Multiple shape registration using constrained optimal control

Sylvain Arguillère, Emmanuel Trélat, Alain Trouvé, Laurent Younes

MULTIPLE SHAPE REGISTRATION USING CONSTRAINED OPTIMAL CONTROL

SYLVAIN ARGUILLÈRE, EMMANUEL TRÉLAT, ALAIN TROUVÉ, AND LAURENT YOUNES

ABSTRACT. Lagrangian particle formulations of the large deformation diffeomorphic metric mapping algorithm (LDDMM) can be numerically challenging when the number of particles is large. In this paper, we introduce and discuss numerical schemes useful for surface and image matching, which are based on representing the Eulerian velocity over a finite-dimensional basis that deforms in time. The method is described within the optimal control formalism, and optimality conditions are derived, together with the equations that are needed to implement gradient-based methods. Experimental results are shown, both with surfaces and images.

1. INTRODUCTION

The large deformation diffeomorphic metric mapping (LDDMM) approach to shape matching is a powerful topology-preserving registration method with an increasing record of successful applications in medical imaging. It was first described in [38] for point sets and in [17, 59, 46, 8] for images and has become widely used in the medical imaging literature and other applications. While deeper understanding and extensions of the underlying theoretical framework was pursued [47, 22, 10, 48, 70, 21, 71, 9, 2] and alternative numerical methods were designed [5, 14, 60, 49, 29, 64, 4, 28, 18], LDDMM has been applied to medical imaging data including brain [45, 77, 52, 11, 54], heart [1, 6] and lung [65] images. This algorithm provides a non-rigid registration method between various types of objects (point sets, curves surfaces, functions, vector fields...) within a unified framework driven by Grenander's concept of deformable templates [25]. It optimizes a flow of diffeomorphisms that transform an initial object (shape) into a target one.

The practical importance of shape registration is underlined by the increasing amount of work that has flourished in the literature over the past few years. LDDMM is one among many methods that have been proposed to perform this task. Several such methods are based on elastic matching energies [7, 16], and other, like LDDMM, inspired by viscous fluid dynamics [12, 57, 62, 3, 63]. For surfaces, which will be our main focus, several authors have developed approaches to find approximate conformal parametrizations with respect to the unit disk or sphere [33, 31, 36, 26, 37, 68, 34, 27, 35]. More recently, quasi-conformal parametrizations based on the minimization of the Beltrami coefficient have been designed [75, 41, 67]. Another class of non-rigid registration methods include those based on optimal mass transportation [30, 32, 39, 40], while [44, 42, 43] introduce comparison methods based on Gromov–Hausdorff or Gromov–Wasserstein distances.

2000 *Mathematics Subject Classification.* 58D05 49N90 49Q10 68E10.

Key words and phrases. Shape analysis; optimal control; deformations; groups of diffeomorphisms.

This work was partially supported by the ONR award N000140810606.

Computational methods based on integer programming and graph optimization have also been recently introduced [76, 78, 23]. We also refer the reader to the survey papers [79, 15, 69, 72] and textbooks [24, 71] for additional entries on the literature.

In this paper, we discuss an extension of the LDDMM framework, in which multiple shapes are registered simultaneously within a deformation scheme involving contact constraints among the shapes. This is represented and solved as a constrained optimal control problem, in the spirit of the general framework recently introduced in [2].

Indeed, one of the characteristics of LDDMM is that it derives shape deformation from a global diffeomorphisms of the whole ambient space considered as a homogeneous medium, and does not allow for a differentiation of the deformation properties assumed by the shapes, or, more precisely, the objects they represent. This crude modeling may provide results that are not realistic in some applications. Consider the situation in which one studies several shapes, representing, for example, different sub-structures of the brain. In this case, if one assumes that all shapes are deformed by a single flow of diffeomorphisms, shapes coming too close to one another will undergo a tremendous deformation, which creates artifacts that can mislead subsequent analyses. One would rather associate a different diffeomorphism to each shape, independent from the others, but the issue is that the resulting collection of diffeomorphisms may not be consistent: the shapes could overlap along the deformation. The solution briefly introduced in [2] and developed in this paper is the following: embed the shapes into a "background", complement of the shapes, deformed by a new, independent deformation, and add *constraints* such that, as all the shapes are simultaneously transformed, their boundary moves with the boundary of the background so that the configuration consistency is preserved. This is the approach that we develop here, focusing on surface registration. Note that a multi-diffeomorphism approach has been recently developed for image matching [55], each diffeomorphism being restricted to a fixed region of the plane. The main (and fundamental) contrast with what we develop here is that, in our case, these subregions are variable and optimized, while they were fixed in [55]. The models along which sliding constraints are addressed in this paper and ours also differ.

This paper is organized as follows. We start by recalling the classical LDDMM algorithm in Section 2, setting the definitions, notation and appropriate framework for the rest of the paper. Then, in Section 3, we introduce rigorously the concept of multishape, describe identity and sliding background constraints, and describe the augmented Lagrangian algorithm for general constraints that will be used for in our numerical simulations. Section 4 follows, specializing the algorithm to the case of identity and sliding constraints in great details. Finally, Section 5 applies our method to synthetic examples to to real data as well.

2. Large Deformation Diffeomorphic Metric Mapping

2.1. **Notation.** In this paper, we define a *shape* as a C^p embedding $q : M \to \mathbb{R}^d$, where M is a compact manifold. We denote by $\mathcal{M}$ the corresponding *shape space*, which is an open subset of the Banach space $\mathcal{Q} = C^p(M, \mathbb{R}^d)$.

Typical examples are as follows:

- $M = \{1, \dots, m\}$ is finite, and q can be identified with a collection $q_1, \dots, q_m$ of distinct points in $\mathbb{R}^d$.
- $M = [0, 1]$ and q is a curve in $\mathbb{R}^d$.
- $M = S^{d-1}$ (the unit sphere in $\mathbb{R}^d$) and q is a hypersurface.

Our goal is to discuss models in which several shapes can deform, while being subject to contact constraints. The deformation process will be similar to the one designed for *large deformation diffeomorphic metric mapping* (LDDMM), which can be formulated as an *optimal control problem*. Before introducing our general framework, it will be easier to start with a description of the now well explored single-shape problem upon which we will build. For this, we let $(V, \|\cdot\|_V)$ be a Hilbert space of vector fields on $\mathbb{R}^d$, assumed to be continuously imbedded in the space $B := C_0^p(\mathbb{R}^d, \mathbb{R}^d)$, the completion of the space of smooth compactly supported vector fields for the norm $\|\cdot\|_{p,\infty}$, which denotes the sum of supremum norms of derivatives of order p or less, with $p \geq 1$. Then V possesses a *reproducing kernel*, that is a mapping $K : (x, y) \mapsto K(x, y)$, defined over $\mathbb{R}^d \times \mathbb{R}^d$, with values in the space of $d \times d$ matrices, such that all partial derivatives with order less than p with respect to each variable exist and

$$K(\cdot, y)a \in V \quad \text{with} \quad \langle K(\cdot, y)a\,,\, w\rangle_V = a^T w(y),$$

for all $(a, y) \in (\mathbb{R}^d)^2$. The LDDMM algorithm uses flows of time-dependent vector fields $v(\cdot) \in L^2(0, 1; V)$.

2.2. **Registering Two Shapes Using LDDMM.**

2.2.1. *General Problem.* The general LDDMM problem is formulated as the infinite-dimensional optimal control problem consisting of minimizing the cost functional

$$F(v) = \frac{1}{2}\int_0^1 \|u(t)\|_V^2\, dt + U(q(1)), \tag{1}$$

subject to the constraint

$$\begin{aligned} \partial_t q(t) &= u(t) \circ q(t) \quad \text{for a.e. } t \in [0, 1], \\ q(0) &= q_{init}. \end{aligned} \tag{2}$$

This differential constraint is a control system, where the control is the time-dependent vector field $u(\cdot) \in L^2(0, 1; V)$, the solution of which is $q(t, \cdot) = \varphi(t, q_{init}(\cdot))$ where φ is the flow of diffeomorphisms generated by $u(\cdot)$, defined as the unique solution of the Cauchy problem $\partial_t\varphi(t) = u(t)\circ\varphi(t)$, $\varphi(0) = \mathrm{id}_{\mathbb{R}^d}$. For every time t, we have $\varphi(t, \cdot) \in \mathit{Diff}^p$, the set of p-times differentiable diffeomorphisms in $\mathbb{R}^d$.

The function U is a matching cost function, that is, a penalization that pushes the solution of (1)-(2) towards a target. It will be assumed to be Fréchet differentiable from $\mathcal{Q}$ to $\mathbb{R}$. To simplify the discussion, and because this covers most of the interesting cases in practice, we will assume that there exists some fixed measure ν_M on M such that its derivative, denoted $dU(q)$ or dU_q

when evaluated at $q \in \mathcal{Q}$, can be expressed in the form $dU_q = z_q \nu_M$ for some (ν_M-measurable) $z_q : M \to \mathbb{R}^d$, meaning that

$$\forall h \in \mathcal{C}^p(M, \mathbb{R}^d), \quad (dU_q \,|\, h) = \int_M h(m) \cdot z_q(x)\, d\nu_M(x).$$

Throughout the paper, for any Banach space X, the notation $(\mu \,|\, v)$ will be used to designate the application $\mu(v)$ of a linear form $\mu \in X^*$ to a vector $v \in X$.

Under these assumptions, one can prove that the gradient of the objective function F defined by (1) (which is a mapping on $L^2([0,1], V)$) is given by

$$\nabla_V \tilde{F}(u)(t, \xi) = u(t, \xi) - \int_M K(\xi, q(t,x))\alpha(t,x)\, d\nu_M(x),$$

where K is the reproducing kernel of V, and $\alpha : [0,1] \times M \to \mathbb{R}^d$ is a time-dependent function defined by $\alpha(1, \cdot) = -z_{q(1)}$ and

$$\partial_t \alpha = -(du \circ q)^T \alpha. \tag{3}$$

This result implies, in particular, that the solutions of (1)-(2) must satisfy the Pontryagin maximum principle (see [51, 58]), which is the following first-order necessary condition for optimality. Let H_u be the Hamiltonian defined by

$$H_u(\rho, q) = (\rho \,|\, u \circ q) - \frac{1}{2}\|u\|_V^2,$$

for every $u \in V$, every $q \in \mathcal{Q}$ and every $\rho \in \mathcal{Q}^*$. If $u(\cdot)$ is an optimal control, solution of the optimal control problem (1)-(2), then it must be such that

$$u(t) = \mathrm{argmax}_w H_w(\rho(t), q(t)), \tag{4}$$

where (ρ, q) are solutions of

$$\begin{cases} \partial_t q(t) = \partial_\rho H_u(\rho(t), q(t)), \\ \partial_t \rho(t) = -\partial_q H_u(\rho(t), q(t)), \end{cases}$$

(ρ is the so-called *co-state*, or *adjoint state*) and $\rho(1) = -dU_{q(1)}$. Indeed, it suffices to take $\rho(\cdot) = \alpha(\cdot)\nu_M$ and to use properties of the reproducing kernel to check that all the conditions are satisfied. Moreover, (4) then implies that $u = \int_M K(\cdot, q(x))\alpha(x)\, d\nu_M$ at every time.

2.2.2. *Examples of shapes and matching cost functions.*

Example 1. To start with a simple example, let $M = \{1, \dots, m\}$ so that a shape $q = (q(1), \dots, q(m))$ is a collection of landmarks, and consider the landmark matching cost function defined by $U(q) = \sum_{k=1}^m |q(k) - y_k|^2$, for fixed $y = (y_1, \dots, y_m) \in \mathbb{R}^N$, in which $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$. We then have

$$(dU_q \,|\, h) = 2 \sum_{k=1}^m (q(k) - y_k)^T h(k)\,,$$

or, to interpret this result in the general form provided above, $dU_q = z\nu_M$ with $z(k) = 2(q(k) - y_k)$ and ν_M the counting measure on $\{1, \dots, m\}$.

Example 2. If μ is a scalar measure on $\mathbb{R}^d$ and z a μ-integrable $\mathbb{R}^d$-valued function defined on $\mathrm{support}(\mu)$, we will denote by $z\mu$ the vector measure such that

$$(z\mu \,|\, w) = \int_{\mathrm{support}(\mu)} z \cdot w \, d\mu,$$

where $z \cdot w$ denotes the standard euclidean dot product.

Vector measures of the form $z\mu$ are continuous linear forms over any space that is continuously imbedded in $C_0^0(\mathbb{R}^d, \mathbb{R}^d)$, and in particular over any reproducing kernel Hilbert space W. For such a space, equipped with a reproducing kernel χ, the operator norm of $z\mu$ is given by

$$\|z\mu\|_\chi^2 = \iint z(x) \cdot (\chi(x,y) z(y)) \, d\mu(x) \, d\mu(y),$$

and more generally, the norm of the difference between two such measures is

$$\begin{aligned}\|z\mu - \tilde{z}\tilde{\mu}\|_\chi^2 = &\iint z(x) \cdot (\chi(x,y) z(y)) \, d\mu(x) \, d\mu(y) \\ &- 2 \iint z(x) \cdot (\chi(x,y) \tilde{z}(y)) \, d\mu(x) \, d\tilde{\mu}(y) + \iint \tilde{z}(x) \cdot (\chi(x,y) \tilde{z}(y)) \, d\tilde{\mu}(x) \, d\tilde{\mu}(y).\end{aligned}$$

Note that W and V have no relationship one to each other, except that both have a continuous inclusion in $C_0^0(\mathbb{R}^d, \mathbb{R}^d)$, so that χ is different from K.

One can deduce from this the surface-matching cost function introduced in [61], in which an oriented surface S is represented as a geometric current and a dual-RKHS norm between currents is used. Identifying surface currents with vector measures, this leads to the representation of S given by $\mu_S = N_S \sigma_S$, where N_S is the unit normal to S and σ_S its volume form. Assume that M (the parameter space) is an oriented 2D manifold so that $S = q(M)$ is a surface, and let η be a positively oriented volume form on M. For $m \in M$ let $N_q(x) \in \mathbb{R}^3$ denote the "area-weighted normal" to $S = q(M)$ at $q(x)$, defined by $N_q(x) = dq_x e_1 \times dq_x e_2$ where (e_1, e_2) is an arbitrary basis of $T_x M$ such that $\eta_x(e_1, e_2) = 1$. Then

$$(\mu_S \,|\, w) = \int_M (w \circ q \cdot N_q) \, d\eta,$$

for every $w \in \mathcal{C}_0^0(\mathbb{R}^3, \mathbb{R}^3)$.

