

# Splines for Diffeomorphisms

Nikhil Singh<sup>a</sup>, François-Xavier Vialard<sup>b</sup>, Marc Niethammer<sup>a</sup>

<sup>a</sup>The University of North Carolina, Chapel Hill, USA

<sup>b</sup>University Paris-Dauphine, Paris, France

---

## Abstract

This paper develops a method for higher order parametric regression on diffeomorphisms for image regression. We present a principled way to define curves with nonzero acceleration and nonzero jerk. This work extends methods based on geodesics which have been developed during the last decade for computational anatomy in the large deformation diffeomorphic image analysis framework. In contrast to previously proposed methods to capture image changes over time, such as geodesic regression, the proposed method can capture more complex spatio-temporal deformations.

We take a variational approach that is governed by an underlying energy formulation, which respects the nonflat geometry of diffeomorphisms. Such an approach of minimal energy curve estimation also provides a physical analogy to particle motion under a varying force field. This gives rise to the notion of the quadratic, the cubic and the piecewise cubic splines on the manifold of diffeomorphisms. The variational formulation of splines also allows for the use of temporal control points to control spline behavior. This necessitates the development of a shooting formulation for splines.

The initial conditions of our proposed shooting polynomial paths in diffeomorphisms are analogous to the Euclidean polynomial coefficients. We experimentally demonstrate the effectiveness of using the parametric curves both for synthesizing polynomial paths and for regression of imaging data. The performance of the method is compared to geodesic regression.

*Keywords:* LDDMM, Diffeomorphisms, Splines, Image Regression, Polynomials, Time Series

---

## 1. Introduction

With the now common availability of longitudinal and time-series image data, models for their analysis are critically needed. In particular, spatial correspondences need to be established through image registration for many medical image analysis tasks. While this can be accomplished by pair-wise image registration to a template image, such an approach neglects spatio-temporal data aspects. Instead, explicitly accounting for spatial *and* temporal dependencies is desirable.

A common way to describe differences in geometry of objects in images is to summarize them using transformations. Transformations are fundamental mathematical objects and have long been known to effectively represent biological changes in organisms (Thompson et al., 1942; Amit et al., 1991). The field of *computational anatomy* (Miller et al., 1997; Grenander and Miller, 1998; Thompson and Toga, 2002; Miller, 2004) provides a rich mathematical setting for statistical analysis of complex geometrical structures seen in 3D medical images. At its core, computational anatomy is based on the representation of anatomical shape and its variability using smooth and invertible transformations that are elements of the nonflat manifold of diffeomorphisms with an associated Riemannian structure. The large deformation (LDDMM) framework of computational anatomy exploits ideas from fluid mechanics and builds maps of diffeomorphisms as flows of smooth

velocity fields (Younes, 2010; Younes et al., 2009).

Research in the last decade provided several methods to represent natural biological variability by modeling them as nonlinear transformations in the manifold of diffeomorphisms. Their focus has primarily been on geodesic models. For example, methods of Fréchet mean (Davis, 2008), geodesic regression (Niethammer et al., 2011) and hierarchical geodesic models (Singh et al., 2013a) are first order models that rely on computing geodesics within the space of diffeomorphisms. While such models have proven to be effective, their use is limited to modeling only “geodesic-like” image data. However, geodesics are not always appropriate for regression modeling of time-series data. In particular, nonmonotonous shape changes seen in time sequence or videos of medical images of periodic breathing, cardiac motion, or shape changes in the human brain during a long age range (10-90 yrs), do generally not adhere to constraints of geodesicity. This necessitates the development of higher order models of regression within the space of diffeomorphic transformations. Computational anatomy has seen very little work on higher-order models of registrations for modeling image time series (Figure 1).

*Contribution.* In this article we propose:

1. an acceleration-controlled model that generalizes the idea of cubic curves to manifold of diffeomorphisms and is capable of modeling nonmonotonic shape

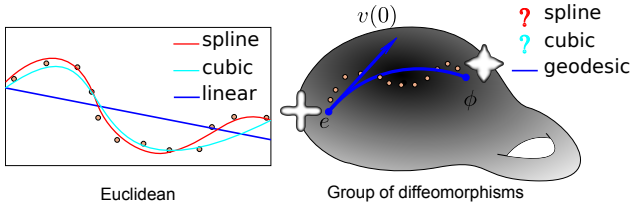


Figure 1: Models of parametric regression to explain data on diffeomorphisms.

changes under the large deformation (LDDMM) setting,

2. a shooting based solution to cubic curves that enables parametrization of the full regression path using only initial conditions,
3. a method of shooting cubic splines as smooth curves to fit complex shape trends while keeping data-independent (finite and few) parameters, and
4. a numerically practical algorithm for regression of “non-geodesic” medical imaging data.

The work described in this manuscript significantly extends our work presented at MICCAI (Singh and Niethammer (2014)). In particular, (1) we make use of a new formulation directly advecting the inverse of a diffeomorphism, (2) we provide extended discussions of the approach, and (3) present a variety of new results to illustrate the behavior of the approach.

### 1.1. Related work

Methods that generalize Euclidean parametric regression models to manifolds have proven to be effective for modeling the dynamics of changes represented in time-series of medical images. For instance, methods of geodesic image regression (Niethammer et al., 2011; Singh et al., 2013b) and longitudinal models on images (Singh et al., 2013a) generalize linear and hierarchical linear models, respectively. Although the idea of polynomials (Hinkle et al., 2014) and splines (Trouvé and Vialard, 2012) on the landmark representation of shapes have been proposed, higher-order extensions for image regression remain deficient. While Hinkle et al. (2014) develop an approach for general polynomial regression and demonstrate it on finite-dimensional Lie groups, infinite dimensional regression is demonstrated only for first-order geodesic image regression.

These parametric regression models are advantageous since their estimated parameters can be used for further statistical analysis. For instance, initial momenta obtained from Fréchet atlas construction of a population of images can be treated as signature representations of shape differences across the group and can be treated as features to train classification and regression models (Singh et al., 2014). Other regression methods include those by Davis et al. (2010) that generalize notion of kernel regression to manifolds. Kernel regression is a nonparametric approach

and hence does not provide a summary representation of the regression fit in terms of a finite set of parameters for further analysis.

The remainder of this article is structured as follows: Section 2 reviews the variational approach to splines in Euclidean space and motivate its shooting formulation for parametric regression. Section 3 then generalizes this concept of shooting splines for diffeomorphic image regression. We discuss experimental results in Section 4, and conclude the article with a discussion of future work in Section 5.

## 2. Shooting-splines in the Euclidean Case

To motivate our formulation for splines on diffeomorphisms it is instructive to first revisit the variational formulation for splines in the Euclidean case. This facilitates a more straightforward presentation of the fundamental approach and allows to make direct connections to the formulation for splines on diffeomorphisms.

### 2.1. Variational formulation.

An acceleration controlled curve with time-dependent states,  $(x_1, x_2, x_3)$  such that,  $\dot{x}_1 = x_2$  and  $\dot{x}_2 = x_3$ , defines a cubic curve in Euclidean spaces for a constant acceleration,  $x_3$ . Here,  $x_1$  denotes position and  $x_2$  velocity. In particular, such a cubic curve minimizes an energy of the form,  $E = \frac{1}{2} \int_0^1 \|x_3\|^2 dt$  subject to the dynamic constraints above. The corresponding constrained optimization problem can be written as

$$\underset{x_1, x_2, x_3}{\text{minimize}} E(x_3) \text{ subject to } x_2 = \dot{x}_1 \text{ and } x_3 = \dot{x}_2. \quad (1)$$

Here  $x_3$  is referred to as the control variable that describes the acceleration of the dynamics in this system. The unconstrained Lagrangian for the above is,

$$E(x_1, x_2, x_3, \mu_1, \mu_2) = \frac{1}{2} \int_0^1 \|x_3\|^2 dt + \int_0^1 \mu_1^T (\dot{x}_1 - x_2) dt + \int_0^1 \mu_2^T (\dot{x}_2 - x_3) dt,$$

where  $\mu_1$  and  $\mu_2$  are the time-dependent Lagrangian variables or the *adjoint* variables (also called duals) that enforce the dynamic constraints. Optimality conditions on the gradients of the above Lagrangian with respect to the states,  $(x_1, x_2, x_3)$ , result in the adjoint system of equations,  $\dot{\mu}_1 = 0$  and  $\dot{x}_3 = -\mu_1$  ( $\mu_2$  gets eliminated). This allows for a relaxation solution to Eq. (1), where the state of the system is the full time-course of states, i.e.,  $(x_1(t), x_2(t), x_3(t))$ , and the condition  $x_3(t) = \text{const.}$  will be fulfilled at convergence. However, we may also formulate this problem with respect to initial conditions alone, amounting to a shooting solution as discussed in Section 2.2.

## 2.2. From relaxation to shooting.

A relaxation solution has originally been proposed for diffeomorphic image registration by Beg et al. (2005). Here, a full-spatio-temporal velocity field was the variable to be estimated. Instead, a shooting reformulation (Vialard et al., 2011) allowed to represent the image registration problems by optimizing over an initial image and an initial momentum. In the scalar-valued setting the shooting-formulation corresponds to optimizing over the initial y-intercept and slope of a line, thereby searching over the space of straight lines instead of converging to a straight line as in the relaxation setting. Shooting thereby allowed the formulation of geodesic regression approaches (Niethammer et al., 2011) where one aims to determine the best geodesic fitting the given data and optimized over the initial conditions specifying the geodesic only. Hence, to allow for splines on diffeomorphisms we also need a shooting formulation to be able to compactly represent splines and to express the equivalent of piecewise cubic curves. In the scalar-valued case such a shooting formulation can be obtained by explicitly adding the evolution of  $x_3$ , obtained by solving the relaxation problem, as a dynamical constraint. This increases the order of the dynamics. Denoting,  $x_4 = -\mu_1$ , results in the classical system of equations for shooting cubic curves,

$$\dot{x}_1 = x_2(t), \quad \dot{x}_2 = x_3(t), \quad \dot{x}_3 = x_4(t), \quad \dot{x}_4 = 0. \quad (2)$$

The states,  $(x_1, x_2, x_3, x_4)$ , at all times, are entirely determined by their initial values  $(x_1^0, x_2^0, x_3^0, x_4^0)$ , and in particular,  $x_1(t) = x_1^0 + x_2^0 t + \frac{x_3^0}{2} t^2 + \frac{x_4^0}{6} t^3$ . Also note that  $x_4$  is the derivative of acceleration,  $x_3$ , and can therefore be interpreted as jerk. For a cubic the jerk is constant. When piecing together multiple cubic curves, as will be described in Section 2.3, the jerk will be allowed to jump.

## 2.3. Shooting-splines with data-independent controls for regression.

We now present our proposed method of regression using cubic splines using the shooting equations. The goal is to define a *smooth* curve that best fits the data in the least-squares sense. Since a cubic polynomial by itself is restricted to only fit ‘‘cubic-like’’ data, we propose to add flexibility to the curve by piecing together piecewise cubic polynomials. In other words, we define controls at pre-decided locations in time where the state  $x_4$  is allowed to jump.

Let,  $y_i$ , for  $i = 1 \dots N$ , denote  $N$  measurements at time-points,  $t_i \in (0, 1)$ . Let  $t_c \in (0, 1)$ , for  $c = 1 \dots C$ , denote  $C$  data-independent fixed control locations. For notational convenience, we assume there are no measurements at the end points,  $\{0, 1\}$ , or at the control locations,  $\{t_c\}$ . The control locations also implicitly define  $C + 1$  intervals or partitions in  $(0, 1)$ . Let us denote these intervals as  $\mathcal{I}_c$ , for  $c = 1 \dots (C + 1)$ . The constrained energy minimization that solves the regression problem with such a data

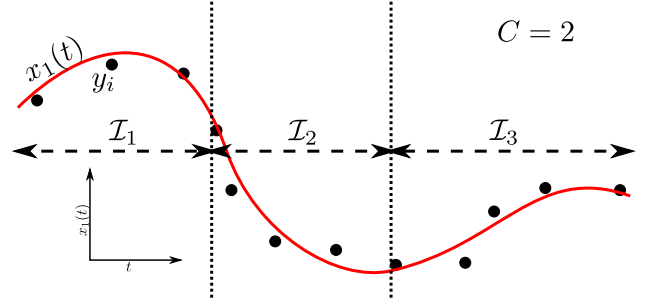


Figure 2: Gluing together cubics to construct a piecewise cubic curve. This example uses three partitions,  $\mathcal{I}_1, \mathcal{I}_2$  and  $\mathcal{I}_3$  defined by placing controls at two locations ( $C = 2$ ).

configuration can be written as,

$$\underset{x_1(0), x_2(0), x_3(0), x_4(0), \{x_4(t_c)\}}{\text{minimize}} \quad \frac{1}{2\sigma^2} \sum_{c=1}^{C+1} \sum_{i \in \mathcal{I}_c} \|x_1(t_i) - y_i\|^2$$

$$\text{s.t.} \quad \dot{x}_1 = x_2(t), \quad \dot{x}_2 = x_3(t), \quad \dot{x}_3 = x_4(t), \quad \dot{x}_4 = 0,$$

(within each interval,  $\mathcal{I}_c$ ), and

$$\text{s.t.} \quad x_1, x_2, \text{ and } x_3 \text{ are continuous across } C.$$

The partitioning of the domain of independent variable for regression for the case of three partitions using two control locations is depicted in Figure 2.

