

DIFFEOMORPHIC SURFACE REGISTRATION WITH ATROPHY CONSTRAINTS

ABSTRACT. Diffeomorphic registration using optimal control on the diffeomorphism group and on shape spaces has become widely used since the development of the Large Deformation Diffeomorphic Metric Mapping (LDDMM) algorithm. More recently, a series of algorithms involving sub-riemannian constraints have been introduced, in which the velocity fields that control the shapes in the LDDMM framework are constrained in accordance with a specific deformation model. Here, we extend this setting by considering, for the first time, inequality constraints, in order to estimate surface deformations that only allow for atrophy, introducing for this purpose an algorithm that uses the augmented lagrangian method. We prove the existence of solutions of the associated optimal control problem, and the consistency of our approximation scheme. These developments are illustrated by numerical experiments on brain data.

1. INTRODUCTION

Over the last couple of decades, multiple studies have provided evidence of anatomical differences between control groups and cognitively impaired groups at the population level, for a collection of diseases, including schizophrenia, depression, Huntington’s or dementia [36, 18, 11, 37, 54, 42, 1, 29, 41, 52, 33]. In the particular case of neuro-degenerative diseases, a repeated objective has been to design anatomical biomarkers, measurable from imaging data, that would allow for individualized detection and prediction. This goal has become even more relevant with the recent emergence of longitudinal studies, involving patients at early stages or “converters” which showed that, when the effect is measured at the population level, anatomical changes caused by diseases like Alzheimer’s or Huntington’s were happening several years before cognitive impairment could be detected on individual subjects.

Shape analyses from medical imaging data rely most of the time on a *registration step* that places the subjects’ anatomies in a common coordinate system, often associated with a template, or average shape, to which all images are co-registered [15, 44, 46, 8, 5, 6, 9, 10, 31, 32, 47, 48]. A large number of registration methods have been proposed in the literature, with an extensive survey proposed in [40], including more than 400 references for image registration alone. Surface registration has also been extensively developed [24, 23, 45, 17, 27, 30, 4, 49, 56, 7, 13, 28, 55, 12].

We focus in this paper on surface registration using the LDDMM algorithm, which has been used extensively to analyze shape variation in regions of interest (ROIs) represented by triangulated surfaces. While one its main advantages is its flexibility and its ability to render smooth, diffeomorphic, free-form shape changes, there are situations in which prior information is available, and should be used to enhance the results of the shape

analysis. In this paper, we focus on situations in which no tissue growth is expected to occur (like with brain longitudinal data). In such contexts, it is natural to ensure that shape analysis should only detect atrophy, even when noise and inaccuracy in the ROI segmentation process may lead in the other direction. (Here, we mean “atrophy” in the general sense of local volume loss.) In this paper, we introduce an *atrophy-constrained* registration algorithm, that include some of the ideas introduced in [2], while extending them to inequality constraints associated to the problem we consider. This algorithm will be described in section 2, with our numerical approach discussed in section 3. Some theoretical results on existence of solutions and consistency of discrete approximations are provided in section 4. An extension of the algorithm to include affine alignment is provided in section 5. Another extension to time series is discussed in section 6. Finally, experimental results are provided in section 7.

2. ATROPHY-CONSTRAINED LDDMM

2.1. Continuous Optimal Control Problem. The LDDMM algorithm implements an “optimal control” strategy in which a template surface S_0 is “driven” toward a target surface S_1 via a time-dependent process $t \mapsto S(t)$, with $S(0) = S_0$. This is achieved by minimizing

$$(1) \quad \frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + D(S(1), S_1)$$

subject to the state equation $dS/dt = v(t, S(t))$ where v is a smooth velocity field on \mathbb{R}^3 . By $\|v\|_V^2$, we mean a functional norm in a reproducing kernel Hilbert space (RKHS) V , that we will assume to be embedded in $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ (the completion, for the standard supremum norm of up to p derivatives, of the space of compactly-supported infinitely differentiable vector fields) with $p \geq 1$. This space can, for example, be defined as the Hilbert completion of

$$\|v\|_V^2 = \int_{\mathbb{R}^3} (Av(x))^T v(x) dx$$

(originally defined for smooth vector fields), where A is a differential operator. A typical example is $A = (\text{Id} - \Delta)^k$, where Id is the identity, Δ is the Laplacian operator and k is large enough to ensure that required Sobolev inclusions hold. In the general case, let $\mathbb{A} : V \mapsto V^*$ be the Hilbert duality mapping, so that $\mathbb{A}v$ is the linear form $w \mapsto \langle v, w \rangle_V$, and let $\mathbb{K} = \mathbb{A}^{-1}$. The reproducing kernel of V , also denoted \mathbb{K} , is a mapping defined on $\mathbb{R}^3 \times \mathbb{R}^3$, with values in $\mathcal{M}_3(\mathbb{R})$ (the set of 3 by 3 real matrices) such that, letting $\mathbb{K}^i(x, y)$ denote the i th column of $\mathbb{K}(x, y)$, the vector field $\mathbb{K}^i(\cdot, y) : x \mapsto \mathbb{K}^i(x, y)$ belongs to V for all $y \in \mathbb{R}^3$ with, for all $v \in V$, $\langle \mathbb{K}^i(\cdot, y), v \rangle_V = v_i(y)$, the i th coordinate of $v(y)$. To simplify the notation, we will assume in this paper that \mathbb{K} is a scalar kernel, i.e., that it takes the form $\mathbb{K}(x, y) = K(x, y) \text{Id}_{\mathbb{R}^3}$ where K is scalar valued.

The function D in (1) is a measure of discrepancy that penalizes the difference between the controlled surface $S(\cdot)$ at the end of its evolution and the target surface S_1 . Among the measures introduced in the literature in combination with the LDDMM algorithm, the most convenient computationally are designed as Hilbert-space norms between surfaces

considered as linear forms over spaces of smooth structures. The “simplest” example, linear forms on smooth scalar functions arising from integrating functions over surfaces, yields the “measure matching” cost introduced in [21, 22]. “Current matching”, introduced in [20, 45] results from integrating smooth differential forms over oriented surfaces. More recently varifold-based cost functions [12] were designed, in which functions defined on $\mathbb{R}^3 \times Gr(2, \mathbb{R}^3)$ (the Grassmannian manifold of 2D spaces in \mathbb{R}^3) are integrated over the surface. Details on these cost functions, their discrete versions on triangulated surfaces, and the computation of their gradient are provided in the cited references. To focus the discussion, we will here assume that D is the current-matching norm for which

$$(2) \quad D(S, S') = \langle S, S \rangle_\chi - 2\langle S, S' \rangle_\chi + \langle S', S' \rangle_\chi$$

where the “dot product” between surfaces is defined by

$$(3) \quad \langle S, S' \rangle_\chi = \int_{S \times S'} \chi(s, s') N(s)^T N(s') d\mathcal{A}_S(s) d\mathcal{A}_{S'}(s')$$

where χ is a positive definite kernel and $\mathcal{A}_S, \mathcal{A}_{S'}$ refer to the area forms on S and S' .

Interpreting (1) in optimal control language, v is the “control”, S is the “state”, and v is optimized in order to bring the state near a desired endpoint. With this construction, each point x_0 in S_0 is registered to a point $x(t) = \varphi(t, x_0)$ in $S(t)$ that evolves according to the differential equation $dx/dt = v(t, x)$, with $x(0) = x_0$. The overall evolution is diffeomorphic, i.e., for each time t , $\varphi(t, \cdot)$ can be extended to a smooth invertible transformation with smooth inverse on \mathbb{R}^3 .

To define our atrophy constraints, we assume that surfaces are closed and oriented. We let $N_0(t, x_0)$ be the outward-pointing unit normal to S_0 . An outward-pointing normal to $S(t)$ at $x = \varphi(t, x_0)$ is then given by

$$(4) \quad N(t, x) = d\varphi(t, x_0)^{-T} N_0(t, x_0)$$

where $d\varphi(t, x_0)$ denote the differential of $y \mapsto \varphi(t, y)$ at $y = x_0$ (a 3 by 3 matrix), with the “ $-T$ ” exponent indicating the transposed inverse. Note that $N(t, \cdot)$ does not necessarily have norm one.

We can express the atrophy constraint by the fact that the surface evolves inward at all points, i.e., by $v(t, x)^T N(t, x) \leq 0$ for all $x \in S(t)$ and $t \in [0, 1]$. Adding this constraint to the original surface-matching LDDMM problem leads to the atrophy-constrained problem of minimizing

$$\frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + D(S(1), S_1)$$

subject to $\partial_t S = v(t, S(t))$, $v(t, x)^T N(t, x) \leq 0$, $x \in S(t)$.

We now reformulate this problem under the assumption that S_0 is parametrized with an embedding $q_0 : M \rightarrow \mathbb{R}^3$, where M is a two-dimensional Riemannian manifold. This is no loss of generality, since one can always take $M = S_0$ and $q_0 = \text{identity}$. It will also be useful when discussing discrete approximations to relax the atrophy constraint in the form

$v(t, x)^T N(t, x) \leq \varepsilon |N(t, x)|$, $x \in S(t)$, which allows for a small amount of expansion, with normal velocity less than ε . Given this, we take parametrizations $q : M \rightarrow \mathbb{R}^3$ as state variables, together with functions $N : M \rightarrow \mathbb{R}^3$ and solve

Problem 1. *Minimize*

$$(5) \quad \frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + D(q(1, M), S_1)$$

$$(6) \quad \text{subject to} \quad \begin{cases} q(0, \cdot) = q_0, N(0, \cdot) = N_0, \\ \partial_t q(t, \cdot) = v(t, q(t, \cdot)), \\ \partial_t N(t, \cdot) = -dv(t, q(t, \cdot))^T N(t, \cdot), \\ v(t, q(t, \cdot))^T N(t, \cdot) \leq \varepsilon |N(t, \cdot)| \end{cases}$$

where, with a slight change of notation, $N_0(m)$ is the outward-pointing unit normal to S_0 at $q_0(m)$. $N(t, m)$ is then an outward-pointing (not necessarily unit) normal to $S(t) = q(t, M) = \varphi(t, S_0)$ at $q(t, m)$. The third equation in the constraints is the time derivative of (4).