Now, given a reproducing kernel χ and a target surface $\tilde{S} = \tilde{q}(M)$, we define the surface-matching cost by

$$U(q) = \|\mu_{q(M)} - \mu_{\tilde{q}(M)}\|_\chi^2.$$

Example 3. This cost function is actually a special case of the most general framework in which one compares compact k-dimensional oriented submanifolds of $\mathbb{R}^d$, which we briefly discuss hereafter. Given such a manifold, S, with a global parametrisation $q : M \to S$, one can associate to any $\omega \in C_0^p(\mathbb{R}^d, (\Lambda^k \mathbb{R}^d)^*)$ (the set of C^p differential k-forms on $\mathbb{R}^d$ that vanish at infinity), its integral

$$(C_S \,|\, \omega) = \int_S \omega = \int_M q^*\omega,$$

where $q^*\omega$ denotes the pull-back of ω on M. An RKHS, $\tilde{W}$, of such forms, is a Hilbert space continuously embedded in $C_0^p(\mathbb{R}^d, (\Lambda^k \mathbb{R}^d)^*)$, with kernel $\tilde{\chi}(x, y)$ taking values in the space of bilinear functions on $\Lambda^k \mathbb{R}^d \times \Lambda^k \mathbb{R}^d$. The linear form C_S then belongs in $\tilde{W}^*$, and is a special form of a *geometric current*, as defined in [20]. If $S = q(M)$ and $\tilde{S}$ is a target manifold, they can be compared using the operator norm

$$U(q) = \|C_{q(M)} - C_{\tilde{S}}\|^2_{\tilde{W}^*}. \tag{5}$$

Now, if we consider $\hat{q} : \hat{M} \doteq (-1, 1) \times M \to \mathbb{R}^d$ a smooth perturbation of q such that $q_\varepsilon = q + \varepsilon \delta q + o(\varepsilon)$ where $q_\varepsilon(x) = \hat{q}(\varepsilon, x)$ for $\varepsilon \in (-1, 1)$ and $x \in M$, we have, for $M_\varepsilon \doteq \{\varepsilon\} \times M$,

$$(dU_q \,|\, \delta q) = \frac{d}{d\epsilon} \left(\int_M q_\epsilon^* \omega \right)_{|\epsilon=0} = \frac{d}{d\epsilon} \left(\int_{M_\epsilon} \hat{q}^* \omega \right)_{|\epsilon=0} = \int_{M_0} \mathcal{L}_{\partial/\partial\varepsilon}(\hat{q}^*\omega),$$

where $\mathcal{L}_{\partial/\partial\varepsilon}$ is the Lie derivative along the vector field $\frac{\partial}{\partial\varepsilon}$ on $\hat{M}$ (which is equal to $(1, 0) \in \mathbb{R} \times T_m M$ at any location $(\varepsilon, x) \in \hat{M}$) and

$$\omega = 2K_{\tilde{W}}(C_{q(M)} - C_{\tilde{S}}), \tag{6}$$

with $K_{\tilde{W}}$ the isometry from $\tilde{W}^*$ to $\tilde{W}$. We next show that

$$(dU_q \,|\, \delta q) = \int_M \alpha_q \cdot \delta q \, \mathrm{vol_M} + \int_{\partial M} \beta_q \cdot \delta q \, \mathrm{vol}_{\partial \mathrm{M}},$$

where vol_M and $\mathrm{vol}_{\partial M}$ are the positive Riemannian volume forms on M and ∂M, and $\alpha_q : M \to \mathbb{R}^d$ (resp. $\beta_q : \partial M \to \mathbb{R}^d$) is such that $\alpha_q(x)$ is normal to $S = q(M)$ (resp. $\beta_q(x)$ is normal to $\partial S = q(\partial M)$) at $q(x)$. Using the Cartan magic formula we get

$$\mathcal{L}_{\partial/\partial\varepsilon}(q^*\omega) = i_{\partial/\partial\varepsilon} d(q^*\omega) + d(i_{\partial/\partial\varepsilon}(q^*\omega)),$$

so that, applying the Stokes theorem,

$$(dU_q \,|\, \delta q) = \int_{M_0} i_{\partial/\partial\varepsilon}(\hat{q}^* d\omega) + \int_{\partial M_0} i_{\partial/\partial\varepsilon}(\hat{q}^*\omega),$$

where $i_{\partial/\partial\varepsilon}$ denotes the contraction operator. Note that, for $\xi_1, \ldots, \xi_k \in T_x M$,

$$\begin{aligned} i_{\partial/\partial\varepsilon}(\hat{q}^* d\omega)_{(0,x)}(\xi_1, \cdots, \xi_k) &= d\omega_{q(x)}(\delta q(x), dq_x\xi_1, \ldots, dq_x\xi_k) \\ &= d\omega_{q(x)}(\delta q^\perp(x), dq_x\xi_1, \ldots, dq_x\xi_k), \end{aligned}$$

where $\delta q^{\perp}(x)$ denotes the projection of $\delta q(x)$ on $T_{q(x)}S^{\perp}$ (since the form vanishes if $\delta q(x) \in T_{q(x)}S = T_{q(x)}q(M)$). Since the set of k-forms is a one-dimensional space on M, this means that we can write, for some function α_q such that $\alpha_q(x) \in (T_{q(x)}S)^{\perp}$, with $x \in M$,

$$i_{\partial/\partial\varepsilon}(q^* d\omega)_{(0,x)} = ((\delta q \cdot \alpha_q)\,\mathrm{vol}_M)_x\,.$$

Similarly,

$$i_{\partial/\partial\varepsilon}(q^*\omega_q)_{(0,x)} = ((\delta q \cdot \beta_q)\mathrm{vol}_{\partial M})_x,$$

for $\beta_q(x) \in (T_{q(x)}\partial S)^{\perp}$.

The data term defined in (5) is derived from this general construction, using the fact that two-forms in $\mathbb{R}^3$ (or $(d-1)$-forms in $\mathbb{R}^d$) can be identified with vector fields via $\omega_x(e_1, e_2) = w(x) \cdot (e_1 \times e_2)$. The current C_S is then identified with the vector measure μ_S. The form ω introduced in (6) becomes, introducing a global parametrizaion $\tilde{q} : M \to \tilde{S}$ of the target $\tilde{S}$, the vector field

$$w(\cdot) = 2\int_{\mathbb{R}^d} \chi(\cdot, y)\, d(\mu_{q(M)} - \mu_{\tilde{q}(M)})(y).$$

With this identification, we have $\alpha_q = \mathrm{div}(w) \circ q\, N_q$ where N_q is the oriented "area-weighted" normal to S at $q(x)$ defined previously, and $\beta_q = \tau_q \times w \circ q$, where τ_q is the oriented "length-weighted" tangent to $q(\partial M)$ given as $\tau_q(x) = dq_x e_1$, where e_1 is the unit positively oriented tangent vector at x along ∂M.

Example 4. Returning to surfaces, the discrete case, in which triangulated surfaces are compared, is, for practical purposes, even more important. We here also consider the case $M = \{1, \ldots, m\}$, with an additional family F of facets, which are ordered triples (i, j, k) with $i, j, k \in M$. (We assume that F is a consistent with a manifold structure: The set V_i of indices that share a facet with i must form a chain and no pair of indices can be included in more than two facets.)

Given a one-to-one mapping $q : M \to \mathbb{R}^3$, define S_q as the collection of triangles $S_q = \{(q(i), q(j), q(k)), (i, j, k) \in F\}$. If $f = (i, j, k)$, let $q(f) = (q(i), q(j), q(k))$, $N_q(f) = (q(j) - q(i)) \times (q(k) - q(i))$ and $c_q(f) = (q(i) + q(j) + q(k))/3$ respectively denote the triangle, area-weighted normal and center associated to the facet f. Following [61], we define the vector measure associated to q by

$$\mu_q = \sum_{f \in F} N_q(f)\delta_{c_q(f)}.$$

Here, δ_x (with $x \in \mathbb{R}^3$) denotes the atomic measure of mass 1 with support $\{x\}$. The (discrete) surface matching cost associated to a target $\tilde{q}$ is then defined by

$$U(q) = \|\mu_q - \mu_{\tilde{q}}\|^2_{\chi}.$$

Note that $\tilde{q}$ does not need to be consistent with q, and can be defined on a different set of indices, $\tilde{M} = \{1, \ldots, \tilde{m}\}$ and triangle structure $\tilde{F}$. One then has $dU_q = \alpha_q \nu_M$, where, as above, ν_M is the counting measure on M, and

$$\alpha_q(i) = \sum_{f \in F, i \in f} (dZ^T_{c_q(f)} N_q(f)/3 + e_q(f, i) \times Z(c_q(f))),$$

with $e_q(f,i) = q(k) - q(j)$ the oriented edge opposed to $q(i)$ in $q(f)$, and

$$Z(\cdot) = \sum_{f \in F} \chi(\cdot, c_q(f)) N_q(f) - \sum_{\tilde{f} \in \tilde{F}} \chi(\cdot, c_{\tilde{q}}(\tilde{f}) N_{\tilde{q}}(\tilde{f}).$$

2.2.3. *Reduced Problem.* Since the optimal control must satisfy

$$v = \int_M K(\cdot, q(t,x))\alpha(t,x)\, d\nu_M \tag{7}$$

for some function α defined on M, it is natural to parametrize v by α and use this function as a new control. We define the inner product

$$\langle \alpha\,,\, \beta \rangle_q = \int_{M\times M} \alpha(x)^T K(q(x), q(\tilde{x}))\beta(\tilde{x})\, d\nu_M(x)\, d\nu_M(\tilde{x})$$

between two measurable functions α and β defined on M. If v is given by (7), the reproducing property of the kernel implies that $\|v\|_V^2 = \|\alpha\|_q^2$. The optimal control problem (1)-(2) is then equivalent to the reduced problem consisting of minimizing the cost functional

$$F(\alpha) = \frac{1}{2}\int_0^1 \|\alpha(t)\|_{q(t)}^2\, dt + U(q(1)), \tag{8}$$

subject to the constraint (control system)

$$\partial_t q(t,x) = \int_M K(q(t,x), q(t,\tilde{x}))\alpha(t,\tilde{x}) d\nu_M(\tilde{x}), \tag{9}$$

almost everywhere over the time interval $[0,1]$.

According to [6, 13, 61, 73], we have $\nabla F(\alpha) = \alpha - p$, where p (the co-state) is a time-dependent vector-valued measurable function on M such that $p(1)\nu_M = -dU_{q(1)}$ and

$$\partial_t p(t) = -\partial_q \left(\langle p(t)\,,\, \alpha(t) \rangle_q - \|\alpha(t)\|_{q(t)}^2/2 \right),$$

where q is defined by (9). Here, the gradient is computed with respect to the inner product $\langle \cdot\,,\, \cdot \rangle_q$.

3. Multiple Shape problems

3.1. **Motivating Examples.** In the previous formulation, the shape evolution was controlled by a single, smooth vector field v, inducing a single diffeomorphism of $\mathbb{R}^d$ restricted to the considered shape. This approach has been successfully used to model variations of single, homogeneous shapes, and led to important applications in computational anatomy, including, among many other examples, the impact of pathologies like Huntington disease [74], schizophrenia [53], and Alzheimer's disease [66, 19, 54] on brain structures. This deformation model, however, is not well adapted in situations in which several shapes interact, or situations in which shapes have heterogeneous parts. Let us review some motivating examples.

(1) Consider a schematic representation of a kite, or a manta ray, composed with a two-dimensional surface, representing the body, and an open curve attached to it representing the tail. When comparing two such objects, the body is assumed to only show small differences in shape, while the tail can vary widely.
(2) Consider a two-dimensional representation of a mouth, with two curves representing the upper and lower lip. Because the mouth can be wide open or closed, it is not possible to consider its deformations as resulting from the restriction of a smooth diffeomorphism of $\mathbb{R}^2$.
(3) Finally, it is natural, when analyzing multiple organs in the human body, to consider multiple shapes, each of them being relatively stable (only subject to small deformations) while their position with respect to each other is subject to larger variations, so that the background (the intersection of their complements) is subject to very large deformations. Here again, modeling the whole process with a single diffeomorphism is not adequate.

These examples suggest using multiple deformations applied to each component of the considered model. Generalizing (1)-(2), consider parameter spaces $M_1, \dots, M_n$ for an n-component model. Each shape, or component, is a mapping $q^{(k)} \in \mathcal{Q}_k : M_k \to \mathbb{R}^d$. The shape space will then be $\mathcal{Q} = \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_n$. To each shape, associate a control $u_k \in V_k$, where V_k is an RKHS embedded in $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ with the state evolution equation $\partial_t q^{(k)} = u_k \circ q^{(k)}$. We can then choose each V_k according to how "wildly" we want to allow the k-th shape to deform. These evolutions, however, must be consistent with each other, implying contact constraints that we will consider in two forms:

- **Identity constraints:** These are constraints that make a subset of the k-th shape stay stitched to a subset of the l-th shape, so that these subsets coincide in $\mathbb{R}^d$ and move identically along the deformation. Given some pair $(k, l) \in \{1, \dots, n\}^2$, and given a one-to-one mapping $g_{kl} : A_{kl} \subset M_k \to A_{lk} = g_{kl}(A_{kl}) \subset M_l$, one has $q^{(k)}(x) = q^{(l)}(g_{kl}(x))$ for every $x \in A_{kl}$.
- **Sliding constraints:** These are constraints that force a closed submanifold of the k-th shape to slide on a corresponding submanifold of the l-th shape along the deformation, for all (k, l). Here, we assume that all parameter spaces are orientable differential manifolds, and that all $q^{(k)}$'s are immersions. Given some pair $(k, l) \in \{1, \dots, n\}$, and a closed submanifold without boundary $A_{kl} \subset M_k$, there exists a diffeomorphism $g_{kl} : A_{kl} \to A_{lk} \subset M_l$ onto a fixed closed submanifold A_{lk} of M_l such that $q^{(k)}(x) = q^{(l)}(g_{kl}(x))$ for every $x \in A_{kl}$.

Let us turn back to the examples mentioned at the beginning of the section. For Example (1), we can take $M_1 = S^2$ and $M_2 = [0, 1]$, and, letting x_0 represent the north pole in S^2, impose $q^{(1)}(x_0) = q^{(2)}(0)$. We can then assign different deformation models to $q^{(1)}$ and $q^{(2)}$ via the metrics on V_1 and V_2.

Example (2) requires a slightly more complex construction, that only imperfectly addresses the issue. Let $M_1 = M_2 = [0, 1]$ and $M_3 = \{1, 2\} \times [0, 1]$. Let $q^{(1)}$ represent the upper lip, $q^{(2)}$ the lower one and $q^{(3)}$ their union. We use the identity constraints $q^{(1)}(1) = q^{(2)}(0)$ and

$q^{(2)}(1) = q^{(1)}(0)$ for the extremities of each lip, and $q^{(3)}(1,\cdot) = q^{(1)}(\cdot)$, $q^{(3)}(2,\cdot) = q^{(2)}(\cdot)$. We take $V_1 = V_2$ and choose V_3 such that the latter allows for large deformations at a small cost. With this model, it becomes easier to "almost" close the mouth, although the deformation inside the mouth must remain diffeomorphic, so that the closing cannot go all the way.

For Example (3), there are n shapes, $n-1$ of which are associated with the organs, and the last of which represents the background. For example, we can take $M_k = S^2$ for $k = 1,\dots,n-1$, and $M_n = \{1,\dots,n-1\} \times S^2$. Assuming that the shapes do not intersect, we can define identity or sliding constraints for the background, enforcing $q^{(n)}(k,\cdot) = q^{(k)}(\cdot)$ or $q^{(n)}(\{k\}\times S^2) = q^{(k)}(S^2) = S_k$ for $k \in \{1,\dots,n-1\}$ during the deformation.

3.2. **Induced Constraints.** The previous constraints can be reformulated as equality constraints involving the state and control. Identity constraints $q^{(k)}(x) = q^{(l)}(g(x))$ are equivalent (taking time derivatives) to $u_k(t, q^{(k)}(t,x)) = u_l(t, q^{(l)}(t, g(x)))$ as soon as the constraints are satisfied at time $t = 0$, which we obviously assume.

Making the same assumption, sliding constraints can be expressed as

$$N^{(k)}(t, q^{(k)}(t,x))^T (u_k(t, q^{(k)}(t,x)) - u_l(t, q^{(k)}(t,x))) = 0, \tag{10}$$

where $N^{(k)}(t, q^{(k)})$ is a $d \times (d - \dim(A_{kl}))$ matrix consisting of independent vectors perpendicular to $T_{q^{(k)}} B_{kl}(t)$ (e.g, a normal frame to B_{kl}), with $B_{kl} = q^{(k)}(A_{kl})$ for every (k,l). Let us briefly justify this statement.

