{c}_0^{All}} D_{ij}$$

$$\nabla_{T_k^W} E = \sum_{i=1}^N \frac{1}{2\sigma_k^2} \nabla_{T_k^W} D_{ik}^W \quad \nabla_{\boldsymbol{c}_0^W} E = \sum_{i=1}^N \sum_{j=1}^M \frac{1}{2\sigma_j^2} \nabla_{\boldsymbol{c}_0^W} D_{ij}$$

$$\nabla_{\boldsymbol{\alpha}_{s0}^{All}} E = \sum_{j=1}^M \frac{1}{2\sigma_j^2} \nabla_{\boldsymbol{\alpha}_{s0}^{All}} D_{sj} + (\Gamma_\alpha^{All})^{-1} \boldsymbol{\alpha}_{s0}^{All}$$

$$\nabla_{\boldsymbol{\alpha}_{s0}^W} E = \sum_{j=1}^M \frac{1}{2\sigma_j^2} \nabla_{\boldsymbol{\alpha}_{s0}^W} D_{sj} + (\Gamma_\alpha^W)^{-1} \boldsymbol{\alpha}_{s0}^W \quad (10)$$

where $D_{ij}^G = ||\Pi(S_{ij} - \phi_{i1}^{All}(T_j^G))||_{W_{\Lambda j}^*}^2$ and $D_{ij}^W = ||\Pi(S_{ij} - \phi_{i1}^{All}(\phi_{i1}^W(T_j^W)))||_{Q_{\Lambda j}^*}^2$ refer to the data terms of grey and white matter structures respectively whereas D_{ij} refers to the data-term of any structure.

To calculate the gradients of the data terms $\{D_{ij}\}$, we need the deformed template complex at the end of the second diffeomorphism for every subject i and for all structures, namely $\phi_{i1}^{All}(\boldsymbol{T}^G)$ and $\phi_{i1}^{All}(\phi_{i1}^W(\boldsymbol{T}^W))$. First, we integrate forward in time $\dot{\boldsymbol{L}}_i^W(t) = F(\boldsymbol{L}_i^W(t))$ (Eq.6) and $\dot{\boldsymbol{T}}_i^W(t) = Z[\boldsymbol{T}_i^W(t), \boldsymbol{L}_i^W(t)]$ (Eq.7). Then, we integrate $\dot{\boldsymbol{L}}_i^{All}(t) = F(\boldsymbol{L}_i^{All}(t))$ (Eq.6) and, using as initial value $\boldsymbol{T}_i^{All}(0) = \boldsymbol{T}_{i1}^W \cup \boldsymbol{T}_{i0}^G$, we integrate also $\dot{\boldsymbol{T}}_i^{All}(t) = Z[\boldsymbol{T}_i^{All}(t), \boldsymbol{L}_i^{All}(t)]$ (Eq.8). After that, we can compute the data term D_{ij} and its gradient with respect to the vertices of $\boldsymbol{T}_i^{All}(1) = \boldsymbol{T}_{i1}^{All}$. Using the calculus of variations, this information is brought back from $t = 1$ to $t = 0$ to update first $\boldsymbol{L}_i^{All}(0) = \{\boldsymbol{c}_0^{All}, \boldsymbol{\alpha}_{i0}^{All}\}$ and $\boldsymbol{T}^G$ and then $\boldsymbol{L}_i^W(0) = \{\boldsymbol{c}_0^W, \boldsymbol{\alpha}_{i0}^W\}$ and $\boldsymbol{T}^W$. The optimisation is based on a set of adjoint equations describing the evolution of four auxiliary variables $\theta_i^{All}, \xi_i^{All} = \{\xi_{\alpha i}^{All}, \xi_{ci}^{All}\}, \theta_i^W, \xi_i^W = \{\xi_{\alpha i}^W, \xi_{ci}^W\}$:

$$\dot{\theta}_i^{All}(t) = -(\partial_{\boldsymbol{T}_i^{All}} Z_i^{All}(t))^T \theta_i^{All}(t) \quad (11)$$

$$\dot{\xi}_i^{All}(t) = -(\partial_{\boldsymbol{L}_i^{All}} Z_i^{All}(t))^T \theta_i^{All}(t) + \quad (12)$$
$$(d_{\boldsymbol{L}_i^{All}} F_i^{All}(t))^T \xi_i^{All}(t)$$

$$\dot{\theta}_i^W(t) = -(\partial_{\boldsymbol{T}_i^W} Z_i^W(t))^T \theta_i^W(t) \quad (13)$$

$$\dot{\xi}_i^W(t) = -(\partial_{\boldsymbol{L}_i^W} Z_i^W(t))^T \theta_i^W(t) + \quad (14)$$

$$(d_{\boldsymbol{L}_i^W} F_i^W(t))^T \xi_i^W(t)$$

$$\text{s.t. } \theta_i^{All}(1) = \nabla_{\boldsymbol{T}_i^{All}(1)} \boldsymbol{D}_i \; , \; \boxed{\theta_i^W(1) = \theta_i^{All,W}(0)}$$