2.2. Discrete Approximations. We now assume that surfaces are triangulated, so that parametrizations are replaced by pairs (\mathbf{q}, F) , where $\mathbf{q} \in (\mathbb{R}^3)^n$ is a set of n vertices (where n depends on S) and $F \subset \{1, \dots, n\}^3$ provides the list of indices of vertices that form the triangular faces.

We assume that the surface is oriented, so that an edge that belongs to two faces is oriented in different directions in each face. If $\mathbf{q} = (q_1, \dots, q_n)$, and $f = (i, j, k) \in F$, we let $S(\mathbf{q}, f)$ denote the closed triangle with vertices (q_i, q_j, q_k) . We also let $c(\mathbf{q}, f) = (q_i + q_j + q_k)/3$ be the center of mass and

$$(7) \quad N(\mathbf{q}, f) = \frac{1}{2} (q_j - q_i) \times (q_k - q_i),$$

be the area weighted normal, this expression being invariant by circular permutation of i , j and k ($N(\mathbf{q}, f)$ is oriented according to the normal to the face, with norm equal to the area of the face). From this we define the area-weighted normal at a vertex q_k by

$$(8) \quad N_k(\mathbf{q}, F) = \sum_{f \in F: k \in f} N(\mathbf{q}, f)/3$$

and we let $\mathbf{N}(\mathbf{q}, F) = (N_1(\mathbf{q}, F), \dots, N_n(\mathbf{q}, F))$. Finally, we let $S(\mathbf{q}, F)$ be the associated piecewise-triangular surface, namely

$$S(\mathbf{q}, F) = \bigcup_{f \in F} S(\mathbf{q}, f).$$

To define a discrete version of Problem 1, we introduce state variables $\mathbf{q} = (q_1, \dots, q_n)$ and $\mathbf{N} = (N_1, \dots, N_n)$, initialized with a initial triangulation (\mathbf{q}_0, F_0) . We also assume a

target surface (\mathbf{q}_1, F_1) . The discrete problem will minimize

$$\frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + D_\delta((\mathbf{q}(1), F_0), (\mathbf{q}_1, F_1)),$$

subject to

$$\begin{cases} \mathbf{q}(0, \cdot) = \mathbf{q}_0, \mathbf{N}(0) = \mathbf{N}(q_0, F_0), \\ \partial_t q_k(t) = v(t, q_k(t)), \\ \partial_t N_k(t) = -dv(t, q_k(t))^T N_k(t), \\ v(t, q_k(t)) \cdot N_k(t) \leq \varepsilon |N_k(t)|, k = 1, \dots, n \end{cases}$$

Here, D_δ is a discrete approximation of D in (2) in which the dot product between triangulated surfaces is replaced by

$$(9) \quad \langle (\mathbf{q}, F), (\mathbf{q}', F') \rangle_{\chi, \delta} = \sum_{f \in F} \sum_{f' \in F'} \chi(c(\mathbf{q}, f), c(\mathbf{q}', f')) N(\mathbf{q}, f)^T N(\mathbf{q}', f')$$

(so that χ is approximated by a constant value on each face.)

Given that the end-point cost and the constraints depends on v only through the configurations $\mathbf{q}(t)$, one shows, using standard RKHS reductions, that the optimal v takes the form

$$v(t, \cdot) = \sum_{k=1}^n K(\cdot, q_k(t)) \alpha_k(t)$$

where K is the reproducing kernel of V . Using this, the previous problem reduces to

Problem 2. *Minimize*

$$(10) \quad \frac{1}{2} \int_0^1 \sum_{k,l=1}^n K(q_k(t), q_l(t)) \alpha_k(t)^T \alpha_l(t) dt + D_\delta((\mathbf{q}(1), F_0), (\mathbf{q}_1, F_1))$$

$$(11) \quad \text{subject to} \quad \begin{cases} \mathbf{q}(0) = \mathbf{q}_0, \mathbf{N}(0) = \mathbf{N}(q_0, F_0), \\ \partial_t q_k(t) = \sum_{l=1}^n K(q_k(t), q_l(t)) \alpha_l(t), \\ \partial_t N_k(t) = - \sum_{l=1}^n \partial_1 K(q_k(t), q_l(t)) N_k(t)^T \alpha_l(t), \\ \sum_{l=1}^n K(q_k(t), q_l(t)) \alpha_l(t)^T N_k(t) \leq \varepsilon |N_k(t)|, k = 1, \dots, n \end{cases}$$

However, in the discrete case, it is possible to avoid the introduction of \mathbf{N} as a state variable and solve instead:

Problem 3. *Minimize*

$$(12) \quad \frac{1}{2} \int_0^1 \sum_{k,l=1}^n K(q_k(t), q_l(t)) \alpha_k(t)^T \alpha_l(t) dt + D_\delta((\mathbf{q}(1), F_0), (\mathbf{q}_1, F_1))$$

(1) Given α , compute the associated state \mathbf{q} via $\partial_t \mathbf{q} = K(\mathbf{q})\alpha$ and evaluate the matrix $C(\mathbf{q})$.

(2) Introduce a co-state $\mathbf{p}(t) = (p_1, \dots, p_n) \in (\mathbb{R}^3)^n$, $t \in [0, 1]$, defined by

$$\mathbf{p}(1) = \nabla_q D_\delta((q, F_0), (\mathbf{q}_1, F_1))|_{q=q(1)} \text{ and}$$

$$\partial_t \mathbf{p} = -\nabla_q \left(\mathbf{p}^T K(\mathbf{q})\alpha - \frac{1}{2} \alpha^T K(\mathbf{q})\alpha - \frac{\mu}{2} \left| \left(C(\mathbf{q})\alpha - \varepsilon |\mathbf{N}| - \frac{\lambda}{\mu} \right)^+ \right|^2 \right)$$

(3) Define $\nabla_\alpha F = K(\mathbf{q})(\alpha - \mathbf{p}) + \mu \left(\left(C(\mathbf{q})\alpha - \varepsilon |\mathbf{N}| - \frac{\lambda}{\mu} \right)^+ \right)^T C(\mathbf{q})$.

Remark. Ensuring that the total volume of the surface decreases (instead of enforcing inward motion at every point) can be done very similarly to the full atrophy constraint, using the single constraint $\sum_{k=1}^n v(t, q_k(t))^T N_k(t) \leq \varepsilon$ or, after reduction

$$\sum_{k,l=1}^n K(q_k(t), q_l(t)) \alpha_l(t)^T N_k(t) \leq \varepsilon,$$

where N_k is given by (8). It is important here to use the area-weighted normal, to discretize the continuous constraint $\int_{S(t)} v(t, \cdot)^T N(t, \cdot) d\mathcal{A}_{S(t)} \leq \varepsilon$, which provides the derivative of the total volume with respect to time. This constraint can be rewritten as $\mathbf{1}_n^T C(\mathbf{q}(t))\alpha(t) \leq \varepsilon$, where $\mathbf{1}_n$ is the n -dimensional vector with all coordinates equal to 1. The reformulation of the augmented Lagrangian method for this scalar constraint is straightforward and left to the reader.

4. EXISTENCE OF SOLUTIONS AND CONVERGENCE

The following theorems address two important properties of the proposed method. The first one shows that solutions of the constrained problem within our space of interest exist, and the second one considers the issue of consistency of the discretization approach, showing that under suitable assumptions, the solutions obtained with discrete surfaces converge to solutions of the continuous problem. These results are proved in the appendix.

Theorem 1. *Assume that V is continuously embedded in $C_0^p(\mathbb{R}^3, \mathbb{R}^3)$ for $p \geq 2$ and that the data attachment term D is such that $\varphi \mapsto D(\varphi(S_0), S_1)$, defined over all C^p -diffeomorphisms φ such that $\varphi - id \in C_0^p$, is continuous for the uniform C^p convergence over compact sets. Then Problems 1, 2 and 3 always have an optimal solution $v \in L^2(0, 1; V)$.*

The assumption on D applies to the current norm (2) as soon as $p \geq 1$.

We now introduce some assumptions to address the consistency of discrete approximations. We say that a sequence of triangulations $((\mathbf{q}^n, F^n))$ converges nicely to a surface S of class C^2 if there exists an $n \geq n_0$ such that the following conditions are satisfied.

(i) The vertices of S^n belong to S : $\mathbf{q}^n = (q_1^n, \dots, q_{m_n}^n)$ with $q_k^n \in S$ for $k = 1, \dots, m_n$.

(ii) The triangulations are nested:

$$\{q_1^{n_1}, \dots, q_{m_{n_1}}^{n_1}\} \subset \{q_1^{n_2}, \dots, q_{m_{n_2}}^{n_2}\}.$$

In particular, we can order the collection of all vertices of the triangulations in a countable sequence (q_1, q_2, \dots) and for each q_j , there exists an integer n_j such that q_j is a vertex in \mathbf{q}^n for all $n \geq n_j$.

(iii) The maximum edge length in (\mathbf{q}^n, F^n) goes to 0.

(iv) If $\eta(\mathbf{q}^n, f)$ is the length of the largest edge of a face f , the ratio

$$\frac{\text{area}(S(\mathbf{q}^n, f))}{\eta(\mathbf{q}^n, f)^2},$$

$f \in F^n$ is bounded away from 0 (uniformly in n).

(v) The triangulations are close enough to S to ensure that every point $x \in S(\mathbf{q}^n, F^n)$ has a unique closest point $\xi(x)$ on S , such that ξ is onto.

These hypotheses imply the convergence of normals as follows [34].

(vi) In addition

$$\max_{f \in F^n, p \in S(\mathbf{q}^n, f)} \left| \frac{N(\mathbf{q}^n, f)}{|N(\mathbf{q}^n, f)|} - N(\xi(p)) \right| \rightarrow 0$$

when n goes to infinity, where N is the unit normal to S .

We have the following result.