We express the sliding constraint as $q^{(k)}(t,x) = q^{(l)}(t, g(t,x))$ for some diffeomorphism $g(t,\cdot) : A_{kl} \to A_{lk}$, assuming a differentiable dependency on time. Taking time derivatives, we get

$$u_k(t, q^{(k)}(t,x)) = u_l(t, q^{(l)}(t, g(t,x))) + dq^{(l)}(t, g(t,x))\partial_t g(t,x), \quad x \in M_k.$$

Since $q^{(l)}(t, g(t,x)) = q^{(k)}(t,x)$, we obtain

$$\begin{aligned} u_k(t, q^{(k)}(t,x)) - u_l(t, q^{(k)}(t,x)) &= dq^{(l)}(t, g(t,x))\partial_t g(t,x) \\ &= dq^{(k)}(t,x) dg(t,x)^{-1} \partial_t g(t,x), \end{aligned}$$

which is tangent to B_{kl} at x. Note that, since the image of $g(t,\cdot)$ is A_{lk} for every time t, we do have $\partial_t g(t,x) \in T_{g(t,x)} A_{lk} = dg(t,x)(T_x A_{kl})$, so $dg(t,x)^{-1}\partial_t g(t,x)$ is well-defined.

Conversely, assume that (10) holds for every $x \in A_{kl}$, with $q^{(k)}(0,x) = q^{(l)}(0, g_0(x))$ for some diffeomorphism $g_0 : A_{kl} \to A_{lk} \subset M_l$. Then for every time t, the mapping

$$w : x \in A_{kl} \mapsto dq^{(k)}(t,x)^{-1} \underbrace{(u_k(t, q^{(k)}(t,x)) - u_l(t, q^{(k)}(t,x)))}_{\in T_{q^{(k)}(x)} B_{kl}} \in T_x A_{kl}$$

defines a time-dependent vector field on A_{kl}. Since A_{kl} is a closed manifold, this vector field is complete, and we denote its flow by $h(t,\cdot) : A_{kl} \to A_{kl}$. Then

$$\begin{aligned} \partial_t q^{(k)}(t, h(t,x)) &= u_k(t, q^{(k)}(t, h(t,x))) + dq^{(k)}(t, h(t,x))\partial_t h(t,x) \\ &= u_l(q^{(k)}(t, h(t,x))), \end{aligned}$$

where the last identity holds for every $x \in A_{kl}$, so that $q^{(k)}(t, h(t,x))$ and $q^{(l)}(t, g_0(x))$ satisfy the same differential equation with the same initial condition and therefore coincide. Hence, letting $g(t,x) = g_0(h^{-1}(t,x))$, we obtain $q^{(l)}(t, g(t,x)) = q^{(k)}(t,x)$ for every $x \in A_{kl}$.

It is possible to extend this construction to the case where A_{kl} is a compact manifold with boundary. In this case, the matrix $N^{(k)}(t)$ must consist of a normal frame along $\partial B_{kl}(t)$ and possesses therefore an extra column, and we get an additional constraint along the boundary of A_{kl}.

We will need the constraints to depend smoothly on q, and therefore we will need a smooth representation of the normal space to q (a smooth map $q \mapsto N(q)$) in order to be able to use (10). When this is not possible (or convenient), one can also use the alternative approach of introducing a new state, say $N^{(k)}(t,x)$, evolving according to

$$\partial_t N^{(k)} = -du_k(q^{(k)})^T N^{(k)}, \tag{11}$$

which ensures that $N^{(k)}(t,x)$ remains perpendicular to $T_{q^{(k)}(t,x)}S_k$ as soon as this holds true at $t = 0$. The constraint $N^{(k)}(t,x)^T(u_k(t, q^{(k)}(t,x)) - u_l(t, q^{(k)}(t,x)))$ is now a smooth function of the extended state.

The two problems that we consider are therefore special cases of the general problem considered in [2], which is the problem of minimizing the cost functional

$$\frac{1}{2}\sum_{k=1}^{n}\int_0^1 \|u_k(t)\|^2_{V_k}\,dt + \sum_{k=1}^{n} U_k(q^{(k)}(1)), \tag{12}$$

subject to the constraints

$$\partial_t q^{(k)}(t) = u_k(t, q^{(k)}(t)), \quad \text{and} \quad C(q(t))u(t) = 0, \tag{13}$$

almost everywhere over the time interval $[0,1]$, where $C : \mathcal{M} \to L(V, \mathcal{Y})$ takes values in the space of bounded linear operators from V to a Banach space $\mathcal{Y}$. Here, we have $V = V_1 \times \cdots \times V_n$ and $q = (q^{(1)}, \ldots, q^{(n)})$.

The study of this constrained optimal control problem, and in particular, the derivation of its first-order optimality conditions (of the type of Pontryagin maximum principle), is challenging in this infinite-dimensional setting. In [2], it is proved that, under some differentiability conditions, and *under the important assumption that $C(q)$ is surjective for every $q \in \mathcal{M}$*, optimal solutions must be such that there exist $p = (p^{(1)}, \ldots, p^{(n)}) \in H^1([0,1], \mathcal{Q}^*)$ and $\lambda \in L^2([0,1], \mathcal{Y}^*)$ that satisfy

$$\begin{cases} \partial_t q^{(k)} = u_k(q^{(k)}), \\ \partial_t p^{(k)} = -\partial_{q^{(k)}}\big(p^{(k)} \,|\, u_k(q^{(k)})\big) - \partial_{q^{(k)}}(\lambda \,|\, C(q)u), \\ \langle u_k, v\rangle_{V_k} = -\big(p^{(k)} \,|\, v \circ q^{(k)}\big) - (\lambda \,|\, C_k(q)v), \quad v \in V_k, \\ \displaystyle\sum_{k=1}^{n} C_k(q)u_k = 0, \end{cases} \tag{14}$$

where $C_k(q)u_k = C(q)(0, \dots, 0, u_k, 0, \dots, 0)$.

Unfortunately, the constraints $C(q)$ that correspond to our identity or contact constraints are, in general, not surjective, and the results of [2] cannot be applied in a fully general infinite-dimensional context. However, surjectivity becomes almost straightforward when these constraints are discretized to a finite number. They are true as soon as the points involved in the constraints are all distinct, which is a mild assumption. We now proceed to the description of a discrete version of this approach.

4. Discrete Approximations

4.1. **Augmented Lagrangian.** As an example, and to simplify the presentation, we detail our implementation for multi-shape problems in which shapes interact (through constraints) with a background, but not directly with each other. Direct interactions between shapes can be handled in a similar way. Our constrained optimization method uses the augmented Lagrangian method (see, e.g., [50]). In a nutshell, in order to minimize a function $u \mapsto F(u)$ subject to multi-dimensional equality constraints $C(u) = 0$, the augmented Lagrangian method consists of considering the functional

$$L(u) = F(u) - \lambda \cdot C(u) + \frac{\mu}{2}|C(u)|^2,$$

in which λ lives in the dual space of the space of constraints Y, and μ is a positive real number. Each iteration of the algorithm consists in minimizing L with fixed λ and μ (our implementation using nonlinear conjugate gradient) until the gradient norm passes below some upper bound, and then in updating λ according to the rule

$$\lambda \leftarrow \lambda - \mu C(u),$$

before running a new minimization of L. The constant μ is increased only if needed, i.e., if the norm of the constraint did not decrease enough during the minimization. More details can be found in [50].

We first apply this to identity constraints, which only require the shapes to be discretized into a sets of points. We will then discuss sliding constraints, which will require more structure in order to define normal frames to the boundary.

4.2. **Identity Constraints.** We consider $n-1$ objects, discretized into point sets, so that M_k is a finite set of indices for each k. Let $x_j^{(k)} = q^{(k)}(j)$ and $x^{(k)} = (x_1^{(k)}, \dots, x_{m_k}^{(k)})$, for $k = 1, \dots, n-1$, with $m_k = |M_k|$. We add as n-th object the background, defined on $M_n = (\{1\} \times M_1) \cup \dots \cup (\{n-1\} \times M_{n-1})$. We let $z_j^{(k)} = q^{(n)}(k, j)$, $z^{(k)} = (z_1^{(k)}, \dots, z_{m_k}^{(k)})$ and $z = (z^{(1)}, \dots, z^{(n-1)})$ (a collection of $m = m_n = m_1 + \dots + m_{n-1}$ points).

Assume that end-point cost functions $U_1(x^{(1)}), \dots, U_{n-1}(x^{(n-1)})$ are defined, typically measuring the discrepancy between each collection of points and an associated target. We assume similar functions $\tilde{U}_1(z^{(1)}), \dots, \tilde{U}_{n-1}(z^{(n-1)})$ for the background, typically using $U_j = \tilde{U}_j$. The associated

constrained optimal control problem consists in minimizing the cost functional

$$\frac{1}{2}\sum_{k=1}^{n}\int_0^1 \|u_k(t)\|^2_{V_k}\,dt + \sum_{k=1}^{n-1} U_k(x^{(k)}(1)) + \sum_{k=1}^{n-1}\tilde{U}_k(z^{(k)}(1)),$$

subject to the constraints (almost everywhere along $[0,1]$)

$$\begin{cases} \partial_t x_j^{(k)} = u_k(x_j^{(k)}) \quad j=1,\dots,m_k,\ k=1,\dots,n-1,\\ \partial_t z_j^{(k)} = u_n(z_j^{(k)}), \quad j=1,\dots,m_k,\ k=1,\dots,n-1,\\ z^{(k)} = x^{(k)}, \quad k=1,\dots,n-1. \end{cases}$$

For y and y' ordered families of points in $\mathbb{R}^d$, let $K^{(k)}(y,y')$ be the matrix formed with all $d\times d$ blocks $K_{V_k}(y_i,y'_j)$, and let $K^{(k)}(y)=K^{(k)}(y,y)$, for $k=1,\dots,n$, where K_{V_k} is the kernel of V_k. Since the problems only depend on the values taken by $u_1,\dots,u_n$ on their corresponding point set trajectories $x^{(1)},\dots,x^{(n-1)},z$, the optimal vector fields take the form

$$\begin{aligned} u_k(\cdot) &= K^{(k)}(\cdot,x^{(k)})\alpha^{(k)}, \quad k=1,\dots,n-1,\\ u_n(\cdot) &= K^{(n)}(\cdot,z)\beta, \end{aligned}$$

for some families of d-dimensional vectors $\alpha^{(1)},\dots,\alpha^{(n-1)},\beta$. The problem can therefore be reduced to the finite-dimensional optimal control problem consisting in minimizing the cost functional

$$E(\alpha,\beta,x,z) = \frac{1}{2}\sum_{k=1}^{n-1}\int_0^1 \alpha^{(k)}\cdot(K^{(k)}(x^{(k)})\alpha^{(k)})\,dt + \frac{1}{2}\int_0^1 \beta\cdot(K^{(n)}(z)\beta)\,dt + \sum_{k=1}^{n-1} U_k(x^{(k)}(1)) + \sum_{k=1}^{n-1}\tilde{U}_k(z^{(k)}(1))$$

subject to the constraints (almost everywhere along $[0,1]$)

$$\begin{cases} \partial_t x^{(k)} = K^{(k)}(x^{(k)})\alpha^{(k)},\\ \partial_t z = K^{(n)}(z)\beta,\\ z^{(k)} = x^{(k)}, \quad k=1,\dots,n-1. \end{cases}$$

Extending E with the augmented Lagrangian method, we introduce coefficients $\lambda^{(k)}$, $k=1,\dots,n-1$ (where $\lambda^{(k)}$ has the same dimension as $x^{(k)}$) and $\mu>0$, defining

$$L(\alpha,\beta,x,z) = \frac{1}{2}\sum_{k=1}^{n-1}\int_0^1 \alpha^{(k)}\cdot(K^{(k)}(x^{(k)})\alpha^{(k)})\,dt + \frac{1}{2}\int_0^1 \beta\cdot(K^{(n)}(z)\beta)\,dt + \sum_{k=1}^{n-1} U_k(x^{(k)}(1)) + \sum_{k=1}^{n-1}\tilde{U}_k(z^{(k)}(1)) - \sum_{k=1}^{n-1}\int_0^1 \lambda^{(k)}\cdot(x^{(k)}-z^{(k)})\,dt + \frac{\mu}{2}\sum_{k=1}^{n-1}\int_0^1 |x^{(k)}-z^{(k)}|^2\,dt,$$

which will be minimized subject to the constraints (almost everywhere along $[0,1]$)

$$\begin{cases} \partial_t x^{(k)} = K^{(k)}(x^{(k)})\alpha^{(k)}, \\ \partial_t z = K^{(n)}(z)\beta. \end{cases}$$

From the constraints, L can be considered as a function of α and β only, and its differential with respect to these variables can be computed via the adjoint method as follows. Denoting the co-states by $p^{x,k}$, $k = 1, \dots, n-1$, and p^z, the associated Hamiltonian is

$$H = \sum_{k=1}^{n-1} \int_0^1 p^{x,k} \cdot K^{(k)}(x^{(k)})(\alpha^{(k)})\, dt + \int_0^1 p^z \cdot K^{(n)}(z)\beta\, dt - L.$$

The computation of the gradient of L follows the same general scheme as the one described in Section 2.2 for the basic LDDMM algorithm. Given α and β and the associated trajectories x and z, one has solve the adjoint equations

$$\begin{aligned} \partial_t p^{x,k} &= -\partial_{x^{(k)}} H, \quad & p^{x,k}(1) &= -\nabla U_k(x^{(k)}(1)), \quad & k = 1, \dots, n-1, \\ \partial_t p^z &= -\partial_z H, & p^{z,k}(1) &= -\nabla \tilde{U}_k(z^{(k)}(1)), & k = 1, \dots, n-1. \end{aligned}$$

The computation of the differential system gives

$$\begin{aligned} \partial_t p_i^{x,k} = &- \sum_{j=1}^{m_k} \nabla_1(p_i^{x,k} \cdot K^{(k)}(x_i^{(k)}, x_j^{(k)})\alpha_j^{(k)}) - \sum_{j=1}^{m_k} \nabla_1(\alpha_i^{(k)} \cdot K^{(k)}(x_i^{(k)}, x_j^{(k)})p_j^{x,k}) \\ &+ 2\sum_{j=1}^{m_k} \nabla_1(\alpha_i^{(k)} \cdot K^{(k)}(x_i^{(k)}, x_j^{(k)})\alpha_j^{(k)}) - (\lambda^{(k)} - \mu(x^{(k)} - z^{(k)})), \end{aligned}$$

and

$$\begin{aligned} \partial_t p_i^{z,k} = &- \sum_{l=1}^{n-1}\sum_{j=1}^{m_l} \nabla_1(p_i^{z,k} \cdot K^{(n)}(z_i^{(k)}, z_j^{(l)})\beta_j^{(l)}) - \sum_{l=1}^{n-1}\sum_{j=1}^{m_l} \nabla_1(\beta_i^{(k)} \cdot K^{(n)}(z_i^{(k)}, z_j^{(l)})p_j^{z,l}) \\ &+ 2\sum_{l=1}^{n-1}\sum_{j=1}^{m_l} \nabla_1(\beta_i^{(k)} \cdot K^{(n)}(z_i^{(k)}, z_j^{(l)})\beta_j^{(l)}) + \sum_{l=1}^{n}(\lambda^{(l)} - \mu(x^{(l)} - z^{(l)})) \end{aligned}$$

The gradient of L with respect to (α, β) is then deduced from the partial differentials of H with respect to these variables, yielding

$$\begin{aligned} \nabla_{\alpha^{(k)}} L &= K^{(k)}(x^{(k)})(\alpha^{(k)} - p^{x,k}), \\ \nabla_\beta L &= K^{(n)}(z)(\beta - p^z). \end{aligned}$$

Alternatively, on may choose to use the gradient relative to the dot product on $V_1 \times \ldots \times V_n$, which is simply given by

$$\nabla_{\alpha^{(k)}} L = \alpha^{(k)} - p^{x,k},$$
$$\nabla_{\beta} L = \beta - p^{z}.$$

The latter choice is simpler, and generally more efficient numerically.

