STOCHASTIC METAMORPHOSIS WITH TEMPLATE UNCERTAINTIES

ALEXIS ARNAUDON, DARRYL D HOLM, AND STEFAN SOMMER

ABSTRACT. In this paper, we investigate two stochastic perturbations of the meta-morphosis equations of image analysis, in the geometrical context of the Euler-Poincaré theory. In the metamorphosis of images, the Lie group of diffeomorphisms deforms a template image that is undergoing its own internal dynamics as it deforms. This type of deformation allows more freedom for image matching and has analogies with complex fluids when the template properties are regarded as order parameters (coset spaces of broken symmetries). The first stochastic perturbation we consider corresponds to uncertainty due to random errors in the reconstruction of the deformation map from its vector field, as introduced in [13]. We also consider a second stochastic perturbation, which compounds the uncertainty in of the deformation map with the uncertainty in the reconstruction of the template position from its velocity field. We apply this general geometric theory to several classical examples, including landmarks, images, and closed curves, and we discuss its use for functional data analysis.

1. INTRODUCTION

Variability in shapes can be modelled using flows of the group G of diffeomorphic deformations of the ambient domain Ω in which the shape is embedded. This is the basis of the large deformation diffeomorphic metric mapping (LDDMM) framework, see [4, 7, 8, 28]. In the LDDMM approach, the shape of an embedded template image $\eta \in N$ in the manifold of embedded shapes $\mathrm{Emb}(N, \Omega)$ changes via the action $g_t.\eta$ of time-dependent diffeomorphisms $g_t \in G$ on $\eta \in N$, through the action of g_t on the domain Ω. The metamorphosis extension [14, 21, 30, 31] of LDDMM introduces a further time-dependent variation η_t of the template to model the combined dynamics $g_t.\eta_t$.

In this paper, we combine the geometrical metamorphosis framework of [14] with recent developments in stochastically perturbed Euler-Poincaré dynamics in fluid dynamics and shape analysis [2, 3, 12], to model evolutions of both shape and template under stochastic perturbations. The resulting framework allows modelling of random evolutions of shape and template simultaneously. A potential application of such an evolution is in modelling the progression of disease using computational anatomy, in which the model would address the analysis of disease progression in both the population average and in the individual. From longitudinal image data, mean evolutions over the population can be inferred. While average template evolutions can be modelled deterministically, models for the dynamics of each individual subject that include stochastic uncertainty are arguably more realistic than models

---

supporting only smooth deterministic trajectories. The stochastic metamorphosis model includes such non-smooth and non-deterministic variations by incorporating stochastic perturbations in shape and template simultaneously. We detail this application and outline further areas of applications where similar generative models of data appear, particularly in the combined modelling of phase and amplitude variation in functional data analysis.
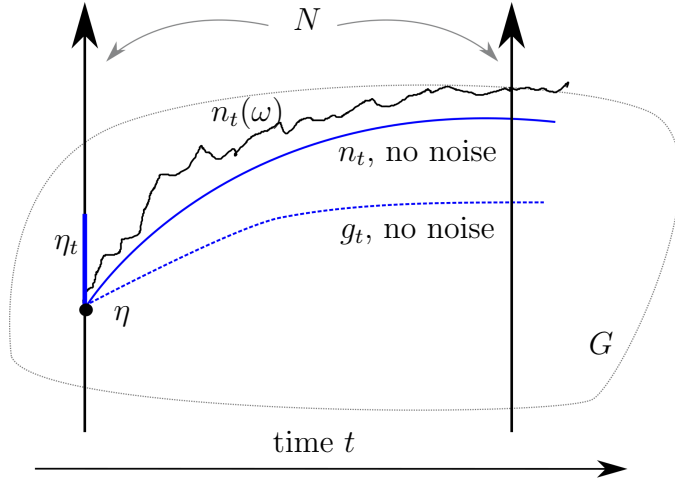

FIGURE 1. Sketch of the deformation group G, the shape space N (vertical arrows), evolution of the deformation variable g_t, template variable η_t, and shape variable n_t without noise ($W_t = 0$, blue), and shape variable n_t with noise ω (black). The shape space is illustrated as being linear (e.g. landmarks, images). However, the framework applies to general non-linear shape spaces (e.g. curves, tensor fields).

1.1. **Background.** The LDDMM framework models the change of a shape $\eta \in N$ by the action of time dependent flows of diffeomorphisms $g_t \in G$ on the embedding space Ω. One lifts the shape trajectory to a time-dependent curve g_t on the diffeomorphisms by setting $n_t = g_t.\eta \in N$. For a right invariant metric on the tangent space of a subgroup G of the diffeomorphism group $\mathrm{Diff}(N)$, N being the shape space, an energy can be defined as $E(g_t) = \int_0^1 \|\partial_t g_t\|_{g_t}^2 \, \mathrm{d}t$. Combined with a data attachment term, this approach allows matching of shapes and image registration [4]. The invariance of $E(g_t)$ under the right action of G implies that the metric descends to a metric structure on the data space N itself. The action of g_t differs between data types, but otherwise, the framework is formally equivalent for different classes of shapes. The use of the flows g_t to model the shape variability is fundamental and the right trivialization $v_t := \partial_t g_t \circ g_t^{-1}$ gives an Eulerian interpretation of the metric. The right invariance of the metric enables Euler-Poincaré reduction of the dynamics to the Lie algebra of G to be performed, and the critical paths for E appear from the reduced dynamics.

Metamorphosis extends the LDDMM setting by letting the template vary in time as well as the deformation, thereby resulting in the flow $g_t.\eta_t$, in which $\eta_0 = \eta$ is the original template. The metamorphosis energy is encoded into a Lagrangian

depending on both the G and N variability, again assuming invariance of the energy to the group action on both G and N. A particular example of metamorphosis dynamics arises in image analysis, where the image I_t changes both by deformation via the right action $g_t.I_t = I_t \circ g_t^{-1}$ and via a pointwise change $\partial_t I_t(x)$ for each pixel/voxel x.

In this formulation of metamorphosis dynamics, an analogy with the flows of complex fluids arises. In complex fluids, a diffeomorphic flow carries an order parameter, defined as a coset space for a broken symmetry of homogeneous fluids, on which the diffeomorphisms act. The order parameter moves with the fluid, but it can also have its own internal dynamics, which in turn is coupled to the fluid motion [10, 11]. A similar combined dynamics of shape and template also appears in the Fshape framework [5].