Proposition 1. *Assume that (\mathbf{q}^n, F^n) converges nicely to S and (\mathbf{q}_1^n, F_1^n) converges nicely to S_1 . Then*

$$D_\delta((\mathbf{q}^n, F^n), (\mathbf{q}_1^n, F_1^n)) \rightarrow D(S, S_1)$$

when n goes to infinity, where D and D_δ are defined in equations (2), (3) (9).

We can now state our consistency theorem.

Theorem 2. *Fix $\varepsilon^\infty > 0$, and assume that V is continuously embedded in $C_0^p(\mathbb{R}^3, \mathbb{R}^3)$ for $p \geq 2$. Let S_0 and S_1 be two C^2 surfaces, and $(S_0^n), (S_1^n)$ be two sequences of triangulations that converge nicely to S_0 and S_1 , respectively. Then, there exists a decreasing sequence of real numbers $\varepsilon^n > 0$ with $\varepsilon^n \rightarrow \varepsilon^\infty$, such that if $v^n \in L^2(0, 1; V)$ solves Problem 2 with $\varepsilon = \varepsilon^n$, initial surface S_0^n and target S_1^n , then the sequence $(v^n)_{n \in \mathbb{N}}$ is bounded in $L^2(0, 1; V)$ and any weak limit point v^* of v^n is a solution to Problem 1 with $\varepsilon = \varepsilon^\infty$.*

5. AFFINE ALIGNMENT

Because the RKHS V is embedded in $C_0^p(\mathbb{R}^3, \mathbb{R}^3)$, vector fields $v \in V$ vanish at infinity. This implies, in particular, that affine transformations do not belong to this Hilbert space, and that the diffeomorphic registration does not incorporate any rigid alignment. If such an alignment is needed, one can include it in the optimal control framework by completing the control with the corresponding vector fields.

Let the registration be computed over $G \times \mathbb{R}^3$, where G is a subgroup of $GL_3(\mathbb{R})$, \times referring to the semi-direct product extending G with translations to obtain affine transformations. Let \mathfrak{g} be the Lie algebra of G , with basis E_1, \dots, E_h . Instead of $v \in V$, we use a control given by $(v, \beta_1, \dots, \beta_h, \tau)$ with $\beta_1, \dots, \beta_h \in \mathbb{R}$ and $\tau \in \mathbb{R}^3$, and the state equation

$$(15) \quad \partial_t q_k(t) = v(t, q(t)) + \sum_{l=1}^h \beta_l(t) E_l q(t) + \tau(t)$$

with associated cost $\frac{1}{2} \int_0^1 \left(\|v(t)\|_V^2 + \sum_{k=1}^h c_k \beta_k(t)^2 + c_0 |\tau(t)|^2 \right) dt$ for some non-negative numbers c_0, c_1, \dots, c_h . The extension of the numerical procedure to this setting is rather straightforward, and not detailed here. Of course, the affine components must be added to v in the atrophy constraint $v \cdot N \leq \varepsilon |N|$.

Consider the special case $G = SO_3$, the rotation group (so that $h = 3$) and assume that Euclidean transformations act as isometries on V , which means that, for all $v \in V$, $R \in SO_3$, $b \in \mathbb{R}^3$, the vector field $\tilde{v} : x \mapsto R^T v(Rx + b)$ also belongs to V and $\|\tilde{v}\|_V = \|v\|_V$. Euclidean-invariant RKHS's of vector fields are extensively described in [19], to which we refer for more details. In the case of scalar kernels $\mathbb{K}(x, y) = K(x, y) Id_{\mathbb{R}^3}$, Euclidean invariance implies that K is a radial kernel, i.e., that $K(x, y) = \gamma(|x - y|^2)$ for some function γ (additional conditions on γ are needed to ensure that K is a positive kernel; see [19]). Assume finally that the end-point cost D is invariant by Euclidean transformation: $D(S, S') = D(RS + b, RS' + b)$. In this case, the optimal control problem (using $c_0 = \dots = c_3 = 0$) that minimizes $\frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + D(S(1), S_1)$ in v, β, τ , subject to

$$(16) \quad \begin{cases} q(0) = q_0, \\ \partial_t q(t) = v(t, q(t)) + \sum_{l=1}^3 \beta_l(t) E_l q(t) + \tau(t) \\ \partial_t N(t, \cdot) = - \left(dv(t, q(t, \cdot)) + \sum_{l=1}^3 \beta_l(t) E_l \right)^T N(t, \cdot) \\ \left(v(t, q(t)) + \sum_{l=1}^3 \beta_l(t) E_l q(t) + \tau(t) \right)^T N(q(t)) \leq \varepsilon |N(q(t))| \end{cases}$$

is equivalent to minimizing $\frac{1}{2} \int_0^1 \|\tilde{v}(t)\|_V^2 dt + D(\tilde{S}(1), R_1 S_1 + b_1)$ in \tilde{v}, R_1, b_1 , subject to

$$(17) \quad \begin{cases} \tilde{q}(0) = q_0, \\ \partial_t \tilde{q}(t) = \tilde{v}(t, \tilde{q}(t)) \\ \partial_t \tilde{N}(t, \cdot) = - d\tilde{v}(t, \tilde{q}(t, \cdot))^T \tilde{N}(t, \cdot) \\ \tilde{v}(t, \tilde{q}(t))^T \tilde{N}(t, \tilde{q}(t)) \leq \varepsilon |\tilde{N}(t, \tilde{q}(t))| \end{cases}$$

via the change of variable $q(t) = R(t)\tilde{q}(t) + b(t)$, $v(t, x) = R(t)^{-1}\tilde{v}(t, Rx + b)$, $N(t) = R(t)\tilde{N}(t)$, with $\partial_t R = \sum_{l=1}^3 \beta_l(t) E_l R$, $\partial_t b = \sum_{l=1}^3 \beta_l(t) E_l b(t) + \tau(t)$, $R_1 = R(1)^{-1}$ and $b_1 = -R(1)^{-1}b(1)$. In other words, Euclidean registration via optimal control using (16) is equivalent to the original atrophy-constrained LDDMM optimizing its target over an orbit under the action of rigid transformations.

Note that $c_0 = \dots = c_h = 0$ should not be used with general affine transformations when non-compact components of the affine group are added to rotations. Intuitively, this would allow small deformations to be scaled up to larger ones at zero cost, and one can conjecture that the associated optimal control problem has no solutions, and admits minimizing sequences of controls with vanishing cost at the limit. The equivalence with a problem in which the target is allowed to vary over its orbit via affine transformations is not true either for groups larger than the Euclidean one, essentially because invariant kernels do not exist in such cases.

Some attention should be paid to the discretization in time t , in particular in the rigid case. In our implementation, we use a simple Euler scheme to discretize the equation $\partial_t q = v(t, q)$, i.e., we take $q(t + \delta t) = q(t) + \delta t v(t, q(t))$. When using rigid registration, however, we take, with $A(t) = \sum_{l=1}^3 \beta_l(t) E_l$ a skew-symmetric matrix

$$q(t + \delta t) = e^{\delta t A(t)} q(t) + \delta t v(t, q(t)) + \delta t \tau(t)$$

to discretize $\partial_t q = v(t, q) + Aq + \tau$, with the explicit expression $e^U = \text{Id} + \frac{\sin c_U}{c_U} U + \frac{1 - \cos c_U}{c_U^2} U^2$, $c_u = \sqrt{-\text{tr}(U^2)}$ for a 3 by 3 skew-symmetric matrix U . This ensures that the rigid registration remains a displacement, even for large values of the coefficients β_l , which are made possible by the absence of cost on this transformation.

6. EXTENSION TO TIME SERIES

Longitudinal studies generally include more than one follow-up scan, and we now extend our algorithm to the case of multiple targets observed at successive times. We still have a baseline surface S_0 and now assume p follow-up surfaces S_1, \dots, S_p . A direct generalization of the previous approach replaces equation (5) in Problem 1 by

$$(18) \quad \int_0^1 \|v(t)\|_V^2 dt + \sum_{k=1}^n D(S(t_k), S_k)$$

where $0 < t_1 \leq \dots \leq t_n = 1$ are attachment times, which can be fixed a priori (e.g., $t_k = k/n$) or also optimized. The expression in (18) should be minimized, starting from a baseline S_0 , with the same constraints as in the single-target case.

As pointed out in [38, 39, 26, 3], this formulation of longitudinal registration (with or without atrophy) gives a special role to the baseline image and creates a bias in the way the information is obtained by breaking the symmetry within the acquisition and segmentation protocols that provided the images or surfaces. We address this by adding an extra error term, $D(S(t_0), S_0)$ with $t_0 = 0$, making the sum start at $k = 0$ in (18). The trajectory baseline, $S(0)$ becomes a new variable to estimate. To ensure that the topology of the estimated trajectory remains consistent with the data (as was originally enforced by $S(0) = S_0$), we assume that $S(0)$ is a diffeomorphic transformation of another surface (a template) of known topology. Letting \bar{S}_0 denote the template, this leads to the

minimization of

$$(19) \quad \lambda \int_0^1 \|\bar{v}(t)\|_V^2 dt + \int_0^1 \|v(t)\|_V^2 dt + \sum_{k=0}^n D(S(t_k), S_k)$$

where λ is a regularization parameter and \bar{v} controls a smooth deformation between the template \bar{S}_0 and the initial point of the trajectory, $S(0)$. More precisely, we let $S(0) = \bar{S}(1)$, where $\bar{S}(t)$ is a template-to-baseline trajectory satisfying $\bar{S}(0) = \bar{S}_0$ and $\partial_t \bar{S} = \bar{v}(t, \bar{S}(t))$. The cost function is then minimized in the pair of chained controls (\bar{v}, v) . While atrophy constraints remain applicable to v , no such constraint is enforced on \bar{v} . This algorithm has as been implemented in Python, with examples of trajectories provided in Figure 6 and Figure 7.

