Statistical Analysis of Functions on Surfaces, with an application to Medical Imaging

Eardi Lila*[1] and John A. D. Aston†[2]

[1]Cambridge Centre for Analysis, University of Cambridge
[2]Statistical Laboratory, DPMMS, University of Cambridge

June 26, 2019

Abstract

In Functional Data Analysis, data are commonly assumed to be smooth functions on a fixed interval of the real line. In this work, we introduce a comprehensive framework for the analysis of functional data, whose domain is a two-dimensional manifold and the domain itself is subject to variability from sample to sample. We formulate a statistical model for such data, here called Functions on Surfaces, which enables a joint representation of the geometric and functional aspects, and propose an associated estimation framework. We assess the validity of the framework by performing a simulation study and we finally apply it to the analysis of neuroimaging data of cortical thickness, acquired from the brains of different subjects, and thus lying on domains with different geometries.

1 Introduction

Advances in medical imaging acquisition are constantly increasing the complexity of data representing anatomical objects. In particular, some of these imaging modalities offer a richer representation of anatomical manifolds, as a geometric object coupled with a function defined on the geometric object itself, i.e. a Function on a Surface (FoS). In this work we focus on Functions on Surfaces (FoSs) that are real functions located on domains that are two-dimensional manifolds, where the domains themselves are subject to variability from sample to sample, as shown in Figure 1. In the applied mathematics literature, these are also known with the name of Functional Shapes (Charon and Trouvé,

*e.lila@maths.cam.ac.uk
†j.aston@statslab.cam.ac.uk

2014). However, as it will be clear from the methodological section of this paper, the proposed framework can be extended to deal with more complex situations, such as vector-valued functions describing features arising from multi-modal imaging techniques or the RGB representation of colors, as done in Yao et al. (2017), with the purpose of inferring the underlying geometry. Further extensions could also include situations where the functions have an inherent time component. For simplicity of exposition, we will concentrate on univariate FoS data in this paper.

The aim of the present paper is the introduction of a comprehensive statistical framework for the analysis of FoSs. To this end, a statistical model is introduced, with the main aim of jointly representing the *geometric variability* and *functional variability* of the data. Suppose there is an underlying true one-to-one correspondence between the points on the geometries of the observed FoSs. By geometric variability we mean variations on the shape of the domains, i.e. variations of the point positions from one FoS to another. By functional variability we mean variations on the amplitude of the functions of the observed FoSs, at the points in correspondence. For instance, it is evident that the three FoSs in Figure 1, show both geometric variability and functional variability. In order to quantify these two types of variability, we introduce estimators of the underlying unknown quantities within the proposed statistical model.

Moreover, as it is clear from their definitions, the study of geometric variability cannot be performed independently from the study of functional variability, as the results of the latter generally depend on the former. This motivates the introduction of a novel diffeomorphic registration algorithm for functional data whose domain is a two-dimensional manifold, which enables the exploitation of the functional information to achieve a better registration.

Many ideas from the literature on image registration (see e.g. Thirion, 1995, 1998; Dupuis et al., 1998) and the literature on landmarked shapes (see e.g. Bookstein, 1997a,b) have been recently extended to the more general setting of surfaces, without a functional component, both from the applied mathematics prospective (Younes, 2010) and from the statistical prospective (Patrangenaru and Ellingson, 2015). It is also natural to contextualize FoSs in the Functional Data Analysis (FDA) framework. However, FDA is generally

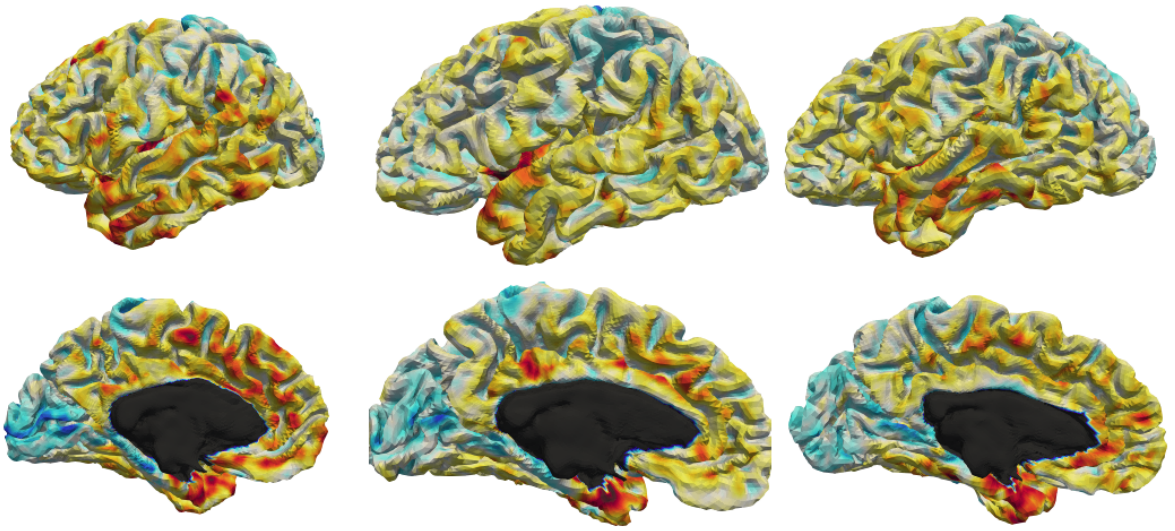

Figure 1: Surface reconstructions of the brain's left hemispheres of three different subjects, with an associated scalar signal representing the cerebral cortex thickness of the subjects. These have been reconstructed from 3D MRI scans of the subjects. The black area is a region which is not part of the brain surface.

performed in controlled environments, where data are assumed to be smooth functions on a fixed interval of the real line (Ramsay and Silverman, 2005), or more generally, smooth functions on a fixed domain. The setting considered here represents a new challenge for this branch of statistics.

More recently, a joint mathematical model for geometric and functional variability has been proposed in Charlier et al. (2017). The approach consists of generalizing the notion of deformation to a notion of metamorphosis, introduced for 2D images in Trouvé and Younes (2005). A metamorphosis includes both a geometric deformation term and an additive functional term. This enables the representation of any FoS as a metamorphosis of a template FoS. The geometric deformation and the functional additive term, to explain a given FoS, can be weighted by two different parameters in the model. In contrast, our approach takes a statistical perspective on the problem of analyzing a set of FoSs, and aims to offer a methodological toolset that can be feasibly applied to the analysis of the brain surfaces shown in Figure 1.

1.1 Motivating application

The motivating application of the proposed model is the study of a collection of FoSs derived from Magnetic Resonance Imaging (MRI). A 2D surface representing the geometry of the cerebral cortex, the outermost layer of the brain, can be extracted from 3D MRI data thanks to fully automated surface-extraction algorithms (Glasser et al., 2013). The cerebral cortex is a highly convoluted thin sheet of 2 to 4 millimeters of thickness which consists of neuronal cell bodies and it is the source of large parts of our neuronal activity. An illustration of the surface-extraction step is shown in Figure 2.

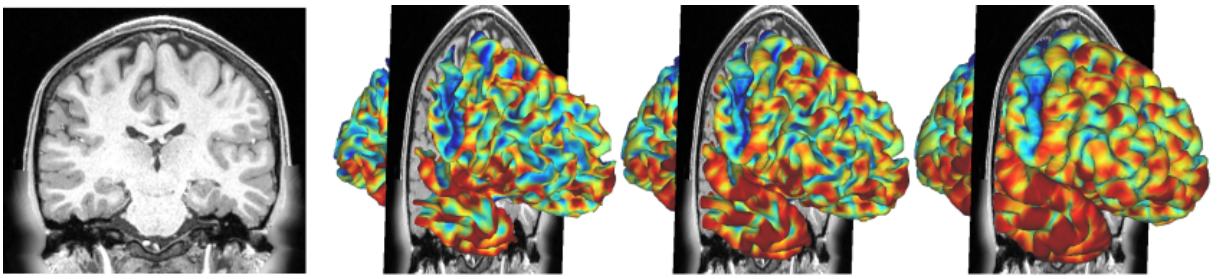

Figure 2: From left to right, in the first panel we have a section of a structural MRI of a single subject. In the next three panels we have, respectively, the estimated inner, mid-thickness and outer surfaces of the cerebral cortex. The inner and outer surfaces enclose the cerebral cortex, the mid-thickness surface interpolates its middle points and it is used to represent its geometry. These images have been produced using Connectome Workbench (Glasser et al., 2013).

Thanks to complementary imaging techniques, like functional MRI, a function can be associated to the estimated cerebral cortex (see, e.g., Hagler et al., 2006), resulting in a FoS. Such functions can be vector-valued functions, where each component represents a feature of the cerebral cortex, extracted from a different imaging technique. However, in this work, the function we consider is the map of thickness measurements of the cerebral cortex. In fact, thanks to the recent improvements of the resolution of MRI scans it is now possible to have an accurate estimation of this thickness map (Lerch and Evans, 2005). Details on the cerebral cortex surface reconstruction and the cortical thickness estimation are covered in Section 4. In Figure 1 we show the FoSs representing cerebral cortex geometry and thickness of three different individuals.

Almost all studies of these kind presume a preprocessing registration step and so do not consider the inherent variability effects that might be induced by the registration step on the functional measurements. Indeed, this issue goes beyond neuroimaging, as the same techniques are often used in a wide variety of medical imaging settings (Audette et al., 2000), as well as computer vision applications (Zaetz and Kurtek, 2015).

The rest of the article is organized as follows. In Section 2 we introduce a generative statistical model which allows for both geometric and functional variability. In Section 3, we propose the statistical estimators of the underlying unknown quantities of the generative model. We then apply the framework introduced to study the relation between geometry and thickness of the human cerebral cortex in Section 4 and draw some concluding remarks in Section 5. Moreover, in the Supplementary Material, we present further details of the proposed methodology and a simulation study on synthetic data to investigate our estimation procedure.

A longer version of this article, where the Supplementary Material is incorporated in the main manuscript, can be found in Lila and Aston (2017).

2 Model for Functions on Surfaces

2.1 Definitions

A set of FoSs, such as the ones in Figure 1, can be mathematically formulated as a collection of pairs $\{(\mathcal{M}_i, Y_i) : i = 1, \ldots, n\}$. The collection $\{\mathcal{M}_i : i = 1, \ldots, n\}$ is a set of topologically equivalent smooth two-dimensional manifolds, embedded in $\mathbb{R}^3$, representing the geometry of the data. The functional aspect of the data is represented by the collection $\{Y_i : i = 1, \ldots, n\}$, where Y_i is an element of the function space $L^2(\mathcal{M}_i)$, i.e. the Hilbert space of square integrable functions on $\mathcal{M}_i$ with respect to the area measure.

Here, we propose a statistical generative model for FoSs, modelled in terms of mathematically more tractable objects. To this end, we define a *deformation operator* φ, such that $\varphi_v : \mathbb{R}^3 \to \mathbb{R}^3$ is parametrized by the elements of a Hilbert space $\{v : v \in \mathcal{V}\}$. Moreover, we assume φ_v is an homomorphism of $\mathbb{R}^3$ for all $v \in \mathcal{V}$ and that $\varphi_0(x) = x$ for all $x \in \mathbb{R}^3$. For each $v \in \mathcal{V}$, $\varphi_v : \mathbb{R}^3 \to \mathbb{R}^3$ represents a deformation of the space $\mathbb{R}^3$,

which means that when φ_v is applied to a point $x \in \mathbb{R}^3$ this is relocated to the location $\varphi_v(x) \in \mathbb{R}^3$. In addition, φ_v being a homomorphism of $\mathbb{R}^3$ implies that, for a fixed $v \in \mathcal{V}$, there is a one-to-one correspondence between each element $x \in \mathbb{R}^3$ and the relocated element $\varphi_v(x) \in \mathbb{R}^3$.