$$\xi_i^{All}(1) = 0 \; , \; \xi_i^W(1) = 0$$

where $Z_i^{All}(t) = Z[\boldsymbol{T}_i^{All}(t), \boldsymbol{L}_i^{All}(t)]$ and $Z_i^W(t) = Z[\boldsymbol{T}_i^W(t), \boldsymbol{L}_i^W(t)]$. The size of θ_i^{All} (resp. θ_i^W) and ξ_i^{All} (resp. ξ_i^W) are the same as the ones of $\boldsymbol{T}$ (resp. $\boldsymbol{T}^W$) and $\boldsymbol{L}_i^{All}$ (resp. $\boldsymbol{L}_i^W$) respectively. We first integrate backward in time Eq.11 and Eq.12 obtaining $\theta_i^{All}(0)$ and $\xi_i^{All}(0) = \{\xi_{\alpha i}^{All}(0), \xi_{ci}^{All}(0)\}$. Then, we use $\theta_i^{All,W}(0)$, which are the initial values of θ_i^{All} relative to the white matter, as final values for θ_i^W and we integrate backward in time Eq.13 and Eq.14 obtaining $\theta_i^W(0)$ and $\xi_i^W(0) = \{\xi_{\alpha i}^W(0), \xi_{ci}^W(0)\}$. From this set of equations, we can notice that the optimisation of the two diffeomorphisms is linked by the constraint $\theta_i^W(1) = \theta_i^{All,W}(0)$. The information given by $\nabla_{\boldsymbol{T}_i^{All}(1)} \boldsymbol{D}_i$, where $\boldsymbol{D}_i = \{D_{ij}\}_{j=1,...,M}$, flows from the second diffeomorphism (All) to the first one (W) and eventually it is used to update all parameters:

$$\nabla_{\boldsymbol{T}^G} E = \sum_{i=1}^N \theta_i^{All,G}(0) \quad \nabla_{\boldsymbol{T}^W} E = \sum_{i=1}^N \theta_i^W(0)$$

$$\nabla_{\boldsymbol{c}_0^{All}} E = \sum_{i=1}^N \xi_{ci}^{All}(0) \quad \nabla_{\boldsymbol{c}_0^W} E = \sum_{i=1}^N \xi_{ci}^W(0)$$

$$\nabla_{\boldsymbol{\alpha}_{s0}^{All}} E = \xi_{\alpha s}^{All}(0) + (\Gamma_\alpha^{All})^{-1} \boldsymbol{\alpha}_{s0}^{All}$$

$$\nabla_{\boldsymbol{\alpha}_{s0}^W} E = \xi_{\alpha s}^W(0) + (\Gamma_\alpha^W)^{-1} \boldsymbol{\alpha}_{s0}^W \quad (15)$$

where $\theta_i^{All,G}$ refers to the values of θ_i^{All} relative to the grey matter structures. A diagram of the optimisation procedure is shown in Fig.3. More details about the computations can be found in the Appendix of [23].
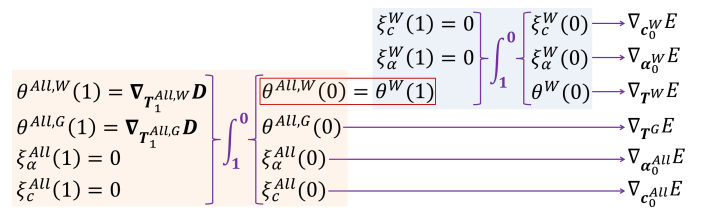

Fig. 3. Diagram of the optimisation procedure. We omit the subject index i for clarity purpose.

F. Atlas Parameters Initialisation

Since we use a gradient descent scheme, we need to initialise the atlas parameters. Control points of both diffeomorphisms are initialised as a regular lattice covering the entire ambient space with an inter-points distance equal to the bandwidth of the diffeomorphic kernel $\boldsymbol{K}_D$. Momenta are initialised to zero. The template of surface meshes is initialised as the average of the population when a vertex-correspondence is available. Otherwise, we use a centred and scaled ellipsoid as in [20]. For the template of fiber bundles, all subjects'

bundles are first gathered together into a single bundle which is then approximated as a set of weighted prototypes. The weights of the prototypes are scaled so that the norm of the template is equal to the average norm of the population.

III. EXPERIMENTS AND RESULTS

In this section, we first describe the dataset used in the following experiments and some numerical aspects of the proposed algorithm. Then, we use both a toy example and real data to compare the registrations based on a single and double diffeomorphism. After that, we present an explanatory toy example where we show how one could use the proposed double diffeomorphism to compare two groups of subjects. Eventually, we assess the effectiveness of our algorithm by showing that it better discriminates between controls and GTS patients than a single diffeomorphism.

A. Materials

The dataset used in this paper contains 76 subjects: 27 controls and 49 GTS patients divided in three sub-groups based on their symptoms: ST=simple-tics (17 patients), CT=complex-tics (15), OCD=complex-tics with Obsessive Compulsive Disorders (12). Anatomical images are acquired using 3D MR T1-w sequences with a voxel size of 1x1x1 mm^3. MR diffusion weighted images are obtained with 50 gradient-directions, a B-factor of 1000 and a voxel size of 2x2x2 mm^3. Artefacts related to spike, motion, susceptibility and eddy currents are corrected using Connectomist-2.0. Diffusion and T1-w data are matched using a mutual information based registration technique. In the experiments, we use the left hemisphere of the cortical surface, left putamen and the fiber bundles connecting them. For the selection of the tracts, we use a specific technique conceived for the cortico-striatal circuits explained in [2]. Cortical surfaces are segmented using FreeSurfer v5.3 followed by a pipeline of BrainVisa v4.3.0 which produces a vertex-correspondence between subjects. Putamens are segmented with FSL. Fiber bundles result from a deterministic tractography algorithm (1 seed per voxel, SDT model) available in Connectomist-2.0. See [2] for more details about the data-set, pre-processing and tractography.

