Geodesic Shape Regression with Multiple Geometries and Sparse Parameters

James Fishbaugh[a,*], Stanley Durrleman[b], Marcel Prastawa[c], Guido Gerig[a]

[a]*Department of Computer Science and Engineering, NYU Tandon School of Engineering, NY, USA*
[b]*Brain and Spine Institute (ICM), INRIA, Paris, France*
[c]*Icahn School of Medicine, Mount Sinai, NY, USA*

Abstract

Many problems in medicine are inherently dynamic processes which include the aspect of change over time, such as childhood development, aging, and disease progression. From medical images, numerous geometric structures can be extracted with various representations, such as landmarks, point clouds, curves, and surfaces. Different sources of geometry may characterize different aspects of the anatomy, such as fiber tracts from DTI and subcortical shapes from structural MRI, and therefore require a modeling scheme which can include various shape representations in any combination.

In this paper, we present a geodesic regression model in the large deformation (LDDMM) framework applicable to multi-object complexes in a variety of shape representations. Our model decouples the deformation parameters from the specific shape representations, allowing the complexity of the model to reflect the nature of the shape changes, rather than the sampling of the data. As a consequence, the sparse representation of diffeomorphic flow allows for the straightforward embedding of a variety of geometry in different combinations, which all contribute towards the estimation of a single deformation of the ambient space. Additionally, the sparse representation along with the geodesic constraint results in a compact statistical model of shape change by a small number of parameters defined by the user. Experimental validation on multi-object complexes demonstrate robust model estimation across a variety of parameter settings. We further demonstrate the utility of our method to support the analysis of derived shape features, such as volume,

---

*Corresponding author: james.fishbaugh@nyu.edu

and explore shape model extrapolation.

Our method is freely available in the software package `deformetrica` which can be downloaded at `www.deformetrica.org`.

1. Introduction

The analysis and monitoring of change over time is fundamental to many problems in medicine, where anatomical change is often driven by a continuous dynamic process, such as in early childhood development, aging, or disease progression. Measuring and understanding change over time is required to assess development. For example, head circumference is measured during pediatric examination and compared to a standardized model to determine if a child is developing along a normative trajectory. In addition to assessment, measuring change over time is essential to monitor treatment, or the effectiveness of drug therapy. Furthermore, the analysis of change can improve our understanding of the time-course of psychiatric disorders or pathologies, which may provide additional information to steer treatment efforts. For example, neuroscientists discovered a decrease of striatum volume (Paulsen et al., 2006, 2010) in Huntington's disease, a finding which has led to research efforts into gene therapy targeted at the putamen (Benraiss & Goldman, 2011).

Magnetic resonance imaging (MRI) is now ubiquitous in clinical practice, and represents a powerful tool to monitor and measure change in anatomical tissue in vivo. Medical imaging technology is improving in fidelity, as well as becoming more readily available, which has led to the proliferation of imaging studies. Such studies may be cross-sectional, where imaging data belongs to a representative population distributed in time, but each subject is scanned only once. More recently, the focus has shifted towards subject-specific and individualized analysis, where data comes from a single subject, or large scale longitudinal studies where serial scans are acquired from the same subjects over time. In any case, such time-indexed imaging databases provide a rich environment for research activity, and are essential for improving our understanding of various disorders and pathologies.

Medical imaging studies, either cross-sectional or longitudinal, rely on medical images which represent measurements sparsely distributed in time.

2

Such images can be thought of as snapshots or frames of the underlying continuous sequence. Further, MR images represent imperfect observations of anatomy, with noise introduced by the scanner, image reconstruction, patient movement, among other possible sources. What is needed are statistical models to capture the trend in the data, which also characterize the underlying continuous anatomical change. Such statistical models can be used to age match subjects, or alternatively, match subjects along disease progression, to temporally align imaging data with clinical scores not taken at the same time as the image acquisition, for interpolation or extrapolation to generate new unobserved shapes, or as a mathematical representation for statistical hypothesis testing. The most ambitious is the construction of normative models of development and aging for comparison and monitoring patient progress.

In many clinical studies, measurements such as volume are extracted and regression models are estimated absent any imaging or geometric information. Typically, linear models are chosen for convenience and simplicity, rather than motivated by the changes in the anatomy. We advocate for modeling at a higher level, where our understanding of anatomical change can be introduced. For example, a shape regression model based on the flow of diffeomorphisms guarantees structures cannot be created, destroyed, holes introduced, or folded over on themselves, which are desirable and required properties for many clinical applications. Furthermore, shape models support traditional analysis of scalar measurements derived from morphometric features, as any value of interest can be extracted continuously from the shape trajectory.

Just as it is important to have an anatomically realistic model of change, it is desirable for a model to include multiple sources of geometric information as a multi-object complex. There is a large variety of geometric information which can be extracted from medical images such as: surfaces, curves, point clouds, and landmarks. As we will see later in the paper, shapes may derive from different modalities, such as subcortical shapes from structural imaging and white matter connections from diffusion tensor imaging. The different sources of geometry may complement each other, giving a more complete description of change over time. Due to the wide variety of shape representations derived from medical imaging data, we desire a model which is independent of a given shape parameterization, instead it must be able to handle numerous shape representations arranged in various combinations.

In the field of medical image analysis, the problem of regression has re-

ceived considerable attention over the last 10 years, as regression is a necessary tool in many longitudinal statistical analysis pipelines (Datar et al., 2012; Durrleman et al., 2013b; Fishbaugh et al., 2012; Hart et al., 2010; Singh et al., 2016). There are a variety of methods introduced in the general Riemannian setting, such as geodesic regression (Fletcher, 2011, 2013). This idea was extended to polynomial regression (Hinkle et al., 2014), with geodesics being a special case. In addition to linear models, other work includes nonlinear regression in the general Riemannian setting (Banerjee et al., 2015). The application of these methods is typically finite dimensional manifolds, most commonly shapes represented in Kendall shape space (Kendall, 1984). There has also been attention towards methods focused on a specific manifold, such as the Grassmannian (Hong et al., 2014, 2016).

Regression on medical images has also been explored, including the extension of kernel regression to images (Davis et al., 2007) and splines for diffeomorphic image regression (Singh et al., 2015). Geodesic regression has been developed for imaging data in Niethammer et al. (2011) which leverages the initial momenta formulation of the EPDiff equation (Vialard et al., 2012). In Niethammer et al. (2011), the momenta are a scalar field of the same dimension of the image, as such, they can be thought of as attached to each voxel. The direction of the initial momenta is orthogonal to the gradient of the deforming baseline image (Miller et al., 2006). Rather than use scalar initial momenta, the work of Singh et al. (2013) introduces a vector formulation to ease the estimation of the baseline image. The optimization procedure need not jointly compute both baseline image and initial momenta, rather only momenta are estimated, and a new baseline image is computed in turn. This leads to faster convergence in terms of the number of iterations of gradient descent.

Regression on geometric structures extracted from imaging has also been explored. In Vialard & Trouvé (2012), a nonparametric spline model is proposed as perturbations of a geodesic path. In Datar et al. (2009), each landmark point in correspondence across the population are assumed to follow a linear trajectory. As with the Riemannian methods, these methods are applicable to shapes represented as landmarks.

Several regression methods have been proposed for multi-object complexes containing a variety of shape representations, such as piecewise-geodesic regression (Durrleman et al., 2009) and regression based on controlled acceleration (Fishbaugh et al., 2011). However, the methods do not provide a solution for estimating a baseline shape, rather the regression is constrained

4

to start from the observation earliest in time. The dimensionality of the models is directly related to the sampling of the data, as the model parameters in Durrleman et al. (2009) and Fishbaugh et al. (2011) are located on the vertices of the shapes. Furthermore, the models are not based on a shooting formulation from initial conditions, instead requiring model parameters at every shape point *and* every time-point in the discretization. As a consequence, many thousands of parameters are needed to describe shape evolution.

To summarize, currently available methods are either limited to specific data types, such as landmarks in correspondence or images, or require a huge number of deformation parameters. The high dimensionality can be due to a shared parameterization between deformation and shape, as is the case of momenta attached to image voxels or shape vertices, or to nonparametric models whose parameters are functions of time. What is lacking is a model of shape change which is flexible to the data representation, which is also a compact generative model which describes shape evolution with a small number of parameters.

In this paper, we present a geodesic shape regression model in the large deformation (LDDMM) framework that incorporates multiple sources of geometry in different combinations as multi-object complexes which drive the estimation of a single continuous deformation of the ambient space. The proposed generative model uses a sparse representation of diffeomorphisms, which describe complex nonlinear changes over time with a small number of model parameters defined by the user. By analogy with simple linear regression, we estimate an intercept as the initial baseline shape configuration, as well as a slope, which in our model is the initial momenta vectors as well as their *location.* We derive the Euler-Lagrange equations and propose a gradient descent algorithm for model estimation as well as systematic experimentation to expand on our previous published conference work (Fishbaugh et al., 2013b).

2. Shape Regression

In its most basic form, regression analysis involves exploring the relationship between a dependent variable and one or more independent variables. The most ubiquitous model is simple linear regression, where we assume a linear relationship between one dependent and one independent variable. Given the parametric form of a line $y = mx + b$, linear regression can be

expressed as

$$E(m,b) = \sum_{i=1}^{N} \Big((mx_i + b) - y_i \Big)^2 \tag{1}$$

given measurements $[y_1, y_2, ..., y_n]$ and corresponding explanatory variables $[x_1, x_2, ..., x_n]$. Model estimation involves finding values of m and b which minimize the regression criterion (1). Specifically, we compute the slope m and y-intercept b which minimize the sum-of-squared distance between the line and the observations.

The concept of regression extends naturally to the case of shape observations, which we call shape regression. The problem of shape regression involves finding correlations between shape configuration and a continuous scalar parameter such as age, disease progression, drug delivery, or cognitive scores. Intuitively, we seek the continuous sequence of shapes that best explains the observed shapes, in a least squares sense to be defined later. The concept of shape regression is illustrated in Figure. 1.

Consider a set of shape observations $\mathbf{O}_{t_i}$ at times t_i. Shape regression can be described by the generic regression criterion

$$E(\psi_t) = \sum_{t_i} \mathrm{D}(\psi_{t_i}(\mathbf{O}_{t_0}), \mathbf{O}_{t_i})^2 - \gamma \mathrm{Reg}(\psi_t) \tag{2}$$

where ψ_t continuously transforms baseline shape $\mathbf{O}_{t_0}$ over time to match shape observations $\mathbf{O}_{t_i}$ with respect to a shape similarity metric D, Reg controls the regularity of the transformation, and γ balances data matching
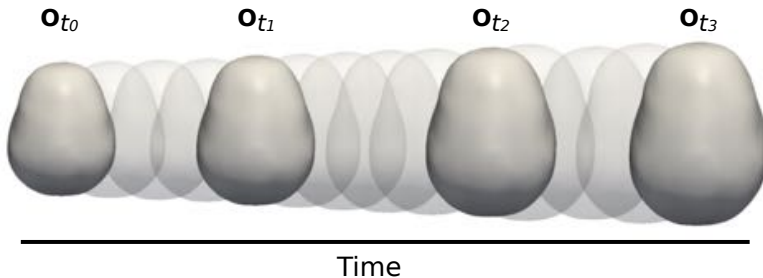

Figure 1: An illustration of shape regression. Four time-indexed observations of the intracranial surface are shown as solid surfaces. The objective of shape regression is to estimate the continuous evolution of shape (transparent surfaces) which best explains the observed data.

with regularity. As with simple linear regression mentioned above, model fitting involves finding parameters of ψ_t which minimizes the criterion (2).

Shape regression models of this form require two essential components:

1. A model of deformation which gives the form and parameterization of the time varying deformation ψ_t.
2. A shape representation and corresponding shape similarity metric D which takes shapes as input and returns a scalar value indicating the similarity (or dissimilarity) of the shapes.

2.1. Deformation model

The deformation model gives the form and parameterization of the time-varying deformations, thus providing the mechanism for computing and applying deformations.

2.1.1. Large deformation model

The large deformation model is commonly referred to as the Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework. The idea is to allow for large deformations while guaranteeing a well behaved transformation. Namely, the transformation is a diffeomorphism which is continuously differentiable with differentiable inverse. This is built on the work of Trouvé (1995) and Dupuis & Grenander (1998), and has been influential in registration and atlas building (Avants et al., 2008; Joshi et al., 2004; Marsland & Twining, 2004). We briefly cover the mathematical foundations of LDDMM.

