Supplementary Material for: '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

April 8, 2019

S1 Geometric Diffeomorphic Registration

Here we present the algorithmic details of the estimation framework introduced in Section 3.

The space of smooth vector fields $\mathcal{V}$, in the geometric registration model (5), is usually constructed as a RKHS (Miller et al., 2015). In detail, let $K_{\mathcal{V}} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ be a bounded symmetric positive definite function. $K_{\mathcal{V}}$ is usually referred to as the kernel of $\mathcal{V}$ and a typical choice for it is the Gaussian isotropic kernel, i.e. $K_{\mathcal{V}}(x,y) = \exp(-\|x-y\|_2^2/(2\sigma_{\mathcal{V}}^2))\mathrm{Id}_{3\times 3}$, with $\mathrm{Id}_{3\times 3}$ denoting a 3×3 identity matrix and $\sigma_{\mathcal{V}}$ reflecting the rigidity of the space. Define the pre-Hilbert space $\mathcal{V}_0 = \mathrm{span}\{K_{\mathcal{V}}(\cdot,x)\omega | x \in \mathbb{R}^3, \omega \in \mathbb{R}^3\}$. Given $f, g \in \mathcal{V}_0$ we can write them as $f = \sum_{i=1}^N K_{\mathcal{V}}(\cdot, x_i)\omega_i$ and $g = \sum_{i=1}^N K_{\mathcal{V}}(\cdot, y_i)z_i$. We thus define the inner product between f and g to be $\langle f, g\rangle_{\mathcal{V}} = \sum_{i,j=1}^N \omega_i^T K_{\mathcal{V}}(x_i, y_j)z_j$. The space $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}})$, defined as the closure of $\mathcal{V}_0$, is a (Reproducing Kernel) Hilbert space of smooth vector fields.

For modeling purposes, the time-variant vector-field v_t, introduced in Section 2.4, is assumed to be of the form (see e.g. Vaillant et al., 2004)

$$v_t(\cdot) = \sum_{k=1}^{k_g} K_{\mathcal{V}}(\phi_v(t, c_k), \cdot)\alpha_k(t), \tag{17}$$

for a set of control points $\{c_k : k = 1, \ldots, k_g\} \subset \mathbb{R}^3$ and the auxiliary variables $\{\alpha_k(t) : \mathbb{R} \to \mathbb{R}^3\}$ called momenta of the deformation. The control points $\{c_k\}$ are commonly chosen to be the nodes of the triangulated representation of the surface to be deformed. ϕ_v denotes the solution of the ODE (3) given the time-variant vector field $\{v_t : t \in [0,1]\}$. The associated deformation energy is defined to be

$$\int_0^1 \|v_t\|_{\mathcal{V}}^2 = \int_0^1 \sum_{k,l=1}^{k_g} \alpha_k(t)^T K_{\mathcal{V}}(\phi_v(t, c_k), \phi_v(t, c_l))\alpha_l(t). \tag{18}$$

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

Denoting with ∇_1 the gradient with respect to the first variable, the vector field v_t generating geodesics, with respect to the energy term $\int_0^1 \|v_t\|_{\mathcal{V}}^2$, can be characterized as the solution of the coupled ODE system, known as the EPDiff equation (Miller et al., 2015)

$$\begin{cases} \frac{\partial c_k(t)}{dt} & = \sum_{l=1}^{k_g} K_{\mathcal{V}}(c_k(t), c_l(t))\alpha_l(t) \\ \frac{\partial \alpha_k(t)}{dt} & = -\frac{1}{2}\left(\sum_{l=1}^{k_g} \nabla_1 K_{\mathcal{V}}(c_k(t), c_l(t))\alpha_l(t)\right)^T \alpha_k(t), \end{cases} \tag{19}$$

for a set of initial conditions $\{\alpha_k = \alpha_k(0)\} \subset \mathbb{R}^3$, parameterizing the initial vector field v_0. This means that the energy minimizing vector fields, generating diffeomorphisms, can be determined by (19) and fully controlled by the initial vector field

$$v_0(\cdot) = \sum_{k=1}^{k_g} K_{\mathcal{V}}(\cdot, c_k)\alpha_k,$$

parametrized in terms of the initial momentum vector $\{\alpha_k : k = 1, \ldots, k_g\}$. Moreover, along a geodesic path the instantaneous deformation energies $\|v_t\|_{\mathcal{V}}$ are constant, meaning that the total deformation energy $\int_0^1 \|v_t\|_{\mathcal{V}}^2 dt$ can be equivalently represented by the initial deformation energy $\|v_0\|_{\mathcal{V}}^2 = \sum_{k,l} \alpha_k K_{\mathcal{V}}(c_k, c_l)\alpha_l$.

Thanks to the finite dimensional representation underlying the element of the RKHS $\mathcal{V}$, the minimization of (5) can be cast in a finite dimensional setting and can be approached, for instance, with a gradient descent algorithm on the initial momentum vector parametrizing the initial velocity field (see, among others, Vaillant et al., 2004).

The MATLAB toolkit fshapesTk (https://github.com/fshapes/fshapesTk) offers an implementation of the described geometric registration algorithm, and its extension to the fshape framework (Charlier et al., 2017).

S2 Registration of Functional Data on a two-dimensional manifold

Here we cover further details of the functional registration algorithm, for functional data whose domain is a two-dimensional manifold, introduced in Section 3.2. The main idea of the proposed algorithm is to perform functional registration as compositions of small diffeomorphisms, each parameterized by a stationary velocity field. This class of algorithms are also known as Diffeomorphic Demons algorithms (Vercauteren et al., 2009a,b). Diffeomorphic Demons were originally introduced for functions on Euclidean domains and an extension to spherical domains has been proposed in Yeo et al. (2010). However, this extension exploits spherical vector spline interpolation theory and cannot be extended to a generic manifold. In the geometric registration problem, as detailed in Section S1, smoothness is imposed by controlling the norm $\|\cdot\|_{\mathcal{V}}$ of the functional space. In fact, in $\mathbb{R}^3$, it is easy to define symmetric definite positive kernels from which we can straightforwardly define $\mathcal{V}$ thanks to the RKHS machinery. This approach does not easily extend to non-linear domains such as $\mathcal{M}_0$.

For this reason, here we rely on a construction of the space of smooth vector fields $\mathcal{W}$ based, instead, on the definition of a differential operator encoding smoothness, as done for instance in the planar 2D case in Beg et al. (2005). However, in the planar 2D case a matrix operator for a vector field $u : \mathbb{R}^2 \to \mathbb{R}^2$ can be defined as the isotropic Laplacian

operator

$$\begin{bmatrix} \Delta & 0 \\ 0 & \Delta \end{bmatrix},$$

where Δ is the Laplacian operator for real valued functions. The isotropic Laplacian applies the Laplacian operator component-wise to a vector field in $\mathbb{R}^2$, exploiting the fact that, in the Euclidean space $\mathbb{R}^2$ there is a global reference system. The introduction of an analogous operator for vector fields on a manifold is not straightforward for the main reason that nearby vectors, living on different tangent spaces, cannot be compared component-wise, as they are expressed in different local basis. The definition of such coordinate independent operator for vector fields requires additional notions of Riemannian geometry. In particular, we rely on the Bochner-Laplacian, which is used to enforce smoothness on the vector fields generating diffeomorphism on a manifold.

S2.1 Differential operators on tangent vectors

Recall that we denote with $T_p\mathcal{M}_0$ the tangent space on the point $p \in \mathcal{M}_0$ and with g_p be the metric on $\mathcal{M}_0$. Moreover we denote with $T\mathcal{M}_0 = \dot{\bigcup}_{p\in\mathcal{M}_0} T_p\mathcal{M}_0$ the tangent bundle, i.e. the disjoint union of tangent spaces. The space of smooth sections of the tangent bundle $T\mathcal{M}_0$, i.e. the space of smooth vector fields on $\mathcal{M}_0$, is denoted with $\Gamma(T\mathcal{M}_0)$.