In [2, 3], a stochastic model of shape evolution was introduced that preserves the Euler-Poincaré theory of the deterministic LDDMM framework. The model is based on the stochastic fluid dynamics model [12] where right-invariant noise is introduced to perturb the reconstruction equation that evolves the flow from the reduced dynamics. In deterministic LDDMM, the reconstruction equation specifies the evolution of the group element by $\partial_t g_t = v_t \circ g_t$ generated by the reduced Eulerian velocity vector field v_t. Stochasticity is introduced as a perturbation to the reconstruction equation, by introducing the stochastic time differential

$$\mathrm{d}g_t\, g_t^{-1} = v_t \mathrm{d}t + \sum_{l=1}^{N} \sigma_l \circ \mathrm{d}W_t^l\,. \tag{1.1}$$

Here W_t^l are standard Wiener processes and σ_l are vector fields on the data domain which characterize the spatial correlation of the noise. As it turns out, the noise in Stratonovich form is denoted conventionally with the same symbol ($\circ$) that denotes composition of maps. This coincidence should not cause any confusion. However, just to be sure, we will write composition of maps as concatenation whenever the two meanings appear in the same equation, as in (1.1). The perturbation of the reduced variable implies that the noise is right-invariant and in a certain sense compatible with the right-invariant LDDMM metric. This approach preserves many of the geometric structures of the deterministic framework. Importantly, the descent of the stochastic model to particular data types is similar to the way the metric descends in the deterministic LDDMM framework.

A stochastic metamorphosis extension of the stochastic Euler-Poincaré framework was introduced in [13]. The stochastic perturbations there were also introduced in the reduced variable influencing the deformation flow from the reconstruction equation. The template evolution η_t is still deterministic. The aim of the present paper is to extend this model to include noise in the template evolution η_t as well. We will make this extension on the reduced template velocity $g_t \partial_t \eta_t$ similarly to the perturbation of the group variable. This procedure results in simultaneous stochastic perturbations of the flow equations for both g_t and η_t.

1.2. **Paper outline.** After a brief survey of the deterministic metamorphosis framework in section (2.1), we formally derive the stochastic model in section (2.2). We

then show in section (2.3) how to derive these equations in the Hamilton-Pontryagin formulation, where the noise appears as a stochastic constraint in the variational principle. We end the theoretical section by deriving the corresponding Hamiltonian stochastic equations in (2.4) to then move to some classical examples of image analysis and computational anatomy in section (3), including landmarks and images. The inclusion of two types of stochastic variations links the framework to combined random phase and amplitude variations in functional data analysis. We provide perspectives of the method to future applications in functional data analysis and computational anatomy in section (4).

2. General stochastic metamorphosis

In this section, we introduce the stochastic deformation of metamorphosis, but first, we recall the basis of this theory, in the context of reduction by symmetry. We will only review what will be needed for our exposition, and we refer to [14] and [13] for more extensive treatments.

2.1. **Deterministic metamorphosis.** The theory of metamorphosis begins with a template N, considered here as a manifold (landmarks, images, etc. . .) upon which a group of diffeomorphism $G = \mathrm{Diff}(N)$ acts. The parameter space of this theory is $G \times N$, with curves $(g_t, \eta_t) \in G \times N$, where g_t is the deformation curve and η_t is the template curve. The image curve will be denoted $n_t = g_t.\eta_t \in N$, where the dot represents the group action. This curve is the total motion of the template, or image N, under both the deformation and its own dynamics. For standard LDDMM, the motion of the image is only $n_t = g_t.\eta$, for a fixed reference template η. This combined action thus allows more freedom in the matching procedure, while remaining compatible with the theory of reduction by symmetries, which we now describe. We first define the two reduced velocity fields

$$u_t := \dot{g}_t g_t^{-1}, \quad \text{and} \quad \nu_t := g_t \dot{\eta}_t. \tag{2.1}$$

The first is the reduced deformation velocity and the second is the reduced template velocity. We then assume that the original Lagrangian of this theory is invariant under the group action of G, so that we may write the reduced Lagrangian in terms of the reduced velocity fields and the image position n_t, i.e.,

$$L(g_t, \dot{g}_t, \eta_t, \dot{\eta}_t) = l(u_t, n_t, \nu_t). \tag{2.2}$$

Because the reduced Lagrangian still depends on the template variable n_t, reduction by the action of the diffeomorphisms will result in a semi-direct product structure, where the template is an advected quantity, in the language of fluid dynamics.

We next compute the variations of the three variables in the reduced Lagrangian, upon introducing the notation $\xi_t = \delta g_t g_t^{-1}$ and $\omega = g\delta\eta$, where δg and $\delta\eta$ are free variations, to obtain

$$\begin{aligned}
\delta u &= \dot{\xi}_t - [u_t, \xi_t], \\
\delta n &= \omega_t + \xi_t \eta_t, \\
\delta \nu &= \dot{\omega}_t + \xi_t \nu_t - u_t \omega_t.
\end{aligned} \tag{2.3}$$

In these formulas, we need to specify what we mean by the multiplication, as the vector fields live in different spaces. In fact, $u_t, \xi_t \in \mathfrak{g} = \mathfrak{X}(N)$ are vector fields; so the Lie bracket is the natural operation. Recall that $\eta_t \in N$, thus $\xi_t \eta_t$ corresponds to the tangent map of the action of G on the manifold N, and similarly for $\nu \in TN$, where the action is on the tangent space of N. We do not need these actions explicitly now, but we will need their 'adjoint action' in the following sense:

$$\langle n^* \diamond m, u \rangle_{\mathfrak{g}} = -\langle n^*, un \rangle_N \,, \tag{2.4}$$

$$\langle u \star \nu^*, \nu \rangle_{TN} = \langle \nu^*, u\nu \rangle_{TN} \,, \tag{2.5}$$

where $N \in N$, $n^* \in N^*$, $\nu \in TN$, $\nu^* \in T^*N$ and $u \in \mathfrak{g}$. The first equality defines the diamond operation ($\diamond$), which will serve as a force term to capture the coupling between the advected quantity n and the main dynamics of the diffeomorphism group. The second equality defines the star operation ($\star$), which is the adjoint of the action of u on TN. That is, it defines the action of u on T^*N.