Let the ambient space be represented by Ω, an open subset of R^d, and let $\mathbf{V}$ represent a Hilbert space of smooth vector fields on $\mathbf{\Omega}$. A norm is assigned to $\mathbf{V}$ through association with a differential operator L

$$\|\mathbf{v}\|_{\mathbf{V}}^{\mathbf{2}} = \int_{\mathbf{\Omega}} (\mathbf{Lv(x)}, \mathbf{v(x)})\mathbf{dx}, \tag{3}$$

where L induces the structure of V and relates velocity v to momenta $\boldsymbol{\alpha} = \mathbf{Lv}$ in dual space $\mathbf{V}^*$. The operator L has inverse $K = L^{-1}$ where K is a smooth kernel operator such that

$$(\mathbf{Kv})(\mathbf{x}) = \int_{\mathbf{\Omega}} \mathbf{K(x,y)v(y)dy}, \tag{4}$$

where K maps momentum vector $\boldsymbol{\alpha}$ to $\mathbf{v} = \mathbf{K\alpha}$. In practice we deal directly with the kernel K, for example the scalar Gaussian kernel $K(x,y) = \exp(-\|x-y\|^2/\sigma^2)$.

The main idea of the LDDMM framework is to parameterize diffeomorphisms of Ω by time-varying velocity fields $v_t : \Omega \rightarrow R^d$, $t \in [0,1]$. The diffeomorphism ϕ_t is then defined by the ordinary differential equation

$$\dot{\phi}_t = v_t(\phi_t), \tag{5}$$

given the construction of V as a Reproducing Kernel Hilbert Space (RKHS) as shown above.

Solving this equation (5) generates a flow of diffeomorphisms $\phi_t : \Omega \rightarrow \Omega$ (transformations of the space to itself) beginning with the identity transformation $\phi_0 = Id$ and ending at $\phi_1 = \phi_0 + \int_0^1 v_t(\phi_t)\, dt$. Here the variable t need not be interpreted as time, but rather it serves as a variable of integration to generate a path in the space of diffeomorphisms. This is an important distinction here, as later in the article the variable t will usually correspond to physical units of time.

Equation (3) defines a Riemannian metric on the space of diffeomorphisms, which provides a way to compute distance between the identity diffeomorphism and an arbitrary diffeomorphism φ as

$$d(Id, \varphi) = \inf \{ \int_0^1 \|v_t\|_V^2 \, dt \ : \ \phi_1^{v_t} = \varphi \}. \tag{6}$$

Equation (6) can then be used to compute the distance between any two diffeomorphisms

$$d(\phi, \varphi) = d(Id, \varphi \circ \phi^{-1}). \tag{7}$$

2.1.2. Sparse Parameterization of Diffeomorphisms

Following the framework of landmark matching in Joshi & Miller (2000), a low dimensional parameterization of diffeomorphisms was introduced in Durrleman et al. (2011). This sparse parameterization of diffeomorphisms has been further explored and applied to problems in image and shape analysis (Durrleman et al., 2012, 2013a, 2014; Fishbaugh et al., 2013a, 2014).

Let $\mathbf{c}(t) = \{c_1(t), ..., c_{N_c}(t)\}$ be the spatial coordinates of a set of N_c control points for each time t. Associated with each control point is a momenta vector, the set of which is denoted $\boldsymbol{\alpha}(t) = \{\alpha_1(t), ...\alpha_{N_c}(t)\}$. Together, the control points and momenta represent the state of the system $\mathbf{S}(t) = \{\mathbf{c}(t), \boldsymbol{\alpha}(t)\}$. The finite set of control point/momenta pairs define the

time-varying velocity field everywhere in space as

$$\dot{x}(t) = v_t(x) = \sum_{p=1}^{N_c} K(x, c_p(t))\alpha_p(t), \tag{8}$$

where K is a Gaussian kernel $K(x, y) = \exp(-||x - y||^2/\sigma_V^2)$ defining the RKHS and corresponding metric properties through the choice of σ_V. We write (8) in short as $\dot{x}(t) = G(x(t), \mathbf{S}(t))$ and refer to this as the *flow* equation.

The time-varying velocity field v_t then builds a flow of diffeomorphisms as in the LDDMM framework by integrating the differential equation

$$\dot{\phi}_t(x(t)) = v_t(\phi_t(x(t))) = \sum_{p=1}^{N_c} K(\phi_t(x(t)), c_p(t))\alpha_p(t) \tag{9}$$

given initial value $x(0)$. Additionally, the location of control points evolve in time according to the flow equation (8) written as

$$\dot{c}_i(t) = \sum_{p=1}^{N_c} K(c_i(t), c_p(t))\alpha_p(t) \tag{10}$$

given initial values $c_i(0)$. The flow equation (8) defines the trajectory of any location in space x, with equation (10) being a special case, evaluated only at control point locations. The trajectory $x(t)$ is computed by solving (9), which is fully defined by the control point and momenta pairs $\mathbf{S}(t)$.

This representation of diffeomorphic flow has two main benefits. First, it provides a sparse representation of dense deformations. A low dimensional parameterization is beneficial for statistical analysis, due to the problem of high dimensionality and low sample size. It is also beneficial to reduce noise in the description of shape variability, as the true variability is likely parameterized by far fewer parameters than the number of image voxels, for instance. The second benefit of the control point framework is it decouples deformation parameters from any specific shape representation. Freeing the control points from the vertices of the shape allows the deformation parameters to be concentrated where the most dynamic changes occur.

Consider the example shown in Figure 2, where a circle is deformed to match an ellipse. In the case of dense deformations, the momenta vectors are initialized on the vertices of circle and remain attached to the shape points after model estimation. In contrast, the sparse parameterization initializes

a collection of momenta vectors in the ambient space, and their position are estimated in addition to their direction. As a consequence, the sparse parameterization is able to characterize the left/right expansion of the circle with 8 parameters, as opposed to 32 in the dense parameterization.

2.2. Shape Representation

The geodesic model presented here separates the deformation model from the shape representation. The residuals measured with respect to the shape metric drive the estimation of model parameters, but the method is not limited to a specific shape metric. Indeed, any shape metric can be added to our algorithm, provided a mechanism to compute the shape metric as well as the gradient. You will note our formulation is written generically with respect to some shape metric D with gradient ∇D. While we focus our experimentation on shapes represented as currents (Vaillant & Glaunès, 2005; Glaunès, 2005), it is trivial to implement for landmarks and straightforward to implement for other metrics such as varifolds (Charon & Trouvé, 2013). Our
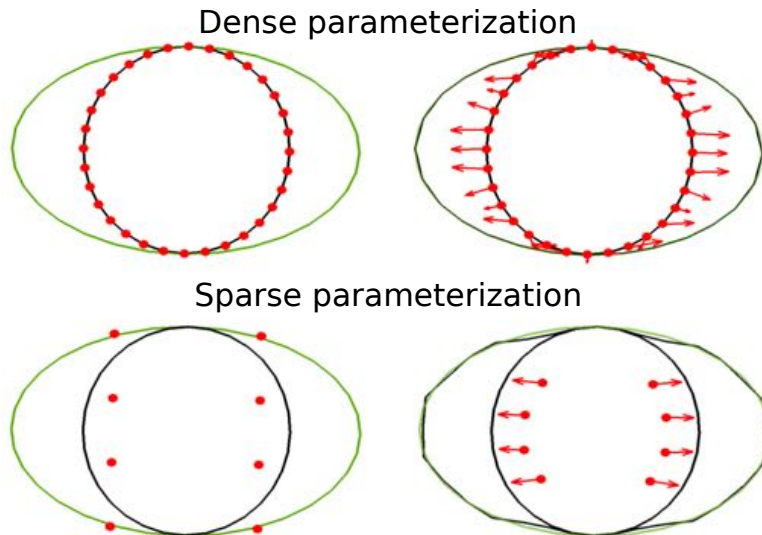

Figure 2: A black circle is deformed to match a green ellipse with initialization (left) and estimated model (right). Top: Deformation parameters are located at the vertices of the circle. Bottom: Deformation parameters are independent of the shape representation; a set of control points are initialized in the ambient space (bottom left). The deformation directions are estimated as well as their locations (bottom right). See Section 3.2 for details regarding the estimation of deformation parameters and their locations. The left/right expansion of the circle is more succinctly captured by the sparse parameterization.

software implementation `deformetrica` (`www.deformetrica.org`) supports landmarks, currents, and varifolds.

3. Geodesic Shape Regression

3.1. Geodesic flow of diffeomorphisms

The geodesic path connecting ϕ_0 to ϕ_1 is the path with constant velocity, which is equivalent to the path which minimizes the total kinetic energy of the the velocity field v_t

$$\frac{1}{2}\int_0^1 \|v_t\|_V^2 \, dt = \int_0^1 \sum_{p=1}^{N_c}\sum_{q=1}^{N_c} \alpha_p(t)^t K(c_p(t), c_q(t))\alpha_q(t) \, dt, \qquad (11)$$

which is defined entirely by the state of the system $\mathbf{S}(t)$. The $\boldsymbol{\alpha}(t)$ which minimize (11) satisfy a set of differential equations defining the evolution of momenta over time (Miller et al., 2006). Combining this with the motion of the control points (10) gives

$$\begin{cases} \dot{c}_i(t) = \sum_{p=1}^{N_c} K(c_i(t), c_p(t))\alpha_p(t), \\ \dot{\alpha}_i(t) = -\sum_{p=1}^{N_c} \alpha_i(t)^t \alpha_p(t)\nabla_1 K(c_i(t), c_p(t)) \end{cases} \qquad (12)$$

where $\nabla_1 K$ is the gradient of K with respect to the first parameter, with initial conditions $\mathbf{S}_0 = \{\mathbf{c}_0, \boldsymbol{\alpha}_0\}$, which we write in short as

$$\dot{\mathbf{S}}(t) = F(\mathbf{S}(t)), \qquad \text{given } \mathbf{S}(0) = \mathbf{S}_0, \qquad (13)$$

which we will refer to as the *shooting* equations, or geodesic shooting.

The shooting equations in (12) provide the continuous evolution of the control points and momenta, and represent the speed and acceleration of control points, respectively. This shows that the entire flow of diffeomorphisms is parameterized completely by the initial state of the system $\mathbf{S}_0$. First, the continuous motion of the control points and momenta can be determined by solving equations (12). One can then determine the velocity at any location and any time using equation (8) and therefore compute the full trajectory of a point x through time by integrating $\dot{\phi}_t(x) = v_t(\phi_t(x))$.

3.2. Method

The goal is to estimate a continuous shape evolution from a discrete set of observed shapes $\mathbf{O}_{t_i}$ at time t_i within the time interval $[t_0, T]$. Here we consider shapes to be generic geometric objects that can be represented as curves, landmark points, or surfaces in $2D$ or $3D$. Shape evolution is modeled as the geodesic flow of diffeomorphisms acting on a baseline shape $\mathbf{X}_0$, defined as $\mathbf{X}(t) = \phi_t(\mathbf{X}_0)$ with t varying continuously within the time interval determined by the observed data. The baseline shape $\mathbf{X}_0$ is continuously deformed over time to match the observation data ($\mathbf{X}(t_i) \sim \mathbf{O}_{t_i}$) with the rigidity of the evolution controlled by a regularity term. This is naturally expressed as a variational problem, described by the regression criterion

$$
\begin{aligned}
E(\mathbf{X}_0, \phi_t) &= \sum_{i=1}^{N_{obs}} ||(\phi_{t_i}(\mathbf{X}_0) - \mathbf{O}_{t_i})||^2 + \text{Reg}(\phi_t) \\
&= \sum_{i=1}^{N_{obs}} D(\mathbf{X}(t_i), \mathbf{O}_{t_i}) + L(\phi_t),
\end{aligned}
\tag{14}
$$

where D represents a squared distance metric and L is a measure of the regularity of the time-varying deformation ϕ_t. Recall from Section 2.2 that our method is generic with respect to shape metric D and that we could interpret the norm here as sum of squared difference, currents, or varifolds, among others.

