Learning the spatio-temporal variability of longitudinal shape data sets: application to Alzheimer's disease progression modeling

Alexandre Bône, Olivier Colliot, Stanley Durrleman

Noname manuscript No.
(will be inserted by the editor)

Learning the spatio-temporal variability of longitudinal shape data sets: application to Alzheimer's disease progression modeling

Alexandre Bône · Olivier Colliot · Stanley Durrleman · The Alzheimer's Disease Neuroimaging Initiative

Received: date / Accepted: date

Abstract In order to diagnose, monitor, and eventually cure a disease, quantitative measures of abnormality are required. If reference growth curves are available for child developmental disorders, building normative scenarii of alterations for neuro-degenerative diseases, which slowly develop over years, remains an open problem. Two main difficulties explains this. First, relevant and objective markers of progression most often come from medical imaging, and are therefore very high-dimensional and structured measurements. As a consequence, the subtle temporal individual changes are easily masked by the large innate and normal variability of the population. Second, a disease like Alzheimer's may start developing at any age and at any pace. Therefore no common time-line is explicitly available to compare individual health records. Based on the large deformation diffeomorphic metric mapping framework, this paper introduces unified modeling approach that jointly learns a mean progression pattern along with temporal and geometrical variance estimates from unaligned temporal progressions of imaging measurements. Normal distributions of shape trajectories are defined as generative and hierarchical statistical models, which are learned by an original calibration algorithm. Based on the estimated normative scenarii, vizualisation, correlation, classification and simulation tasks are naturally defined and carried out. Proposed algorithms are validated on a simulated data set, illustrated on video sequences of emotive faces, and applied on a medical data set of Alzheimer's patients. Atrophy-protective genetic, biological and environmental cofactors are identified.

Alexandre Bône, Olivier Colliot, Stanley Durrleman
Institut du Cerveau et de la Moelle épinière, Inserm, CNRS, Sorbonne Université, Paris, France
Inria, Aramis project-team, Paris, France
{alexandre.bone, olivier.colliot, stanley.durrleman}@icm-institute.org

1 Introduction

1.1 Motivation

In order to diagnose, monitor, and eventually cure a disease, quantitative measures of abnormality are required. Newborns for instance, who develop at a rapid pace, are subject to a particular attention: among other metrics, their size and weight are regularly measured and consigned. Individual progressions are compared to normative scenarii of growth, and offer in turn powerful information to caregivers. If size and weight are relatively relevant markers of child growth, they are much poorer markers of adult aging, and absolutely unrelated with neurodegenerative diseases such as Alzheimer's – which remains today largely misunderstood despite its high social and economic cost. A normative scenario of more advanced features is required, medical images at the top of them all. Building this scenario is however a challenge, because of two fundamental differences with our introductive example: (i) unlike size or weight, medical images are very high-dimensional and structured measurements; (ii) unlike birth, Alzheimer's disease can start at any age. If each challenge taken separately could be tackled with existing technical solutions, a method able to jointly align the individual progressions and quantitatively assess the remaining inter-subject differences is still missing. Following this intuition, we propose in this paper a unified modelling approach able to learn a relevant normative scenario from unaligned temporal progressions of imaging measurements.

1.2 Related work

High-dimensional structured data can advantageously be represented as elements of curved spaces, in order to intrinsically model their structure. Either defined by invariance [30,

49, 50] or topology-preserving properties [8, 13, 17, 29, 45], shape spaces originally aim at defining relevant distance metrics adapted to the geometry of a well-identified class of objects, such as brain magnetic resonance images or segmented organs. Based on those structured spaces, generalized mean-variance analysis [3, 23, 24, 44, 58] learns the geometrical distribution of a cross-sectional data set in terms of an average shape, and variability-encoding parameters. Typical healthy or pathological configurations can be summarized in this manner, opening the way to automatic diagnosis at the individual level. Time-series data sets, consisting in the repeated observation of the same object at successive time-points, can be advantageously described by generalized regression approaches on the same shape spaces [7, 22,23,25,35,43]. A time-continuous scenario of geometrical transformation is then estimated, offering in turn individualized interpolation and extrapolation models.

Classical shape spaces usually have a differential structure of infinite dimension. In particular, the large deformation diffeomorphic metric mapping (LDDMM) define shape spaces as orbits of template shapes under the action of an infinite-dimensional parametric group of diffeomorphisms of the 2D/3D ambient space [56]. With this approach, the geometrical differences between two objects are captured by estimating the diffeomorphic transformation that warps one into the other. Recent works developed LDDMM-based sparse approaches: [57] uses truncated Fourier transforms to build a finite-dimensional Lie algebra, when [16] constructs a finite-dimensional Riemannian manifold based on a set of self-interacting particles, untying the construction of deformations from the shapes on which those deformations act.

Such structures are favorable to the analysis of longitudinal data sets because they naturally offer the parallel transport operator [36], which allows to compare tangent-space vectors at distant points in a relevant manner. In [49,50], trajectories on manifolds are analyzed by parallel-transporting their orientation vectors back to some privileged point of the manifold. A similar approach is followed in [47] with the co-adjoint transport, as well as in [32] which generalizes the linear approach in [31] dedicated to voxel-wise analysis of large medical images. In [46] the authors define the exp-parallelization operator which adapts the notion of parallel lines to curved spaces. The works [9, 33] build on this operator to analyze dynamic networks and shape objects respectively. Exceptions to this trend in the recent literature would be for instance [42] which defines the so-called Sasaki metric on the manifold of geodesic trend parameters, or [10] which works directly in a space of trajectories, well-defined for a fixed number of observations per subject.

If parallel transport allows to spatially align manifold-valued trajectories, a temporal alignment mechanism is also needed for data sets with variability in the individual progression dynamics. For instance, two patients developing the same neurological pathology have no reason to reach the same disease stage at the same age, nor to remain in synchrony along the disease course. The common solution is to use time-warp functions, which define a mapping between an abstract common reference time frame and the individual time lines [9, 17, 32, 33, 46, 49, 50]. In [49, 50], the authors build on the square-root velocity fields framework to quotient the space of spatio-temporal paths by diffeomorphic time-warps. In the other cited works, simple affine functions are used to synchronize individual time-lines.

1.3 Contributions

In this paper, we propose a method that learns a mean progression pattern along with temporal and geometrical variance estimates from a longitudinal data set of shapes. This paper extends and further validate the conference paper [9], with finer modeling of the variability in the individual paces of progression, an original optimization method for accelerated model calibration, and three experimental settings that demonstrate the descriptive power of the proposed approach.

Section 2 defines the concept of shape spatio-temporal reference frame, which allows to introduce in Section 3 our statistical model of trajectories as a natural generalization of the mean-variance analysis. Section 4 details the calibration, personalization and simulation algorithms, which are evaluated and illustrated in Section 5 on varied data sets.

2 Shape spatio-temporal reference frame

2.1 Rationale

Traditional morphometry positions shape objects with respect to a reference shape, often called atlas or template. A coordinate system is defined in the tangent-space at the template shape. We propose here to replace the template (which is a single shape) by a curve (i.e. a shape trajectory), and the coordinate system by a tubular spatiotemporal coordinate system centered around the template trajectory. We first review the usual construction of a static template shape before extending it to the spatiotemporal case.