Applying the variational calculus to the action $S = \int l\,\mathrm{d}t$, we obtain the Euler-Poincaré formulation of the metamorphosis equation in the form

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta u} + \mathrm{ad}^*_{u_t}\frac{\delta l}{\delta u} + \frac{\delta l}{\delta n} \diamond n + \frac{\delta l}{\delta \nu} \diamond \nu &= 0 \,, \\
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta l}{\delta \nu} + u_t \star \frac{\delta l}{\delta \nu} - \frac{\delta l}{\delta n} &= 0 \,,
\end{aligned} \tag{2.6}$$

together with the reconstruction equation

$$\dot{n} = u_t n_t + \nu_t \,. \tag{2.7}$$

We refer to [13, 14] for the details of this derivation, which we will do in the context of Hamilton-Pontryagin with noise in the next section.

From here, a choice of Lagrangian and data N will reduce the system to particular cases, some of which we discuss in the applications section 3.

2.2. Formal derivation of the stochastic equations.
We will first derive the equation informally, using 'stochastic variations', then show a more straightforward derivation using the Hamilton-Pontryagin principle. The second derivation also has the advantage of revealing the effects of the noise more transparently.

In order to introduce a noise compatible with the Euler-Poincaré equation, we need to perturb the theory at its core, which is in this case the definition of the reduced velocities in (2.1). Indeed, the variations were computed from these definitions, and the deterministic Euler-Poincaré equation emerged. Although a single relation is used in the Euler-Poincaré equation (2.7), we will split it into two parts, and perturb them with two different noise components as follows,

$$\begin{aligned}
\mathrm{d}g_t g_t^{-1} &= u_t(x)\mathrm{d}t + \sum_{l=1}^{K^u} \sigma_l^u(x) \circ \mathrm{d}W_t^l =: \mathrm{d}u_t(x) \,, \\
g_t \mathrm{d}\eta &= \nu_t \mathrm{d}t + \sum_{k=1}^{K^\nu} \sigma_k^\nu \circ \mathrm{d}W_t^k =: \mathrm{d}\nu_t \,.
\end{aligned} \tag{2.8}$$

In a slight abuse of notation, $\mathrm{d}u_t(x)$ and $\mathrm{d}\nu_t$ are written as stochastic processes. Here $\sigma_l^u : N \to \mathfrak{g}$ are a set of K^u vector fields on the domain Ω, and $\sigma_l^\nu \in TN$ are another set of K^ν tangent vectors on the template. We also denote by W_t^l or W_t^k the $K^u + N^\nu$ independent standard Weiner processes. In addition, we denote by $x_0 \in \Omega$ the Lagrangian labels upon which g_t acts, so that the first equation can be written equivalently as

$$\mathrm{d}g_t = u_t(g_t x_0)\mathrm{d}t + \sum_{l=1}^{K^u} \sigma_l^u(g_t x_0) \circ \mathrm{d}W_t^l \,.$$

The second equation for η in (2.8) does not have any x_0 dependence, as it is an equation for the template itself. Thus, σ_k^ν are not functions of N; rather, they are tangent vectors to N.

With the notation for $\mathrm{d}u_t$ and $\mathrm{d}\nu_t$ in (2.8), we have the complete reconstruction relation for the stochastic image template n_t

$$\mathrm{d}n_t = \mathrm{d}u_t\, n_t + \mathrm{d}\nu_t \,. \tag{2.9}$$

Because $n_t \in N$, the concatination $\mathrm{d}u_t\, n_t$ means the composition $\mathrm{d}u_t(n_t)$. In (2.9), the noise in the u_t vector field was introduced in [13], based on the stochastic fluid dynamics model of [12], whereas the noise in the ν_t field is new. The first noise term in (2.9) corresponds to random errors in the reconstruction of the diffeomorphism path from its velocity field, while the second one represents random errors for the reconstruction of the template position from its velocity field. In stochastic metamorphosis, the two noise terms will affect the dynamical equations differently.

From these stochastic perturbations of the reconstruction relation, we can formally compute the variations and obtain

$$\begin{aligned} \delta u &= \mathrm{d}\xi_t + [\xi, \mathrm{d}u_t] \,, \\ \delta \nu &= \mathrm{d}\omega + \xi \mathrm{d}\nu_t - \mathrm{d}u_t \omega \,. \end{aligned} \tag{2.10}$$

These are convenient expressions, but they introduce the variations as stochastic processes; so they should not be taken at face value without further analysis. We will see in the next section how to re-derive these equations without introducing stochastic variations, by using the Hamilton-Pontryagin principle. Because the results are identical for the two methods, we can proceed formally here by using these variations as we did in the deterministic variational principle to obtain the following stochastic reduced metamorphosis equations in Euler-Poincaré form,

$$\begin{aligned} \mathrm{d}\frac{\delta l}{\delta u} + \mathrm{ad}_{\mathrm{d}u_t}^* \frac{\delta l}{\delta u} + \frac{\delta l}{\delta n} \diamond n \mathrm{d}t + \frac{\delta l}{\delta \nu} \diamond \mathrm{d}\nu_t &= 0 \,, \\ \mathrm{d}\frac{\delta l}{\delta \nu} + \mathrm{d}u_t \star \frac{\delta l}{\delta \nu} - \frac{\delta l}{\delta n}\mathrm{d}t &= 0 \,, \end{aligned} \tag{2.11}$$

as well as equation (2.9), all to be compared with the deterministic case in equations (2.6) and (2.7).

2.3. Derivation using the Hamilton-Pontryagin principle.
We now show how to rederive the stochastic metamorphosis equations more transparently, without introducing stochastic variations (2.10). For this purpose, we will use the stochastic Hamilton-Pontryagin approach and closely follow the exposition of [13].

The deterministic Hamilton-Pontryagin principle is a variational principle with the following constrained action

$$S(u_t, n_t, \dot{n}_t, \nu_t, g_t, \dot{g}_t) = \int_0^1 l(u_t, n_t, \nu_t)\mathrm{d}t + \int_0^1 \langle M_t, (\dot{g}_t g_t^{-1} - u_t)\rangle \mathrm{d}t$$
$$+ \int_0^1 \langle \sigma_t, (\dot{n}_t - \nu_t - u_t n_t)\rangle \mathrm{d}t \,, \tag{2.12}$$

where $M_t \in \mathfrak{X}^*(N)$ and $\sigma_t \in T^*N$ are generalised Lagrange multipliers to enforce the constraint of the reconstruction relations. Taking free variations for all the variables yields the deterministic reconstruction relation (2.7) and the deterministic Euler-Poincaré equation (2.6). We refer to [13] for more details of the derivation. The crucial point here is to allow free variations, by introducing constraints into the variational principle, and not in the variations as in the standard Euler-Poincaré reduction theory. An alternative approach would be to use the Clebsch constrained variational method used for fluid dynamics in [12].