Moreover, we introduce $\mathcal{M}_0$, a smooth two-dimensional manifold topologically equivalent to $\{\mathcal{M}_i\}$, which represents a fixed template geometric object. Given a FoS, the geometric template together with the deformation operator offers an alternative representation of the geometry of the FoS in hand as: $\varphi_v \circ \mathcal{M}_0$, for a particular choice of $v \in \mathcal{V}$. Here, $\varphi_v \circ \mathcal{M}_0$ is the geometric object obtained by deforming $\mathcal{M}_0$ through the map φ_v, and specifically, by relocating each point $x \in \mathcal{M}_0$ to the new location $\varphi_v(x)$, to resemble the target manifold. For this reason, we will informally say that the element $v \in \mathcal{V}$ encodes the geometry, or the shape, of a FoS, as in fact v defines the deformation φ_v, which defines the geometry $\varphi_v \circ \mathcal{M}_0$. The choice of the deformation operator is driven by the particular problem in hand. We first introduce the generative model and subsequently discuss different choices of this operator.

2.2 The model

Let now $\{v_i : i = 1, \ldots, n\}$ be a set of random samples of a zero-mean and finite second moment $\mathcal{V}$-valued random function V and $\{Z_i : i = 1, \ldots, n\}$ be a set of random smooth samples of a zero-mean and finite second moment random real function Z with values in $L^2(\mathcal{M}_0)$. We assume the following generative model for the ith observation $(\mathcal{M}_i, Y_i)$:

$$\begin{cases} \mathcal{M}_i &= \varphi_{v_i} \circ \mathcal{M}_0, \\ X_i &= \mu + \delta Z_i, \\ Y_i &= X_i \circ \varphi_{v_i}^{-1}, \end{cases} \tag{1}$$

where $\mu \in L^2(\mathcal{M}_0)$ is a fixed function, modelling the common function behavior between the different samples, and δ is a coefficient representing the magnitude of the function variations around the mean μ. In addition, we assume the objects in Model 1 are subject to a discretization error, which is considered in the estimation process. This formulation generalizes an often used model for the one-dimensional functional registration problem

(see, e.g. Tang and Muller (2008)).

Model 1 achieves the goal of representing FoSs as a collection of more tractable objects, decomposing the generation of the ith FoS into three main steps. In the first step, the geometry $\mathcal{M}_i$ of the ith object is generated by the deformation φ_{v_i} applied to the template $\mathcal{M}_0$, where v_i is a random sample from V. In the second step, a random function X_i, on the template, is generated as the sum of the fixed function μ and a stochastic term δZ_i. In the third step the generated function X_i is transported on the manifold $\mathcal{M}_i$, defining Y_i. This is done through the equation $Y_i = X_i \circ \varphi_{v_i}^{-1}$, which means that for all $x \in \mathcal{M}_0$, $Y_i(\varphi_{v_i}(x)) = X_i(x)$, or informally that the functional value $X_i(x)$ is 'transported' with the deformation to the location $\varphi_{v_i}(x) \in \mathcal{M}_i$.

We now describe the FoSs generation process from Model 1, for different choices of the deformation operator:

- *Shift operator*: Let $\mathcal{V} = \mathbb{R}^3$, we define φ_v to be such that $\varphi_v(x) = x + v$ for all $v \in \mathcal{V}, x \in \mathbb{R}^3$. Clearly, in this case, $\{\mathcal{M}_i = \varphi_{v_i} \circ \mathcal{M}_0\}$ in Model 1 would generate a collection of surfaces shifted in the directions specified by $\{v_i\}$.

- *Identity operator*: Let $\mathcal{V}$ be the space of smooth functions $v : \mathcal{M}_0 \to \mathbb{R}^3$ and let $\varphi_v(x) = x + v(x)$ for all $x \in \mathcal{M}_0$. In this case, $\{\mathcal{M}_i\}$ would be a collection of smoothly deformed versions of the template $\mathcal{M}_0$. Note however, that the maps being only smooth and not homeomorphic, it cannot be guaranteed that every choice of $v \in \mathcal{V}$ preserves the topology of $\mathcal{M}_0$. Nevertheless, this choice might still represent a valid option in a small deformations setting.

  To solve this problem, we could think of restricting $\mathcal{V}$ to contain only smooth and homeomorphic functions, however, in this way, the linearity of the space $\mathcal{V}$ is lost, and this is a property of fundamental importance to the subsequent analysis, given that we want to apply linear statistics on the random function V, which takes values on $\mathcal{V}$.

- *Diffeomorphic operator*: Let $\mathcal{V}$ be a Sobolev space of sufficiently smooth vector fields from $\mathbb{R}^3$ to $\mathbb{R}^3$ vanishing, with their derivatives, at infinity. Let φ be a diffeomorphic deformation operator, i.e. an operator such that φ_v is a diffeomorphism of $\mathbb{R}^3$ for

all $v \in \mathcal{V}$. Then, for different choices of v, Model 1 would generate a collection of surfaces that are diffeomorphic (and thus homeomorphic) deformations of the template $\mathcal{M}_0$. More importantly, these deformations are parametrized by the linear space $\mathcal{V}$, where linear statistics can be applied. For this choice, an illustration of the generative process is shown in Figure 3. The diffeomorphic deformation operator can be defined by means of an Ordinary Differential Equation (ODE). Details of this are described in Section 2.4.

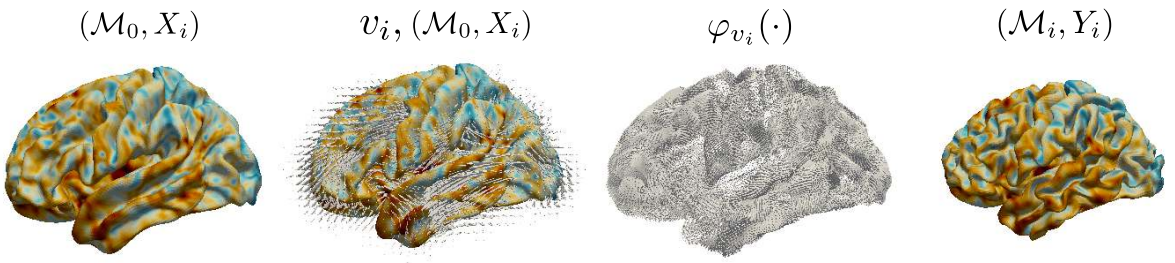

Figure 3: An illustration of the generation of a FoS through Model 1 with φ the diffeomorphic deformation operator. From left to right, in the first panel we have a functional sample X_i on the geometric template $\mathcal{M}_0$. In the second panel we have a vector field $v_i \in \mathcal{V}$ sampled from the random function V, and evaluated on a uniform grid in $\mathbb{R}^3$. This is shown together with $(\mathcal{M}_0, X_i)$. In the third panel we have the diffeomorphic deformation φ_{v_i}, obtained from v_i as described in Section 2.4, here displayed as the set of vectors $\{\varphi_{v_i}(\xi_k)\} \subset \mathbb{R}^3$ with $\{\xi_k\}$ the nodes of the triangulated surface representing the template $\mathcal{M}_0$. In the fourth panel, we have the FoS $(\mathcal{M}_i, Y_i)$ obtained by applying the deformation φ_{v_i} to $\mathcal{M}_0$ and 'transporting' the functional values with it.

More complicated generative models could be built from Model 1. For example, the functions $\{v_i\}$ and $\{X_i\}$, representing respectively geometries and functions, could be modelled in terms of conditional expectation of different sources of information on the subjects such as age, status disease or other subject-specific explanatory variables, as done, in the case of functional data located on 1D domains, in Hadjipantelis et al. (2015). However, Model 1 is the simplest model enabling a comprehensive study of the relation between geometric and functional variability.

2.3 Geometric and Functional variability

Here we formalize the geometric and functional variability relationship. Recalling the definition of geometric and functional variability, given in the introduction, we can notice that in Model 1 we have that $\{v_i\}$ describe the geometric variability in the data, while $\{X_i\}$ describe the functional variability in the data. The key idea of this work is to formalize geometric and functional variability by means of functional Principal Component Analysis (fPCA), so that geometric variability can be represented in terms of the Principal Components (PCs) of the random function V, generating the samples $\{v_i\}$, and functional variability can be represented in terms of the PCs of the random function $X = \mu + \delta Z$ generating the random samples $\{X_i\}$. Under the hypothesis that a finite number of PCs is sufficient to represent V and X, we can than use classical multivariate statistics, such as multivariate regression or canonical correlation analysis, to model the relation between the PCs of V and the PCs of X, ultimately formalizing the concept of geometric and functional variability being related. This should also further clarify the choice to introduce a deformation operator φ. In fact, as already mentioned, the deformation operator allows us to parametrize the space of deformations through the linear space $\mathcal{V}$, and thus linear fPCA can be applied on the $\mathcal{V}$-valued random variable V.

More formally, under typical assumptions on V, thanks to fPCA, V can be expanded in terms of the orthonormal sequence of eigenfunctions $\{\psi_j^G\}$ of the covariance operator of V, as

$$V = \sum_{j=1}^{\infty} a_j^G \psi_j^G,$$

where $a_1^G, a_2^G, \ldots$ are uncorrelated real random variables, with variances in decreasing order $\kappa_1^G, \kappa_2^G, \ldots$. The collection $\{\psi_j^G\}$ defines the strongest modes of variation of the random function V and these are called PC functions. We refer to ψ_j^G as the jth mode of geometric variation, or alternatively the jth geometric PC function. This represents variations of the type $c\psi_j^G$ around the mean of V, with $c \in \mathbb{R}$. The PC function ψ_j^G is thus associated to the geometric deformations $\varphi_{c\psi_j^G}$ of $\mathbb{R}^3$, that applied to the geometric template correspond to the geometries described by $\varphi_{c\psi_j^G} \circ \mathcal{M}_0$. In practice, we visualize the jth mode of geometric variation by visualizing the associated geometries for some

specific choice of c, e.g. $\varphi_{\pm\sqrt{\kappa_j^G}\psi_j^G} \circ \mathcal{M}_0$. An example of this visualization is given in Figure 8. PCA have been previously used in a similar fashion in Vaillant et al. (2004) and Tward et al. (2017), with φ the diffeomorphic deformation operator, to represent anatomical geometries.

With analogous considerations, the random variable X can be expanded, using the associated orthonormal eigenfunctions $\{\psi_j^F\}$, of the covariance operator of X, as

$$X = \mu + \sum_{j=1}^{\infty} a_j^F \psi_j^F,$$

where the real random variables $a_1^F, a_2^F, \ldots$ are uncorrelated with variances, in decreasing order, $\kappa_1^F, \kappa_2^F, \ldots$. We refer to ψ_j^F as the jth mode of functional variation.

FPCA basis expansions have the fundamental property of separating the discrete set of stochastic terms from the functional terms. Hence, the relation between the geometry and the functional terms can be formalized in terms of the random variables $\{a_j^G\}$ and $\{a_j^F\}$. We assume that only a finite number of the PC functions are necessary to describe the phenomenon in hand and denote with $\mathbf{a}^G$ the associated K^G-dimensional random vector $(a_1^G, \ldots, a_{K^G}^G)$ and with $\mathbf{a}^F$ the K^F-dimensional random vector $(a_1^F, \ldots, a_{K^F}^F)$.

Different multivariate statistical models can be applied at this stage, to formalize the geometric and functional variability relation in terms of the relation between the random vectors $(a_1^G, \ldots, a_{K^G}^G)$ and $(a_1^F, \ldots, a_{K^F}^F)$. A first possible formalization of the geometric and functional variability relation is

$$\mathbb{E}[X|V] = \mu + \sum_{j=1}^{K^F} \mathbb{E}[a_j^F|\mathbf{a}^G]\psi_j^F. \tag{2}$$

Under linear assumptions on the dependency, the conditional expectation term can be modelled as

$$\mathbb{E}[a_j^F|\mathbf{a}^G] = \boldsymbol{\beta}_j' \mathbf{a}^G,$$

with $\boldsymbol{\beta}_j$ the K^G-dimension vector of the regression coefficients of the jth functional mode of variation.