The unconstrained Lagrangian enforcing shooting and continuity constraints using time-dependent adjoint states,  $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ , and duals,  $(\nu_1, \nu_2, \nu_3)$ , is

$$\begin{aligned} E(x_1^0, x_2^0, x_3^0, x_4^0, x_4^{t_c}, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = & \\ & \frac{1}{2\sigma^2} \sum_{c=1}^{C+1} \sum_{i \in \mathcal{I}_c} \|x_1(t_i) - y_i\|^2 \\ & + \int_0^1 (\lambda_1^T (\dot{x}_1 - x_2) + \lambda_2^T (\dot{x}_2 - x_3) + \lambda_3^T (\dot{x}_3 - x_4) \\ & + \lambda_4^T \dot{x}_4) dt + \nu_1 (x_1^-(t_c) - x_1^+(t_c)) \\ & + \nu_2 (x_2^-(t_c) - x_2^+(t_c)) + \nu_3 (x_3^-(t_c) - x_3^+(t_c)). \end{aligned}$$

The conditions of optimality on the gradients of the above Lagrangian result in the adjoint system of equations,  $\dot{\lambda}_1 = 0, \dot{\lambda}_2 = -\lambda_1, \dot{\lambda}_3 = -\lambda_2, \dot{\lambda}_4 = -\lambda_3$ . The gradients with respect to the initial conditions for states  $x_l^0$  for  $l = 1, \dots, 4$  are,  $\delta_{x_1^0} E = -\lambda_1(0)$ ,  $\delta_{x_2^0} E = -\lambda_2(0)$ ,  $\delta_{x_3^0} E = -\lambda_3(0)$  and  $\delta_{x_4^0} E = -\lambda_4(0)$ . The jerks at controls,  $x_4^{t_c}$ , are updated using,  $\delta_{x_4^{t_c}} E = -\lambda_4(t_c)$ . The values of adjoint variables required in these gradients are computed by integrating backward the adjoint system. Note that  $\lambda_1, \lambda_2$  and  $\lambda_3$  are continuous at joins, but  $\lambda_1$  jumps at the data-point location as per,  $\lambda_1(t_i^+) - \lambda_1(t_i^-) = \frac{1}{\sigma^2} (x_1(t_i) - y_i)$ . During backward integration,  $\lambda_4$  starts from zero at each interval at  $t_{c+1}$  and the accumulated value at  $t_c$  is used for the gradient update of  $x_4(t_c)$ .

It is critical to note that, along the time,  $t$ , such a formulation guarantees that: (a)  $x_4(t)$  is piecewise constant, (b)

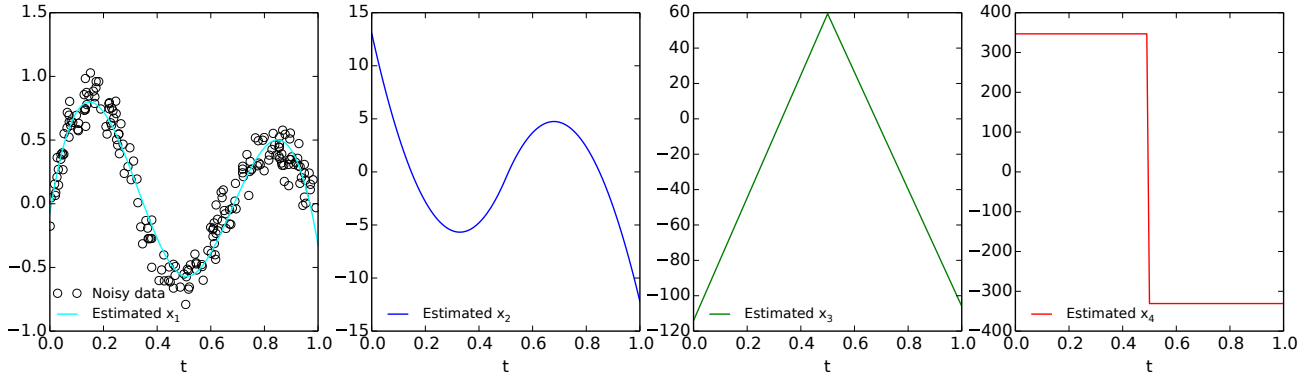


Figure 3: States for splines regression in Euclidean space with one control at  $t=0.5$ .

$x_3(t)$  is piecewise linear, (c)  $x_2(t)$  is piecewise quadratic, and (d)  $x_1(t)$  is piecewise cubic. Thus, this results in a cubic-spline curve. Figure 3 demonstrates this shooting spline fitting on scalar data. While it is not possible to explain this data with a simple cubic curve alone, it suffices to allow one control location to recover the meaningful underlying trend. The state,  $x_4$ , experiences a jump at the control location that integrates up thrice to give a  $C^2$ -continuous evolution for the state,  $x_1$ .

### 3. Shooting-splines for Diffeomorphisms

Our goal is to generalize the variational approach to cubic splines to the group of diffeomorphism to define splines that can capture for example complex image deformations over time. Our approach is completely analogous to the scalar-valued case described in earlier sections.

#### 3.1. Notations and preliminaries.

We denote the group of diffeomorphisms by  $G$  and its elements by  $g$ ; the tangent space at  $g$  by  $T_g G$ ; and the Lie algebra,  $T_e G$ , by  $\mathfrak{g}$ . Let  $\Omega$  be the coordinate space of the image,  $I$ . A diffeomorphism,  $g(t)$ , is constructed by integrating an ordinary differential equation (ODE) on  $\Omega$  defined via a smooth, time-indexed velocity field,  $v(t)$ . The deformation of an image  $I$  by  $g$  is defined as the action of the diffeomorphism, given by  $g \cdot I = I \circ g^{-1}$ . The choice of a self-adjoint differential operator,  $L$ , determines the right-invariant Riemannian structure on the collection of velocity fields with the norm defined as,  $\|v\|_{\mathfrak{g}}^2 = \int_{\Omega} (Lv(x), v(x)) dx$ . The velocity,  $v \in \mathfrak{g}$ , maps to its dual deformation momenta,  $m \in \mathfrak{g}^*$ , via the operator  $L$  such that  $m = Lv$  and  $v = K \star m$ . The operator  $K : \mathfrak{g}^* \rightarrow \mathfrak{g}$  denotes the inverse of  $L$ . For a thorough review of the Riemannian structure on the group of diffeomorphisms, please refer to [Younes \(2010\)](#), [Arnol'd \(1966\)](#) and [Younes et al. \(2009\)](#).

#### 3.2. Variational formulation.

Let us introduce curves of minimal acceleration on a general Riemannian manifold ([Noakes et al., 1989](#); [Camarinha et al., 1995](#)) which are needed for the image case. As a boundary value problem, Riemannian cubic splines are defined as curves that minimize the following energy  $E(g) = \frac{1}{2} \int_0^1 \|\nabla_{\dot{g}} \dot{g}\|_{T_g G}^2 dt$  subject to boundary constraints  $g(0) = g_0, \dot{g}(0) = v_0$  and  $g(1) = g_1, \dot{g}(1) = v_1$ . Here,  $\nabla$  denotes the Levi-Civita connection associated with the Riemannian metric denoted by  $\|\cdot\|_{T_g G}$ . The quantity,  $\nabla_{\dot{g}} \dot{g}$ , is the generalization of the idea of acceleration to Riemannian manifolds. Note that the associated Euler-Lagrange equation is in the form

$$\frac{D^3}{Dt^3} \dot{x} - R(\dot{x}, \frac{D}{Dt} \dot{x}) \dot{x} = 0, \quad (3)$$

which involves the curvature tensor  $R$  associated with the metric. Another way to define the spline is by defining a time-dependent *control* that forces the curve  $g(t)$  to deviate from being a geodesic ([Trouvé and Vialard, 2012](#)). Such a control or a forcing variable,  $u(t)$  is then integrated using the formula  $\nabla_{\dot{g}(t)} \dot{g}(t) = u(t)$ . Notice,  $u(t) = 0$  implies that  $g(t)$  is a geodesic.

Taking this idea forward to the group of diffeomorphisms,  $G$ , we propose to include a time-dependent forcing term that describes how much the ‘geodesic requirement’ deviates. Thus, we define the control directly on the known momenta EPDiff evolution equation for geodesics ([Younes et al., 2009](#)), obtained using the right invariant metric to give the evolution in the Lie algebra (in fact, the tangent space at the identity deformation),  $\mathfrak{g}$ , as  $\dot{m} + \text{ad}_v^* m = 0$ . Here, the operator  $\text{ad}^*$  is the adjoint of the Jacobi-Lie bracket ([Bruveris et al., 2011](#); [Younes et al., 2009](#)). After adding this control, the dynamics take the form,  $\dot{m} + \text{ad}_v^* m = u$ , where  $u \in \mathfrak{g}^*$ . Thus we allow the geodesic to deviate from satisfying the EPDiff constraints and constrain it to minimize an energy of the form,  $E = \frac{1}{2} \int_0^1 \|u(t)\|_{\mathfrak{g}^*}^2 dt$ . It is important to note that such a formulation will avoid direct computation of curvature. We bypass it when we control the EPDiff in  $\mathfrak{g}$  instead of controlling  $\nabla_{\dot{g}(t)} \dot{g}(t)$  in  $T_{g(t)} G$ .



The associated constrained energy minimization problem for splines is then,

$$\underset{u}{\text{minimize}} \quad \frac{1}{2} \int_0^1 \|u(t)\|_{\mathfrak{g}}^2 dt \quad (4)$$

$$\dot{m}(t) - u(t) + \text{ad}_{K m(t)}^* m(t) = 0, \quad m(0) = m_0 \quad (5)$$

$$\dot{h}(t) + (Dh)Km = 0, \quad h(0) = id \quad (6)$$

Here we use,  $h(t) = g^{-1}(t)$  for ease of notations and  $id$  denotes the identity map. The constraint in (5) is the control and (6) is a deformation advection constraint, which describes how the the inverse of the diffeomorphism evolves over time. This equation will also play a central role for the matching terms as it allows to match for example, images, landmarks, surfaces, etc.

Similar to the Euclidean case, the Euler-Lagrange equations for the above optimization problem give an adjoint system that explains the evolution of  $u$ , such that,

$$\dot{u} - (Dh)^\top p - K^{-1} \text{ad}_{Km} Ku + \text{ad}_{Ku}^* m = 0, \quad (7)$$

$$\dot{p} + \nabla \cdot (p \otimes Km) = 0. \quad (8)$$

Here,  $p$  is the adjoint variable corresponding to the deformation evolution constraint on  $g^{-1}$  in Eq. (6). The details of the derivation of the above Euler-Lagrange equations are presented in [Appendix A](#). Note however that the existence of a minimizer to this variational problem is still an open problem whereas the shooting splines solutions introduced in the next section are well-defined.

### 3.3. From relaxation to shooting.

Notice that the above discussion is analogous to the discussion of the relaxation formulation for the Euclidean case in the sense that the Euclidean states  $(x_1, x_2, x_3) \mapsto (h, m, u)$  in diffeomorphisms. We now convert the adjoint state,  $p(t)$  to a primal state to form a forward shooting system. Analogous to the Euclidean case, this increases the order of the system by one. The shooting system for acceleration controlled motion is,

$$\left. \begin{aligned} \dot{h} + (Dh)Km &= 0, \\ \dot{m} - u + \text{ad}_K^* m &= 0, \\ \dot{u} - (Dh)^\top p - K^{-1} \text{ad}_{Km} Ku + \text{ad}_{Ku}^* m &= 0, \\ \dot{p} + \nabla \cdot (p \otimes Km) &= 0. \end{aligned} \right\} \quad (9)$$

The image evolves (equivalently advects) as per the group action of  $g$  on the initial image  $I^0$ , i.e.,  $I(t) = I^0 \circ g^{-1}(t) = I^0 \circ h(t)$ . Here, the vector quantity,  $p$  is analogous to  $x_4$ .

### 3.4. Shooting-splines with data-independent controls for regression.

Similar to the data configuration in the Euclidean example, in the context of regression, let,  $J_i$ , for  $i = 1 \dots N$ , denote  $N$  measured images at timepoints,  $t_i \in (0, 1)$ . The goal now is to define finite and relatively fewer points than the number of measurements in the interval,  $(0, 1)$  where

$p$  is allowed to jump. In other words,  $p$  does not jump at every measurement but instead, is allowed to be free at predefined time-points that are decided independently of the data. Thus, we construct a curve  $g(t)$ , similar to the Euclidean case, in  $G$ , along the time,  $t$ , such that it guarantees, (a)  $p$  may jump at predefined time-points only, (b)  $u$  is  $C^0$ -continuous, (c)  $m$  is  $C^1$ -continuous, and (d)  $g$  is  $C^2$ -continuous.

