Diffeomorphic 3D Image Registration via Geodesic Shooting Using an Efficient Adjoint Calculation

**François-Xavier Vialard · Laurent Risser ·
Daniel Rueckert · Colin J. Cotter**

Abstract In the context of *large deformations by diffeomorphisms*, we propose a new diffeomorphic registration algorithm for 3D images that performs the optimization directly on the set of geodesic flows. The key contribution of this work is to provide an accurate estimation of the so-called initial momentum, which is a scalar function encoding the optimal deformation between two images through the Hamiltonian equations of geodesics. Since the initial momentum has proven to be a key tool for statistics on shape spaces, our algorithm enables more reliable statistical comparisons for 3D images.

Our proposed algorithm is a gradient descent on the initial momentum, where the gradient is calculated using standard methods from optimal control theory. To improve the numerical efficiency of the gradient computation, we have developed an integral formulation of the adjoint equations associated with the geodesic equations.

We then apply it successfully to the registration of 2D phantom images and 3D cerebral images. By comparing

our algorithm to the standard approach of Beg et al. (Int. J. Comput. Vis. 61:139–157, 2005), we show that it provides a more reliable estimation of the initial momentum for the optimal path. In addition to promising statistical applications, we finally discuss different perspectives opened by this work, in particular in the new field of longitudinal analysis of biomedical images.

1 Introduction

In this section, we first present diffeomorphic registration in *Computational Anatomy* and the framework of *Large Deformation by Diffeomorphisms Metric Mappings* for 3D images. We then introduce our methodology in 1.3.

1.1 Diffeomorphic Registration in *Computational Anatomy*

In the field of *Computational Anatomy*, the analytical and statistical study of biological shape variability has been actively developed in the past fifteen years Pennec (2009) or Grenander and Miller (1998). In particular, substantial efforts (Klein et al. 2009) have been made to propose efficient non-rigid diffeomorphic registration algorithms for volumetric 3D images (Beg et al. 2005; Ashburner 2007; Avants et al. 2008; Vercauteren et al. 2009). A crucial point for *Computational Anatomy* is the usability of the registration for statistical analysis. To this end, the *Large Deformation by Diffeomorphism Metric Mapping* (*LDDMM*) framework has strong mathematical foundations which have been built since the pioneering work of Trouvé (1998),

F.-X. Vialard (✉)
Institute for Mathematical Science, Imperial College London,
53 Prince's Gate, SW7 2PG London, UK
e-mail: f.vialard@imperial.ac.uk

L. Risser
Biomedical Image Analysis Group and Institute for Mathematical
Sciences, Imperial College London, 53 Prince's Gate,
SW7 2PG London, UK

D. Rueckert
Department of Computing, Imperial College London,
53 Prince's Gate, SW7 2PG London, UK

C.J. Cotter
Department of Aeronautics, Imperial College London,
53 Prince's Gate, SW7 2PG London, UK

Dupuis et al. (1998). The key idea underlying the *LDDMM* framework (Miller and Younes 2001; Miller et al. 2002; Grenander 1993) is the study of shape spaces under the action of a diffeomorphism group endowed with a right-invariant Riemannian metric (see Lee 1997 for a comprehensive introduction to Riemannian geometry). In this context, the registration problem is similar to finding a shortest path or geodesic on the infinite-dimensional Riemannian manifold of diffeomorphisms. More importantly, as explained in Trouvé (1998), this action also induces a Riemannian metric on the orbit of the initial shape, which is the set of all the possible deformed shapes using any diffeomorphism in the group. Consequently this registration problem is also equivalent to finding a shortest path on the orbit of the initial shape. Therefore, many tools from Riemannian geometry can be directly used or, if necessary, extended to the infinite dimensional setting (Miller et al. 2006). Among those, the exponential map encodes the geodesic flow through the tangent space at a given point, i.e. a given shape. Based on this geodesic representation, statistical tools for non-linear spaces have been recently developed in (Fletcher et al. 2004, 2008; Fletcher 2004), where the authors provide, in particular, a generalization of the classical PCA to Riemannian manifolds.

1.2 *LDDMM* for 3D Image Registration

Motivated by these statistical developments, it was proposed to practically use the exponential map and hence to optimize on the initial tangent vector to solve the shape matching problem. The finite dimensional case of landmark, originally developed in Joshi and Miller (2000), is the most accessible one. To this end, the Hamiltonian formulation of geodesics (Allassonnière et al. 2005), as comprehensively explained in Lee (1997), has been deeply used, allowing statistical studies of the deformations (Vaillant et al. 2004). These advances have been pushed further in Allassonnière et al. (2006, 2007), which also include template estimation. In the same direction, Hamiltonian shooting methods have been developed in Marsland and McLachlan (2007), Allassonnière et al. (2005). All of them are based on finite dimensional approximations that reduce the problem to the landmark case in a Lagrangian setting. For the case of surfaces, Ma et al. (2010) uses the adjoint equations based on landmarks. Interestingly some very recent work has been done in Du et al. (2011) to simultaneously unifies images and set of points in the *LDDMM* framework. Compared with landmarks, volumetric images can encode shapes with spatial uncertainty through partial volume effects and probability maps. They also often require fewer pre-processing or manual interventions which can imply a loss of spatial accuracy. In contrast to this Lagrangian setting, the algorithm introduced in Beg et al. (2005) presents a strategy to match volumetric images from an Eulerian point of view.

Briefly, we now describe the *LDDMM* framework in the case of 3D images. Based on the minimization of an energy functional such as (1), it provides a smooth and invertible deformation that transforms the source image I in order to match the target image J. In Beg et al. (2005), a gradient descent on the whole set of curves in the diffeomorphism group is implemented for the functional

$$\mathcal{E}(v) \doteq \int_0^1 \frac{\lambda}{2} |v(t)|_V^2 \, dt + \frac{1}{2} \| I \circ \phi_{0,1}^{-1} - J \|_{L^2}^2. \tag{1}$$

The second term of $\mathcal{E}$ is the similarity measure and is chosen as the square of the L^2 difference between the deformed image $I \circ \phi_{1,0}$ and the target image J. The first term represents the energy of the deformation. The time-dependent velocity field $v \in L^2([0, 1], V)$ encodes the path of diffeomorphisms $\phi_{0,t}$ through the ODE $\dot{\phi}_{0,t}(x) = v(t) \circ \phi_{0,t}$ with initial condition $\phi_{0,0} = Id$. The velocity fields are sufficiently smooth to produce diffeomorphisms and V is a Reproducing Kernel Hilbert Space (RKHS) of vector fields defined on the domain of interest Ω such that, for a positive constant M: $|du|_\infty \leq M \|u\|_V$ for any $u \in V$. This condition ensures the existence of a minimizer for the functional $\mathcal{E}$ as proven in Beg et al. (2005). Theoretically, any minimizer of $\mathcal{E}$ gives rise to a geodesic path on the group of diffeomorphisms, as well as a geodesic path on the template orbit.

From a practical perspective, we would expect the gradient descent in Beg et al. (2005) to give a geodesic path after convergence. With hindsight on the practical use of Beg et al. (2005), even though the matching between the compared shapes is obtained after relatively few iterations, the convergence of the diffeomorphisms path towards the optimal path of deformations can be very slow for complex deformations. This is due to the large number of degrees of freedom of the problem (the number of timesteps times the dimension of a 3D vector field), as well as the use of a gradient descent method. In this context, even if the optimized path is not geodesic, the initial momentum can be estimated. Unfortunately in this non-geodesic case, this initial momentum does not encode the entire path. In practice, exact geodesics are indeed not retrieved with Beg et al. (2005), as shown in Sect. 4, even after a large number of gradient descent iterations.

As a matter of fact, the algorithm (Beg et al. 2005) is currently used to perform PCA analysis on the initial momentum, for instance Wang et al. (2006) or more recently Singh et al. (2010). Such statistical studies would gain accuracy from reliable estimations of initial momenta. In order to address this problem, the computation of Jacobi fields on the group of diffeomorphisms has been developed in Younes (2007). In this paper, an optimization algorithm on the initial momentum is proposed, which requires the computation of Jacobi fields on the group of diffeomorphisms. As described hereafter, we will follow a versatile mathematical

method, which mathematically differs from Younes (2007) since its formulation is completely independent of any computation on the underlying diffeomorphisms group. In fact, we do compute the adjoint Jacobi fields on the space of images itself.