The model above describes how the main modes of geometric variation explain each mode of functional variation, implying that we expect the geometry to influence the functions. This might be the case of neurodegenerative disease, where we expect the

functional activity (the function) to adapt to the disease progression (the geometry). However, the reverse roles of geometry on functions is also plausible in some cases. For instance, through a comparative study between taxi drivers and bus drivers, it has been shown that the different functional activation patterns influence the growth of the gray matter volume, and thus the brain geometry (Maguire et al., 2006). Moreover, given that in model (2) each mode of functional variations is explained thorugh a linear combination of the modes of geometric variability, the interpretability of the overall model strongly relies on the interpretability of the singular functional main modes of variations.

A second possible formalization of the geometric and functional variability relationship might consist of simply examining the maximal directions of correlation between geometry and function. This is equivalent to performing a Canonical Correlation Analysis (CCA). A CCA analysis on the coefficients of the fPCA basis expansion is equivalent to finding a new basis expansions for V and X as a linear combination of the respective fPCA basis. However, the elements of the new basis are ordered in a way that maximizes the correlation between their coefficients, i.e. the interdependency between geometry and function, representing how the geometric variability associates with the functional variability and vice versa.

2.4 The diffeomorphic deformation operator

The deformation operator, introduced in the Section 2.1, has to be chosen in such a way that it is flexible enough to represent the observed surfaces, as a deformation of the template surface. Clearly, the shift operator is not sufficient to capture the variations in geometry of the FoSs in Figure 1, in terms of deformation of the template. However, this operator should only include 'sensible' deformations, in the sense that the deformation operator should have its image contained in the set of diffeomorphic deformations from $\mathbb{R}^3$ to $\mathbb{R}^3$. This choice is driven by the fact that diffeomorphic deformations are smooth deformations that preserve the topological properties of the shapes and that avoid two separate points on the template collapsing to one point on the observed surface.

For this reason, we rely on the idea of constructing diffeomorphic deformations as flows of an ODE (Dupuis et al., 1998), which can be parameterized by a Hilbert function

space. Specifically, let now $\mathcal{V}$ be a Sobolev space of sufficiently smooth vector fields from $\mathbb{R}^3$ to $\mathbb{R}^3$ vanishing, with their derivatives, at infinity. Let $v : [0,1] \times \mathbb{R}^3 \to \mathbb{R}^3$ be a *time dependent* vector field in $L^2([0,1], \mathcal{V})$, the space of vector fields with finite (squared) norm $\int_0^1 \|v_t\|^2_{\mathcal{V}} dt$. Then, for a given v, the solution $\phi_v : [0,1] \times \mathbb{R}^3 \to \mathbb{R}^3$ of the ODE

$$\frac{\partial \phi_v}{\partial t}(t,x) = v_t \circ \phi_v(t,x) \qquad t \in [0,1], x \in \mathbb{R}^3. \tag{3}$$

with initial conditions $\phi_v(0,x) = x$, is a smooth diffeomorphic map in $\mathrm{Diff}(\mathbb{R}^3)$, at each fixed time t (see, e.g., Younes, 2010). The ODE (3) is intuitively defining the solution ϕ_v to be a function such that, for all $t \in [0,1]$, the time-derivative $\frac{\partial \phi_v}{\partial t}(t,x)$ (i.e. the velocity field at time t) is given by the vector field $v_t \circ \phi_v(t,x)$. In other terms ϕ_v represents the 'flow' described by the velocity vector field $\{v_t : t \in [0,1]\}$.

Note that we use φ and ϕ to represent two different object and their relation is defined as follows.

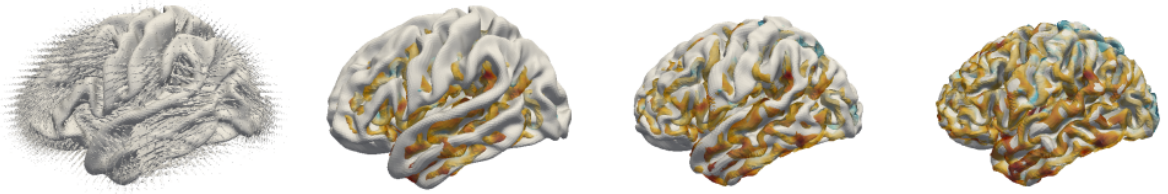

Figure 4: From left to right, in the first panel we have an initial vector field $v_0 \in \mathcal{V}$ and in gray the template $\mathcal{M}_0$. In the consecutive panels, we show the solution ϕ_v of the ODE at the times $t = 0, 0.5, 1$ (which are diffeomorphic deformations of $\mathbb{R}^3$), as deformations of the template $\mathcal{M}_0$. In this specific case, the initial vector field v_0 has been chosen in such a way that the surface $\varphi_{v_0} \circ \mathcal{M}_0$ is a close approximation of a target surface, i.e. the colored surface in the figure.

Given an initial vector field v_0, we define $\{v_t : t \in [0,1]\}$ to be the time-variant vector field which minimizes the quantity $\int_0^t \|v_t\|^2_{\mathcal{V}} dt$. For this choice, the vector field $\{v_t : t \in [0,1]\}$ can be derived from v_0 through the resolution of the EPDiff equation (Miller et al., 2006). Finally, the deformation operator can be defined to be $\varphi_{v_0}(x) = \phi_v(1,x)$, where $v_0 \in \mathcal{V}$ is the initial vector field generating $\{v_t : t \in [0,1]\}$, through the EPDiff equation, and ϕ_v is the solution of the ODE (3). The choice to define $\varphi_{v_0}(x)$ with the solution of the ODE (3) at time $t = 1$ is arbitrary, in fact any other choice of a fixed $t > 0$

would have been equivalent, given that the $\phi_v(t, x)$ is guaranteed to be diffeomorphism of $\mathbb{R}^3$ for any $t > 0$.

A summary of the main elements necessary to define φ is given by the following

$$\underbrace{v_0 \xrightarrow[\text{EPDiff}]{} \{v_t : t \in [0,1]\} \xrightarrow[\text{ODE (3)}]{} \phi_v \to \varphi_{v_0} := \phi_v(1, \cdot)}_{\varphi_{v_0}:\mathbb{R}^3 \to \mathbb{R}^3} . \tag{4}$$

In Figure 4 we show the solution of the ODE (3) for a given initial vector field v_0. We emphasize that the ODE (3) is not used here to model the phenomenon in hand, but it is just a convenient tool to generate a diffeomorphism of $\mathbb{R}^3$ from a smooth vector field v_0 belonging to the linear space $\mathcal{V}$.

3 Estimation framework

The arguments made in the previous section are formalized in terms of quantities derived from the underlying unknown random variables modelling the data generation. However, in practice, only a set of observed noisy FoSs is available, and those quantities have to be estimated from the data. In this section, we mostly work with the set of idealized FoSs $\{(\mathcal{M}_i, Y_i) : i = 1, \ldots, n\}$. Instead, when the specific computer representation is of importance to the proposed algorithms, we work with the associated collection of pairs denoted with $\{(\mathcal{M}_i^{\mathcal{T}}, Y_i^{\mathcal{T}}) : i = 1, \ldots, n\}$, each composed by a triangulated surface $\mathcal{M}_i^{\mathcal{T}} \subset \mathbb{R}^3$, approximating the underlying smooth two-dimensional manifold $\mathcal{M}_i \subset \mathbb{R}^3$, and a real piecewise linear function $Y_i^{\mathcal{T}} \in L^2(\mathcal{M}_i^{\mathcal{T}})$ representing a noisy approximation of the underlying smooth function $Y_i \in L^2(\mathcal{M}_i)$.

In this section, we outline the estimation procedures applied to the data to recover the different quantities in Model 1. A flow chart summarizing the main steps is shown in Figure 5. The implementation details are covered in the Supplementary Material.

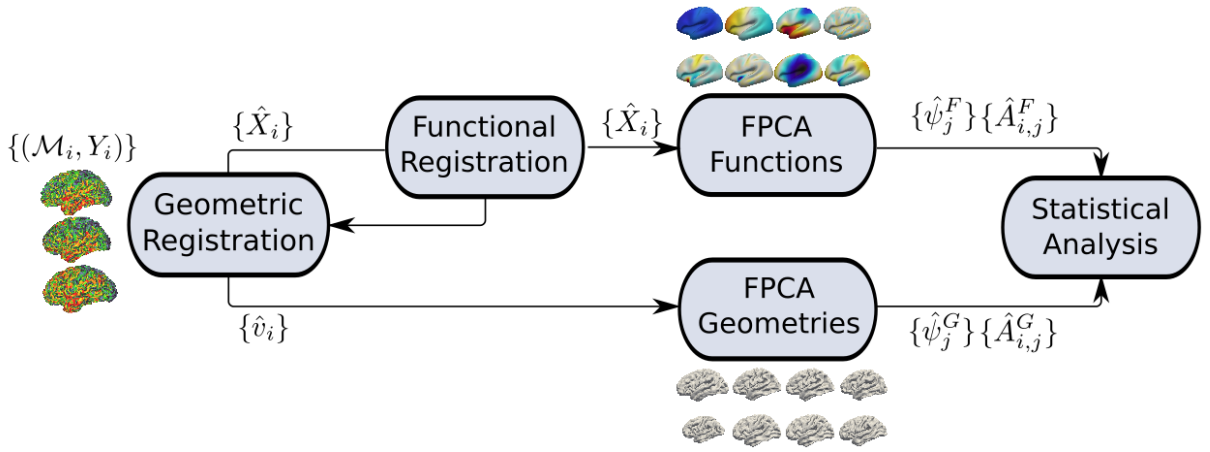

Figure 5: A flow chart summarizing the main steps of the estimation procedure proposed in Section 3. In the Geometric Registration step, as described in Section 3.1, the FoSs are registered to a template $\mathcal{M}_0$, i.e. for each FoS we estimate a vector field $\hat{v}_i$ representing the geometric object $\mathcal{M}_i$ as the deformed template $\varphi_{\hat{v}_i} \circ \mathcal{M}_0$. Each function $\hat{X}_i$ is obtained by transporting Y_i on the template through the estimated registration. In the Functional Registration step, as described in Section 3.2, the functional information is used to achieve a more accurate registration leading to a corrected version of the functions $\{\hat{X}_i\}$ and the vector fields $\{\hat{v}_i\}$. In the FPCA Functions and FPCA Geometries steps, as described in Section 3.3, fPCA is performed on the functions $\{\hat{X}_i\}$ and the vector fields $\{\hat{v}_i\}$ respectively, leading to the estimation of a set of PC functions and scores representing functional variability and geometric variability. Finally, as described in Section 3.4, classical statistical analysis is performed on the PC scores $\hat{A}_{i,j}^F$ and $\hat{A}_{i,j}^G$ to study the relation between functional and geometric variability.

3.1 Geometric Registration and Linear representation of shapes

In practice the computation of a diffeomorphic deformation between the template $\mathcal{M}_0$ and the surface $\mathcal{M}_i$ is achieved by solving a minimization problem of the form

$$\hat{v}_i = \underset{v_i \in \mathcal{V}}{\arg\min}\, D^2(\varphi_{v_i} \circ \mathcal{M}_0, \mathcal{M}_i) + \lambda \|v_i\|_{\mathcal{V}}^2, \tag{5}$$

where $D^2(\varphi_{v_i} \circ \mathcal{M}_0, \mathcal{M}_i)$, the shape similarity function, is a measure of the amount of mismatching between the deformed template surface and the target surface. The constant λ is a weighting parameter between the data-fidelity term and the term $\|v_i\|_{\mathcal{V}}^2$, which could be regarded as a measure of the amount of deformation induced by φ_{v_i}. The functions

$\{\hat{v}_i\}$ are an estimation of $\{v_i\}$ in Model 1. In Figure 4 we show an example of a vector field in $\mathcal{V}$, estimated by solving (5), with the aim of representing a target surface as a deformation of a template.