The geodesic flow of diffeomorphisms ϕ_t is parameterized by N_c control points and momenta vectors $\mathbf{S}_0 = \{\mathbf{c}_0, \boldsymbol{\alpha}_0\}$, which act as initial conditions for the shooting equations (12). The baseline shape $\mathbf{X}_0$ can then be deformed by the flow equation (8). Therefore we seek to estimate the position of the control points, initial momenta, and position of the points on the baseline shape such that the resulting geodesic flow of the baseline shape best matches the observed data. An overview of our control point formulation of geodesic shape regression is shown in Figure 3. With all elements of our framework defined, geodesic shape regression can now be described by the specific regression criterion

$$
E(\mathbf{X}_0, \mathbf{S}_0) = \sum_{i=1}^{N_{obs}} \frac{1}{2\gamma^2} D(\mathbf{X}(t_i), \mathbf{O}_{t_i}) + L(\mathbf{S}_0),
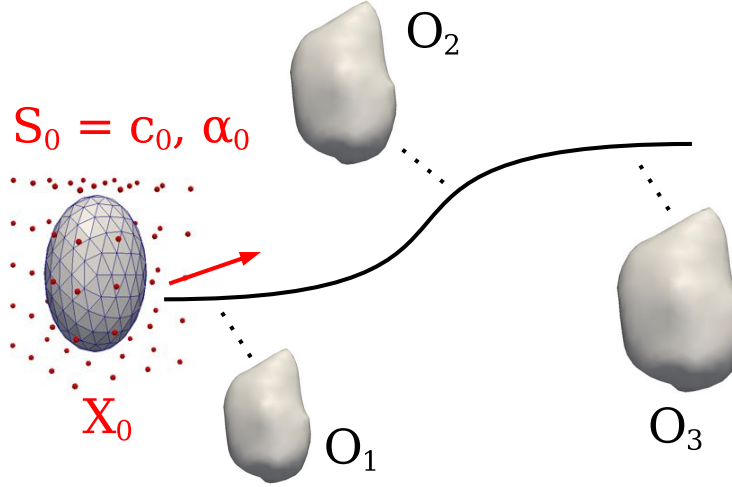\tag{15}
$$

12

Figure 3: Overview of geodesic shape regression. An initial baseline configuration $\mathbf{X}_0$ is deformed over time to match shape observations. The flow of diffeomorphisms is constrained to be a geodesic, parameterized by initial momenta $\boldsymbol{\alpha}_0$ located at $\mathbf{c}_0$. Parameters of the model which must be estimated are shown in red.

subject to

$$\begin{cases} \dot{\mathbf{S}}(t) = F(\mathbf{S}(t)) & \text{with } \mathbf{S}(0) = \{\mathbf{c}_0, \boldsymbol{\alpha}_0\}, \\ \dot{\mathbf{X}}(t) = G(\mathbf{X}(t), \mathbf{S}(t)) & \text{with } \mathbf{X}(0) = \mathbf{X}_0, \end{cases} \tag{16}$$

where γ^2 is used to balance the importance of the data term and regularity. The regularity term is given by the kinetic energy of the control points

$$L(\mathbf{S}_0) = \sum_{p=1}^{N_c} \sum_{q=1}^{N_c} \alpha_p(0)^t K(c_p(0), c_q(0)) \alpha_q(0). \tag{17}$$

The first part of (16) describes the trajectory of the control points and momenta as in the shooting equations (12). The second equation of (16) represents flowing the baseline shape along the deformation defined by $\mathbf{S}(t)$ as in (8).

As shown in the appendix, the gradients of the criterion (15) are

$$\nabla_{\mathbf{S}_0} E = \xi(0) + \nabla_{\mathbf{S}_0} L \qquad \nabla_{\mathbf{X}_0} E = \theta(0), \tag{18}$$

where the auxiliary variables $\theta(t)$ and $\xi(t) = \{\xi^c, \ \xi^\alpha\}$ satisfy the ODEs:

$$\dot{\theta}(t) = -\partial_1 G(\mathbf{X}(t), \mathbf{S}(t))^t \theta(t) + \sum_{i=1}^{N_{Obs}} \nabla_{\mathbf{X}(t_i)} D(\mathbf{X}(t_i), \mathbf{O}_{t_i}) \delta(t - t_i) \quad \theta(T) = 0,$$

$$\dot{\xi}(t) = -(\partial_2 G(\mathbf{X}(t), \mathbf{S}(t))^t \theta(t) + d_{\mathbf{S}(t)} F(\mathbf{S}(t))^t \xi(t)) \qquad\qquad \xi(T) = 0.$$
$$\tag{19}$$

where ∂_i denotes the partial derivative with respect to the ith parameter and $\delta(t - t_i) = 1$ when $t = t_i$ and is zero otherwise.

3.3. Algorithm

We implement an adaptive step size gradient descent algorithm, summarized in Algorithm 1. The gradient is computed by first integrating equations (12) forward in time to construct the flow of diffeomorphisms. The deformations are then applied to the baseline shape by integrating forward in time equation (8). With the full trajectory of the deformed baseline shape, one can compute the gradient of the data term $\nabla_{\mathbf{X}(t_i)} D(\mathbf{X}(t_i), \mathbf{O}_{t_i})$, corresponding to each observation.

The ODEs (19) are then integrated backwards in time, with the gradients of the data term acting as jump conditions at observation time-points, which pull the geodesic towards target data. The final values of the auxiliary variables $\theta(0)$ and $\xi(0)$ are then used to update the location of the control points, the initial momenta, and the location of the points on the baseline shape. All ODEs are solved using an Euler scheme with prediction correction, which is equivalent to second order Runge-Kutta.

This method has been implemented in the freely available software package `deformetrica` which can be downloaded at `www.deformetrica.org`.

3.3.1. Initialization

There are several parameters which require initialization. The simplest of which is the initial momenta $\boldsymbol{\alpha}_0$, which are initialized to 0, corresponding to no deformation. It is possible to initialize momenta by some preprocessing, such as registration between the earliest and latest time-points. However, this is not necessary, as the algorithm computes a reasonable estimate for momenta after the first iteration.

The algorithm also requires an initial baseline shape. For surfaces, one choice for initialization is an ellipsoid for each connected component of the observed shapes. The ellipsoid serves as a reasonable blobby approximation of many anatomical shapes. The ellipsoid(s) defines the number of shape

Algorithm 1: Geodesic shape regression

Input: $\mathbf{X}_0$ (initial baseline shape), $\mathbf{O}_{t_i}$ (observed shapes), t_0 (start time), T (end time), γ (tradeoff), σ_V (std. dev. of deformation kernel), σ_W (std. dev. of currents metric)

Output: $\mathbf{X}_0, \mathbf{c}_0, \boldsymbol{\alpha}_0$

1 $\boldsymbol{\alpha}_0 = \mathbf{0}$
2 Initialize control points $\mathbf{c}_0$ on regular grid with spacing σ_V
3 **repeat**

    // Compute path of control points and momentum (forward integration)

4   $c_i(t) = c_i(0) + \int_{t_0}^{T} \sum_{p=1}^{N_c} K(c_i(s), c_p(s))\alpha_p(s)ds$

5   $\alpha_i(t) = \alpha_i(0) - \int_{t_0}^{T} \sum_{p=1}^{N_c} \alpha_i(s)^t \alpha_p(s) \nabla_1 K(c_i(s), c_p(s))ds$

    // Trajectory of deformed baseline shape (forward integration)

6   $x_k(t) = x_k(0) + \int_{t_0}^{T} \sum_{p=1}^{N_c} K(x_k(s), c_j(s))\alpha_j(s)ds$

    // Compute the gradient of the data term for each observation

7   $\nabla_{\mathbf{X}(t_i)} D(t_i)$

    // Compute auxiliary variable $\theta(t)$ (backward integration)

8   $\theta_k(t) = \theta_k(T) + \int_{T}^{t} \sum_{p=1}^{N_c} \alpha_p(s)^t \theta_k(s) \nabla_1 K(x_k(s), c_p(s)) - \sum_{i=1}^{N_{obs}} \nabla_{x_k(t_i)} D\delta(s - t_i)ds$

    // Compute auxiliary variable $\xi^c(t)$ (backward integration)

9   $\xi_k^c(t) = \xi_k^c(T) - \int_{T}^{t} \sum_{p=1}^{N_x} \alpha_k(s)^t \theta_p(s) \nabla_1 K(c_k(s), x_p(s)) + (\partial_c F^c)\xi_k^c(s) + (\partial_c F^\alpha)\xi_k^\alpha(s)ds$

    // Compute auxiliary variable $\xi^\alpha(t)$ (backward integration)

10   $\xi_k^\alpha(t) =$
  $\xi_k^\alpha(T) - \int_{T}^{t} \sum_{p=1}^{N_x} K(c_k(s), x_p(s))\theta_p(s) + (\partial_\alpha F^c)\xi_k^c(s) + (\partial_\alpha F^\alpha)\xi_k^c(s)ds$

    // Compute gradients

11   $\nabla_{\mathbf{c}_0} E = \xi^{\mathbf{c}}(0) + \nabla_{\mathbf{c}_0} L$

12   $\nabla_{\boldsymbol{\alpha}_0} E = \xi^{\boldsymbol{\alpha}}(0) + \nabla_{\boldsymbol{\alpha}_0} L$

13   $\nabla_{\mathbf{X}_0} E = \theta(0)$

    // Update control points, momenta, and baseline shape

14   $c_i(0) = c_i(0) - \varepsilon \nabla_{c_i} E$

15   $\alpha_i(0) = \alpha_i(0) - \varepsilon \nabla_{\alpha_i} E$

16   $x_i(0) = x_i(0) - \varepsilon \nabla_{x_i} E$ 15

17 **until** *Convergence*
18 **return** $\mathbf{X}_0, \mathbf{c}_0, \boldsymbol{\alpha}_0$

points as well as the connectivity, which is preserved during optimization. Figure 4 shows an example initial baseline shape with connectivity. It is also reasonable to choose one of the observed shapes for the initial baseline shape, the earliest observed shape for example. This will improve the speed of convergence, but may slightly bias the estimation towards that observation.

The location of the control points $\mathbf{c}_0$ must also be initialized. For our implementation, we initialize control points on a regular grid defined by a bounding box around the observed data. Figure 4 shows initial control points surrounding the initial baseline shape. The spacing is determined by σ_V and consequently determines the number of control points and therefore the dimensionality of the model parameters. One can also provide an explicit list of initial control point locations, perhaps motivated by some prior knowledge. It is also possible to initialize control point locations at the vertices of the baseline shape. However, this goes against the spirit of the control point formulation, which decouples the deformation parameters from the shape representation. For that reason, we prefer initialization on a regular grid.
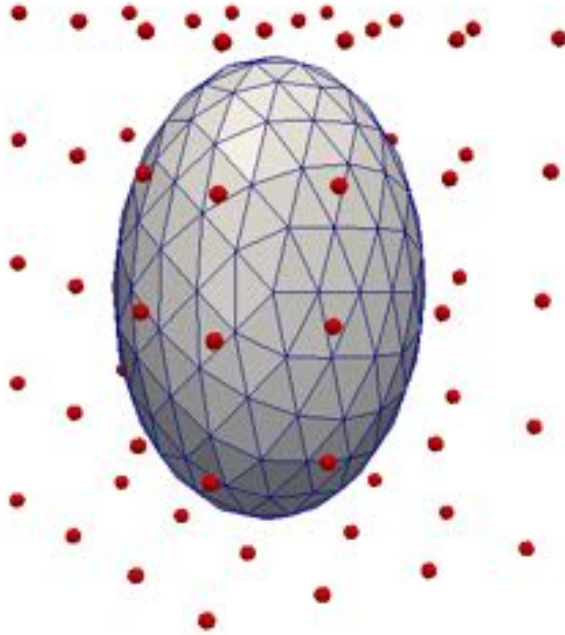

Figure 4: Initialization of control points on a regular grid in red. The baseline shape here is initialized as an ellipse, which defines the number of shape points and connectivity.

4. Impact of Parameter Selection via Cross-Validation

There are three main parameters which influence model estimation:

- σ_V: the size of the kernel which defines the deformation. It is the distance at which points move in a correlated way. Higher values result in mostly rigid deformation, while lower values allow each point to move independently.

- σ_W: the size of the kernel which defines the metric on currents. This parameter allows you to tune the metric properties of the space of currents to suit your application. Intuitively, this parameter is the scale at which shape differences are considered noise.

- γ: the trade-off between data-matching and regularity.

In this section, we explore the impact of these three parameters on the estimated geodesic model by cross validation.

Data: We have a cross-sectional dataset of 53 healthy children clustered at 6, 12, and 24 months of age. The 53 subjects are distributed as follows: 18 subjects around six months old, 18 subjects around 12 months old, and 17 subjects around 24 months old. For each subject, we have extracted the left caudate and left putamen. The population variability is summarized in Figure 5, which shows caudate and putamen volume for all subjects. For our experiments, we consider the caudate and putamen together as a multi-object complex rather than two independent shapes. The caudate and putamen are in close proximity to each other in shared anatomical space, which motivates a *joint* analysis, as the interplay between the two structures is potentially important.

Experimental setup: We perform 5-fold cross validation by partitioning the 53 subjects into 5 unique groups of size 11, 11, 11, 10, and 10 subjects. A geodesic model is built by leaving out one of the groups, and using all remaining data in model estimation. The data used in model estimation will be referred to as training data, with the data left out referred to as testing data. This is repeated until every group has eventually been left out of model estimation.