4.3. **Sliding Interface.** Assume that the parameter sets M_k are vertices of pure oriented, simplicial complexes T_k of dimension $r_k < d$ (we will however only provide implementation details for codimension $d - r_k = 1$). We let F_k denote the set of facets of the k-th complex. We also assume that $T_1, \ldots, T_{n-1}$ are disjoint and that T_n is their union, $T_n = \bigcup_{k=1}^{n-1} T_k$. We also let $F = \bigcup_{k=1}^{n-1} F_k$ (disjoint union).

The associated shape space is formed by functions $q_k : M_k \to \mathbb{R}^d$ such that $q_k(f)$ is not degenerate (i.e., has maximal dimension) for all $f \in F_k$. Each object is allowed to slide against the background. We will write $x^{(k)} = q^{(k)}(M_k)$, $k = 1, \ldots, n-1$, and $z^{(k)} = q^{(n)}(M_k)$, $z = (z^{(1)}, \ldots, z^{(n-1)})$, in accordance with our previous notation. If $f \in F_k$ is a facet in $T_k \subset T_n$, we discretize (10) into

$$N^{(n)}(f) \cdot \left(\sum_{j \in f} (u_k(z_j^{(k)}) - u_n(z_j^{(k)})) \right) = 0, \tag{15}$$

where $N^{(n)}(f)$ is a $d \times (d - r_k)$ matrix spanning the normal space to $q^{(n)}(f)$, assumed to be defined as a smooth function of $q^{(n)}$. If $r_k = d - 1$, this is always possible, since $N^{(n)}$ is a vector that can be taken as the cross product of $z_{f,2} - z_{f,1}, \ldots, z_{f,d} - z_{f,1}$ where $z_{f,1}, \ldots, z_{f,d}$ is any labeling of the vertices of $q^{(n)}(f)$ ordered consistently with the orientation.

We now restrict to this case, with $d = 3$, so that shapes are triangulated surfaces in $\mathbb{R}^3$, as discussed in Section 2.2. For $f \in F_k$ and $j \in f$, we denote by $e_{j,f}$ the edge $(z_{j''}^{(k)} - z_{j'}^{(k)})$ where j' and j'' are the other two vertices of f such that (j, j', j'') is positively oriented. Similarly, let $e'_{j,f} = (z_{j'}^{(k)} - z_j^{(k)})$ and $e''_{j,f} = (z_{j''}^{(k)} - z_j^{(k)})$ be the two edges stemming from $z_j^{(k)}$ so that

$$e'_{j,f} \times e''_{j,f} = 2 \operatorname{area}(q^{(n)}(f)) N^{(n)}(f) =: \tilde{N}^{(n)}(f)$$

is the area-weighted positively oriented normal to f in $q^{(n)}(M_k)$. Note that $e_{j,f} = e''_{j,f} - e'_{j,f}$.

With this notation, we can rewrite the constraint in the form

$$\sum_{j \in f} \det(e'_{j,f}, e''_{j,f}, u_k(z_j^{(k)}) - u_n(z_j^{(k)})) = 0.$$

holding for all $f \in F_k$ and $k = 1, \ldots, n-1$.

Introducing a Lagrange multiplier λ_f for each of these constraints, after reduction of the vector fields, which proceeds similarly to the identity constraints case, the augmented Lagrangian takes

the form

$$L(\alpha,\beta,x,z)=\frac{1}{2}\sum_{k=1}^{n-1}\int_0^1 \alpha^{(k)}\cdot(K^{(k)}(x^{(k)})\alpha^{(k)})\,dt+\frac{1}{2}\int_0^1\beta\cdot(K^{(n)}(z)\beta)\,dt+\sum_{k=1}^{n-1}U_k(x^{(k)}(1))$$
$$+\sum_{k=1}^{n-1}\tilde{U}_k(z^{(k)}(1))-\sum_{k=1}^{n-1}\sum_{f\in F_k}\int_0^1(\lambda_f\Gamma_f^{(k)}(x^{(k)},z)-\frac{\mu}{2}\Gamma_f^{(k)}(x^{(k)},z)^2)\,dt,$$

with

$$\Gamma_f^{(k)}(x^{(k)},z):=\sum_{j\in f}\det\left(e'_{j,f},e''_{j,f},K^{(k)}(z_j^{(k)},x^{(k)})\alpha^{(k)}-K^{(n)}(z_j^{(k)},z)\beta\right).$$

We now compute the evolution equations for the co-states, as done with identity constraints. For $f\in F_k$ and $i\in M_k$, we have

$$\partial_{x_i^{(k)}}\Gamma_f^{(k)}=\sum_{j\in f}\nabla_1(\alpha_i^{(k)}\cdot K^{(k)}(x_i^{(k)},z_j^{(k)})\tilde{N}^{(n)}(f)). \tag{16}$$

Denoting

$$\delta^{(k)}(f):=\sum_{j\in f}(u_k(z_j^{(k)})-u_n(z_j^{(k)})),$$

if $i\in f\in F_k$, then

$$\partial_{z_i^{(k)}}\Gamma_f^{(k)}=-e_{i,f}\times\delta^{(k)}(f)-\sum_{l=1}^{n-1}\sum_{j=1}^{m_l}\nabla_1(\tilde{N}^{(n)}(f)\cdot K^{(n)}(z_i^{(k)},z_j^{(l)})\beta_j^{(l)}) \tag{17}$$
$$+\sum_{j=1}^{m_k}\nabla_1(\tilde{N}^{(n)}(f)\cdot K^{(k)}(z_i^{(k)},x_j^{(k)})\alpha_j^{(k)})-\sum_{l=1}^{n-1}\sum_{j=1}^{m_l}\nabla_1(\beta_i^{(k)}\cdot K^{(n)}(z_i^{(k)},z_j^{(l)})\tilde{N}^{(n)}(f)).$$

Let $p^{x,1},\dots,p^{x,n-1}$ and $p^z=(p^{z,1},\dots,p^{z,n-1})$ be the co-states. Let $\gamma_f^{(k)}=\lambda_f-\mu\Gamma_f^{(k)}$. For $i\in M_k$, let $F_k(i)=\{f\in M_k : i\in f\}$. Then

$$\partial_t p_i^{x,k}=-\sum_{j=1}^{N_k}\nabla_1(p_i^{x,k}\cdot K^{(k)}(x_i^{(k)},x_j^{(k)})\alpha_j^{(k)})-\sum_{j=1}^{N_k}\nabla_1(\alpha_i^{(k)}\cdot K^{(k)}(x_i^{(k)},x_j^{(k)})p_j^{x,k})$$
$$+2\sum_{j=1}^{N_k}\nabla_1(\alpha_i^{(k)}\cdot K^{(k)}(x_i^{(k)},x_j^{(k)})\alpha_j^{(k)})-\sum_{f\in F_k}\gamma_f^{(k)}\partial_{x_i^{(k)}}\Gamma_f^{(k)},$$

and

$$\partial_t p_i^{z,k} = -\sum_{j=1}^N \nabla_1(p_i^{z,i}\cdot K^{(n)}(z_i^{(k)},z_j)\beta_j) - \sum_{j=1}^N \nabla_1(\beta_i^{(k)}\cdot K^{(n)}(z_i^{(k)},z_j)p_j^z) + 2\sum_{j=1}^N \nabla_1(\beta_i^{(k)}\cdot K^{(n)}(z_i^{(k)},z_j)\beta_j) - \sum_{f\in F_k}\gamma_f^{(k)}\partial_{z_i^{(k)}}\Gamma_f^{(k)},$$

where $\partial_{x_i^{(k)}}\Gamma_f^{(k)}$ and $\partial_{z_i^{(k)}}\Gamma_f^{(k)}$ are given by (16) and (17).

For $f\in F_k$ ($k=1,\dots,n-1$), we have

$$\partial_{\alpha^{(k)}}\Gamma_f^{(k)} = \sum_{j\in f} K^{(k)}(x^{(k)},z_j^{(k)})\tilde{N}_j^{(n)}(f),$$
$$\partial_{\beta}\Gamma_f^{(k)} = -\sum_{j\in f} K^{(n)}(z,z_j^{(k)})\tilde{N}_j^{(n)}(f),$$

Letting $\theta_j^{(k)} = \sum_{f\in F_k: j\in f}\gamma_f^{(k)}N_j^{(n)}(f)$, the gradient of L in α and in β is then given by

$$\nabla_{\alpha^{(k)}}L = K^{(k)}(x^{(k)})(\alpha^{(k)}-p^{x,k}) - K^{(k)}(x^{(k)},z)\theta,$$
$$\nabla_{\beta}L = K^{(n)}(z)(\beta-p^z) + K^{(n)}(z,z)\theta$$

or, taking the Hilbert gradient,

$$\nabla_{\alpha^{(k)}}L = \alpha^{(k)} - p^{x,k} - K^{(k)}(x^{(k)})^{-1}K^{(k)}(x^{(k)},z)\theta,$$
$$\nabla_{\beta}L = \beta - p^z + K^{(n)}(z,z)\theta.$$

In spite of it requiring the inversion of a linear system in the first equation, we found the latter version preferable to the L^2 gradient in our experiments.

4.4. **Remarks.**

Existence of constrained solutions. It is important to note that, according to [2, Theorem 1], there always exists at least one solution of (12)-(13) satisfying the constraints.

Convergence to surfaces. A question naturally arising is whether our discrete approximation using triangulations converges to the smooth setting as triangles get smaller and smaller. More precisely, assume that smooth initial surfaces $S_{init}^k = q_{init}^{(k)}(M_k)$ are triangulated, with increasingly fine triangulations $q_{init}^{(k,\ell)} : M_{k,\ell}\to\mathbb{R}^3$, $\ell=1,2,\dots$, where $M_{k,\ell}$ labels the vertices of a simplicial complex $T_{k,\ell}$ whose faces are $F_{k,\ell}$. We discuss whether minimizers $(u_{k,\ell}, k=1,\dots,n)$ of the discrete problems have a subsequence that converges to a minimizer $(u_k, k=1,\dots,n)$ of the limit problem.

Assume that the following condition holds for the sequence of triangulations:

(i) We assume that for all k and ℓ, and for every $f \in F_{k,\ell}$, there exists an embedding $\psi^f_{k,\ell} : t^f_{k,\ell} \to S^{(k)}_{init} = q^{(k)}_{init}(M_k)$ (where $t^f_{k,\ell}$ is the interior of the triangle $q_{k,\ell}(f)$) such that $(\psi^f_{k,\ell}(t^f_{k,\ell}), f \in F_{k,\ell})$ partitions $S^{(k)}_{init}$ up to a negligible set and $\max_f \|\psi^f_{k,\ell} - \mathrm{id}_{t_{k,\ell}}\|_{1,\infty} \to 0$ when $\ell \to \infty$.

This conditions ensure that data attachment terms like those described in section 2.2.2 computed at diffeomorphic transformations $\varphi_{k,\ell} \circ q^{(k,\ell)}_{init}$ converge to the same term computed at $\varphi_k \circ q^{(k)}_{init}$ as soon as $\varphi^{(k,\ell)}$ converges to $\varphi^{(k)}$ in $C^1(\mathbb{R}^3)$. Given this, we sketch the argument leading to the consistency of the discrete approximations.

For identity constraints, one can use [2, Proposition 5], which proves that, if the triangulations are nested (every vertex at step ℓ lies on the limit surface and is also a vertex at step $\ell+1$), then one can extract, from a corresponding sequence of identity-constrained optimal vector fields, a subsequence that converges towards an identity-constrained solution of (12)-(13).

For sliding constraints, one cannot directly apply [2, Proposition 5], because the constraints are not nested, even when the triangulations are. To obtain a consistent approximation, we need to relax the discrete problems. More precisely, let $t \mapsto (u_1(t), \dots, u_n(t)) \in V_1 \times \cdots \times V_n$ be a minimizer of the continuous problem with sliding constraints, and let $(\varphi_1, \dots, \varphi_n)$ denote the corresponding flow with $q^{(k)}(t) = \varphi_k(t) \circ q^{(k)}_{init}$ the corresponding deformation of $q^{(k)}_{init}$, and $N^{(k)}(t,x) = d\varphi_k(t)^T N^{(k)}_{init}(x)$ a normal to $S_k(t) = q^{(k)}(t)(M_k)$ at $q^{(k)}(t,x)$. In particular we have, for every $k = 1, \dots, n$, every $x \in M_k$, and almost every time t,

$$N^{(k)}(t,x) \cdot (u_k(t, q^{(k)}(t,x)) - u_n(t, q^{(k)}(t,x)) = 0.$$

Moreover, as $\sum_{k=1}^n \|u_k(t)\|^2$ is constant, both $u_k(t)$ and $du_k(t)$ are α-Lipschitz for some positive constant α that does not depend on t or k.

Now let $q^{(k,\ell)}(t) = \varphi_k(t) \circ q^{(k,\ell)}_{init}$ be the corresponding deformation of the discretization at step ℓ. Recall that $N^{(k,\ell)}(t,f)$ denotes the unit normal to the triangle $q^{(k,\ell)}(t,f)$. We will prove:

(ii) The discretized deformations at step ℓ satisfy the following relaxed sliding constraints

$$\left| N^{(k,\ell)}(f) \cdot \left(\sum_{j \in f} (u_k(z^{(k,\ell)}_j) - u_n(z^{(k,\ell)}_j) \right) \right| \leq \varepsilon_\ell$$

at almost every t and for every face f in $F_{k,\ell}$, for a suitably chosen sequence $\varepsilon_\ell > 0$ going to 0 as ℓ goes to ∞.

Indeed, fix a face f and an integer ℓ. Define

$$\bar{N}^{(k,\ell)}(t,f) = \frac{d\varphi_k(t)^T N^{(k,\ell)}(0,f)}{|d\varphi_k(t)^T N^{(k,\ell)}(0,f)|}$$

for every time t. Note that assumption (i) implies that

$$|N^{(k,\ell)}(t,f) - \bar{N}^{(k,\ell)}(t,f)| \leq \gamma_\ell,$$

for some sequence γ_ℓ, independent of f and t, and going to 0 as ℓ goes to infinity. Assumption (i) also implies that there exists a sequence η_ℓ, independent of f and going to 0 as $\ell \to \infty$ such that for $y \in \psi^f_{k,\ell}(f)$, and $q^{(k)}_{init}(x) = y$, we have

$$|N^{(k)}_{init}(x) - N^{(k,\ell)}(0, f)| + \sum_{j \in f} |q^{(k)}_{init}(x) - z^{(k,\ell)}_j(0)| \leq \eta_\ell.$$

Some triangle inequalities, Gronwall's lemma, and the fact that u_k and du_k are $\alpha-$Lipschitz then imply

$$\left| N^{(k,\ell)}(t, f) \cdot \left(\sum_{j \in f} (u_k(t, z^{(k,\ell)}_j) - u_n(t, z^{(k,\ell)}_j) \right) \right| \leq \varepsilon_\ell,$$

with $\varepsilon_\ell = \alpha\eta_\ell e^\alpha + \gamma_\ell$ going to 0 as ℓ goes to infinity.

Consequently, if $(u_{1,\ell}, \dots, u_{k,\ell})$ is a sequence of minimizers of the discretized problem at step ℓ with relaxed sliding constraints

$$\left| N^{(k,\ell)}(t, f) \cdot \left(\sum_{j \in f} (u_{k,\ell}(t, z^{(k,\ell)}_j(t)) - u_{n,\ell}(t, z^{(k,\ell)}_j(t)) \right) \right| \leq \varepsilon_\ell,$$

we see that the infimum limit over ℓ of the respective discretized costs of $(u_{1,\ell}, \dots, u_{n,\ell})$ is smaller than or equal to the cost of a minimizer of the continuous problem with sliding constraints. So to prove that a limit point of that sequence is a minimizer of the cost for the continuous problem with sliding constraints, all we need is to check that any such limit point does satisfy the constraints.