The unconstrained Lagrangian for spline regression for a fixed initial image  $I^0$  takes the form:

$$E(h, m, u, p, p^{t_c}, \lambda_h, \lambda_m, \lambda_u, \lambda_p) = \quad (10)$$

$$\begin{aligned} & \frac{1}{2\sigma_I^2} \sum_{c=1}^{C+1} \sum_{i \in \mathcal{I}_c} d^2(I^0 \circ h(t_i), J_i) + \frac{1}{2\sigma_u^2} \int_0^1 \langle u, Ku \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{hc}, \dot{h}_c + (Dh_c)Km_c \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{mc}, \dot{m}_c - u_c + \text{ad}_{Km_c}^* m_c \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{uc}, \dot{u}_c - (Dh_c)^\top p_c - K^{-1} \text{ad}_{Km_c} Ku_c \\ & + \text{ad}_{Ku_c}^* m_c \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{pc}, \dot{p}_c + \nabla \cdot (p_c \otimes Km_c) \rangle_{L^2} dt, \text{ and} \end{aligned}$$

subject to continuity of  $h, m, u$ , and  $p$  at  $C$  joins.

Notice the second term is the relaxation energy term,  $\|u\|_{\mathfrak{g}}^2$ , which acts as a regularizer on the force along the full path in diffeomorphisms.

### 3.5. Gradients.

The optimality conditions on the gradients of the above energy functional show that the adjoint variables,  $\lambda_h, \lambda_m, \lambda_u$ , are continuous at all  $C$  joins. The gradients with respect to the initial conditions are,

$$\delta_{m_0} E = -\lambda_{m1}(0), \quad \delta_{u_0} E = -\lambda_{u1}(0), \quad \delta_{p_0} E = -\lambda_{p1}(0).$$

We compute the gradients by integrating the adjoint system of equations within each interval backward in time,

$$\begin{aligned} \dot{\lambda}_{hc} - \nabla \cdot (p_c \otimes \lambda_{uc} - \lambda_{hc} \otimes Km_c) &= 0 \\ \dot{\lambda}_{mc} - \text{ad}_{Km_c} \lambda_{mc} + K \text{ad}_{\lambda_{mc}}^* m_c - K \text{ad}_{Ku_c}^* K^{-1} \lambda_{uc} \\ - \text{ad}_{Ku_c} \lambda_{uc} + K((D\lambda_{pc})^\top p_c) - K((Dh_c)^\top \lambda_{hc}) &= 0 \\ \dot{\lambda}_{uc} - \frac{1}{\sigma_p^2} Ku_c + \lambda_{mc} + K \text{ad}_{Km_c}^* K^{-1} \lambda_{uc} \\ + K \text{ad}_{\lambda_{uc}}^* m_c &= 0 \\ \dot{\lambda}_{pc} + (Dh_c) \lambda_{uc} + (D\lambda_{pc}) Km_c &= 0 \end{aligned}$$

The details of the derivation of the above Euler-Lagrange equations are presented in [Appendix B](#).

All variables start from zero as their initial conditions for this backward integration. Similar to the Euclidean case, we add jumps in  $\lambda_h$  as,

$$\lambda_{hc}(t_i^+) - \lambda_{hc}(t_i^-) = \frac{1}{\sigma_I^2} (I^0 \circ h(t_i) - J_i) \nabla I^{t_i},$$

at measurements,  $t = t_i$  if we use a sum-of-squared differences similarity measure for  $d^2(\cdot, \cdot)$  between the measured images  $\{J_i\}$  and the estimated images  $\{I^0 \circ h(t_i)\}$ . More general similarity measures could easily be used and would only change these jump conditions. We ensure the continuity of  $\lambda_{hc}$ ,  $\lambda_{mc}$ , and  $\lambda_{uc}$  at the joins and  $\lambda_{pc}$  starts from zero at every join. We use the accumulated  $\lambda_{pc+1}$  to update the jerk,  $p_c(t_k)$ , at the control location with,

$$\delta_{p_{c+1}(t_c)} E = -\lambda_{pc+1}(t_c).$$

Note this is the ‘data independent’ control that we motivated our formulation with. This determines the initial condition of the forward system for each interval and needs to be estimated numerically. Also note that other regularizers can be added on the initial momenta,  $m_1^0$ , initial jerk,  $p_1^0$ , and jerks at controls,  $p_{c+1}^{t_c}$ , by restricting their Sobolev norms. In this case, the gradient includes additional terms of the form,  $Km_1^0$ ,  $Kp_1^0$ , and  $Kp_{c+1}^{t_c}$ , respectively. The estimate for the force term,  $u_0$ , does not need to be regularized since minimizing the norm on  $u(t)$  along the path itself acts as a regularizer.

## 4. Results

We evaluate our proposed model using synthetic data and two real time-sequence imaging data sets. One of the real imaging data examples is from cellular imaging of snapshots acquired for a deforming cell imaged using atomic force microscopy. The other is from the Sunnybrook cardiac MR database (Radau et al., 2009). In these experiments, the kernel,  $K$ , corresponds to the invertible and self-adjoint Sobolev operator,  $L = -a\nabla^2 - b\nabla(\nabla\cdot) + c$ , with  $a = 0.2$ ,  $b = 0.2$ , and  $c = 0.001$ . We use fourth order Runge-Kutta to integrate the primal states forward and to integrate the corresponding adjoint states backwards. We use a line search with gradient descent to estimate optimal initial states of spline curves and the controls. We fix the initial image,  $I(0)$ , and estimate initial states,  $m(0)$ ,  $u(0)$ ,  $p(0)$  and  $p(t_c)$  at control locations that completely determine the spline curve  $g(t)$  from  $t = 0$  to  $t = 1$ . In our validations, we first experimentally demonstrate that the shooting equations for evolving cubic curves in diffeomorphisms as per the set of ODE’s in Eq. (9) are indeed analogous to the classical system of shooting cubic curves in Euclidean as given in Eq. (2) and to the cubic equation,  $y = ax^3 + bx^2 + cx + d$ . We discuss these results in Sec. 4.1. Next, we present our experiments for regression with splines using synthetic data, the deforming cell imaging data and the cardiac imaging data in Sec. 4.2.

### 4.1. Assessment of shooting higher order curves

We first study the interpretation of the new states proposed in this paper: the *force* denoted by  $u(t)$  and the *jerk* denoted by  $p(t)$  for the evolution of curves in diffeomorphisms. It is informative to study the simplest case of regression first: the image matching problem between a set of two images, such that it solves the variational problem in Eq. (10) in the absence of any control locations, i.e.,  $C = 0$ . Since, there are only two points, the energy minimizing curve is a geodesic path but one that allows changes in velocity and in acceleration. The fourth derivative, jerk, however must conserve mass during its transport. Also note that if we only constrain the jerk state to be zero at all times, such that  $p(t) = 0$ , the diffeomorphism evolves in a ‘quadratic’ form. Along with the jerk, if we also constrain the force state to be zero at all times, such that  $u(t) = 0$ , the diffeomorphism evolves in the standard geodesic form.

To investigate this, we decompose our analysis into three simpler experiments. We first solve the problem of matching two images by constraining the force and the jerk states to be zero at all times and only estimate the initial momenta that describe the geodesic. We call this the *momenta only* matching (Figure 4 top row). Images on the left are the initial states: image, momenta, force and jerk at  $t = 0$ , and on the right are the final states: image, momenta, force and jerk at  $t = 1$ . This is simply solving the classical image matching problem using the geodesic shooting equations. In this case, the momenta evolution satisfies the standard EPDiff evolution. In the view of classical mechanics and particle motion, it is also analogous to describing the motion of a particle with constant velocity under the absence of any external force. Next, we solve the problem of matching the same images by constraining the starting momenta to be zero at  $t = 0$ , and the jerk states to be zero at all times, and estimate only the initial force. We call this the *force only* matching (Figure 4 middle row). This is analogous to the motion of a stationary particle under constant force such that it starts from zero velocity and then *constantly accelerates*. Finally, we solve the problem of matching the same two images by constraining the starting momenta and force to be zero at  $t = 0$ , and estimate only the initial jerk. We call this the *jerk only* matching (Figure 4 bottom row). This is analogous to describing the motion of a stationary particle with a *continuous* impulse such that it starts moving from zero velocity, and zero acceleration and then its *acceleration increases constantly* and its velocity increases in second order. The simple gradient descent with line search optimization converges for these three experiments. As expected, the momenta at the end point of the matching path for the jerk only matching are larger than that observed at the end point for force only matching. Also, the final force state for jerk only matching ends up being larger in magnitude than the force state for force only matching. Note that for all three curves the start point (the identity

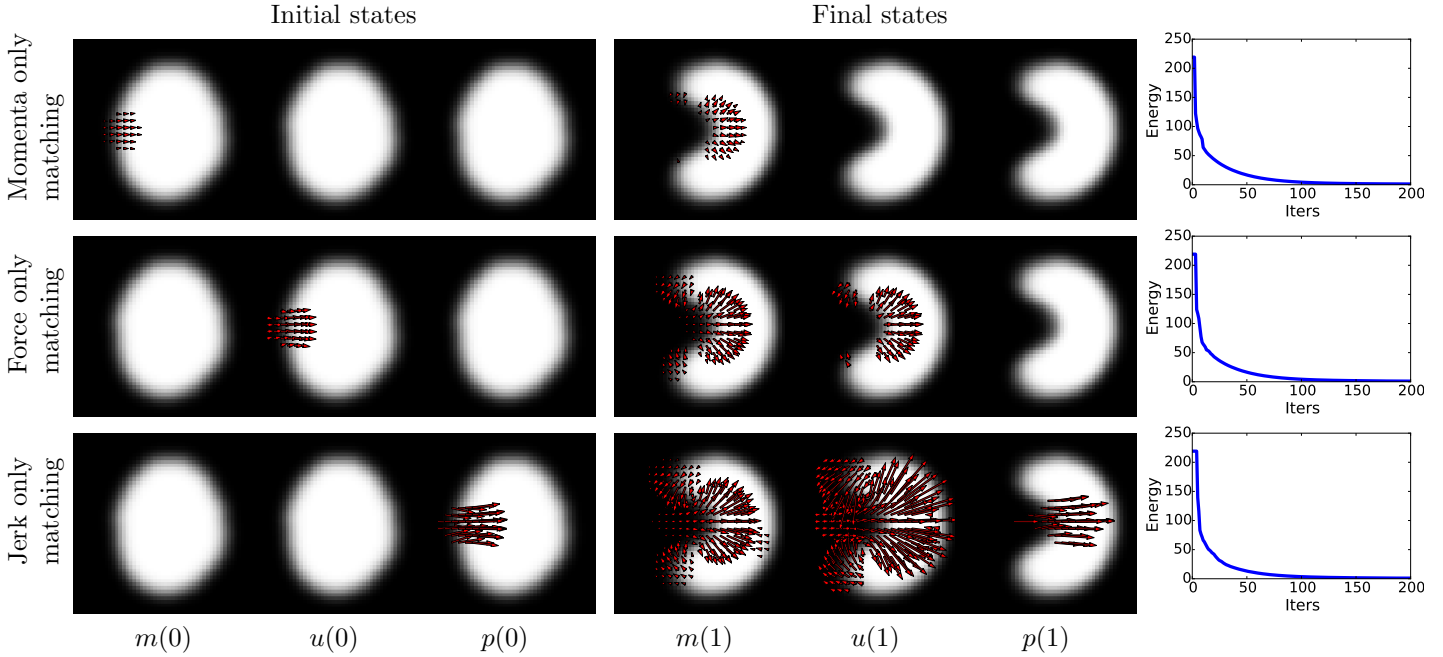


Figure 4: Higher order matching of two images. (a) Top row is the momenta only matching which is equivalent to the usual geodesic matching of images such that  $u(t) = 0$  and  $p(t) = 0$  along the matching path in diffeomorphisms. This matching only estimates initial momenta,  $m(0)$ . (b) Middle row is for force only matching to estimate  $u(0)$  such that  $p(t) = 0$  along the path but  $m(t)$  accelerates from zero initial condition, i.e.,  $m(0) = 0$ . This matching only estimates initial force,  $u(0)$ . (c) Bottom row correspond to jerk only matching such that both  $m(t)$  and  $u(t)$  accelerate from zero initial conditions, i.e.,  $m(0) = u(0) = 0$ . This matching only estimates initial jerk,  $u(0)$ . The length of arrows is proportional to their scales. On the right is the convergence of gradient descent with line search for the three optimization problems.

Table 1: Total squared error of fit

	Spline fit (one control)	Spline fit (no control)	Geodesic fit
Synthetic	40.50	139.77	162.39
Cell data	70.68	71.36	307.18
Cardiac data	962.75	975.15	1440.58

deformation) and the end points (the best matching deformation) are identical. The matching only differs in the order of motion along the path the curve traces in diffeomorphisms.