For each set of training data, we estimate a number of models by varying the parameter values. For each fold of our cross validation experiment, we
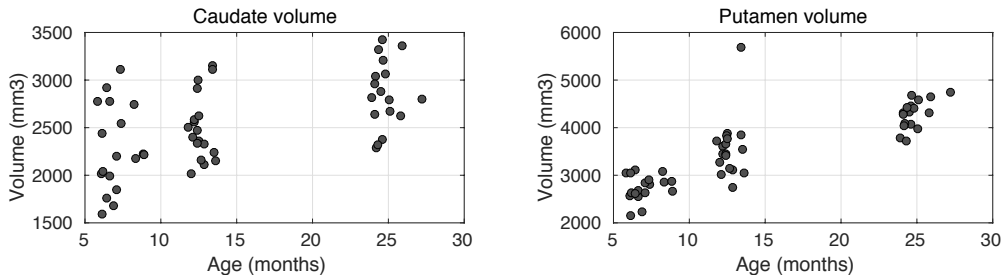
17

Figure 5: Caudate and putamen volume for a cross-sectional dataset of 53 healthy children.

estimate 120 models by exploring the range of values $\sigma_V = [50, 40, 20, 10, 5, 3]$ mm, $\sigma_W = [50, 20, 10, 3]$ mm, and $\gamma = [100, 10, 1, 0.1, 0.01]$.

For each fold, we have a geodesic model for every combination of parameters, along with a unique set of testing data left out during model estimation. We therefore explore how well the estimated models match the testing data. As the shape complexes are not in anatomical correspondence across subjects, we use `MeshValmet` (Gerig et al., 2001) to compute surface to surface distance by dense sampling. For each left out shape complex, we get a distribution of sample error, which we accumulate for all testing data. From the full sample error distribution, we can compute summary statistics such as mean, standard deviation, min and max.

4.1. Impact of deformation kernel width σ_V

The deformation kernel width controls the dimensionality of the parameterization of the resulting diffeomorphic flow, and therefore the degrees of freedom of the system. Intuitively, lowering σ_V will often result in better model fit which comes with a corresponding increase in the number of model parameters. Figure 6 summarizes the impact of σ_V with a fixed $\sigma_W = 3$ and $\gamma = 1.0$, by showing the distribution of error with different parameter combinations. We see a significant increase in model fit by decreasing σ_V from 50 mm to 20 mm, however a further decrease to 5 mm does not have a large impact. In fact, lowering σ_V from 20 mm to 5 mm decreases the average error from 0.4826 to 0.4141, while the number of model parameters increases from 8 to 240; a minimal increase in model accuracy which increases the complexity of the model by 2900%. Lowering σ_V further to 3 mm has no meaningful impact on model fit, but increases the number of model parameters to 960.

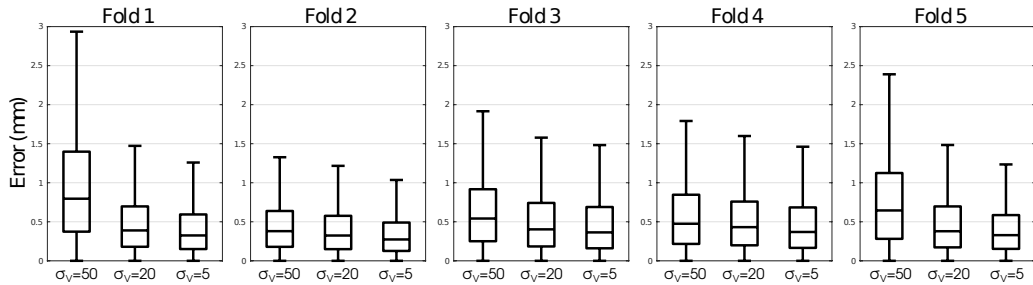This can also be seen in Figure 7, which shows the typical error between

18

Figure 6: The impact of σ_V is explored by measuring model error with respect to testing data with fixed $\sigma_W = 3$ and $\gamma = 1$. Each fold contains a unique partition of training and testing data.

the model and a caudate/putamen pair from the testing data. Lowering σ_V from 50 mm to 20 mm significantly improves model fit, reducing matching error particularly on the edges of the shapes. Lowering σ_V further to 5 mm results in small improvements, but greatly increases model complexity in terms of the number of model parameters. The additional model parameters could be considered noise or redundancy in the model description.

Indeed, smaller values of σ_V lead to higher accuracy in matching observed data. However, at a certain point one will see diminishing returns. While model accuracy may increase slightly, the gain is made at the cost of over-parameterization and redundancy in the parameterization. For certain
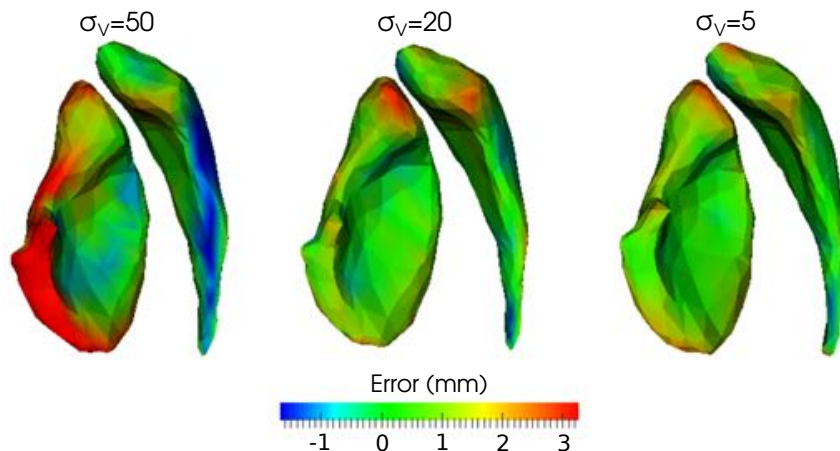

Figure 7: A typical example of model error shown on the surface with varying σ_V, and fixed $\sigma_W = 3$ and $\gamma = 1$. Note the minor improvement from $\sigma_V = 20$ to $\sigma_V = 5$.

19

applications, very high accuracy may be desirable regardless of the increase in model dimensionality. For example, measuring subtle changes of anatomy in response to drug treatment.

4.2. Impact of shape matching kernel width σ_W

The shape matching kernel width defines the space of currents and intuitively controls the scale of shape features which should be matched. Two structures whose local shape features differ by less than σ_W are considered equivalent. The geodesic model presented here is independent of the shape metric used. Indeed, any metric can be used within this framework given a way to compute the metric and its gradient. However, we include an analysis of σ_W for completeness.

Figure 8 summarizes the impact of σ_W with a fixed $\sigma_V = 20$ and $\gamma = 1$ by showing the distribution of surface errors over a range of values of σ_W. Generally, surface error decreases as σ_W decreases. However, in this experiment, the decrease in error is quite small. That is likely because the observations are rather smooth anatomical shapes without small, sharp features of interest. In such cases where shapes contain prominent local features, different values of σ_W may result in considerably different estimated models.

Unlike σ_V, σ_W does not have an impact on the dimensionality of the estimated model. Rather, σ_W allows the user to incorporate domain knowledge into model building, by denoting the scale of shape differences which should be ignored. In that sense, it is hard to conclude that smaller values of σ_W are better in general. In the case where data contains small, spurious features of interest, one must choose σ_W carefully to ensure that the features are captured by the model.
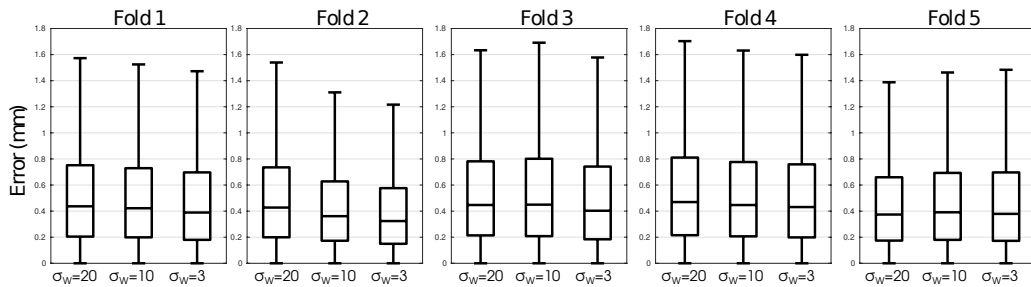

Figure 8: The impact of σ_W is explored by measuring model error with respect to testing data with fixed $\sigma_V = 5$ and $\gamma = 1$. Each fold contains a unique partition of training and testing data.

4.3. Impact of regularity weight γ

In our cross-validation experiments, the value of γ did not have a large impact on the estimated model. The difference between the error for a given σ_V and σ_W over the range of values for γ was less than 1%. One possible explanation is that the geodesic constraint is already a strong regularizer on model estimation. It is also likely the estimation problem itself requires little regularization as the observations are smooth anatomical shapes.

5. Impact of Missing Data

The previous section explored the ability of the geodesic model to match a cross-sectional population with considerable inter-subject variability. Here, we focus on the geodesic model applied to estimating subject-specific growth trajectories from longitudinal data. In such cases, only a few observations sparsely distributed in time are available. Further, the time between observations is on the order of months or even years. There can potentially be dramatic changes and differences in observations over the span of years, particularly in the case of developing children or the study of disease. Therefore, it is important we understand how the geodesic model 'fills in the gaps' between observations, to better inform when the geodesic model is a suitable choice.

To assess the impact of missing MRI data on the geodesic model, we leverage a unique longitudinal dataset of a child scanned 16 times from around 4 to 8 years old. For each time-point, we segment three subcortical pairs (left/right): hippocampus, caudate, and putamen. The left of Figure 9 shows the subcortical shapes at the earliest time-point of 4.2 years. Segmentation is done independently for each time-point, by nonlinear alignment to a template (Gouttard et al., 2007). The right of Figure 9 shows the volume of the structures over time, showing the distribution of observations as well as highlighting the variability in the extracted shapes, with noise introduced during image acquisition and segmentation. The noisy observations represent a significant modeling challenge, where the goal is not to match observations as closely as possible, but rather the model should capture the overall trend.

With such a dense sampling across time, we use the 16 observations as a proxy for ground truth. We can therefore leave a subset of observations out during model estimation, and evaluate how well the model matches the excluded data. We estimate models using 2 observations, 3, and so forth, up to all 16 observations, always choosing the kept observations as evenly spaced
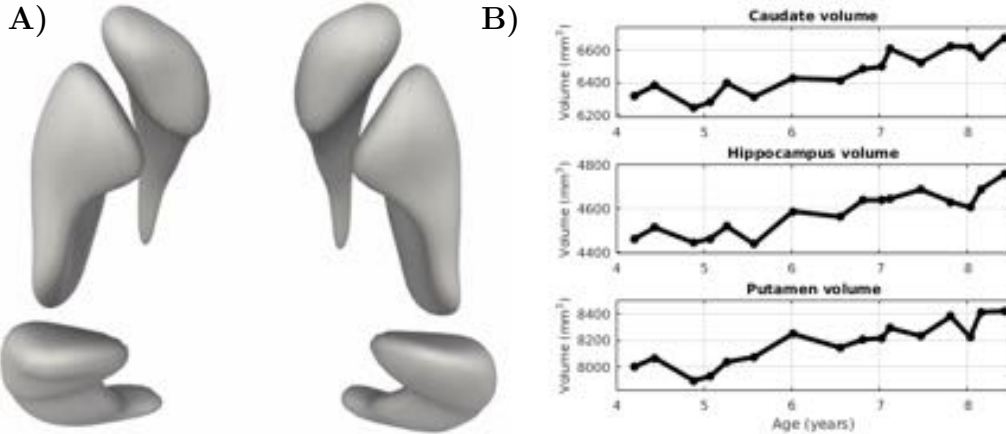
Figure 9: **A)** Hippocampus, caudate, and putamen observations segmented from MRI for a child at 4.2 years old. **B)** Volume of subcortical shapes measured from 16 observations of the same child.

in time as possible. Models are estimated jointly on the shape-complex consisting of 6 subcortical shapes with parameters $\sigma_V = 10$ mm, $\sigma_W = 3$ mm, and $\gamma = 1.0$. Each model is therefore parameterized by 294 momenta vectors. For comparison, we also estimate a corresponding set of piecewise-geodesic models, as the piecewise-geodesic model is also based on flows of diffeomorphisms, and therefore has the same parameters $\sigma_V = 10$ mm and $\sigma_W = 3$ mm. It also has a regularity parameter which we set to 0.0001. The regularity parameter cannot be directly compared between methods, but we choose each value so the data-matching portion is 5 orders of magnitude larger than the regularity term in order to prioritize accurate matching in each model.

Figure 10 summarizes the results of the leave-several-out experiments on the hippocampus, caudate, and putamen. We see the coefficient of determination R^2 as a function of the number of observations used in model estimation. R^2 is a measure of how well a model fits observed data, with a value of 1 indicating a perfect fit between model and data. The coefficient of determination can be computed as

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \tag{20}$$

where y_i are the observations, f_i are the modeled values, and $\bar{y}$ is the mean of the observations. Intuitively, R^2 is the sum of squared residuals normalized by the variance of the data.