1.3 Overview of the Proposed Methodology

Following an Eulerian viewpoint as in Beg et al. (2005), we propose a simple optimization algorithm on the initial momentum using standard approaches in variational data assimilation (Le Dimet and Talagrand 1986; Sasaki 1970). In the field of variational data assimilation, optimal control methods have extensively been applied, especially for weather forecast purposes. For a brief but complete overview of the method used here, we may refer the reader to Sect. 3 of Le Dimet et al. (1997). In order to apply this method, the first step consists of constraining the dynamic of the paths to be geodesic. To this end, an additional variable P, i.e. the momentum, is needed for the shooting equations (2). This ODE completely encodes the whole time-dependent deformation in the initial momentum P_0. The second step consists in rewriting the functional $\mathcal{E}$ in terms of this initial momentum: this is done in Sect. 2 with the functional $\mathcal{S}$ defined in (4). The deformed image is now given by the solution of the Hamiltonian equations of geodesics. To find the initial momentum that minimizes (4) subject to the constraint (2), we use a gradient descent method. To compute the gradient of the functional with respect to the initial momentum, the adjoint system to the Hamiltonian equations (2) needs to be solved. From a mathematical viewpoint, it appears that these adjoint equations are exactly the adjoint Jacobi fields for the geodesics on the space of images. In comparison with Younes (2007), we do not work on the whole space of diffeomorphisms but only on the cotangent space of images. The interest of such a formulation is shown in Hart et al. (2009), where the authors reformulate the gradient descent of Beg et al. (2005) on the cotangent space of images in a coordinate-free form. Their key contribution is to show the numerical feasibility of this direct approach using WENO schemes (Jiang and Shu 1996). Contrary to Hart et al. (2009) and Beg et al. (2005), we aim at enforcing the geodesicity at each step of the optimization process. As expected, solving the equations in this situation is more involved, since the adjoint Jacobi fields equations are of second order. To address the numerical issues raised by these second-order equations, we have developed an integral formulation which relies on the computation of the whole path of diffeomorphisms at each step of the optimization process. This formulation is a crucial point to obtain an accurate estimation for the gradient of the functional.

In Sect. 2, we present the Hamiltonian formulation of the geodesic equations on the space of images and give

the formulation of the adjoint equations associated with this Hamiltonian system. In order to make the numerical calculation of the adjoint equations feasible, we propose an integral formulation of these equations and give the details of the new registration algorithm in Sect. 3. Our method is evaluated in Sect. 4 and compared to the reference algorithm (Beg et al. 2005). Section 5 finally opens a discussion on the method and proposes a few perspectives.

2 Hamiltonian Equations of Geodesics and Associated Adjoint Equations

Numerous efforts have been made to build the mathematical foundations of the LDDMM framework. Here, we give a brief overview of the necessary material for our work and present some related questions of direct interest. In Beg et al. (2005), the image to be deformed is assumed to be in the Sobolev space $H^1(\Omega)$, where $\Omega \subset \mathbb{R}^d$ is a bounded domain. The Euler-Lagrange equation is computed in this case and gives the existence of a momentum in $L^2(\Omega)$ that actually belongs to $H^1(\Omega)$ provided that the target image is also in $H^1(\Omega)$. When the images can be treated as SBV (Special Bounded Variations) functions, i.e. discontinuous images, the mathematical analysis of the Euler-Lagrange equation of (1) is more subtle. This has been treated in Vialard and Santambrogio (2009), it is proven that the momentum does not belong to L^2 since it also contains a singular part supported by the jump set of the source image. Since this complication would force us to mix both Lagrangian and Eulerian approaches, we only work with sufficiently smooth source and target images in order to carry on with the Eulerian approach. In this case, the Hamiltonian equations of geodesics are written in Younes (2007). In this infinite dimensional situation, the geodesic equations involve the momentum P and the image I which are both scalar functions. Omitting the time and space dependences for the image I and the momentum P, those geodesic equations read

$$\begin{cases} \partial_t I + \nabla I \cdot v = 0 \\ \partial_t P + \nabla \cdot (Pv) = 0 \\ v + K \star \nabla I P = 0, \end{cases} \qquad (2)$$

subject to initial conditions at time 0, $I(0), P(0) \in H^1(\Omega)$. Here, K is the translation-invariant kernel of the chosen RKHS and $\star$ denotes the convolution operator. The velocity field v evolves in time according to the last equation of (2). Remark that the momentum P is a scalar function dragged as a density along the flow which pointwise multiplies the gradient of the current image I. In the advection equation for the image I (first equation in (2)), ∇ denotes the gradient and in the continuity equation for the momentum P (second equation in (2)), $\nabla\cdot$ denotes the divergence operator.

Remark 1 The two first equations in (2) are dual to each other. Consequently, the same system describes the geodesic flow on the space of densities, as remarked in Younes et al. (2009). In this situation, the corresponding momentum would be I and the density would be P. Since this case is sufficiently close to the image matching case from the Eulerian point of view developed in this article, we have treated the case of densities in Appendix A.

Remark 2 These geodesic equations can be seen as a particular case of geodesics for the "metamorphosis" approach originally introduced in Trouvé and Younes (2005), and then developed further in Holm et al. (2009). This approach enables the grey scale value of each voxel in the image to be changed within a Riemannian setting. The advection equation on the image is replaced with

$$\partial_t I + \nabla I \cdot v = \sigma^2 P, \tag{3}$$

as written in Younes et al. (2009), where σ is a real parameter. When $\sigma = 0$, the geodesic equations (2) are recovered.

The final image $I(1)$ can be obtained from the initial image $I(0)$ by solving the system (2) of PDEs, given an initial momentum $P(0)$. Hence, deformations of $I(0)$ can be encoded from initial conditions for P. For this initial value problem, existence and uniqueness of solutions have been proven in Vialard (2009) and rely on existence and uniqueness of the initial value problem for EPDiff, for a particular set of initial conditions, that can be found in Younes (2010). A related proof in the case of metamorphosis is given in Trouvé and Younes (2005).

Noting that the energy in this autonomous Hamiltonian system is conserved, the time integral in the functional $\mathcal{E}$ may be replaced by the Hamiltonian of the system at time 0. Hence, we may rewrite the functional $\mathcal{E}$ as:

$$\mathcal{S}(P(0)) = \frac{\lambda}{2} \langle \nabla I(0) P(0), K \star \nabla I(0) P(0) \rangle_{L^2}$$
$$+ \frac{1}{2} \|I(1) - J\|_{L^2}^2. \tag{4}$$

The final diffeomorphism does not appear in this modified formulation since the final deformed image is given by solving the geodesic equations (2) until time 1. Obviously, remark that one can not retrieve in general the diffeomorphism from the deformed image at time 1.

In order to compute the gradient of $\mathcal{S}$, we use the standard method of the adjoint equations in optimal control. We introduce a time-dependent Lagrange multiplier in $\mathcal{S}$ to enforce the geodesic constraints (2) on the variables (I, P). We then obtain the gradient of the functional $\mathcal{S}$ by solving the system of adjoint equations (5), given below.

Theorem 1 *The gradient of $\mathcal{S}$ is given by*:

$$\nabla_{P(0)}\mathcal{S} = \lambda \nabla I(0) \cdot K \star (P(0) \nabla I(0)) - \hat{P}(0),$$

where $\hat{P}(0)$ is given by the solution of the following PDE solved backward in time:

$$\begin{cases} \partial_t \hat{I} + \nabla \cdot (v\hat{I}) + \nabla \cdot (P\hat{v}) = 0, \\ \partial_t \hat{P} + v \cdot \nabla \hat{P} - \nabla I \cdot \hat{v} = 0, \\ \hat{v} + K \star (\hat{I}\nabla I - P\nabla \hat{P}) = 0, \end{cases} \tag{5}$$

subject to the initial conditions:

$$\begin{cases} \hat{I}(1) = J - I(1), \\ \hat{P}(1) = 0, \end{cases} \tag{6}$$

and $P(t), I(t)$ are the solution of the shooting system (2) for the initial conditions $I(0), P(0)$.

Remark 3 As is standard in optimal control, the computation of the gradient in Theorem 1 is divided into two steps: first, solving the geodesic shooting equations (2) forward in time; second, solving the adjoint equations (5) backward in time. This requires knowledge of the whole path $(I(t), P(t))$ for $t \in [0, 1]$. This path can be either stored, recomputed using the shooting equations backwards, or by a mixed strategy, called check-pointing, to obtain a more efficient implementation in terms of memory allocation and computational time, as shown in Ngodock (2005).

Remark 4 We underline that this approach is valid for a large range of similarity measures satisfying some smoothness assumptions. More precisely, rewriting the similarity measure of the functional $\mathcal{E}$ as $M(I) \doteq \frac{1}{2}\|I - I_{target}\|_{L^2}^2$, we obtain $\hat{I}(1) = -\frac{\delta M}{\delta I}(I(1))$ as initial condition for the adjoint equations.

The proof of Theorem 1 relies on classical optimal control calculus. It consists in introducing a time dependent Lagrange multiplier to enforce the geodesicity encoded via the shooting equations (2). Since we are dealing with images which are essentially modeled in infinite dimensional spaces, an infinite-dimensional version of Lagrange multipliers is desirable. A convenient way consists in recasting this problem using an ODE on a Hilbert space, which is the point of view adopted in Appendix B. Then one can apply a general version of Lagrange multipliers on Hilbert spaces (Zeidler 1995).