To further understand the Euclidean analogy of our proposed shooting equations in Eq. (9) we combine the three estimates and observe the resulting evolution. In particular, the above three experiments result in the three coefficients that follow similar scaling rules as the standard cubic curves,  $y = ax^3 + bx^2 + cx + d$ , where  $a$  is equivalent to  $p(0)$  estimated for jerk only matching,  $b$  to  $u(0)$  estimated for force only matching, and  $c$  to  $m(0)$  obtained for momenta only matching. We can now conveniently synthesize different parametric curves using the scaling of these coefficients. In Figure 5, we demonstrate the quadratic and the cubic polynomial curves in diffeomorphisms synthesized using these estimates obtained from matching.

To simulate a curve similar to the Euclidean quadratic,

$y = x^2 - x$ , we integrate Eq. (9) starting from the initial momenta of  $m(0)$  and initial force of  $-u(0)$ . We observe that the diffeomorphic path traces a quadratic path which is analogous to the motion of a particle under constant force. This is similar to the motion of a particle with a given initial velocity at  $t = 0$  but opposite force such that the particle decelerates initially, comes to a rest state exactly at  $t = 0.5$  and then accelerates to return back to the exact initial position at  $t = 1.0$  (Figure 5 (a)). The shooting equations are accurate such that the diffeomorphisms end at the identity transformation at  $t = 0$ . Another way to visualize this path in diffeomorphisms is to observe the motion of a pixel at the boundary of the image as it deforms from  $t = 0$  to  $t = 1$ . For this we display stacked-up 1-D cross sections (middle row of the image). This forms a 2D matrix displayed as a picture in Figure 5 (c), such that the rows are stacked up in increasing order of time from bottom to top. We notice that the pixels trace a quadratic

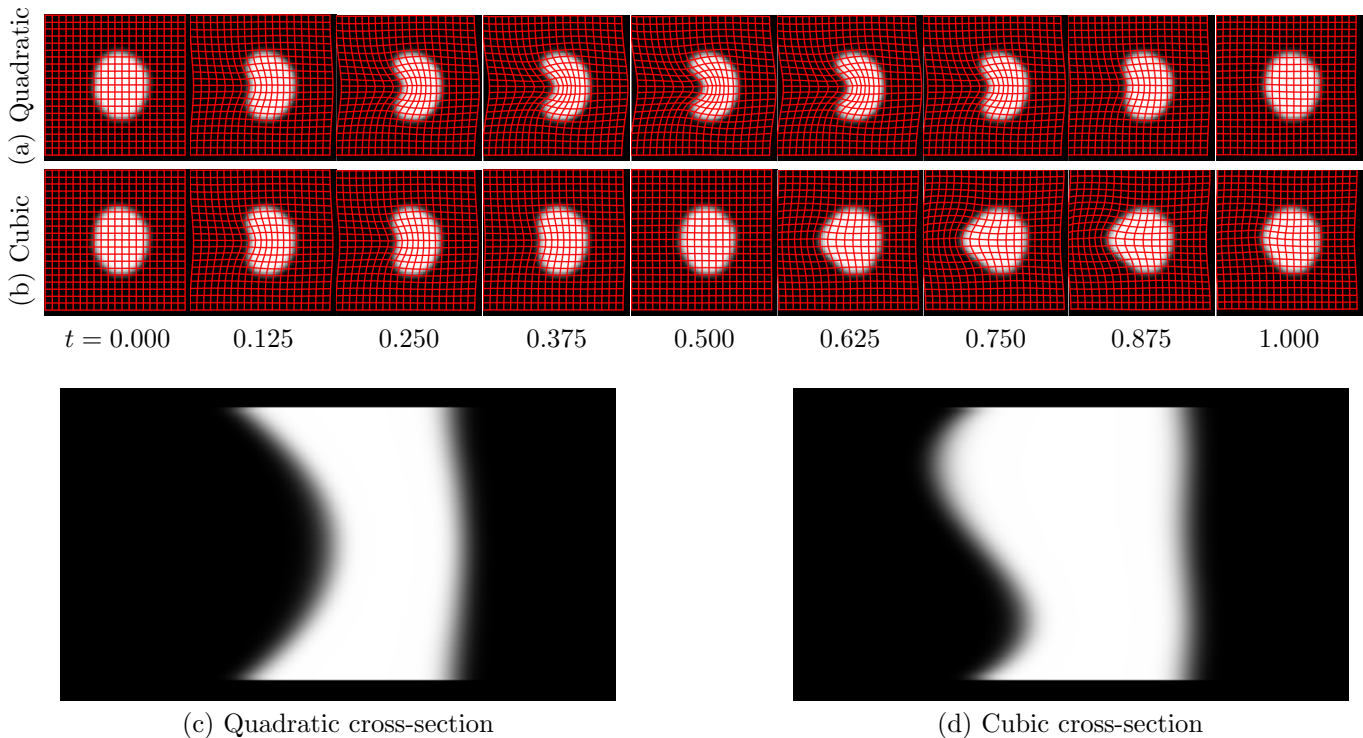


Figure 5: Quadratic and cubic polynomial paths in diffeomorphisms. Row (a) corresponds to the diffeomorphic path for quadratic evolution obtained by shooting along  $m(0)$  and  $-u(0)$ . This is analogous to the motion of a particle with finite initial velocity under constant decelerating force. Row (b) corresponds to the cubic evolution obtained by shooting along coefficients scaled as per the Bernoulli coefficients, i.e.,  $m(0)$ ,  $-3u(0)$  and  $2p(0)$ . Bottom row, (c) and (d) are the cross-section visualization of the generated images along these paths. The trajectory of boundary points follow quadratic and cubic paths, respectively.

Note that the base coefficients,  $m(0)$ ,  $u(0)$  and  $p(0)$  that were scaled, are the initial conditions estimated by solving the image registration problem with momenta only, force only and jerk only constraints, respectively, as per Figure 4.

curve as the image deforms along this path.

To demonstrate a cubic-like behavior, we scale the coefficients to generate a Bernoulli polynomial of degree 3 that takes the form,  $y = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x$ . For this, we integrate Eq. (9) starting with scaled initial conditions and use the initial momenta of  $\frac{1}{2}m(0)$ , initial force of  $-\frac{3}{2}u(0)$  and the initial jerk of  $p(0)$ . We observe that the diffeomorphic path traces a cubic path such that the shape first compresses in one direction and then expands and reaches back to a state that it started with and continues to expand until it finally shrinks back (Figure 5 (b)). This is analogous to the behavior a Euclidean Bernoulli polynomial of degree 3 follows with the exact same coefficients. Also, similar to the visualization of the quadratic curve, the cross sectional boundary pixel visualization of this curve results in a Bernoulli path that resembles a Bernoulli cubic (Figure 5 (d)).

#### 4.1.1. Quadratic and cubic regression

Next, to assess the strength of regression using these parametric shooting equations, we generate ( $N = 9$ ) sampled shapes to simulate *non-monotonic* quadratic and cubic like dynamics from  $t = 0$  to  $t = 1$  (Figure 6 (a) and (c), respectively) using the shooting methods described above. The corresponding quadratic and cubic regression fit are

shown in Figure 6 (b) and (d), respectively.

*A note on initialization.* A good initialization is necessary since the variational problem is non-convex. A possible strategy for the initialization is to first compute the initial momenta,  $\tilde{m}$ , for matching the first image (source image) with the next image in sequence that looks to be the most deformed image (target image) relative to the initial image. In the case of quadratic data in row (a), a good candidate for the target image could be the one at  $t = 0.5$  while for the cubic data in row (c), a possible candidate for the target image could be the one at  $t = 0.25$ . A good initialization for the quadratic regression gradient descent optimization problem could then use this momenta direction estimated only from the data to initialize both the initial quadratic states, for example, using approximately scaled  $\tilde{m}$ :  $m^k(0) = 2\tilde{m}$  and  $u^k(0) = -4\tilde{m}$  at first iteration for  $k = 1$ . Similarly, for the cubic regression variational problem, such an initialization could be used to also initialize all the initial states of momenta, force and jerk for the first iteration at  $k = 1$  of the gradient descent, for example, using approximately scaled:  $m^k(0) = 2\tilde{m}$ ,  $u^k(0) = -12\tilde{m}$  and  $p^k(0) = 24\tilde{m}$ , where  $\tilde{m}$  denotes the momenta corresponding to the geodesic matching problem of matching initial image with the image at  $t = 0.25$ . Note these fac-



tors correspond to quadratic and Bernoulli cubics respectively along with additional scalings to compensate for the shorter length of the path for which  $\hat{m}$  is computed.

We emphasize that the above strategies use only the data to decide initial conditions of a curve that could be close to the possible minimizing least square polynomial curve. This provides a practical way to find good initialization for the fitting problem for any given data set. We notice that using such initializations the optimization converges faster for both the quadratic and the cubic regression variational problems (Figure 6 (f) and (h)).

For the quadratic regression, we notice that the estimated fit captures the trend and results in a smoothly shrinking followed by expanding grids along the regression path to closely match the data. The boundary pixel also traces a quadratic curve when the image deforms (Figure 6 (e)). Similarly, for the cubic regression, the estimated fit captures the two inflections of the motion to best fit the data. The boundary pixel also traces a quadratic curve when the image deforms (Figure 6 (e)).

#### 4.2. Assessment of spline regression.

In this section, we investigate the performance of spline regression on synthetic and real data. For the synthetic data, we generate the data such that a cubic-like dynamic alone is not sufficient to explain the trends in shape changes and therefore necessitates adding a control for the spline fit. For the real data, we perform spline regression experiments using cell and cardiac images.

##### 4.2.1. Synthetic data.

To assess the strength of spline regression for non-geodesic image data, we create a synthetic sequence of  $N = 13$  shapes to simulate non-monotonic changes with *more than two* inflection points from  $t = 0$  to  $t = 1$  (Figure 7, (a)). The synthetic shape first shrinks and then expands till  $t = 0.5$ , and then again shrinks and finally expands back again till it reaches the end point, at  $t = 1.0$ . Using such data, we attempt to simulate dynamics such that a cubic alone would not be sufficient to trace through the inflection. For optimization, we follow a similar strategy we discussed in Section 4.1.1 for cubic regression to initialize momenta, force and jerk at  $t = 0$  but just switch the sign on the jerk state,  $p$  at the control location to add another inflection.

We report a quantitative comparison of the three fits in the first row of Table 1. The reported error of fit corresponds to  $L^2$  image residual as per the data-likelihood in Eq. (10). We observe that adding a single spline control at the mid point results in the best fit that summarizes the smooth dynamics of change (Figure 7 (b)). The estimated diffeomorphism successfully captures the two cycles of trends in the shrinking shape, followed by expansion (Figure 7 (c)). Without adding any control, the resulting spline trend, even though a cubic, fails to capture the dynamics and fails to recover the inflection points in the

rate of shape change. Finally, being the most inflexible, geodesic regression performs worst, and barely captures any real spatio-temporal trend (Figure 7 (d)). The visualization of the motion of boundary pixels also confirms (last column) the flexible diffeomorphism fit obtained for the spline regression with one control.

##### 4.2.2. Deforming cell data.

The cell time-sequence data corresponds to 11 snapshots at equal intervals a deforming cell imaged using an atomic force microscope. We preprocessed the images using total variation denoising (Chambolle, 2004) to preserve edges. The images depict a trend in which the shape of the cell deforms such that its left boundary first bends inward and then resumes back to its original shape (Figure 8 (a)).

A visual assessment of the cell images suggests that this data should have an inflection in the dynamics of shape changes. The regression fit using one control and no control result in a very similar fit (Figure 8 (b) and (c)). However, the geodesic fit only results in a monotonous compression of the cell and fails to capture the expansion in the last half of the dynamics (Figure 8 (d)). We also notice in Table 1 that the spline regressions with one control and without a control result in comparable  $L^2$  error of fit as per the data likelihood. The geodesic fit, however, clearly performs worst in terms of the residual error of fit.

##### 4.2.3. Cardiac data.

The cardiac time-sequence data corresponds to 20 snapshots at equal intervals of the beating heart of a normal individual with age=63 years (Subject Id: SCD0003701). We cropped all the axial images to a common rectangular region around the heart followed by histogram matching to align intensities of all the timepoints to the image at  $t = 0$ . Figure 9 shows the original scans (first row) and the result of regression models (second to fourth row). We only display half of the timepoints of the ones actually used to fit the model. Similar to the synthetic data, we observe that the original data exhibit a non-geodesic and non-monotonic trend in changing shape of the beating heart. The comparison in terms of the error of fit for all models suggests that both spline curves perform better than the geodesic. Although visually the fits look similar for splines, we obtained a marginal improvement of fit for the spline curve with single control when compared to the spline curve without any control (Table 1). The geodesic fit again performs worst out of the three models. The dynamics of the beating heart for these models are best seen in the multimedia file in the supplementary material. The CPU/GPU implementations of the shooting splines for 2D and 3D image sequences is at <https://nikhilsingh@bitbucket.org/nikhilsingh/diffeosplines.git>.