In the present context, we enforce the stochastic reconstruction relations (2.8) via the following stochastic Hamilton-Pontryagin principle

$$S(u_t, n_t, \mathrm{d}n_t, \nu_t, g_t, \mathrm{d}g_t) = \int_0^1 l(u_t, n_t, \nu_t) + \int_0^1 \langle M_t, (\mathrm{d}g_t g_t^{-1} - \mathrm{d}u_t)\rangle$$
$$+ \int_0^1 \langle \sigma_t, (\mathrm{d}n_t - \mathrm{d}\nu_t - \mathrm{d}u_t n_t)\rangle \,, \tag{2.13}$$

or, more explicitly, upon substituting for $\mathrm{d}u_t$ and $\mathrm{d}\nu_t$ from (2.8), we have

$$S(u_t, n_t, \mathrm{d}n_t, \nu_t, g_t, \mathrm{d}g_t) = \int_0^1 l(u_t, n_t, \nu_t)\mathrm{d}t$$
$$+ \int_0^1 \left\langle M_t, \mathrm{d}g_t g_t^{-1} - u_t \mathrm{d}t - \sum_{l=1}^{K^u} \sigma_l^u(x) \circ \mathrm{d}W_t^l \right\rangle$$
$$+ \int_0^1 \left\langle \sigma_t, \mathrm{d}n_t - \nu_t \mathrm{d}t - \sum_{k=1}^{K^\nu} \sigma_k^\nu(x) \circ \mathrm{d}W_t^k \right\rangle \tag{2.14}$$
$$- \int_0^1 \left\langle \sigma_t, \left(u_t \mathrm{d}t + \sum_{l=1}^{K^u} \sigma_l^u(x) \circ \mathrm{d}W_t^l\right) n_t \right\rangle \,.$$

Proposition 2.1. *The stochastic variational principle $\delta S = 0$ with action (2.14) yields the stochastic Euler-Poincaré equation (2.11) with stochastic reconstruction relation (2.8) and (2.9).*

Proof. The proof is a direct computation by taking free variations. We will show the key steps below. First, the variations with respect to M_t and σ_t yield the reconstruction relations (2.8) and (2.9). Then, the variations with respect to u_t, n_t and ν_t specify

$$\frac{\delta l}{\delta u_t} = M_t + \sigma_t \diamond n_t \,, \quad \text{and} \quad \frac{\delta l}{\delta \nu_t} = \sigma_t \,. \tag{2.15}$$

We also have, for the n_t variations,

$$\frac{\delta l}{\delta n_t}\mathrm{d}t = \mathrm{d}\sigma_t + u_t \star \sigma_t \mathrm{d}t + \sum_{l=1}^{K^u} \sigma_l^u(x) \star \sigma_t \circ \mathrm{d}W_t^l \,. \tag{2.16}$$

Finally, for $\xi = \delta g g^{-1}$ vanishing at the endpoints, we have

$$\delta(\mathrm{d}g_t g_t^{-1}) = \mathrm{d}\xi - \left[u_t \mathrm{d}t + \sum_{l=1}^{K^u} \sigma_l^u(x) \circ \mathrm{d}W_t^l, \xi \right] \,. \tag{2.17}$$

From this computation, we have the last term in the calculus of variations which reads

$$\mathrm{d}M_t = - \mathrm{ad}_{u_t}^* M_t - \sum_{l=1}^{K^u} \mathrm{ad}_{\sigma_l^u(x)}^* M_t \circ \mathrm{d}W_t^l \,. \tag{2.18}$$

Finally, substituting the values of M_t and σ_t of (2.15) in equation (2.16) and (2.18) yields the stochastic metamorphosis equation (2.11) after a few more manipulations (see Corollary 3 of [13]). $\qquad\square$

2.4. **Hamiltonian formulation.** Provided that the Lagrangian is hyperregular, the stochastic metamorphosis equation (2.11) can be written as a stochastic Hamiltonian equation with Hamiltonian obtained via the reduced Legendre transform,

$$h(\mu, \sigma, n) = \langle \mu, u \rangle + \langle \sigma, \nu \rangle - l(u, \nu, n) \,, \tag{2.19}$$

in which μ and σ_t are the conjugate variables of u_t and ν_t, respectively. The noise is encoded into the stochastic potentials

$$\Phi_l^u(\mu_t) = \langle \mu_t, \sigma_l^u \rangle_{\mathfrak{g} \times \mathfrak{g}^*} \,, \quad \text{and} \quad \Phi_k^\nu(\sigma_t) = \langle \sigma_t, \sigma_k^\nu \rangle_{TN \times T^*N} \,, \tag{2.20}$$

such that the stochastic equation of motion has a Hamiltonian drift term with h and stochastic terms obtained via the same Hamiltonian structure, but with stochastic potentials. Notice that the two potentials have a different pairing, one on the Lie algebra of the diffeomorphism group, and the other on the tangent space of the template manifold. The Hamiltonian structure is given in [13] and we will only display here the Hamiltonian equations

$$\mathrm{d}\mu_t + \mathrm{ad}_{\frac{\delta h}{\delta\mu}}^* u \mathrm{d}t + \sigma \diamond \frac{\delta h}{\delta\sigma}\mathrm{d}t + \frac{\delta h}{\delta n} \diamond n \mathrm{d}t$$
$$+ \sum_l \mathrm{ad}_{\frac{\delta\Phi_l^u}{\delta\mu}}^* u \circ \mathrm{d}W_t^l + \sum_l \sigma_t \diamond \frac{\delta\Phi_l^\nu}{\delta\sigma} \circ \mathrm{d}W_t^l = 0 \,, \tag{2.21}$$
$$\mathrm{d}\sigma_t + \frac{\delta h}{\delta\mu} \star \sigma_t \mathrm{d}t - \frac{\delta h}{\delta n} + \sum_l \frac{\delta\Phi_l^u}{\delta\mu} \star \sigma_t \circ \mathrm{d}W_t^l = 0 \,.$$

In the examples in the next section, we will use this formulation to derive the stochastic equations of motion. Taking the Hamiltonian approach turns out to be more transparent than the Lagrangian description.