2.2 Positioning a shape with respect to a static atlas

Positioning a target shape y with respect to a static reference y_0 is called the registration problem. Deformation-based morphometry solves it by estimating a diffeomorphism ϕ_1 of the ambient space $\mathbb{R}^d$ ($d = 2$ or 3) that transforms y_0 into y, which we note $\phi_1 \star y_0 = y$. In the context of large deformation diffeomorphic metric mapping (LDDMM), diffeomorphisms are constructed by following the streamlines of

dynamic vector fields $t \to v_t \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d)$ over $[0,1]$:

$$\partial_t \phi_t = v_t \circ \phi_t \quad \text{with} \quad \phi_0 = \text{Id}. \tag{1}$$

Following the approach in finite-dimension of [18], we further assume that any v_t writes as the Gaussian convolution of p momentum vectors $m_t = m_t^{(1)}, ..., m_t^{(p)} \in \mathbb{R}^d$ over a corresponding set of control points $c_t = c_t^{(1)}, ..., c_t^{(p)} \in \mathbb{R}^d$:

$$v_t : x \in \mathbb{R}^d \to \sum_{k=1}^{p} g\big[c_t^{(k)}, x\big] \cdot m_t^{(k)} \in \mathbb{R}^d \tag{2}$$

with $g : x, x' \in \mathbb{R}^d \to \exp \|x' - x\|^2_{\ell^2} / \sigma^2$ the Gaussian kernel function of kernel width $\sigma > 0$. Many other diffeomorphisms constructed in this manner might actually transform y_0 into $\phi_1 \star y_0$: we call solution of the registration problem the most regular transformation i.e. that minimizes its "kinetic" energy:

$$\frac{1}{2}\int_{t=0}^{1} \|v_t\|^2_{G_{c_t}} = \frac{1}{2}\int_{t=0}^{1} m_t^\top \cdot G_{c_t} \cdot m_t = \frac{1}{2}\, \|v_0\|^2_{G_{c_0}} \tag{3}$$

where $\forall t \in [0,1]$, G_{c_t} is the $p \times p$ "kernel" symmetric positive-definite matrix of general term $g[c_t^{(k)}, c_t^{(l)}]$, and $(.)^\top$ is the matrix transposition. The last equality of equation (3) is proven in [41], where the author also demonstrate that paths of minimal energy verify the Hamiltonian equations:

$$\dot{c}_t = G_{c_t} \cdot m_t \quad ; \quad \dot{m}_t = -\frac{1}{2}\nabla_{c_t}\left\{m_t^\top \cdot G_{c_t} \cdot m_t\right\} \tag{4}$$

where $\nabla_x(.)$ is the gradient operator with respect to x. Assuming that there exist a diffeomorphism ϕ_1 constructed according to equations (1, 2) such that $\phi_1 \star y_0 = y$, this last result allows to compactly represent the positioning of y with respect to y_0 with a set of p control points c_0 and attached momenta m_0. In practice, the perfect registration constraint $\phi_1 \star y_0 = y$ is relaxed, and we call solution to the registration problem the extremal-path diffeomorphism ϕ_1 that warps y_0 as close as possible to y, for some extrinsic error measure $d_{\mathcal{E}}(y_0, y)$. In this paper, the following choices are considered for $d_{\mathcal{E}}$, depending on the nature of y_0 and y:

- the ℓ^2 noise model for meshes with point-to-point correspondence,
- the current [52] or varifold [12] metrics for oriented surface meshes without point-to-point correspondence.

Noting y_0 as a collection $y_0^{(1)}, ..., y_0^{(K)}$ of K points of $\mathbb{R}^d$, ϕ_1 acts independently and directly on each point $y_0^{(k)}$ according to $\phi_1 \star y_0^{(k)} = \phi_1 \circ y_0^{(k)}$. Note that if this work focuses on mesh shapes (typically obtained by segmentation of raw images), it could be seamlessly adapted to the case of images.

2.3 Riemannian structure

Let c_0 be a set of p control points. We define:

$$\mathcal{D}_{c_0} = \big\{\phi_1 \,|\, \partial_t\phi_t = v_t \circ \phi_t,\ \phi_0 = \text{Id},\ v_t = \text{Conv}(c_t, m_t) \\ (\dot{c}_t, \dot{m}_t) = \text{Ham}(c_t, m_t),\ m_0 \in \mathbb{R}^{p\times d}\big\} \tag{5}$$

where Conv(.,.) and Ham(.,.) are compact notations for the convolution operator defined by equation (2) and the Hamiltonian equations given by equation (4) respectively. Equipped at any $\phi \in \mathcal{D}_{c_0}$ with the local metric $G^{-1}_{\phi(c_0)}$, $\mathcal{D}_{c_0}$ has the structure of a Riemannian manifold of dimension $p \times d$. The tangent-space at ϕ is the set of velocity fields obtained by convolving any set of momentum on $\phi(c_0)$:

$$T_\phi\mathcal{D}_{c_0} = \big\{\text{Conv}(\phi(c_0), m) \,|\, m \in \mathbb{R}^{p\times d}\big\}. \tag{6}$$

The geodesics of $\mathcal{D}_{c_0}$ are the curves $t \to \phi_t$ of constant kinetic energy (see eq. 3) i.e. such that the corresponding control points and momenta trajectories $t \to c_t, m_t$ are Hamiltonian (see eq. 4). We define the exponential operator on $\mathcal{D}_{c_0}$:

$$\text{Exp}_\phi^{t_0,t} : v_0 \in T_\phi\mathcal{D}_{c_0} \to \phi_t \in \mathcal{D}_{c_0} \tag{7}$$

where ϕ_t is the diffeomorphism reached at time t by the geodesic path obtained by integration from some reference time $t_0 \in \mathbb{R}$ with initial conditions $\phi(c_0)$, m_0 such that $v_0 = \text{Conv}(\phi(c_0), m_0)$, and $\phi_0 = \phi$. The particular case $\text{Exp}_\phi^{0,1}$ corresponds to the classical Riemannian exponential map and will simply be noted Exp_ϕ. Diffeomorphisms $\phi \in \mathcal{D}_{c_0}$ act on shapes of the ambient space y through the action $\star$ previously defined. Let y_0 be a reference shape. We define its orbit under the action $\star$:

$$\mathcal{S}_{y_0,c_0} = \mathcal{D}_{c_0} \star y_0 = \{\phi \star y_0 \,|\, \phi \in \mathcal{D}_{c_0}\}. \tag{8}$$

$\mathcal{S}_{y_0,c_0}$ is a submanifold of the extrinsic shape space $\mathcal{E}$ in which is defined the distance $d_{\mathcal{E}}$.