The procedure described is also referred to as the registration step, as in fact the estimated map $\varphi_{\hat{v}_i}$, up to approximation error, defines a one-to-one smooth correspondence between the points of the target manifold $\mathcal{M}_i$ and the template $\mathcal{M}_0$. Thus, the function $\hat{X}_i \in L^2(\mathcal{M}_0)$, obtained by registering Y_i to the template, can be defined as the element $\hat{X}_i$ such that $\hat{X}_i(x) = Y_i(\varphi_{\hat{v}_i}(x))$ for all $x \in \mathcal{M}_0$. The registered maps $\{\hat{X}_i\}$ can be regarded as a first approximation of $\{X_i\}$ in Model 1. In practice, there might be a small approximation error between $\varphi_{v_i} \circ \mathcal{M}_0$ and $\mathcal{M}_i$, which might contrast with the definition $\hat{X}_i(x) = Y_i(\varphi_{\hat{v}_i}(x))$ for all $x \in \mathcal{M}_0$, as $\varphi_{\hat{v}_i}(x)$ might not exactly belong to $\mathcal{M}_i$. However, we assume that $\varphi_{\hat{v}_i}(x)$ is close enough to $\mathcal{M}_i$, for all $x \in \mathcal{M}_0$, and in practice define $\hat{X}_i(x) = Y_i(y)$ with $y \in \mathcal{M}_i$ the nearest neighbor of $\varphi_{\hat{v}_i}(x)$.

The implementation of the registration algorithm (5) requires the definition of a shape similarity function D. As already mentioned, the geometry of a FoS is in practice encoded as a triangulated mesh, we thus define the similarity function D between triangulated surfaces. We should differentiate between two possible settings at this point. In the first setting, we suppose that a correspondence between the points of the triangulated surfaces $\mathcal{M}_0^{\mathcal{T}}$ and $\mathcal{M}_i^{\mathcal{T}}$ is known for all $i = 1, \ldots, n$. In other terms, we suppose that $\mathcal{M}_0^{\mathcal{T}}$ and $\mathcal{M}_i^{\mathcal{T}}$ have already been registered and thus there is a set of landmarks $\{x_l, y_l : x_l \in \mathcal{M}_0^{\mathcal{T}}, y_l \in \mathcal{M}_i^{\mathcal{T}}\}$ in correspondence between them. In this case, a simple mismatching functional is given by the Euclidean distance between the correspondent landmarks i.e.

$$D^2(\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}, \mathcal{M}_i^{\mathcal{T}}) = \sum_l \|\varphi_{v_i}(x_l) - y_l\|^2_{\mathbb{R}^3}. \tag{6}$$

This choice has been adopted for instance in Joshi and Miller (2000).

This situation is frequent in neuroimaging, a field that has developed their own ad hoc registration algorithms and where diffeomorphic constraints are explicitly imposed without the necessity to use a diffeomorphic deformation operator. In this case the estimates $\{\hat{X}_i\}$ are already provided, given that $\mathcal{M}_0^{\mathcal{T}}$ and $\mathcal{M}_i^{\mathcal{T}}$ have already been registered, nevertheless, the framework introduced here is still of relevance, in fact, we still need to estimate $\{\hat{v}_i\} \subset \mathcal{V}$, in equation (5), to represent the given registration maps (i.e. defor-

mation maps), and thus the geometries, in terms of elements of a linear space, which is a fundamental property to the subsequent analysis.

In the second setting, we suppose that a registration step has not been performed yet. In this situation, registration and linear representation can be performed jointly by choosing an appropriate shape similarity function D not based on landmarks, but for instance, proximity. An example of such similarity function is proposed in Vaillant and Glaunès (2005) and Vaillant et al. (2007), and is defined as follows. Let $K_{\mathcal{Z}} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ be a Gaussian isotropic kernel of variance $\sigma_{\mathcal{Z}}^2$, i.e. $K_{\mathcal{Z}}(x,y) = \exp(-\|x-y\|_2^2/(2\sigma_{\mathcal{Z}}^2))\mathrm{Id}_{3\times 3}$, with $\mathrm{Id}_{3\times 3}$ denoting a 3×3 identity matrix. Indeed, such kernel can be any symmetric positive definite kernel, however it is common to choose a Gaussian kernel. Denote with $c(l)$ and $\eta(l)$, respectively, the center point and the normal vector of the lth triangle of the mesh $\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}$. Denote with $c_i(q)$ and $\eta_i(q)$, respectively, the center point and the normal vector of the qth triangle of the mesh $\mathcal{M}_i^{\mathcal{T}}$. Moreover, let the triangles of the mesh $\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}$ be indexed by l and g and the triangles in $\mathcal{M}_i^{\mathcal{T}}$ be indexed by q and r. The resulting shape similarity function has the form

$$\begin{aligned} D^2(\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}, \mathcal{M}_i^{\mathcal{T}}) = &\sum_l \sum_g K_{\mathcal{Z}}(c(l), c(g))\eta(l)\cdot\eta(g) \\ &- 2\sum_l \sum_q K_{\mathcal{Z}}(c(l), c_i(q))\eta(l)\cdot\eta_i(q) \\ &+ \sum_q \sum_r K_{\mathcal{Z}}(c_i(q), c_i(r))\eta_i(q)\cdot\eta_i(r), \end{aligned} \tag{7}$$

with $\cdot$ denoting the scalar product in $\mathbb{R}^3$. Intuitively, the first and last terms measure deformations to the local geometry within the two surfaces, and the middle term measures the mismatch in local geometry between the two surfaces.

Thanks to the procedure outlined in this section, given a set of FoSs, we are able to register them to a fixed template $\mathcal{M}_0$. As a result, the information regarding the geometry of the data is stored in terms of the estimates $\{\hat{v}_i\} \subset \mathcal{V}$ of $\{v_i\}$ in Model 1. These are estimated so that $\varphi_{\hat{v}_i} \circ \mathcal{M}_0$ resembles the geometry of the ith FoS. Moreover, we obtain a set of functions $\{\hat{X}_i\}$ on the fixed template, that are a first estimate of the functions $\{X_i\}$ in Model 1.

In practice the space of smooth functions $\mathcal{V}$ is implemented as a Reproducing Kernel Hilbert Space (RKHS), as described in the Supplementary Material.

3.2 Functional Registration

The aim of this section is the introduction of a novel functional registration algorithm for functional data whose domain is a fixed two-dimensional manifold. The functional registration algorithm can then be applied to align the set of functions $\{\hat{X}_i : \hat{X}_i \in L^2(\mathcal{M}_0)\}$, estimated in Section 3.1, by registering them to a template function $X_0 \in L^2(\mathcal{M}_0)$, which can be in first instance approximated by the cross-sectional sample mean of $\{\hat{X}_i\}$. The rationale for such a procedure is that, as well known in FDA, the functions $\{\hat{X}_i\}$ on $\mathcal{M}_0$ should in principle be able to drive a better registration, on the assumption that the underlying functions $\{X_i\}$ in Model 1 have a preponderant mean effect, with respect to its second order variation.

In fact, each estimated function $\hat{X}_i$, strongly depends on the associated deformation map $\varphi_{\hat{v}_i}$, whose estimation is usually driven only by geometric features. Hence, a systematic mis-registration, due to a naive approximation of the deformation maps, could introduce fictitious functional variability on the functions $\{\hat{X}_i\}$, which in fact should be accounted for by geometric variability, in particular in a setting where obvious landmarks are not available and the deformations $\{\varphi_{\hat{v}_i}\}$ are estimated while ignoring the functional information. The functional registration algorithm can be regarded as a correction step to $\{\hat{X}_i\}$, and thus $\{\hat{v}_i\}$, estimated from Section 3.1.

A review on the registration of functional data can be found in Marron et al. (2015). However, most of the FDA literature treats only the case of functions whose domain is an interval of the real line. Registration of 2D images has also been well studied (see e.g., Zitová and Flusser, 2003, for a review). Methods that preserve invertibility of the deformation have also been proposed for 2D/3D Euclidean images (Vercauteren et al., 2009) and extended to functions with spherical domains in Yeo et al. (2010). However, to the best of our knowledge, these methods are not able to deal with the registration of a collection of functions whose domain is a fixed generic two-dimensional manifold embedded in $\mathbb{R}^3$.

Alternatively, in the case of landmark based registration, functional information can be introduced into the registration process, by modifying the algorithm that provides the landmarks, to account for function similarity. In the case where landmarks are not avail-

able functional information can be introduced by equipping the shape similarity functional (7) with a functional similarity term, as done in Charon and Trouvé (2014) and Charlier et al. (2017).

3.2.1 Definitions

Let $T_p\mathcal{M}_0$ be the tangent space on the point $p \in \mathcal{M}_0$ and let g_p be the metric on $\mathcal{M}_0$, i.e. a scalar product on the tangent space $T_p\mathcal{M}_0$. In our case it is natural to consider the scalar product induced by the Euclidean embedding space $\mathbb{R}^3$, i.e. the first fundamental form. Define the tangent bundle to be the disjoint union of tangent spaces $T\mathcal{M}_0 = \dot{\bigcup}_{p\in\mathcal{M}_0} T_p\mathcal{M}_0 = \bigcup_{p\in\mathcal{M}_0} \{p\} \times T_p\mathcal{M}_0$. A section of the tangent bundle $T\mathcal{M}_0$ is the formalization of the concept of a vector field on $\mathcal{M}_0$, an example of which is shown in Figure 6. We denote with $L^2(T\mathcal{M}_0)$ the Hilbert space of square integrable sections of $T\mathcal{M}_0$. Moreover, let Δ_{BL} be the Bochner-Laplacian operator. The Bochner-Laplacian of a smooth vector field v, i.e. $\Delta_{BL}v$, is a vector field on $\mathcal{M}_0$, whose L^2 norm gives a measure of the smoothness of the vector field v. i.e. low values for smooth vector fields v and high values for rough vector fields. A more formal definition of the Bochner-Laplacian operator, from the Levi-Civita operator, is given in the Supplementary Material.

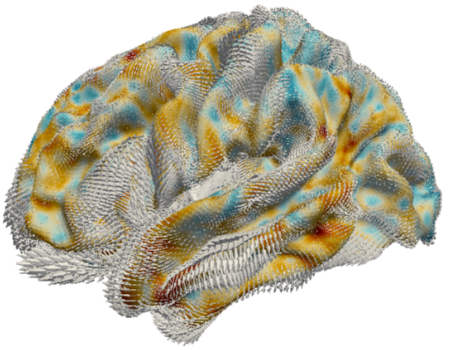

Figure 6: A section of the tangent bundle $T\mathcal{M}_0$, which has been computed by minimizing the linearized version of the equation (9).

3.2.2 Estimation

The registration of $\{\hat{X}_i\}$ is performed in an iterative fashion, which means that each function $\hat{X}_i$ is aligned to the function X_0 by composition of small diffeomorphic deformations.

Let $\{s_i : \mathcal{M}_0 \to \mathcal{M}_0\}$ be the set deformation maps estimated from the previous iterations of the algorithm, such that $\hat{X}_i \circ s_i$ is a registered version of $\hat{X}_i$ to the function X_0. The functions $\{s_i\}$ can be the set of identity maps in the first iteration. Moreover, let $\{p_j : j = 1, \ldots, S\} \subset \mathcal{M}_0$ be a collection of S control points where the functions $\{\hat{X}_i\}$ are sub-sampled. In practice, these will be the nodes of the triangulation $\mathcal{M}_0^{\mathcal{T}}$, i.e. the points where the functions are actually observed.