3. Applications

Following [14], we explicitly provide the stochastic metamorphosis equations for a few classical examples, including landmarks and images, and leave other applications such as closed planar curves, densities or tensor fields for later works.

3.1. Landmarks and peakons.

Consider the case when the template manifold N is the space of n landmarks $\mathbf{q} = (q_1, \ldots, q_n) \in \Omega^n$ with momenta $\mathbf{p} = (p_1, \ldots, p_n) \in T_\mathbf{q}\Omega^n \cong \Omega^n$. One needs to specify a Lagrangian for this system, and the simplest is

$$l(u, n, \nu) = \frac{1}{2}\|u\|_K^2 + \frac{\lambda^2}{2}\sum_{i=1}^n |p_i|^2, \tag{3.1}$$

where the first norm depends on the kernel $K(x)$ and the second norm is the vector norm of the momenta multiplied by a constant λ^2. In this case, we interpret the momenta as the conjugate variables to the template deformation vector field ν in order to have an equation only in term of the position and momenta of the landmarks. The derivation of the landmark equation is rather standard. Hence, we will only show it on the Hamiltonian side. We refer, for example, to [14] for more details of the deterministic derivation, or to [3] and [15] for discussions of the stochastic landmark dynamics.

Recall that the landmark Hamiltonian is

$$h_K(\mathbf{p}, \mathbf{q}) = \frac{1}{2}\sum_{ij} p_i \cdot p_j K(q_i - q_j), \tag{3.2}$$

and the metamorphosis Hamiltonian is thus

$$h(q_i, p_i) = h_K(\mathbf{q}, \mathbf{p}) + \frac{\lambda^2}{2}\sum_{i=1}^n |p_i|^2. \tag{3.3}$$

The stochastic potentials (2.20) become in this case

$$\Phi_l^u(\mathbf{q}, \mathbf{p}) = \sum_i p_i \cdot \sigma_l^u(q_i) \qquad \text{and} \qquad \Phi_i^\nu(\mathbf{p}) = p_i \cdot \sigma_i^\nu. \tag{3.4}$$

Notice that the stochastic potential Φ^ν is described by a fixed vector, where σ_i^ν is the amplitude of the noise for the landmark i. However, for the stochastic potential Φ^u, we have to specify space (or $\mathbf{q}$) dependent functions $\sigma_l^u(\mathbf{q})$. This simple form comes from the fact that we used a discrete set of points and $\nu = \mathbf{p}$ for the template deformation, and the summation over k becomes a summation over the landmark index. In addition, a sum of two Wiener process is another Wiener process with the sum of the amplitude (if it is additive and in Itô form). From this observation, one can see that the general equation $\Phi_k^\nu(\mathbf{p}) = \sum_i p_i \cdot \sigma_k^\nu$ is equivalent to a change of amplitudes σ_k^ν and $i = k$.

We compute the stochastic Hamiltonian equations for landmarks to arrive at

$$dq_i = \frac{\partial h_K}{\partial p_i} dt + \sum_l \sigma_l^u \circ dW_t^l + \lambda^2 p_i dt + \sigma_i^\nu dW_t^i ,$$

$$dp_i = -\frac{\partial h_K}{\partial q_i} dt + \sum_l \partial_{q_i} (p_i \cdot \sigma_l^u) \circ dW_t^l ,$$

(3.5)

in which we can use the Itô integral for the ν-noise, as it is additive.

Notice that setting $\lambda = 0$ recovers the standard landmark dynamics, but with an additive noise in the position equation. This is different from the conventional physical perspective, in which additive noise often appears in the momentum equation, as in [20, 29, 33].

3.2. **Images.** The present stochastic metamorphosis framework can be directly applied to images, by taking the template space N to be the space of smooth functions from the domain $\Omega \subset \mathbb{R}^2$ to $\mathbb{R}$. We set $u_t \in \mathfrak{X}(\Omega)$ the deformation vector field and $\rho \in TN \cong N$ the template vector field. As before, the Lagrangian must have two parts, and the simplest non-trivial one is the sum of kinetic energies written as

$$l(u, n, \nu) = \frac{1}{2} \|u_t\|_K^2 + \frac{\lambda^2}{2} |\rho_t|_{L^2}^2 ,$$

(3.6)

where the first norm depends on the kernel K and the second norm is the standard L^2 norm over Ω. By choosing a L^2 norm we can identify ρ_t with its dual in the case $\lambda = 1$. We will thus not distinguish between σ_t and ν_t of the general framework.

Thus, as before, we use the Hamiltonian formulation of the stochastic metamorphosis equations with the stochastic potentials,

$$\Phi_l^u(m_t) = \int_\Omega \langle m_t(x), \sigma_l^u(x) \rangle dx \quad \text{and} \quad \Phi_k^\nu(\sigma_t) = \int_\Omega \langle \rho_t(x), \sigma_l^\nu(x) \rangle dx .$$

(3.7)

Notice that in this case, both σ_l^u and σ_l^ν are functions of the domain Ω, and they encode spatial correlation structure of the stochastic perturbations.

Then, because the Hamiltonian structure has three sorts of terms, the ad^*, the $\diamond$ and the $\star$ terms defined in equation (2.5), which in this case are

$$\mathrm{ad}_{u_t}^* m_t = (u_t \cdot \nabla) m_t + (m_t \cdot \nabla) u_t + \mathrm{div}(u_t) m_t ,$$

$$\sigma_t \diamond \nu_t = \sigma_t \cdot \nabla \nu_t ,$$

$$u_t \star \sigma_t = \nabla \cdot (\sigma_t u_t) ,$$

we arrive at the following set of stochastic PDEs (for any λ)

$$dm_t + \mathrm{ad}_{u_t}^* m_t dt + \sum_l \mathrm{ad}_{\sigma_l^u}^* m_t \circ dW_t^l = \lambda^2 \rho_t \cdot \nabla \rho_t dt + \sum_k \rho_t \cdot \nabla \sigma_k^\nu \circ dW_t^k ,$$

$$d\rho_t + \nabla \cdot (\rho_t u_t) dt + \sum_l \nabla \cdot (\rho \sigma_l^u) \circ dW_t^l = 0 .$$

(3.8)

Another important equation is the reconstruction relation (2.9), which now reads

$$dg_t = u_t(g_t) dt + \sum_l \sigma_l^u(g_t) \circ dW_t^l + \rho_t dt + \sum_k \sigma_k^\nu \circ dW_t^k .$$

(3.9)

Notice that if we set $\lambda = 1$, the effect of the density, or template motion on the momentum m only appears via the noise term, similarly to the landmark case.