22

The R^2 values in Figure 10 are computed with respect to the currents metric. The currents metric is the reason the values are so close to 1, as shape features smaller than 3mm considered equivalent. The key takeaway here is the trend in R^2 as more and more observations are used in model estimation. For piecewise-geodesic models, R^2 steadily increases as more observations are utilized. The geodesic model gets increasing more accurate up to about 6 observations, at which point the addition of further observations does not greatly influence model estimation. In this case, the piecewise-geodesic model is overfitting the noisy observations while the geodesic model captures the overall trend.

The piecewise-geodesic shape sequence undergoes instantaneous change of direction, effectively matching the variability in the observations. The estimated geodesic model results in a smooth shape sequence which captures the overall trend without being greatly influenced by outliers. This can also be seen clearly by tracing out the path of shape points over time, as shown in Figure 11. The trajectories estimated by piecewise-geodesic regression are highly irregular, and do not represent a believable or realistic model of smooth anatomical change. In contrast, trajectories from geodesic regression represent smooth and reasonable anatomical growth, while achieving very high data-matching ($R^2 \approx 0.99$) with respect to the currents metric.

5.1. Modeling of Extracted Shape Features

Many analysis pipelines extract a measurement, such as volume, from imaging data and perform subsequent analysis on the measurement, now isolated from the original anatomical context. One main benefit of modeling shape evolution is the variety of measurements which can be simply extracted
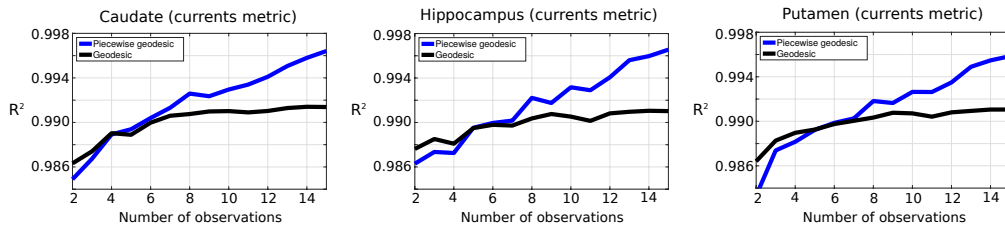

Figure 10: Coefficient of determination R^2 as a function of the number of observations used in model estimation. R^2 here is measured with respect to the currents metric, which explains why the values are nearly 1, as the currents metric is insensitive to shape features smaller than σ_W.
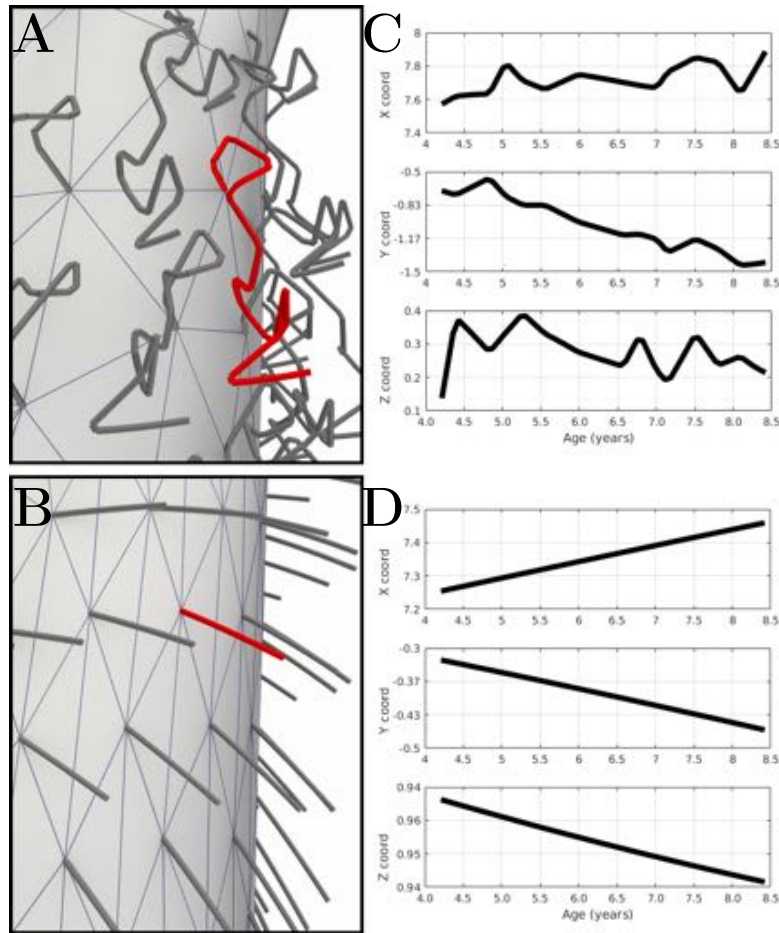
Figure 11: Point trajectories from 4.2 to 8.4 years estimated by piecewise-geodesic regression (**A**) and geodesic regression (**B**) for a zoomed in section of the left caudate. The trajectory in red is decomposed into coordinates in **C** and **D**. The models were estimated using all 16 observations, which highlights overfitting in piecewise-geodesic regression. Geodesic regression produces trajectories which more reasonably model smooth anatomical change.

from the resulting shape evolution. Here we explore the application of the geodesic model to extracted measurements, namely volume, as it is a common clinical measurement and biomarker in many applications.

Here, we perform the same leave-several-out experiments on the 16 time-point data set as in the previous section, by estimating piecewise-geodesic and geodesic models. From the resulting models, we extract continuous volume

curves and compare to the volume of the observations. Figure 12 shows R^2 as a function of the number of observations used in model estimation. We observe the general behavior of the two models as in the previous section, with the piecewise geodesic model showing increasing R^2 as the number of observations increases, while the geodesic model shows relatively little change in R^2 over the full range of experiments. In this experiment, we can also more clearly see the shape matching accuracy which the currents metric used in the previous section did not illustrate. When all observations are used, volume extracted from the piecewise geodesic model very closely matches the observations, with an R^2 over 0.9 for all structures. Conversely, the geodesic model results in an R^2 between 0.7 and 0.8 for the three structures.

However, as before, the data-matching accuracy increases from piecewise-geodesic regression comes with overfitting, as additional observations greatly alter model estimation. This suggests that the piecewise-geodesic model more closely resembles interpolation than regression, preferring to match observations closely rather than capture the overall trend. The tendency to overfit at the cost of model generalizability is shown in Figure 13, which shows observed putamen volume as well as volume curves extracted from piecewise-geodesic and geodesic regression models. Volume extracted from piecewise-geodesic regression does not follow a realistic time course, while the geodesic model produces a smooth and anatomically reasonable volume curve. Furthermore, the putamen volume curve extracted after geodesic regression is similar to an exponential regression on the volume measurements themselves. However, the shape model was built using the left/right caudate, putamen, and hippocampus. Modeling the shapes jointly allows for possible interactions between structures which was not considered in a regression analysis of putamen volume alone. Furthermore, from a single model of shape
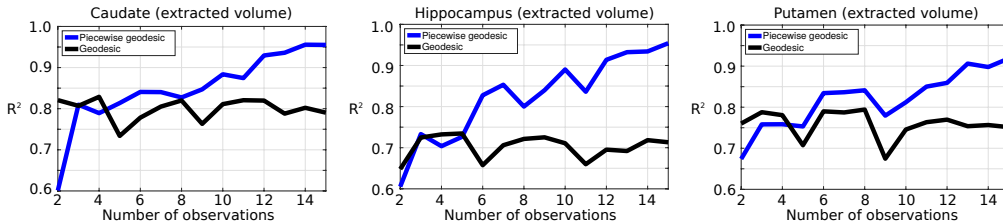

Figure 12: Coefficient of determination R^2 as a function of the number of observations used in model estimation. R^2 here is measured with respect to volume extracted from the estimated shape regression models.
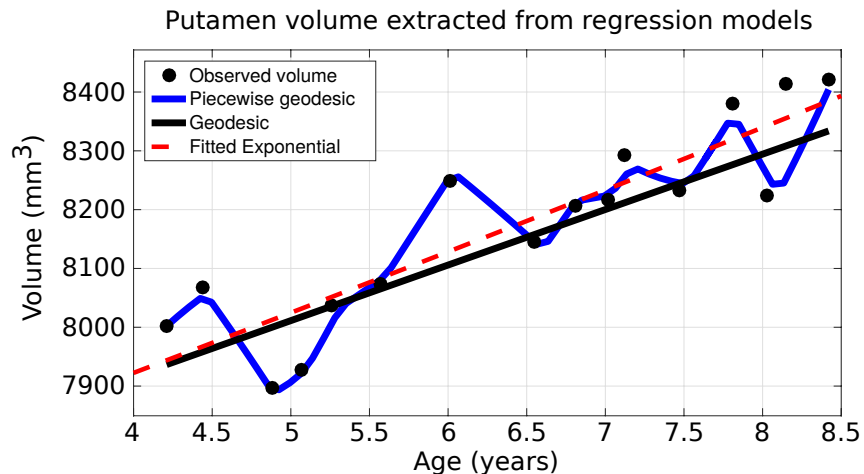
Figure 13: Putamen volume extracted continuously after piecewise-geodesic and geodesic shape regression using all 16 observations. Also shown is a curve from exponential regression on the raw observed volumes. Volume from piecewise-geodesic regression closely follows the observations, but does not generalize the trend in the observations. Geodesic regression produces a smooth volume trajectory which is similar to that from exponential regression.

change we can extract other shape features, such as surface area, thickness, curvature, among numerous others. This experiment illustrates that spatiotemporal shape modeling fits naturally into traditional analysis pipelines for scalar measurements.

5.2. Extrapolation

The geodesic model is a generative parametric model which allows for extrapolation beyond the observation period. In contrast, the piecewise-geodesic model has no mechanism for extrapolation. Here we explore the extrapolation properties of the geodesic model on the 16 time-point data set much in the same form as the previous section. We build models using an increasing number of observations, always selecting the observations earliest in time. For each model, we then extrapolate the remaining time interval. For example, the model estimated with 2 observations uses the first 2 observations at 4.2 and 4.4 years, and then extrapolated until 8.4 years. From each extrapolated shape model spanning the whole time interval, we then extract volume as we did previously, and compare with the raw volume observations.

Figure 14 shows how the extracted putamen volume sequences match the observed raw volumes for geodesic models estimated using 3, 7, and 11 ob-
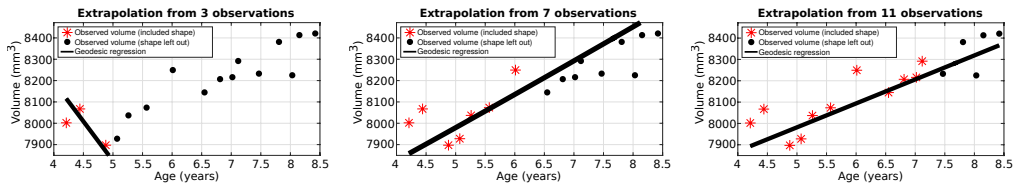
Figure 14: Putamen volume estimated from geodesic regression on a subset of observations and extrapolated to the end of the time interval. The shape model itself is extrapolated, and volume is extracted continuously from the resulting shape sequence. The appearance of linearity here is coincidental and data-driven; the volume trends are densely sampled nonparametric sequences.
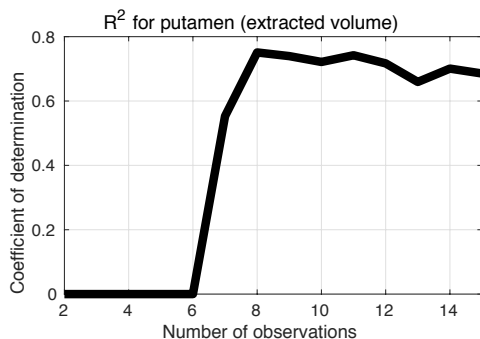

Figure 15: R^2 for volume extracted from putamen shape models, built from subset of observations earliest in time and then extrapolated to the full time interval. Note that the model is estimated on shape observations, and extrapolation is also performed to extend the shape sequence to the end of the observation period. Volume is then measured continuously from the shape models. Negative values are clamped to 0, denoting the extracted volume is a worse fit than the simple average of the observed volumes.

servations. For this dataset, the first 3 observation poorly capture the overall trend, and therefore the extrapolated sequence is a very poor model. However, the models estimated using 7 and 11 observations capture the overall trend, and the extrapolated sequence is a reasonable match for the remaining observations. Note that the volume trends, though they appear linear in Figure 14, are in fact nonparametric sequences measured from extrapolated shape models. There are no constraints about linearity of any derived shape feature built into model estimation.