2.4 Positioning a shape with respect to a dynamic atlas

Instead of positioning shapes with respect to a static atlas y_0, shapes are positioned with respect to a shape geodesic $t \to \gamma(t) \star y_0$, where γ is a geodesic of $\mathcal{D}_{c_0}$ of the form $\gamma(t) = \text{Exp}_{\text{Id}}^{t_0,t}(v_0)$ with $v_0 = \text{Conv}(c_0, m_0)$. Similarly to cylindrical coordinates in Euclidian spaces, under some conditions (see [26,39]) a shape $y \in \mathcal{S}_{y_0,c_0}$ admits unique Fermi coordinates $t \in \mathbb{R}$ and $v \in T_{\gamma(t_0)}\mathcal{D}_{c_0}$ such that $v \perp \dot{\gamma}(t_0)$:

$$y = \text{ExpP}_\gamma^v(t) \star y_0 \ \text{ with } \ \text{ExpP}_\gamma^v(t) = \text{Exp}_{\gamma(t)}\big[\text{P}_\gamma^v(t)\big] \tag{9}$$

where $\text{P}_\gamma^v(t)$ denotes the parallel transport of v along γ from t_0 to t. The curves γ and $\eta : t \to \text{ExpP}_\gamma^v(t)$ are said exp-parallel, and the mapping $\gamma \to \eta$ is called exp-parallelization along v. Figures 1 and 2 illustrate this concept of longitudinal registration. In other words, a choice of y_0, c_0, m_0, t_0

defines a spatio-temporal reference frame, with respect to which a shape y can be unambiguously positioned in terms of a time t and a velocity field v orthogonal to $v_0 = \dot{\gamma}(t_0)$. This decomposition can be seen as the orthogonal projection of y onto the one-dimensional submanifold $\gamma \star y_0$, hence the condition $v \perp v_0$. The temporal reference t_0 plays no particular role, in the sense that y can be described in the same manner for any other choice t_0'; bijective reference change formulae can actually be simply derived:

$$t' = t + t_0' - t_0 \quad \text{and} \quad v' = \mathrm{P}_\gamma^v(t_0'). \tag{10}$$

In general, the target shape y might not exactly belong to $\mathcal{S}_{y_0,c_0}$. Similarly to the static atlas case, equation (9) is relaxed and we call solution to the longitudinal registration problem the pair t, v such that y_0 is warped as close as possible to y, in the sense of the extrinsic metric $d_{\mathcal{E}}$.

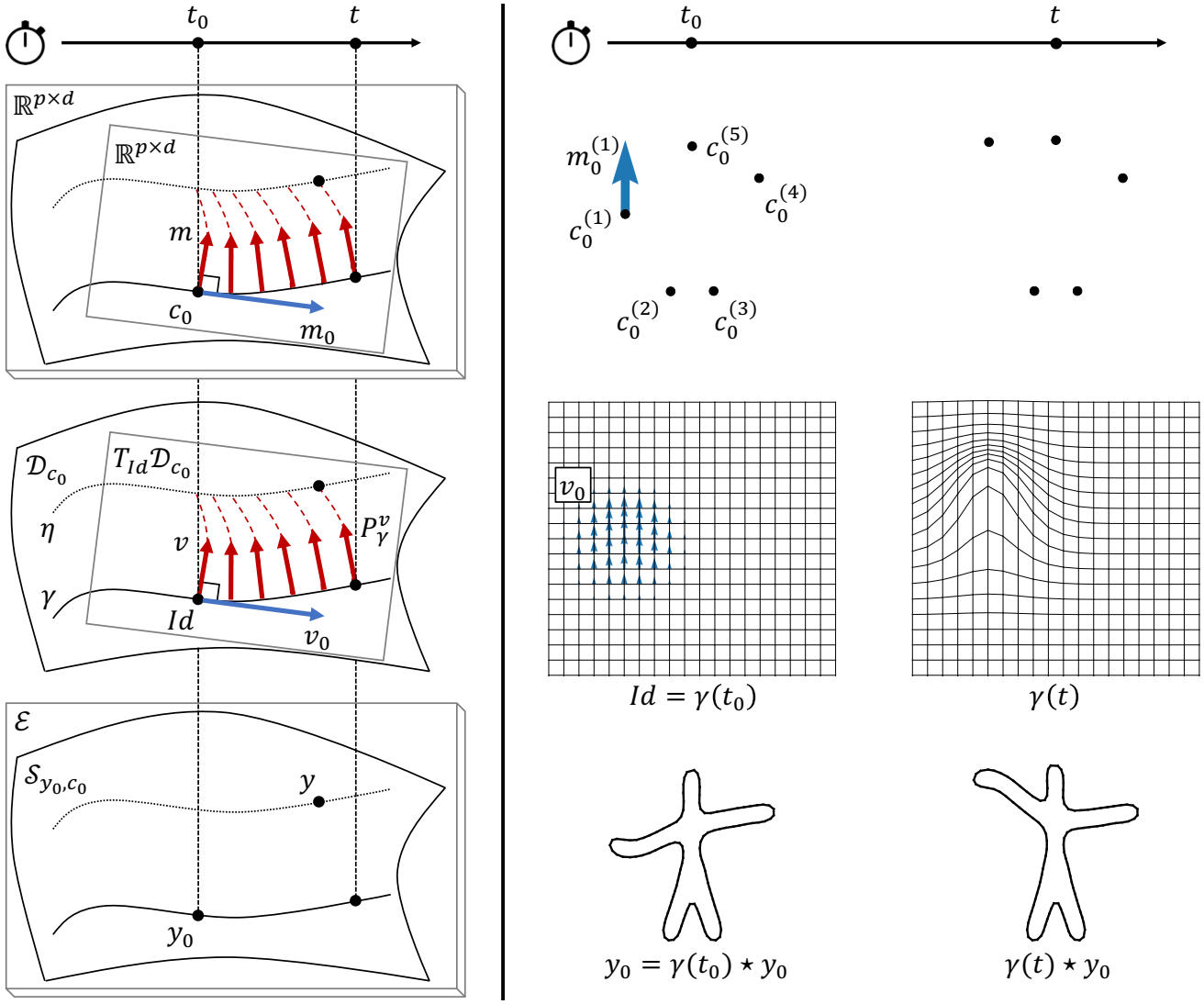


Fig. 1: **[Left].** Shape spatio-temporal reference frame y_0, c_0, m_0, t_0 with respect to which a shape y admits coordinates $t \in \mathbb{R}$, $v \in v_0^\perp \subset T_{\gamma(t_0)}\mathcal{D}_{c_0}$. Three structural layers are involved: the landmark submanifold of $\mathbb{R}^{p\times d}$ which hosts the control points c_0, the manifold of diffeomorphisms $\mathcal{D}_{c_0}$ which hosts the curves γ and η, and the shape submanifold $\mathcal{S}_{y_0,c_0}$ of the extrinsic shape space $\mathcal{E}$. The tangent bundles of the landmark manifold and $\mathcal{D}_{c_0}$ are dual spaces, both isomorphic to the vector space $\mathbb{R}^{p\times d}$. They respectively host the momenta m_0, m and the corresponding velocity fields v_0, v. The latter is parallel-transported along the geodesic γ by the operator $t \to P_\gamma^v(t)$.
[Right]. Illustrations of the manifolds abstractly depicted on the left side of the figure. Each line plots elements of the corresponding geodesic (solid black lines on the left panel). The two columns correspond to the times t_0 and t.