Proof In order to prove Theorem 1, we use the standard method of a time dependent Lagrange multiplier to derive the adjoint equations. Let $\hat{I}, \hat{P}, \hat{V}$ be the adjoint variables for respectively I, P, v (the variable $\hat{v}$ is in fact redundant

but it makes the calculation clearer). We also replace the notation $\partial_t f$ with $\dot{f}$ for any function f. We then extend the functional $\mathcal{S}$ with the Lagrange multipliers to get:

$$\hat{\mathcal{S}} \doteq \mathcal{S} + \int_0^1 \langle \hat{P}, \dot{P} + \nabla \cdot (Pv) \rangle_{L^2} \, dt$$

$$+ \int_0^1 \langle \hat{I}, \dot{I} + \nabla I \cdot v \rangle_{L^2} \, dt$$

$$+ \int_0^1 \langle \hat{V}, v + K \star \nabla I P \rangle_{L^2} \, dt. \tag{7}$$

Before taking variations with respect to the whole paths of $I(t), P(t), v(t)$, we recall useful integration by part formulas. For a smooth vector field v defined on Ω the divergence theorem reads

$$\int_\Omega f \nabla \cdot v \, dx + \int_\Omega \langle \nabla f, v \rangle_{\mathbb{R}^d} \, dx = \int_{\partial \Omega} \langle v, n \rangle_{\mathbb{R}^d} \, ds, \tag{8}$$

with n denoting the normalized outward-pointing normal vector to the boundary $\partial \Omega$. In our situation, we assume that the kernel K vanishes on $\partial \Omega$ such that for any vector field v satisfying the last equation of System (2) we get

$$\langle f, \nabla \cdot (vg) \rangle_{L^2(\Omega)} = -\langle v \cdot \nabla f, g \rangle_{L^2(\Omega)} \tag{9}$$

We now take variations of the augmented functional (7) using the previous formula (9) to get

$$\delta \hat{\mathcal{S}} = \lambda \langle \nabla I(0) \delta P(0), K \star \nabla I(0) P(0) \rangle_{L^2}$$

$$+ \langle \delta I(1), I(1) - J \rangle_{L^2}$$

$$+ \int_0^1 \langle -\dot{\hat{P}} - \nabla \hat{P} \cdot v, \delta P \rangle_{L^2} \, dt + \langle \hat{P}(1), \delta P(1) \rangle_{L^2}$$

$$- \langle \hat{P}(0), \delta P(0) \rangle_{L^2}$$

$$+ \int_0^1 \langle -\dot{\hat{I}} - \nabla \cdot (Iv), \delta I \rangle_{L^2} \, dt$$

$$+ \langle \hat{I}(1), \delta I(1) \rangle_{L^2} - \langle \hat{I}(0), \delta I(0) \rangle_{L^2}$$

$$+ \int_0^1 (\langle \hat{V}, \delta v + K \star \nabla \delta I P + K \star \nabla I \delta P \rangle_{L^2}$$

$$- \langle P \nabla \hat{P}, \delta v \rangle_{L^2} + \langle \hat{I} \nabla I, \delta v \rangle_{L^2}) \, dt. \tag{10}$$

The first term is obtained by calculating the variations of the energy of the momentum at time 0 (first term of the functional (4)).

Since we have

$$\langle \hat{V}, K \star P \nabla \delta I + K \star \nabla I \delta P \rangle_{L^2}$$

$$= \langle -\nabla \cdot P K \star \hat{V}, \delta I \rangle + \langle \nabla I \cdot K \star \hat{V}, \delta P \rangle_{L^2}, \tag{11}$$

the adjoint equations in terms of $\hat{I}, \hat{P}, \hat{V}$ read

$$\begin{cases} \dot{\hat{P}} + \nabla \hat{P} \cdot v - \nabla I \cdot K \star \hat{V} = 0, \\ \dot{\hat{I}} + \nabla \cdot (Iv) + \nabla \cdot P K \star \hat{V} = 0, \\ \hat{V} + \hat{I} \nabla I - P \nabla \hat{P} = 0, \end{cases} \tag{12}$$

subject to boundary conditions

$$\begin{cases} \hat{I}(1) + I(1) - J = 0, \\ \hat{P}(1) = 0. \end{cases} \tag{13}$$

If we can find a solution of the adjoint equations with the boundary conditions, we can vary $\hat{S}$ with respect to a variation of $P(0)$ since we have $\delta I(0) = 0$ to obtain:

$$\delta \hat{\mathcal{S}} = \lambda \langle \nabla I(0) \delta P(0), K \star \nabla I(0) P(0) \rangle_{L^2}$$

$$- \langle \hat{P}(0), \delta P(0) \rangle_{L^2}. \tag{14}$$

It ends to prove Theorem 1 where we used the shorter notation $\hat{v} = K \star \hat{V}$. □

Next, we address the questions of existence and uniqueness of solutions to the adjoint equations (5) and describe our algorithm.

3 Diffeomorphic Registration Algorithm via Geodesic Shooting

3.1 Diffeomorphic Registration Algorithm

The computation of the gradient described in Theorem 1 requires the solution of the PDE (5). In this section, we present an integral formulation of the adjoint equations using the whole path of diffeomorphisms obtained by integrating the velocity v. From a mathematical point of view, this integral formulation of the system (5) not only gives an existence and uniqueness result but also a practical algorithm for implementing the gradient descent on the initial momentum. Although this result may not appear surprising, it is worth presenting it since it gives some insight on the spatial and time regularity of the solutions, which has important consequences for numerical accuracy.

Theorem 2 *Let $I(0), J \in H^2(\Omega, \mathbb{R})$ be two images and K be a C^2 kernel on Ω. For any $P(0) \in L^2(\Omega)$, we consider the unique solution (I, P) of the shooting problem (2) with initial conditions $I(0), P(0)$. Then, the corresponding adjoint equations (5) with initial conditions (6) have a unique*

solution $(\hat{I}, \hat{P})$ *in* $C^0([0, 1], H^1(\Omega) \times H^1(\Omega))$. *The solution satisfies the integral relation*:

$$\begin{cases} \hat{P}(t) = \hat{P}(1) \circ \phi_{t,1} - \int_t^1 [\nabla I(s) \cdot \hat{v}(s)] \circ \phi_{t,s} \, ds, \\ \hat{I}(t) = \mathrm{Jac}(\phi_{t,1})\hat{I}(1) \circ \phi_{t,1} \\ \qquad\qquad + \int_t^1 \mathrm{Jac}(\phi_{t,s})[\nabla \cdot (P(s)\hat{v}(s))] \circ \phi_{t,s} \, ds, \end{cases} \quad (15)$$

with:

$$\begin{cases} \hat{v}(t) = K \star [P(t)\nabla \hat{P}(t) - \hat{I}(t)\nabla I(t)], \\ P(t) = \mathrm{Jac}(\phi_{t,0})P(0) \circ \phi_{t,0}, \\ I(t) = I(0) \circ \phi_{t,0}, \end{cases} \quad (16)$$

where $\phi_{s,t} \doteq \phi_{0,t} \circ \phi_{0,s}^{-1}$ *is defined by the flow of the time-dependent velocity field* $v(t) = -K \star P(t)\nabla I(t)$.

Remark If the source and target images are smooth, then the initial momentum path is regular in time and space. For instance, the initial momentum belongs to $H^1(\Omega)$ if the source and target images are given in $H^2(\Omega)$. This theoretical remark justifies the numerical feasibility of our approach.

The proof of Theorem 2 is given in Appendix B and is based on the Banach fixed point theorem. After introducing the variables $\tilde{P}$ and $\tilde{I}$ defined as $\tilde{P}(t) = \hat{P}(t) \circ \phi_{0,t}$ and $\tilde{I}(t) = \mathrm{Jac}(\phi_{0,t})\hat{I}(t) \circ \phi_{0,t}$, we have:

$$\begin{cases} \tilde{P}(t) = \tilde{P}(1) - \int_t^1 [\nabla I(s) \cdot \hat{v}(s)] \circ \phi_{0,s} \, ds, \\ \tilde{I}(t) = \tilde{I}(1) + \int_t^1 \mathrm{Jac}(\phi_{0,s})[\nabla \cdot (P(s)\hat{v}(s))] \circ \phi_{0,s} \, ds. \end{cases} \quad (17)$$

Using (24), we have implemented the following algorithm to compute the gradient of the functional:

Algorithm 1 The computation of $\nabla_{P(0)}\mathcal{S}$ is divided into two parts:

Shooting algorithm: Initialize $I(0)$, $\phi_{0,0} = Id$ and $P(0)$. Compute $v(0) = -K \star P(0)\nabla I(0)$ and estimate the number of time-steps $h = [\frac{\max |v(0)|}{\beta}]$ with typically $\beta = 0.1$ to ensure the Courant Friedrichs Lewy (CFL) condition at the first time step (here $[\cdot]$ stands for the floor of the real value). Discretize the time interval $[0, 1]$ in $(t_i)_{i=0,h}$ defined as $t_i = \frac{i}{h}$.
While the CFL condition is satisfied, for each timestep:

1. Update $P(t_i)$ and $I(t_i)$ with $I(t_i) = I(0) \circ \phi_{t_i,0}$ and $P(t_i) = \mathrm{Jac}(\phi_{t_i,0})P_0 \circ \phi(t_i, 0)$.
2. Update $v(t_i)$ with $v(t_i) = -K \star P(t_i)\nabla I(t_i)$.