Given a tangent vector $w \in T_p\mathcal{M}_0$, a vector field $u \in \Gamma(T\mathcal{M}_0)$ and a smooth function $f : \mathcal{M}_0 \to \mathbb{R}$, a covariant derivative ∇ is an operator $\nabla_w u \in T_p\mathcal{M}_0$ that is linear in both w and u and is such that it satisfies the Leibniz rule, namely

$$\nabla_w(fu) = df(w)u + f\nabla_w u.$$

For a manifold $\mathcal{M}_0$ embedded in an Euclidean space, by requiring that the affine connection ∇ must preserve the metric and must be torsion free, we have that this can be uniquely determined. Under these hypotheses, ∇ is called the Levi-Civita connection. In

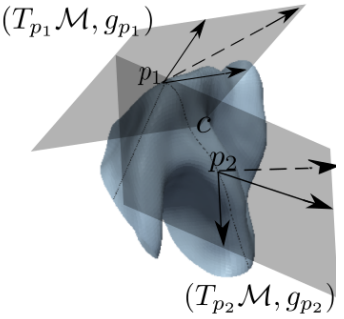

Figure 12: The figure is a pictorial representation of the parallel transport of the striped arrow from p_1 to p_2. Note that because of the different reference systems in p_1 and p_2, expressing the vector as a linear combination of the basis element in $T_{p_2}\mathcal{M}_0$ with the same coefficients as in $T_{p_1}\mathcal{M}_0$ would yield to a different result.

practice a connection defines a way to generalize parallel transport on a manifold. In fact, the parallel transport of a vector $u \in T_p\mathcal{M}_0$ along a curve c can be defined as the collection of vectors along the curve c such that $\nabla_{c'(s)}u = 0$, where $c'(s) \in T_{c(s)}\mathcal{M}_0$. A pictorial representation of this is given in Figure 12. Finally, we can define the Bochner-Laplacian operator, of a smooth section of $v \in \Gamma(T\mathcal{M}_0)$, as

$$\Delta_{BL} = \nabla^*\nabla \tag{20}$$

where ∇^* is the L^2 adjoint of ∇.

S2.1.1 Functional Registration Model

Let now $M, F : \mathcal{M}_0 \to \mathbb{R}$ be respectively a 'moving' and 'fixed' image. We recall here the objective function of the functional registration model (9), in terms of M and F:

$$E_{\mathcal{M}_0}(u) = \sum_{j=1}^{S} \left(F(p_j) - M \circ s \circ \phi_u(p_j)\right)^2 + \lambda \|\Delta_{BL} u\|^2_{L^2(T\mathcal{M}_0)}, \tag{21}$$

with $\{p_j, j = 1, \ldots, S\} \subset \mathcal{M}_0$ the set of control points on the template and ϕ_u denoting the solution of the ODE (8) for the vector field u, at time $t = 1$.

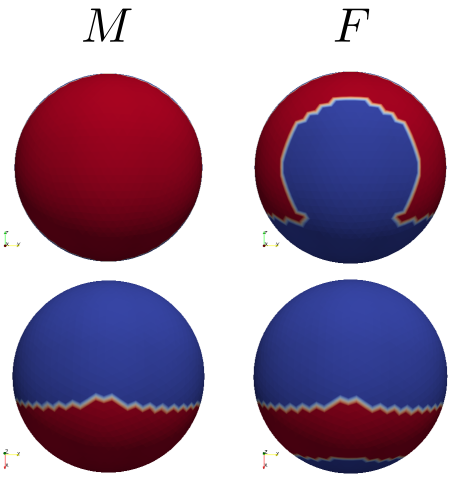

Figure 13: On the left two views of a semi-circle image on the unit sphere, representing the moving image M, while on the right two views of a C-shaped image on the unitary sphere, representing the fixed image F.

The term $M \circ s \circ \phi_u$ in such equation is then linearized with respect to u. This results in the approximation

$$M \circ s \circ \phi_u \approx M \circ s + L_u,$$

where L_u is a first order approximation of $M \circ s \circ \phi_u - M \circ s$. In practice L_u is chosen to be of the form

$$L_u(p) = g_p(J(p), u(p)), \qquad p \in \mathcal{M}_0,$$

with $J(p) \in T_p\mathcal{M}_0$ for all $p \in \mathcal{M}_0$. Two classical choices for J, in the planar case, are $J = \nabla_D(M \circ s)$ and $J = \frac{1}{2}(\nabla_D(M \circ s) + \nabla_D(F))$ (Vercauteren et al., 2009b), where ∇_D denotes a discrete estimate of the gradient. Plugging the linearized term in (21) we obtain the objective function

$$E_{\mathcal{M}_0}(u) = \sum_{j=1}^{S} \left(F(p_j) - (M \circ s)(p_j) - g_{p_j}(J(p_j), u(p_j))\right)^2 + \lambda \|\Delta_{BL} u\|^2_{L^2(T\mathcal{M}_0)}. \tag{22}$$

The minimization of (21) can be achieved by iteratively minimizing the associated problem (22) and updating the current deformation s with $s \to s \circ \phi_u$, with ϕ_u denoting the solution of the ODE (8) at time $t = 1$.

S2.2 Problem reformulation

To minimize the objective function in (22) we opt for a finite elements discretization approach. Finite element discretization has been previously applied to the discretization of FDA problems on manifolds, for instance, in Ettinger et al. (2016) and Lila et al. (2016). Here, we extend the methodology to the estimation of smooth vector fields on a generic two-dimensional manifold. To this end, we first reformulate the minimization of (22) in terms of the Euler-Lagrange equation associated to this minima problem.

Define now the space of smooth vector fields on the template to be $\mathcal{W} = \{u \in L^2(T\mathcal{M}_0) | \Delta_{BL} u \in L^2(T\mathcal{M}_0)\}$. Let the vector field $u \in \mathcal{W}$, in the functional (22), be perturbed by an ε amount along the arbitrary direction $\varphi \in \mathcal{W}$. The minimization problem is reformulated by imposing the Gateaux derivative $\partial_\varphi E_{\mathcal{M}_0}(u)$ of the energy functional to be 0 for all $\varphi \in \mathcal{W}$.

This leads to the problem reformulation: find $\hat{u} \in \mathcal{W}$ such that

$$\begin{aligned} \sum_{j=1}^{S} g_{p_j}(\varphi(p_j), J(p_j)) g_{p_j}(\hat{u}(p_j), J(p_j)) + \lambda \langle \Delta_{BL}\varphi, \Delta_{BL}\hat{u} \rangle_{L^2} = \\ \sum_{j=1}^{S} g_{p_j}(\varphi(p_j), J(p_j))(F(p_j) - M \circ s(p_j)) \end{aligned} \tag{23}$$

for every $\varphi \in \mathcal{W}$. Moreover, equation (23) can be reformulated as the problem of finding $(\hat{f}, \hat{h}) \in \mathcal{W} \times L^2(T\mathcal{M}_0)$ that satisfies

$$\begin{cases} \langle \Delta_{BL}\hat{u}, v \rangle_{L^2} - \langle \hat{h}, v \rangle_{L^2} = 0 \\ \lambda \langle \hat{h}, \Delta_{BL}\varphi \rangle_{L^2} + \sum_{j=1}^{S} g_{p_j}(\varphi(p_j), J(p_j)) g_{p_j}(\hat{u}(p_j), J(p_j)) = \sum_{j=1}^{S} g_{p_j}(\varphi(p_j), J(p_j))(F(p_j) - M \circ s(p_j)) \end{cases} \tag{24}$$

for all $(\varphi, v) \in \mathcal{W} \times L^2(T\mathcal{M}_0)$. In this last reformulation, we have introduced the auxiliary function $\hat{h}$, which has been imposed to be equal, in a weak sense, to $\Delta_{BL}\hat{u}$. Now, asking the auxiliary function v and the test functions φ to be such that $v, \varphi \in \mathcal{W}^1 = \{u \in L^2(T\mathcal{M}_0) | \nabla u \in L^2(T^*\mathcal{M}_0 \otimes T\mathcal{M}_0)\}$, and by exploiting the definition of the Bochner-Laplacian, we can rewrite the problem only in terms of the connection operator ∇, and consequently be able to formulate it in a finite dimensional space involving only first order polynomials, as done in equation (26).