With a slight abuse of notation, let the diffeomorphic function $\phi_u : \mathcal{M}_0 \to \mathcal{M}_0$ be the solution generated at time $t = 1$ by the ODE

$$\begin{cases} \frac{\partial \phi_u}{\partial t}(t, x) = u \circ \phi_u(t, x) & t \in [0, 1], x \in \mathcal{M}_0, \\ \phi_u(0, x) = x & x \in \mathcal{M}_0 \end{cases} \tag{8}$$

where u is a sufficiently smooth vector field on $\mathcal{M}_0$. If $\mathcal{M}_0$ has a boundary, than we assume u vanishes, with their derivatives, on the boundary. Such an ODE is used here as a tool to generate a diffeomorphic function ϕ_u from a vector field u that needs only to be smooth. Then, we propose to estimate a set of functional registration maps, each aligning $\hat{X}_i \circ s_i$ to X_0 by minimizing

$$E_{\mathcal{M}_0}(u_i) = \sum_{j=1}^{S} \left(X_0(p_j) - \hat{X}_i \circ s_i \circ \phi_{u_i}(p_j)\right)^2 + \lambda \|\Delta_{BL} u_i\|^2_{L^2(T\mathcal{M}_0)}, \tag{9}$$

where $\|\Delta_{BL} u_i\|^2_{L^2(T\mathcal{M}_0)}$ is the L^2 norm of the vector field $\Delta_{BL} u_i$, which imposes smoothness on u_i. The constant λ is a weighting coefficient between the data fidelity term, i.e. how well aligned we want $\hat{X}_i \circ s_i \circ \phi_{u_i}$ to be to X_0, and the smoothing term, i.e. how smooth we want the vector field u_i to be.

The term $\hat{X}_i \circ s_i \circ \phi_{u_i}$ in Equation (9), is then linearized with respect to u_i. This results in the approximation

$$\hat{X}_i \circ s_i \circ \phi_{u_i} \approx \hat{X}_i \circ s_i + L_{u_i},$$

where L_{u_i} is a first order approximation of $\hat{X}_i \circ s_i \circ \phi_{u_i} - \hat{X}_i \circ s_i$. By means of Vector Finite Elements, an approximate solution $\hat{u}_i$, at the nodes of $\mathcal{M}_0^{\mathcal{T}}$, can be characterized in terms of the solution of a linear system. An approximate vector field $\hat{u}_i$ on the triangulation $\mathcal{M}_0^{\mathcal{T}}$ is then computed by linear interpolation of the solution found at the nodes of the

Algorithm 1 Functional Registration Algorithm

1: Initialization:

(a) Initialize $\{s_i^0 \leftarrow Id : i = 1, \ldots, n\}$ to be the identity functions on $\mathcal{M}_0$

(b) Initialize $\{\hat{X}_i : i = 1, \ldots, n\}$ to be the functions estimated from Section 3.1

(c) Initialize the functional template to be $X_0 \leftarrow \frac{1}{n}\sum_i \hat{X}_i$

2: Compute $\{\hat{u}_i^k : i = 1, \ldots, n\}$, the solution at the kth iteration, from the second order functional

$$E_{\mathcal{M}_0}(u_i) = \sum_{j=1}^{S} \left(X_0(p_j) - \hat{X}_i \circ s_i^{k-1}(p_j) - L_{u_i}(p_j)\right)^2 + \lambda \|\Delta_{BL} u_i\|^2_{L^2(T\mathcal{M}_0)},$$

3: Compute the registration maps $\{\phi_{\hat{u}_i^k} : i = 1, \ldots, n\}$ by solving the ODE

$$\begin{cases} \frac{\partial \phi_{u_i^k}}{\partial t}(t, x) = u_i^k \circ \phi_{u_i^k}(t, x) & t \in [0, 1], x \in \mathcal{M}_0, \\ \phi_{u_i^k}(0, x) = x & x \in \mathcal{M}_0 \end{cases}$$

4: Update current registration maps and functional template:

$\{s_i^k \leftarrow s_i^{k-1} \circ \phi_{\hat{u}_i^k} : i = 1, \ldots, n\}$

$X_0 \leftarrow \frac{1}{n}\sum_i \hat{X}_i \circ s_i^k$

5: Output and analysis (e.g. fPCA) of the result of the current iteration:

$\{\hat{X}_i \circ s_i^k : i = 1, \ldots, n\}$

6: Repeat Steps 2–5 until until a stopping criterion is satisfied

triangulation. Details of this procedure can be found in the Supplementary Material. The main steps of the functional registration algorithm are summarized in Algorithm 1.

Each iteration of the functional registration algorithm result in a newly estimated set of functions $\{\hat{X}_i \circ s_i^k\}$, representing a re-aligned correction of the maps $\{\hat{X}_i\}$. The composition $\hat{X}_i \circ s_i^k$ means that for all $x \in \mathcal{M}_0$ the functional value $\hat{X}_i(x)$ is, after k iterations, relocated on the point $(s_i^k)^{-1}(x) \in \mathcal{M}_0$. Thus, the functional registration also has the effect of correcting the overall geometric deformations $\{\varphi_{\hat{v}_i} : \mathcal{M}_0 \to \mathcal{M}_i\}$,

estimated in Section 3.1, to be

$$\varphi_{\hat{v}_i} \circ (s_i^k)^{-1}, \qquad i = 1, \ldots, n. \tag{10}$$

The functional registration model introduced in this section, as opposed to the geometric registration model in Section 3.1, is based on the composition of small deformations, where at the kth iteration, $\{\hat{X}_i \circ s_i^k : i = 1, \ldots, n\}$ represent the re-aligned versions of $\{\hat{X}_i \circ s_i^{k-1} : i = 1, \ldots, n\}$. The constant λ in the model, controls the change between the functions $\{\hat{X}_i \circ s_i^{k-1}\}$ and those estimated at the next iteration, as in fact large values of λ privilege small deformations. This has the advantage that the fPCA analysis can be re-performed on the functions $\{\hat{X}_i \circ s_i^k\}$ at each iteration k. The output of this analysis can provide useful information for the next step of the functional registration algorithm, as for instance a stopping criterion in a similar fashion to Kneip and Ramsay (2008).

In summary, we have introduced a method that exploits the functional information to achieve a better registration by updating the functional estimates $\{\hat{X}_i\}$ to be $\{\hat{X}_i \circ s_i^{k_{\text{stop}}}\}$, and the diffeomorphic geometric deformations $\varphi_{\hat{v}_i}$ to be $\{\varphi_{\hat{v}_i} \circ (s_i^{k_{\text{stop}}})^{-1}\}$, where k_{stop} denotes the iteration where the functional algorithm is stopped. As pointed out at many stages in this work, it is however important to have a representation of the final update of the deformations maps $\{\varphi_{\hat{v}_i} \circ (s_i^{k_{\text{stop}}})^{-1}\}$ in terms of elements of the linear Hilbert space $\mathcal{V}$, so that we can perform linear statistics. To this purpose, we can estimate such elements by applying the geometric deformation model in Section 3.1, i.e. by solving

$$\hat{v}_i^{k_{\text{stop}}} = \underset{v_i \in \mathcal{V}}{\arg\min}\, D^2(\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}, \varphi_{\hat{v}_i} \circ (s_i^{k_{\text{stop}}})^{-1} \circ \mathcal{M}_0^{\mathcal{T}}) + \lambda \|v_i\|_{\mathcal{V}}^2,$$

where D^2 denotes the landmark distance defined in equation (6).

The overall procedure in this section results in a set of corrected estimates $\{\hat{X}_i \circ s_i^{k_{\text{stop}}}\}$ and $\{\hat{v}_i^{k_{\text{stop}}}\}$, that exploit functional information, estimating respectively $\{X_i\}$ and $\{v_i\}$ in Model 1. To ease the notation, in the next section, we drop the index on the number of iterations of the functional registration algorithm, denoting with $\{\hat{X}_i\}$ and $\{\hat{v}_i\}$ the corrected estimates of functions and geometric deformations respectively.

3.2.3 Remarks on computational times

It is also important to highlight that the idea of alternating between each iteration of the functional registration algorithm and the fPCA analysis on the functions is ultimately enabled by the computational efficiency of the proposed functional registration algorithm and fPCA algorithm on the functions. In the case of the application, in Section 4, each FoS is represented by a 32K nodes triangulated surface and the associated 32K functional values on the nodes. In this setting, the computational time of one iteration of the functional registration algorithm, applied between two functions, is in the order of 2 minutes on a Intel Core i5-3470 3.20GHz workstation, with 4 GB of RAM. The computational time for a singular PC of the functions is 15 seconds, on the same workstation, with the fPCA implementation proposed in Section 3.3. Instead, the landmark driven geometric registration of the 32K nodes template to a 32K nodes surface, representing a cerebral cortex, takes approximately 3 hours on a cluster's node equipped with a Dell T620 server and a NVIDIA K20 GPU.

3.3 Functional Principal Component Analysis

In Sections 3.1-3.2 we introduced the estimation procedure for the objects $\{\hat{v}_i : i = 1, \ldots, n\}$ representing the geometries and $\{\hat{X}_i : i = 1, \ldots, n\}$ representing the functions, from a set of n FoSs. In this section, the aim is to outline the estimation procedure to the empirical PC component functions from the observed objects $\{\hat{v}_i\}$ and $\{\hat{X}_i\}$, in analogy to what proposed in Section 2.3, in terms of PCs of the underlying random functions V and X.

3.3.1 Geometric variability

The empirical PC functions are in practice computed from the eigen-decomposition of the empirical covariance operator $\hat{C}_{\mathcal{V}}$, defined as

$$\hat{C}_{\mathcal{V}}(v) = \frac{1}{n}\sum_{i=1}^{n}\langle v, \hat{v}_i - \bar{v}\rangle_{\mathcal{V}}(\hat{v}_i - \bar{v}), \qquad v \in \mathcal{V}, \tag{11}$$

where $\bar{v} = \frac{1}{n}\sum_{i=1}^{n}\hat{v}_i$ and $\langle\cdot,\cdot\rangle_{\mathcal{V}}$ denotes the scalar product in $\mathcal{V}$. An explicit solution of this eigenvalue problem can be derived by expanding v and $\hat{v}_i$ in (11) over a basis of $\mathcal{V}$ or

discretizing the problem over a fine grid of $\mathbb{R}^3$. Since the number of observations in this setting is small with respect to the size of the space, an appropriate choice of the basis is given by the collection of the actually observed vector fields $\hat{v}_i$ (Ramsay and Silverman, 2005). Thus, the eigenvalue problem $\hat{C}_{\mathcal{V}}(\psi_j^G) = \kappa_j^G \psi_j^G$ can be re-formulated to a discrete eigenvector problem in terms of the basis expansion coefficients, leading to the empirical PC functions estimates $\{\hat{\psi}_j^G\}$ and empirical variance estimates $\{\hat{\kappa}_j^G\}$. The empirical PC scores vectors can be estimated by projecting $\{\hat{v}_i\}$ on the estimated PC functions, i.e. the ith element of the jth scores vector is given by

$$\hat{A}_{i,j}^G = \langle \hat{v}_i - \bar{v}, \hat{\psi}_j^G \rangle_{\mathcal{V}} \qquad i = 1, \ldots, n,\ j = 1, \ldots, K^G.$$

The empirical jth mode of geometric variation is thus represented by the PC function $\hat{\psi}_j^G$, which is associated to the deformations $\varphi_{\pm\sqrt{\hat{\kappa}_j^G}\hat{\psi}_j^G}$ of $\mathbb{R}^3$ that applied to the geometric template correspond to the change of geometry described by $\varphi_{\pm\sqrt{\hat{\kappa}_j^G}\hat{\psi}_j^G} \circ \mathcal{M}_0$. The observed vector fields can be finally expressed in terms of the basis expansion, also known as the Karhunen-Loève expansion:

$$\hat{v}_i \approx \bar{v} + \sum_{j=1}^{K^G} \hat{A}_{i,j}^G \hat{\psi}_j^G. \tag{12}$$

Equation (12) emphasizes the fact that the matrix $(\hat{A}_{i,j}^G)_{ij}$ is such that the ith row is a compact description of the vector field $\hat{v}_i$.