B. Numerical aspects

In the following experiments, cortical surfaces are modelled with landmarks, putamens as varifolds with λ_W=3mm and fiber bundles, approximated with weighted prototypes, as weighted currents with λ_g=7mm, λ_a=10mm (cortex) and λ_b=5mm (putamen). The bandwidths of both diffeomorphic kernels are equal to 11mm, which produce 804 control points. The maximum number of iterations fo the atlas construction is 120 and the computations are performed on a Intel Xeon, 32 cores, CPU E5-2650, 2.60GHz with a graphic card NVIDIA Quadro 5000. The code is written in C++ and CUDA and it is an extension of the freely available software suite *deformetrica* (www.deformetrica.org). The computational time for an atlas with 10 subjects is about 37 hours. All shape complexes are previously rigidly registered to a reference shape complex.

C. Toy example - registration

In Fig.4, we compare the registrations of a toy template complex (blue) towards a toy subject complex (red) based on a single (first row) and double diffeomorphism (second row). Both complexes are composed of a *pseudo* cortical surface, sub-cortical structure and fiber bundle linking them, all modelled as varifolds. We use the same parameters for the two deformation schemes. Grey matter components have a similar shape but they do not share the same structural connectivity. As it is possible to notice, a single diffeomorphism can not correctly put into correspondence all structures. On the contrary, a double diffeomorphism makes first the fiber bundle move, keeping fixed the grey matter structures, and then it accurately registers all structures with the second diffeomorphism. In this way, it is possible to disentangle the differences in structural connectivity, captured by the first diffeomorphism, from the global morphological changes, captured by the second diffeomorphism.
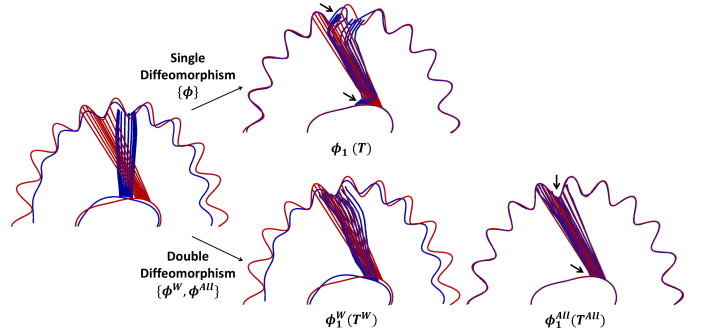

Fig. 4. Registration between a toy-template complex (blue) and a toy-subject complex (red) using either a single or a double diffeomorphism. Black arrows indicate the areas where only the double diffeomorphism can correctly put into correspondence all structures.

D. Real data - registration

In Fig.5, we compare the results of a single and double diffeomorphism using real data. We match a control subject to a Gilles de la Tourette patient. A double diffeomorphism better aligns both white and grey matter (see Fig. 4 in the Appendix) than a single diffeomorphism, capturing at the same time the variations in structural connectivity (i.e. ϕ^W).

E. Toy Example - group differences

We present here an explanatory toy example of the proposed atlas construction procedure based on a toy data-set constituted of 6 pseudo shape complexes representing two different populations (3 controls and 3 GTS patients). They are shown in Fig.6 where it is possible to notice that the complexes of population A have a different organization and shape with respect to the ones of population B. The Bayesian atlas construction results in a final template complex and in the covariance matrices of the momenta of both diffeomorphisms. The template shows the characteristics common to both populations. The two covariance matrices describe the
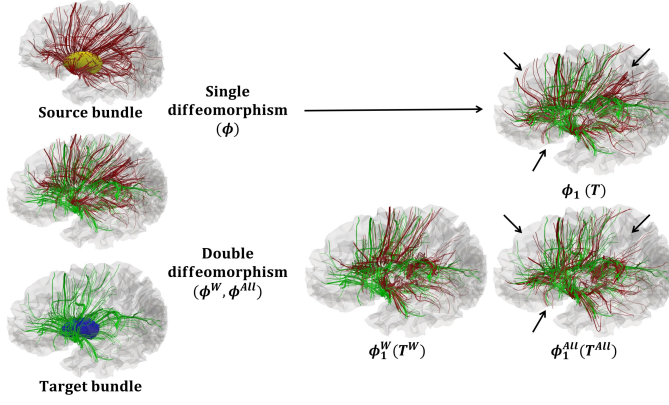
Fig. 5. Comparison between a single and double diffeomorphic registration using real data. Source and target bundles belong to a control and a GTS patient respectively. Black arrows indicate the areas where a single diffeomorphism can not correctly match the fiber bundles.
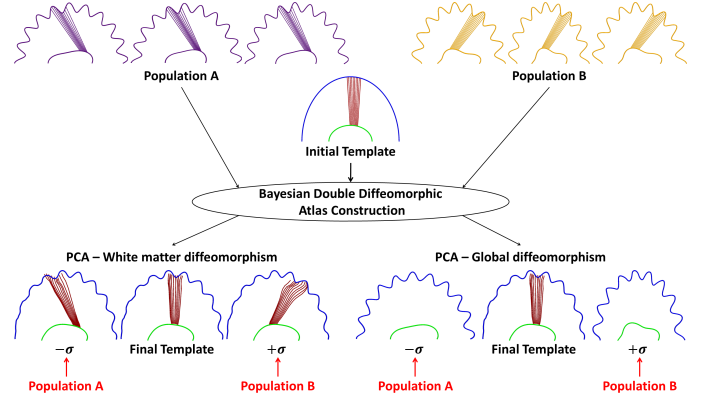