3. Compute and store $\phi_{0,t_{i+1}} = \phi_{0,t_i} + (t_{i+1} - t_i)v(t_i) \circ \phi_{0,t_i}$ (linear interpolation of v_{t_i} is used here).
4. From the advection equation on the diffeomorphism,

   $$\partial_t \phi_{t,0} + d\phi_{t,0}(v_t) = 0,$$

   compute and store $\phi_{t_{i+1},0} = \phi_{t_i,0} + F(\phi_{t_i,0}, v_{t_i})$ where F represents the standard finite volume scheme using the MinMod limiter for advection problems, as described in Le Veque (2002).

If the CFL condition is not satisfied at a certain time-step, increase the number of steps h accordingly.
Gradient computation: Since $\tilde{P}(0) = \hat{P}(0)$ we have

$$\nabla_{P(0)}\mathcal{S} = -\tilde{P}(0) + \nabla I(0) \cdot K \star (P(0)\nabla I(0)). \quad (18)$$

The integral formula is computed using the forward Euler discretization for the formula (17) as

$$\tilde{P}(t_i) = \tilde{P}(1) - \sum_{j=i+1}^h [\nabla I(t_j) \cdot \hat{v}(t_j)] \circ \phi_{0,t_j}, \quad (19)$$

with $\hat{v}(t_i) = K \star [P(t_i)\nabla \hat{P}(t_i) - \hat{I}(t_i)\nabla I(t_i)]$. The forward Euler discretization is also used to compute $\tilde{I}(t_i)$. Then, $\hat{P}(t_i)$ and $\hat{I}(t_i)$ are computed from the formulas below, using linear interpolation:

$$\hat{P}_{t_i} = \tilde{P}_{t_i} \circ \phi_{t_i,0},$$
$$\hat{I}_{t_i} = \mathrm{Jac}(\phi_{0,t_i})\tilde{I}_{t_i} \circ \phi_{t_i,0}.$$

Several improvements to the accuracy of this approach can be made, but a fast algorithm is required for practical applications. We underline that the CFL condition still has to be satisfied when computing $\phi_{t,0}$. This is directly related to the norm of P_0, hence requiring more time steps for large deformations. Other second-order limiters such as the so-called SuperBee limiter produced equivalent results; the first order upwind method was too diffusive.

Remark 5 When it comes to solving the adjoint equations, one can choose between solving the numerical adjoint of the shooting equations and solving the adjoint model of the continuous forward model. Here, the integral formulation gives the solution to the adjoint of the continuous forward equations; the numerical efficiency justifies not solving the numerical adjoint model. Indeed, the adjoint equations have CFL conditions corresponding to propagation by both the forward and adjoint velocities. When the discrepancy between the target and mapped images is large, this corresponds to a large adjoint momentum and hence a large adjoint velocity, which in our experiments was more demanding than the CFL condition for the geodesic shooting equations.
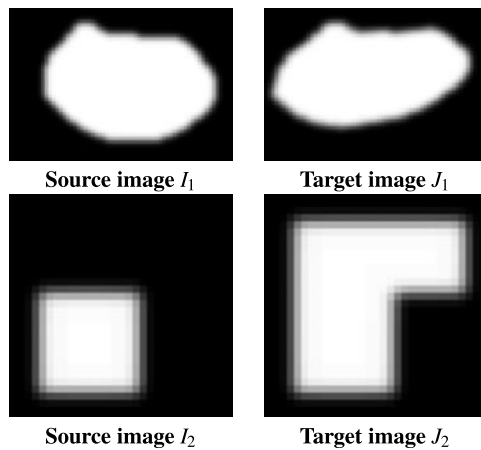
Fig. 1 Source and target images registered in the tests of Sect. 4.1. Results are presented in Figs. 2 and 3

4 Evaluation

4.1 Momentum Map Estimation Accuracy

The approach of Beg et al. (2005) is particularly interesting since at has been shown efficient for the comparison of real 3D images Beg et al. (2004), Helm et al. (2006), Singh et al. (2010). Our goal is to show that the proposed geodesic shooting strategy is more accurate than Beg et al. (2005) to estimate initial momenta.

We compare in this subsection our geodesic shooting strategy with the standard LDDMM technique of (Beg et al. 2005) on 2D phantom images. To do so, we register the two image pairs of Fig. 1. The first image pair (I_1, J_1) only requires a simple deformation to have a good matching, while the second one (I_2, J_2) requires a relatively large deformation. Note also that the shapes of (I_2, J_2) have a low resolution. The square I_2 has indeed a size of 8×8 pixels. It is then numerically hard to find a diffeomorphism to compare I_2 and J_2.

We used isotropic Gaussian kernels K having standard deviations of 5 pixels to register I_1 on J_1 and of 3 pixels to register I_2 on J_2. We emphasize that the goal of our registration method is not to compute the best matching between the images but to find optimal statistical descriptors which compare them according to the metric defined by the kernel K. As a consequence, the matching of sharp corners is not necessarily accurate. Deeper insights about the choice of the kernel can be found in Risser et al. (2010, 2011).

4.1.1 Parameters Influence

In all cases the algorithms converged in less than 100 gradient descent iterations. Note that, contrary to the initial momenta estimated using (Beg et al. 2005), those estimated using geodesic shooting have already a statistical meaning before the algorithm convergence. To register I_1 on J_1, only
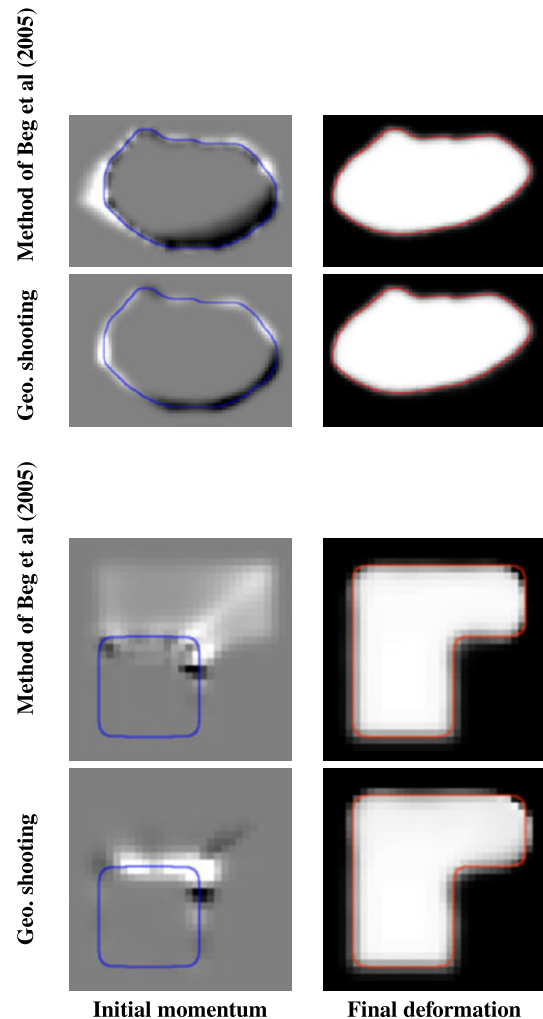

Fig. 2 Estimation of the initial momentum using (Beg et al. 2005) and our geodesic shooting algorithm to compare the source and target images of Fig. 1. *The two first rows* represent the registration of I_1 on J_1 and two last ones the registration of I_2 on J_2. In the images of initial momenta, the blue curves show the boundary of the source images. In the images of the final deformations, *the red curves* show the boundary of the target images

$h = 10$ time-steps between $t = 0$ and $t = 1$ were enough to respect the CFL condition since the deformations are relatively simple. For more complex deformations between I_2 on J_2, $h = 20$ time-steps were required. Note that in real applications, the number of time-steps is automatically increased when the CFL condition is not satisfied. It is also important to note that to compare two images, the larger the Gaussian kernel K, the more numerically stable the shooting and therefore the fewer time-steps are required.