S2.3 Vector Finite Element discretization

Here we introduce a linear finite element space for vector fields on a triangulated surface, where we seek for the discrete solution of the problem (24). To this end, consider the triangulated surface $\mathcal{M}_0^{\mathcal{T}}$, approximated representation of the manifold $\mathcal{M}_0$. $\mathcal{M}_0^{\mathcal{T}}$ is not a smooth surface, so it is not even clear what the tangent space on a vertex of the triangulation is. For this reason, we use elements of computer graphics to define an interpolation basis on the triangulated surface, as done for instance in Zhang et al. (2006); Knöppel et al. (2013).

Let now $\xi_1, \ldots, \xi_K$ be the vertices of $\mathcal{M}_0^{\mathcal{T}}$. For each vertex ξ_k consider the subset of $\mathcal{M}_0^{\mathcal{T}}$ composed by the triangles adjacent to ξ_k, that we call here one-ring. Following the approach in Knöppel et al. (2013), the one-ring surface is idealized by normalizing the

sum of the angles incident to the vertex ξ_k to add up to 2π, i.e. by 'flattening' the vertex and uniformly distributing that curvature to the flat triangles of the one-ring. To the vertex ξ_k they associate a unit vector basis (e_k^1, e_k^2) representing a reference orientation, so that an element of the tangent vector $u_k \in T_{\xi_k}\mathcal{M}_0^{\mathcal{T}}$ will be represented by its coefficients $\mathbf{u}_k \in \mathbb{R}^2$ respect to the local basis. Then, an interpolation basis can be defined on the idealized one-ring of the vertex ξ_k by parallel transporting through geodesics (e_k^1, e_k^2) to the interior points of the one-ring and by scaling them with a piecewise linear function which takes value 1 on ξ_k and 0 one the other vertices of the one-ring (see Knöppel et al., 2013, for details).

What is important to this work is that the outlined procedure leads to a basis of K functions, whose kth function has support localized on the triangles adjacent to ξ_k, and that we denote here with the function $\boldsymbol{\psi}_k = (\psi_k^1, \psi_k^2)$, with ψ_k^1 and ψ_k^2 vector fields on $\mathcal{M}_0^{\mathcal{T}}$. For this basis functions the FE matrices $\langle \boldsymbol{\psi}_k, \boldsymbol{\psi}_{k'} \rangle_{L^2}$ and $\langle \nabla\boldsymbol{\psi}_k, \nabla\boldsymbol{\psi}_{k'} \rangle_{L^2}$ are provided.

We can finally define the FE function space $\mathcal{W}_h$ to be

$$\mathcal{W}_h = \Big\{ u_h = \sum_{k=1}^{K} \boldsymbol{\psi}_k' \mathbf{u}_k | \mathbf{u}_k \in \mathbb{R}^2 \Big\}. \tag{25}$$

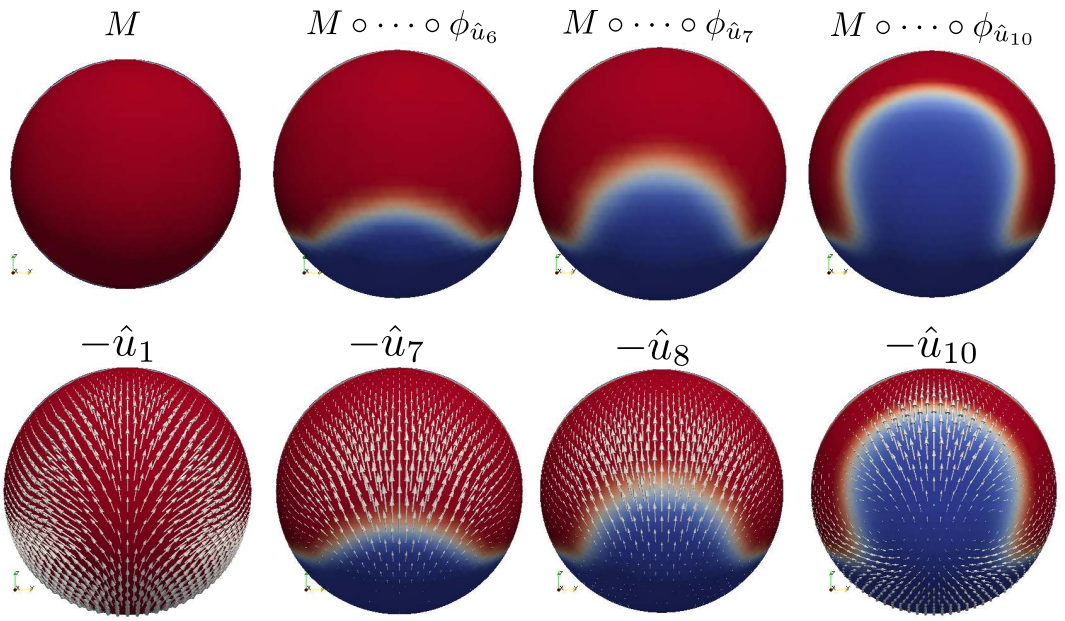

Figure 14: From left to right, the estimated vector fields, and associated deformations of M, at 4 different iterations of the functional registration algorithm. The target is the C-shaped image F.

The solution in the restricted space $\mathcal{W}_h$ is finally given by the discrete approximations $\hat{u}_h, \hat{h}_h \in \mathcal{W}_h$, obtained by solving

$$\begin{cases} \langle \nabla\hat{u}_h, \nabla\varphi_h \rangle_{L^2} - \langle \hat{h}_h, \varphi_h \rangle_{L^2} = 0 \\ \lambda \langle \nabla\hat{h}_h, \nabla v_h \rangle_{L^2} + \sum_{j=1}^{S} g_{p_j}(v_h(p_j), J(p_j)) g_{p_j}(\hat{u}(p_j), J(p_j)) = \sum_{j=1}^{S} g_{p_j}(v_h(p_j), J(p_j))(F(p_j) - M \circ s(p_j)) \end{cases} \tag{26}$$

for all $\varphi_h, v_h \in \mathcal{W}_h$.

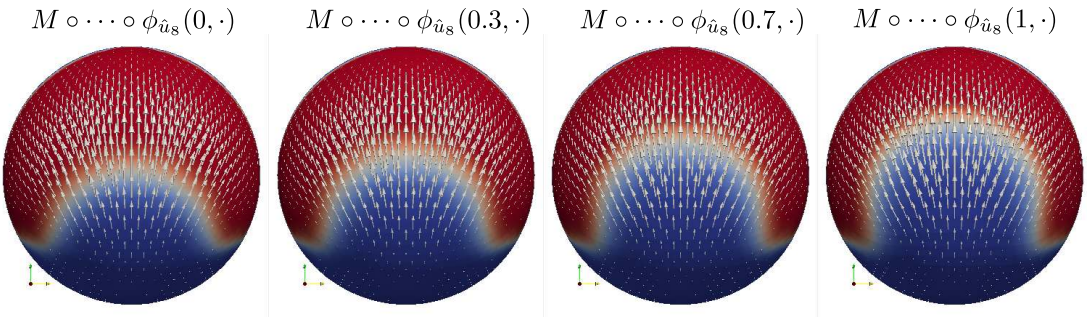

Figure 15: From left to right, evolution of the flow through the ODE (8) for a fixed vector field. The vector field is obtained by the minimization of the linearized objective function (22) at the 8th iteration.