7. EXPERIMENTAL RESULTS

7.1. Comparing two Shapes. Fig. 1 provides four examples of segmented hippocampus surfaces taken from the BIOCARD longitudinal study [33, 52]. For each subject, a baseline image is compared to a follow-up taken approximately eight years later. The color map is proportional to the total normal displacement during the estimated deformation. The deformation patterns with and without the atrophy constraints generally agree on where atrophy lies, but expansion is ignored by the latter as expected.

7.2. Time Series. We now present experiments that compare time series, as described in section 6. Our shapes are once again extracted from the BIOCARD dataset and correspond to a selection of subjects with one baseline image and four follow-up. We ran our registration methods with and without the atrophy constraint, with results on three sequences provided in Figures 2 and 3. It is interesting to notice when comparing the estimated sequences with and without atrophy that the “monotonicity” constraint seems to have a regularization effect, with an evolution pattern fairly visible in the atrophy-constrained case, but less apparent when constraints are relaxed. This is particularly apparent in the example shown in the second rows of Figures 2 and 3. We also computed atrophy/expansion rates associated to these sequences, using linear regression separately at each vertex. These are summarized in Figure 4.

8. DISCUSSION

In this paper, we have introduced a new approach targeting atrophy in the registration of longitudinal data, in view of applications to the analysis of neuro-degenerative diseases. We have formulated the problem via constrained optimal control, and proposed an algorithm that combined the adjoint method in optimal with augmented lagrangian methods. We also provided theoretical results on the existence of solutions for the problems, and on the consistency of our numerical scheme for approaching the solution.

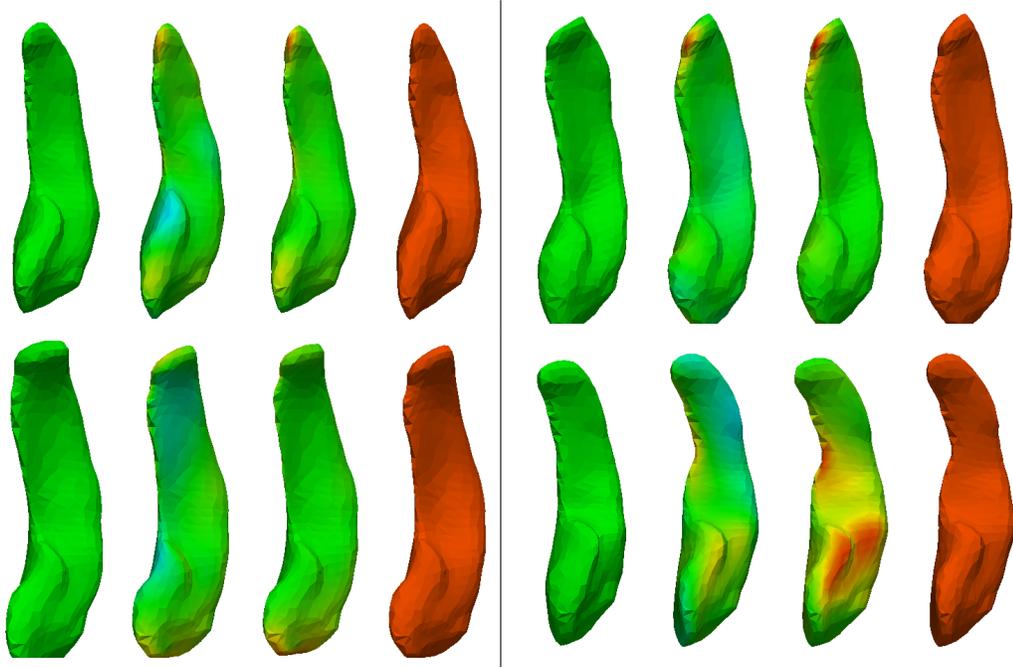


FIGURE 1. Four registration examples from biocard subjects. From left to right: baseline; LDDMM surface registration (no constraint); registration with atrophy constraint; target (eight years later). Colors on registered surfaces reflect normal displacements, from red (atrophy) to blue (expansion), with green being neutral. Color maps are on the same symmetrical scale from -2.9mm to $+2.9\text{mm}$.

ACKNOWLEDGEMENT

This paper is supported in part by grants from the National Institutes of Health: U01-AG03365, P50-AG005146, R01-EB017638, R01-NS084957, P41-EB015909, the Office of Naval Research ONR-90048203, and the National Science Foundation ACI-1053575 for the Computational Anatomy Gateway via Extreme Science and Engineering Discovery Environment (XSEDE).

APPENDIX A. PROOF OF THEOREM 1

This theorem is proved along the same lines as similar statements addressing the existence of solutions for LDDMM problems [43, 16, 50, 51, 2]. We only provide a proof for Problem 1, since Problems 2 and 3 are addressed in a very similar way.

Let v^n be a minimizing sequence. The boundedness of v^n in $L^2(0, 1; V)$ implies that (extracting a subsequence if needed) v^n converges weakly in this Hilbert space to a limit v (with a norm in $L^2(0, 1; V)$ no larger than the liminf of the norms of v^n). This, in turn, implies that the associated flows of diffeomorphisms φ^n (associated to v^n) and their

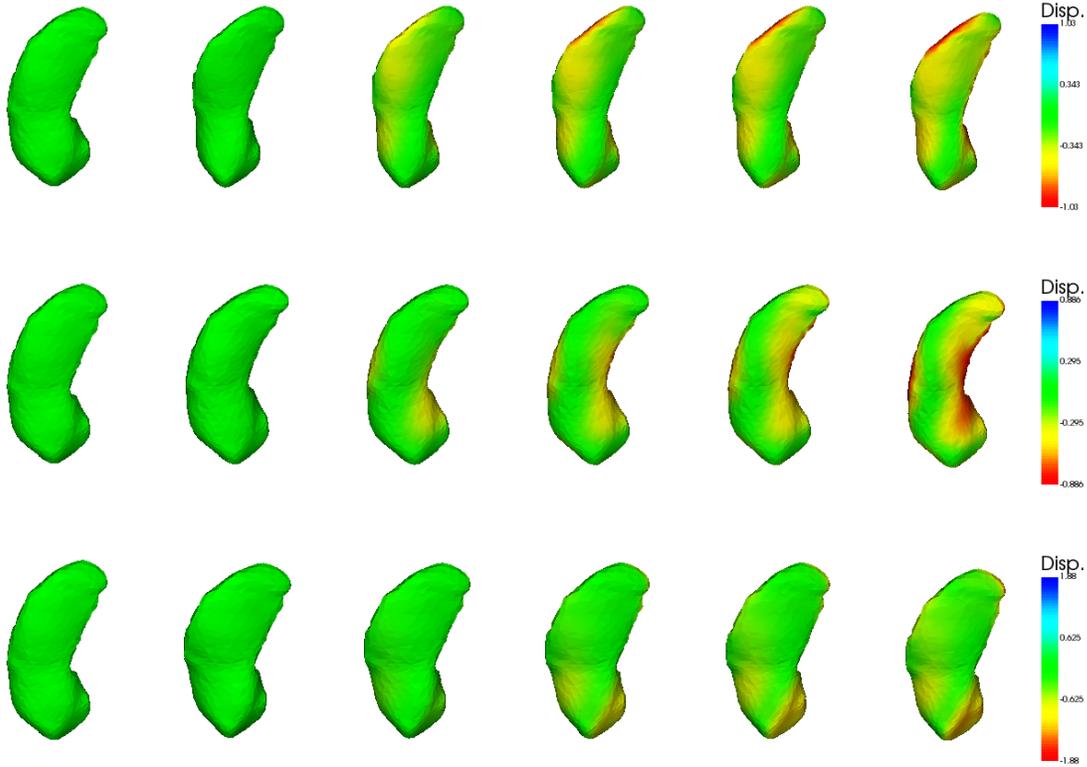


FIGURE 2. Estimated trajectories via LDDMM time series with atrophy constraints. Each row is a different subject. The first shape (left) is the template, which is fixed. Subsequent shapes provide baseline and follow-up for each subject. The color mapped on the surface is proportional to the normal displacement during the evolution starting with the baseline. Warm colors (yellow-red) correspond to atrophy and cold colors (blue) to expansion, while green is neutral. As a result of the constraint, no expansion is observed in these sequences. The maximum computed atrophy is 1.03mm, 0.89mm and 1.88mm respectively, and the scale varies between these values and their negative.

first p spatial derivatives converge uniformly (in time and space) over compact sets to φ (associated to v) and its first p derivatives [16, 50]. One gets from this that $D(\varphi^n(1, S_0), S_1)$ converges to $D(\varphi(1, S_0), S_1)$.

Letting (q^n, N^n) denote the state in Problem 1 associated with the control v^n , we have $q^n(t) = \varphi^n(t, q_0)$, $N^n(t) = d\varphi^n(t, q_0)^{-T} N_0$ therefore converging to $q(t) = \varphi(t, q_0)$ and $N(t) = d\varphi(t, q_0)^{-T} N_0$ when $n \rightarrow \infty$. This and the upper bound of the $L^2(0, 1; V)$ -norm of v implies that the cost function is minimized at v .