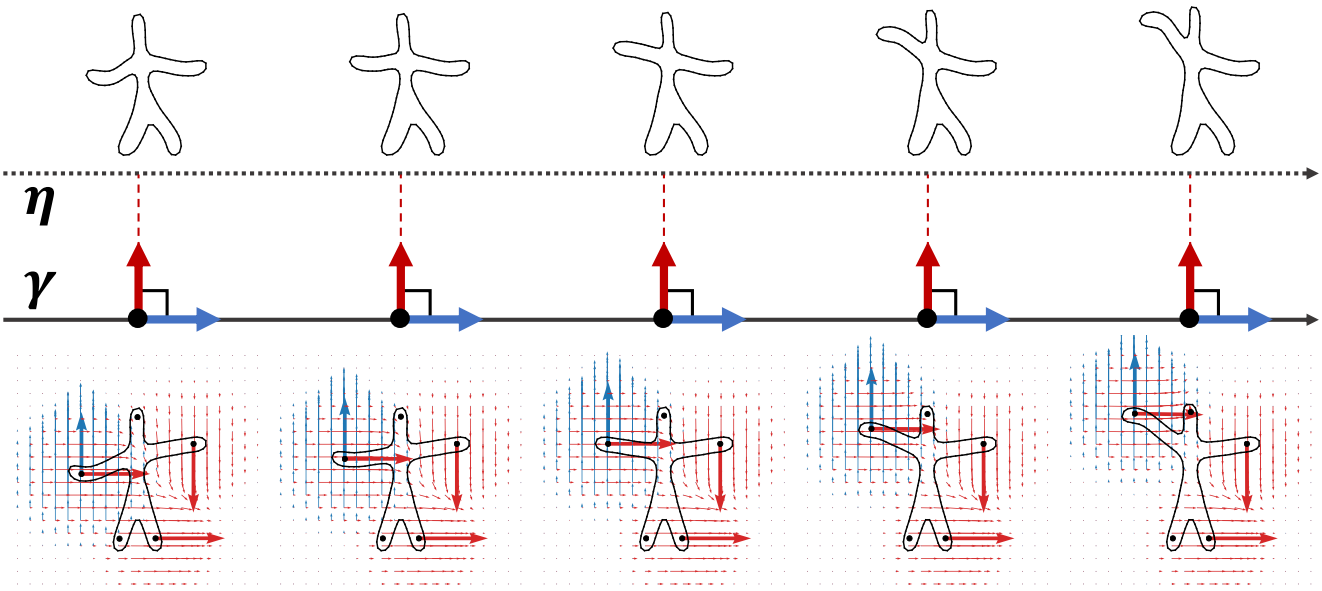


Fig. 2: **[Bottom].** Illustration of a shape geodesic $t \to \gamma(t) \star y_0$: the man-like shape (solid black contour) raises his left arm. This geodesic is fully parametrized encoded by a single set of control points c_0 (black dots) and attached momentum vectors m_0 (bold blue arrows), to which corresponds the velocity field v_0 (light blue arrows). A second single set of momentum vector m (bold red arrows) attached to the same control points c_0 parametrizes the exp-parallelization of this shape geodesic. **[Top].** Exp-parallel shape curve $t \to \eta(t) \star y_0$ to the shape geodesic $\gamma \star y_0$: the man-like shape (solid black contour) also raises his left arm, in a way that respects its own geometry.

3 Statistical model for longitudinal data sets of shapes

3.1 The generative and hierarchical statistical model

Let $(y, t) = (y_{i,j}, t_{i,j})_{i,j}$ be a longitudinal data set of shapes: y is structured as the collection of repeated individual measurements $y_{i,1}, ..., y_{i,n_i}$ for $i = 1, ..., n$, and to each $y_{i,j}$ corresponds a time $t_{i,j} \in \mathbb{R}$. Measurements are modeled as sample points along individual trajectories described in a common shape spatio-temporal reference frame y_0, c_0, m_0, t_0 which will be estimated so as to minimize the variance of spatio-temporal coordinates of the individual progressions, thus generalizing classical mean-variance analyses to time-indexed shapes. Noting $v_0 = \text{Conv}(c_0, m_0)$ and $\gamma : t \to \text{Exp}_{c_0}^{t_0,t}(v_0)$ the template geodesic, the statistical model writes:

$$
\begin{aligned}
&\text{ExpP}_{\gamma}^{v_i}\left[\psi_i(t_{i,j})\right] \star y_0 \overset{\text{iid}}{\sim} \mathcal{N}_{\mathcal{E}}\left(y_{i,j},\, \sigma_\epsilon^2\right), \\
&\text{where} \left|
\begin{array}{l}
\psi_i : t \to \alpha_i \cdot (t - \tau_i) + t_0, \\
v_i = \text{Conv}(c_0, m_i), \quad m_i = A_{0,m_0^\perp} \cdot s_i,
\end{array}
\right. \qquad (11) \\
&\text{and} \left|
\begin{array}{l}
\alpha_i \overset{\text{iid}}{\sim} \mathcal{N}_{[0,+\infty[}(1, \sigma_\alpha^2), \quad \tau_i \overset{\text{iid}}{\sim} \mathcal{N}(t_0, \sigma_\tau^2), \\
s_i \overset{\text{iid}}{\sim} \mathcal{N}(0, 1).
\end{array}
\right.
\end{aligned}
$$

Model (11) is hierarchical in the sense that individual trajectories $t \to \text{ExpP}_{\gamma}^{v_i} \circ \psi_i(t)$ are independently defined as spatio-temporal transformations of a common, population-level geodesic $t \to \gamma(t)$. The time-warp functions ψ_i encode the temporal variability of the observed individual trajectories in terms of pace of progression α_i and onset time τ_i. Those parameters allow to synchronize the individual progressions with the template geodesic γ. The "spatial" or geometrical variability across individuals is encoded by the space-shifts $v_i \in v_0^\perp \subset T_{\gamma(t_0)}\mathcal{D}_{c_0}$ along which γ is exp-parallelized. Those space-shifts admit dual representations under the form of momenta m_i, which are assumed to derive from q source parameters $s_i = s_i^{(1)}, ..., s_i^{(q)}$, in the spirit of independent component analysis (ICA) [27]. The orthogonality $v_i \perp v_0$, necessary for the identifiability of the model, is ensured by the projection of each column of the $(p \cdot d) \times q$ mixing matrix A_0 onto the hyperplane $m_0^\perp$ of $\mathbb{R}^{p \times d}$ for the cometric G_{c_0}, noted $A_{0,m_0^\perp}$. The individual parameters are modeled as independent samples from normal distributions:

- a truncated normal distribution with fixed mean for the acceleration factor α_i, ensuring the identifiability of m_0;
- a normal distribution for the onset time τ_i;
- a normal distribution with fixed mean and variance, ensuring the identifiability of y_0 and $A_{0,m_0^\perp}$ respectively.

In a word, individual trajectories are considered as random spatiotemporal perturbations of a population trajectory. The spatial and temporal transformations commute, in the sense that $\forall t \in \mathbb{R}$, $\text{ExpP}_{\gamma \circ \psi_i}^{v_i} = \text{ExpP}_{\gamma}^{v_i} \circ \psi_i$. The population trajectory is fully parametrized by the template shape y_0, the control points c_0, the momenta m_0 and the reference time

t_0. The individual variability is unambiguously represented by two reduced sets of scalar parameters: the acceleration α_i and the onset time τ_i in one hand for the temporal part, and the sources $s_i^{(1)}, ..., s_i^{(q)}$ on the other hand for the spatial part. In practice, it is possible to choose a number of sources $q \ll p \times d$ much lower than the dimension of the tangent-space $T_{\gamma(t_0)}\mathcal{D}_{c_0}$ while still capturing a satisfyingly representative ratio of the geometrical variability of the data. The longitudinal measurements $(y_{i,j})$ are finally seen as sample points at times $(t_{i,j})$ along the individual trajectories.