Exploiting the representation (25) of functions in $\mathcal{W}_h$ we can rewrite (26) as a linear system as follows. Let $\hat{\mathbf{u}}$ be a $2K$ vector obtained from the vectorization of the set coefficients $\{\mathbf{u}_i\}$. In the same way let $\hat{\mathbf{h}}$ be the vectorization of the coefficients of $\hat{h}_h$ in (26). Now, introduce the $2K \times S$ matrix Θ_1 and the $2K \times 2K$ matrix Θ_2, such that

$$\mathbf{v}'\Theta_1\mathbf{z} = \sum_{j=1}^{S} g_{p_j}(v_h(p_j), J(p_j))(F(p_j) - M \circ s(p_j))$$
$$\mathbf{v}'\Theta_2\hat{\mathbf{u}} = \sum_{j=1}^{S} g_{p_j}(v_h(p_j), J(p_j))g_{p_j}(\hat{u}(p_j), J(p_j)),$$

with $\mathbf{z}$ the vector of length S such that its jth element is $(F(p_j) - M \circ s(p_j))$ and $\mathbf{v}$ the $2K$ vector obtained from the vectorization of the set coefficients of v_h. These sparse matrices are defined in Section S2.4, together with the $2K \times 2K$ mass and stiffness matrices R_0 and R_1, such that

$$\hat{\mathbf{h}}' R_0 \boldsymbol{\varphi} = \langle \hat{g}_h, \varphi_h \rangle_{L^2}$$
$$\hat{\mathbf{h}}' R_1 \mathbf{v} = \langle \nabla \hat{g}_h, \nabla v_h \rangle_{L^2},$$

where $\boldsymbol{\varphi}$ is a $2K$ vector obtained from the vectorization of the set coefficients φ_h.

Finally, the coefficients $\hat{\mathbf{u}}, \hat{\mathbf{h}}$, of $\hat{u}_h, \hat{h}_h$ are given by the solution of the linear system

$$\begin{bmatrix} \Theta_2 & \lambda R_1 \\ \lambda R_1 & -\lambda R_0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{h}} \end{bmatrix} = \begin{bmatrix} \Theta_1 \mathbf{z} \\ \mathbf{0} \end{bmatrix}, \tag{27}$$

where $\mathbf{0}$ is a $2K$ length zero-vector.

The coefficients $\{\hat{\mathbf{u}}_k\}$ extracted from their vectorization $\hat{\mathbf{u}}$ in (27) represent the approximated tangent vectors on the vertices $\{\xi_k\}$. They are then linearly interpolated to define a solution on $\mathcal{M}_0^{\mathcal{T}}$. This linear piecewise solution on $\mathcal{M}_0^{\mathcal{T}}$ is then used to generate a diffeomorphic transformation through the ODE (8), which is itself approximated with the Euler method. At each step of the Euler method the image of the solution is re-projected on $\mathcal{M}_0^{\mathcal{T}}$. Finally, the current registration is updated by composition with the newly estimated deformation as $s \leftarrow s \circ \phi_{\hat{u}}$, where $\phi_{\hat{u}}$ denotes the solution of the time $t = 1$ given by the Euler method.

In Figure 13, we show an example of a moving image M, which is a semicircle indicator function, and a fixed image F, which is a C-shaped indicator function. They both live on

the same spherical domain. This example tries to replicate the C-shaped planar registration problem, where image registration algorithms are usually tested, as for instance done in Vercauteren et al. (2009a). In Figure 14 we show the vector fields estimated at four different iterations of the Algorithm 1. While in Figure 15, for one particular iteration, we show the evolution of the flow generated by the ODE (8). In this specific example, the domain is chosen to be spherical for visualization purposes, however it can be any smooth two-dimensional manifold, as for instance, in Section 4. The performances of the algorithm, with these synthetic data, are excellent. In fact, only 12 iterations are necessary to register the semicircled indicator function to the C-shaped indicator function.

Finally, it could be argued that being the proposed approximation of the vector field $\hat{u}$ only piecewise linear, and not of higher regularity, this could lead to deformations that are not diffeomorphic. However, the use of reasonably fine triangulated meshes $\mathcal{M}_0^{\mathcal{T}}$ should solve the problem. After all, in practice, even for higher regularity vector fields, the computer resolution of the ODE relies on a finite number of sampled values from the vector field, and thus on a non smooth vector field.

S2.4 Finite element matrices

Assume, for simplicity, that the points $\{p_j\}$ coincide with the nodes $\{\xi_k : 1, \dots, K\}$ of the mesh $\mathcal{M}_0^{\mathcal{T}}$. The non-zero entries of the matrices Θ_1 and Θ_2 are

$$\{\Theta_1\}_{2k,k} = g_{\xi_k}(J(\xi_k), e_1^k),$$
$$\{\Theta_1\}_{2k+1,k} = g_{\xi_k}(J(\xi_k), e_2^k)$$

and

$$\{\Theta_2\}_{2k,2k} = g^2_{\xi_k}(J(\xi_k), e_1^k), \qquad \{\Theta_2\}_{2k,2k+1} = g_{\xi_k}(J(\xi_k), e_1^k) g_{\xi_k}(J(\xi_k), e_2^k),$$
$$\{\Theta_2\}_{2k+1,2k} = g_{\xi_k}(J(\xi_k), e_1^k) g_{\xi_k}(J(\xi_k), e_2^k), \qquad \{\Theta_2\}_{2k+1,2k+1} = g^2_{\xi_k}(J(\xi_k), e_2^k)$$

with the matrices indexed from zero and $k = 0, \dots, K-1$. The computation of the entries $g_{\xi_k}(J(\xi_k), e_1^k)$ can be performed by representing the tangent vectors $J(\xi_k)$ and e_1^k as vectors in $\mathbb{R}^3$ and computing the $\mathbb{R}^3$ Euclidean scalar product between them, as in fact the manifold $\mathcal{M}_0$, and its associated triangulated mesh $\mathcal{M}_0^{\mathcal{T}}$, are embedded in $\mathbb{R}^3$. The entries of the $2K \times 2K$ matrices R_0 and R_1 in (26) are computed in (Knöppel et al., 2013, Section 6.1.1), for the purpose of computing eigen-vectors of the Bochner-Laplacian operator.

S2.5 Boundary Conditions

The deformations generated by the functional registration algorithm are by definition constrained to be maps with their image on the template surface, since the ODE (8) is defined on the manifold itself. However, if the template is a manifold with a boundary, as in the simulations performed in Section S3, the vector might generate deformations that transport the functions outside the boundary. This can be avoided by imposing homogeneous Dirichlet boundary conditions on the estimated vector field. Dirichlet boundary conditions can be implemented in different ways. Here, we opt for applying them after the linear system (27) has been built. In particular given a boundary node k, we add a large constant M to the entries $2k, 2k$ and $2k+1, 2k+1$ of the left hand side matrix and set to 0 the entries $2k$ and $2k+1$ of the right hand side vector. As a consequence, the vector

fields estimated from the modified linear system will smoothly vanish as approaching the boundary.

S3 Analysis of a synthetic dataset

In this section, we validate the estimation framework introduced in Section 3, by performing a study on a dataset generated from Model 1. We thus proceed with defining the unknown quantities of such model. We will not use different notation for the theoretical objects and their respective computer representations, unless necessary.