In the one dimensional case, the metamorphosis equation is known to reduce to the so-called CH2 system, which is equation coupling the Camassa-Holm equation with a density advection equation for $\rho_t = \nu_t$. We refer to [6, 14] and references therein for more details about this equation and its complete integrability in the deterministic case. A similar reduction holds for both stochastic deformations, and we have the following stochastic CH2 equation

$$\mathrm{d}m + (u\partial_x m + 2m\partial_x u)\mathrm{d}t$$

$$= -\rho\partial_x\rho\mathrm{d}t - \sum_k \rho\partial_x\sigma_k^\nu \circ \mathrm{d}W_t^k - \sum_l \left(\sigma_l^u\partial_x m + \sum_l 2m\partial_x\sigma_l^u\right)\circ \mathrm{d}W_t^l, \quad (3.10)$$

$$\mathrm{d}\rho + \partial_x(\rho u)\mathrm{d}t + \partial_x(\rho\sigma_l^u)\circ \mathrm{d}W_t^l = 0\,.$$

Compared to the landmark example, the noise associated to the template dynamics is described by a set of functions of the image, not a set of fixed vectors. The difference between the nature of these two types of noise is thus less apparent, apart from how they appear in the equation.

4. Perspectives

4.1. **Computational Anatomy.** Estimation of population atlases and longitudinal analysis of anatomical changes caused by disease progression constitute integral parts of computational anatomy [34]. The relation between these problems and the stochastic metamorphosis model presented here can be illustrated by the analysis of longitudinal brain MR-image data of patients suffering from Alzheimer's disease. The data manifold N is here a vector space of images as described above with $\Omega \subseteq \mathbb{R}^3$.

Focusing first on template estimation, in medical imaging commonly denoted atlas estimation, the aim is to find a population average of data assumed observed at a fixed time point. In the literature, this is for example pursued by minimizing the total sum of the regularized LDDMM energies of deterministic geodesic trajectories that deform the atlas to match the observed data [16]. For k data points $n^1, \ldots, n^k$ and with data matching term $S : N \times N \to \mathbb{R}$, the template η is then estimated by joint minimization of

$$\min_{(\eta, v_t^1, \ldots, v_t^k)} \sum_{i=1}^{k} \int_0^1 \|v_t^i\|^2 \mathrm{d}t + S(\phi_T^i.\eta, n^i)\,, \quad (4.1)$$

where the deformations ϕ_T^i each are endpoints of the integral of the vector fields v_t^i on an interval $[0, T]$.

A different approach to atlas estimation is to perform inference in statistical models where observations are assumed random perturbations of a template and inference of the template is performed via maximum-likelihood or maximum-a-posteriori estimation. This approach is pursued, for example, in a [1, 24, 35]. See also the discussion below.
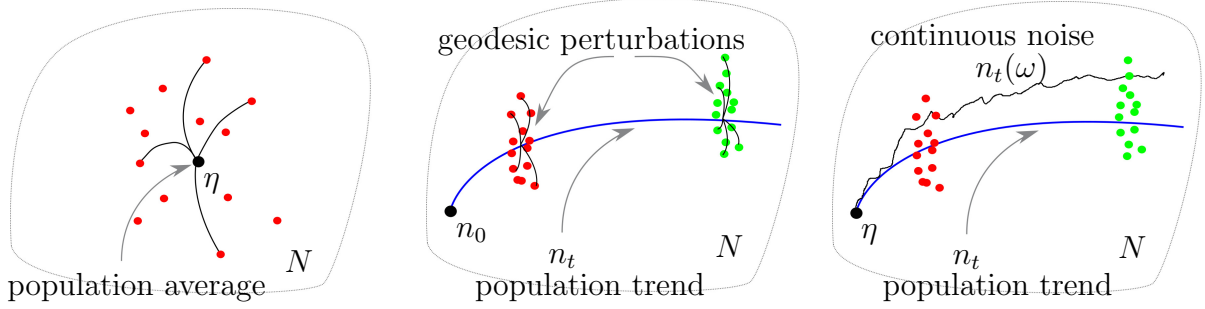
FIGURE 2. (left) Template estimation in the form (4.1) aims at finding a single descriptor η for the population average of the observed shapes $n^1, \ldots, n^k$ (red dots) in the non-linear shape space N. The variational principle (4.1) corresponds to assuming n^i arise from geodesic perturbations of η. (center) Geodesic regression models a population trend as a geodesic n_t. Observations at different time points ($n^i_{t_1}$ red, $n^i_{t_2}$ green) arise as perturbations of the points n_{t_1} and n_{t_2} by random geodesics. (right) Stochastic metamorphosis models the evolution of the population trend n_t deterministically while observations $n^i_{t_j} = n_{t_j}(\omega^i)$ appear from individual noise realizations ω^i. The perturbations are time continuous and apply to each case i individually making the model natural for modelling longitudinal evolutions with noise.

Longitudinal analysis aims at capturing the average time evolution of the brain shape caused by the disease [22, 23]. A common approach here is to estimate a general deterministic trend that is perturbed by noise at discretely observed time points in order to describe the observed images [9]. For example, the noise can take the form of random initial velocity vectors for geodesics emanating from the deterministic trajectory.

The stochastic metamorphosis framework proposed here combines deterministic longitudinal evolution of the template in both shape, represented by the deformations g_t, and in the template image, $n_t = g_t.\eta_t$. We can assume longitudinal observations $n^i_{t_j}$, $i = 1, \ldots, k$, $j = 1, \ldots, t_l$ at l time points are realizations of the stochastic model with time-continuous noise process drawn for each subject i. The stochastic perturbations are thus tied to each subject affecting the dynamics simultaneously with the evolution of the deterministic flow. The relation between this model, geodesic regression models, and atlas estimation is illustrated in Figure 2.

Because of the randomness, algorithms for inference of the template η and its evolution $n_t = g_t.\eta_t$ from data can naturally be formulated by matching statistics of the data, e.g. by matching moments or by maximum-likelihood as done for the landmark case of stochastic EPDiff equations in [3]. Development of such inference schemes constitutes natural future research directions.