3.3.2 Functional variability

From similar arguments, we can build an estimator for the PC functions and PC scores vectors for the functions $\{\hat{X}_i\}$. The estimated functions $\{\hat{X}_i\}$ are noisy estimates of the realization of the underlying unobserved random function X. A pre-smoothing of the noisy functions could be considered, however here we rely on the fPCA algorithm proposed in Lila et al. (2016), where the regularization term is applied directly to the PC functions to be estimated.

In fact, the PC functions $\{\psi_j^F\}$ of the centered random function $X - \mu$, satisfy the following property

$$\{\psi_m^F\}_{m=1}^M = \underset{(\{\psi_m\}_{m=1}^M : \langle \psi_m, \psi_l \rangle_{L^2(\mathcal{M}_0)} = \delta_{ml})}{\arg\min} \mathbb{E} \int_{\mathcal{M}_0} \left\{ X - \mu - \sum_{m=1}^{M} \langle X - \mu, \psi_m \rangle_{L^2(\mathcal{M}_0)} \psi_m \right\}^2, \tag{13}$$

where $\int_{\mathcal{M}_0}$ denotes the surface integral over $\mathcal{M}_0$ and $\langle \cdot, \cdot \rangle_{L^2(\mathcal{M}_0)}$ denotes the scalar product in $L^2(\mathcal{M}_0)$. In (13) we can see that the PC functions minimize the loss of information caused by the truncation of the series expansion to the first M components. Let $\{p_j : j = 1, \ldots, S\} \subset \mathcal{M}_0$ be a collection of S points where the estimated functions $\{\hat{X}_i\}$ are sub-sampled. In practice, these will be the nodes of the triangulation $\mathcal{M}_0^{\mathcal{T}}$, i.e. the points where the functions are actually observed. Let Δ be the Laplace-Beltrami operator (see e.g. Chavel, 2006). The Laplace-Beltrami operator of a smooth function $f \in L^2(\mathcal{M}_0)$ is a function in $L^2(\mathcal{M}_0)$ that gives a measure of the local curvature of the function f.

The first PC function $\hat{\psi}_1^F \in L^2(\mathcal{M}_0)$ and associated first scores vector $(\hat{A}_{1,1}^F, \ldots, \hat{A}_{n,1}^F)$ are estimated by minimizing the following regularized empirical version of (13):

$$(\hat{\psi}_1^F, \{\hat{A}_{i,1}^F\}_{i=1}^n) = \underset{\psi_1, \{A_{i,1}\}_{i=1}^n}{\arg\min} \sum_{i=1}^{n} \sum_{j=1}^{S} \left(\hat{X}_i(p_j) - \bar{X}(p_j) - A_{i,1}\psi_1(p_j)\right)^2 + \lambda \|\Delta\psi_1\|_{L^2(\mathcal{M}_0)}, \quad (14)$$

where $\bar{X}$ denotes the sample mean function of $\{\hat{X}_i\}$ and λ is a weighting coefficient between the empirical and regularizing term. The regularization term imposes smoothness on the estimated PC function $\hat{\psi}_1^F$, coherently with the structure of the manifold $\mathcal{M}_0$. Subsequent PCs can be estimated by reapplying (14) to the residuals. Details of the implementation and an application to functional Magnetic Resonance Imaging can be found in Lila et al. (2016).

The observed functions $\{\hat{X}_i\}$ can be finally expressed in terms of the basis expansion

$$\hat{X}_i \approx \bar{X} + \sum_{j=1}^{K^F} \hat{A}_{i,j}^F \hat{\psi}_j^F. \quad (15)$$

The matrix $(\hat{A}_{i,j}^F)_{ij}$ is such that the ith row is a compact description of the function $\hat{X}_i$.

3.4 Geometric and Functional variability relation

The matrices $(\hat{A}_{i,j}^G)_{ij}$ and $(\hat{A}_{i,j}^F)_{ij}$, computed in Section 3.3, are such that their ith row represents a compact description of the geometry and functions of the ith FoS $(\mathcal{M}_i, Y_i)$. Each row of these matrices could also be regarded as the estimated empirical ith realization of the random vector $(a_1^G, \ldots, a_{K^G}^G)$ and $(a_1^F, \ldots, a_{K^F}^F)$ defined in Section 2.3. As outlined in that section, the matrices $(\hat{A}_{i,j}^G)_{ij}$ and $(\hat{A}_{i,j}^F)_{ij}$ can then be used to study the relation between geometric variability and functional variability of the given collection of FoSs.

To this end we can perform, for instance, a linear regression analysis where we try to explain the jth mode of functional variability as a linear combination of the K^G modes of geometric variation.

Alternatively, we could perform CCA, and look for the lth mode of co-variation $(\hat{\mathbf{w}}^{G,l}, \hat{\mathbf{w}}^{F,l})$, representing the lth maximally correlated linear combination $\hat{\mathbf{w}}^{G,l} \in \mathbb{R}^{K^G}$, of the K^G modes of geometric variation with the linear combination $\hat{\mathbf{w}}^{F,l} \in \mathbb{R}^{K^F}$ of the K^F modes of functional variation. The lth mode of co-variation $(\hat{\mathbf{w}}^{G,l}, \hat{\mathbf{w}}^{F,l})$ can be visualized as the sequence of FoSs

$$\begin{cases} \mathcal{M}_{\text{CCA},l} = \varphi_{c\,\hat{\psi}^G_{\text{CCA},l}} \circ \mathcal{M}_0, \\ Y_{\text{CCA},l} = c\,\hat{\psi}^F_{\text{CCA},l} \circ \varphi^{-1}_{c\,\hat{\psi}^G_{\text{CCA},l}}, \end{cases} \tag{16}$$

obtained by varying $c \in \mathbb{R}$ in an interval containing 0, with $\hat{\psi}^G_{\text{CCA},l} = \sum_{j=1}^{K^G} \hat{w}_j^{G,l} \hat{\psi}_j^G$ and $\hat{\psi}^F_{\text{CCA},l} = \sum_{j=1}^{K^F} \hat{w}_j^{F,l} \hat{\psi}_j^F$, where $\{\hat{\psi}_j^G\}$ and $\{\hat{\psi}_j^F\}$ are the estimated geometric and functional PC component functions, while $\hat{w}_j^{G,l}$ and $\hat{w}_j^{F,l}$ denote the jth element of $\hat{\mathbf{w}}^{G,l}$ and $\hat{\mathbf{w}}^{F,l}$ respectively. An example of such a visualization is shown in Figures 10-11, for the real application.

3.5 Choice of the hyper-parameters

In the proposed models, various hyper-parameters have to be chosen. In particular, in the geometric registration step in Section 3.1, we have to choose the regularization weighting parameter λ. The regularization weighting parameter, in our analysis, does not play a large role. In fact, if the surfaces were noisy reconstructions, its choice would have been more delicate. However, in practice, the surfaces are extracted from a regularized segmentation process of 3D images, and thus are smooth. For this reason, the regularization weighting parameter λ, in the geometric registration, is chosen to be small.

As previously mentioned, $\mathcal{V}$ is in practice a RKHS. Important to the registration problem is the choice of $\sigma_{\mathcal{V}}$, the size of the kernel of the RKHS $\mathcal{V}$ (see Supplementary Material). In fact a RKHS with a large kernel size $\sigma_{\mathcal{V}}$ is able to better capture large deformations (e.g. size differences), while under-fitting local differences. A RKHS with a small kernel size has an opposite behaviour. Following the approach of Bruveris et al.

(2012), we take a sum of two Gaussian kernels, which allows the space $\mathcal{V}$ to account for both large and small deformations.

The functional registration has also a regularization weighting parameter λ, which determines how slowly the algorithm approaches an optimal solution. As in Kneip and Ramsay (2008), after some experimentation, we choose the value λ that achieves a smooth variation on functional PC functions, obtained from the functional variability analysis, between each iteration. To determine the number of iterations needed, we examine the eigenvalue plots (scree plots) to determine when stability of these plots has been reached in a analogous manner to Kneip and Ramsay (2008). Finally, the regularization weighting parameter of the fPCA algorithm applied to the functions, has been chosen by K-fold cross-validation, with $K = 5$, details of which can be found in Lila et al. (2016).

On a more general note, choosing the hyper-parameters of the registration algorithms, in a data-driven fashion, is admittedly a very difficult problem and it has been very little explored in the current literature, even in simpler situations such as for functions on the real line. The above only represents one possible method of choosing them, which appears to work well in our application, although further work would be needed for very different settings.

4 Application

The publicly available data set considered in this work has been collected by the Human Connectome Project Consortium (HCP, Essen et al., 2012), with the ultimate goal of elucidating the understanding of the brain functions, by collecting multi-modal neuroimaging data such as structural scans, resting-state and task-based functional MRI scans, and diffusion-weighted MRI scans from a large number of healthy volunteers. A minimal preprocessing pipeline have been applied to the dataset (Glasser et al., 2013).

4.1 Preprocessing

A 3D structural MRI scan has been performed for each individual, returning a 3D image describing the internal structure of the brain. A slice of the 3D image is shown on the

left panel of Figure 2. The cerebral cortex is the outermost layer of the brain, mostly consisting of neuronal cell bodies. With automatic segmentation techniques, it is possible to separate the cerebral cortex from the other parts of the brain. Subsequently the two surfaces enclosing the cerebral cortex can be computed. The inner surface represents the boundary between the cerebral cortex and the white matter (second panel in Figure 2), while the outer surface corresponds to the boundary between the cerebral cortex and the cerebrospinal fluid (fourth panel in Figure 2). The resolution of the MRI images (0.7 mm isotropic, in this study) and the effectiveness of the segmentation algorithm determine the level of details at which such surfaces can be reconstructed.

The geometry of the cerebral cortex is generally represented by the mid-thickness surface, which is the surface fitting the middle-points of the inner and outer surfaces, an example of which is shown on the third panel of Figure 2. Thus, it is natural to expect the resulting surfaces to have wider sulci and thinner gyri than what we could observe from a picture of the brain surface. Moreover, the mid-thickness surface can be equipped with a function representing the thickness of the cerebral cortex, computed from the inner and outer surface, as described in Fischl and Dale (2000). A comparison of the various methods for the cerebral cortex thickness estimation can be found in Lerch and Evans (2005). In Figure 1, we show the reconstructed (mid-thickness) surfaces of the left hemisphere of 3 different subjects with the associated cerebral cortex thickness maps. Each surface is represented by a 32K nodes mesh, and at each node of the mesh an evaluation of the function is available.

The mid-thickness surfaces of the collected cohort are pre-registered to the Conte69 template, on the left in Figure 7, through a surface-based registration algorithm driven by geometric features that describe measures of cortical shape folding, such as sulcal depth or local curvature (Fischl et al., 1999; Glasser et al., 2013). Registrations are ensured to be one-to-one by introducing, in the objective function, a term related to the metric distortion of the registration maps and a term that enforces the positivity of the signed areas of the triangles on the surfaces (see Fischl et al., 1999, for details). Such a procedure defines a one-to-one correspondence between the 32K nodes of the template and the 32K nodes of each of the mid-thickness surfaces, which can be regarded as a set of 32K landmarks.