## 5. Discussion

In this article, we developed a theory for higher order curves that generalizes the notion of parametric curves



such as the quadratic, the cubic and the piecewise cubics to the manifold of diffeomorphisms. We provided a principled way to define curves with nonzero acceleration and nonzero jerk, which is the natural next step of extension to geodesic-based methods developed during the last decade for computational anatomy in the large deformation diffeomorphic image analysis framework. We took a variational approach that is governed by an underlying energy formulation, which respects the nonflat geometry of diffeomorphisms. Such an approach of minimal energy curve estimation also provides a physical analogy with particle motion under a varying force field.

As a consequence, the initial conditions of our variational quadratics and cubics are interpretable similar to the initial conditions of their corresponding Euclidean parametric counterparts. To validate this, we demonstrated that evolving the curves according to scaled initial conditions also results in the same behavior as a scalar Euclidean parametric curve would exhibit under these scalings. We tested this for different scalings and presented the results for quadratic parabolic scalings and for Bernoulli cubic scalings.

Our proposed system of evolution equations for higher order curves can be used for regression. The benefit of using these forward shooting equations in an optimal control setting is that the solution to the resulting regression problem is given only in terms of a few initial conditions. We emphasize that in all our experiments, the full diffeomorphic paths and the evolution of all states along the estimated curves are completely parameterized by very few parameters (four for spline fits with one control and three for spline fits without any control) that are independent of the study size.

### 5.1. Open questions and extensions.

Regression models are expected to fit better with increasing number of control points, which will necessitate model selection methods. Due to the limitations of our current optimization method, i.e., gradient descent, our spline estimation experiments used at most one control point only. A possible future work would be to explore better optimization strategies and to develop second order methods utilizing limited memory for the optimization of splines (Byrd et al., 1995). We expect improved convergence for example by using a quasi-Newton methods such as LBFGS.

We developed the regression problem for a fixed initial image. Adding template estimation would add another parameter to the estimation problem. With better optimization strategies, it should also be possible to develop an alternate optimization algorithm for template estimation (Singh et al., 2013b).

The position of control points is also a modeling choice. The uniform placement of control points on the axis of the regressor variable is convenient and facilitates model comparisons and interpretability in medical imaging population studies. However, it remains an open problem to

also optimize for the locations of control points. The possible challenges will include investigating the differentiability of the energy functional in Eq. (10) with respect to the location of the jump in the jerk state.

Another aspect could be investigating possibilities of combining our model on diffeomorphisms with the higher order models on shapes (Gay-Balmaz et al., 2012; Trounev and Vialard, 2012).

One of the most critical contribution in this research is the ability to trace a path in diffeomorphisms with nonzero acceleration. This could be immensely useful for medical studies of growth or decline where the rate of change gets affected and the emphasis is on accelerated tissue growth or decline. One future possible application for the quadratic models could be to study differences in aging of individuals with or without dementia and investigating the ages and local region in the brain exhibiting the most accelerated atrophy. This would add second order information on tissue atrophy to the information currently being obtained using contemporary first order geodesic regression methods and deformation based morphometry analysis.

Another use of higher order models is for the recently proposed longitudinal models for diffeomorphisms such as hierarchical geodesic models (HGM) (Singh et al., 2013a). Even though for brain studies, a geodesic-like trend is expected to be a good approximation of changes in the brain for a single individual when the measurements are taken within a span of five years, the geodesic assumption on the average group trend of the entire population data for staggered designs may not be the best modeling choice. Thus, our proposed higher order curves would provide a better model for the longitudinal summary of a group spanning a wide range of ages from, say, 60 to 90 years.

Our method of shooting splines in diffeomorphisms lays a foundation to model flexible dynamics of shape changes seen in time series of medical images, and also opens the possibility to model periodic data by adding periodicity constraints.

## 6. Acknowledgments

The authors would like to thank Jacob Hinkle for his valuable input during the development of the initial ideas of the acceleration controlled curves in Riemannian manifolds. We would also like to thank Richard Superfine and his students for providing the cell images using the atomic force microscope for testing our regression models. This research is supported by the grants, NSF EECS-1148870, NSF EECS-0925875, and NIH R01-MH091645.

## Appendix A. Euler-Lagrange for relaxation problem

We determine the Euler-Lagrange equation in the context of the regression problem. We write the problem of regression for the fixed initial image  $I^0$  and use  $h(t)$  to

denote the inverse deformation  $g^{-1}(t)$ . For readability we drop the argument,  $t$ , for all the time dependent states,  $h(t)$ ,  $m(t)$  and  $u(t)$ . We denote the Jacobian operator by  $D(\cdot)$  and the pixel-wise Kronecker product of two vectors by  $\otimes$ . The constrained energy minimization for the relaxation problem of minimizing the elastic energy for the force controlled curve takes the form,

$$\begin{aligned} \tilde{E}(h, m, u, \lambda_h, \lambda_m, \lambda_u) &= \frac{1}{2\sigma_I^2} \sum_{i=1}^N d^2(I^0 \circ h(t_i), J_i) \\ &+ \frac{1}{2} \int_0^1 \langle u, K^u u \rangle_{L^2} dt + \int_0^1 \langle \lambda_h, \dot{h} + (Dh)K^m m \rangle_{L^2} dt \\ &+ \int_0^1 \langle \lambda_m, \dot{m} - u + \text{ad}_{K^m m}^* m \rangle_{L^2} dt, \end{aligned}$$

where  $K^u$  and  $K^m$  correspond to the different time independent metric kernels for  $u$  and  $m$ , respectively, and  $d(\cdot, \cdot)$  is the metric on images. Note that the metric on images can also be the metric on  $G$ , here we will keep this simpler, and derive our results for the  $L^2$  metric.

#### Appendix A.1. Variations

In what follows, unless specified otherwise, all inner products correspond to the  $L^2$  pairing. Computing variations of  $\tilde{E}$  with respect to adjoint variables will give us the dynamic constraints back. Computing variations of  $\tilde{E}$  with respect to all state variables gives,

$$\begin{aligned} \delta_h \tilde{E} &= \langle \frac{1}{\sigma_I^2} \nabla_{h(t_i)} d_h(I^0 \circ h(t_i), J_i), \delta h(t_i) \rangle \\ &+ \langle \lambda_h(1), \delta h(1) \rangle - \langle \lambda_h(0), \delta h(0) \rangle \\ &+ \int_0^1 \langle -\dot{\lambda}_h - \nabla \cdot (\lambda_h \otimes K^m m), \delta h \rangle dt \\ \delta_m \tilde{E} &= \langle \lambda_m(1), \delta m(1) \rangle - \langle \lambda_m(0), \delta m(0) \rangle \\ &+ \int_0^1 \langle -\dot{\lambda}_m + K^m((Dh)^\top \lambda_h) + \text{ad}_{K^m m} \lambda_m \\ &- K^m \text{ad}_{\lambda_m}^* m, \delta m \rangle dt \\ \delta_u \tilde{E} &= \int_0^1 \langle K^u u - \lambda_m, \delta u \rangle dt \end{aligned}$$

At the optimum the above must vanish. This results in the following adjoint system,

$$\begin{aligned} \dot{\lambda}_h + \nabla \cdot (\lambda_h \otimes K^m m) &= 0 \\ \dot{\lambda}_m - K^m((Dh)^\top \lambda_h) - \text{ad}_{K^m m} \lambda_m + K^m \text{ad}_{\lambda_m}^* m &= 0 \\ K^u u - \lambda_m &= 0 \end{aligned}$$

Notice that, similar to the Euclidean case where the adjoint variable corresponding to the second state,  $x_2$ , turned out to be equal to the third state  $x_3$ , here  $\lambda_m = K^u u$ .

$$\begin{aligned} \dot{\lambda}_h + \nabla \cdot (\lambda_h \otimes K^m m) &= 0 \\ K^u \dot{u} - K^m((Dh)^\top \lambda_h) - \text{ad}_{K^m m} K^u u + K^m \text{ad}_{K^u u}^* m &= 0 \end{aligned}$$

Again, analogous to the Euclidean state, where the adjoint variable corresponding to the first state,  $x_1$ , was renamed as the primal state,  $x_4$ , for the shooting evolution, we rename  $\lambda_h$  to be  $p$  to describe the evolution of  $u$  for shooting cubics to give:

$$\begin{aligned} \dot{p} + \nabla \cdot (p \otimes K^m m) &= 0 \\ K^u \dot{u} - K^m((Dh)^\top p) - \text{ad}_{K^m m} K^u u + K^m \text{ad}_{K^u u}^* m &= 0 \end{aligned}$$

Thus the forward evolution of the cubic system is determined by the following set of four PDE's:

$$\begin{aligned} \dot{h} + (Dh)K^m m &= 0 \\ \dot{m} - u + \text{ad}_{K^m}^* m &= 0 \\ \dot{u} - K^{-1u} K^m((Dh)^\top p) - K^{-1u} \text{ad}_{K^m m} K^u u \\ &+ K^{-1u} K^m \text{ad}_{K^u u}^* m = 0 \\ \dot{p} + \nabla \cdot (p \otimes K^m m) &= 0 \end{aligned}$$

For the case when the metric is same, i.e.,  $K^m = K^u = K$ , the above equations further simplify to,

$$\begin{aligned} \dot{h} + (Dh)K m &= 0 \\ \dot{m} - u + \text{ad}_K^* m &= 0 \\ \dot{u} - (Dh)^\top p - K^{-1} \text{ad}_{K m} K u + \text{ad}_{K u}^* m &= 0 \\ \dot{p} + \nabla \cdot (p \otimes K m) &= 0 \end{aligned}$$

Next we discuss how to use these shooting equations for the cubic evolution to define the cubic regression problem.

#### Appendix B. Euler-Lagrange for shooting problem

Let,  $J_i$ , for  $i = 1 \dots N$ , denote  $N$  measured images at timepoints,  $t_i \in (0, 1)$ . Let us assume there are no measurements at the end points, i.e., neither at  $t = 0$ , nor at  $t = 1$ . Let  $t_c \in (0, 1)$ , for  $c = 1 \dots C$ , denote  $C$  data-independent fixed control locations. The control locations also implicitly define  $C + 1$  intervals or partitions in  $(0, 1)$ . Let us denote these intervals as  $\mathcal{I}_c$ , for  $c = 1 \dots (C + 1)$ .

For regression on such a data configuration, the least-squares energy takes the form,

$$E(h) = \frac{1}{2\sigma^2} \sum_{c=1}^{C+1} \sum_{i \in \mathcal{I}_c} d^2(I^0 \circ h(t_i), J_i),$$

subject to the dynamic constraints.

We first write the constrained energy minimization problem as,

$$\underset{I}{\text{minimize}} \quad E(I) \tag{B.1}$$

$$\left. \begin{aligned} \dot{h} + (Dh)K m &= 0 \\ \dot{m} - u + \text{ad}_K^* m &= 0 \\ \dot{u} - (Dh)^\top p - K^{-1} \text{ad}_{K m} K u + \text{ad}_{K u}^* m &= 0 \\ \dot{p} + \nabla \cdot (p \otimes K m) &= 0 \end{aligned} \right\} \text{for each interval, } \mathcal{I}_c$$

subject to continuity of  $h$ ,  $m$ , and  $u$  at  $C$  joins.

The unconstrained Lagrangian for spline regression for a fixed initial image  $I^0$  takes the form:

$$\begin{aligned} \tilde{E}(h, m, u, p, p^{t_c}, \lambda_h, \lambda_m, \lambda_u, \lambda_p) = & \\ & \frac{1}{2\sigma_I^2} \sum_{c=1}^{C+1} \sum_{i \in \mathcal{I}_c} d^2(I^0 \circ h(t_i), J_i) + \frac{1}{2\sigma_u^2} \int_0^1 \langle u, Ku \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{h_c}, \dot{h}_c + (Dh_c)Km_c \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{m_c}, \dot{m}_c - u_c + \text{ad}_{Km_c}^* m_c \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{u_c}, \dot{u}_c - (Dh_c)^\top p_c - K^{-1} \text{ad}_{Km_c} Ku_c \\ & + \text{ad}_{Ku_c}^* m_c \rangle_{L^2} dt \\ & + \sum_{c=1}^{C+1} \int_0^1 \langle \lambda_{p_c}, \dot{p}_c + \nabla \cdot (p_c \otimes Km_c) \rangle_{L^2} dt, \text{ and} \\ & \text{subject to continuity of } h, m, u, \text{ and } p \text{ at } C \text{ joins.} \end{aligned}$$

The second term is the relaxation energy term,  $\|u\|_{\mathfrak{g}}^2$ . This term acts as a regularizer on the force along the full path in diffeomorphisms.