Thus, we denote with $\mathcal{M}_0$, the template temporal lobe shown in Figure 16. We set

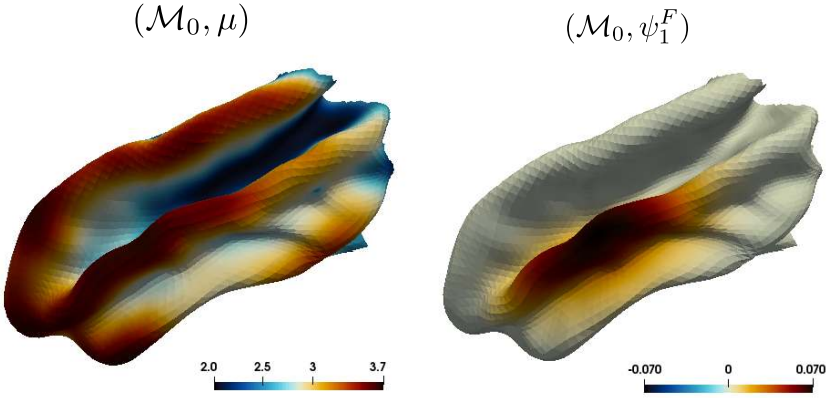

Figure 16: On the left, a template of the temporal lobe $\mathcal{M}_0$ with an associated cortical thickness map μ. On the right, the function ψ_1^F used to generate subject-specific functional variability.

the deformation operator φ to be the diffeomorphic deformation operator introduced in Section 2.4. We then choose two orthonormal vector fields $\psi_1^G, \psi_2^G \in \mathcal{V}$, visualized in Figure 17 as the deformations $\varphi_{\pm c\psi_1^G}, \varphi_{\pm c\psi_2^G}$ applied to the template $\mathcal{M}_0$, where $c \in \mathbb{R}$ is a constant regulating the norm of the two orthonormal vector fields, for visualization purposes. The vector field ψ_1^G encodes a change in the length of the temporal lobe, while the vector field ψ_2^G encodes a change in the size of temporal lobe.

We set the mean function $\mu \in L^2(\mathcal{M}_0)$, to be the thickness maps in Figure 17, which is a sharpened version of the cross-sectional average thickness of 100 real subjects. Note that despite it being computed from real data, this plays the role of an unknown quantity of the model. Further details on the real data are left to Section 4. Moreover, we introduce localized functional variability through the single mode of variation $\psi_1^F \in L^2(\mathcal{M}_0)$, this also visualized in Figure 16.

We then generate $n = 50$ FoSs $(\mathcal{M}_1, Y_1), \ldots, (\mathcal{M}_n, Y_n)$ by

$$\begin{cases} \mathcal{M}_i &= \varphi_{a_{i1}\psi_1^G + a_{i2}\psi_2^G} \circ \mathcal{M}_0, \\ X_i &= \mu + \delta a_{i2}\psi_1^F, \\ Y_i &= X_i \circ \varphi^{-1}_{a_{i1}\psi_1^G + a_{i2}\psi_2^G}, \end{cases} \tag{28}$$

where a_{i1}, a_{i2} are independent random variables distributed as $a_{il} \sim N(0, \sigma_l^2)$, with $\sigma_1 = 15$ and $\sigma_2 = 10$. The constant $\delta = 0.1$ determines the scale that relates variations in the functional term $\delta a_{i2}\psi_1^F$ and variations in the geometric term $a_{i2}\psi_2^G$. Finally, normally

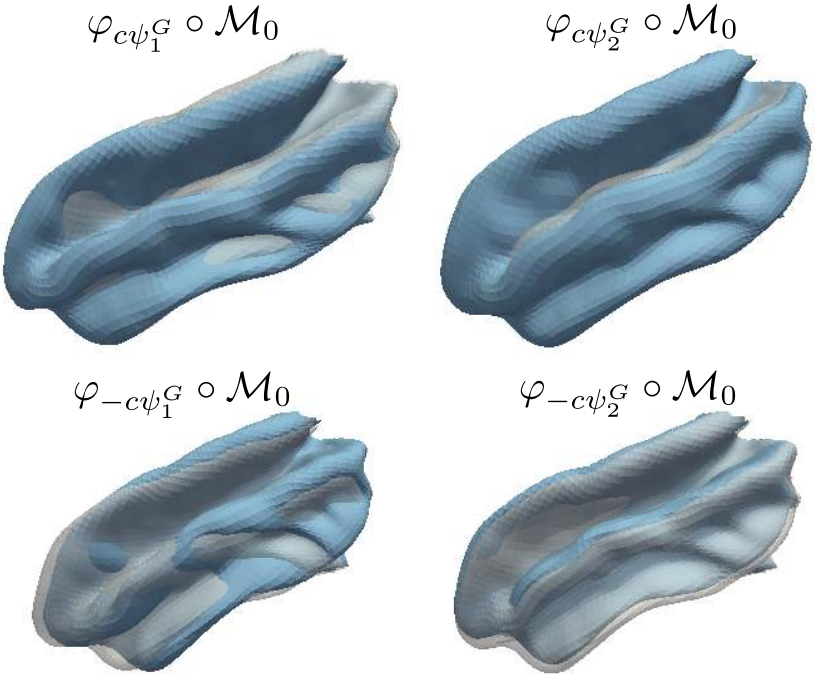

Figure 17: From left to right, first and second geometric modes of variation of the generated FoSs, here visualized as $\varphi_{\pm c\psi_1^G} \circ \mathcal{M}_0, \varphi_{\pm c\psi_2^G} \circ \mathcal{M}_0$, where $c \in \mathbb{R}$ is a constant regulating the magnitude for visualization purposes.

distributed noise with variance $\sigma = 0.3$, is added to each node of the mesh where the function is observed. The generative model proposed here is a simplistic implementation of the one proposed in Model 1, with $v_i = a_{i1}\psi_1^G + a_{i2}\psi_2^G$ and $Z_i = a_{i2}\psi_1^F$.

The generative model (28) seeks to reproduce a situation where the FoSs have two modes of geometric variation. The first one is a mode of variation which is not correlated with a variation in the functions. The second one, which encodes the size of the temporal temporal lobe, has an effect on the function, formalized with a linear relation between the scores of the second geometric mode of variation ψ_2^G and the scores of the functional mode of variation ψ_1^F. The generated FoSs are such that larger temporal lobe have larger cortical thickness in proximity of the central gyrus of the cerebral cortex, independently of the first geometric mode of variation. We hope to recover this relation through the approximation pipeline introduced in Section 3.

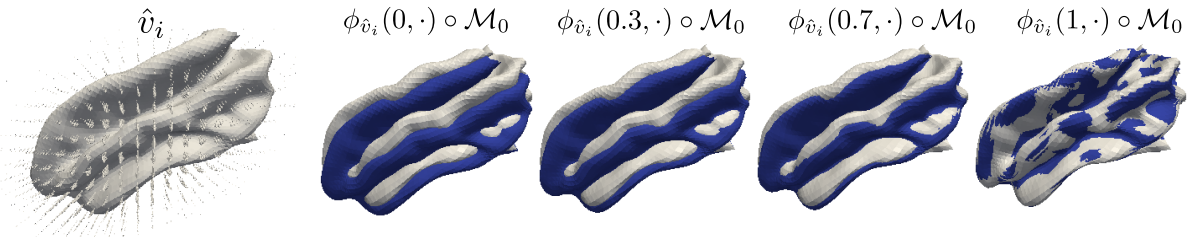

Figure 18: On the left, the template $\mathcal{M}_0$ with an estimated vector field $\hat{v}_i \in \mathcal{V}$ generating the diffeomorphic deformation $\varphi_{\hat{v}_i}$ that registers the template to the ith subject surface. Next, the evolution of the flow generating the diffeomorphic deformations $\phi_{\hat{v}_i}(t,\cdot)$ through the ODE (3), which registers the template to the target surface at time $t = 1$.

Figure 19: Two vector fields estimated from the functional registration algorithm, generating, for two different subjects, the flow which aligns two different functions to the cross-sectional mean function.

In particular, we perform non-landmarked diffeomorphic registration of the template to the single surfaces, resulting in the estimated vector fields $\{\hat{v}_i : i = 1, \dots, n\}$. The ith vector field $\hat{v}_i$ is such that $\varphi_{\hat{v}_i} \circ \mathcal{M}_0$ resembles the geometry $\mathcal{M}_i$ of the ith FoS, with φ the diffeomorphic deformation operator. In Figure 18, we show an estimated vector field $\hat{v}_i \in \mathcal{V}$ and the ODE's (3) flow $\phi_{\hat{v}_i}(t, \cdot)$, generated from the estimated vector field, which deforms the template to match the target.