Figure 15 summarizes how well the extracted volumes match the observations, by showing R^2 as a function of the number of observations. For up to 6 observations used in model estimation, the value of R^2 is actually

negative, though it is shown clamped to 0 for display purposes. This means that the volume extracted from the extrapolated shape models is a worse fit than a horizontal line representing the average volume. When the number of observations is greater than 6, the value of R^2 is approximately 0.7, showing that the extrapolation is a reasonable fit in this case. It appears that extrapolation captures the trend, given enough observations to overcome the noisy nature of the observed shapes.

This experiment demonstrates that extrapolating shape models estimated using a limited number of observations must be done with care. Given a limited data-set consisting of noisy observations, extrapolation can diverge considerably from the real trend. However, the problem lies more in the temporal inconsistency of the observations than it does with the geodesic regression model. No generic model is able to accurately extrapolate into the future when the first several observations deviate significantly from the overall trend. Indeed, the noisy nature of the observations suggests a need for 4D segmentation schemes (Wang et al., 2012).

6. Multimodal shape regression analysis

The control point formulation of diffeomorphic flow separates the deformation parameterization from any specific shape representation. As a consequence, we can embed several shapes with different representations (i.e. points, curves, meshes, etc.) into the same ambient space without impacting the dimensionality or parameterization of the geodesic model. By including multiple sources of geometric information in the analysis, we get a more complete picture than is possible from any single source. Here, we explore multimodal modeling by combining shape information extracted from diffusion tensor imaging (DTI) and structural MRI (sMRI).

Data We have longitudinal observations of the same subject at 6, 12, and 24 months which include both DTI and sMRI (T1W and T2W). The sMRI is rigidly co-registered and the left/right caudate and putamen are segmented (Shi et al., 2011). DTI observations are non-linearly aligned to a 12-month old DTI atlas using `dti-reg` (https://github.com/NIRALUser/DTI-Reg). In the atlas space, we extract fiber geometry by tractography using `3DSlicer` (Fedorov et al., 2012) with a seed region covering the entire corpus callosum. Tractography is done in the atlas space to overcome the relatively poor quality of infant DTI and to provide topologically consistent
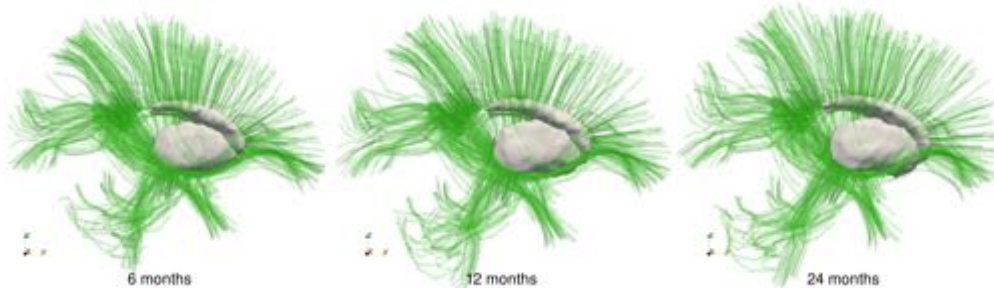
Figure 16: Fiber tracts and subcortical shapes at 6, 12, and 24 months old.

fiber structure across time, which cannot be guaranteed by independent fiber tracking performed on each observation. Fiber geometries are propagated to the individual DTI spaces by the inverse transformations which align observations to the atlas space (Goodlett et al., 2009). The goal here is not to capture differences at the scale of individual fiber bundles, such as longitudinal changes of fiber dispersion, bifurcations or crossings. Rather, we wish to model the general size and shape changes of fiber bundles over time, inspired by the tract-based analysis framework of Goodlett et al. (2009). Finally, mean diffusivity images are rigidly aligned with structural T2W images, as they have a similar appearance. This allows to align fiber geometry and subcortical structures for each time-point, as well as to provide alignment across time. The aligned fiber tracts and subcortical shapes are shown in Figure 16.

6.1. Regression of fibers and shapes

We estimate a comprehensive geodesic model using the 4 subcortical shapes as well as the fiber curves with parameters $\sigma_V = 8$ mm, $\sigma_W = 6$ mm for the fibers, $\sigma_W = 3$ mm for the subcortical shapes, and $\gamma = 0.1$. Figure 17 shows 3 views of the shapes estimated from geodesic regression at 24 months, along with grids which show the amount of deformation from 6 to 24 months. The model captures a large amount of torsion, as the top half and bottom half of the fibers move in opposite directions. Also captured is a pose change in the subcortical structures as well as non-linear growth of each structure.

The contribution of the fibers in addition to shapes in model estimation is clear. The fibers cover a large region surrounding the subcortical shapes, giving additional geometric information that expands far beyond the boundaries of the shapes. However, the reverse is not as obvious; is is not
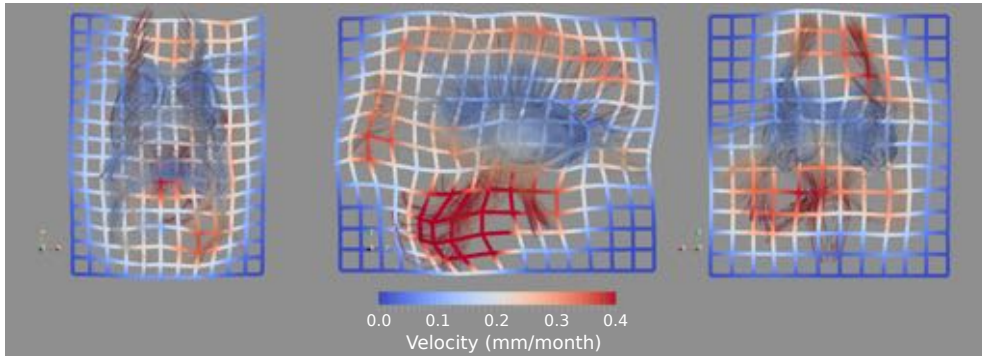
Figure 17: Fibers and subcortical shapes at 24 months estimated from geodesic regression. The same shape-complex is shown from three orthogonal angles (top, side, front) with a grid showing the total deformation from 6 to 24 months. Color denotes the magnitude of velocity.

immediately clear what is gained by including subcortical shapes in addition to fiber geometry. For comparison, we estimate a geodesic model using only fiber geometry with the same parameter settings as before.

Figure 18 shows an axial slice through the subcortical structures for the model estimated on fibers alone (top) and the model built with fibers and subcortical shapes (bottom). The grid shows the total amount of deformation over the time interval from 6 to 24 months. The inclusion of subcortical shapes in model estimation results in considerably more deformation to the ambient space which is shared by the fibers and subcortical structures. This example serves as an illustration that multimodal models estimated from a variety of geometric sources result in a more comprehensive model which captures changes not possible by any isolated shape taken out of anatomical context.

7. Conclusions

In this paper, we detailed a sparse representation of diffeomorphisms, where momenta are located at discrete control points. From the discrete momenta, dense deformations of the whole space can be computed. Then, through geodesic shooting, a geodesic flow of diffeomorphisms can be constructed and used to deform various shapes embedded in the ambient space. This machinery became the foundation around which we developed a dedicated algorithm for geodesic shape regression. Indeed, the control point
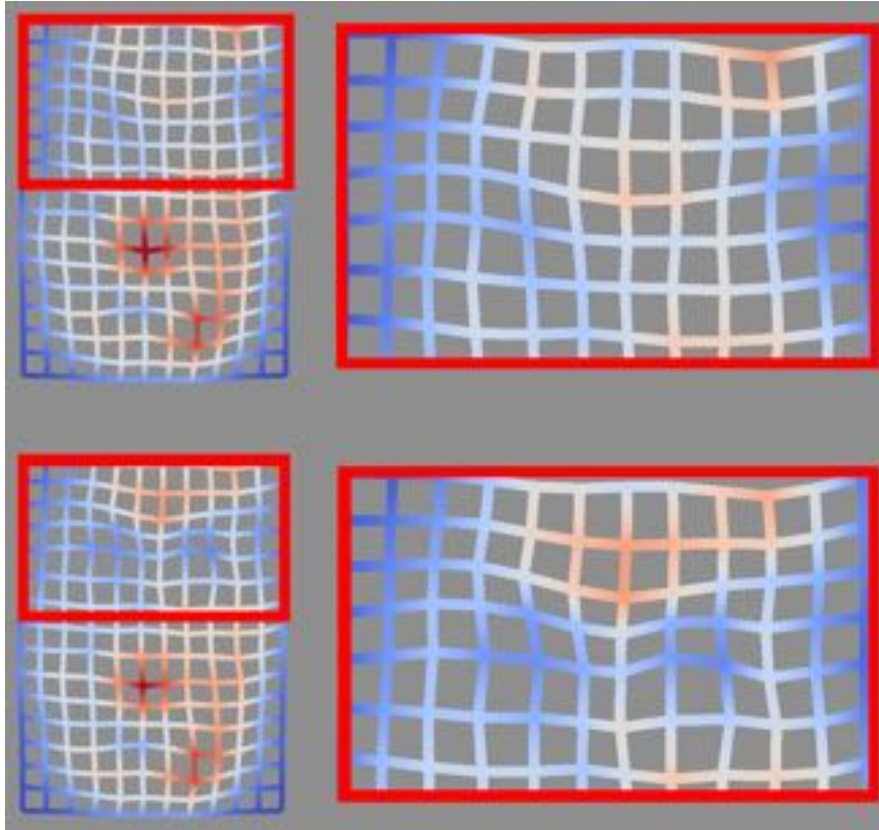
Figure 18: Grid looking down from the top, which shows the total deformation from 6 to 24 months. Top) Model estimated on fiber geometry alone. Bottom) Model estimated jointly on fiber geometry and subcortical shapes. Note the additional deformation present in the region of the subcortical shapes in the multimodal model, showing the impact of including additional sources of information in model estimation.

framework provided the flexibility of parameterization to include a variety of shapes in any combination. In addition to the estimation of a baseline shape configuration and initial momenta, we also estimate the locations of the momenta. This provides a compact statistical representation of dense deformations, particularly in the case of multi-object complexes which may contain thousands of shape points. Our method has been implemented in the software package `deformetrica` available to download at `www.deformetrica.org`

We saw on real medical data robust model estimation across a variety of parameter settings. Rather than producing vastly different results, the parameters of the model allow a user to fine tune the estimated models to

31

suit their needs and to inject their domain knowledge. The deformation kernel width σ_V balances the complexity of the model with the degrees of freedom, with lower values requiring more parameters to capture local non-linear changes and higher values tending towards rigid deformation with fewer parameters. The kernel width defining the currents metric σ_W controls the local size of shape features that are important for a given application. One can make σ_W lower for more detailed matching without impacting model complexity. The choice of regularity trade-off parameter γ is more nuanced, as there are no physical units to help aid selection, as there are for σ_V and σ_W. However, our experiments showed that γ can be chosen to favor accurate data-matching without causing strange deformations to appear, as the geodesic constraint is already a strong regularizer.

We demonstrated how geodesic shape regression can support traditional scalar regression analysis by exploring shape features extracted from shape trajectories. Shape regression models are estimated using multiple shapes which share a common ambient space, from which any shape feature of interest can be extracted. Volume extracted from geodesic models captured the overall trend in raw volumes using only a few observations, and closely resembled exponential regression on the volume measurements themselves. Additionally, one powerful aspect of the parametric geodesic model is the ability to extrapolate beyond the observation time window. We showed how extrapolation may well approximate future changes, though it has to be used carefully when only a few time-points are available. It remains future work to better understand the extrapolation properties of geodesic regression.

We showed an experiment on a single subject which uses fiber geometry from diffusion tensor imaging as well as multiple subcortical shapes derived from structural MRI. The results show that the two modalities provide complementary information. Models estimated using fibers and subcortical shapes jointly capture dynamic changes which are not possible in models on fibers or subcortical shapes alone.

Continuous models of shape change provide many important utilities. The continuous shape trajectory can be sampled at any time-point, allowing us to align shape observations with clinical measurements and scores not acquired at the same time, a procedure which is vital in studies where clinical observations occur more frequently than image acquisitions. Along these same lines, we can age match subjects not observed at the same time, or align a subject to a 4D normative trajectory in order to determine if a subject is developing abnormally. Future work will focus on large scale longitudinal

studies, where the goal is to detect subtle differences between populations.