### Appendix B.1. Variations

We discuss each piece of this optimization separately and combine the result in the end of this section. It is convenient to first derive it for the case with no controls, i.e.,  $C = 0$ . The unconstrained Lagrangian takes the form,

$$\begin{aligned} \tilde{E}(h, m, u, p, \lambda_h, \lambda_m, \lambda_u, \lambda_p) = & \frac{1}{2\sigma_I^2} \sum_{i=1}^N d^2(I^0 \circ h(t_i), J_i) \\ & + \frac{1}{2} \int_0^1 \langle u, Ku \rangle_{L^2} dt + \int_0^1 \langle \lambda_h, \dot{h} + (Dh)Km \rangle_{L^2} dt \\ & + \int_0^1 \langle \lambda_m, \dot{m} - u + \text{ad}_{Km}^* m \rangle_{L^2} dt \\ & + \int_0^1 \langle \lambda_u, \dot{u} - (Dh)^\top p - K^{-1} \text{ad}_{Km} Ku + \text{ad}_{Ku}^* m \rangle_{L^2} dt \\ & + \int_0^1 \langle \lambda_p, \dot{p} + \nabla \cdot (p \otimes Km) \rangle_{L^2} dt \end{aligned}$$

The variations with respect to dual adjoint variables give the dynamic constraints back. We write the variations

with respect to all the primals as:

$$\begin{aligned} \delta_h \tilde{E} = & \left\langle \frac{1}{\sigma_I^2} \nabla_{h(t_i)} d_h(I^0 \circ h(t_i), J_i), \delta_h(t_i) \right\rangle \\ & + \langle \lambda_h(1), \delta h(1) \rangle - \langle \lambda_h(0), \delta h(0) \rangle \\ & + \int_0^1 \langle -\dot{\lambda}_h + \nabla \cdot (p \otimes \lambda_u - \lambda_h \otimes Km), \delta h \rangle dt \\ \delta_m \tilde{E} = & \langle \lambda_m(1), \delta m(1) \rangle - \langle \lambda_m(0), \delta m(0) \rangle \\ & + \int_0^1 \langle -\dot{\lambda}_m + K((Dh)^\top \lambda_h) + \text{ad}_{Km} \lambda_m - K \text{ad}_{\lambda_m}^* m \\ & + K \text{ad}_{Ku}^* K^{-1} \lambda_u + \text{ad}_{Ku} \lambda_u - K((D\lambda_p)^\top p), \delta m \rangle dt \\ \delta_u \tilde{E} = & \langle \lambda_u(1), \delta u(1) \rangle - \langle \lambda_u(0), \delta u(0) \rangle \\ & + \int_0^1 \langle -\dot{\lambda}_u + \frac{1}{\sigma_u^2} Ku - \lambda_m \\ & - K \text{ad}_{Km}^* K^{-1} \lambda_u - K \text{ad}_{\lambda_u}^* m, \delta u \rangle dt \\ \delta_p \tilde{E} = & \langle \lambda_p(1), \delta p(1) \rangle - \langle \lambda_p(0), \delta p(0) \rangle \\ & + \int_0^1 \langle -\dot{\lambda}_p - (Dh)\lambda_u - (D\lambda_p)Km, \delta p \rangle dt \end{aligned}$$

At the optimum, the above must vanish. This results in the following adjoint system:

$$\begin{aligned} \dot{\lambda}_h - \nabla \cdot (p \otimes \lambda_u - \lambda_h \otimes Km) &= 0 \\ \dot{\lambda}_m - \text{ad}_{Km} \lambda_m + K \text{ad}_{\lambda_m}^* m - K \text{ad}_{Ku}^* K^{-1} \lambda_u \\ &\quad - \text{ad}_{Ku} \lambda_u + K((D\lambda_p)^\top p) - K((Dh)^\top \lambda_h) = 0 \\ \dot{\lambda}_u - \frac{1}{\sigma_u^2} Ku + \lambda_m + K \text{ad}_{Km}^* K^{-1} \lambda_u + K \text{ad}_{\lambda_u}^* m &= 0 \\ \dot{\lambda}_p + (Dh)\lambda_u + (D\lambda_p)Km &= 0 \end{aligned}$$

The boundary conditions are,

$$\begin{aligned} \delta_{h(0)} E &= -\lambda_h(0) & \delta_{h(1)} E &= \lambda_h(1) \\ \delta_{m(0)} E &= -\lambda_m(0) & \delta_{m(1)} E &= \lambda_m(1) \\ \delta_{u(0)} E &= -\lambda_u(0) & \delta_{u(1)} E &= \lambda_u(1) \\ \delta_{p(0)} E &= -\lambda_p(0) & \delta_{p(1)} E &= \lambda_p(1) \end{aligned}$$

Notice that we can also rewrite the adjoint system by removing the  $K$  operator and using their duals,  $v = Km$  and  $f = Ku$ , and the conjugate operator,  $\text{ad}_X^\dagger$  is  $\text{ad}_X^\dagger = K \star \text{ad}_X^* L$ , such that:

$$\begin{aligned} \dot{\lambda}_h - \nabla \cdot (p \otimes \lambda_u - \lambda_h \otimes v) &= 0 \\ \dot{\lambda}_m - \text{ad}_v \lambda_m + \text{ad}_{\lambda_m}^\dagger v - \text{ad}_f^\dagger \lambda_u \\ &\quad - \text{ad}_f \lambda_u + K((D\lambda_p)^\top p) - K((Dh)^\top \lambda_h) = 0 \\ \dot{\lambda}_u - f + \lambda_m + \text{ad}_v^\dagger \lambda_u + \text{ad}_{\lambda_u}^\dagger v &= 0 \\ \dot{\lambda}_p + (Dh)\lambda_u + (D\lambda_p)v &= 0 \end{aligned}$$

Furthermore, the adjoint in terms of duals for  $\lambda_f = K^{-1}\lambda_u$  and  $\lambda_v = K^{-1}\lambda_m$  is:

$$\dot{\lambda}_f - u + \lambda_v + \text{ad}_v^* \lambda_f + \text{ad}_{\lambda_u}^* m = 0$$

*Appendix B.1.1. For the data fit constraints*

The gradient of the data match term with respect to  $h$  is:

$$\begin{aligned} & \frac{1}{2\sigma_I^2} \langle \nabla_{h(t_i)} d_h(I^0 \circ h(t_i), J_i), \delta_h(t_i) \rangle \\ &= \frac{1}{\sigma_I^2} \langle (I^0 \circ h(t_i) - J_i), \nabla_{h(t_i)} (I^0 \circ h(t_i)) \delta_h(t_i) \rangle \\ &= \frac{1}{\sigma_I^2} \langle (I^0 \circ h(t_i) - J_i), \nabla I^{t_i} \delta_h(t_i) \rangle \end{aligned}$$

Note there is no application of the chain rule in the last step above for the term,  $\nabla_{h(t_i)} (I^0 \circ h(t_i))$  because the gradient is computed with respect to  $h(t_i)$ . This is equivalent to the substitution of  $h(t_i)$  as  $y$ , which then is equivalent to taking the derivative of  $I(y)$  with respect to  $y$ . Thus, this gradient of the data fit constraints with respect to  $h$ , results in jumps during the backward integration of the adjoint variable,  $\lambda_h$ , as,

$$\lambda_h(t_i^+) - \lambda_h(t_i^-) = \frac{1}{\sigma_I^2} (I^0 \circ h(t_i) - J_i) \nabla I^{t_i}.$$

*Appendix B.1.2. For the continuity constraints*

We derive the variations of the energy functional for the joins and study the continuity of the adjoint system in the interval  $(0, 1)$ . For this analysis, we introduce again the subscripts  $c$  to denote the intervals. We first rewrite the above functional here as,

$$\begin{aligned} \tilde{E} &= \text{fit of the data within each interval, } \mathcal{I}_c \\ &+ \text{dynamics within each interval, } \mathcal{I}_c \\ &+ \sum_{c=1}^C \langle \alpha_{hc}, h_c(t_c) - h_{c+1}(t_c) \rangle_{L^2} \\ &+ \sum_{c=1}^C \langle \alpha_{mc}, m_c(t_c) - m_{c+1}(t_c) \rangle_{L^2} \\ &+ \sum_{c=1}^C \langle \alpha_{uc}, u_c(t_c) - u_{c+1}(t_c) \rangle_{L^2} \end{aligned}$$

Notice that the Lagrangian adjoint variables,  $\alpha_{hc}$ ,  $\alpha_{mc}$  and  $\alpha_{uc}$ , do not vary with time. We write the variations with respect to all the primals at joins as:

$$\begin{aligned} \delta_{hc(t_c)} E &= \lambda_{hc}(t_c) + \alpha_{mc} \\ \delta_{hc+1(t_c)} E &= -\lambda_{hc+1}(t_c) - \alpha_{mc} \\ \delta_{mc(t_c)} E &= \lambda_{mc}(t_c) + \alpha_{mc} \\ \delta_{mc+1(t_c)} E &= -\lambda_{mc+1}(t_c) - \alpha_{mc} \\ \delta_{uc(t_c)} E &= \lambda_{uc}(t_c) + \alpha_{uc} \\ \delta_{uc+1(t_c)} E &= -\lambda_{uc}(t_c) - \alpha_{uc} \end{aligned}$$

Equating all the above variations to zero and algebraically eliminating variables, we get,

$$\begin{aligned} \lambda_{hc}(t_c) &= \lambda_{hc+1}(t_c) \\ \lambda_{mc}(t_c) &= \lambda_{mc+1}(t_c) \\ \lambda_{uc}(t_c) &= \lambda_{uc+1}(t_c) \end{aligned}$$

This proves that  $\lambda_{hc}$ ,  $\lambda_{mc}$  and  $\lambda_{uc}$  are continuous at the boundaries of the control point locations.

*Appendix B.1.3. Gradient computation using backward integration*

We summarize the optimization here. The gradients with respect to the initial conditions are,

$$\begin{aligned} \delta_{m_1(0)} E &= -\lambda_{m1}(0) \\ \delta_{u_1(0)} E &= -\lambda_{u1}(0) \\ \delta_{p_1(0)} E &= -\lambda_{p1}(0) \end{aligned}$$

We compute the gradients by integrating the adjoint system of equations within each interval backward in time,

$$\begin{aligned} \dot{\lambda}_{hc} - \nabla \cdot (p_c \otimes \lambda_{uc} - \lambda_{hc} \otimes K m_c) &= 0 \\ \dot{\lambda}_{mc} - \text{ad}_{K m_c} \lambda_{mc} + K \text{ad}_{\lambda_{mc}}^* m_c - K \text{ad}_{K u_c}^* K^{-1} \lambda_{uc} \\ &- \text{ad}_{K u_c} \lambda_{uc} + K((D\lambda_{pc})^\top p_c) - K((Dh_c)^\top \lambda_{hc}) = 0 \\ \dot{\lambda}_{uc} - \frac{1}{\sigma_p^2} K u_c + \lambda_{mc} + K \text{ad}_{K m_c}^* K^{-1} \lambda_{uc} \\ &+ K \text{ad}_{\lambda_{uc}}^* m_c = 0 \\ \dot{\lambda}_{pc} + (Dh_c) \lambda_{uc} + (D\lambda_{pc}) K m_c &= 0 \end{aligned}$$

All variables start from zero as their initial conditions for this backward integration. We add jumps in  $\lambda_h$  as  $\lambda_h(t_i^+) - \lambda_h(t_i^-) = \frac{1}{\sigma_I^2} (I^0 \circ h(t_i) - J_i) \nabla I^{t_i}$  at measurements,  $t = t_i$ . We ensure the continuity of  $\lambda_{hc}$ ,  $\lambda_{mc}$ , and  $\lambda_{uc}$  at the joins. However,  $\lambda_{pc}$  starts from zero at every join.

The accumulated  $\lambda_{pc+1}$  is used to update the  $p_c(t_k)$  as per the gradient,

$$\delta_{p_{c+1}(t_c)} E = -\lambda_{pc+1}(t_c)$$

Note this is the ‘data independent’ control that we motivated our formulation with. This determines the initial condition of the forward system for each interval and needs to be estimated numerically.

- Amit, Y., Grenander, U., Piccioni, M., 1991. Structural image restoration through deformable templates. *Journal of the American Statistical Association* 86, 376–387.
- Arnol’d, V.I., 1966. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l’hydrodynamique. *Ann. Inst. Fourier* 16, 319–361.
- Beg, M.F., Miller, M.I., Trounev, A., Younes, L., 2005. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision* 61, 139–157.
- Bruveris, M., Gay-Balmaz, F., Holm, D., Ratiu, T., 2011. The momentum map representation of images. *Journal of Nonlinear Science* 21, 115–150. doi:[10.1007/s00332-010-9079-5](https://doi.org/10.1007/s00332-010-9079-5).
- Byrd, R., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16, 1190–1208. doi:[10.1137/0916069](https://doi.org/10.1137/0916069).
- Camarinha, M., Leite, F.S., Crouch, P., 1995. Splines of class  $ck$  on non-euclidean spaces. *IMA Journal of Mathematical Control and Information* 12, 399–410.
- Chambolle, A., 2004. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision* 20, 89–97. doi:[10.1023/B:JMIV.0000011325.36760.1e](https://doi.org/10.1023/B:JMIV.0000011325.36760.1e).
- Davis, B.C., 2008. *Medical Image Analysis via Fréchet Means of Diffeomorphisms*. ProQuest.