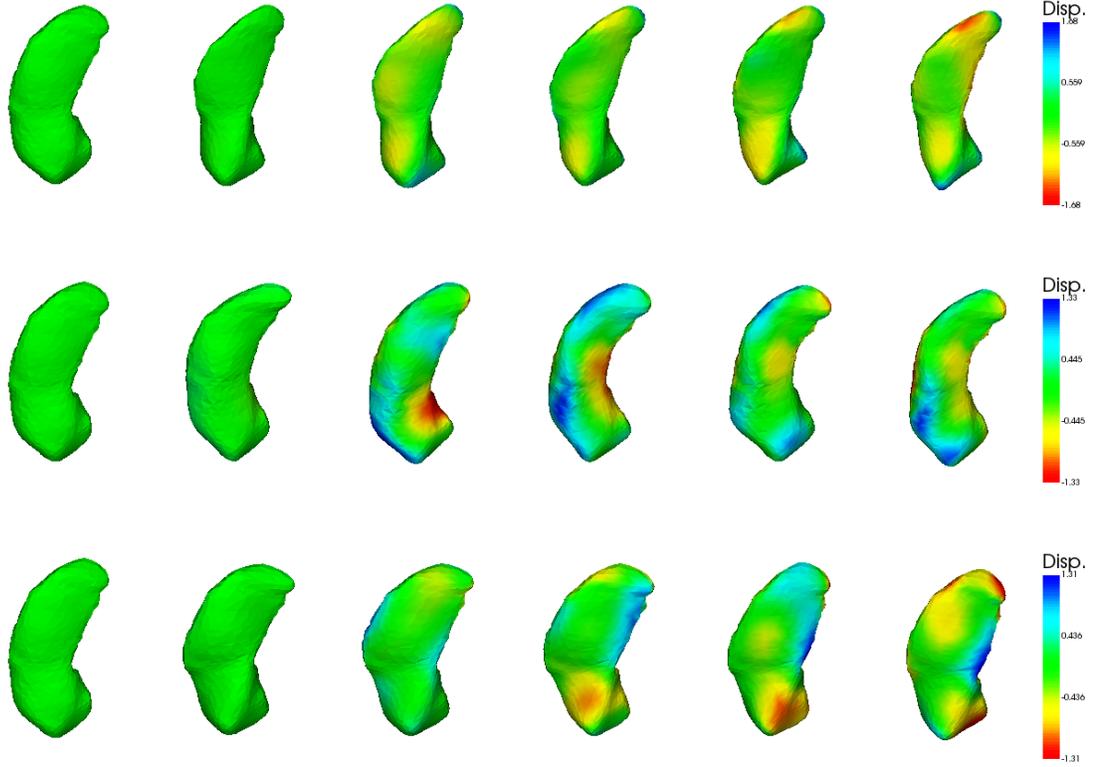


FIGURE 3. The display is similar to the one described in Figure 2, with the same sequences registered this time without the atrophy constraint. Several expansion zones (blue) now appear, with estimated normal expansion around 1.3mm on each subject. The atrophy levels and pattern are similar to the one observed in Figure 2 even though they look reduced because of the expansion of the color maps, who range, here, between 1.68mm, 1.33mm ,1.31mm (from top to bottom) and their negative.

We now show that $v(t, q(t, x))^T N(t, x) \leq \varepsilon |N(t, x)|$ for all $x \in M$ and almost all $t \in [0, 1]$. Fixing x , and writing

$$\begin{aligned} v^n(t, q^n(t, x))^T N^n(t, x) &= v^n(t, q(t, x))^T N(t, x) \\ &\quad + (v^n(t, q^n(t, x)) - v^n(t, q(t, x)))^T N(t, x) + v^n(t, q^n(t, x))^T (N^n(t, x) - N(t, x)) \end{aligned}$$

one easily deduces from the facts that v^n converges weakly to v , dv^n is uniformly bounded, q^n and N^n converge uniformly to q and N , that $v^n(t, q^n(t, x))^T N^n(t, x)$ converge weakly to $v(t, q(t, x))^T N(t, x)$ in $L^2(0, 1; \mathbb{R})$. This implies that $v(t, q(t, x))^T N(t, x) \leq \varepsilon |N(t, x)|$ for almost all t (the set of non-positive a.e functions is a closed convex set in L^2 and therefore

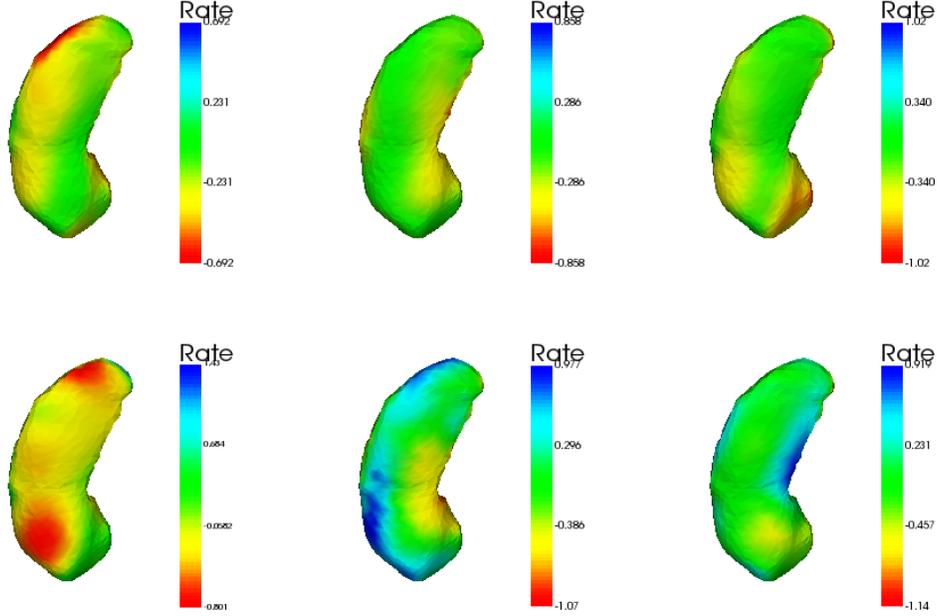


FIGURE 4. Rate of atrophy/expansion estimated from previous trajectories. Color maps are proportional to millimeters per year in normal displacement, red for atrophy and blue for expansion. The first row provides the rates obtained under atrophy constraint, and the second row corresponds to an estimation without constraint. The color maps are scaled over symmetric intervals of radii 0.692, 0.858, 1.02 in the first row. The intervals for the second row are $[-0.801, 1.43]$, $[-1.07, 0.977]$ and $[-1.14, 0.919]$.

also weakly closed). By considering a countable dense set of x 's, and using the fact that $x \mapsto v(t, q(t, x))^T N(t, x)$ is continuous, we find that $v(t, q(t, x))^T N(t, x) \leq \varepsilon |N(t, x)|$ for all $x \in M$ and almost all $t \in [0, 1]$.

APPENDIX B. PROOF OF THEOREM 2

We first prove Proposition 1. Let ξ^n and ξ_1^n denote the closest-point maps from $S(\mathbf{q}^n, F^n)$ to S and $S(\mathbf{q}_1^n, F_1^n)$ to S_1 . For $f \in F^n$, let $S^n(f) = \xi^n(S(\mathbf{q}^n, f)) \subset F^n$ be the image of the triangle associated to f by ξ^n . Define similarly $S_1^n(f) \subset F_1^n$ for $f \in F_1^n$. Let $c^n(f) = \xi^n(c(\mathbf{q}^n, f))$, $c_1^n(f) = \xi_1^n(c(\mathbf{q}_1^n, f))$ be the images of the face centers on the surface. Since the surfaces are smooth and compact and the maximal edge length of the triangulations goes to 0, we have

$$\max_{f \in F^n} \max_{s \in f} |N(s) - N(c^n(f))| \rightarrow 0, \quad \max_{f \in F_1^n} \max_{s \in f} |N_1(s) - N_1(c_1^n(f))| \rightarrow 0,$$

and

$$\max_{f \in F^n, f' \in F_1^n} \max_{s \in f, s' \in f'} |\chi(s, s') - \chi(c^n(f), c_1^n(f'))| \rightarrow 0$$

when $n \rightarrow \infty$. This implies that

$$\begin{aligned} \langle S, S_1 \rangle_\chi &= \sum_{f \in S} \sum_{f' \in S_1} \int_{S^n(f) \times S_1^n(f')} \chi(s, s') N(s)^T N_1(s') d\mathcal{A}_S(s) d\mathcal{A}_{S_1}(s') \\ (20) \quad &= \sum_{f \in S} \sum_{f' \in S_1} \chi(c^n(f), c_1^n(f')) N(c^n(f))^T N_1(c_1^n(f')) \text{area}(S^n(f)) \text{area}(S_1^n(f')) + \eta_n \end{aligned}$$

where $\eta_n \rightarrow 0$. Note that the first identity uses the assumption that ξ and ξ_1 are bijective. Using results from [34], we can deduce from the convergence assumptions that

$$\max_{f \in F^n} |c^n(f) - c(\mathbf{q}_n, f)|, \quad \max_{f \in F^n} \left| N(c^n(f)) - \frac{N(\mathbf{q}_n, f)}{|N(\mathbf{q}_n, f)|} \right|$$

and

$$\max_{f \in F^n} |\text{area}(S^n(f)) / \text{area}(S(\mathbf{q}^n, F^n)) - 1|$$

all converge to 0, with obviously an identical statement for S_1 . This implies that the right-hand side in (20) can be replaced by $\langle (\mathbf{q}^n, F^n), (\mathbf{q}_1^n, F_1^n) \rangle_{\chi, \delta} + \eta'_n$ with $\eta'_n \rightarrow 0$. This and similar arguments for the other dot products defining D and D_δ imply Proposition 1.

Remark. It is easy to check that, for some positive constant C and for any diffeomorphism φ of \mathbb{R}^3 ,

$$(21) \quad |D(\varphi(S_0), S_1) - D(\varphi(\mathbf{q}^n), \mathbf{q}_1)| \leq C |D(S_0, S_1) - D(\varphi(\mathbf{q}^n), \mathbf{q}_1)| \max_{x \in S_0} (|d\varphi(x)|, |d\varphi(x)^{-1}|).$$

We now prove Theorem 2. Let $(S_0^n = (\mathbf{q}_0^n, F_0^n))$ and $(S_1^n = (\mathbf{q}_1^n, F_1^n))$ be sequences of triangulations converging nicely to smooth surfaces S_0 and S_1 . We let $E(v)$ denote the cost for the continuous problem 1 (Equation (5)), and by $E^n(v)$ the one associated to the discrete problem at step n .