3.2 Mixed-effects and Bayesian modeling

We further specify the formulation of the model (11) to fit the framework of mixed-effects models. We distinguish:

- the fixed-effects $\theta = (\theta_1, \theta_2)$ with $\theta_1 = (t_0, \sigma_\tau, \sigma_\alpha, \sigma_\epsilon)$ and $\theta_2 = (y_0, c_0, m_0, A_0)$, also simply called the model parameters,
- the random effects $z = (z_i)_i$ where $z_i = (\alpha_i, \tau_i, s_i)$.

Note that the population parameters correspond to the fixed effects, and the individual ones to the random effects. In spirit, such a modeling embodies the idea that the observed data set is a random realization of a deeper mechanism, that will be captured by averaging over the full random effects distribution. We also choose to work in a Bayesian framework, in order to theoretically ensure the existence of the maximum a posteriori (MAP) estimate of the parameters θ^{m}. Such priors also regularize and guide the estimation procedure thanks to reasonable and mild prior assumptions on the optimal fixed effects values. The following standard conjugate distributions are selected as Bayesian priors on the model parameters:

$$
\begin{aligned}
t_0 &\sim \mathcal{N}(\overline{t_0}, \varsigma_t^2), & y_0 &\sim \mathcal{N}(\overline{y_0}, \varsigma_y^2),\\
\sigma_\tau^2 &\sim \mathcal{IG}(m_\tau, \varsigma_\tau^2), & c_0 &\sim \mathcal{N}(\overline{c_0}, \varsigma_c^2),\\
\sigma_\alpha^2 &\sim \mathcal{IG}(m_\alpha, \varsigma_\alpha^2), & m_0 &\sim \mathcal{N}(\overline{m_0}, \varsigma_m^2),\\
\sigma_\epsilon^2 &\sim \mathcal{IG}(m_\epsilon, \varsigma_\epsilon^2), & A_0 &\sim \mathcal{N}(\overline{A_0}, \varsigma_A^2),
\end{aligned}
$$

where $\mathcal{IG}(.,.)$ denotes the inverse-gamma distribution.

4 Algorithms: calibration, personalization, simulation

4.1 Objectives

Given a longitudinal data set of shapes (y, t), we formulate three algorithmic objectives:

- *Calibration*, which consists in computing the maximum a posteriori (MAP) parameters θ^{m}, unconditionally to any random effect z:

$$\theta^{\mathrm{m}} = \mathrm{argmax}_\theta \int p(y, z, \theta\,; t) \cdot \mathrm{d}z. \tag{12}$$

- *Personalization*, which consist in computing the MAP random effects z^{m} given the calibrated model θ^{m}:

$$z^{\mathrm{m}} = \mathrm{argmax}_z\ p(y, z, \theta^{\mathrm{m}}\,; t). \tag{13}$$

- *Simulation*, which consist in generating a new data set y^{s} that resembles the original data set y.

In a word, we propose to capture, analyze and reproduce the variability of a longitudinal data set of shapes y by sequentially solving those sub-problems. Code will be made available upon publication for all algorithms.

4.2 The complete log-likelihood sub-algorithm

Computing the joint likelihood $p(y, z, \theta\,; t)$ (or rather its logarithm, for numerical stability reasons) is central for both calibration and personalization algorithms. Behind this statistical notation hides the geometrical modelling introduced in Section 2, which relies on computationally-intensive numerical integration schemes. Algorithm 1 details the procedure, where $|\mathcal{E}|$ denotes the dimension of the extrinsic

Algorithm 1: Compute the complete log-likelihood.

input : Longitudinal data set of shapes $y, t = (y_{i,j}, t_{i,j})_{i,j}$.
Population parameters $\theta = y_0, c_0, m_0, a_0, t_0, \sigma_\tau, \sigma_\alpha, \sigma_\epsilon$.
Individual parameters $z = (z_i)_i$ with $z_i = \alpha_i, \tau_i, s_i$.
output: The complete log-likelihood value $Q = \log p(y, z, \theta\,; t)$.
Set $Q = 0$. `// initialization`

`/* compute the squared residuals` $\epsilon_{i,j}^2$ `for each visit */`
Compute the initial velocity field $v_0 = \mathrm{Conv}(c_0, m_0)$.
Compute the geodesic $\gamma : t \to \mathrm{Exp}_{\mathrm{Id}}^{t_0,t}(v_0)$. `// see [18]`
for the geometrical component index $l = 1$ **to** q
  Compute the l-th column of $A_{0,m_0^\perp}$, projecting $\mathrm{Col}_l(A_0)$ on $m_0^\perp$.
  Compute the initial velocity field $w_l = \mathrm{Conv}\big[c_0,\ \mathrm{Col}_l(A_{0,m_0^\perp})\big]$.
  Compute the parallel transport $w_l : t \to \mathrm{P}_\gamma^{w_l}(t)$. `// see [37]`
end
for the individual index $i = 1$ **to** n
  for the visit index $j = 1$ **to** n_j
    Compute the time-warped age $\psi_{i,j} = \alpha_i \cdot (t_{i,j} - \tau_i) + t_0$.
    Compute the initial velocity field $v_{i,j} = \sum_{l=1}^{q} s_i^{(l)} \cdot w_l(\psi_{i,j})$.
    Compute $\phi_{i,j} = \mathrm{Exp}_{\gamma(\psi_{i,j})}(v_{i,j}) \circ \gamma(\psi_{i,j})$. `// see [18]`
    Compute the squared residual $\epsilon_{i,j}^2 = d_{\mathcal{E}}(y_{i,j},\ \phi_{i,j} \star y_0)^2$.
    `/* add the model log-likelihood` $\log p(y_{i,j}|z_i, \theta\,; t_{i,j})$ `*/`
    Update $Q \leftarrow Q - \frac{1}{2}\big\{\ |\mathcal{E}| \cdot \log \sigma_\epsilon^2 + \epsilon_{i,j}^2/\sigma_\epsilon^2\big\}$.
  end
  `/* add the random effects log-likelihood` $\log p(z_i|\theta)$ `*/`
  Update $Q \leftarrow Q - \frac{1}{2}\big\{\ \log \sigma_\tau^2 + (\tau_i - t_0)^2/\sigma_\tau^2 + \|s_i\|_{\ell^2}^2 + \log \sigma_\alpha^2 + \log(1 - F(-1/\sigma_\alpha))^2 + (\alpha_i - 1)^2/\sigma_\alpha^2\big\}$.
end

`/* add the Bayesian prior log-likelihood` $\log p(\theta)$ `*/`
Update $Q \leftarrow Q - \frac{1}{2}\big\{(t_0 - \overline{t_0})^2/\varsigma_t^2 + m_\tau\,(\log \sigma_\tau^2 + \varsigma_\tau^2/\sigma_\tau^2) + m_\alpha\,(\log \sigma_\alpha^2 + \varsigma_\alpha^2/\sigma_\alpha^2) + m_\epsilon\,(\log \sigma_\epsilon^2 + \varsigma_\epsilon^2/\sigma_\epsilon^2)\big\}$. `// log` $p(\theta_1)$
Update $Q \leftarrow Q - \frac{1}{2}\big\{\ \|y_0 - \overline{y_0}\|_{\ell^2}^2/\varsigma_y^2 + \|c_0 - \overline{c_0}\|_{\ell^2}^2/\varsigma_c^2 + \|m_0 - \overline{m_0}\|_{\ell^2}^2/\varsigma_m^2 + \big\|A_0 - \overline{A_0}\big\|_{\ell^2}^2/\varsigma_A^2\big\}$. `// log` $p(\theta_2)$

shape shape $\mathcal{E}$, and $F(.)$ the cumulative distribution function (CDF) of the standard Gaussian.