4.1.2 Matching and Momentum Accuracy

We show in Fig. 2 the momenta and deformations $I_1 \circ \phi^{-1}$ and $I_2 \circ \phi^{-1}$ obtained by using the geodesic shooting and Beg et al. (2005) with $h = 20$ time-steps and 100 gradient
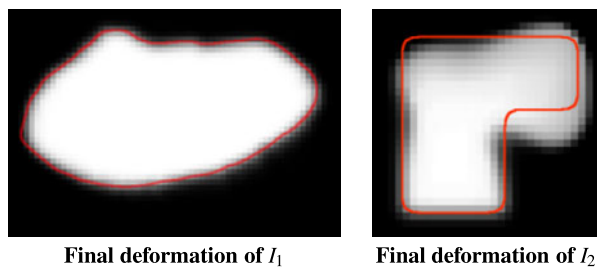
Final deformation of I_1 **Final deformation of** I_2

Fig. 3 Deformations obtained by integrating the geodesic shooting equation (2) from the initial momenta estimated indirectly (using Beg et al. 2005) as shown in Fig. 2. *The red curves* show the boundary of the target images

descent iterations, where ϕ denotes the estimated deformation between the images at $t = 1$. Using both techniques, the matching between $I_1 \circ \phi^{-1}$ and J_1 is very accurate. Between $I_2 \circ \phi^{-1}$ and J_2 the matching is also accurate except at the shape corners in J_2, with a small advantage for Beg et al. (2005). The lack of accuracy at the corners is naturally due to the choice of the metric as discussed earlier in this subsection. However the initial momentum indirectly estimated using (Beg et al. 2005) appears coarse compared to the one directly estimated using the geodesic shooting strategy.

In order to evaluate the use of the initial momenta estimated using Beg et al. (2005) to describe the estimated deformations, we integrated the geodesic shooting equation (2) using these momenta. Results are shown Fig. 3. Interestingly, the matching $I_1 \circ \phi^{-1}$ and J_1 is as accurate as the one obtained using the geodesic shooting directly. This is however not the case between $I_2 \circ \phi^{-1}$ and J_2. In this test, the momentum estimated using Beg et al. (2005) is therefore not fine enough to encode in details the deformation. Note that even by refining the time steps of the shooting, similar deformations are obtained. However, there may be errors caused by the use of a different spatial discretization for the gradient descent (namely, semi-Lagrangian advection) than for the geodesic shooting equation (namely, a Eulerian upwind finite volume method). The variational integrator implied by the optimum obtained Beg's algorithm would have a rather complicated and implicit equation for the conjugate momentum. To conclude here, it is not practical to use the estimated initial momentum maps to reconstruct estimated deformations using the gradient descent of Beg et al. (2005), especially when encoding relatively large or fine deformations.

4.2 Geodesic Shooting to Compare Complex Organ Anatomies

In this subsection, we show the usability of the geodesic shooting algorithm to quantify the temporal evolution of organ anatomies represented in volumetric images. Our technique is applied to the longitudinal estimation of the early brain development on two 3D MR brain scans acquired at
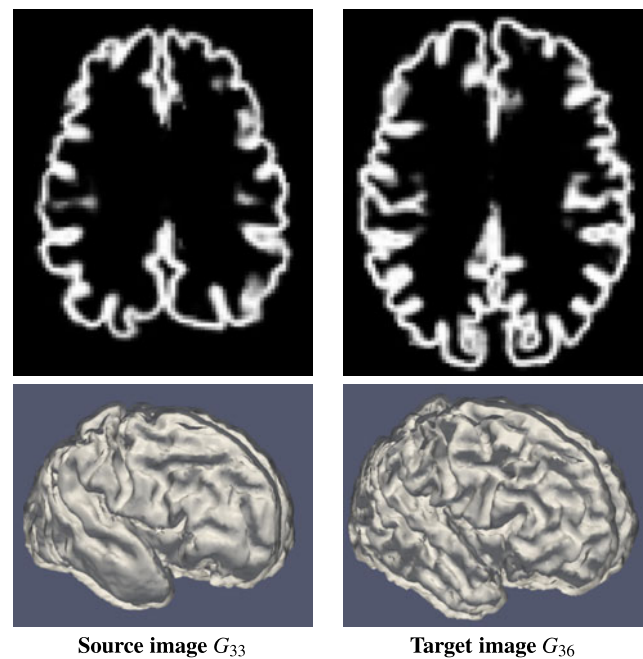

Source image G_{33} **Target image** G_{36}

Fig. 4 Registered segmented cortex out of MR images of pre-born babies at 33 and 36 weeks of gestational age, denoted by G_{33} and G_{36} respectively. (*Top*) Slices out of the segmented images re-used in Fig. 6. (*Bottom*) Internal face of the volumetric images surface

33 and 36 weeks of gestational age with a spatial resolution of 0.85 mm. Bias field correction was first performed using N3 (Sled et al. 1998) and the images were rigidly aligned. In order to extract the brain grey matter, probabilistic segmentation of this structure was then performed using the method of Murgasova et al. (2011), as presented in Fig. 4. We then compare in this subsection the extracted gray matter of follow-up images, denoted as G_{33} and G_{36} for the images acquired at 33 and 36 weeks respectively.

In order to emphasize the influence of the metric when comparing complex structures like the cortex, we used different kernels to register the follow-up images. We first registered G_{33} and G_{36} by using isotropic Gaussian kernels $K_{1.5}$, K_3, K_5, K_{10} and K_{25} with standard deviations 1.5, 3, 5, 10 and 25 mm respectively. We then used the multikernel strategy of Risser et al. (2010, 2011) by defining the last kernel *MK* as the sum of two isotropic Gaussians of standard deviations 1.5 and 25 mm. We used $h = 20$ timesteps, which was also sufficient to properly match I_2 and J_2 with large deformations in Sect. 4.1.2. A multi-resolution strategy was also adopted: the images were registered at half and then full resolution (50 and 15 iterations respectively) with proportional kernels. The geodesic shooting algorithm required about 150 minutes and 955 MB to register the $124 \times 124 \times 124$ voxel images.[1] As a comparison

---

[1] Linux 64-bit machine with a 3.6 Ghz Intel Core Duo, C++ language.

the algorithm of Beg et al. (2005) required 95 minutes and 1.3 GB for the same task with the same parameters. We show in Table 1 the average amplitude of the deformations (AOD) and determinant of the Jacobians (detJ) at the cortical surface and the normalized Sum of Squared Differences (SSD) between $G_{33} \circ \phi^{-1}$ and G_{36}. In Fig. 5, we also show the initial momenta estimated using different kernels at the cortical surface. Finally, we present the initial momentum and deformation estimated using the kernel *MK* in Fig. 6.

As expected, the estimated initial momenta are strongly related to the metric as shown in Fig. 5. These variations have an influence on the registration quality, as quantified in Table 1: the smaller the standard deviation of the isotropic kernels, the smaller the average AOD, the larger the average detJ and the better the matching. This is due to the complex anatomical growth of the cortex. The skull growth implies a general growth of the cortex outwards of the brain while the sulci deepening implies a local cortical growth inwards of the brain. Large kernels therefore mainly capture the general brain growth while small ones capture a mixture of the growths in opposite directions. As a result, the matching is not accurate using large kernel and the estimated deformations are meaningless using small kernels. Quantitatively, the overlap between $G_{33} \circ \phi^{-1}$ and G_{36} is never higher than 94% using a Gaussian kernel. Using the multi-kernel *MK* then appears as a good strategy here since it captures properly the general and local deformations: the average AOD is similar to those observed using large Gaussian kernels and the Jacobian detJ is similar to what observed using small Gaussian kernels. Moreover the best matching is found us-

ing *MK* with deformations that seem plausible, as observed in Fig. 6. Here the overlap between $G_{33} \circ \phi^{-1}$ and G_{36} is about 95%. Note that these results are similar to those obtained in Risser et al. (2011) by using Beg et al. (2005).

Finally, it is interesting to highlight the obvious relation between the initial momentum and the deformation in Fig. 6. It is clear that the largest deformations are observed in the areas with positive or negative momenta. Interestingly, positive momenta push the shape outward of its boundaries while negative momenta push it inside, with a spatial correlation directly related to the metric. Note that the color bars of the 3D representations of AOD and momenta in Fig. 6 have been inverted, in order to avoid confusion due to the fact that the internal face of the cortical surface is represented.

To conclude in this subsection, the initial momentum estimated using geodesic shooting could appear as a pertinent descriptor to compare the 3D anatomy of complex organs. One must however be careful with the choice of the metric, which has a strong influence on the estimated initial momentum and its meaning. Moreover, its meaning can not be disconnected from the gradient of the initial image. Therefore, one may prefer the initial velocity field or even the Jacobian map itself to perform statistics. These are interesting research directions that will be addressed elsewhere.