Using the same approach as the one used to prove Theorem 1, it is not too difficult to prove that, if v^n is a sequence of solutions of Problem 2 for the n -th triangulation and $\varepsilon = \varepsilon^n$, then (v^n) is bounded in $L^2(0, 1; V)$, and any weak limit point v^* of this sequence satisfies the constraints of Problem 1 with $E(v^*)$ no larger than the limit inferior of the sequence $E(v^n)$. This is done in steps 4 and 5 below. For the proof to be complete, one still need to show that v^* is an optimal solution for Problem 1. A natural approach for this would be to show that any solution, say v_c^* , of the continuous problem is admissible for the relaxed discrete problems, i.e., that it satisfies its constraints (because, if this is proved, then $E^n(v_c^*)$ is larger than $E^n(v^n)$, while converging to $E(v_c^*)$, so that the latter cannot be smaller than $E(v^*)$). The difficulty is that the constraints involve the values of $v_c^*(t, \cdot)$ along the trajectories, and cannot be controlled without a uniform bound on the supremum norm of $v_c^*(t, \cdot)$, which could be achieved, in our case, via a uniform bound on $\|v_c^*(t, \cdot)\|_V$. Formally, one could use optimality conditions for v_c^* (a Pontryagin maximum

principle) to obtain such a bound, but such principles with infinite dimensional constraints are rare and difficult to prove. To work around this, our proof considers a sequence v_c^n of solutions of the continuous problem in which constraints are enforced only over the discrete sets defined by the corresponding triangulation. We use the maximum principle, which is valid in this case, to prove that $\|v_c^n(t, \cdot)\|_V$ is bounded (uniformly in t and n) and that v_c^n is an admissible solution for the discrete problem, with a suitable relaxation of the constraints (steps 1 to 3 below). We can then wrap up the proof by noticing that $E(v_c^n)$ is necessarily smaller than the optimal cost of the continuous problem (fewer constraints), but also larger (up to a negligible difference) than the optimal discrete cost. We now pass to the detailed proof.

Step 1: Reduction to continuous shape with finite number of constraints. Fix $\varepsilon_\infty > 0$, and denote by E^∞ the minimal value of the cost functional in Problem 1 with initial surface S_0 and $\varepsilon = \varepsilon^\infty$. Let

$$B_n = \{s_1, \dots, s_{k_n}\} \subset M$$

be the increasing sequence of subsets in the parametric space M such that $q_0(s_k) = q_{0,k}^n$, with k_n the number of vertices in the n -th triangulation. Note that the s_k are two by two distincts. For each n , consider a time-dependent vector field v_c^n with flow φ_c^n that minimizes

$$E(v) := \frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + D(q(1, M), S_1)$$

subject to

$$\begin{cases} \mathbf{q}(0, \cdot) = \mathbf{q}_0, N(0, \cdot) = N_0, \\ \partial_t q(t, \cdot) = v(t, q(t, \cdot)), \\ \partial_t N(t, \cdot) = -dv(t, q(t, \cdot))^T N(t, \cdot), \\ v(t, q(t, s))^T N(t, s) \leq \varepsilon_\infty |N(t, s)|, \quad s \in B_n. \end{cases}$$

Let $q_c^n(t) = \varphi_c^n(t) \circ q_0$, $S_c^n(t) = \varphi_c^n(t, S_0)$ and $N_c^n(t) = d\varphi_c^n(t) \circ q_0^{-T} N_0$ (the existence of such a v_c^n is proved just as described in Appendix A). Note that any solution v with flow φ of Problem 1 also satisfies these constraints, so that

$$\begin{aligned} E^\infty &= \frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + D(\varphi(1, S_1)) \\ &\geq \frac{1}{2} \int_0^1 \|v_c^n(t)\|_V^2 dt + D(q_c^n(1, M), S_1) = E(v_c^n). \end{aligned}$$

Step 2: Pontryagin maximum principle and uniform boundedness of v_c^n . The constraints $v(t, q(t, s))^T N(t, s) \leq \varepsilon_\infty |N(t, s)|$, $s \in B_n$ are regular constraints (and are in finite number), that is, for each parametrized C^1 -embedded surface $q \in \text{Emb}^1(M, \mathbb{R}^3)$ (where $\text{Emb}^1(M, \mathbb{R}^3)$, the space of C^1 embeddings of M in \mathbb{R}^3 is an open subset of $\mathcal{C}^1(M, \mathbb{R}^3)$) and any vector field $N \in \mathcal{C}^0(M, \mathbb{R}^3 \setminus \{0\})$, the mapping

$$v \in V \mapsto (v(q(s_1))^T N(s_1), \dots, v(q(s_{k_n}))^T N(s_{k_n})) \in \mathbb{R}^{k_n}$$

is surjective since the value of v can be chosen freely at the points $q(s_1), \dots, q(s_{n_k})$. The Pontryagin maximum principle can therefore be applied as follows (see [25, Theorems 4.1 and 4.2] for the general statement). Define the Hamiltonian of the constrained problem

$$H_c^n : \text{Emb}^1(M, \mathbb{R}^3) \times \mathcal{C}^0(M, \mathbb{R}^3 \setminus \{0\}) \times \mathcal{C}^1(M, \mathbb{R}^3)^* \times \mathcal{C}^0(M, \mathbb{R}^3)^* \times V \times \mathbb{R}_+^{k_n} \longrightarrow \mathbb{R}$$

by

$$\begin{aligned} H_c^n(q, N, p, \nu, v, \lambda) &= (p \mid v(q(\cdot))) - (\nu \mid dv(q(\cdot))^T N(\cdot)) - \frac{1}{2} \|v\|_V^2 \\ &\quad - \sum_{k=1}^{k_n} \lambda_k (v(q(s_k))^T N(s_k) - \varepsilon^\infty |N(s_k)|), \end{aligned}$$

in which the notation $(\mu \mid w)$ denotes the evaluation of a linear form μ at a vector w , i.e., $(\mu \mid w) = \mu(w)$. Also, define the convex set Γ_c^n by

$$\Gamma_c^n = \{v \in V, \forall k = 1, \dots, k_n, v(q_c^n(t, s_k))^T N_c^n(t, s_k) \leq \varepsilon^\infty |N_c^n(t, s_k)|\}.$$

Then because v_c^n is optimal, the Pontryagin maximum principle states the following.

Lemma 1. *There exists $p_c^n : [0, 1] \rightarrow \mathcal{C}^1(M, \mathbb{R}^3)^*$ and $\nu_c^n : [0, 1] \rightarrow \mathcal{C}^0(M, \mathbb{R}^3)^*$ of Sobolev class H^1 , and $\lambda_c^n \in L^2(0, 1; \mathbb{R}_+^{k_n})$ such that $p_c^n(1) = -\nabla_q D(q_c^n(1), S_1)$, $\nu_c^n(1) = 0$, and, for almost every t , and every $k = 1, \dots, k_n$,*

$$(22) \quad \dot{p}_c^n(t) = -\partial_q H_c^n(t),$$

$$(23) \quad \dot{\nu}_c^n(t) = -\partial_N H_c^n(t),$$

$$(24) \quad 0 = \lambda_k(t) (v_c^n(t, q_c^n(t, s_k))^T N_c^n(t, s_k) - \varepsilon^\infty |N_c^n(t, s_k)|), \quad k = 1, \dots, k_n$$

$$(25) \quad H_c^n(t) = \max_{v \in \Gamma_c^n} H_c^n(q_c^n(t), N_c^n(t), p_c^n(t), \nu_c^n(t), v, \lambda_c^n(t))$$

$$(26) \quad = \max_{v \in \Gamma_c^n} (p_c^n(t) \mid v(q_c^n(t, \cdot))) - (\nu_c^n(t) \mid dv(q_c^n(t, \cdot))^T N_c^n(t, \cdot)) - \frac{1}{2} \|v\|_V^2$$

$$(27) \quad H_c^n(t) = H_c^n(1) = \max_{v \in \Gamma_c^n} \int_{S_c^n(1)} -\nabla_q D(q_c^n(1), S_1)(s)^T v(s) ds - \frac{1}{2} \|v\|_V^2.$$

Here, we wrote $H_c^n(t) = H_c^n(q_c^n(t), N_c^n(t), p_c^n(t), \nu_c^n(t), v_c^n(t), \lambda_c^n(t))$ for short.

For almost every t , the set Γ_c^n is convex and contains 0. The maximized function in (26) is of the form $(P \mid v) - \frac{1}{2} \|v\|_V^2$ for some $P \in V^*$. We then have that, for any $v \in \Gamma_c^n$, $(P \mid v - v_c^n(t)) - \langle v_c^n(t), v - v_c^n(t) \rangle_V \leq 0$. Applying this to $v = 0$ yields

$$\begin{aligned} &-(P \mid v_c^n(t)) + \|v_c^n(t)\|_V^2 \leq 0 \\ \implies &\|v_c^n(t)\|_V^2 \leq (P \mid v_c^n(t)) \\ \implies &\frac{1}{2} \|v_c^n(t)\|_V^2 \leq (P \mid v_c^n(t)) - \frac{1}{2} \|v_c^n(t)\|_V^2 = H_c^n(t) = H_c^n(1) \end{aligned}$$

by (27). Hence, for almost every t , we get $\|v_c^n(t)\|_V \leq \sqrt{2H_c^n(1)}$. But,

$$\begin{aligned} H_c^n(1) &\leq \max_{v \in V} \int_{S_c^n(1)} -\nabla_q D(q_c^n(1), S_1)(s) \cdot v(s) ds - \frac{1}{2} \|v\|_V^2 \\ &= \frac{1}{2} \iint_{S_c^n(1) \times S_c^n(1)} K(s, s') \nabla_q D(q_c^n(1), S_1)(s)^T \nabla_q D(q_c^n(1), S_1)(s') ds' ds \quad . \end{aligned}$$

Now, note that the sequence (v_c^n) is bounded in $L^2(0, 1; V)$ by $\sqrt{2D(S_0, S_1)}$, so that the family $(\varphi_c^n(t))_{n \in \mathbb{N}, t \in [0, 1]}$ is bounded in $C^p(\mathbb{R}^3, \mathbb{R}^3)$ (just use Gronwall's lemma). Looking at the formula for D , we easily deduce that $H_c^n(1)$ is bounded independently of n . In particular, for some positive constant C , $\|v_c^n(t)\|_V \leq C$ for every n and almost every t in $[0, 1]$.