4.3 Calibration

4.3.1 The MCMC-SAEM-GD algorithm

Calibration is a computationally-intensive task, for two fundamental reasons. First, the optimized variable θ is of high dimension $|\theta| = 4 + |y_0| + d \cdot p \cdot (2 + q)$ where d is the dimension of the ambient space, p the number of control points, q the number of considered geometrical components, and $|y_0|$ the number of coordinates or pixels necessary to describe the template mesh or image respectively. Second, the optimized function cannot be directly computed, as we access $p(y, z, \theta\,; t)$ by solving sets of ordinary differential equations (see Algorithm 1). In this paper, we propose to address this computational challenge by combining the MCMC-SAEM algorithm with gradient descent (GD). The backbone of this algorithm is the SAEM algorithm [14], which is a stochastic approximation (SA) of the classical expectation-maximization (EM) algorithm [15]: are alternated a stochastic simulation step $z^{[k]} \sim p(z|y, \theta^{[k-1]})$ of the latent variables followed by a deterministic update of the model parameters $\theta^{[k]} \leftarrow \theta^\star(z^{[k]})$. In [4], the authors introduce the MCMC-SAEM algorithm, where the simulation step is replaced by an Monte-Carlo Markov chain (MCMC) step while still preserving the theoretical convergence properties. In this paper, an analytical update rule $\theta^\star$ cannot be found for all the parameters θ: we use a gradient descent approach to overcome this difficulty, and we name MCMC-SAEM-GD the global resulting algorithm. Algorithm 2 gives a high-level pseudo-code of the proposed procedure. The mentioned sufficient statistics write:

$$S_t = \frac{1}{n}\sum_{i=1}^{n} \tau_i, \qquad S_\alpha = \frac{1}{n}\sum_{i=1}^{n} (\alpha_i - 1)^2, \tag{14}$$

$$S_\tau = \frac{1}{n}\sum_{i=1}^{n} \tau_i^2, \qquad S_\epsilon = \frac{1}{|\mathcal{E}| \cdot n \cdot \langle n_i \rangle_i} \sum_{i=1}^{n}\sum_{j=1}^{n_i} \epsilon_{i,j}^2,$$

where $\epsilon_{i,j}^2 = d_{\mathcal{E}}\{y_{i,j},\ \mathrm{ExpP}_\gamma^{v_i}\big[\psi_i(t_{i,j})\big] \star y_0\}^2$ and $\langle n_i \rangle_i$ is the average number of longitudinal observations per subject. The update rules write:

$$t_0^\star = \left[\varsigma_t^2\, S_t + \frac{\sigma_\tau^{\star 2}}{n}\,\overline{t_0}\right] \cdot \left[\varsigma_t^2 + \frac{\sigma_\tau^{\star 2}}{n}\right]^{-1} \tag{15}$$

$$\sigma_\tau^\star = \left[S_\tau - 2\,t_0^\star\, S_t + t_0^{\star 2} + \frac{m_\tau}{n}\,\varsigma_\tau^2\right]^{\frac{1}{2}} \cdot \left[1 + \frac{m_\tau}{n}\right]^{-\frac{1}{2}} \tag{16}$$

$$\sigma_\alpha^\star = \left[S_\alpha + \frac{m_\alpha}{n}\,\varsigma_\alpha^2\right]^{\frac{1}{2}} \cdot \left[1 - \frac{f(-1/\sigma_\alpha^\star)/\sigma_\alpha^\star}{1 - F(-1/\sigma_\alpha^\star)} + \frac{m_\alpha}{n}\right]^{-\frac{1}{2}} \tag{17}$$

$$\sigma_\epsilon^\star = \left[S_\epsilon + \frac{m_\epsilon}{|\mathcal{E}|\ n\,\langle n_i \rangle_i}\,\varsigma_\epsilon^2\right]^{\frac{1}{2}} \cdot \left[1 + \frac{m_\epsilon}{|\mathcal{E}|\ n\,\langle n_i \rangle_i}\right]^{-\frac{1}{2}} \tag{18}$$

where $f(.)$ is the probability density function of the standard normal distribution. Both the coupled set of equations (15-16) and the implicit equation (17) can easily be solved by iterative replacement.

Algorithm 2: Calibration with MCMC-SAEM-GD.

input : Dataset y. Initial parameters $\theta^{[0]}$ and $z^{[0]}$.
Sequence of step-sizes $(\rho^{[k]})_k$. Sampling variances $(\sigma^{(b)})_b$.
output: Estimation of $\theta^{\mathrm{m}} \approx \theta^{[k]}$.

Set $k = 0$ and $S_1^{[0]} = S_1(z^{[0]})$. `// initialization, eq. 14`
repeat
  Set $k \leftarrow k + 1$.
  `/* block Gibbs symmetric random walk sampling */`
  foreach block index b **do**
    Draw a candidate $z^{(b)} \sim \mathcal{N}(z^{[k-1](b)}, \sigma_b^2)$ for the block b.
    Let $z = (z^{[k](1)}, ..., z^{[k](b-1)}, z^{(b)}, z^{[k-1](b+1)}, ...)$.
    Compute the ratio $\omega = \log \frac{p(y,z,\theta^{[k-1]};t)}{p(y,z^{[k-1]},\theta^{[k-1]};t)}$. `// alg. 1`
    Draw u according to the uniform distribution $u \sim \mathcal{U}(0,1)$.
    if $\log u < \omega$ **then** $z^{[k](b)} \leftarrow z^{(b)}$ **else** $z^{[k](b)} \leftarrow z^{[k-1](b)}$.
  end
  Adapt the proposal variances $(\sigma^{(b)})_b$. `// see [6]`
  `/* analytical update rule for θ1 (classical SAEM) */`
  Set $S_1^{[k]} \leftarrow S_1^{[k-1]} + \rho^{[k]} \cdot \big[S_1(z^{[k]}) - S_1^{[k-1]}\big]$. `// eq. 14`
  Set $\theta_1^{[k]} \leftarrow \theta_1^\star(S^{[k]})$ `// eqs. 15 - 18`
  `/* gradient-descent-based update heuristic for θ2 */`
  Compute $\theta_2^\star = \mathrm{argmax}_{\theta_2} p(y, z^{[k]}, \theta_1^{[k]}, \theta_2\,; t)$ by GD. `// alg. 1`
  Set $\theta_2^{[k]} \leftarrow \theta_2^{[k-1]} + \rho^{[k]} \cdot \Big[\theta_2^\star - \theta_2^{[k-1]}\Big]$. `// heuristic`
until convergence

4.3.2 Implementation details

The sequence of $\rho^{[k]}$ evoked in Algorithm 2 is chosen to be constantly equal to 1 in a first memory-free phase of the calibration procedure, and then geometrically decreases with the iterations. The fanning numerical scheme is used to compute the parallel transport along geodesics in a scalable manner [37, 38, 55]. A block Metropolis-Hasting-within-Gibbs approach is used for the MCMC sampling step, where each variable α_i, τ_i and s_i are successively sampled. Several transition kernels can be chained in order to decrease the correlation between $z^{[k-1]}$ and $z^{[k]}$. Proposal variances are dynamically tuned along the estimation in a vanishing manner, ensuring in practice acceptation rates close to 30% [6]. The optimization problem for the update of θ_2 is solved by steepest gradient descent. The gradients of the highly non-linear optimized function are obtained thanks to the autodifferentiation library offered by the PyTorch project. The PyKeops library [11] solved the key issue of memory overflows encountered when naively using autodifferentiation packages with convolution-intensive programs.