- Davis, B.C., Fletcher, P.T., Bullitt, E., Joshi, S., 2010. Population shape regression from random design data. *International journal of computer vision* 90, 255–266.
- Gay-Balmaz, F., Holm, D.D., Meier, D.M., Ratiu, T.S., Vialard, F.X., 2012. Invariant higher-order variational problems II. *J. Nonlinear Science* 22, 553–597.
- Grenander, U., Miller, M.I., 1998. Computational anatomy: an emerging discipline. *Q. Appl. Math.* LVI, 617–694.
- Hinkle, J., Fletcher, P., Joshi, S., 2014. Intrinsic polynomials for regression on Riemannian manifolds. *Journal of Mathematical Imaging and Vision* , 1–21doi:10.1007/s10851-013-0489-5.
- Miller, M., Banerjee, A., Christensen, G., Joshi, S., Khaneja, N., Grenander, U., Matejic, L., 1997. Statistical methods in computational anatomy. *Statistical Methods in Medical Research* 6, 267–299.
- Miller, M.I., 2004. Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. *NeuroImage* 23, 19–33.
- Niethammer, M., Huang, Y., Vialard, F.X., 2011. Geodesic regression for image time-series, in: Fichtinger, G., Martel, A., Peters, T. (Eds.), *MICCAI 2011*. Springer Berlin Heidelberg. volume 6892 of *LNCIS*, pp. 655–662. doi:10.1007/978-3-642-23629-7\_80.
- Noakes, L., Heinzinger, G., Paden, B., 1989. Cubic splines on curved spaces. *IMA Journal of Mathematical Control and Information* 6, 465–473.
- Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A., Wright, G., 2009. Evaluation framework for algorithms segmenting short axis cardiac MRI. *MIDAS* .
- Singh, N., Fletcher, P., Preston, J.S., King, R.D., Marron, J.S., Weiner, M.W., Joshi, S., 2014. Quantifying anatomical shape variations in neurological disorders. *Medical Image Analysis* 18, 616–633. doi:10.1016/j.media.2014.01.001.
- Singh, N., Hinkle, J., Joshi, S., Fletcher, P., 2013a. A hierarchical geodesic model for diffeomorphic longitudinal shape analysis, in: *IPMI*. volume 7917, pp. 560–571.
- Singh, N., Hinkle, J., Joshi, S., Fletcher, P., 2013b. A vector momenta formulation of diffeomorphisms for improved geodesic regression and atlas construction, in: *ISBI*, pp. 1219–1222. doi:10.1109/ISBI.2013.6556700.
- Singh, N., Niethammer, M., 2014. Splines for diffeomorphic image regression, in: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (Eds.), *Medical Image Computing and Computer-Assisted Intervention MICCAI 2014*. Springer International Publishing. volume 8674 of *Lecture Notes in Computer Science*, pp. 121–129. URL: [http://dx.doi.org/10.1007/978-3-319-10470-6\\_16](http://dx.doi.org/10.1007/978-3-319-10470-6_16), doi:10.1007/978-3-319-10470-6\_16.
- Thompson, D.W., et al., 1942. On growth and form. *On Growth and Form* .
- Thompson, P.M., Toga, A.W., 2002. A framework for computational anatomy.
- Trouvé, A., Vialard, F.X., 2012. Shape splines and stochastic shape evolutions: A second order point of view. *Quarterly of Applied Mathematics* 70, 219–251.
- Vialard, F.X., Risser, L., Rueckert, D., Cotter, C., 2011. Diffeomorphic 3D image registration via geodesic shooting using an efficient adjoint calculation. *International Journal of Computer Vision* , 1–1310.1007/s11263-011-0481-8.
- Younes, L., 2010. *Shapes and Diffeomorphisms*. volume 171. Springer Berlin.
- Younes, L., Arrate, F., Miller, M.I., 2009. Evolution equations in computational anatomy. *NeuroImage* 45, S40–S50.



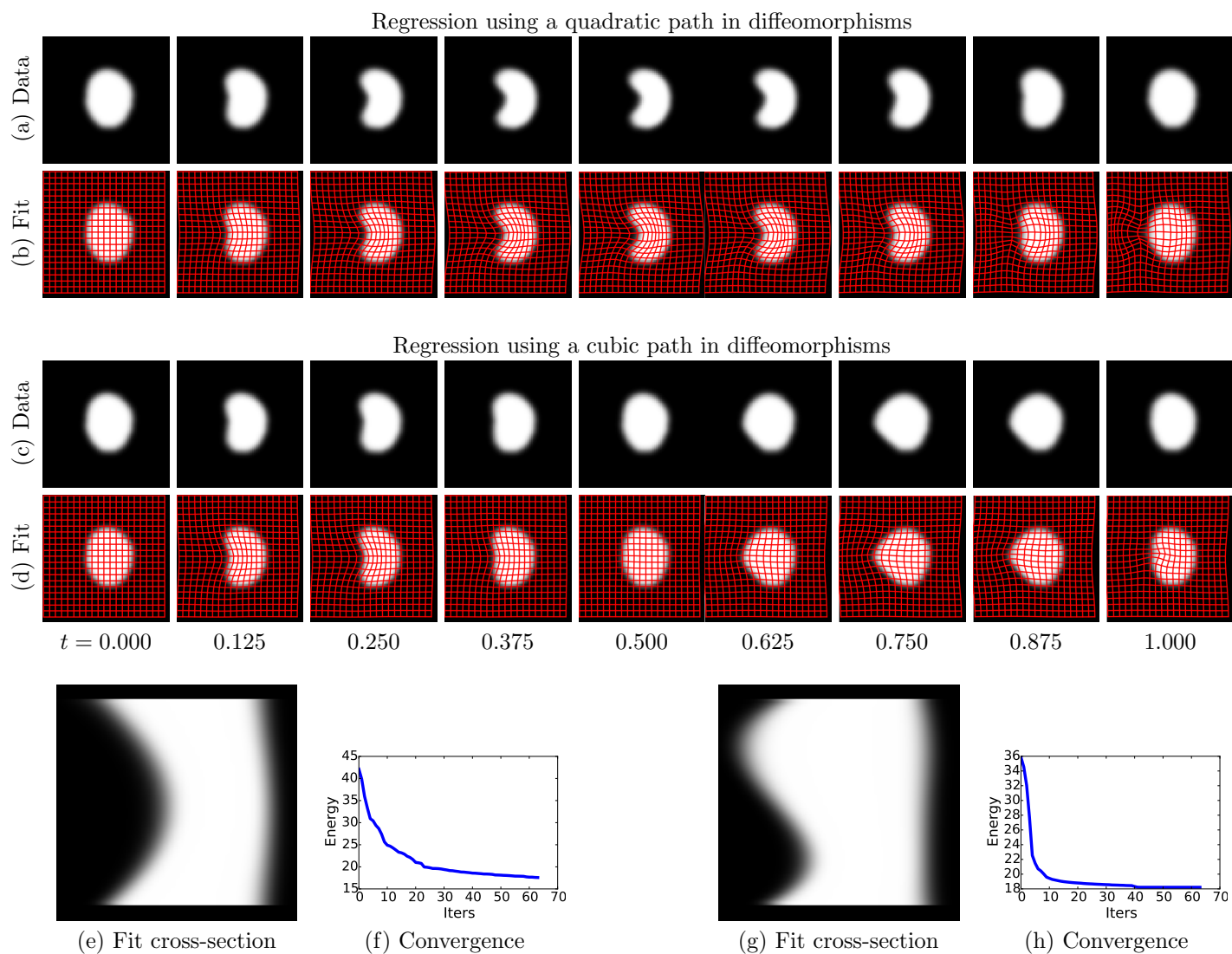
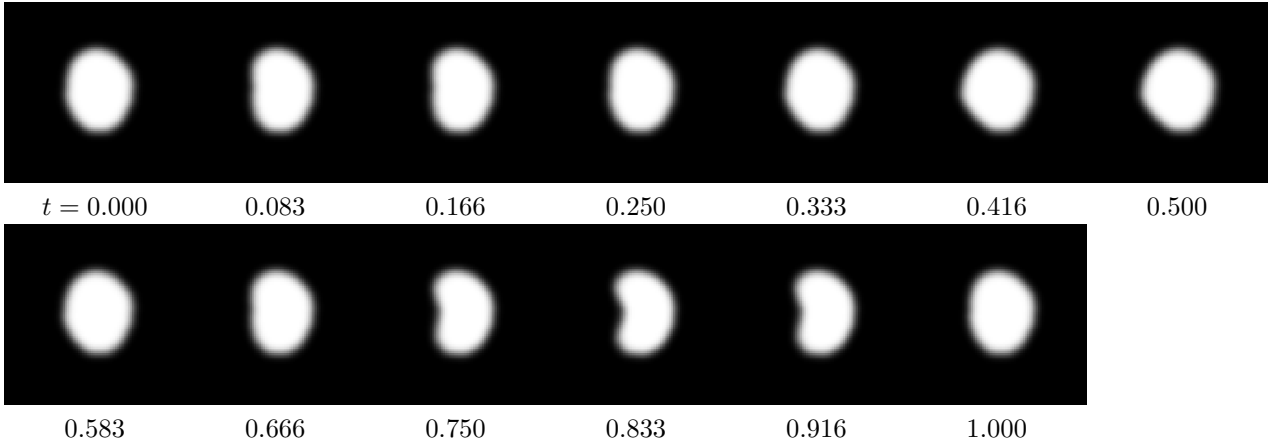
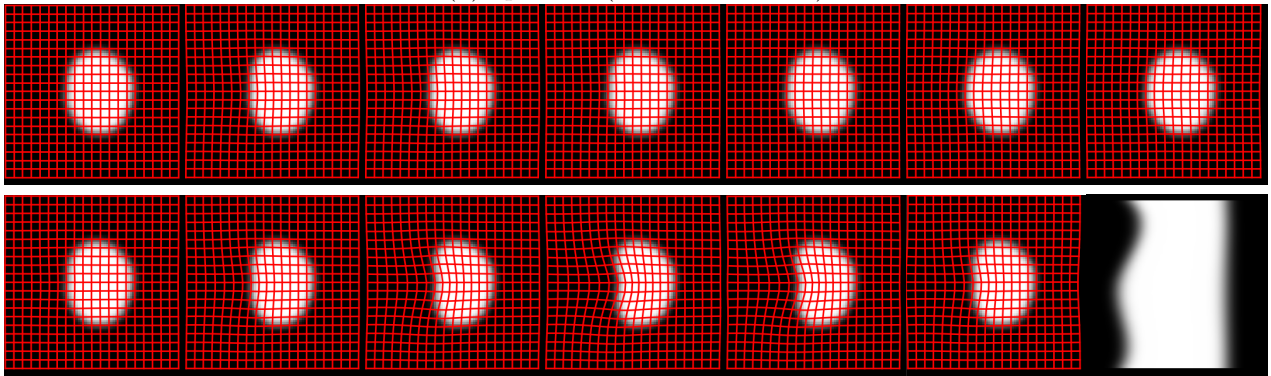


Figure 6: Regression of quadratic and cubic like data. Top two rows (a) and (b) detail the quadratic like sparsely sampled data and the corresponding quadratic regression fit, respectively, and (e) and (f) show the cross-sectional view of movement of a boundary point along the quadratic fit and the convergence of the optimization, respectively. Similarly, rows (c) and (d) detail the cubic-like sparsely sampled data and the corresponding cubic regression fit, respectively, and (g) and (h) show the cross-sectional view of movement of a boundary point along the cubic fit and its convergence, respectively.