The estimated diffeomorphic deformations $\{\varphi_{\hat{v}_i} = \phi_{\hat{v}_i}(1, \cdot)\}$ are then used to transport the functions $\{Y_i\}$ on the template surface, thus leading to the estimates $\{\hat{X}_i\}$. Subsequently, the cross-sectional mean map of $\{\hat{X}_i\}$ is computed and each function $\hat{X}_i$ is iteratively registered to it through the functional registration algorithm presented in Section 3.2. In Figure 19, we show the template surface, with the tangential vector fields that generate the deformations that align two different functions to the cross-sectional mean function.

At each iteration of the functional registration algorithm, the cross-sectional mean and the first 2 functional PCs, from the functionally aligned versions of $\{\hat{X}_i\}$, are computed. The results are shown in Figure 20. We can notice that while the cross-sectional mean does not change from iteration to iteration of the functional registration algorithm, the estimates of the PC functions do. In particular, the first PC function is supposed to capture ψ_1^F. However, where no functional registration is applied, the first estimated PC component is a mix of the ψ_1^F and fictitious variability due to misalignment, while the second PC function is a flat and corrupted version of ψ_1^F. After only one iteration of the functional registration algorithm, the estimated first PC function starts resembling the shape of ψ_1^F, shifting the misalignment component to the second PC function. With the subsequent iteration the first estimated PC function becomes a sharper estimation of ψ_1^F, while the misalignment component disappears also from the second component, in favour of a flat PC function, which is a regularized PC function of the noise added to the functions.

Subsequently, we perform fPCA on the estimated vector fields $\{\hat{v}_i\}$ representing the overall deformation, due to both geometric and functional registration. In Figure 21 we show the estimated main modes of variation before the functional registration has been applied. By comparison with Figure 17, we can see that the first two PCs capture the main geometric modes of variations introduced in the generative process of the FoSs. The estimated geometric PC function do not change, in a visible manner, from iteration to iteration of the functional registration algorithm, because the functional registration brings only small deformations.

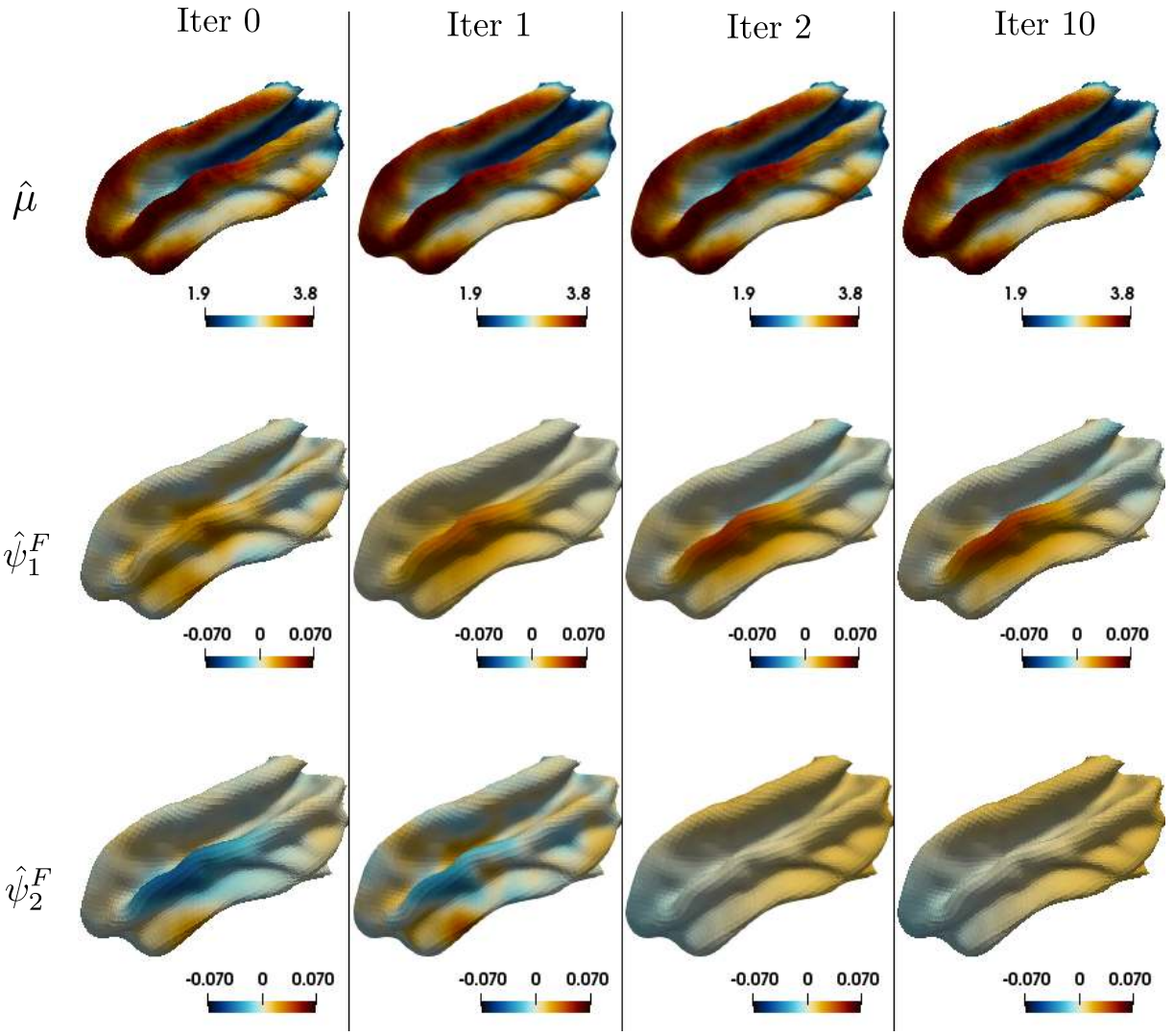

Figure 20: From left to right, the mean and first two functional PC functions estimates of the functions, computed after $0, 1, 2$ and 10 iterations of the functional registration algorithm.

We finally plot, in Figure 22, the scores associated to the PCs describing the geometric variability and those describing the functional variability, for the estimated quantity without functional registration and after seven iterations of the functional registration. Note that without performing functional registration, not only is the first functional mode of variation a spurious version of the true underlying component, but this is also correlated to the geometric mode of variations, which might lead to misleading conclusions. Functional registration removes from the first PC the misalignment effect, bringing to light the true underlying linear dependence between the functional mode of variation and the second geometric mode of variation.

In practice, the above procedure is particularly useful if the discovered PCs have biological interpretations. However, in practice, the discovered PCs tend to vary, depending for instance on the pre-registration method applied or on the scalar product adopted to impose orthogonality between the PC functions. For these reasons, if the aim is to study the relation between geometry and function, we advocate CCA (see Section 3.4). We perform a CCA on the estimated scores of the geometric and functional variability,

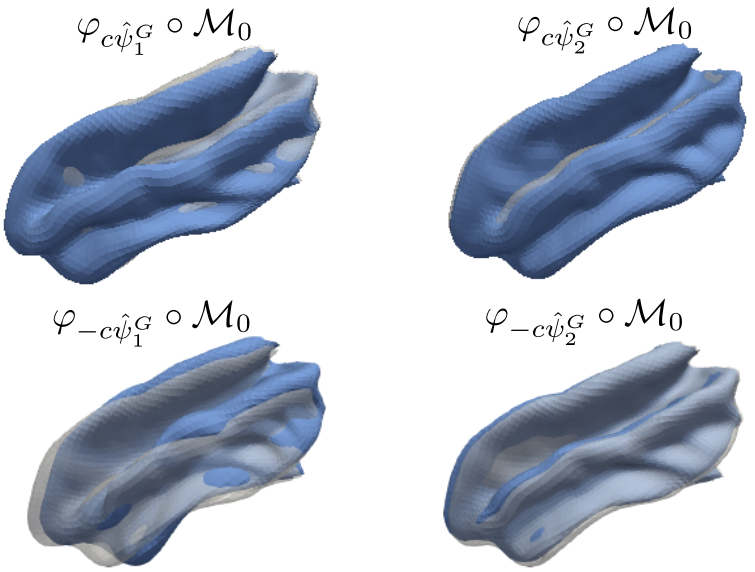

Figure 21: From left to right, the first two geometric PC functions computed on the space of initial vector fields. These are visualized as $\varphi_{\pm c\hat{\psi}_j^G} \circ \mathcal{M}_0$, where $\hat{\psi}_j^G$ is the estimated jth geometric PC function.