Fig. 6. At the top, we present two toy populations characterised by a different cortex, sub-cortical nucleus and structural connectivity. In the middle, we show the initial template. The final estimated template is presented at the bottom. It is deformed at $\pm\sigma$ along the first modes of two PCA computed with Γ_α^W and Γ_α^{All}. The endpoints of the two modes, at $-\sigma$ and $+\sigma$, reproduce the structural connectivity and the morphological characteristics of the two groups respectively.

organisational and global *morphological* variability within the 6 subjects respectively. We compute a Principal Component Analysis (PCA) for each covariance matrix and we deform the final template complex at $\pm\sigma$ (standard deviation) along the first modes of both PCAs. The main variations captured by the first diffeomorphisms $\{\phi_i^W\}$, which affect only the fiber bundles, explain the principal differences in structural connectivity between the two populations. The positions of the fiber bundle at $-\sigma$ and $+\sigma$ are the ones of population A and B respectively. The first mode of the second PCA describes instead the main global morphological variations. We can notice that the grey matter structures at $-\sigma$ and $+\sigma$ reproduce the morphological characteristics of population A and B respectively. This example shows the exploratory potential of the proposed method and it is based on a simple toy data-set where the intra-group variations are definitely smaller than the inter-group ones. This is probably exaggerated compared to a real-data example. Nevertheless, given the important structural changes that are likely to occur in syndromes such as GTS, we may assume that controls and patients create distinct clusters. In the next section, we will exploit this hypothesis by looking for the discriminant hyperplane that separates the two groups.

F. Real data - Classification

Here, we use the estimated initial momenta of the two diffeomorphisms, $\boldsymbol{\alpha}_{i0}^W$ and $\boldsymbol{\alpha}_{i0}^{All}$, as features to discriminate between controls and patients. Then, we compare the resulting classification scores with the ones obtained using the initial momenta of a single diffeomorphism.

First of all, we build an atlas with 10 subjects (5 controls and 5 patients). Since we use subjects from both groups, the final template should be positioned in between them in the shape space. The estimated template is successively warped to all the *remaining* $J = 66$ subjects by minimizing a cost function similar to Eq.5 where we do not sum over all subjects i and where we fix the controls points, Γ_α^W, Γ_α^{All} and σ_j^2 to the values estimated during the initial atlas. The resulting initial momenta, $\boldsymbol{\alpha}_{i0}^{All}$ and $\boldsymbol{\alpha}_{i0}^W$, represent the input features of the classifier. We employ a Linear Discriminant Analysis (LDA) with a leave-one-out cross validation strategy. We assume that the class-conditional densities of the initial momenta are Gaussian with a covariance matrix equal to the one estimated during the initial atlas. This can be seen as a regularised LDA since the covariance matrix is estimated as in Eq.9. We separately test the discriminative power of the two diffeomorphisms by using either only $\boldsymbol{\alpha}_{i0}^{All}$ or $\boldsymbol{\alpha}_{i0}^W$. Moreover, we compare these results with the ones obtained using the initial momenta of a single diffeomorphism where we employ either only the fiber bundles or all structures from both grey and white matter. Resulting sensibility, sensitivity and balanced accuracy are shown in Table I where we separately use either all patients or each sub-group alone. We assess the statistical significance of the classification scores with a randomization test (1000 permutations). It is possible to notice that the classification scores based on the first (white) diffeomorphism, especially for the most severe patients (CT and OCD), are definitely better than using a single diffeomorphism.

Due to the variability of the results, we also investigate the sampling distributions of sensitivity, sensibility and balanced accuracy within the group of patients with a bootstrap analysis. More precisely, we perform it on the top of the previous leave-one-out cross validation classification. At each iteration, we pick a random sample (with replacement) of the 44 patients which is classified, together with the 22 controls, using LDA. We repeat this process 1000 times. The histograms of balanced accuracy for the double and single diffeomorphic approach are shown in Fig.7. The average sensitivity and specificity is respectively: 74% and 51% for the global diffeomorphism, 73% and 64% for the white diffeomorphism, 64% and 48% for the single diffeomorphism, considering both white and grey matter, and 70% and 52% for the single diffeomorphism, using only the fiber bundles.