So let $(u_{1,\ell}, \dots, u_{n,\ell})$ be a sequence of minimizers of the relaxed discrete problem at step ℓ[1] that weakly converges to $(u_1, \dots, u_n)$ in $V_1 \times \cdots \times V_n$ (which is true for at least one subsequence of any minimizing sequence), then the associated flows $\varphi_{k,\ell}$ and their first two space derivatives converge uniformly in time and space to the flows φ_k and their first two derivatives.

From this it is easy to see that any sequence approximating $q^{(k)}_{init}(x)$ as in assumption (i) is such that $q^{(k,\ell)}(t, j_\ell) \to q^{(k)}(t, x)$ and $N^{(k,\ell)}(t, j_\ell) \to N^{(k)}(t, x)$ at all times. Moreover, such a minimizing sequence must be, like all geodesics, such that $\sum_{k=1}^n \|u_{k,\ell}(t)\|^2_{V_k}$ is constant in time, and smaller than the cost function associated to, say, $u_k = 0$ for all k. This implies that the vector fields $u_{k,\ell}(t)$ are continuous uniformly in k, ℓ and t, which, combined with the continuity of the evaluation functionals in an RKHS implies that $u_{k,\ell}(q^{(k,\ell)}(t, j_\ell)) \to u_k(q^{(k)}(t, x))$ for all times. Consequently, one easily checks that each $(u_{1,\ell}, \dots, u_{n,\ell})$ satisfy a relaxed version of the continuous version of the constraints, with a precision that goes to 0 as ℓ goes to infinity. This finishes proving that the constraints are satisfied with exactitude at the limit.

Kernel derivatives. Expressions similar to $\nabla_1(n \cdot K(x, y)\alpha)$ appear at multiple times in the previous computation (for some vectors n and α). For radial kernels ($K(x, y) = G(|x - y|^2)\mathrm{Id}_{\mathbb{R}^d}$), we have

$$\nabla_1(n \cdot K(x, y)\alpha) = 2G'(|x - y|^2)(n \cdot \alpha)(x - y),$$

[1]It is easy to prove that such minimizers exist using the same method as that of [2], and replacing equality constraints with inequality constraints.

which (slightly) simplifies the expressions.

Sliding Interface – Alternate Version. As discussed in Section 3, the sliding constraint can also be handled by introducing a new state variable N that tracks a vector (or frame) normal to the interface via (11). In the discrete case, one can discretize this equation by introducing states $N(f), f \in F$, indexed by the facets of M, and evolving according to

$$\partial_t N(f) = -\frac{1}{|f|} \sum_{i \in f} du^{(n)}(z_i)^T N(f),$$

where $|f|$ is the number of vertices in f. The sliding constraints are now expressed in terms of the state variables in a more direct way, but with a new co-state variable for the normals, bringing in an extra degree of complexity and increasing the computational cost. Note that the finite-dimensional reduction is still possible in this case, so that $u^{(n)}(\cdot) = K^{(n)}(\cdot, z)\beta$, and the evolution of the normals can be expressed in a form involving the differential of the kernel. This yields an adjoint system involving second derivatives of the kernel. We will not detail the computations in this paper, since they follow the same pattern as the other two that were already discussed (see [56] for more examples on how higher-order variables can be handled in similar contexts). Note that this alternate version of the sliding constraints is slightly more general than the one discussed in the previous section, since it does not require a definition of a normal field that smoothly depends on the manifolds.

5. Experimental Results

5.1. **Synthetic Example.** The first example is described in Figure 1. In this synthetic example, the template has two identical balls initially close to each other. In the target, the first ball (referred to as "Ball A") gets bigger, and "impacts" the other one (referred to as "Ball B"), which assumes an oblong, non-convex shape (the target shapes slightly overlap, so that an exact homeomorphic match cannot be achieved).

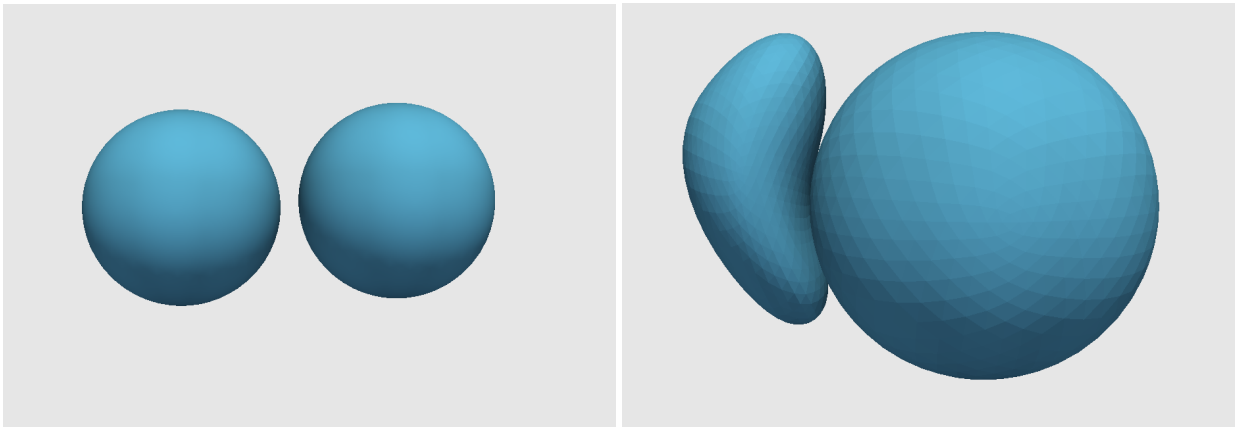

FIGURE 1. Template and target shapes for synthetic example

Our results, provided in Figures 2 to 5, illustrate our multishape deformation method, and use two complementary deformation indexes:

(i) the tangent Jacobian, which is the Jacobian determinant of the surface-to-surface transformations, and which measures the ratio between the areas of elementary surface patches at each point before and after deformation;
(ii) the normal Jacobian, which is the ratio of the Jacobian determinant (of the 3D diffeomorphism) to the tangent Jacobian, and which measures the ratio between the length of an infinitesimal line element normal to the surface after and before transformation.

These indexes are mapped on the deformed template image, which is close to the target.

Figure 2 compares the normal Jacobian of the shape and background deformations when using identity constraints. While shape diffeomorphisms characterize each shape transformation (uniform variation for Ball A, expansion at the top and compression otherwise for Ball B), the effect of compressing the space is clearly visible in the background deformation, when the two shapes get close to each other.

Figure 3 provides the corresponding tangent Jacobian, which is identical for shape and background transformation since we are using identity constraints.

Figures 4 and 5 compare the normal and tangent Jacobians for the synthetic experiment with sliding constraints. Regarding the former, the most notable difference is with Ball B, which shows an expansion pattern at the tips in its shape diffeomorphism which is inverse of the one observed with identity constraints. One plausible explanation is that sliding constraints allow the two shapes to use translation-like motion to position themselves differently, without the need for limiting the amount of shear in the background that would have resulted from identity constraints. The second notable difference can be noted in the background diffeomorphism, in which compression is mostly observed with Ball B. In contrast with the identity constraints, the tangent Jacobians are very different between shape and background diffeomorphisms. Note that Figure 5 uses two different color scales for the left and right panels because of the strong difference between the ranges of the Jacobians in each case. The background deformation, in particular, has a huge tangent expansion around the impact location, which cannot be observed in the shape deformations. Note that both patterns in the sliding case are very different from the one that was observed in the identity case.

For comparison purposes, Figure 6 provides the result of the LDDMM algorithm using a single diffeomorphism. One observes a very strong compression effect for the normal jacobian resulting in an expansion observed on the tangent jacobian on Ball B, that was not observed in any of the constrained examples. The nice uniform expansion in Ball A that could be observed in the sliding constraint case is not observed either.

5.2. **Subcortical Structures.** We now describe an example mapping a group of three subcortical structures: hippocampus, amygdala and entorhinal cortex (ERC). The template and target sets

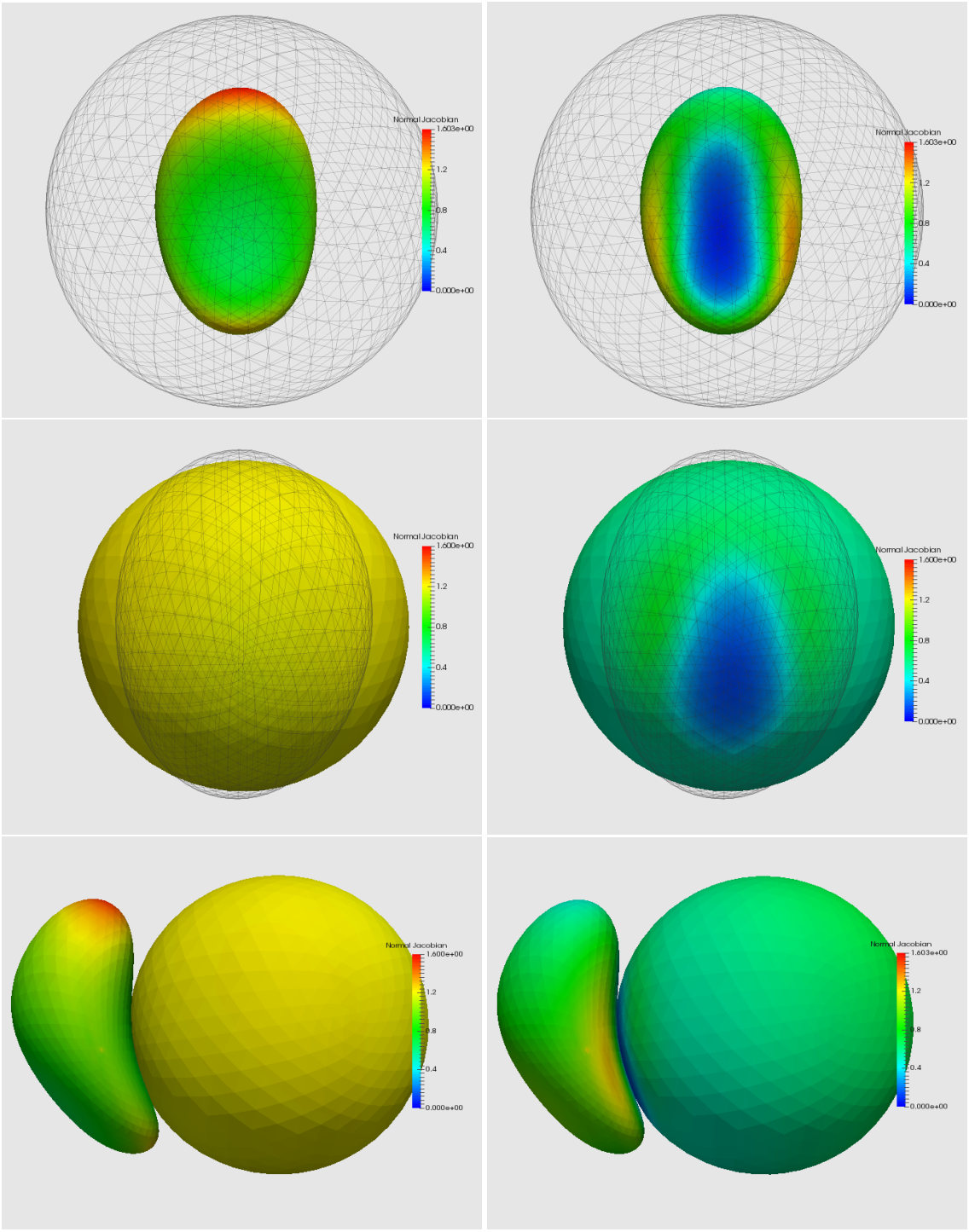

FIGURE 2. Three views of the normal Jacobian with identity constraints: shape diffeomorphisms (left) and background diffeomorphism (right).

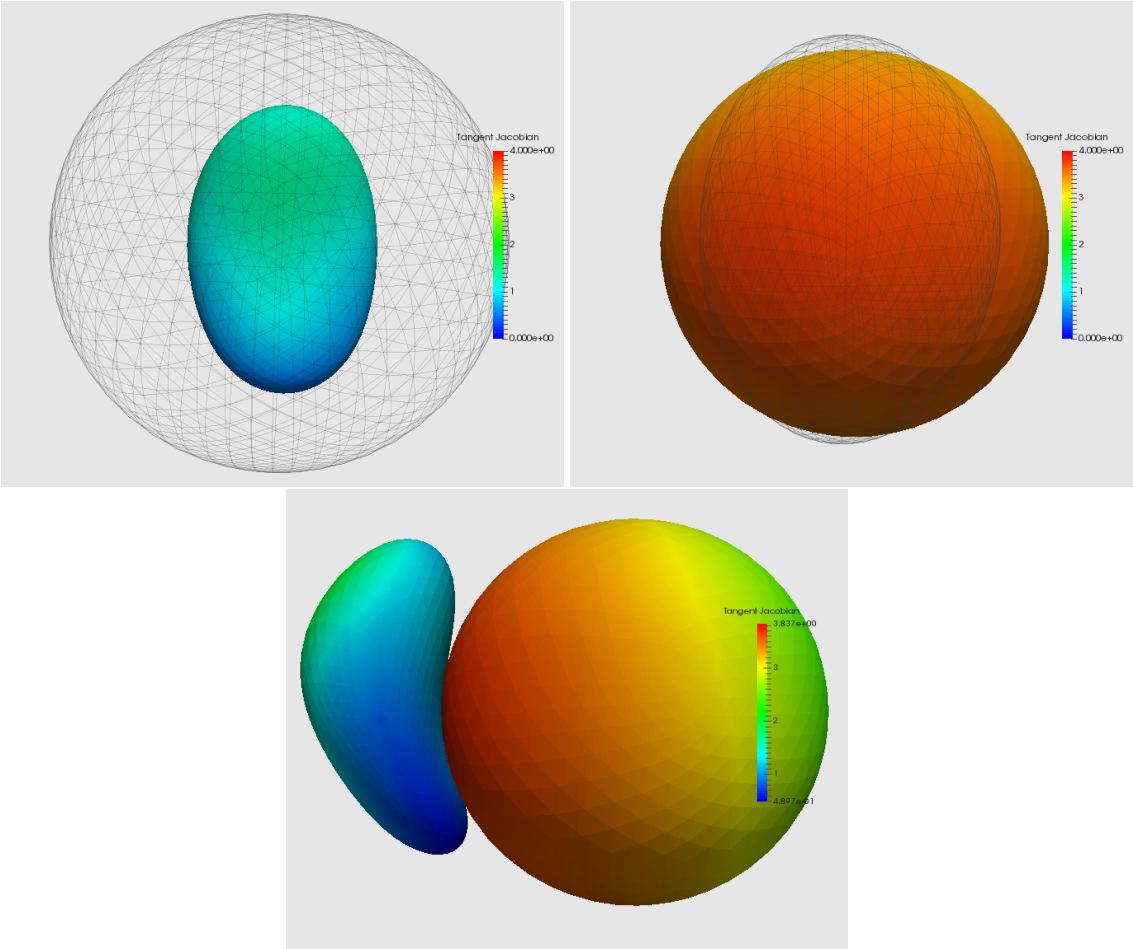

FIGURE 3. Tangential Jacobian: shape and background diffeomorphisms (identity constraints).

are represented in Figure 7. One can observe shape changes in each structure, combined with a significant displacement of the ERC relative to the other two structures when comparing template to target. Because the structures were segmented independently, there is some overlap between the target hippocampus and amygdala.