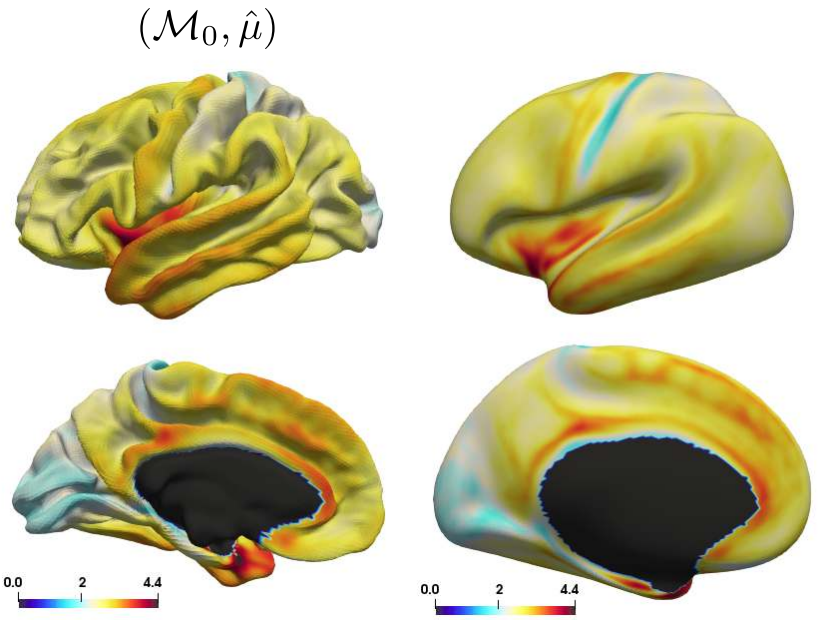

Figure 7: On the left, the Conte69 template, used as a template surface for the registration of the individual surfaces. This is equipped with the cross-sectional mean function computed post-geometric registration. On the right, the cross-sectional mean function visualized on an inflated version of the template.

4.2 Analysis

The relation between geometric features of the brain has raised great interest in the recent years, since it can potentially help us understand the principles underlying brain development. Classically, these studies have been confined to correlation studies on variables summarizing particular geometric features. For instance, in Im et al. (2008), for each subject, the average cortical volume and absolute mean curvature, among other, are computed. This set of real variables are then compared to the average cerebral cortex thickness computed on each subject. Moreover, a more localized analysis is performed by parcellating each cortical surface in the 4 lobes. Subsequently, the analysis is performed independently on each of lobe. However, there are two limitations of such approach. Firstly, the description of the geometric properties through summary statistics is in general incomplete. Secondly, the parcellation of the cortical surfaces determines a priori which areas of the cortical surface can have a different behaviour.

The fact that a geometric registration has already been performed on the HCP data, without relying on the diffeomorphic registration framework in Section 3.1, is not in

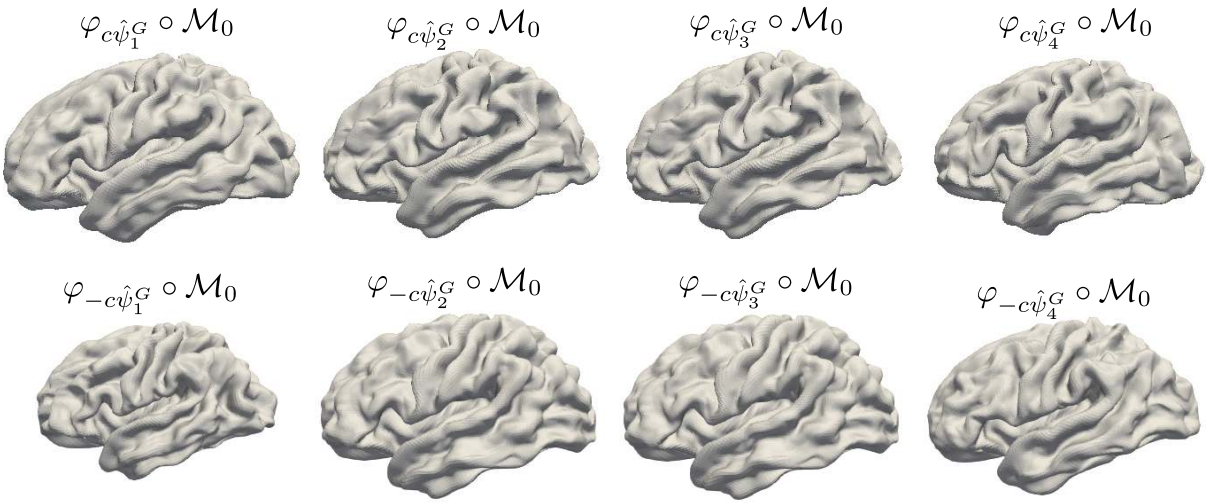

Figure 8: From left to right, the first four geometric PC functions, computed on the space of initial vector fields. These are visualized as $\varphi_{\pm c\hat{\psi}_j^G} \circ \mathcal{M}_0$, where $\hat{\psi}_j^G$ is the jth geometric PC function.

contrast with the proposed analysis. In fact, diffeomorphic-like constraints can be imposed in many different ways when it comes to the estimation of registration maps. However, if the aim is the estimation of a low-dimensional subspace of the diffeomorphic space, these alternative approaches cannot be extended to this more general problem. For this reason, we use the landmarks defined by the pre-processing geometric registration to estimate the vector fields that represent such registrations and then perform fPCA on the estimated vector fields, as described in Section 3.3. The estimated first four geometric PCs are shown in Figure 8. Not surprisingly, they are mostly related to the size of the brain or the size of sub-parts of the brain.

We then perform fPCA on the functions registered on the Conte69 template. The results are shown in the top two rows of Figure 9. Subsequently, we perform functional registration of the functions on the Conte69 template and recompute the functional modes of variation at each iteration. In the bottom two rows of Figure 9 we show the PC functions after 2 iterations of the functional registration algorithm.

We finally perform a CCA on the first eight geometric and functional PC functions scores. The resulting first two main modes of co-variation, the only significant ones from a likelihood ratio test (details of which can be found in the Supplementary Material), are shown in Figures 10-11. From the left to the right panel of Figure 10, we can see

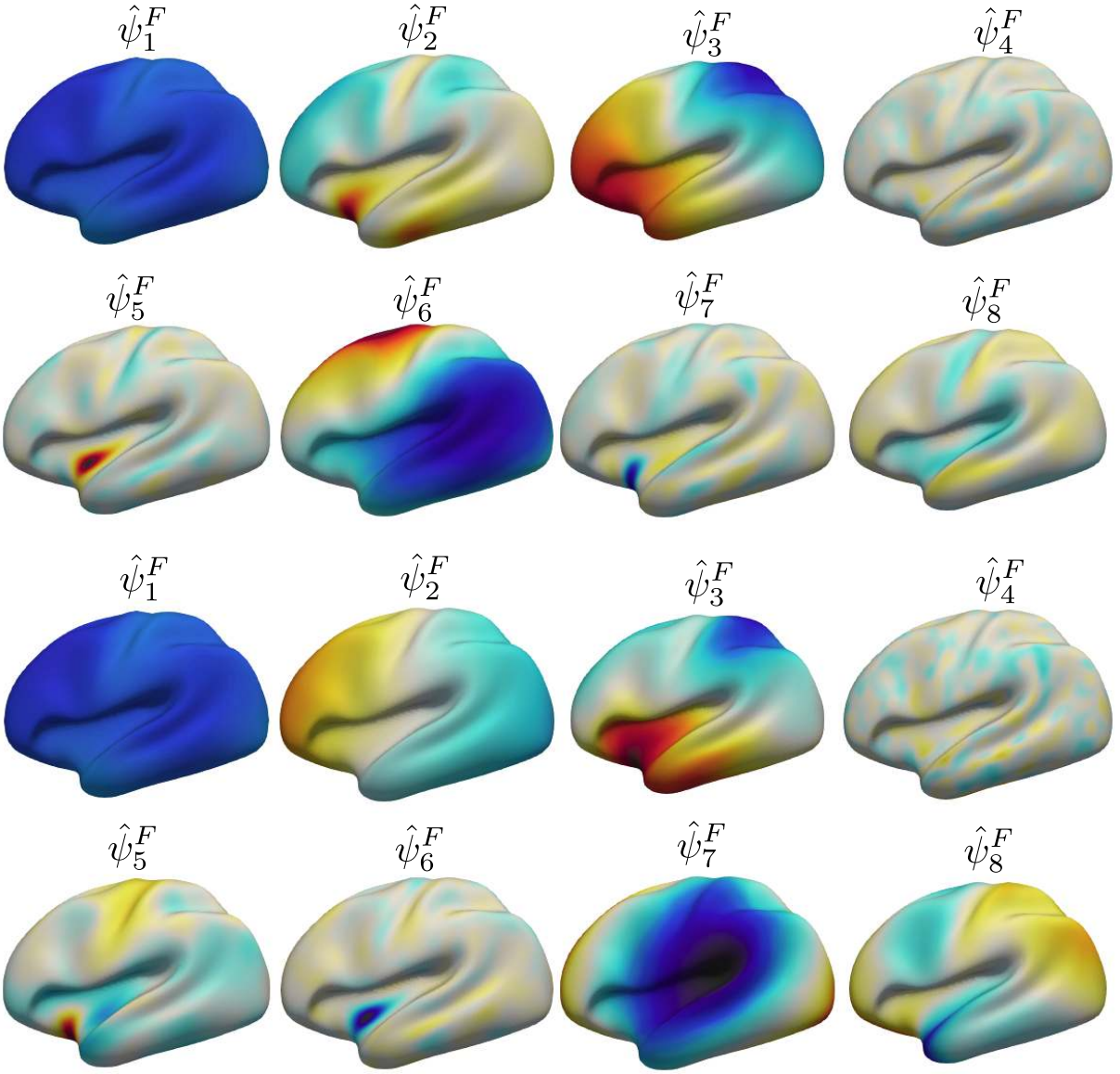

Figure 9: Results of the fPCA on the functions. On the top two rows, the first eight functional modes of variation computed without performing functional registration. On the bottom two rows, the first eight functional modes of variation computed after performing functional registration. These are shown on an inflated version of the template to easy their visualization.

the presence of a correlation between a decrease in thickness in the frontal lobe and an increase in size of the entire brain, while in the temporal lobe, an increase in thickness seems associated to an increase in size of the entire brain. Moreover, in the second main mode of co-variation a more localized phenomenon is captured in proximity of the high average cortical thickness area on the lateral sulcus (see Figure 7), where an association

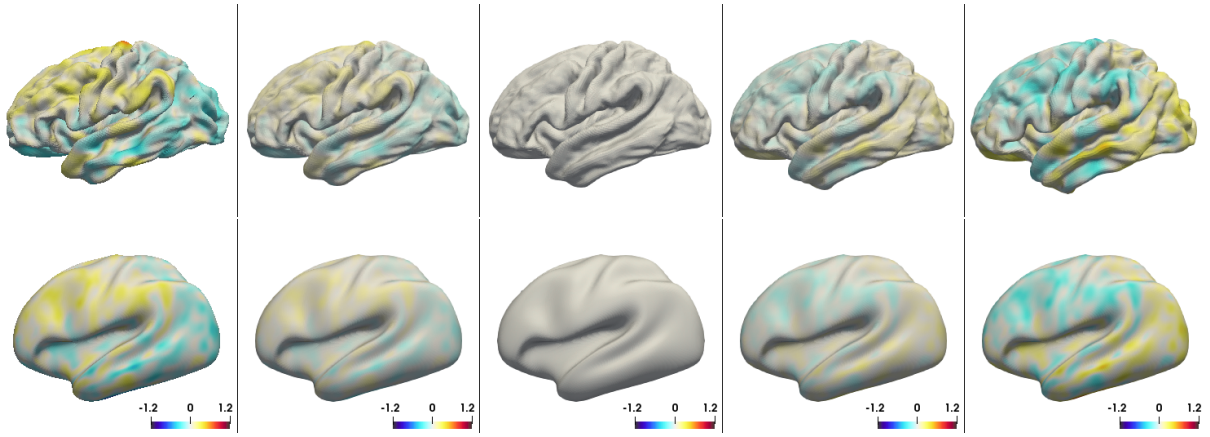

Figure 10: A representation of the first main mode of co-variation of the geometric and functional components of the CCA analysis, representing the most correlated linear combinations of the first eight geometric and functional PC functions. From left to right, this is visualized by plotting the FoS in (16) for a sequence of constants c. To ease visualization, a video showing the first main mode of co-variation can be found in the Supplementary Material.