TABLE I
CLASSIFICATIONS SCORES

| Single Diffeomorphism - White and Grey Matter			
	Sensitivity %	Specificity %	Balanced Accuracy %
ST	12	36	24
CT	33	64	48
OCD	58	59	59
CT+OCD	52	64	58
ST+CT+OCD	54	41	48
Single Diffeomorphism - Only White Matter			
ST	53	54	54
CT	33	45	39
OCD	50	54	52
CT+OCD	59	59	59
ST+CT+OCD	66	45	56
Double Diffeomorphism - First (white) diffeomorphism			
ST	47	59	53
CT	67	77	72*
OCD	50	82	66*
CT+OCD	74	64	69*
ST+CT+OCD	73	41	57
Double Diffeomorphism - Second (global) diffeomorphism			
ST	29	50	40
CT	40	45	43
OCD	50	68	59
CT+OCD	52	68	60
ST+CT+OCD	70	50	60

* : p-value < 0.05
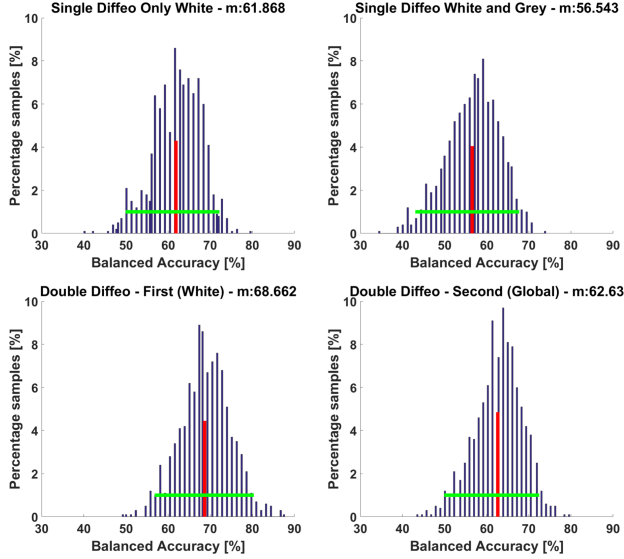
Fig. 7. Bootstrap analysis of 1000 iterations performed on the top of a LDA with a leave-one-out cross validation. Each sample of the histogram represents the classification score obtained using 44 patients chosen randomly (with replacement) among all sub-groups and 22 fixed controls. Red and green lines show the average and the 95% confidence interval respectively.

G. Most discriminative deformation axis

Eventually, we also compute the organizational and morphological characteristics proper to each group by deforming the template complex along the most discriminative deformation axis. We estimate the best linear decision boundary (i.e. $\boldsymbol{\alpha}^T \boldsymbol{w}^* - b^*$) with all the J test subjects (22 controls and

44 patients) using either $\boldsymbol{\alpha}_{i0}^{All}$ or $\boldsymbol{\alpha}_{i0}^{W}$. The typical configurations of patients and controls are found by deforming the template complex at $\mu - w^*$ and $\mu + w^*$ respectively, where $\mu = \frac{1}{2}(\mu_c + \mu_p)$ and $||w^*|| = ||\mu_c - \mu_p||$ with μ_c and μ_p equal to the averages of initial momenta of controls and patients respectively. In Fig. 8, we compare the typical structural connectivity of the two groups. The main differences are in the supplementary motor, premotor, superior frontal areas, insula and in the dorsal and ventro-lateral part of the putamen. These results are in line with those reported in the literature [2]. In Fig. 9, we compare the typical grey matter configurations of controls and patients. In this case, there is mainly a compression in the premotor and frontal area of the cortex, insula and occipital lobe. About the putamen, the main variations are in the fronto-dorsal and posterio-ventral areas. In Fig. 5 of the Appendix, we show for comparison the main variations in structural connectivity and morphology only within the population of controls.
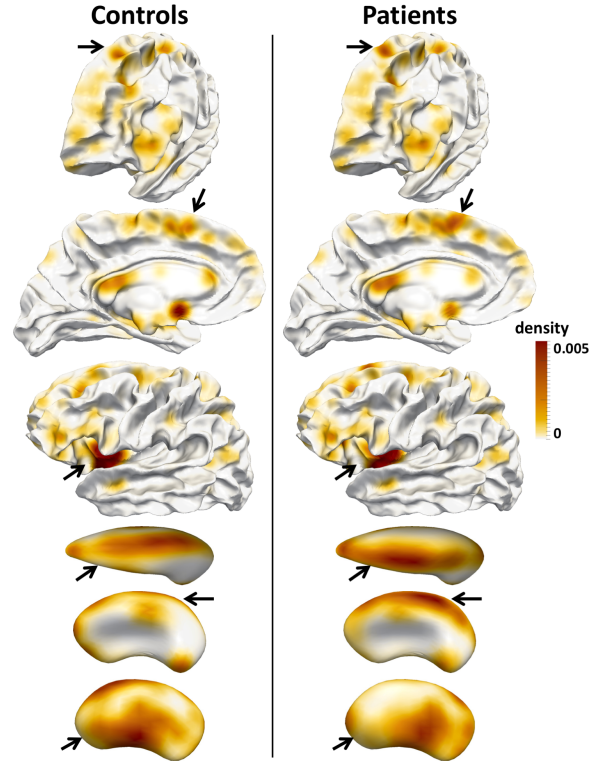

Fig. 8. Typical structural connectivity of controls and patients obtained by deforming the fiber bundle of the template complex along the most discriminative deformation axis in the space of the initial momenta of the first diffeomorphism $\boldsymbol{\alpha}_{i0}^{W}$. Grey matter structures are kept fixed. Colours refer to the density of the extremities of the fiber bundle onto the grey matter.