4.2. **Phase and Amplitude in Functional Data Analysis.** While images exhibit variations in both intensity and shape of the image domain, signals in functional data analysis often exhibit combined variation in amplitude and phase. For a signal
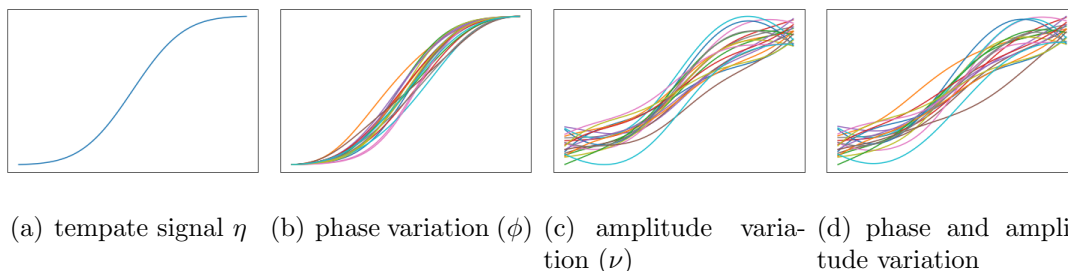
(a) tempate signal η (b) phase variation (ϕ) (c) amplitude varia- (d) phase and ampli-
                                                          tion (ν)   tude variation

FIGURE 3. A template signal (a) can be perturbed by (b) variation in phase, in (4.2) denoted ϕ; (c) variation in amplitude, ν in (4.2); (d) phase and amplitude simultaneously.

$f : I \to N$ defined on an interval I, amplitude variations refer to variations of the values $f(s)$ in N for each fixed $s \in I$ while phase variation covers changes in the parametrization of the domain I. This is illustrated with $N = \mathbb{R}$ in Figure 3. An example of such combined phase and amplitude variations is provided in the growth curves of children and young adults; in which phase variation is connected to the absolute height of the subject while phase variation arise from growth and growth spurts occurring at different ages for different children.

Recent literature covers multiple approaches for identifying, separating and performing inference in situations with combined phase and amplitude variation [19, 25, 32]. One example of a generative model in this settings is the mixed-effects model [18, 25]

$$f(s) = \eta(\phi^{-1}(s)) + \nu(s) + \epsilon , \ s \in I , \tag{4.2}$$

where the average signal η is deformed in phase by the action $\phi.\eta = \eta \circ \phi^{-1}$ of a deformation ϕ of the interval I, and in amplitude by the additive term ν. Here η is considered a fixed, non-random effect while both ϕ and ν are random. Illustrated with the growth curve case above, η models the population average growth curve for each age s, while ϕ controls the timing of the growth process for the individual children and ν the absolute height difference to the population average. One observes that the model (4.2) is non-linear, because of the coupling between ϕ and η. In addition, a model for the deformations ϕ is needed, and the randomness appearing in both ϕ and ν must be specified.

Whereas the LDDMM model is widely used in image analysis, this framework has not yet seen many applications for modelling deformations in functional data analysis, such as the phase variation appearing in (4.2). Instead, works such as [25] use a small-deformation model $\phi(s) = s + v(s)$ with random vector field v modelling displacements on I. On the other hand, large-deformation flow models such as LDDMM traditionally have not integrated random variation directly into the dynamics. Natural families of probability distributions and generative models taking values in non-linear spaces such as deformation spaces are generally non-trivial to construct. However, the model proposed in this paper achieves exactly that.

A direct metamorphosis equivalent of the mixed-effects model (4.2) has $\eta = \eta_0$ the population average η, sets $u_0 = \nu_0 = 0$ and encodes the random effects ϕ and ν in (4.2) in the stochastic increments du_t and $d\nu_t$. The action of g_t on the signal is the right action $g_t.f = f \circ g_t^{-1}$ as in (4.2). Now $d\nu_t$ models pure amplitude variation, du_t phase variation, and the combined stochastic evolution of the signal is $df_t = dn_t$. We then assume the observed signal is $f = f_T$ for a fixed end time T of the stochastic process. Spatial correlation in both the deformation increments du_t and the amplitude increments $d\nu_t$ is encoded in the fields σ_l^u and σ_k^ν respectively.

In the above model, the template is stationary in time when disregarding the stochasticity. However, allowing non-zero initial momenta u_0 and ν_0 in both phase and amplitude allows the template to vary with time and thereby gives a non-linear generalization of a standard multivariate regression model with one latent variable for phase and one for amplitude. This in particular allows modelling of trends over populations where subjects are affected by both the population trend and individual stochastic perturbations.

4.3. **Statistical nonlinear modelling.** It may initially seem overly complicated to use the metamorphosis framework for a simple regression model. However, statistical models that in linear space seem completely standard are often inherently difficult to generalize to non-linear spaces. In general, the lack of vector space structure makes distributions and generative models hard to specify, see e.g. [26, 27] for examples of the geometric complexities of generalizing the Euclidean normal distribution.

In Euclidean space, random vectors can model random perturbations. In non-linear spaces, the lack of vector space structure prevents this and random perturbations are often most naturally expressed with sequences of infinitesimal steps. Vectors are thus replaced with tangent bundle valued sequences that, when integrated over time, give rise to stochastic flows. When modelling both deterministic and random variations, stochasticity generally couples non-trivially with the deterministic evolution. In addition, perturbations and correlation structure must be specified with respect to a frame of reference. While Euclidean space provides a global coordinate system allowing this, a model of transport must be specified in non-linear spaces. The stochastic metamorphosis model is an example of a model coupling deterministic and stochastic evolution and using right-invariance to provide reference frames for the perturbations and correlation structure. An example of a related but different approach is [17] where parallel transport is used to link covariance between tangent spaces.

References

[1] Stephanie Allassonnière, Yali Amit, and Alain Trouvé, *Towards a coherent statistical framework for dense deformable template estimation*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **69** (2007), no. 1, 3–29.

[2] Alexis Arnaudon, Alex L Castro, and Darryl D Holm, *Noise and dissipation on coadjoint orbits*, To appear in JNLS, arXiv preprint arXiv:1601.02249 (2016).

[3] Alexis Arnaudon, Darryl D Holm, and Stefan Sommer, *A geometric framework for stochastic shape analysis*, arXiv preprint arXiv:1703.09971 (2017).