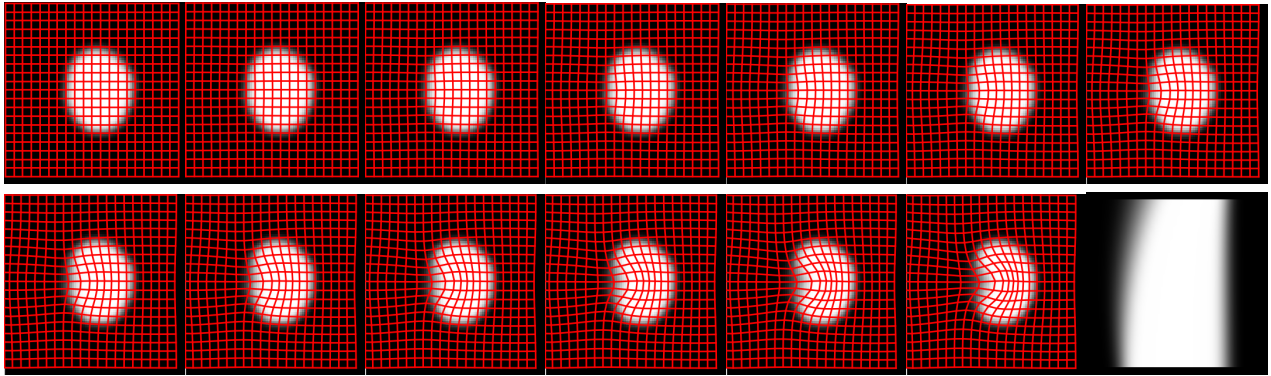
(a) Original data



(b) Spline fit (control at  $t = 0.5$ )



(c) Spline fit (no control)



(d) Geodesic fit

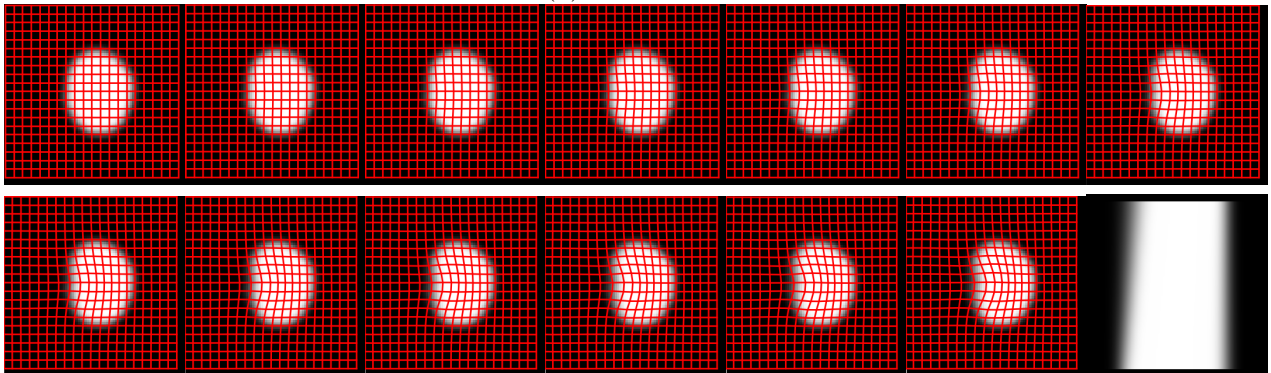


Figure 7: Comparison of regression models on the synthetic data. (a) details the original synthetic shapes that goes through non-monotonic deformations simulating three inflection points. (b) details the spline fit using a control at location  $t = 0.5$  overlaid with the corresponding deformed grid. The last image in the sequence visualizes the motion of a pixel at the boundary, in the middle 1-D horizontal slice of the deforming 2-D image, as we move along the regression fit in diffeomorphisms. (c) and (d) depict the spline and the geodesic fit, respectively.

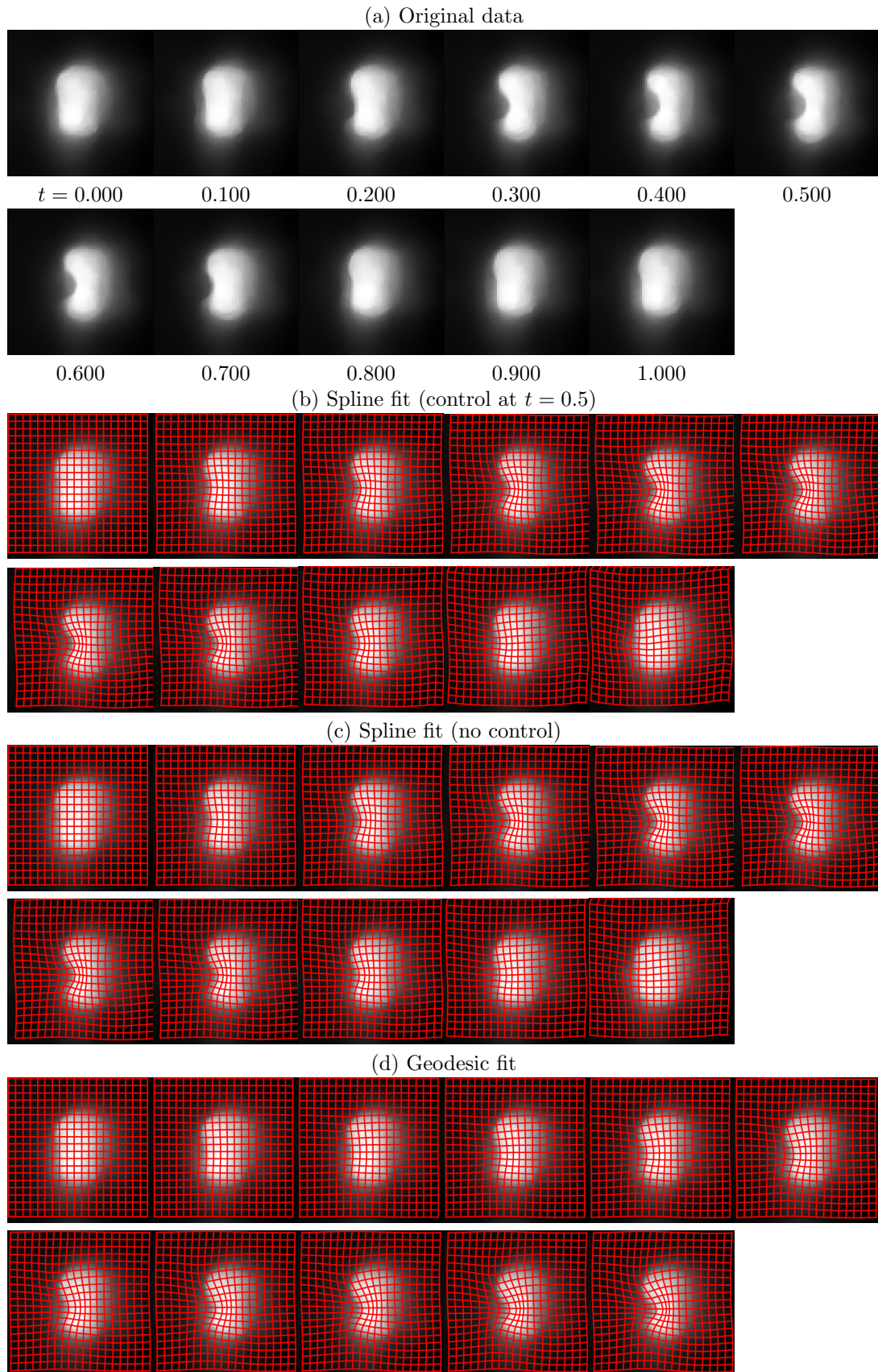
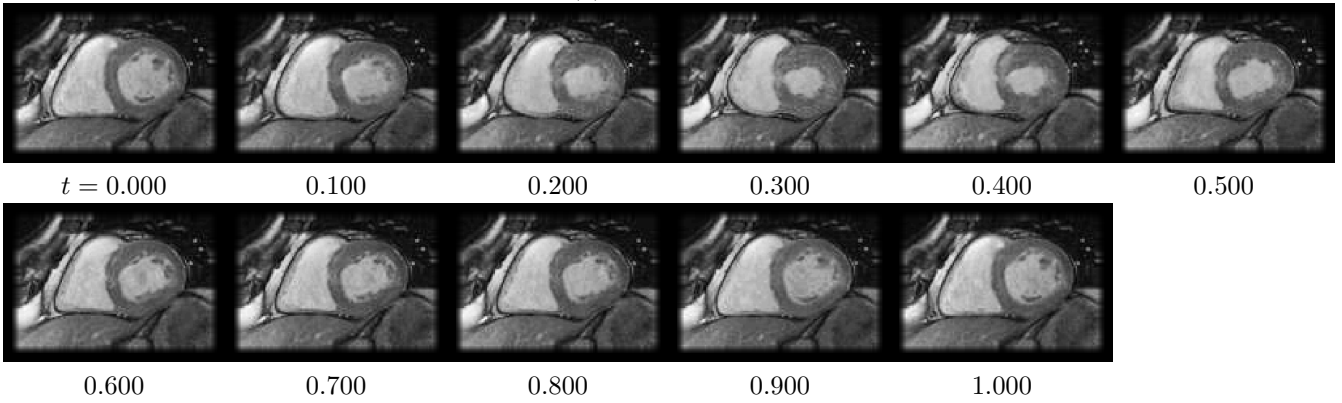


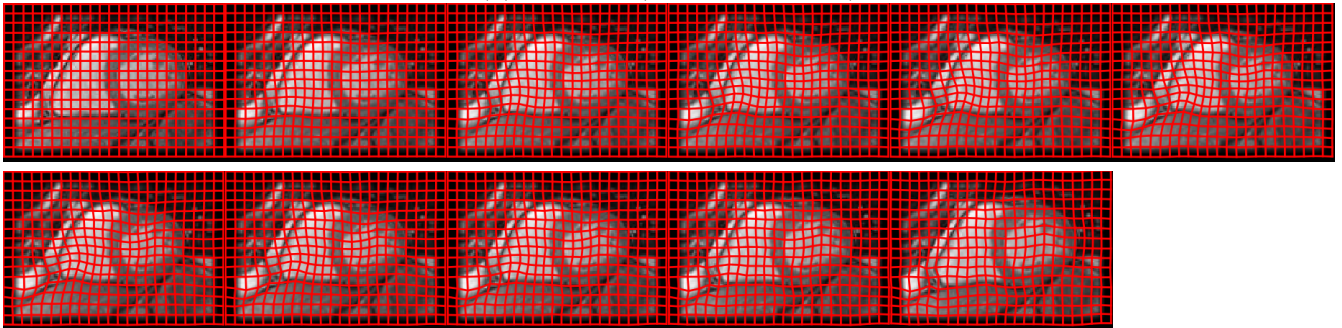
Figure 8: Comparison of regression models on the deforming cell data. (a) details the original cell images that shrinks and then expands back. (b) details the spline fit using a control at location  $t = 0.5$  overlaid with the corresponding deformed grid. (c) and (d) depict the spline with no control and the geodesic fit, respectively.



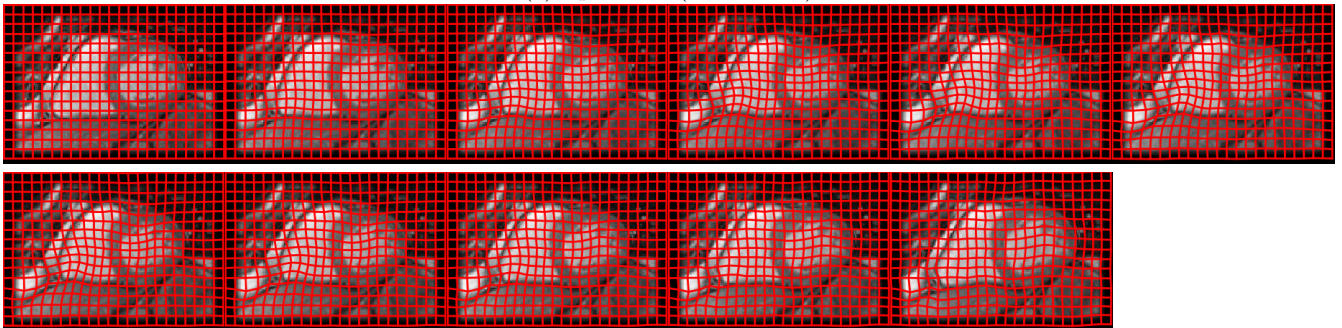
(a) Original data



(b) Spline fit (control at  $t = 0.5$ )



(c) Spline fit (no control)



(d) Geodesic fit

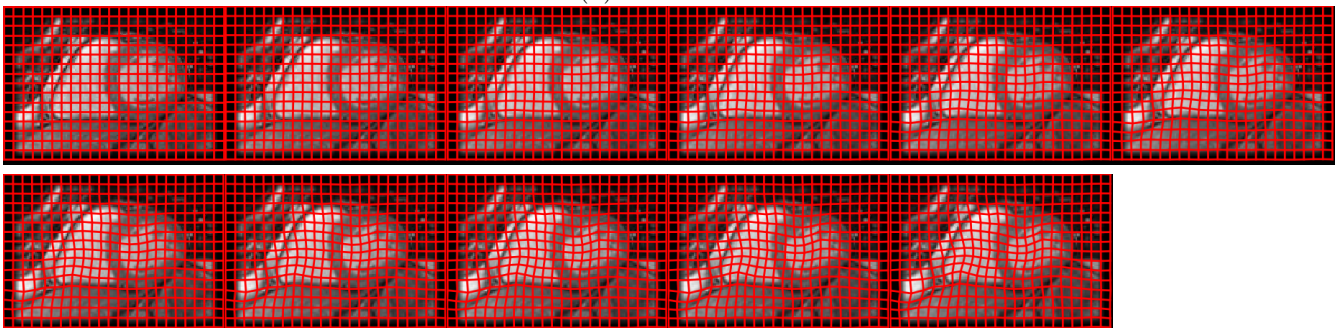


Figure 9: Comparison of regression models on the cardiac motion data. (a) details the original cardiac snapshots. (b) details the spline fit using a control at location  $t = 0.5$  overlaid with the corresponding deformed grid. (c) and (d) depict the spline with no control and the geodesic fit, respectively.