IV. DISCUSSION AND CONCLUSIONS

We presented a double-diffeomorphic mesh-based atlas construction method. In contrast to standard single-diffeomorphic registrations, the cascade of two diffeomorphisms can put into correspondence anatomical complexes characterised by a different structural connectivity. We showed that this approach makes it possible to *characterise*, *localise* and *quantify* both organisational and morphological pathological anomalies altering grey and white matter structures.
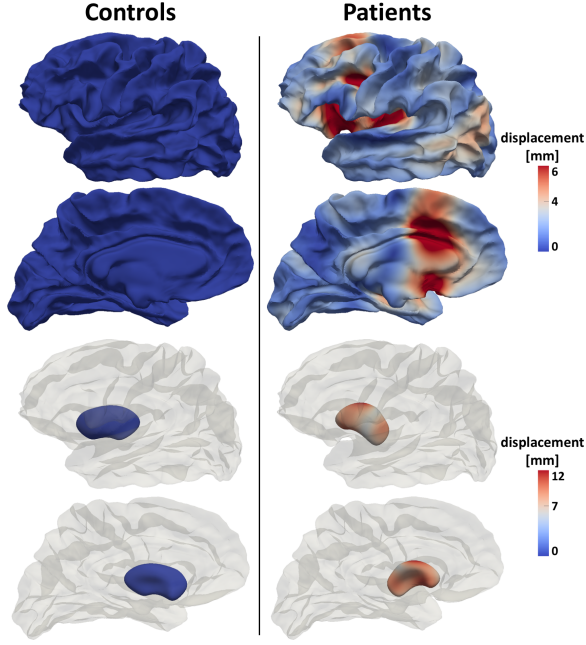
Controls   Patients

Fig. 9. Typical grey matter configurations for controls and patients. They are obtained by deforming the grey matter structures of the template complex along the most discriminative deformation axis in the space of the initial momenta of the second diffeomorphism $\boldsymbol{\alpha}_{i0}^{All}$. Colours refer to the displacement of the configuration of patients from the one of controls.

It is important to notice that it is fundamental to *first* deform the white matter of the template complex and *then* the grey matter in order to retrieve the main variations in structural connectivity. In fact, the first diffeomorphisms $\{\phi_i^W\}$ are comparable across subjects since they are all computed with respect to the same reference frame, namely the *fixed* grey matter of the template complex. If one changed the order, deforming first the grey matter and then the white matter, it would not be possible to compare the variations in structural connectivity since the reference frame, given by the grey matter, would be different across subjects. A diagram describing these two approaches can be found in the Appendix.

White matter fiber bundles are not constrained to always stay in contact with the grey matter during the deformation. We only enforce, by modelling streamlines as weighted currents, that they will be close to the grey matter *at the end* of the second diffeomorphism (See Eq.5). To note that, the two diffeomorphisms are not explicitly weighted during the optimization procedure in Eq.11 - 14. However, they both depend on the gradients of the data-terms, and therefore on the parameters of their corresponding computational models. Furthermore, the precision and flexibility of each diffeomorphism depend on its kernel bandwidth. In this work, since we aim to correctly match both white and grey matter structures, we implicitly gave the same weight to ϕ^W and ϕ^{All} by choosing the same kernel bandwidth (i.e. $\boldsymbol{K}_D$).

A question that naturally arises using the proposed method is about the uniqueness of the decomposition into two diffeomorphisms in regions containing only white matter structures. In these areas, fiber bundles could be deformed into two different but equivalent ways. Using a kernel bandwidth of

11mm for the second diffeomorphism, the deformation of the white matter is correlated to the one of the grey matter. This makes the model identifiable with a unique decomposition of the two diffeomorphisms all over the ambient space.

Both diffeomorphisms are parametrised with control points which define the dimension of the initial momenta. These can be used as input features in a classification task, as in Sec.III-F. In [20], the authors used a single-diffeomorphic atlas construction method similar to the one proposed here. They demonstrated that the statistical performance of a linear classifier augments by decreasing the number of control points until a certain threshold. It seems therefore reasonable to expect the same behaviour for the proposed method. This brings to another question which is how to choose the position and number of the control points. A possible solution was presented in [42]. The authors proposed to integrate in the optimization the selection of the best control points using a penalty similar to Group-Lasso. They started from a regular grid which was trimmed by keeping only the control points that participate to the deformations of all subjects. It would be of interest to integrate this approach to the proposed model.

Another interesting extension might be the use of sparse multi-scale diffeomorphisms such as in [28], [29]. This would probably complicate the statistical analysis but it might also reduce the computational time, using for instance a coarse-to-fine approach as in [29], remove the need for scale tuning of $\boldsymbol{K}_D$, produce compact representations of deformation parameters at different scale and increase registration accuracy.