4.3.3 Initialization pipeline for model calibration

A good choice of initial parameters $\theta^{[0]}$ and latent variables $z^{[0]}$ improves the convergence speed of the calibration algorithm. We propose in this section an initialization pipeline that combines several elementary shape analysis tools. Given a longitudinal data set of shapes $(t_{i,j}, y_{i,j})_{i,j}$:

1. estimate a Bayesian atlas model (see [24]) from the baseline shapes $(y_{i,1})_i$, to get an approximate population-level mean geometry y_0', c_0' as well as n space-shift momenta m_i' mapping this geometry to the baseline observations, and an estimate of the noise level σ_ϵ';
2. for $i = 1, ..., n$, estimate a geodesic regression model (see [22]) from the individual time-series $(y_{i,j})_j$, then parallel transport (see [37]) the computed individual initial momenta back to the mean geometry y_0', c_0' along the corresponding space-shift m_i', and finally compute the Euclidean average of those w_i' to get an approximate population-level mean momenta $m_0' = \langle w_i' \rangle_i$;
3. for $i = 1, ..., n$, initialize the individual temporal parameters with $\tau_i = \langle t_{i,j} \rangle_j$, $\alpha_i^2 = \frac{w_i' \cdot G_{c_0'} \cdot m_0'}{m_0' \cdot G_{c_0'} \cdot m_0'}$ if this value is positive and $\alpha_i = 1$ otherwise, then compute σ_τ' and σ_α' according to equations 16 and 17 respectively;
4. solve a standard ICA problem with q components from the collection of space-shift momenta $w'_{i,m_0'^\perp}$ preliminarily projected on the orthogonal space to m_0', and set A_0' as the estimated mixing matrix;
5. shoot forward the mean geometry y_0', c_0' in the direction m_0' with length $t_0'' - t_0'$ where $t_0' = \langle t_{i,1} \rangle_i$ and $t_0'' = \langle t_{i,j} \rangle_{i,j}$ to get longitudinally centered estimates c_0'', y_0'', m_0'', and parallel-transport the q columns of A_0' along the same geodesic to obtain A_0'';
6. personalize the model given by the initial parameters $\theta^{[0]} = (y_0'', c_0'', m_0'', A_0'', t_0'', \sigma_\tau', \sigma_\alpha', \sigma_\epsilon')$ to obtain $z^{[0]}$.

4.4 Personalization

Once a mean progression pattern is learned by calibrating the model (11) on a data set y, the natural following question is: how each individual differ from this population average? Personalization answers this question, by estimating low-dimensional individual parameters $z = (z_i)_i$ that parametrize the best spatio-temporal transformations of the mean model onto the individual sampled progressions. Solving the personalization problem (13) is a relatively easy optimization task since all individuals can be treated independently: equation (13) is equivalent to the $i = 1, ..., n$ subproblems $z_i^{\mathrm{m}} = \mathrm{argmax}_{z_i} \log p(y_i, z_i, \theta^{\mathrm{m}}; t_i)$. We use the L-BFGS gradient-based method for this purpose [34]. Required gradients are automatically computed thanks to the PyTorch autodifferentiation library.

4.5 Simulation

From a longitudinal data set of shapes (y, t), the calibration and personalization algorithms estimated a mean progression model and individual warps, respectively encoded by the parameters θ^{m} and $z^{\mathrm{m}} = (z_i^{\mathrm{m}})_i$. We note $\widetilde{p}(z^{\mathrm{m}}, t)$ the empirical joint distribution of those individual parameters and of the visit ages t. The simulation task consists in generating a new data set y^{s} that resembles the original data set y according to some statistic function ζ (most often not sufficient, similarly to [40]):

$$\zeta(y^{\mathrm{s}}) \approx \zeta(y), \text{ with } \begin{cases} y^{\mathrm{s}} \overset{\text{iid}}{\sim} p(y | z^{\mathrm{s}}, \theta^{\mathrm{m}}; t^{\mathrm{s}}) \\ (t^{\mathrm{s}}, z^{\mathrm{s}}) \overset{\text{iid}}{\sim} \widetilde{p}(t, z^{\mathrm{m}}) \end{cases} . \tag{19}$$

For visualization purposes, the experimenter can choose to ignore the calibrated variance of noise $\sigma_\epsilon^{\mathrm{m}}$ and replace it with the degenerated value $\sigma_\epsilon = 0$. This paradigm of simulation generates smoother shapes, and will be called simulation "without noise".

5 Experiments

5.1 Validation on synthetic 2D data

In this section, the calibration, personalization and simulation algorithms are validated on a synthetic data set. The simulation algorithm is actually first used in Section 5.1.1 to generate this synthetic data from a chosen ground truth model. Its ability to reproduce this original distribution of observations without knowledge of the true parameters is

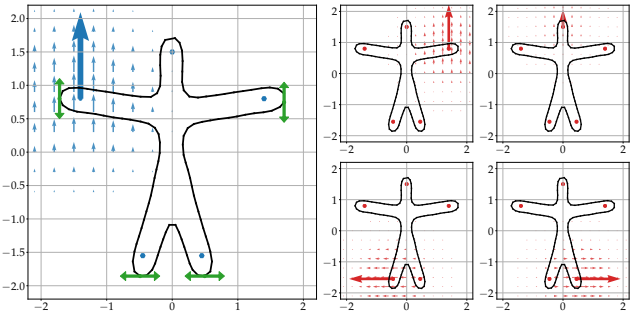


Fig. 3: Visualization of the parameters $\theta_1 = (y_0, c_0, m_0, A_0)$. The template shape y_0 is in solid black, the control points c_0 are the five dot points in either blue or red, the momenta m_0 is the bold blue arrow, and the four columns of A_0 are the bold red arrows. The velocity fields corresponding to the momenta m_0 or the geometrical components of A_0 are respectively represented with light blue or light red arrows. The green dots and arrows on the top figure mark the four landmark positions that will be considered for the statistic ζ, in order to validate the simulation algorithm in Section 5.1.4.

evaluated in Section 5.1.4. The performance and the stability of the calibration algorithm is evaluated in varied settings in Section 5.1.2. Those calibrated models are personalized in Section 5.1.3, and the learned individual parameters are compared to the true values.