[4] M Faisal Beg, Michael I Miller, Alain Trouvé, and Laurent Younes, *Computing large deformation metric mappings via geodesic flows of diffeomorphisms*, International journal of computer vision **61** (2005), no. 2, 139–157.

[5] Benjamin Charlier, Nicolas Charon, and Alain Trouvé, *The Fshape Framework for the Variability Analysis of Functional Shapes*, Foundations of Computational Mathematics **17** (2017), no. 2, 287–357.

[6] Ming Chen, Youjin Zhang, et al., *A two-component generalization of the camassa-holm equation and its solutions*, Letters in Mathematical Physics **75** (2006), no. 1, 1–15.

[7] Gary E. Christensen, Richard Rabbitt, and Michael I. Miller, *Deformable templates using large deformation kinematics*, Image Processing, IEEE Transactions on **5** (1996), no. 10.

[8] Paul Dupuis, Ulf Grenander, and Michael I. Miller, *Variational Problems on Flows of Diffeomorphisms for Image Matching*, Quarterly of applied mathematics (1998).

[9] P. Thomas Fletcher and Miaomiao Zhang, *Probabilistic Geodesic Models for Regression and Dimensionality Reduction on Riemannian Manifolds*, Riemannian Computing in Computer Vision, 2016, pp. 101–121.

[10] François Gay-Balmaz and Tudor S Ratiu, *The geometric structure of complex fluids*, Advances in Applied Mathematics **42** (2009), no. 2, 176–275.

[11] Darryl D Holm, *Euler-Poincaré dynamics of perfect complex fluids*, Geometry, mechanics, and dynamics (2002), 169–180.

[12] _____, *Variational principles for stochastic fluid dynamics*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **471** (2015), no. 2176, 20140963.

[13] _____, *Stochastic metamorphosis in imaging science*, arXiv preprint arXiv:1705.10149 (2017).

[14] Darryl D Holm, Alain Trouvé, and Laurent Younes, *The Euler-Poincaré theory of metamorphosis*, Quarterly of Applied Mathematics **67** (2009), no. 4, 661–685.

[15] Darryl D Holm and Tomasz M Tyranowski, *Variational principles for stochastic soliton dynamics*, Proc. r. soc. a, 2016, pp. 20150827.

[16] Sarang Joshi, Brad Davis, Matthieu Jomier, and Guido Gerig, *Unbiased diffeomorphic atlas construction for computational anatomy*, NeuroImage **23** (2004), 151–160.

[17] Line Kühnel and Stefan Sommer, *Stochastic Development Regression on Non-linear Manifolds*, Information Processing in Medical Imaging, June 2017, pp. 53–64.

[18] Line Kühnel, Stefan Sommer, Aksshha Pai, and Lars Raket, *Most Likely Separation of Intensity and Warping Effects in Image Registration*, SIAM Journal on Imaging Sciences **10** (2017), no. 2, 578–601.

[19] J. S. Marron, James O. Ramsay, Laura M. Sangalli, and Anuj Srivastava, *Functional Data Analysis of Amplitude and Phase Variation*, Statistical Science **30** (2015), no. 4, 468–484. MR3432837

[20] Stephen Marsland and Tony Shardlow, *Langevin equations for landmark image registration with uncertainty*, SIAM Journal on Imaging Sciences **10** (2017), no. 2, 782–807.

[21] Michael I. Miller and Laurent Younes, *Group actions, homeomorphisms, and matching: A general framework*, International Journal of Computer Vision **41** (2001), no. 1-2, 61–84.

[22] Prasanna Muralidharan and P. Thomas Fletcher, *Sasaki Metrics for Analysis of Longitudinal Data on Manifolds*, Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition **2012** (2012), 1027–1034.

[23] Marc Niethammer, Yang Huang, and Franois-Xavier Vialard, *Geodesic Regression for Image Time-Series*, Medical Image Computing and Computer-Assisted Intervention MICCAI 2011, 2011, pp. 655–662.

[24] Akshay Pai, Stefan Sommer, Lars Lau Raket, Line Kühnel, Sune Darkner, Lauge Srensen, and Mads Nielsen, *A Statistical Model for Simultaneous Template Estimation, Bias Correction, and Registration of 3d Brain Images*, 2016.

[25] Lars Lau Raket, Stefan Sommer, and Bo Markussen, *A nonlinear mixed-effects model for simultaneous smoothing and registration of functional data*, Pattern Recognition Letters **38** (2014), 1–7.

[26] Stefan Sommer, *Anisotropic Distributions on Manifolds: Template Estimation and Most Probable Paths*, Information Processing in Medical Imaging, 2015, pp. 193–204.

[27] Stefan Sommer and Anne Marie Svane, *Modelling anisotropic covariance using stochastic development and sub-Riemannian frame bundle geometry*, Journal of Geometric Mechanics **9** (2017), no. 3, 391–410 (en).

[28] Alain Trouvé, *An infinite dimensional group approach for physics based models in pattern recognition*, preprint (1995).

[29] Alain Trouvé and François-Xavier Vialard, *Shape splines and stochastic shape evolutions: a second order point of view*, Quarterly of Applied Mathematics **70** (2012), no. 2, 219–251.

[30] Alain Trouvé and Laurent Younes, *Local geometry of deformable templates*, SIAM journal on mathematical analysis **37** (2005), no. 1, 17–59.

[31] _____, *Metamorphoses through lie group action*, Foundations of Computational Mathematics **5** (2005), no. 2, 173–198.

[32] J. Derek Tucker, Wei Wu, and Anuj Srivastava, *Generative models for functional data using phase and amplitude separation*, Computational Statistics & Data Analysis **61** (2013), no. Supplement C, 50–66.

[33] François-Xavier Vialard, *Extension to infinite dimensions of a stochastic second-order model associated with shape splines*, Stochastic Processes and their Applications **123** (2013), no. 6, 2110–2157.

[34] Laurent Younes, Felipe Arrate, and Michael I. Miller, *Evolutions equations in computational anatomy*, NeuroImage **45** (2009), no. 1, Supplement 1, S40–S50.

[35] Miaomiao Zhang, Nikhil Singh, and P Thomas Fletcher, *Bayesian estimation of regularization and atlas building in diffeomorphic image registration*, International conference on information processing in medical imaging, 2013, pp. 37–48.

AA, DH: Department of Mathematics, Imperial College, London SW7 2AZ, UK

SS: Department of Computer Science (DIKU), University of Copenhagen, DK-2100 Copenhagen E, Denmark