Figures 8 and 9 provide the normal and tangent Jacobian obtained with identity constraints, while Figures 10 and 11 provide this information for sliding constraints. The two types of constraints provide similar deformation indices, especially for the normal jacobians (Figures 8 and

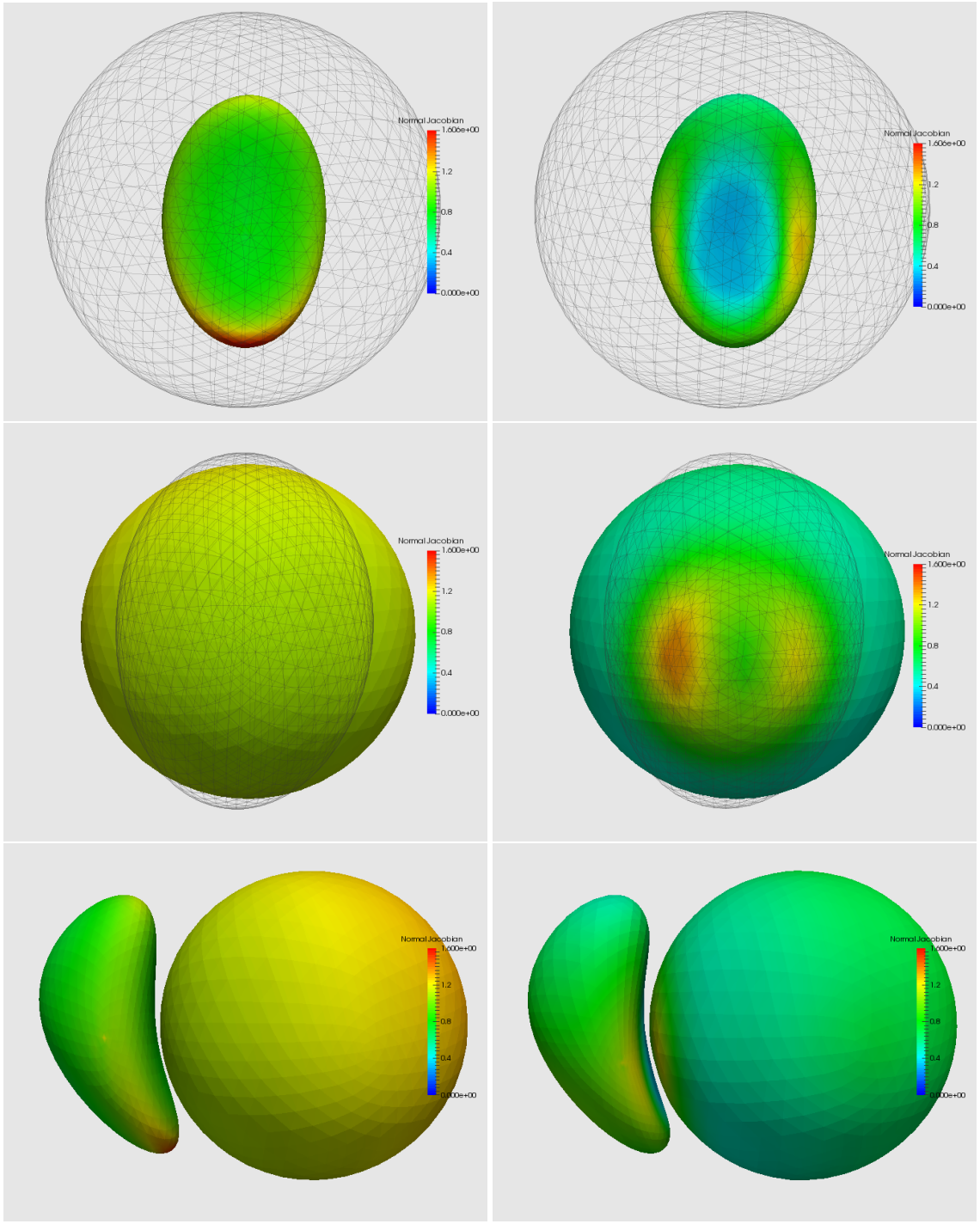

FIGURE 4. Three views of the normal Jacobian with sliding constraints: shape diffeomorphisms (left) and background diffeomorphism (right).

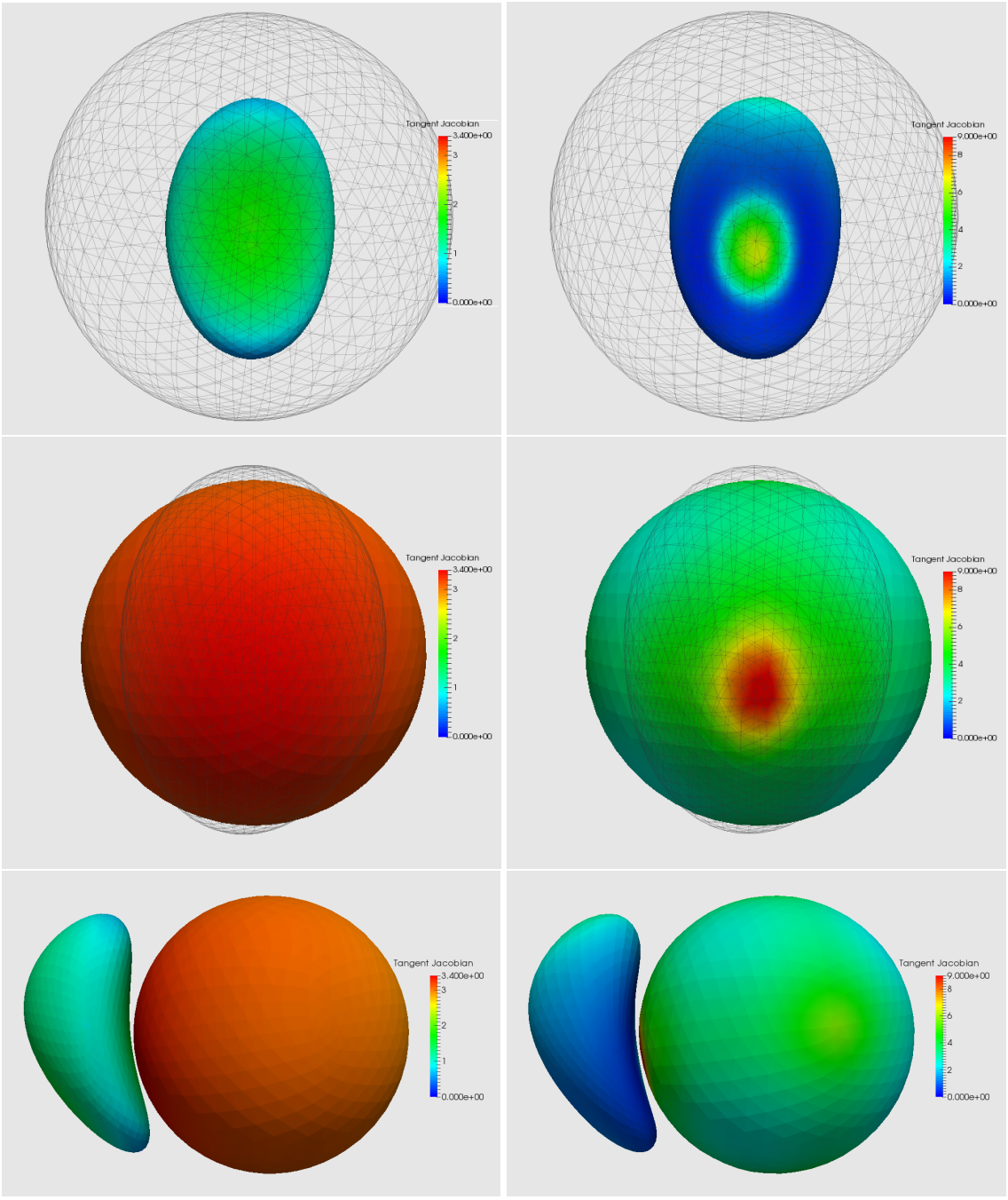

FIGURE 5. Three views of the tangent Jacobian with sliding constraints: shape diffeomorphisms (left) and background diffeomorphism (right).

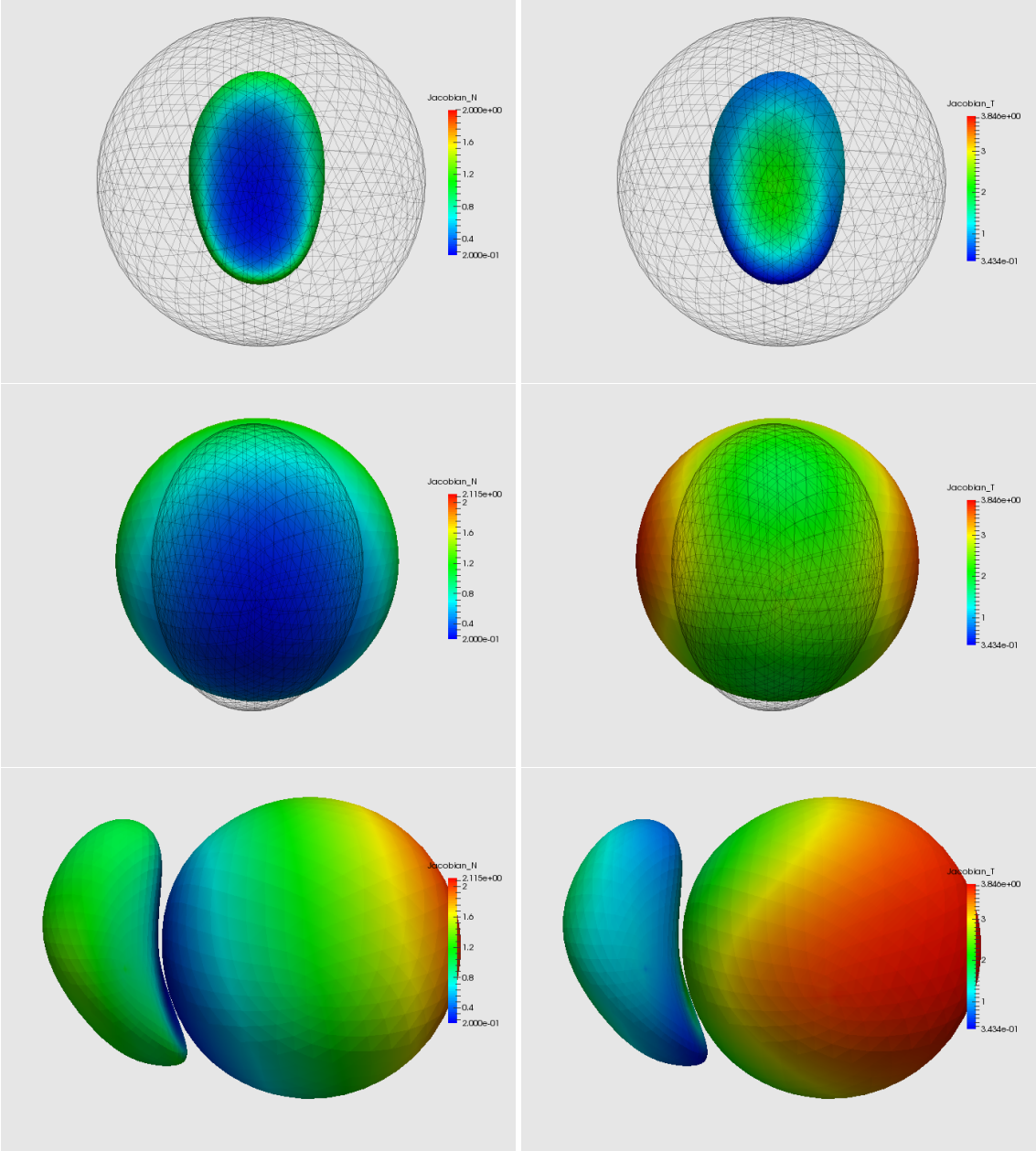

FIGURE 6. Three views of the Normal (left) and tangential Jacobians (right) when using a single diffeomorphism.

10). Minor differences in the tangent jacobian can be observed (Figures 9 and 11). The deformation patterns associated to using a single diffeomorphism (Figure 12) are significantly different, though, exhibiting very strong compression, for example, where shapes are close to each other.

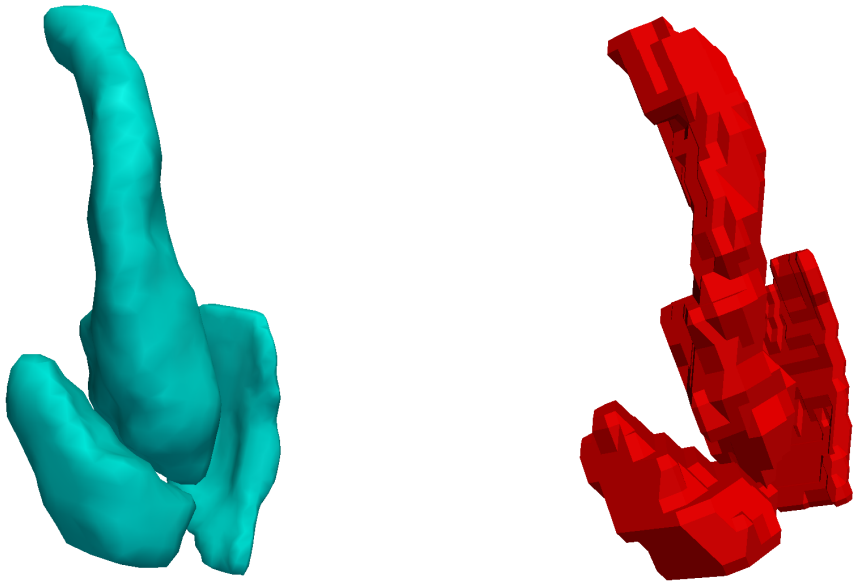

FIGURE 7. Template (blue) and target (red) shapes for subcortical structures. The hippocampus is the central shape, with the amygdala on its left and the ERC on its right.

6. DISCUSSION

The previous approach provides a solution, using constrained optimal control, of the important issue of dealing with multiple objects with varying deformation properties for registration. We have focused on surface matching, numerically dealing with constraints using an augmented Lagrangian method. Note that a similar approach was introduced for plane curves in [2].

The formulation is quite general and can accommodate constraints in various forms, including the examples discussed in Section 3. The investigation of these additional applications will be the subject of future work. One of the limitations of the present implementation is the slow convergence of the augmented Lagrangian procedure, for which each minimization step is, in addition, high dimensional and computationally demanding. One possible alternative can be based on solving the optimality conditions (14) (which hold in the discrete case) by means of a numerical shooting method. This approach has, however, its own numerical challenges, because solving (14) requires the determination of λ such that the last equation (constraint) is satisfied, and this leads to a possibly ill-posed problem for systems in large dimension (see [2] for additional details).