5 Discussion and Outlook

We have proposed a new diffeomorphic registration algorithm for 3D images that encodes the optimal deformation by a scalar function called the initial momentum. Its main contribution is to provide an accurate estimation of the initial momentum in comparison with Beg et al. (2005), which is currently used to perform statistical studies on this quantity. Consequently, it would lead to more sound and reliable statistics on the variability of anatomical shapes. Along with the creation of atlases, this is a direct perspective of our work. In particular, the local structure of the initial momentum will be extremely useful to achieve pertinent statistics in

Table 1 Average amplitude of the deformation (AOD) and determinant of the Jacobians (detJ) estimated between G_{33} and G_{36} at the cortical surface using different kernels and normalized Sum of Squared Differences (SSD) between $G_{33} \circ \phi^{-1}$ and G_{36}

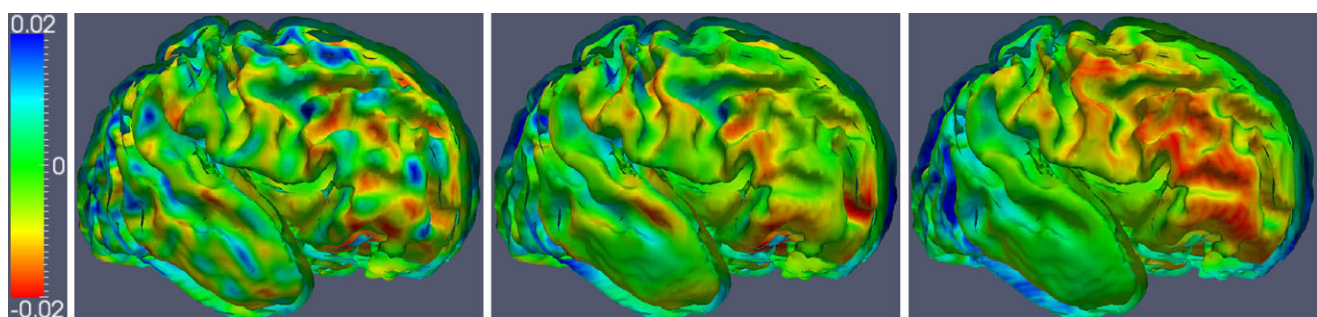|	$K_{1.5}$	K_3	K_5	K_{10}	K_{25}	*MK*
Aver. AOD	1.44	1.85	2.16	2.56	2.57	2.30
Aver. detJ	1.31	1.22	1.16	1.14	1.12	1.38
Norm. SSD	0.012	0.018	0.022	0.025	0.029	0.0061

Fig. 5 Momenta estimated at the cortical surface using *(from left to right)* the isotropic Gaussian kernels $K_{1.5}$, K_5 and K_{25}
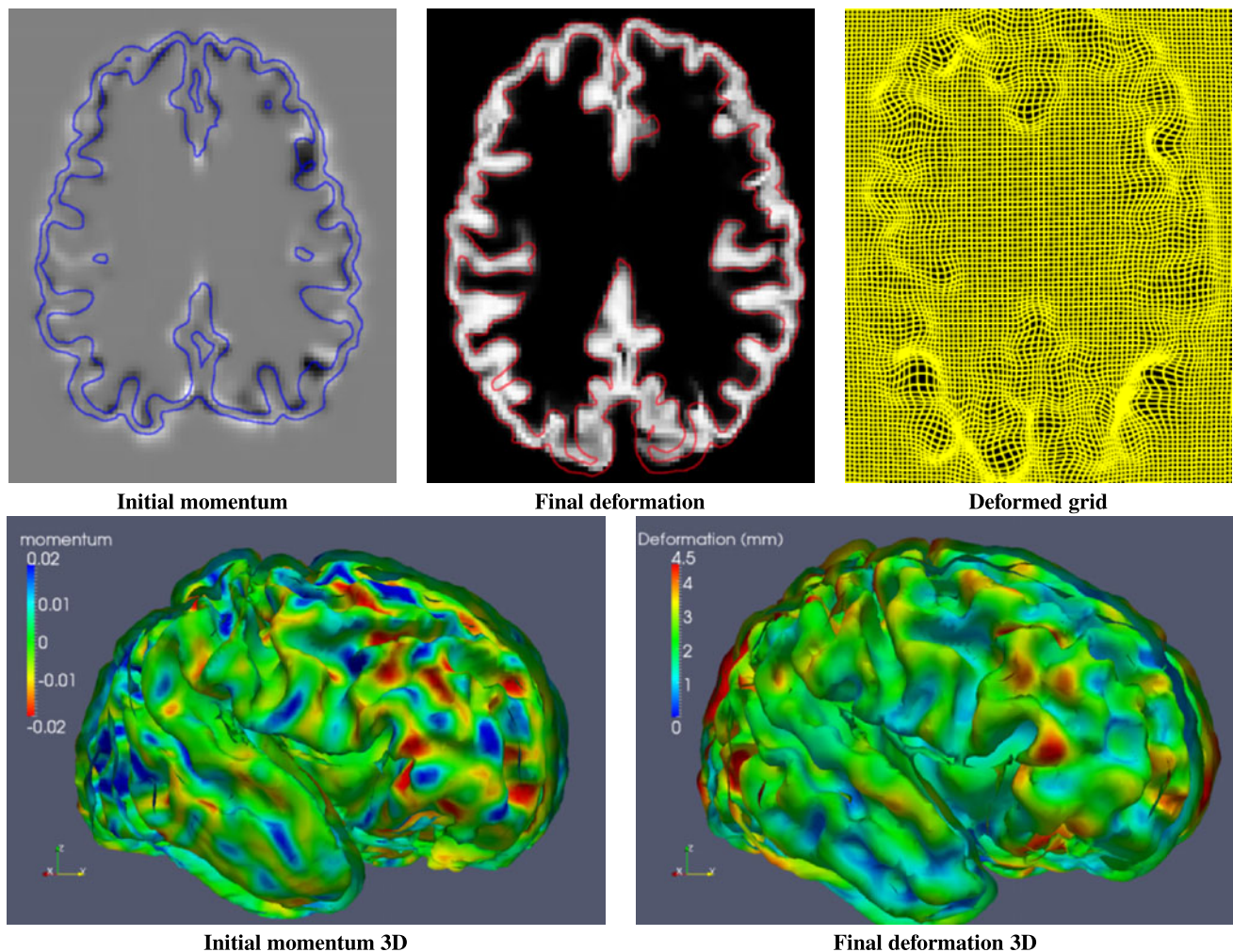
Initial momentum   **Final deformation**   **Deformed grid**

Initial momentum 3D   **Final deformation 3D**

Fig. 6 Estimated initial momentum and matching quality when registering G_{33} on G_{36} using *MK*. (*Top–left*) Initial momentum in a slice of the 3D volume. *The blue curve* represents the cortical surface in G_{33}. (*Top–centre*) $G_{33} \circ \phi^{-1}$ in the same slice. *The red curve* represents the cortical surface in G_{36}. (*Top–centre*) Corresponding deformed grid.

(*Bottom–left*) Estimated initial momentum in the internal cortical surface of G_{33}. (*Bottom–Right*) Internal cortical surface of $G_{33} \circ \phi^{-1}$. *The colors* represent the amplitude of the deformations through the flow of diffeomorphism ϕ_t between $t = 0$ and $t = 1$

comparison with statistics on global descriptors. Especially, we expect meaningful results when the analysis of the initial momentum is performed with a multi-scale approach on the kernel, in the spirit of Risser et al. (2010).

There are several numerical considerations that led us to select the formulation described in this paper. To begin with, we chose to compute with the diffeomorphism itself as a prognostic variable, and to evaluate the diffeomorphism applied to the image. One alternative, as implemented in Hart et al. (2009), is to compute only with the image and corresponding conjugate momentum, solving the advection/continuity equations in Eulerian form for both image and momentum. One reason for choosing our formulation is that the adjoint variables are not computed from Eulerian advection/continuity equations: they are instead computed directly from the diffeomorphism and hence are not them-

selves subject to CFL-type conditions that impose timestep constraints on explicit advection methods. In our experiments we have found that the CFL condition on the adjoint equations (5) is more demanding than the CFL condition on the forward shooting model.

A second reason for choosing our formulation is that in many cases the image is not smooth (binary images, for example) whilst the diffeomorphism is smooth. If one is numerically solving the Eulerian advection equation for a non-smooth image, then it is necessary to use a monotonic/bounded high-order advection scheme, such as finite volume methods with flux limiters (Le Veque 2002), high-order WENO schemes (Shu 2009), or high-order discontinuous Galerkin methods with appropriate slope limiters (Cockburn and Shu 2001). Since the diffeomorphism remains smooth, numerical oscillations are much less likely

to appear and less numerical dissipation will be required to control them. In particular, cheaper, less sophisticated high-order schemes (such as the Min-Mod limiter used in this paper) can be used.