All experiments shown in this paper were based on a single fiber bundle. However, the neural circuits of the brain are composed of several fiber bundles which could be affected by different pathological alterations. This means that every fiber bundle should be deformed in an independent way with respect to the others. The proposed approach would not be appropriate since the first (white) diffeomorphism would act simultaneously on all fiber bundles. A possible solution would be to substitute the first diffeomorphism with N diffeomorphisms, where N would be equal to the number of fiber bundles. Every bundle would be then independently deformed by a diffeomorphism. In this way, we could capture the variations in structural connectivity proper to each bundle and the global morphological changes associated to the entire neural circuit.

In the proposed method, we assumed that the initial momenta of the two diffeomorphisms are independent, that is to say that $p(\boldsymbol{\alpha}_i^{All}, \boldsymbol{\alpha}_i^W) = p(\boldsymbol{\alpha}_i^{All})p(\boldsymbol{\alpha}_i^W)$, even if the update rule for $\boldsymbol{\alpha}_i^{All}$ and $\boldsymbol{\alpha}_i^W$ are related as explained in Sec.II-E. It would seem more reasonable to take that into account by modelling directly $p(\boldsymbol{\alpha}_i^{All}, \boldsymbol{\alpha}_i^W)$ without the assumption of independence. We could model, for instance, their joint distribution as a single Gaussian distribution. However, the statistical relationship between $\boldsymbol{\alpha}_i^{All}$ and $\boldsymbol{\alpha}_i^W$ is highly complex since they are related by the linearised ODEs shown in Sec.II-E and we have not found yet a satisfactory solution to model their joint distribution. This is left as future work.

Nevertheless, we demonstrated that the proposed double diffeomorphic approach captures useful and relevant information since it better discriminates between controls and patients than a single diffeomorphism. In particular, we observed that

the information about structural connectivity might play an important role in the characterisation of the pathophysiological mechanisms underlying GTS.

Acknowledgment

References

[1] K. Martinu and O. Monchi, "Cortico-basal ganglia and cortico-cerebellar circuits in Parkinson's disease: Pathophysiology or compensation?" *Behavioral Neuroscience*, vol. 127, no. 2, pp. 222–236, 2013. 1

[2] 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. 1, 2, 7, 9

[3] M. Niethammer, M. Reuter, F.-E. Wolter, S. Bouix, N. Peinecke, M.-S. Koo, and M. E. Shenton, "Global Medical Shape Analysis Using the Laplace-Beltrami Spectrum," *MICCAI*, no. 10, pp. 850–857, 2007. 1

[4] A. Qiu, T. Brown, B. Fischl, J. Ma, and M. Miller, "Atlas Generation for Subcortical and Ventricular Structures With Its Applications in Shape Analysis," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1539–1547, Jun. 2010. 1

[5] K. Gorczowski, M. Styner, J. Y. Jeong, J. Marron, J. Piven, H. Hazlett, S. Pizer, and G. Gerig, "Multi-Object Analysis of Volume, Pose, and Shape Using Statistical Discrimination," *Pattern Analysis and Machine Intelligence*, vol. 32, no. 4, pp. 652–661, Apr. 2010. 1

[6] G. Postelnicu, L. Zollei, and B. Fischl, "Combined Volumetric and Surface Registration," *IEEE Transactions on Medical Imaging*, vol. 28, no. 4, pp. 508–522, 2009. 1

[7] 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 *MICCAI*, 2012, no. 7511, pp. 57–65. 1, 2

[8] J. Du, L. Younes, and A. Qiu, "Whole brain diffeomorphic metric mapping via integration of sulcal and gyral curves, cortical surfaces, and images," *NeuroImage*, vol. 56, no. 1, pp. 162–173, 2011. 1

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active Shape Models-Their Training and Application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995. 1

[10] F. L. Bookstein, *Morphometric Tools for Landmark Data. Geometry and Biology.*, 1997. 1

[11] M. Styner, I. Oguz, S. Xu, C. Brechbühler, D. Pantazis, J. J. Levitt, M. E. Shenton, and G. Gerig, "Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM," *The insight journal*, no. 1071, pp. 242–250, 2006. 1

[12] 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. 1, 2

[13] J. Cates, P. T. Fletcher, M. Styner, M. Shenton, and R. Whitaker, "Shape Modeling and Analysis with Entropy-Based Particle Systems," in *IPMI*, 2007, no. 4584, pp. 333–345. 1

[14] H. Hufnagel, X. Pennec, J. Ehrhardt, N. Ayache, and H. Handels, "Computation of a Probabilistic Statistical Shape Model in a Maximum-a-posteriori Framework:," *Methods of Information in Medicine*, vol. 48, no. 4, pp. 314–319, Jun. 2009. 1

[15] P. Fletcher, C. Lu, S. Pizer, and S. Joshi, "Principal geodesic analysis for the study of nonlinear statistics of shape," *IEEE Transactions on Medical Imaging*, vol. 23, no. 8, pp. 995–1005, Aug. 2004. 1

[16] U. Grenandner, *General pattern theory - A mathematical study of regular structures*, 1993. 1

[17] C. Davatzikos, M. Vaillant, S. M. Resnick, J. L. Prince, S. Letovsky, and R. N. Bryan, "A computerized approach for morphological analysis of the corpus callosum," *Journal of Computer Assisted Tomography*, vol. 20, no. 1, pp. 88–97, Feb. 1996. 2