We have illustrated our examples using deformation markers derived from the jacobian determinant. This markers are routinely used in shape analysis studies and led to important conclusion in computational anatomy. When dealing with multiple shapes, however, figures 6 and 12 show that, when using the classical LDDMM method with multiple shapes, these markers becomes as much, if not more, influenced by interactions between the shapes as by the changes in the shapes

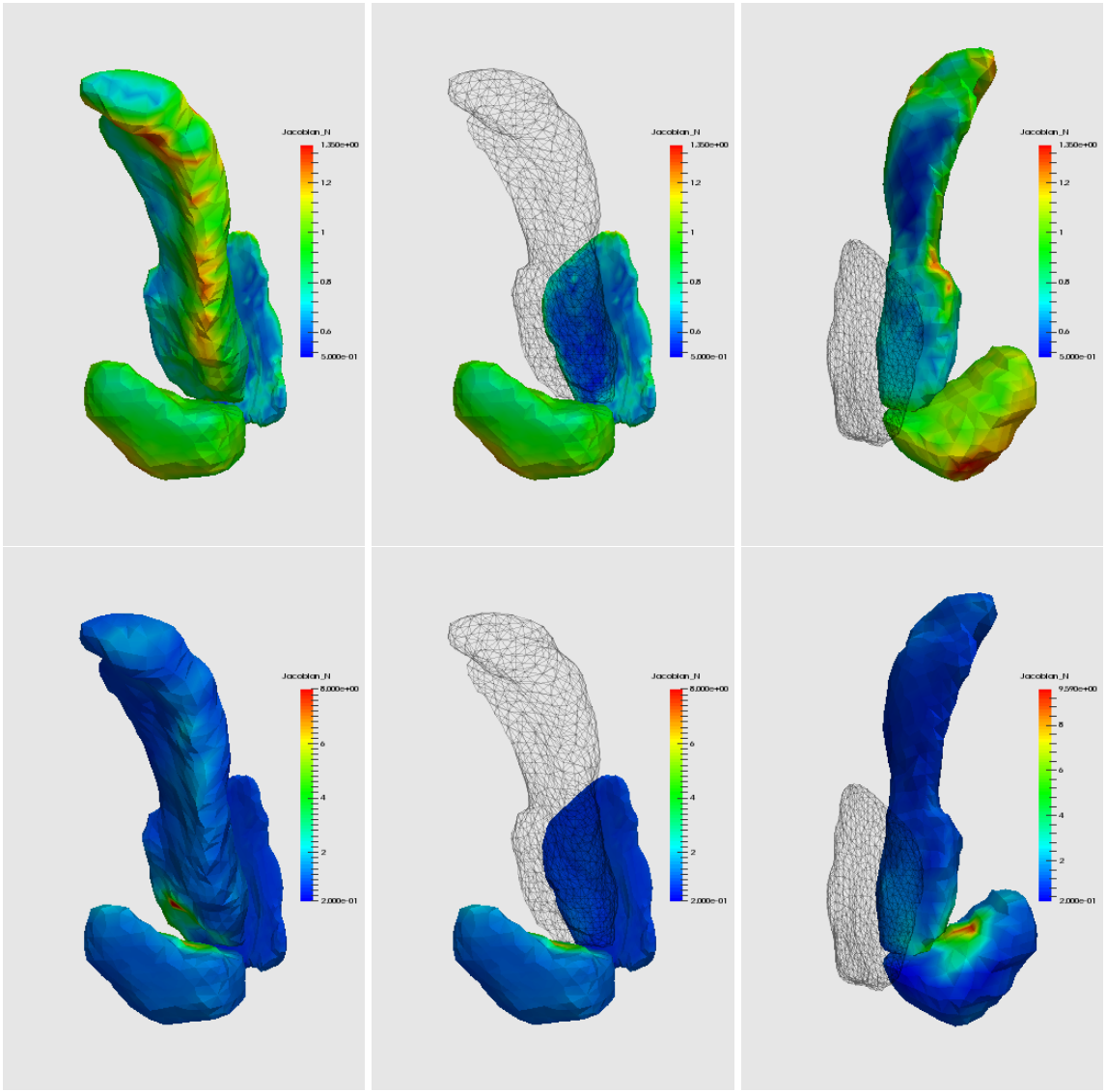

FIGURE 8. Three views of the normal Jacobian with identity constraints: First Row: Shape diffeomorphisms; Second Row: Background diffeomorphism.

themselves. For this reason, multi-shape computational anatomy studies have applied registration methods separately to each shape, without ensuring that the obtained diffeomorphisms are consistent with each other. This limitation is addressed in the present paper, in which we exhibit deformation markers that are meaningful in describing tangential and normal surface stretching, while being consistently associated to a global transformation of the space.

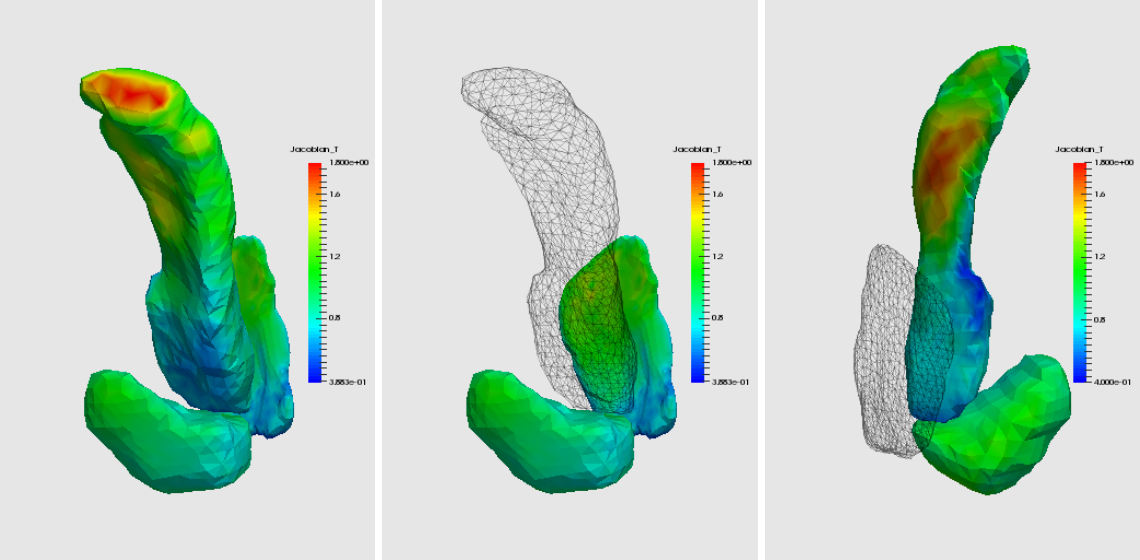

FIGURE 9. Three views of the tangent Jacobian with identity constraints: Shape and background diffeomorphisms.

REFERENCES

[1] Siamak Ardekani, Robert G Weiss, Albert C Lardo, Richard T George, Joao AC Lima, Katherine C Wu, Michael I Miller, Raimond L Winslow, and Laurent Younes. Cardiac motion analysis in ischemic and non-ischemic cardiomyopathy using parallel transport. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 899–902. IEEE, 2009.

[2] Sylvain Arguillère, Emmanuel Trélat, Alain Trouvé, and Laurent Younes. Shape deformation analysis from the optimal control viewpoint. *preprint arXiv:1401.0661, to appear in J. Math. Pures Appl.*, 2014.

[3] John Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007.

[4] John Ashburner and Karl J Friston. Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation. *Neuroimage*, 55(3):954–967, 2011.

[5] Brian B Avants, P Thomas Schoenemann, and James C Gee. Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. *Medical image analysis*, 10(3):397–412, 2006.

[6] 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.

[7] Ruzena Bajcsy, Robert Lieberson, and Martin Reivich. A computerized system for the elastic matching of deformed radiographic images to idealized atlas images. *Journal of Computer Assisted Tomography*, 7(4):618–625, 1983.

[8] M Faisal Beg, Michael I Miller, Alain Trouvé, and Laurent Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.

[9] Martins Bruveris, François Gay-Balmaz, Darryl D Holm, and Tudor S Ratiu. The momentum map representation of images. *Journal of nonlinear science*, 21(1):115–150, 2011.

[10] Yan Cao, Michael I Miller, Raimond L Winslow, and Laurent Younes. Large deformation diffeomorphic metric mapping of vector fields. *IEEE Transactions on Medical Imaging*, 24(9):1216–1230, 2005.

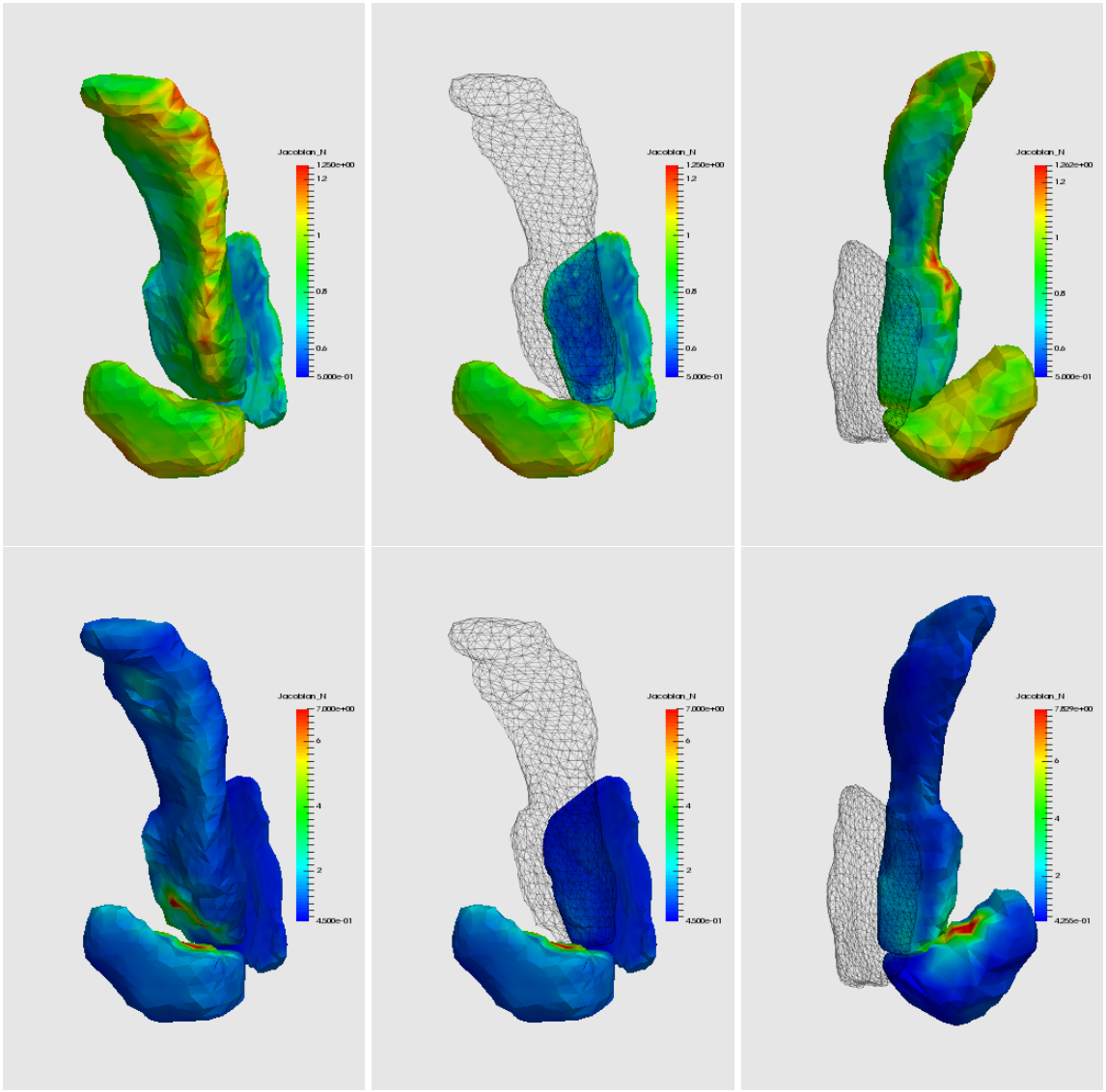

FIGURE 10. Three views of the normal Jacobian with sliding constraints: First Row: Shape diffeomorphisms; Second Row: Background diffeomorphism.

[11] Can Ceritoglu, Kenichi Oishi, Xin Li, Ming-Chung Chou, Laurent Younes, Marilyn Albert, Constantine Lyketsos, Peter van Zijl, Michael I Miller, and Susumu Mori. Multi-contrast large deformation diffeomorphic metric mapping for diffusion tensor imaging. *Neuroimage*, 47(2):618–627, 2009.

[12] Gary E. Christensen, Richard D. Rabbitt, and Michael I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, 1996.

[13] 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.

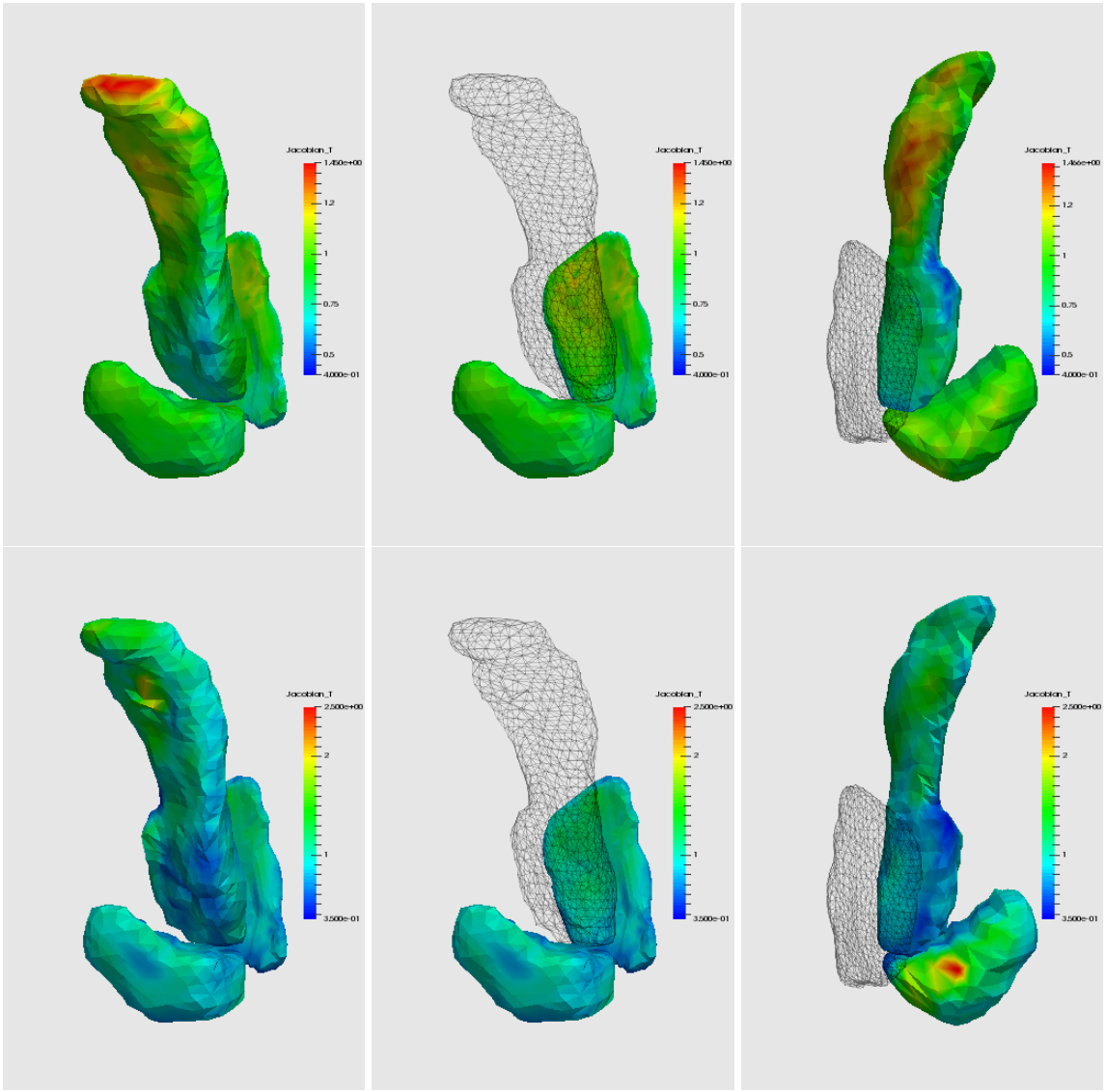

FIGURE 11. Three views of the tangent Jacobian with sliding constraints: First Row: Shape diffeomorphisms; Second Row: Background diffeomorphism.

[14] Colin J Cotter and Darryl D Holm. Singular solutions, momentum maps and computational anatomy. *arXiv preprint nlin/0605020*, 2006.

[15] WR Crum, T Hartkens, and DLG Hill. Non-rigid image registration: theory and practice. *British Journal of Radiology*, 77(suppl 2):S140–S153, 2004.

[16] Marc Droske and Martin Rumpf. A variational approach to nonrigid morphological image registration. *SIAM Journal on Applied Mathematics*, 64(2):668–687, 2004.