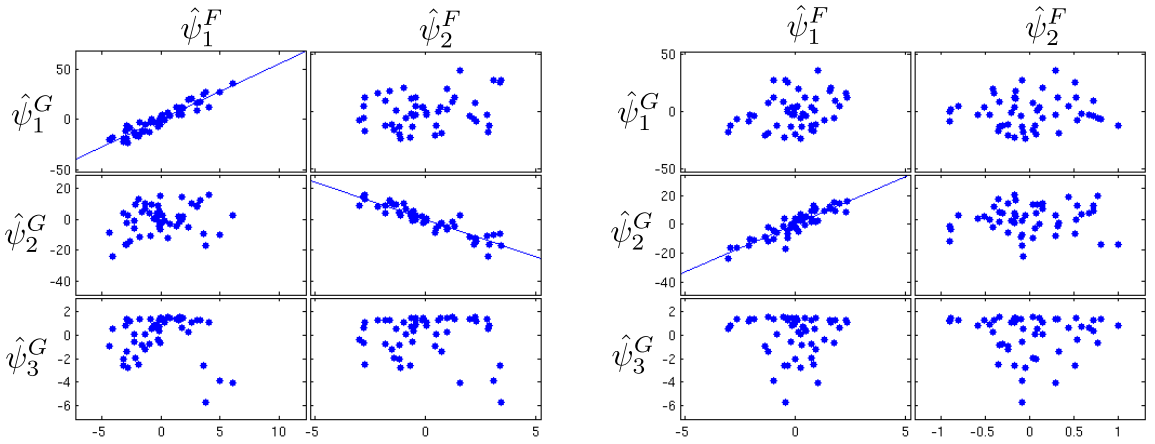

Figure 22: From left to right, scatter plots of the scores obtained from the fPCA on the function and the geometric fPCA, respectively without and with functional registration. After functional registration, these show only the linear dependence imposed between the first PC function on the functions ψ_1^F and the second geometric PC function ψ_2^G. Without functional registration, also the spurious PC function, due to misalignment, is correlated with the first geometric PC function.

after seven iterations of the functional registration algorithm. In detail, we construct a $n \times 3$ matrix $\mathbb{X}^F$ with the scores of the first three components of the fPCA applied to the functions. Moreover, we construct a $n \times 5$ matrix $\mathbb{X}^G$ with the scores of the first five components of the fPCA applied to the deformations. The lth canonical correlation component is the pair of vectors $\hat{\mathbf{w}}^{F,l} \in \mathbb{R}^3$ and $\hat{\mathbf{w}}^{G,l} \in \mathbb{R}^5$. The resulting main mode of

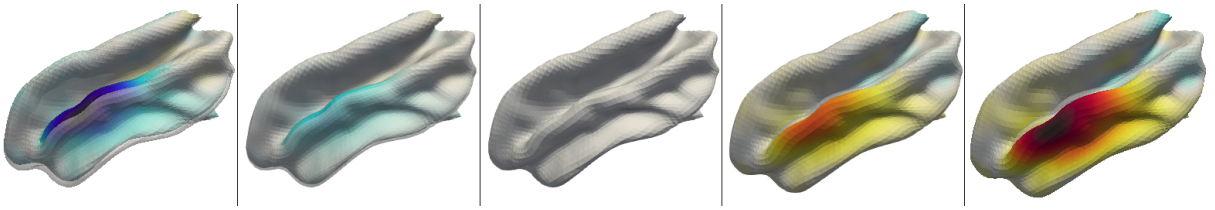

Figure 23: First main mode of co-variation of geometric and functional components of the CCA analysis, representing the most correlated linear combinations of the first four geometric modes of variation and first three functional modes of variation. From left to right, this is visualized by plotting the FoS in (16) for a sequence of constants c.

co-variation $(\hat{\mathbf{w}}^G, \hat{\mathbf{w}}^F) = (\hat{\mathbf{w}}^{G,1}, \hat{\mathbf{w}}^{F,1})$ is visualized in Figure 23 as

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

with $\hat{\psi}^G_{\text{CCA}} = \sum_{j=1}^3 \hat{w}^G_j \hat{\psi}^G_j$ and $\hat{\psi}^F_{\text{CCA}} = \sum_{j=1}^5 \hat{w}^F_j \hat{\psi}^F_j$, where $(\hat{\psi}^G_j)$ and $(\hat{\psi}^F_j)$ are the estimated functional and geometric PC components. $c \in \mathbb{R}$ is a constant varied for visualization purposes in an interval containing 0. As we can see in Figure 23, the dependence between the magnitude and the thickening of the function is captured.

Moreover, we test for the statistical significance of the obtained modes of co-variation. Specifically, we test the hypotheses

$$H_0^l : \hat{\rho}_1 \neq 0, \hat{\rho}_2 \neq 0, \ldots, \ldots, \hat{\rho}_l \neq 0, \hat{\rho}_{l+1} = \ldots = 0, \tag{29}$$

with $\hat{\rho}_l = \text{corr}(\mathbb{X}^G \hat{\mathbf{w}}^{G,l}, \mathbb{X}^F \hat{\mathbf{w}}^{F,l})$. According to a likelihood ratio test, with the Bartlett χ^2 approximation of the test statistic distribution (see Johnson and Wichern, 2007, Chapter 10.6), only the sample correlation between the first canonical correlation variables, i.e. $\mathbb{X}^G \hat{\mathbf{w}}^{G,1}$ and $\mathbb{X}^F \hat{\mathbf{w}}^{F,1}$, is significantly different from zero (p-value $5e-19$), while for $l = 2, 3$ we get p-values 0.7759 and 0.9587 respectively.

S4 Further Simulations

As previously mentioned, the functional registration algorithm introduced in Section 3.2 is not the only option to account for functional information in the registration process. Here we compare our methodology to the joint functional and geometric registration algorithm proposed in Charon and Trouvé (2014), where the shape similarity functional (7) is extended to include a functional similarity term.

Suppose now that the template mesh $\mathcal{M}_0^{\mathcal{T}}$ is equipped with a functional object $\mu^{\mathcal{T}} : \mathcal{M}_0^{\mathcal{T}} \to \mathbb{R}$, which in first instance can be the cross-sectional mean of the functions $\hat{X}_i$ estimated after the geometric registration described in Section 3.1. We briefly recall the notation in Section 3.1, introduced to define (7). We define $K_{\mathcal{Z}} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ to be a Gaussian isotropic kernel of variance $\sigma_{\mathcal{Z}}$, i.e. $K_{\mathcal{Z}}(x, y) = \exp(-\|x - y\|_2^2/(2\sigma_{\mathcal{Z}}^2))\text{Id}_{3\times 3}$, with $\text{Id}_{3\times 3}$ denoting a 3×3 identity matrix. Additionally, we introduce a scalar Gaussian kernel $K_{\mathcal{F}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ of the form $K_{\mathcal{F}}(x, y) = \exp(-(x - y)^2/(2\sigma_{\mathcal{F}}^2))$.