Appendix A. Derivation of gradients

Gradient in matrix form

The regression criterion is

$$E(\mathbf{X}_0, \mathbf{S}_0) = \sum_{i=1}^{N_{obs}} \frac{1}{2\gamma^2} D(\mathbf{X}(t_i), \mathbf{O}_{t_i}) + L(\mathbf{S}_0), \tag{A.1}$$

subject to

$$\begin{cases} \dot{\mathbf{S}}(t) = F(\mathbf{S}(t)) & \text{with } \mathbf{S}(0) = \{\mathbf{c}_0, \boldsymbol{\alpha}_0\}, \\ \dot{\mathbf{X}}(t) = G(\mathbf{X}(t), \mathbf{S}(t)) & \text{with } \mathbf{X}(0) = \mathbf{X}_0, \end{cases} \tag{A.2}$$

where first part of (A.2) describes the trajectory of the control points and momenta as in the shooting equations (12). The second equation of (A.2) represents flowing the baseline shape along the deformation defined by $\mathbf{S}(t)$ as in (8). L represents regularity given by equation (17). We also introduce $D(t_i)$, $F(t)$, $G(t)$ as notation for $D(\mathbf{X}(t_i), \mathbf{O}_{t_i})$, $F(\mathbf{S}(t))$, and $G(\mathbf{X}(t), \mathbf{S}(t))$.

Consider a perturbation $\delta\mathbf{S}_0$ to the initial state of the system $(\mathbf{c}_0, \boldsymbol{\alpha}_0)$, which leads to a perturbation of the motion of the control points $\delta\mathbf{S}(t)$, a perturbation of the template shape trajectory $\delta\mathbf{X}(t)$, and a perturbation of the criterion δE

$$\delta E = \sum_{i=1}^{N_{obs}} \left((\nabla_{\mathbf{X}(t_i)} D(t_i))^t \delta\mathbf{X}(t_i) \right) + (\nabla_{\mathbf{S}_0} L)^t \delta\mathbf{S}_0. \tag{A.3}$$

The perturbations $\delta\mathbf{S}(t)$ and $\delta\mathbf{X}(t)$ satisfy the ODEs:

$$\begin{aligned} \delta\dot{\mathbf{S}}(t) &= d_{\mathbf{S}(t)} F(t) \delta\mathbf{S}(t) & \delta\mathbf{S}(0) &= \delta\mathbf{S}_0 \\ \delta\dot{\mathbf{X}}(t) &= \partial_1 G(t) \delta\mathbf{X}(t) + \partial_2 G(t) \delta\mathbf{S}(t) & \delta\mathbf{X}(0) &= \delta\mathbf{X}_0. \end{aligned} \tag{A.4}$$

33

where ∂_i denotes the partial derivative with respect to the ith parameter.

The first ODE is a linear homogeneous ODE with well known solution

$$\delta \mathbf{S}(t) = \exp\left(\int_0^t d_{\mathbf{S}(u)} F(u) du\right) \delta \mathbf{S}_0. \tag{A.5}$$

The second ODE is a linear inhomogeneous ODE that can be solved by the method of variation of parameters

$$\delta \mathbf{X}(t) = \int_0^t \exp\left(\int_u^t \partial_1 G(s) ds\right) \partial_2 G(u) \delta \mathbf{S}(u) du \\ + \exp\left(\int_0^t \partial_1 G(s) ds\right) \delta \mathbf{X}_0. \tag{A.6}$$

Plugging (A.5) into (A.6) gives

$$\delta \mathbf{X}(t) = \int_0^t \exp\left(\int_u^t \partial_1 G(s) ds\right) \partial_2 G(u) \exp\left(\int_0^u d_{\mathbf{S}(v)} F(v) dv\right) \delta \mathbf{S}_0 du \\ + \exp\left(\int_0^t \partial_1 G(s) ds\right) \delta \mathbf{X}_0. \tag{A.7}$$

To simplify notation, let $R_{st} = \exp\left(\int_s^t d_{\mathbf{S}(u)} F(u) du\right)$ and $V_{st} = \exp\left(\int_s^t \partial_1 G(u) du\right)$ which gives

$$\delta \mathbf{X}(t_i) = \left(\int_0^{t_i} V_{ut_i} \partial_2 G(u) R_{0u} du\right) \delta \mathbf{S}_0 + V_{0t_i} \delta \mathbf{X}_0 \tag{A.8}$$

and now plugged into (A.3)

$$\delta E = \sum_{i=1}^{N_{obs}} \ (\nabla_{\mathbf{X}(t_i)} D(t_i))^t \left[\left(\int_0^{t_i} V_{ut_i} \partial_2 G(u) R_{0u} du\right) \delta \mathbf{S}_0 + V_{0t_i} \delta \mathbf{X}_0\right] \\ + (\nabla_{\mathbf{S}_0} L)^t \delta \mathbf{S}_0. \tag{A.9}$$

By rearranging terms, we can write the variation of the criterion as

$$\delta E = \sum_{i=1}^{N_{obs}} \left[\int_0^{t_i} R_{0u}{}^t \partial_2 G(u)^t V_{ut_i}{}^t \nabla_{\mathbf{X}(t_i)} D(t_i) du\right]^t \delta \mathbf{S}_0 \\ + [\nabla_{\mathbf{S}_0} L]^t \delta \mathbf{S}_0 \tag{A.10} \\ + \sum_{i=1}^{N_{obs}} \left[V_{0t_i}{}^t \nabla_{\mathbf{X}(t_i)} D(t_i)\right]^t \delta \mathbf{X}_0$$

which leads to

$$
\begin{cases}
\nabla_{\mathbf{S}_0} E = \displaystyle\int_0^{t_i} R_{0u}{}^t \partial_2 G(u)^t \sum_{i=1}^{N_{obs}} V_{ut_i}{}^t \nabla_{\mathbf{X}(t_i)} D(t_i) \mathbf{1}_{\{u \le t_i\}} du + \nabla_{\mathbf{S}_0} L \\
\nabla_{\mathbf{X}_0} E = \displaystyle\sum_{i=1}^{N_{obs}} V_{0t_i}{}^t \nabla_{\mathbf{X}(t_i)} D(t_i),
\end{cases}
\tag{A.11}
$$

where $\mathbf{1}_{\{u \le t_i\}} = 1$ whenever $u \le t_i$ and zero otherwise.

For further notational convenience, we introduce $\theta(t) = \sum_{i=1}^{N_{obs}} V_{tt_i}{}^t \nabla_{\mathbf{X}(t_i)} D(t_i) \mathbf{1}_{\{t \le t_i\}}$, $g(t) = \partial_2 G(t)\theta(t)$, and $\xi(t) = \int_t^{t_i} R_{tu}{}^t g(u) du$. The gradients are now

$$
\begin{cases}
\nabla_{\mathbf{S}_0} E = \displaystyle\int_0^{t_i} R_{0u}{}^t g(u) du + \nabla_{\mathbf{S}_0} L = \xi(0) + \nabla_{\mathbf{S}_0} L \\
\nabla_{\mathbf{X}_0} E = \theta(0).
\end{cases}
\tag{A.12}
$$

To compute $\theta(t)$, note that for any time greater than the latest observation $T > t_f$, the value of $\mathbf{1}_{\{t \le t_i\}} = 0$ and therefore $\theta(T) = 0$ and $\dot\theta(t) = -\partial_1 G(t)^t \theta(t) + \sum_{i=1}^{N_{Obs}} V_{tt_i} \nabla_{\mathbf{X}(t_i)} D(t_i) \delta(t - t_i)$. Note that in the second term we have $V_{tt_i} = 1$ whenever $t = t_i$ and $\delta(t - t_i)$ is nonzero only when $t = t_i$. Therefore we write

$$
\begin{cases}
\theta(T) = 0 \\
\dot\theta(t) = -\partial_1 G(t)^t \theta(t) + \displaystyle\sum_{i=1}^{N_{Obs}} \nabla_{\mathbf{X}(t_i)} D(t_i) \delta(t - t_i).
\end{cases}
\tag{A.13}
$$

To compute $\xi(t)$, note that for any time greater than the latest observation $T > t_f$, the value of $\xi(T) = 0$. Also notice that $R_{ts} = \mathrm{ID} - \int_t^s \frac{dR_{us}}{du} du = \mathrm{ID} + \int_t^s R_{us} d_{\mathbf{S}(u)} F(u) du$. Using Fubini's theorem gives

$$
\begin{aligned}
\xi(t) &= \int_t^{t_i} R_{tu}{}^t g(u) du \\
&= \int_t^{t_i} g(u) + d_{\mathbf{S}(u)} F(u)^t \int_u^{t_i} R_{uv}{}^t g(v) dv du \\
&= \int_t^{t_i} g(u) + d_{\mathbf{S}(u)} F(u)^t \xi(u) du,
\end{aligned}
\tag{A.14}
$$

and finally

$$
\begin{cases}
\xi(T) = 0 \\
\dot\xi(t) = -(\partial_2 G(t)^t \theta(t) + d_{\mathbf{S}(t)} F(t)^t \xi(t)).
\end{cases}
\tag{A.15}
$$

35

Gradient in coordinates

Expanding the variables $\mathbf{S}(t) = \{\mathbf{c}_{0,k}(t), \alpha_{0,k}(t)\}$, $\mathbf{X}(t) = \{X_k(t)\}$, $\theta(t) = \{\theta_k(t)\}$, and $\xi(t) = \{\xi_k^c(t), \xi_k^\alpha(t)\}$ gives

$$\nabla_{\mathbf{c}_{0,k}} E = \xi_k^c(0) + \nabla_{\mathbf{c}_{0,k}} L(\mathbf{S}_0) \tag{A.16}$$

$$\nabla_{\alpha_{0,k}} E = \xi_k^\alpha(0) + \nabla_{\alpha_{0,k}} L(\mathbf{S}_0) \tag{A.17}$$

where the gradient of the regularity term is written

$$\nabla_{c_k} L = 2 \sum_{p=1}^{N_c} \alpha_p{}^t \alpha_k \nabla_1 K(\mathbf{c}_k, \mathbf{c}_p) \tag{A.18}$$

$$\nabla_{\alpha_k} L = 2 \sum_{p=1}^{N_c} K(\mathbf{c}_k, \mathbf{c}_p) \alpha_p. \tag{A.19}$$

Computation of $\dot{\theta}(t)$

The term $\partial_1 G(\mathbf{X}(t), \mathbf{S}(t))$ is a block-matrix of size $3N_c \times 3N_x$ whose $(k, p)th$ 3×3 block is given as

$$d_{X_k} G(\mathbf{X}(t), \mathbf{S}(t))_p = \sum_{j=1}^{N_c} \alpha_j(t) \nabla_1 K(X_p(t), c_j(t))^t \delta(p - k) \tag{A.20}$$

so that the vector $\theta(t)$ is updated according to

$$\dot{\theta}_k(t) = - \left[\sum_{p=1}^{N_c} \alpha_p(t)^t \theta_k(t) \nabla_1 K(X_k(t), c_p(t)) \right] + \sum_{i=1}^{N_{Obs}} \nabla_{X_k(t_i)} D\delta(t - t_i). \tag{A.21}$$

Computation of $\dot{\xi}(t) = (\dot{\xi}^c(t), \dot{\xi}^\alpha(t))$

The terms $\partial_{\mathbf{c}} G(\mathbf{X}(t), \mathbf{S}(t))$ and $\partial_\alpha G(\mathbf{X}(t), \mathbf{S}(t))$ are both matrices of size $3N_x \times 3N_c$, whose (k, p) block is given by

$$d_{ck} G_p = \alpha_k (\nabla_1 K(c_k, X_p))^t \tag{A.22}$$

$$d_{\alpha k} G_p = K(c_k, X_p) I_3. \tag{A.23}$$

The differential of the function $F(\mathbf{S}) = \begin{pmatrix} F^c(\mathbf{c}, \boldsymbol{\alpha}) \\ F^\alpha(\mathbf{c}, \boldsymbol{\alpha}) \end{pmatrix}$ can be decomposed into 4 blocks as

$$d_{\mathbf{S}(t)}F = \begin{pmatrix} \partial_{\mathbf{c}}F^c & \partial_{\boldsymbol{\alpha}}F^c \\ \partial_{\mathbf{c}}F^\alpha & \partial_{\boldsymbol{\alpha}}F^\alpha \end{pmatrix}. \tag{A.24}$$

Therefore, the update rules for the auxiliary variables $\xi^c(t)$ and $\xi^\alpha(t)$ are

$$\begin{cases} \dot{\xi}_k^c(t) = -\left[\sum_{p=1}^{N_x} \alpha_k(t)^t \theta_p(t) \nabla_1 K(c_k(t), X_p(t)) \right. \\ \qquad\qquad\qquad\qquad \left. + (\partial_{\mathbf{c}}F^c)^t \xi_k^c(t) + (\partial_{\mathbf{c}}F^\alpha)^t \xi_k^\alpha(t) \right] \\ \dot{\xi}_k^\alpha(t) = -\left[\sum_{p=1}^{N_x} K(c_k(t), X_p(t)) \theta_p(t) \right. \\ \qquad\qquad\qquad\qquad \left. + (\partial_{\alpha}F^c)^t \xi_k^c(t) + (\partial_{\alpha}F^\alpha)^t \xi_k^\alpha(t) \right] \end{cases} \tag{A.25}$$

with

$$
\begin{aligned}
(\partial_{\mathbf{c}}F^c)^t \xi_k^c(t) &= \sum_{p=1}^{N_c} \left(\alpha_p(t)^t \xi_k^c(t) + \alpha_k(t)^t \xi_p^c(t) \right) \nabla_1 K(c_k(t), c_p(t)) \\
(\partial_{\mathbf{c}}F^\alpha)^t \xi_k^\alpha(t) &= \sum_{p=1}^{N_c} \alpha_k(t)^t \alpha_p(t) \nabla_{1,1} K(c_k(t), c_p(t))^t (\xi_p^\alpha(t) - \xi_k^\alpha(t)) \\
(\partial_{\alpha}F^c)^t \xi_k^c(t) &= \sum_{p=1}^{N_c} K(c_k(t), c_p(t)) \xi_j^c(t) \\
(\partial_{\alpha}F^\alpha)^t \xi_k^\alpha(t) &= \sum_{p=1}^{N_c} \nabla_1 K(c_k(t), c_p(t))^t \left(\xi_p^\alpha(t) - \xi_k^\alpha(t) \right) \alpha_p(t).
\end{aligned}
\tag{A.26}
$$

If the kernel is a scalar isotropic kernel of the form $K = f(||x - y||^2)\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, then

$$
\begin{aligned}
\nabla_1 K(x, y) &= 2f'(||x - y||^2)(x - y) \\
\nabla_{1,1} K(x, y) &= 4f''(||x - y||^2)(x - y)(x - y)^t + 2f'(||x - y||^2)\mathbf{I}.
\end{aligned}
\tag{A.27}
$$

References

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.