[18] S. Joshi and M. Miller, "Landmark matching via large deformation diffeomorphisms," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1357–1370, Aug. 2000. 2, 6

[19] B. Avants and J. C. Gee, "Geodesic estimation for large deformation anatomical shape averaging and interpolation," *NeuroImage*, vol. 23, Supplement 1, pp. S139–S150, 2004. 2

[20] 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. 2, 5, 6, 10

[21] R. C. Craddock, S. Jbabdi, C.-G. Yan, J. T. Vogelstein, F. X. Castellanos, A. Di Martino, C. Kelly, K. Heberlein, S. Colcombe, and M. P. Milham, "Imaging human connectomes at the macroscale," *Nature Methods*, vol. 10, no. 6, pp. 524–539, Jun. 2013. 2

[22] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, Mar. 2009. 2

[23] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman, "Joint Morphometry of Fiber Tracts and Gray Matter Structures Using Double Diffeomorphisms," in *IPMI*, Jun. 2015, no. 9123, pp. 275–287. 2, 6

[24] S. Durrleman, M. Prastawa, G. Gerig, and S. Joshi, "Optimal Data-Driven Sparse Parameterization of Diffeomorphisms for Population Analysis," in *IPMI*, 2011, no. 6801, pp. 123–134. 2, 4

[25] L. Risser, F.-X. Vialard, H. Y. Baluwala, and J. A. Schnabel, "Piecewise-diffeomorphic image registration: Application to the motion estimation between 3d CT lung images with sliding conditions," *Medical Image Analysis*, vol. 17, no. 2, pp. 182–193, Feb. 2013. 2

[26] D. Pace, S. Aylward, and M. Niethammer, "A Locally Adaptive Regularization Based on Anisotropic Diffusion for Deformable Image Registration of Sliding Organs," *IEEE Transactions on Medical Imaging*, vol. 32, no. 11, pp. 2114–2126, Nov. 2013. 2

[27] S. Arguillère, E. Trélat, A. Trouvé, and L. Younes, "Registration of multiple shapes using constrained optimal control," *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, pp. 344–385, 2016. 2

[28] S. Sommer, F. Lauze, M. Nielsen, and X. Pennec, "Sparse Multi-Scale Diffeomorphic Registration: The Kernel Bundle Framework," *J Math Imag Vis*, vol. 46, no. 3, pp. 292–308, 2013. 2, 10

[29] M. Tan and A. Qiu, "Large Deformation Multiresolution Diffeomorphic Metric Mapping for Multiresolution Cortical Surfaces: A Coarse-to-Fine Approach," *IEEE TIP*, vol. 25, no. 9, pp. 4061–4074, 2016. 2, 10

[30] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, F. D. V. Fallani, M. Chavez, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman, "Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles," *IEEE Trans Med Imag*, 2016. 2, 4

[31] P. A. Yushkevich, H. Zhang, T. Simon, and J. C. Gee, "Structure-Specific Statistical Mapping of White Matter Tracts," *NeuroImage*, vol. 41, no. 2, pp. 448–461, 2008. 2

[32] N. Charon and A. Trouvé, "The varifold representation of non-oriented shapes for diffeomorphic registration," *SIAM Journal on Imaging Sciences*, vol. 6, no. 4, pp. 2547–2580, 2013. 2, 4

[33] I. Rekik, G. Li, P.-T. Yap, G. Chen, W. Lin, and D. Shen, "A hybrid multishape learning framework for longitudinal prediction of cortical surfaces and fiber tracts using neonatal data," in *MICCAI*, 2016, vol. 23, pp. 210–218. 2

[34] M. Vaillant and J. Glaunès, "Surface Matching via Currents," in *IPMI*, 2005, no. 3565, pp. 381–392. 2

[35] S. Allassonnière, Y. Amit, and A. Trouvé, "Towards a Coherent Statistical Framework for Dense Deformable Template Estimation," *JRSS B*, vol. 69, no. 1, pp. 3–29, Jan. 2007. 3, 4

[36] 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 *MICCAI*, 2013, no. 8149, pp. 267–274. 3

[37] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman, "A bayesian framework for joint morphometry of surface and curve meshes in multi-object complexes," *Medical image analysis*, vol. 35, pp. 458–474, 2017. 3, 4

[38] J. Ma, M. I. Miller, A. Trouvé, and L. Younes, "Bayesian template estimation in computational anatomy," *NeuroImage*, vol. 42, no. 1, pp. 252–261, Aug. 2008. 3

[39] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache, "A statistical model of white matter fiber bundles based on currents," in *IPMI*, 2009, pp. 114–125. 3, 4

[40] 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. 4

[41] P. Gori, "Statistical models to learn the structural organisation of neural circuits from multimodal brain images," Ph.D. dissertation, Université Paris 6, 2016. 4, 5

[42] S. Allassonnière, S. Durrleman, and E. Kuhn, "Bayesian Mixed Effect Atlas Estimation with a Diffeomorphic Deformation Model," *SIAM Journal on Imaging Sciences*, pp. 1367–1395, Jan. 2015. 10