[17] 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.

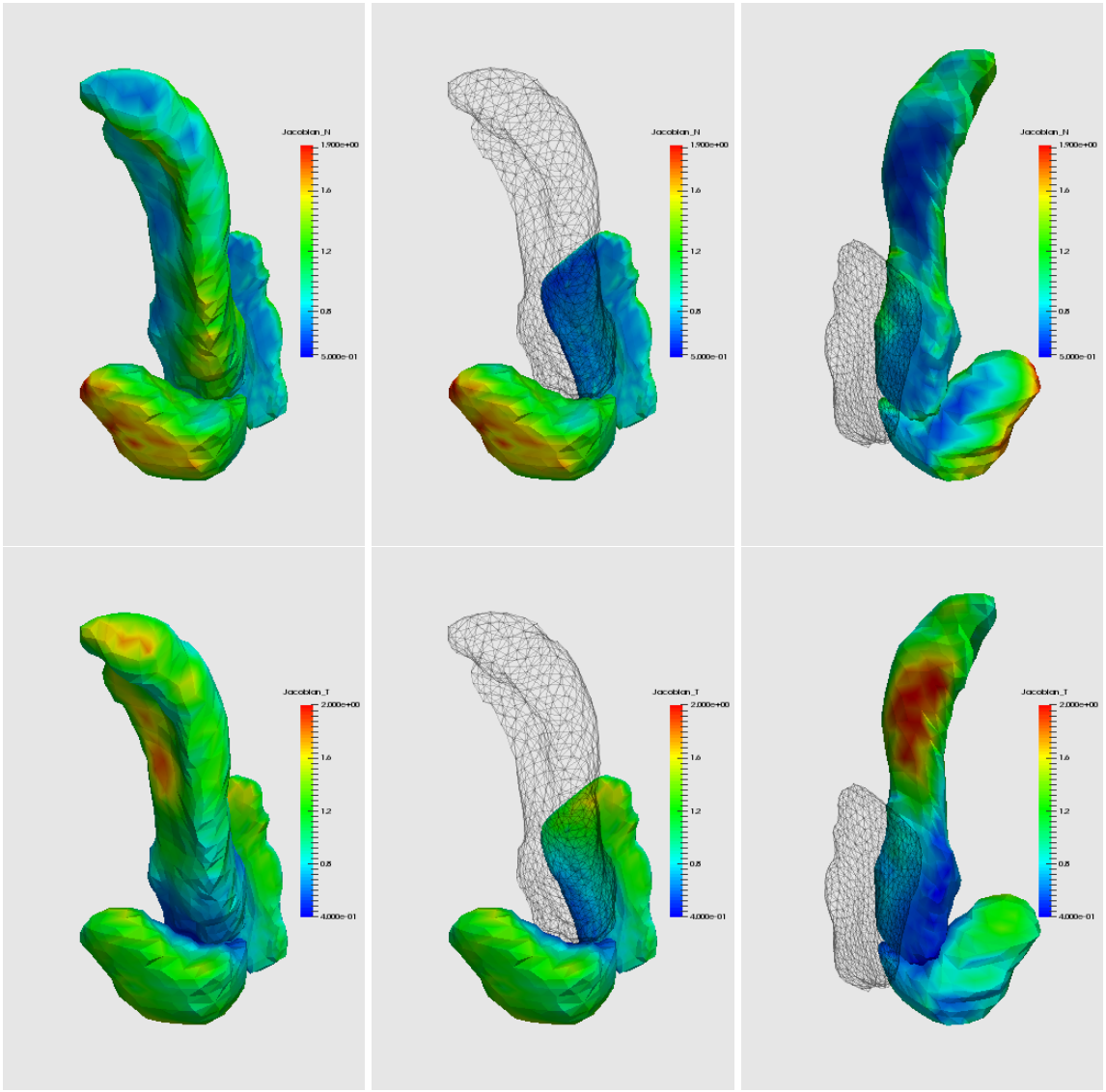

FIGURE 12. Three views of the normal (up) and tangent Jacobians (down) when using a single diffeomorphism.

[18] Stanley Durrleman, Stéphanie Allassonnière, and Sarang Joshi. Sparse adaptive parameterization of variability in image ensembles. *International Journal of Computer Vision*, 101(1):161–183, 2013.

[19] Stanley Durrleman, Xavier Pennec, Alain Trouvé, Guido Gerig, and Nicholas Ayache. Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 297–304. Springer, 2009.

[20] Herbert Federer and Herbert Federer. *Geometric measure theory*, volume 1996. Springer New York, 1969.

[21] 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.

FIGURE 13. Midpoint of the optimal deformation with multishape identity constraints (left), multishape sliding constraints (center) and single diffeomorphism (right).

[22] Joan Glaunès, Marc Vaillant, and Michael I Miller. Landmark Matching via Large Deformation Diffeomorphisms on the Sphere. *Journal of Mathematical Imaging and Vision*, 20:179–200, 2004.

[23] Ben Glocker, Nikos Komodakis, Georgios Tziritas, Nassir Navab, and Nikos Paragios. Dense image registration through mrfs and efficient linear programming. *Medical image analysis*, 12(6):731–741, 2008.

[24] A Ardeshir Goshtasby. *Image Registration: Principles, Tools and Methods*. Springer, 2012.

[25] Ulf Grenander. *General pattern theory: A mathematical study of regular structures*. Clarendon Press Oxford, 1993.

[26] 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. In *Information Processing in Medical Imaging*, pages 172–184. Springer, 2003.

[27] 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. *IEEE Transactions on Medical Imaging*, 23(8):949–958, 2004.

[28] Andreas Günther, Hans Lamecker, and Martin Weiser. Flexible shape matching with finite element based lddmm. *International Journal of Computer Vision*, pages 1–16, 2012.

[29] Andreas Günther, Hans Lamecker, Martin Weiser, et al. Direct lddmm of discrete currents with adaptive finite elements. In *Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy-Geometrical and Statistical Methods for Modelling Biological Shape Variability*, pages 1–14, 2011.

[30] Eldad Haber, Gallagher Pryor, John Melonakos, Allen Tannenbaum, et al. 3d nonrigid registration via optimal mass transport on the gpu. *Medical image analysis*, 13(6):931–940, 2009.

[31] Steven Haker, Sigurd Angenent, Allen Tannenbaum, Ron Kikinis, Guillermo Sapiro, and Michael Halle. Conformal surface parameterization for texture mapping. *IEEE Transactions on Visualization and Computer Graphics*, 6(2):181–189, 2000.

[32] Steven Haker, Lei Zhu, Allen Tannenbaum, and Sigurd Angenent. Optimal mass transport for registration and warping. *International Journal of Computer Vision*, 60(3):225–240, 2004.

[33] Monica K Hurdal, Philip L Bowers, Ken Stephenson, L Sumners De Witt, Kelly Rehm, Kirt Schaper, and David A Rottenberg. Quasi-conformally flat mapping the human cerebellum. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI99*, pages 279–286. Springer, 1999.

[34] Monica K Hurdal and Ken Stephenson. Cortical cartography using the discrete conformal approach of circle packings. *Neuroimage*, 23:S119–S128, 2004.

[35] Monica K Hurdal and Ken Stephenson. Discrete conformal methods for cortical brain flattening. *Neuroimage*, 45(1):S86–S98, 2009.

[36] Monica K Hurdal, Ken Stephenson, Phil Bowers, De Witt Sumners, and David A Rottenberg. Coordinate systems for conformal cerebellar flat maps. *Neuroimage*, 11(5):S467, 2000.

[37] Miao Jin, Yalin Wang, S-T Yau, and Xianfeng Gu. Optimal global conformal surface parameterization. In *Visualization, 2004. IEEE*, pages 267–274. IEEE, 2004.

[38] Sarang C. Joshi and Michael I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9:1357–1370, 2000.

[39] Yaron Lipman and Ingrid Daubechies. Surface comparison with mass transportation. *arXiv preprint arXiv:0912.3488*, 2009.

[40] Yaron Lipman and Ingrid Daubechies. Conformal wasserstein distances: Comparing surfaces in polynomial time. *Advances in Mathematics*, 227(3):1047–1077, 2011.

[41] 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.

[42] Facundo Mémoli. On the use of gromov-hausdorff distances for shape comparison. In *Eurographics symposium on point-based graphics*, pages 81–90. The Eurographics Association, 2007.

[43] Facundo Mémoli. Gromov–wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11(4):417–487, 2011.

[44] Facundo Mémoli and Guillermo Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 5(3):313–347, 2005.

[45] Michael I Miller, M Faisal Beg, Can Ceritoglu, and Craig Stark. Increasing the power of functional maps of the medial temporal lobe by using large deformation diffeomorphic metric mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9685–9690, 2005.

[46] Michael I Miller, Alain Trouvé, and Laurent Younes. On the metrics and euler-lagrange equations of computational anatomy. *Annual review of biomedical engineering*, 4(1):375–405, 2002.

[47] Michael I Miller, Alain Trouvé, and Laurent Younes. The metric spaces, euler equations, and normal geodesic image motions of computational anatomy. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 2, pages II–635. IEEE, 2003.

[48] Michael I Miller, Alain Trouvé, and Laurent Younes. Geodesic shooting for computational anatomy. *Journal of mathematical imaging and vision*, 24(2):209–228, 2006.

[49] Marc Niethammer, Yang Huang, and François-Xavier Vialard. Geodesic regression for image time-series. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 655–662. Springer, 2011.

[50] Jorge Nocedal and Stephen J Wright. *Numerical Optimization, Second Edition*. Springer New York, 2006.

[51] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt. Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962.

[52] Anqi Qiu and Michael I Miller. Cortical hemisphere registration via large deformation diffeomorphic metric curve mapping. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 186–193. Springer, 2007.

[53] Anqi Qiu, Lei Wang, Laurent Younes, Michael P Harms, J Tilak Ratnanather, Michael I Miller, and John G Csernansky. Neuroanatomical asymmetry patterns in individuals with schizophrenia and their non-psychotic siblings. *Neuroimage*, 47(4):1221–1229, 2009.

[54] Laurent Risser, F Vialard, Robin Wolz, Maria Murgasova, Darryl D Holm, and Daniel Rueckert. Simultaneous multi-scale registration using large deformation diffeomorphic metric mapping. *IEEE Transactions on Medical Imaging*, 30(10):1746–1759, 2011.

[55] Laurent Risser, François-Xavier Vialard, Habib Y. Baluwala, and Julia A. Schnabel. Piecewise-diffeomorphic image registration: Application to the motion estimation between 3d {CT} lung images with sliding conditions. *Medical Image Analysis*, 17(2):182 – 193, 2013.

[56] Stefan Sommer, Mads Nielsen, Sune Darkner, and Xavier Pennec. Higher-order momentum distributions and locally affine lddmm registration. *SIAM Journal on Imaging Sciences*, 6(1):341–367, 2013.

[57] Jean-Philippe Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical image analysis*, 2(3):243–260, 1998.

[58] Emmanuel Trélat. *Contrôle optimal*. Mathématiques Concrètes. [Concrete Mathematics]. Vuibert, Paris, 2005. Théorie & applications. [Theory and applications].

[59] Alain Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.

[60] Alain Trouvé, Stanley Durrleman, Xavier Pennec, and Nicholas Ayache. Sparse approximation of currents for statistics on curves and surfaces. In *Proceedings of MICCAI 2008*, 2008.

[61] Marc Vaillant and Joan Glaunes. Surface Matching via Currents. In *Information Processing in Medical Imaging*, pages 381–392, 2005.

[62] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Non-parametric diffeomorphic image registration with the demons algorithm. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 319–326. Springer, 2007.

[63] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Diffeomorphic demons: Efficient non-parametric image registration. *Neuroimage*, 45(1):S61–S72, 2009.

[64] François-Xavier Vialard, Laurent Risser, Daniel Rueckert, and Colin J Cotter. Diffeomorphic 3D Image Registration via Geodesic Shooting Using an Efficient Adjoint Calculation. *International Journal of Computer Vision*, pages 1–13, 2011.

[65] Camille Vidal, Joshua Hewitt, Stephanie Davis, Laurent Younes, Sanjay Jain, and Bruno Jedynak. Template registration with missing parts: Application to the segmentation of m. tuberculosis infected lungs. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 718–721. IEEE, 2009.

[66] Lei Wang, Faisal Beg, Tilak Ratnanather, Can Ceritoglu, Laurent Younes, John C Morris, John G Csernansky, and Michael I Miller. Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the alzheimer type. *IEEE Transactions on Medical Imaging*, 26(4):462–470, 2007.

[67] Yalin Wang, Wei Dai, Xianfeng Gu, Tony F Chan, Shing-Tung Yau, Arthur W Toga, and Paul M Thompson. Teichmüller shape space theory and its application to brain morphometry. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 133–140. Springer, 2009.

[68] Yalin Wang, Xianfeng Gu, Tony F Chan, Paul M Thompson, and Shing-Tung Yau. Intrinsic brain surface conformal mapping using a variational method. In *Medical Imaging 2004*, pages 241–252. International Society for Optics and Photonics, 2004.

[69] Medha V Wyawahare, Pradeep M Patil, Hemant K Abhyankar, et al. Image registration techniques: an overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(3):11–28, 2009.

[70] Laurent Younes. Jacobi fields in groups of diffeomorphisms and applications. *Quarterly of applied mathematics*, 65(1):113–134, 2007.

[71] Laurent Younes. *Shapes and diffeomorphisms*, volume 171. Springer, 2010.

[72] Laurent Younes. Spaces and manifolds of shapes in computer vision: An overview. *Image and Vision Computing*, 30(6):389–397, 2012.

[73] Laurent Younes, Felipe Arrate, and Michael I Miller. Evolutions equations in computational anatomy. *Neuroimage*, 45(1):S40–S50, 2009.

[74] 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.

[75] 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*, pages 2457–2464. IEEE, 2011.

[76] 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*, pages 382–389. IEEE, 2010.

[77] Jiangyang Zhang, Linda J Richards, Michael I Miller, Paul Yarowsky, Peter van Zijl, and Susumu Mori. Characterization of mouse brain and its development using diffusion tensor imaging and computational techniques. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 2252–2255. IEEE, 2006.

[78] Darko Zikic, Ben Glocker, Oliver Kutter, Martin Groher, Nikos Komodakis, Ali Kamen, Nikos Paragios, and Nassir Navab. Linear intensity-based image registration by markov random fields and discrete optimization. *Medical image analysis*, 14(4):550–562, 2010.

[79] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

S. ARGUILLÈRE: CENTER FOR IMAGING SCIENCE AND DEPARTMENT OF APPLIED MATHEMATICS AND STATISTICS, JOHNS HOPKINS UNIVERSITY, 3400 N. CHARLES ST. BALTIMORE MD 21218

E-mail address: sarguillere@gmail.com

E. TRÉLAT: SORBONNE UNIVERSITÉS, UPMC UNIV PARIS 06, CNRS UMR 7598, LABORATOIRE JACQUES-LOUIS LIONS, AND INSTITUT UNIVERSITAIRE DE FRANCE, F-75005, PARIS, FRANCE.

E-mail address: emmanuel.trelat@upmc.fr

A. TROUVÉ: ECOLE NORMALE SUPÉRIEURE DE CACHAN, CENTRE DE MATHÉMATIQUES ET LEURS APPLICATIONS, CMLA, 61 AV. DU PDT WILSON, F-94235 CACHAN CEDEX, FRANCE

E-mail address: trouve@cmla.ens-cachan.fr

L. YOUNES: CENTER FOR IMAGING SCIENCE AND DEPARTMENT OF APPLIED MATHEMATICS AND STATISTICS, JOHNS HOPKINS UNIVERSITY, 3400 N. CHARLES ST. BALTIMORE MD 21218

E-mail address: laurent.younes@jhu.edu