Step 3: Candidate for ε^n . Let us now return to our sequence of triangulations \mathbf{q}_0^n . Using Gronwall's estimates on the ODE $\partial_t \varphi_c^n = v_c^n(t, \varphi)$ one can estimate the differences

$$\left| v_c^n(t, \varphi_c^n(t, \mathbf{q}_{0,k}^n))^T \left(\frac{N_c^n(t, s_k)}{|N_c^n(t, s_k)|} - \frac{(d\varphi_c^n)^{-T}(t, \mathbf{q}_{0,k}^n) N_{0,k}^n}{|(d\varphi_c^n)^{-T}(t, \mathbf{q}_{0,k}^n) N_{0,k}^n|} \right) \right| \leq \eta_n''$$

and

$$\left| v_c^n(t, \varphi_c^n(t, \mathbf{q}_{0,k}^n))^T \frac{(d\varphi_c^n)^{-T}(t, \mathbf{q}_{0,k}^n) N_{0,k}^n}{|(d\varphi_c^n)^{-T}(t, \mathbf{q}_{0,k}^n) N_{0,k}^n|} - v_c^n(t, \varphi_c^n(t, \mathbf{q}_{0,k}^n))^T N_k(\varphi_c^n(\mathbf{q}_0^n), F_0) \right| \leq \eta_n'''$$

with $\eta_n'', \eta_n''' \rightarrow 0$ when $n \rightarrow \infty$. Therefore, v_c^n satisfies the constraints of Problem 2 with $\varepsilon = \varepsilon^\infty + \eta_n'' + \eta_n''' =: \varepsilon^n$ which converges to ε^∞ as n goes to infinity. Moreover, we can choose $(\eta_n'' + \eta_n''')$ so that (ε^n) is decreasing.

Step 4: Optimality of weak limit points. Now, let us return to v^n , the sequence of solutions of Problem 2 with $\varepsilon = \varepsilon^n$ for each triangulation, φ^n the flow of diffeomorphisms associated to v^n and $\mathbf{q}^n = \varphi^n(\mathbf{q}_0^n)$. Since v^n is optimal,

$$(28) \quad E^n := \frac{1}{2} \int_0^1 \|v^n(t)\|_V^2 dt + D_\delta((\mathbf{q}^n(1), F_0), (\mathbf{q}_1^n, F_1)) \leq D_\delta((\mathbf{q}_0^n, F_0), (\mathbf{q}_1^n, F_1)),$$

as the upper-bound is the value of the objective function at $v = 0$. Since the triangulations converge nicely, Proposition 1 implies that this upper-bound can be bounded independently of n , so that $(v^n)_{n \in \mathbb{N}}$ is bounded in $L^2(0, 1; V)$ and therefore contained in a weak compact subset. On the other hand, as v_c^n satisfies the discrete constraints, we get

$$\begin{aligned} E^n &\leq \frac{1}{2} \int_0^1 \|v_c^n(t)\|_V^2 dt + D_\delta(\varphi_c^n(1, \mathbf{q}_0^n), F_0), (\mathbf{q}_1^n, F_1) \\ &= E_c^n + D_\delta((\varphi_c^n(1, \mathbf{q}_0^n), F_0), (\mathbf{q}_1^n, F_1)) - D(S_c^n(1), S_1) \\ &\leq E^\infty + D_\delta((\varphi_c^n(1, \mathbf{q}_0^n), F_0), (\mathbf{q}_1^n, F_1)) - D(S_c^n(1), S_1). \end{aligned}$$

Using (21), the nice convergence of the triangulations and the boundedness of $(\varphi_c^n(1))$ and its derivatives, we see that, as n goes to infinity

$$D_\delta((\varphi_c^n(1, \mathbf{q}_0^n), F_0), (\mathbf{q}_1^n, F_1)) - D(S_c^n(1), S_1) \rightarrow 0,$$

so that the right-hand side of this equation goes to E^∞ , and $\liminf_{n \rightarrow \infty} (E^n) \leq E^\infty$.

Now, let v^* be a weak limit of (v^n) , therefore the limit of a subsequence that we still denote $(v^n)_{n \in \mathbb{N}}$. Let φ^* denote the flow associated to v^* and $\mathbf{q}^*(t) = \varphi(t, \mathbf{q}_0)$. Let also $S^*(t) = \varphi^*(t, S_0)$. As mentioned in the proof of Theorem 1, the weak convergence of (v^n) to v^* implies that $\int_0^1 \|v^*(t)\|_V^2 dt \leq \liminf_{n \rightarrow \infty} \int_0^1 \|v^n(t)\|_V^2 dt$, and that φ^n and its first $p-1$ spatial derivatives converges to φ^* uniformly on compact sets. Using the fact that this property also holds for the inverse of φ^n , it is then easy to show that $(\mathbf{q}^n(t), F_0)$ converges nicely to $S^*(t)$ for all $t \in [0, 1]$, uniformly in t . In particular $D_\delta((\mathbf{q}^n(1), F_0), (\mathbf{q}_1^n, F_1))$ converges to $D(S^*(1), S_1)$.

Therefore,

$$\frac{1}{2} \int_0^1 \|v^*(t)\|_V^2 dt + D(S^*(1), S_1) \leq \liminf (E^n) \leq E^\infty.$$

Therefore, all that is left is to show that v^* is an admissible control for Problem 1 with $\varepsilon = \varepsilon_\infty$.

Step 5: admissibility of v^ .* Let $\mathbf{x}_0 = (x_{0,1}, x_{0,2}, \dots)$ be the union of all the points in \mathbf{q}_0^n for $n \geq 0$. By assumption, for every $k \geq 1$, there exists an $n_k > 0$ such that for all $n \geq n_k$, there exists $j = j_{n,k}$ such that $x_{0,k} = q_{0,j}^n$. Letting $\mathbf{x}^n(t) = \varphi^n(t, \mathbf{x}_0)$ and $\mathbf{x}^*(t) = \varphi^*(t, \mathbf{x}_0)$ we can prove, as done in Theorem 1, that $v^n(t, x_k^n(t))^T N^n(t, x_k^n(t)) / |N_k^n(t)|$ converges weakly to $v^*(t, x_k^*(t))^T N^*(t, x_k^*(t))$ for almost every t and every k , where $N^n(t, x_k^n(t)) = N_{j_{k,n}}(\mathbf{q}^n(t), F_0)$ and N^* is the unit normal to S^* .

But for every integers $n \geq m$, we have that $t \mapsto v^n(t, x_k^n(t))^T N^n(t, x_k^n(t)) / |N_k^n(t)|$ belongs to the closed convex subset $L^2(0, 1; (-\infty, \varepsilon^m])$, which is also weakly closed. Hence, $v^*(t, x_k^*(t))^T N^*(t, x_k^*(t))$ belongs to

$$\bigcap_{n \geq 0} L^2(0, 1; (-\infty, \varepsilon^n]) = L^2(0, 1; (-\infty, \varepsilon^\infty]).$$

We conclude that the constraints are satisfied by v^* for almost every t and all $x_k^*(t)$. This in turn implies that they are satisfied for all $x \in S^*$ by continuity in x of $v^*(t, \cdot)$ and $N^*(t, \cdot)$ since the sequence x_k^* is dense in S^* .

REFERENCES

- [1] Liana G Apostolova, Paul M Thompson, Amity E Green, Kristy S Hwang, Charleen Zoumalan, Clifford R Jack, Danielle J Harvey, Ronald C Petersen, Leon J Thal, Paul S Aisen, et al. 3d comparison of low, intermediate, and advanced hippocampal atrophy in mci. *Human brain mapping*, 31(5):786–797, 2010.
- [2] Sylvain Arguillere, Emmanuel Trélat, Alain Trounev, and Laurent Younes. Shape deformation analysis from the optimal control viewpoint. *To Appear in Journal de Mathématiques Pures et Appliquées (arXiv preprint arXiv:1401.0661)*, 2014.

- [3] John Ashburner and Gerard R Ridgway. Symmetric diffeomorphic modeling of longitudinal structural mri. *Frontiers in neuroscience*, 6, 2012.
- [4] Robert Azencott, Roland Glowinski, Jiwen He, Aarti Jajoo, Yipeng Li, Andrey Martynenko, Ronald HW Hoppe, Sagit Benzekry, and Stuart H Little. Diffeomorphic matching and dynamic deformable surfaces in 3d medical imaging. *Comput. Methods Appl. Math.*, 10(3):235–274, 2010.
- [5] Serdar K Balci, Polina Golland, Martha Shenton, and William M Wells. Free-form b-spline deformation model for groupwise registration. In *Medical image computing and computer-assisted intervention: MICCAI... International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 10, page 23. NIH Public Access, 2007.
- [6] Serdar K Balci, Polina Golland, and W Wells. Non-rigid groupwise registration using b-spline deformation model. *Open source and open data for MICCAI*, pages 105–121, 2007.
- [7] Martin Bauer, Martins Bruveris, et al. A new riemannian setting for surface registration. In *Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy-Geometrical and Statistical Methods for Modelling Biological Shape Variability*, pages 182–193, 2011.
- [8] Mirza Faisal Beg and Ali Khan. Computing an average anatomical atlas using lddmm and geodesic shooting. In *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on*, pages 1116–1119. IEEE, 2006.
- [9] Kanwal K Bhatia, Paul Aljabar, James P Boardman, Latha Srinivasan, Maria Murgasova, Serena J Counsell, Mary A Rutherford, Joseph V Hajnal, A David Edwards, and Daniel Rueckert. Groupwise combined segmentation and registration for atlas construction. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2007*, pages 532–540. Springer, 2007.
- [10] Kanwal K Bhatia, Jo Hajnal, Alexander Hammers, and Daniel Rueckert. Similarity metrics for groupwise non-rigid registration. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2007*, pages 544–552. Springer, 2007.
- [11] E Cavado, M Boccardi, R Ganzola, E Canu, A Beltramello, C Caltagirone, PM Thompson, and GB Frisoni. Local amygdala structural differences with 3t mri in patients with alzheimer disease. *Neurology*, 76(8):727–733, 2011.
- [12] N. Charon and A. Trouvé. The varifold representation of nonoriented shapes for diffeomorphic registration. *SIAM Journal on Imaging Sciences*, 6(4):2547–2580, 2013.
- [13] Matthew J Clarkson, M Jorge Cardoso, Gerard R Ridgway, Marc Modat, Kelvin K Leung, Jonathan D Rohrer, Nick C Fox, and Sébastien Ourselin. A comparison of voxel and surface based cortical thickness estimation methods. *Neuroimage*, 57(3):856–865, 2011.
- [14] CJ Cotter and DD Holm. Discrete momentum maps for lattice epdiff. In Temam and Tribbia, editors, *Handbook of Numerical Analysis*, pages 247–278. North-Holland, 2009.
- [15] William R Crum, Oscar Camara, Daniel Rueckert, Kanwal K Bhatia, Mark Jenkinson, and Derek LG Hill. Generalised overlap measures for assessment of pairwise and groupwise image registration and segmentation. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2005*, pages 99–106. Springer, 2005.
- [16] P Dupuis, U Grenander, and MI Miller. Variation Problems on Flows of Diffeomorphisms for Image Matching. *Quarterly of Applied Mathematics*, LVI(4):587–600, 1998.
- [17] Stanley Durrleman, Xavier Pennec, Alain Trouvé, Nicholas Ayache, et al. A forward model to build unbiased atlases from curves and surfaces. In *2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, pages 68–79, 2008.
- [18] Andreia V Faria, Alexander Hoon, Elaine Stashinko, Xin Li, Hangyi Jiang, Ameneh Mashayekh, Kazi Akhter, John Hsu, Kenichi Oishi, Jiangyang Zhang, et al. Quantitative analysis of brain pathology based on mri and brain atlases applications for cerebral palsy. *Neuroimage*, 54(3):1854–1861, 2011.
- [19] Joan Glaunès and Mario Micheli. Matrix-valued kernels for shape deformation analysis. *Geometry, Imaging and Computing*, 1(1):57–139, 2014.
- [20] Joan Glaunès, Anqi Qiu, Michael I Miller, and Laurent Younes. Large deformation diffeomorphic metric curve mapping. *International journal of computer vision*, 80(3):317–336, 2008.