Moreover, we denote with $c(l)$ and $\eta(l)$, respectively, the center point and the normal vector of the lth triangle of the mesh $\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}$. Instead, we denote with $c_i(q)$ and $\eta_i(q)$, respectively, the center point and the normal vector of the qth triangle of the mesh $\mathcal{M}_i^{\mathcal{T}}$.

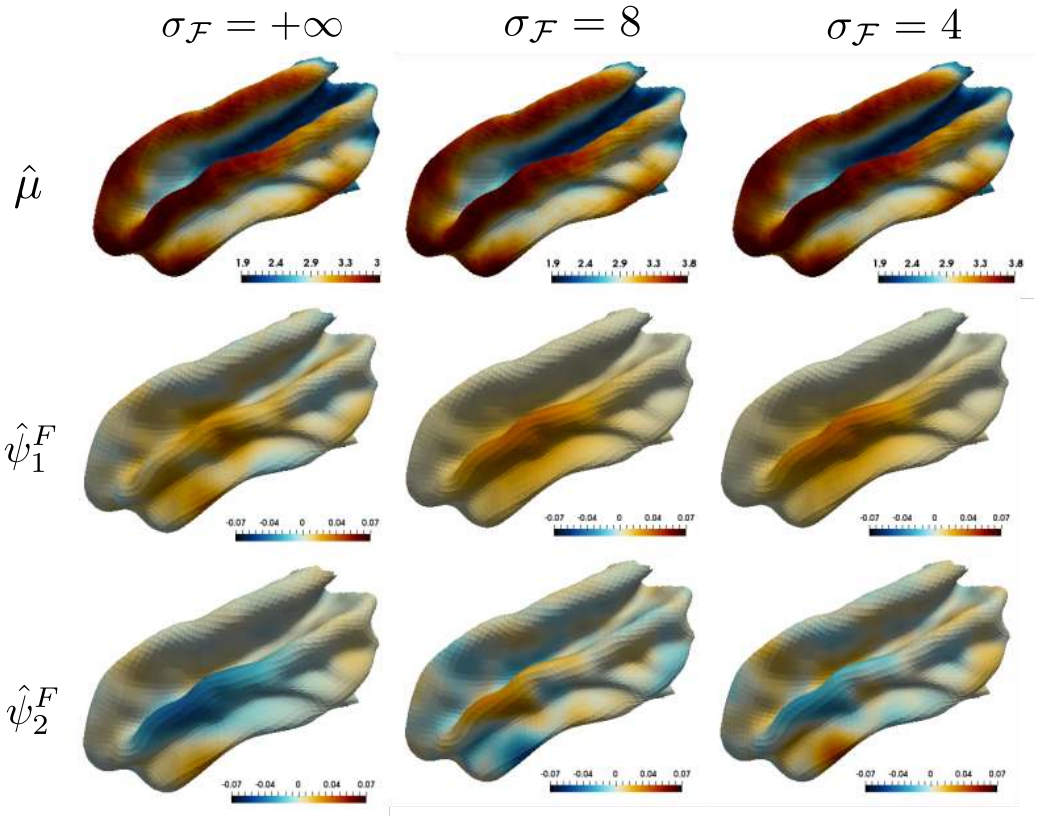

Figure 24: From left to right, the mean and first two functional PC functions estimates of $\{\hat{X}_i\}$ estimated by using the registration maps computed by solving (5) with the extended matching function in (30), for different choices of $\sigma_{\mathcal{F}}$.

Additionally, we introduce $y(l)$, denoting the functional value $\mu^{\mathcal{T}}$, associated to the mesh $\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}$, at the center point of the lth triangle. We denote with $y_i(q)$ the functional value associated to the ith FoSs at the center point of the qth triangle of the mesh $\mathcal{M}_i$.

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 shape similarity functional (7) can be extended to include functional informations as follows (Charon and Trouvé, 2014).

$$
\begin{aligned}
D^2\big((\varphi_{v_i} \circ \mathcal{M}_0^{\mathcal{T}}, \mu^{\mathcal{T}} \circ \varphi_{v_i}^{-1}), (\mathcal{M}_i^{\mathcal{T}}, Y_i^{\mathcal{T}})\big) = \\
&\sum_l \sum_g K_{\mathcal{F}}(y(l), y(g)) K_{\mathcal{Z}}(c(l), c(g)) \eta(l) \cdot \eta(g) \\
&- 2 \sum_l \sum_q K_{\mathcal{F}}(y(l), y_i(q)) K_{\mathcal{Z}}(c(l), c_i(q)) \eta(l) \cdot \eta_i(q) \\
&+ \sum_q \sum_r K_{\mathcal{F}}(y_i(q), y_i(r)) K_{\mathcal{Z}}(c_i(q), c_i(r)) \eta_i(q) \cdot \eta_i(r),
\end{aligned}
\tag{30}
$$

with $\cdot$ denoting the scalar product in $\mathbb{R}^3$. Each term now, measures not only differences in geometry but also differences in the functional values between the template and the target FoS.

Subsequently, given the FoSs $\{(\mathcal{M}_i^{\mathcal{T}}, Y_i^{\mathcal{T}})\}$ generated as described in Section S3, we perform the landmark-free geometric registration by minimizing the objective function in (5), with the shape similarity functional (7), which is equivalent to the similarity functional (30) with $\sigma_{\mathcal{F}} = +\infty$. Thanks to the estimated registration maps we can estimate the

functions $\{\hat{X}_i\}$ and compute the cross-sectional mean function $\mu^{\mathcal{T}}$. Subsequently a second registration step can be performed, by minimizing the objective function in (5) but this time with the similarity functional (30). We have performed this for different choice of $\sigma_{\mathcal{F}}$. The smaller $\sigma_{\mathcal{F}}$, the more we are weighting the functional matching term as opposed to the geometric matching term.

In Figure 24 we show the results of the fPCA applied to the functions $\{\hat{X}_i\}$ for different choices of $\sigma_{\mathcal{F}}$. These need to be compared with the results in Figure 20, obtained by applying the iterative functional registration algorithm in Section 3.2. On the left panel of Figure 24 we can see the mean and first two PC functions estimated when functional information is ignored, which coincide with the one showed on the left panel of Figure 20, as they are computed in the same way. On the other two panels of Figure 24 we can see the mean and first two PC functions estimated when functional information is introduced. As we can see the estimated first PC function resembles the true underlying first PC function, but some fictitious variability is left on the second estimated PC function.

Trying to further decrease $\sigma_{\mathcal{F}}$, to remove the residual fictitious variability, resulted in estimated registration maps failing to bring the template in geometric correspondence to the target surface. Such problem has been the limiting factor in successfully applying the same method to the data in the real application, where the differences in geometries between the template and the target FoSs are much bigger. In fact, this is one of the motivations underlying the introduction of the functional registration algorithm in Section 3.2, where the 'moving' functions are instead 'constrained' to lie in the predefined geometry.

References

M. F. M. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.

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.

B. Ettinger, S. Perotto, and L. M. Sangalli. Spatial regression models over two-dimensional manifolds. *Biometrika*, 103(1):71–88, 2016.

R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis (6th Edition)*. Pearson, 6 edition, Apr. 2007. ISBN 0131877151.

F. Knöppel, K. Crane, U. Pinkall, and P. Schröder. Globally optimal direction fields. *ACM Trans. Graph.*, 32(4):59:1–59:10, July 2013.

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.

M. I. Miller, A. Trouvé, and L. Younes. Hamiltonian Systems and Optimal Control in Computational Anatomy: 100 Years Since D'Arcy Thompson. *Annual Review of Biomedical Engineering*, 17(1):447–509, dec 2015.

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

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

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

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.

E. Zhang, K. Mischaikow, and G. Turk. Vector field design on surfaces. *ACM Trans. Graph.*, 25(4):1294–1326, Oct. 2006.