A possible extension to this approach is to use a Gauss-Newton-Krylov method to solve the nonlinear least squares problem. This consists of outer iterations which require the full nonlinear system (2), combined with inner Krylov iterations in which the linearized of (2) around this trajectory is solved forwards, followed by the adjoint equations solved backwards. It was shown in Gratton et al. (2007) that using under-converged Krylov solves for inverting the Jacobian matrix can result in efficient convergent algorithms under quite general conditions, and so it should be possible to replace the adjoint solve by the approximation developed in this paper. A similar approach could also be used for the forward linear solve, producing forward and backward iterations that form an exact adjoint pair. This could result in a very efficient algorithm since these solves would not be subject to the CFL constraint and so approximate matrix inversions could be obtained very quickly. This will be the subject of future investigations. Further improvements to our algorithm could also include higher-order timestepping schemes (such as the structure preserving Runge-Kutta schemes used in Cockburn and Shu (2001)), and nonlinear conjugate gradient methods rather than steepest descent.

We have shown that our registration algorithm can achieve efficient results in comparison with the algorithm of Beg et al. (2005) with the advantage of geodesicity at every iteration of the optimization process. The matching performances of our algorithm is most of the time sufficient for anatomical images. We underline that this algorithm has not been designed to compete with the best registration algorithms in terms of quality of matching. In fact, this matching quality could be easily improved at the expense of the statistical power of this method.

As detailed in appendix, this method can be adapted in a straightforward manner to match densities instead of images and the corresponding numerical algorithm would follow the same steps. Not unexpectedly, one could also apply the variational data assimilation method to the metamorphosis framework. However, the numerical integration of the adjoint equations would be the most important issue. This metamorphosis model may lead to promising applications for the quantitative characterization of longitudinal biological evolutions that are non-diffeomorphic such as tumor growth.

Maybe even more promisingly, this work also opens up the way for second-order models to interpolate longitudinal evolutions of biological images. Modeling shape evolutions and biological growth is now a topic of growing interest Grenander et al. (2007, 2006) or Khan and Beg (2008), Durrleman et al. (2009). The current state of the art is the piecewise-geodesic interpolation proposed in Niethammer et al. (2009), Khan and Beg (2008) in the context of *LD-DMM* or kernel interpolation on the time variable for shape regression Davis et al. (2010). A second-order model has been recently presented in Trouvé and Vialard (2010) in the case of landmarks, which introduces the physical idea of minimizing the acceleration of the evolution in order to obtain smooth temporal interpolations of time-series of shapes. Our registration algorithm is the first step towards a generalization to 3D images of such a second-order model, which deeply relies on the Hamiltonian formulation of the geodesic evolution.

Note finally that the proposed algorithm is available in the project uTIlzReg on sourceforge http://sourceforge.net/projects/utilzreg/.

Appendix A: Adjoint Equations in the Case of Densities

As written in Younes (2007), the geodesic shooting equations are dual of the system (2) in the sense that P and I are exchanged, we denote by D the density to be matched and P the corresponding momentum.

$$
\begin{cases}
\partial_t D + \nabla \cdot (Dv) = 0, \\
\partial_t P + \nabla P \cdot v = 0, \\
v + K \star D\nabla P = 0.
\end{cases}
\tag{20}
$$

The functional to be minimized can be defined as

$$
\begin{aligned}
\mathcal{S}(P(0)) = \frac{\lambda}{2} \langle \nabla P(0)D(0), K \star \nabla P(0)D(0)\rangle_{L^2} \\
+ M(D(1)),
\end{aligned}
\tag{21}
$$

where M is a similarity measure on the deformed density $D(1)$. The corresponding version of Theorem 1 is, exchanging again P, $\hat{P}$ to D, $\hat{D}$ and I, $\hat{I}$ to P, $\hat{P}$

Theorem 3 *The gradient of $\mathcal{S}$ is given by*:

$$
\nabla_{P(0)}\mathcal{S} = -\lambda \nabla \cdot (D(0) \cdot K \star (D(0)\nabla P(0))) - \hat{P}(0)
$$

where $\hat{P}(0)$ is given by the solution of the following PDE solved backward in time:

$$
\begin{cases}
\partial_t \hat{P} + \nabla \cdot (v\hat{P}) + \nabla \cdot (D\hat{v}) = 0, \\
\partial_t \hat{D} + v \cdot \nabla \hat{D} - \nabla P \cdot \hat{v} = 0, \\
\hat{v} + K \star (\hat{P}\nabla P - D\nabla \hat{D}) = 0,
\end{cases}
\tag{22}
$$

subject to the initial conditions:

$$
\begin{cases}
\hat{D}(1) = \dfrac{\delta M}{\delta D}(D(1)), \\
\hat{P}(1) = 0,
\end{cases}
\tag{23}
$$

and $P(t)$, $D(t)$ are the solution of the shooting system (20) *for the initial conditions $D(0)$, $P(0)$.*

Obviously the corresponding integral formulation in Theorem 2 is

$$
\begin{cases}
\hat{D}(t) = \hat{D}(1) \circ \phi_{t,1} - \int_t^1 [\nabla P(s) \cdot \hat{v}(s)] \circ \phi_{t,s}\, ds, \\
\hat{P}(t) = \mathrm{Jac}(\phi_{t,1})\hat{P}(1) \circ \phi_{t,1} \\
\qquad + \int_t^1 \mathrm{Jac}(\phi_{t,s})[\nabla \cdot (D(s)\hat{v}(s))] \circ \phi_{t,s}\, ds.
\end{cases}
\tag{24}
$$

with:

$$
\begin{cases}
\hat{v}(t) = K \star [D(t)\nabla\hat{D}(t) - \hat{P}(t)\nabla P(t)], \\
D(t) = \mathrm{Jac}(\phi_{t,0})D(0) \circ \phi_{t,0}, \\
P(t) = P(0) \circ \phi_{t,0},
\end{cases}
\tag{25}
$$

using the same notations of Theorem 2. Last, let us underline that the case of metamorphosis can be treated in the same way, however the numerical treatment appears more involved.

Appendix B: Integral Formulation

We give here the main points for a proof of Theorem 2 since it follows from usual arguments.

Proof The verification of the integral formula (24) is a straightforward calculation. We first prove the existence and uniqueness of solutions for small times. To shorten the notations, we denote by $z \doteq \begin{pmatrix} \hat{I} \\ \hat{P} \end{pmatrix}$ and the action of the diffeomorphism $\phi_{t,s}$ on a generic z is denoted by: $\phi_{t,s} \cdot z \doteq \begin{pmatrix} \hat{I} \circ \phi_{t,s} \\ \mathrm{Jac}(\phi_{t,s})\hat{P} \circ \phi_{t,s} \end{pmatrix}$ and we introduce Ψ the linear morphism in $C^0([0,1], H^1 \times H^1)$

$$
[\Psi(z)](t) = \phi_{t,1} \cdot z(1) + \int_t^1 \phi_{t,s} \cdot G_t(z(t))\, ds,
\tag{26}
$$

where $G_t(z(t)) \doteq \begin{pmatrix} \nabla \cdot (P\hat{v}) \\ -\nabla I \cdot \hat{v} \end{pmatrix}$. Since $\hat{v} + K \star \hat{I}\nabla I - P\nabla\hat{P} = 0$, Ψ is a linear map. Remark that if the kernel is C^2 and $I, P \in H^2(\Omega)$ which is implied by the assumptions of Theorem 2, we get that $G_t(z(t)) \in H^1(\Omega) \times H^1(\Omega)$ so that the image of Ψ lies in $C^0([0,1], H^1(\Omega) \times H^1(\Omega))$, which is a Banach space for the sup norm on $H^1(\Omega) \times H^1(\Omega)$.

Remark that we have

$$
\|G_t(z)\|_{H^1} \le M\|z\|_{H^1},
\tag{27}
$$

for a constant M that only depends on the sup norm in H^2 of I and P which is bounded on the whole time interval $[0,1]$. As a consequence, the map Ψ is a contraction when

restricting it to $C^0([s, s+t_0], H^1(\Omega) \times H^1(\Omega))$ for t_0 small enough and for any $s \in [0, 1-t_0]$. An iterative application of the Banach fixed point theorem gives the desired result. $\square$

References

Allassonnière, S., Trouvé, A., & Younes, L. (2005). Geodesic shooting and diffeomorphic matching via textured meshes. In *LNCS: Vol. 3757. Proc. of EMMCVPR*. Berlin: Springer.

Allassonnière, S., Kuhn, E., Trouvé, A., & Amit, Y. (2006). Generative model and consistent estimation algorithms for non-rigid deformable models. In *ICASSP 2006 Proceedings* (pp. 14–19).

Allassonnière, S., Amit, Y., & Trouvé, A. (2007). Towards a coherent statistical framework for dense deformable template estimation. *Journal of the Royal Statistical Society. Series B*, 69(1), 3–29.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, 38, 95–113.

Avants, B. B., Epstein, C. L., Grossman, M., Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12, 26–41.