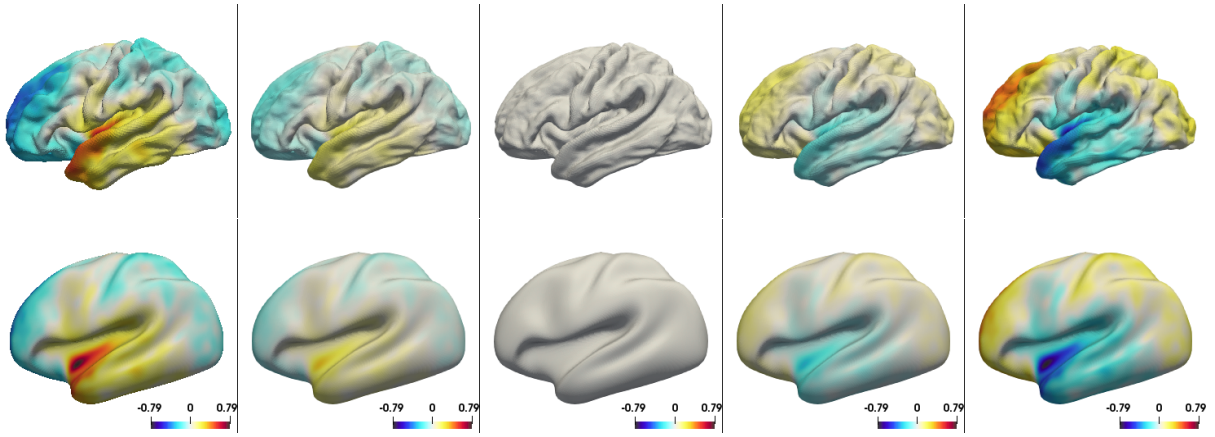

Figure 11: A representation of the second main mode of co-variation of the geometric and functional components of the CCA analysis, representing the second most correlated linear combinations of the first eight geometric and functional PC functions. From left to right, this is visualized by plotting the FoS in (16) for a sequence of constants c. To ease visualization, a video showing the second main mode of co-variation can be found in the Supplementary Material.

between an increase in the cortical thickness and an increase in the size of the brain is suggested. Note that such local effect would have not been captured by a study confined

to study individually each lobe of the brain, and such relation would have probably been ascribed to the entire lobe containing that area.

5 Conclusions and Prospectives

In this paper, motivated by the analysis of neuroimaging data, we introduce a framework for the analysis of FoSs. In particular, a statistical model describing the phenomenon is formulated, and the estimators of the unknown quantities of the model are introduced. The construction of such estimators is complicated by the necessity of the resulting estimates to lie in the non-linear subspace of 'sensible' solutions, here taken to be deformations. Moreover, in such high dimensional setting, it is fundamental for the estimator to incorporate prior information on the geometry and the smoothness of the data, achieved by regularizing the estimates through differential operators. Motivated by simulation studies, we address the necessity of using the functional information to achieve a better registration, a well known fact in FDA, by introducing a novel diffeomorphic registration algorithm for functional data on a two-dimensional manifold.

While the main motivation of this paper was taken from a neuroimaging application into assessing the inherent variabilities of cortical thickness, the methodology has wider applications in medical imaging as a whole, where FoS appear in cardiovascular (e.g. Huang et al. (2016)), muskuloskeletal (e.g. Treece and Gee (2015)) and many other imaging areas. More generally, this methodology is an example of the use of differential operators as regularisers in statistics, a field where not only statistical but also numerical techniques are needed to facilitate solutions.

A future interesting aspect is the exploration of the applicability of the Optimal Transport framework to the registration problem, as suggested in Panaretos and Zemel (2016) in a discrete context, and its links with the diffeomorphic deformation framework. This is of potential interest in the surface registration framework, where we usually lack physical models that can describe the phenomena, and thus a 'least action' approach could well be effective.

Supplementary Material and Acknowledgements

In the Supplementary Material, we provide further details on the geometric and functional registration models. Moreover, we show the results of simulation studies on synthetic data, to investigate our estimation procedure. A link to the full code and data can be found in the `code_link.pdf` file, while a version of the code complete with toy examples is available in the file `FoSs_Code_ToyExamples.tar.gz`. Videos of the main modes of variation are also available as supplementary material.

The authors greatly appreciate the really useful comments of the AE and two referees, which helped considerably strengthen the paper. JA was supported by the Engineering and Physical Sciences Research Council (EP/K021672/2 and EP/N014588/1). EL was supported by the EPSRC grant EP/L016516/1.

References

M. A. Audette, F. P. Ferrie, and T. M. Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4(3):201 – 217, 2000.

F. L. Bookstein. Shape and the information in medical images: A decade of the morphometric synthesis. *Computer Vision and Image Understanding*, 66(2):97 – 118, 1997a.

F. L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge, 1997b.

M. Bruveris, L. Risser, and F.-X. Vialard. Mixture of kernels and iterated semidirect product of diffeomorphisms groups. *Multiscale Modeling & Simulation*, 10(4):1344–1368, 2012.

B. Charlier, N. Charon, and A. Trouvé. The Fshape Framework for the Variability Analysis of Functional Shapes. *Foundations of Computational Mathematics*, 17(2):287–357, apr 2017.

N. Charon and A. Trouvé. Functional Currents: A New Mathematical Tool to Model

and Analyse Functional Shapes. *Journal of Mathematical Imaging and Vision*, 48(3): 413–431, mar 2014.

I. Chavel. *Riemannian Geometry: A Modern Introduction*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2006.

P. Dupuis, U. Grenander, and M. I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, 56(3):1–20, 1998.

D. V. Essen, K. Ugurbil, E. Auerbach, D. Barch, T. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. Curtiss, S. D. Penna, D. Feinberg, M. Glasser, N. Harel, A. Heath, L. Larson-Prior, D. Marcus, G. Michalareas, S. Moeller, R. Oostenveld, S. Petersen, F. Prior, B. Schlaggar, S. Smith, A. Snyder, J. Xu, and E. Yacoub. The human connectome project: A data acquisition perspective. *NeuroImage*, 62(4):2222 – 2231, 2012.

B. Fischl and A. M. Dale. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, 97(20): 11050–11055, 2000.

B. Fischl, M. I. Sereno, and A. M. Dale. Cortical surface-based analysis ii: inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207, feb 1999.

M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. V. Essen, and M. Jenkinson. The minimal preprocessing pipelines for the human connectome project. *NeuroImage*, 80 (0):105 – 124, 2013.

P. Z. Hadjipantelis, J. A. D. Aston, H. G. Müller, and J. P. Evans. Unifying Amplitude and Phase Analysis: A Compositional Data Approach to Functional Multivariate Mixed-Effects Modeling of Mandarin Chinese. *Journal of the American Statistical Association*, 110(510):545–559, apr 2015.

D. J. Hagler, A. P. Saygin, and M. I. Sereno. Smoothing and cluster thresholding for

cortical surface-based group analysis of fmri data. *NeuroImage*, 33(4):1093 – 1103, 2006.

Y. Huang, Z. Teng, M. Elkhawad, J. M. Tarkin, N. Joshi, J. R. Boyle, J. R. Buscombe, T. D. Fryer, Y. Zhang, A. Y. Park, I. B. Wilkinson, D. E. Newby, J. H. Gillard, and J. H. F. Rudd. "high structural stress and presence of intraluminal thrombus predict abdominal aortic aneurysm 18f-fdg uptake: Insights from biomechanics". *"Circulation. Cardiovascular imaging"*, 9, 2016.

K. Im, J. M. Lee, O. Lyttelton, S. H. Kim, A. C. Evans, and S. I. Kim. Brain size and cortical structure in the adult human brain. *Cerebral Cortex*, 18(9):2181–2191, 2008.

S. Joshi and M. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, 2000.

A. Kneip and J. O. Ramsay. Combining Registration and Fitting for Functional Models. *Journal of the American Statistical Association*, 103(483):1155–1165, sep 2008.

J. P. Lerch and A. C. Evans. Cortical thickness analysis examined through power analysis and a population simulation. *NeuroImage*, 24(1):163–173, jan 2005.

E. Lila and J. A. D. Aston. Statistical Analysis of Functions on Surfaces, with an application to Medical Imaging. *arXiv:1707.00453*, 2017.

E. Lila, J. A. D. Aston, and L. M. Sangalli. Smooth principal component analysis over two-dimensional manifolds with an application to neuroimaging. *Ann. Appl. Stat.*, 10 (4):1854–1879, 2016.

E. A. Maguire, K. Woollett, and H. J. Spiers. London taxi drivers and bus drivers: A structural mri and neuropsychological analysis. *Hippocampus*, 16(12):1091–1101, 2006.

J. S. Marron, J. O. Ramsay, L. M. Sangalli, and A. Srivastava. Functional data analysis of amplitude and phase variation. *Statist. Sci.*, 30(4):468–484, 11 2015.

M. I. Miller, A. Trouvé, and L. Younes. Geodesic Shooting for Computational Anatomy. *Journal of Mathematical Imaging and Vision*, 24(2):209–228, mar 2006.

V. M. Panaretos and Y. Zemel. Amplitude and phase variation of point processes. *Ann. Statist.*, 44(2):771–812, 2016.

V. Patrangenaru and L. Ellingson. *Nonparametric Statistics on Manifolds and Their Applications to Object Data Analysis*. CRC Press, Inc., Boca Raton, FL, USA, 1st edition, 2015. ISBN 1439820503, 9781439820506.

J. O. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer, 2nd edition, 2005. ISBN 038740080X.

R. Tang and H.-G. Muller. Pairwise curve synchronization for functional data. *Biometrika*, 95(4):875–889, nov 2008.

J.-P. Thirion. Fast Non-Rigid Matching of 3D Medical Images. Research Report RR-2547, INRIA, 1995.

J.-P. Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3):243 – 260, 1998.

G. Treece and A. Gee. Independent measurement of femoral cortical thickness and cortical bone density using clinical CT. *Medical Image Analysis*, 20(1):249 – 264, 2015.

A. Trouvé and L. Younes. Metamorphoses Through Lie Group Action. *Foundations of Computational Mathematics*, 5(2):173–198, apr 2005.

D. Tward, M. Miller, A. Trouvé, and L. Younes. Parametric surface diffeomorphometry for low dimensional embeddings of dense segmentations and imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1195–1208, June 2017.

M. Vaillant and J. Glaunès. Surface matching via currents. *Information processing in medical imaging : proceedings of the ... conference*, 19:381–392, 2005.

M. Vaillant, M. I. Miller, L. Younes, and A. Trouvé. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23 Suppl 1:S161–9, 2004.

M. Vaillant, A. Qiu, J. Glaunès, and M. I. Miller. Diffeomorphic metric surface mapping in subregion of the superior temporal gyrus. *NeuroImage*, 34(3):1149–1159, 2007.

T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Diffeomorphic demons: Efficient non-parametric image registration. *NeuroImage*, 45(1, Supplement 1):S61 – S72, 2009. Mathematics in Brain Imaging.

K. D. Yao, V. Patrangenaru, and D. Lester. 3D mean Projective Shape Difference for Face Differentiation from Multiple Digital Camera Images. *ArXiv e-prints*, Apr. 2017.

B. Yeo, M. Sabuncu, T. Vercauteren, N. Ayache, B. Fischl, and P. Golland. Spherical Demons: Fast Diffeomorphic Landmark-Free Surface Registration. *IEEE Transactions on Medical Imaging*, 29(3):650–668, mar 2010.

L. Younes. *Shapes and Diffeomorphisms*, volume 171. Springer, first edition, May 2010. ISBN 3642120547.

J. Zaetz and S. Kurtek. A novel riemannian framework for shape analysis of annotated surfaces. In *Proceedings of International Workshop on Differential Geometry in Computer Vision for Analysis of Shapes, Images and Trajectories*, pages 3.1–3.11. BMVA Press, September 2015.

B. Zitová and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.