Banerjee, M., Chakraborty, R., Ofori, E., Vaillancourt, D., & Vemuri, B. C. (2015). Nonlinear regression on riemannian manifolds and its applications to neuro-image analysis, . (pp. 719–27). doi:`10.1007/978-3-319-24553-9_88`.

Benraiss, A., & Goldman, S. A. (2011). Cellular therapy and induced neuronal replacement for huntington's disease. *Neurotherapeutics*, *8*, 577–90. doi:`10.1007/s13311-011-0075-8`.

Charon, N., & Trouvé, A. (2013). The varifold representation of non-oriented shapes for diffeomorphic registration. *CoRR*, *abs/1304.6108*. URL: `http://arxiv.org/abs/1304.6108`.

Datar, M., Cates, J., Fletcher, P., Gouttard, S., Gerig, G., & Whitaker, R. (2009). Particle based shape regression of open surfaces with applications to developmental neuroimaging. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)* (pp. 167–74). volume 5762 of *LNCS*. doi:`10.1007/978-3-642-04271-3_21`.

Datar, M., Muralidharan, P., Kumar, A., Gouttard, S., Piven, J., Gerig, G., Whitaker, R., & Fletcher, P. (2012). Mixed-effects shape models for estimating longitudinal changes in anatomy. In S. Durrleman, P. T. Fletcher, G. Gerig, & M. Niethammer (Eds.), *Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data* (pp. 76–87). Springer Berlin / Heidelberg volume 7570 of *LNCS*. doi:`10.1007/978-3-642-33555-6_7`.

Davis, B., Fletcher, P., Bullitt, E., & Joshi, S. (2007). Population shape regression from random design data. In *Proc. of ICCV* (pp. 1–7).

Dupuis, P., & Grenander, U. (1998). Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, *LVI*, 587–600.

Durrleman, S., Allassonniĕre, S., & Joshi, S. (2013a). Sparse adaptive parameterization of variability in image ensembles. *International Journal of Computer Vision (IJCV)*, *101*, 161–83. doi:`10.1007/s11263-012-0556-1`.

Durrleman, S., Pennec, X., Trouvé, A., Braga, J., Gerig, G., & Ayache, N. (2013b). Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data. *International journal of computer vision*, *103*, 22–59.

Durrleman, S., Pennec, X., Trouvé, A., Gerig, G., & Ayache, N. (2009). Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets. In G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, & C. Taylor (Eds.), *Medical Image Computing and Computer Assisted Intervention (MICCAI)* (pp. 297–304). Springer volume 5761 of *LNCS*.

Durrleman, S., Prastawa, M., Charon, N., Korenberg, J., Joshi, S., Gerig, G., & Trouvé, A. (2014). Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *NeuroImage*, . doi:`10.1016/j.neuroimage.2014.06.043`.

Durrleman, S., Prastawa, M., Gerig, G., & Joshi, S. (2011). Optimal data-driven sparse parameterization of diffeomorphisms for population analysis. In *Information Processing in Medical Imaging (IPMI)* (pp. 123–34). volume 6801/2011 of *LNCS*. doi:`10.1007/978-3-642-22092-0_11`.

Durrleman, S., Prastawa, M., Korenberg, J. R., Joshi, S. C., Trouvé, A., & Gerig, G. (2012). Topology preserving atlas construction from shape data without correspondence using sparse parameters. In *Medical Image Computing and Computer Assisted Intervention (MICCIA)* LNCS (pp. 223–30). Springer.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M. et al. (2012). 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, .

Fishbaugh, J., Durrleman, S., & Gerig, G. (2011). Estimation of smooth growth trajectories with controlled acceleration from time series shape data. In G. Fichtinger, & T. Peters (Eds.), *Medical Image Computing and Computer Assisted Intervention (MICCIA)* (pp. 401–8). Springer volume 6982 of *LNCS*.

Fishbaugh, J., Prastawa, M., Durrleman, S., Piven, J., & Gerig, G. (2012). Analysis of longitudinal shape variability via subject specific growth mod-

eling. In N. Ayache, H. Delingette, P. Golland, & K. Mori (Eds.), *Medical Image Computing and Computer Assisted Intervention (MICCIA)* (pp. 731–8). Springer volume 7510 of *LNCS*.

Fishbaugh, J., Prastawa, M., Gerig, G., & Durrleman, S. (2013a). Geodesic image regression with a sparse parameterization of diffeomorphisms. In *Geometric Science of Information (GSI)* (pp. 95–102). volume 8085.

Fishbaugh, J., Prastawa, M., Gerig, G., & Durrleman, S. (2013b). Geodesic shape regression in the framework of currents. In *Information Processing in Medical Imaging (IPMI)* (pp. 718–29). volume 23 of *LNCS*.

Fishbaugh, J., Prastawa, M., Gerig, G., & Durrleman, S. (2014). Geodesic regression of image and shape data for improved modeling of 4D trajectories. In *International Symposium on Biomedical Imaging (ISBI)* (pp. 385–8). IEEE.

Fletcher, P. (Session 03: Statistical analysis on Riemannian Manifolds) (2011). Geodesic Regression on Riemannian Manifolds. In X. Pennec, S. Joshi, & M. Nielsen (Eds.), *MICCAI MFCA* (pp. 75–86).

Fletcher, P. T. (2013). Geodesic regression and the theory of least squares on riemannian manifolds. *International Journal of Computer Vision (IJCV)*, *105*, 171–85.

Gerig, G., Jomier, M., & Chakos, M. (2001). Valmet: A new validation tool for assessing and improving 3d object segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 516–23). Springer Berlin Heidelberg.

Glaunès, J. (2005). *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique*. Thèse de sciences (PhD Thesis) Université Paris.

Goodlett, C. B., Fletcher, P. T., Gilmore, J. H., & Gerig, G. (2009). Group analysis of dti fiber tract statistics with application to neurodevelopment. *Neuroimage*, *45*, S133–42.

Gouttard, S., Styner, M., Joshi, S., Smith, R. G., Hazlett, H. C., & Gerig, G. (2007). Subcortical structure segmentation using probabilistic atlas priors.

In *Medical Imaging* (pp. 65122J–). International Society for Optics and Photonics.

Hart, G., Shi, Y., Zhu, H., Sanchez, M., Styner, M., & Niethammer, M. (2010). DTI longitudinal atlas construction as an average of growth models. In G. Gerig, P. Fletcher, & X. Pennec (Eds.), *MICCAI STIA*.

Hinkle, J., Fletcher, P., & Joshi, S. (2014). Intrinsic polynomials for regression on Riemannian manifolds. *Journal of Mathematical Imaging and Vision*, (pp. 1–21).

Hong, Y., Kwitt, R., Singh, N., Davis, B., Vasconcelos, N., & Niethammer, M. (2014). Geodesic regression on the grassmannian. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part II* (pp. 632–46). Cham: Springer International Publishing. URL: `http://dx.doi.org/10.1007/978-3-319-10605-2_41`. doi:`10.1007/978-3-319-10605-2_41`.

Hong, Y., Singh, N., Kwitt, R., Vasconcelos, N., & Niethammer, M. (2016). Parametric regression on the grassmannian. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, . URL: `http://arxiv.org/abs/1505.03832`.

Joshi, S., Davis, B., Jomier, M., & Gerig, G. (2004). Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, *23*, 151–60.

Joshi, S., & Miller, M. (2000). Landmark matching via large deformation diffeomorphisms. *Transactions on Image Processing*, *9*, 1357–70.

Kendall, D. G. (1984). Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, *16*, 81–121. doi:`10.1112/blms/16.2.81`.

Marsland, S., & Twining, C. J. (2004). Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images. *Transactions on Medical Imaging*, *23*, 1006–20.

Miller, M. I., Trouvé, A., & Younes, L. (2006). Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, *24*, 209–28.

Niethammer, M., Huang, Y., & Vialard, F. (2011). Geodesic regression for image time-series. In *Medical Image Computing and Computer Assisted Intervention (MICCIA)* (pp. 655–62). volume 6892 of *LNCS*.

Paulsen, J. S., Magnotta, V. A., Mikos, A. E., Paulson, H. L., Penziner, E., Andreasen, N. C., & Nopoulos, P. C. (2006). Brain structure in preclinical huntington's disease. *Biological Psychiatry*, *59*, 57–63. doi:`10.1016/j.biopsych.2005.06.003`.

Paulsen, J. S., Nopoulos, P. C., Aylward, E., Ross, C. A., Johnson, H., Magnotta, V. A., Juhl, A., Pierson, R. K., Mills, J., Langbehn, D., & et al. (2010). Striatal and white matter predictors of estimated diagnosis for huntington disease. *Brain research bulletin*, . URL: `https://www.ncbi.nlm.nih.gov/pmc/articles/pmc2892238/`.

Shi, F., Yap, P.-T., Wu, G., Jia, H., Gilmore, J. H., Lin, W., & Shen, D. (2011). Infant brain atlases from neonates to 1- and 2-year-olds. *PLoS ONE*, *6*, 1–11. URL: `http://dx.doi.org/10.1371%2Fjournal.pone.0018746`.

Singh, N., Hinkle, J., Joshi, S., & Fletcher, P. (2013). A vector momenta formulation of diffeomorphisms for improved geodesic regression and atlas construction. In *International Symposium on Biomedical Imaging (ISBI)* (pp. 1219–22). IEEE. doi:`10.1109/ISBI.2013.6556700`.

Singh, N., Hinkle, J., Joshi, S., & Fletcher, P. T. (2016). Hierarchical geodesic models in diffeomorphisms. *International Journal of Computer Vision*, *117*, 70–92.

Singh, N., Vialard, F.-X., & Niethammer, M. (2015). Splines for diffeomorphisms. *Medical Image Analysis*, .

Trouvé, A. (1995). Diffeomorphisms groups and pattern matching in image analysis.

Vaillant, M., & Glaunès, J. (2005). Surface matching via currents. In *Information Processing in Medical Imaging (IPMI)* (pp. 381–92). volume 3565 of *LNCS*.

Vialard, F., & Trouvé, A. (2012). Shape splines and stochastic shape evolutions: A second-order point of view. *Quarterly of Applied Mathematics*, *70*, 219–51.

Vialard, F.-X., Risser, L., Rueckert, D., & Cotter, C. J. (2012). Diffeomorphic 3D image registration via geodesic shooting using an efficient adjoint calculation. *International Journal of Computer Vision (IJCV)*, *97*, 229–41. doi:`10.1007/s11263-011-0481-8`.

Wang, L., Shi, F., Li, G., & Shen, D. (2012). 4d segmentation of longitudinal brain mr images with consistent cortical thickness measurement. In S. Durrleman, T. Fletcher, G. Gerig, & M. Niethammer (Eds.), *Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data: Second International Workshop, STIA 2012, Held in Conjunction with MICCAI 2012, Nice, France, October 1, 2012. Proceedings* (pp. 63–75). Berlin, Heidelberg: Springer Berlin Heidelberg. URL: `http://dx.doi.org/10.1007/978-3-642-33555-6_6`. doi:`10.1007/978-3-642-33555-6_6`.