Beg, M. F., Helm, P. A., McVeigh, E., Miller, M. I., & Winslow, R. L. (2004). Computational cardiac anatomy using MRI. *Magnetic Resonance in Medecine*, 52(5), 1167–1174.

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

Cockburn, B., & Shu, C. W. (2001). Runge-Kutta discontinuous Galerkin methods for convection-dominated problems. *Journal of Scientific Computing*, 16(3), 173.

Davis, B. C., Fletcher, P. T., Bullitt, E., & Joshi, S. (2010). Population shape regression from random design data. *International Journal of Computer Vision*, 90(2), 255–266.

Du, J., Younes, L., & Qiu, A. (2011). Whole brain diffeomorphic metric mapping via integration of sulcal and gyral curves, cortical surfaces, and images. *NeuroImage*, 56(1), 162–173.

Dupuis, P., Grenander, U., & Miller, M. (1998). Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*

Durrleman, S., Pennec, X., Trouvé, A., Gerig, G., & Ayache, N. (2009). Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets. In *MICCAI* (Vol. 5761, pp. 297–304).

Fletcher, P. T. (2004). *Statistical variability in nonlinear spaces: application to shape analysis and dt-mri*. PhD thesis, Department of Computer Science, University of North Carolina.

Fletcher, P. T., Lu, C., Pizer, M., & Joshi, S. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions Medical Imaging*, 995–1005.

Fletcher, P. T., Venkatasubramanian, S., & Joshi, S. (2008). Robust statistics on Riemannian manifolds via the geometric median. In *Computer Vision and Pattern Recognition* (pp. 1–8).

Gratton, S., Lawless, A., & Nichols, N. (2007). Approximate Gauss-Newton methods for nonlinear least squares problems. *SIAM Journal on Optimization*, 18(1), 106–132.

Grenander, U. (1993). *General pattern theory*. Oxford: Oxford Science Publications.

Grenander, U. & Miller, M. I. (1998). Computational anatomy: an emerging discipline. *Quarterly of Applied Mathematics*, *LVI*(4), 617–694.

Grenander, U., Srivastava, A., & Saini, S. (2006). Characterization of biological growth using iterated diffeomorphisms. In *ISBI* (pp. 1136–1139).

Grenander, U., Srivastava, A., & Saini, S. (2007). A pattern-theoretic characterization of biological growth. *IEEE Transactions on Medical Imaging*, *26*(5), 648–659.

Hart, G. L., Zach, C., & Niethammer, M. (2009). An optimal control approach for deformable registration. In *Computer Vision and Pattern Recognition, Workshop* (pp. 9–16).

Helm, P. A., Younes, L., Beg, M. F., Ennis, D. B., Leclercq, C., Faris, O. P., McVeigh, E., Kass, D., Miller, M. I., & Winslow, R. L. (2006). Evidence of structural remodeling in the dyssynchronous failing heart. *Circulation Research*, *98*, 125–132.

Holm, D. D., Trouvé, A., & Younes, L., (2009, to appear). The Euler Poincaré theory of metamorphosis. *Quarterly of Applied Mathematics*.

Jiang, G. S., & Shu, C. W. (1996). Efficient implementation of weighted ENO schemes. *Journal of Computational Physics, 126*.

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

Khan, A., & Beg, M. (2008). Representation of time-varying shapes in the large deformation diffeomorphic framework. In *ISBI* (pp. 1521–1524).

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M. C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R., Mann, J., & Parseya, R. V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, *46*(3), 786–802.

Le Dimet, F. X., & Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus Series A*, *38*, 97–110.

Le Dimet, F. X., Ngodock, H. E., & Navon, I. M. (1997). Sensitivity analysis in variational data assimilation. *Journal of the Meteorological Society Japan*, 145–155.

Le Veque, R. J. (2002). *Finite volume methods for hyperbolic problems*. Cambridge: Cambridge University Press.

Lee, J. M. (1997). *Riemannian manifolds an introduction to curvature. Graduate texts in mathematics* Berlin: Springer.

Ma, J. Miller, M. I. Younes, L. (2010). A Bayesian generative model for surface template estimation. *International Journal of Biomedical Imaging*, *2010*, 14.

Marsland, S., & McLachlan, R. I. (2007). A Hamiltonian particle method for diffeomorphic image registration. In *Lecture Notes in Computer Science: Vol. 4548. Proceedings of Information Processing in Medical Images* (pp. 396–407). Berlin: Springer.

Miller, M., & Younes, L. (2001). Group actions, homeomorphisms, and matching: a general framework. *International Journal of Computer Vision*, *41*, 61–84.

Miller, M. I., Trouvé, A., & Younes, L. (2002). On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, *4*, 375–405.

Miller, M. I., Trouvé, A., & Younes, L. (2006). Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, *24*(2), 209–228.

Murgasova, M., Aljabar, P., Srinivasan, L., Counsell, S. J., Doria, V., Serag, A., Gousias, I. S., Boardman, J. P., Rutherford, M. A., Edwards, A. D., Hajnal, J. V., & Rueckert, D. (2011). A dynamic 4d probabilistic atlas of the developing brain. *NeuroImage*, *54*(4), 2750–63.

Ngodock, H. E. (2005). Efficient implementation of covariance multiplication for data assimilation with the representer method. *Ocean Modelling*, *8*(3), 237–251.

Niethammer, M., Hart, G., & Zach, C. (2009). An optimal control approach for the registration of image time-series. 2427–2434.

Pennec, X. (2009). Statistical computing on manifolds: From Riemannian geometry to computational anatomy. *Emerging Trends in Visual Computing*, *5416*, 347–386.

Risser, L., Vialard, F. X., Wolz, R., Holm, D. D., & Rueckert, D., (2010). Simultaneous fine and coarse diffeomorphic registration: application to the atrophy measurement in Alzheimer's disease. In *International conference on medical image computing and computer assisted intervention (MICCAI). Lecture notes in computer science*.

Risser, L., Vialard, F., Wolz, R., Murgasova, M., Holm, D., & Rueckert, D. (2011). Simultaneous multiscale registration using large deformation diffeomorphic metric mapping *IEEE Transactions on Medical Imaging*, *99*, 1. doi:10.1007/s11263-011-0481-8

Sasaki, Y. (1970). Some basic formalism in numerical variational. *Monthly Weather Review*, *98*, 875–883.

Shu, C. W. (2009). High order weighted essentially non-oscillatory schemes for convection dominated problems. *SIAM Review*, *51*(1), 82.

Singh, N., Fletcher, P. T., Preston, J. S., Ha, L., King, R., Marron, J. S., Wiener, M., & Joshi, S. (2010). Multivariate statistical analysis of deformation momenta relating anatomical shape to neuropsychological measures. In T. Jiang, N. Navab, J. P. Pluim, & Viergever M. A. (Eds.), *MICCAI* (pp. 529–537). Berlin: Springer.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, *17*(1), 87–97.

Trouvé, A. (1998). Diffeomorphism groups and pattern matching in image analysis. *International Journal of Computer Vision*, *28*(3), 213–221.

Trouvé, A., & Vialard, F. X. (2010, accepted for publication). Shape splines and stochastic shape evolutions: a second order point of view. *Quarterly of Applied Mathematics*.

Trouvé, A., & Younes, L. (2005). Local geometry of deformable templates. *SIAM Journal of Mathematical Analysis*.

Vaillant, M., Miller, M. I., Trouvé, A., Younes, L. (2004). Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, *23*(S1), S161–S169.

Vercauteren, T., Pennec, X., Perchant, A., & Ayache, N. (2009). Diffeomorphic demons: efficient non-parametric image registration. *NeuroImage*, *45*(1), S61–S72.

Vialard, F. X. (2009). *Hamiltonian approach to shape spaces in a diffeomorphic framework: from the discontinuous image matching problem to a stochastic growth model*. PhD thesis, Ecole Normale Supérieure de Cachan.

Vialard, F. X., & Santambrogio, F. (2009). Extension to BV functions of the large deformation diffeomorphisms matching approach. *Comptes Rendus Mathematique*, *347*(1-2), 27–32.

Wang, L., Beg, M. F., Ratnanather, J. T., Ceritoglu, C., Younes, L., Morris, J., Csernansky, J., & Miller, M. I. (2006). Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type. *IEEE Transactions on Medical Imaging*, *26*, 462–470.

Younes, L. (2007). Jacobi fields in groups of diffeomorphisms and applications. *Quarterly of Applied Mathematics*, *65*, 113–134.

Younes, L. (2010). *Shapes and diffeomorphisms*. Berlin: Springer.

Younes, L., Arrate, F., & Miller, M. I. (2009). Evolutions equations in computational anatomy. *NeuroImage*, *45*(1), S40–S50. Mathematics in Brain Imaging (Supplement 1)

Zeidler, E. (1995). Applied functional analysis: main principles and their applications. *Applied Mathematical Sciences*, 109.