- [21] Joan Glaunès, Alain Trouvé, and Laurent Younes. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In *Proceedings of CVPR'04*, volume 2, 2004.
- [22] Joan Glaunès, Alain Trouvé, and Laurent Younes. Modeling planar shape variation via hamiltonian flows of curves. In *Statistics and analysis of shapes*, pages 335–361. Birkhäuser Boston, 2006.
- [23] Xianfeng Gu, Yalin Wang, Tony F. Chan, Paul M. Thompson, and Shing-Tung Yau. Genus zero surface conformal mapping and its application to brain surface mapping. *Medical Imaging, IEEE Transactions on*, 23(8):949–958, 2004.
- [24] Steven Haker, Sigurd Angenent, Allen Tannenbaum, Ron Kikinis, Guillermo Sapiro, and Michael Halle. Conformal surface parameterization for texture mapping. *Visualization and Computer Graphics, IEEE Transactions on*, 6(2):181–189, 2000.
- [25] Richard F. Hartl, Suresh P. Sethi, and Raymond G. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev.*, 37(2):181–218, 1995.
- [26] Kelvin K Leung, Gerard R Ridgway, Sébastien Ourselin, Nick C Fox, Alzheimer’s Disease Neuroimaging Initiative, et al. Consistent multi-time-point brain atrophy estimation from the boundary shift integral. *Neuroimage*, 59(4):3995–4005, 2012.
- [27] Yaron Lipman and Ingrid Daubechies. Surface comparison with mass transportation. *arXiv preprint arXiv:0912.3488*, 2009.
- [28] Yaron Lipman and Ingrid Daubechies. Conformal wasserstein distances: Comparing surfaces in polynomial time. *Advances in Mathematics*, 227(3):1047–1077, 2011.
- [29] Marco Lorenzi, Nicholas Ayache, Giovanni Frisoni, Xavier Pennec, et al. 4d registration of serial brain’s mr images: a robust measure of changes applied to alzheimer’s disease. In *Spatio Temporal Image Analysis Workshop (STIA), MICCAI*. Citeseer, 2010.
- [30] Lok Ming Lui, Tsz Wai Wong, Paul Thompson, Tony Chan, Xianfeng Gu, and Shing-Tung Yau. Shape-based diffeomorphic registration on hippocampal surfaces using Beltrami holomorphic flow. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 323–330. Springer, 2010.
- [31] Jun Ma, Michael I Miller, Alain Trouvé, and Laurent Younes. Bayesian template estimation in computational anatomy. *NeuroImage*, 42(1):252–261, 2008.
- [32] Jun Ma, Michael I Miller, and Laurent Younes. A bayesian generative model for surface template estimation. *Journal of Biomedical Imaging*, 2010:16, 2010.
- [33] Michael I Miller, Laurent Younes, J Tilak Ratnanather, Timothy Brown, Huong Trinh, David S Lee, Daniel Tward, Pamela B Mahon, Susumu Mori, Marilyn Albert, et al. Amygdalar atrophy in symptomatic alzheimer’s disease based on diffeomorphometry: the biocard cohort. *Neurobiology of aging*, 2014.
- [34] Jean-Marie Morvan and Boris Thibert. On the approximation of a smooth surface with a triangulated mesh. *Computational Geometry*, 23(3):337–352, 2002.
- [35] Jorge Nocedal and Stephen J Wright. *Numerical Optimization, Second Edition*. Springer New York, 2006.
- [36] Kenichi Oishi, Kazi Akhter, Michelle Mielke, Can Ceritoglu, Jiangyang Zhang, Hangyi Jiang, Xin Li, Laurent Younes, Michael I Miller, Peter CM van Zijl, et al. Multi-modal mri analysis with disease-specific spatial filtering: initial testing to predict mild cognitive impairment patients who convert to alzheimer’s disease. *Frontiers in neurology*, 2, 2011.
- [37] Stéphane P Poulin, Rebecca Dautoff, John C Morris, Lisa Feldman Barrett, and Bradford C Dickerson. Amygdala atrophy is prominent in early alzheimer’s disease and relates to symptom severity. *Psychiatry Research: Neuroimaging*, 194(1):7–13, 2011.
- [38] Martin Reuter and Bruce Fischl. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage*, 57(1):19–21, 2011.
- [39] Martin Reuter, Nicholas J Schmansky, H Diana Rosas, and Bruce Fischl. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage*, 61(4):1402–1418, 2012.

- [40] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on*, 32(7):1153–1190, 2013.
- [41] Xiaoying Tang, Dominic Holland, Anders M Dale, Laurent Younes, and Michael I Miller. Baseline shape diffeomorphometry patterns of subcortical and ventricular structures in predicting conversion of mild cognitive impairment to alzheimer’s disease. *Journal of Alzheimer’s Disease*, 2014.
- [42] Duygu Tosun, Sarang Joshi, and Michael W Weiner. Neuroimaging predictors of brain amyloidosis in mild cognitive impairment. *Annals of neurology*, 74(2):188–198, 2013.
- [43] Alain Trouvé. Action de groupe de dimension infinie et reconnaissance de formes. *Comptes rendus de l’Académie des sciences. Série 1, Mathématique*, 321(8):1031–1034, 1995.
- [44] Carole J Twining, Tim Cootes, Stephen Marsland, Vladimir Petrovic, Roy Schestowitz, and Chris J Taylor. A unified information-theoretic approach to groupwise non-rigid registration and model building. In *Information Processing in Medical Imaging*, pages 1–14. Springer, 2005.
- [45] Marc Vaillant and Joan Glaunès. Surface Matching via Currents. In *Information Processing in Medical Imaging*, pages 381–392, 2005.
- [46] Fei Wang, Baba C Vemuri, and Anand Rangarajan. Groupwise point pattern registration using a novel cdf-based jensen-shannon divergence. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 1283–1288. IEEE, 2006.
- [47] Qian Wang, Liya Chen, Pew-Thian Yap, Guorong Wu, and Dinggang Shen. Groupwise registration based on hierarchical image clustering and atlas synthesis. *Human brain mapping*, 31(8):1128–1140, 2010.
- [48] Guorong Wu, Hongjun Jia, Qian Wang, and Dinggang Shen. Sharpmean: groupwise registration guided by sharp mean image and tree-based registration. *NeuroImage*, 56(4):1968–1981, 2011.
- [49] BT Thomas Yeo, Mert R Sabuncu, Tom Vercauteren, Nicholas Ayache, Bruce Fischl, and Polina Golland. Spherical demons: fast diffeomorphic landmark-free surface registration. *Medical Imaging, IEEE Transactions on*, 29(3):650–668, 2010.
- [50] Laurent Younes. *Shapes and diffeomorphisms*, volume 171. Springer, 2010.
- [51] Laurent Younes. Constrained diffeomorphic shape evolution. *Foundations of Computational Mathematics*, 12(3):295–325, 2012.
- [52] Laurent Younes, Marilyn Albert, and Michael I Miller. Inferring changepoint times of medial temporal lobe morphometric change in preclinical alzheimer’s disease. *NeuroImage: Clinical*, 2014.
- [53] Laurent Younes, Felipe Arrate, and Michael I Miller. Evolutions equations in computational anatomy. *NeuroImage*, 45(1):S40–S50, 2009.
- [54] Laurent Younes, J Tilak Ratnanather, Timothy Brown, Elizabeth Aylward, Peg Nopoulos, Hans Johnson, Vincent A Magnotta, Jane S Paulsen, Russell L Margolis, Roger L Albin, et al. Regionally selective atrophy of subcortical structures in prodromal hd as revealed by statistical shape analysis. *Human brain mapping*, 2012.
- [55] Wei Zeng and Xianfeng David Gu. Registration for 3d surfaces with large deformations using quasi-conformal curvature flow. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2457–2464. IEEE, 2011.
- [56] Yun Zeng, Chaohui Wang, Yang Wang, Xianfeng Gu, Dimitris Samaras, and Nikos Paragios. Dense non-rigid surface registration using high-order graph matching. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 382–389. IEEE, 2010.