5.1.1 Simulating synthetic data from a ground truth model

Fully specifying a normal distribution of shape trajectories in the sense proposed by this paper simply amounts to choose the model fixed effects $\theta = (y_0, c_0, m_0, A_0, t_0, \sigma_\tau, \sigma_\alpha, \sigma_\epsilon)$. The chosen geometrical parameters y_0, c_0, m_0, A_0 are exposed in Figure 3. In addition, we choose $t_0 = 0$, $\sigma_\tau = 2$, $\sigma_\alpha = 0.2$ and $\sigma_\epsilon \in \{0.00, 0.01, 0.02, 0.03, 0.05\}$. A total of $n \in \{50, 100, 200\}$ individual trajectories are simulated and are sampled at the corresponding times $(t_{i,j})_{j=1}^{n_i}$. We choose to draw the number of visits $n_i \geq 2$ according to a shifted Poisson law, entirely determined by its expectancy $E(n_i) \in \{3, 5, 7, 9\}$. We finally impose that the individual times $(t_{i,j})_j$ are uniformly distributed in the observation interval $[t_{i,1}, t_{i,n_i}] = [t_{i,0} - \Delta t_i/2, t_{i,0} + \Delta t_i/2]$, where both

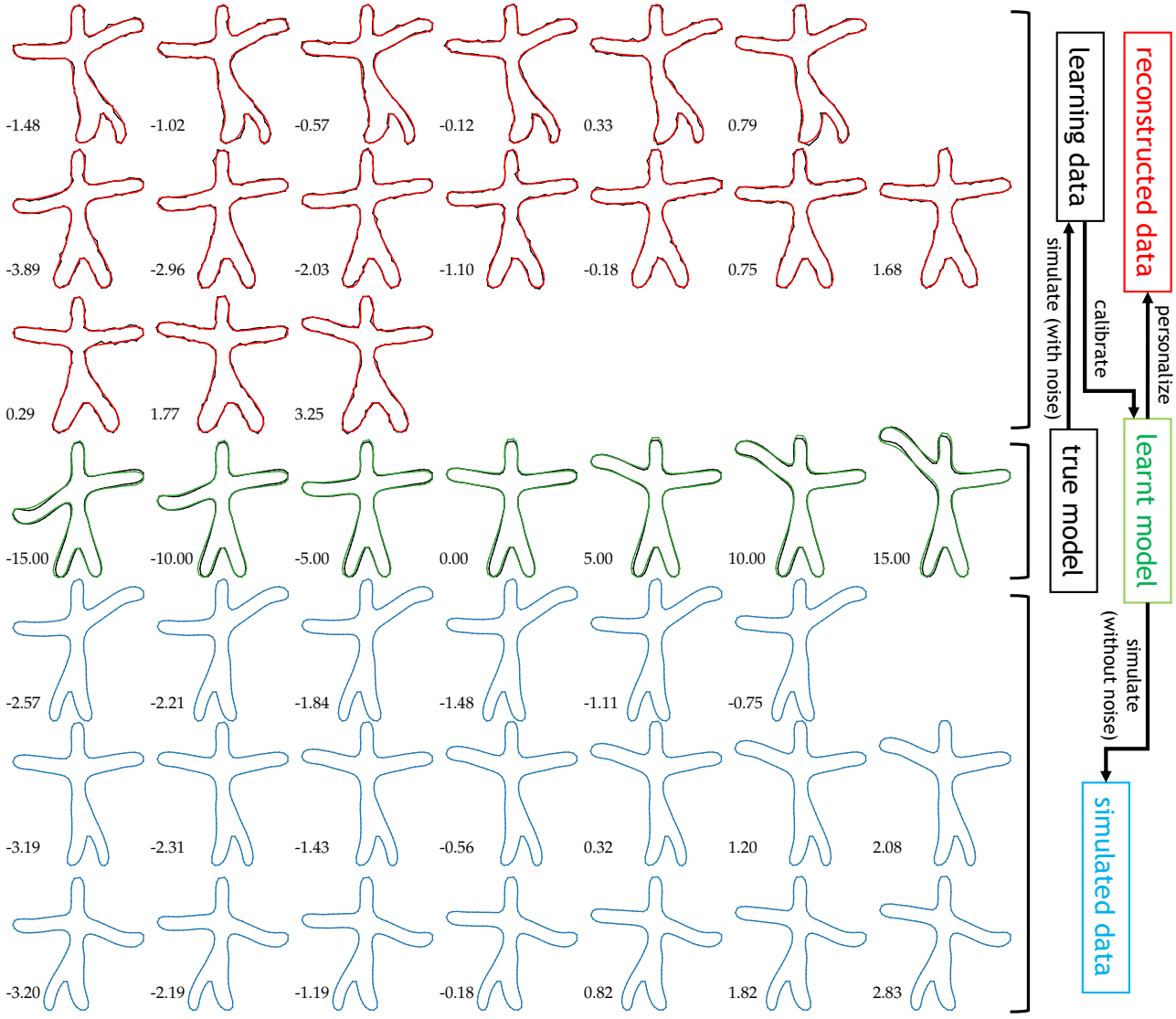


Fig. 4: Illustration of the evaluation procedure for model calibration, and subsequent personalization or simulation from the learned model. The population geodesic of MAP model is plotted in black on the central line. From this ground truth model are simulated $n = 100$ individual spatio-temporal trajectories: three randomly-picked samples are plotted in black on the top lines. The population geodesic of the calibrated model is plotted in green on the central line, and is almost perfectly superimposed with the MAP geodesic. This calibrated model can then be personalized to the training observations as plotted in red, or leveraged to simulate new spatio-temporal trajectories that resemble the original data set.

the observation time window $\Delta t_i = t_{i,n_j} - t_{i,1}$ and the baseline age $t_{i,1}$ are drawn according to normal distributions: $\Delta t_i \overset{\text{iid}}{\sim} \mathcal{N}\big(E(n_i) - 2, \sigma_\tau^2\big)$ and $t_{i,0} \overset{\text{iid}}{\sim} \mathcal{N}(t_0, \sigma_\tau^2)$. Figure 4 displays some generated data points, in the reference case where $\sigma_\epsilon = 0.02$ and $E(n_i) = 7$.

5.1.2 Model calibration

The model calibration outputs are the estimated population parameters $\theta = (\theta_1, \theta_2)$ with $\theta_1 = (t_0, \sigma_\tau, \sigma_\alpha, \sigma_\epsilon)$ and $\theta_2 = (y_0, c_0, m_0, A_0)$, and are expected to be reasonably close to the MAP estimate θ^{m} defined by equation (12).

Computing the MAP estimate. In the case of model calibration, the MAP estimate θ^{m} does not correspond exactly to the ground truth parameters. The MAP is computed easily for θ_1 with equations (15-18), and is approximated for θ_2 with a steepest gradient descent approach, initialized at the true parameters (see Algorithm 2). The MAP parameters will be called the "true" ones in the following.

Normalized error metrics. The error metric for the scalar parameters $\theta_1 = (t_0, \sigma_\tau, \sigma_\alpha, \sigma_\epsilon)$ simply is the absolute difference. The error metric on t_0 is normalized by the characteristic population observation window, that we define as $2 \cdot (1 + \sigma_\alpha) \cdot [E(\Delta t_i)/2 + \sigma_\tau]$. The remaining metrics on $\sigma_\tau, \sigma_\alpha, \sigma_\epsilon$ are respectively normalized by the true standard deviations $\sigma_\tau = 2$, $\sigma_\alpha = 0.2$, and the estimated noise level by the Bayesian atlas model [24] computed during the initialization pipeline described in Section 4.3.3. The error on the template shape y_0 is also quite simply assessed as the maximum point-to-point residual distance, and normalized by the conservative value of 3 spatial units for the characteristic size of the considered shape (see Figure 3). The control points c_0 and the momenta m_0 are jointly evaluated through the ℓ^2 distance of the estimated velocity field $v_0 = \text{Conv}(c_0, m_0)$ to the true value, normalized by the ℓ^2 norm of this true velocity field. Finally, the convergence of modulation matrix A_0 is assessed by computing the spectral norm of the composition of the projector matrices over the sub-spaces defined by the components of the estimated and true modulation matrix minus the true projector, normalized by the largest eigenvalue of the true projector matrix. The estimated modulation matrix is re-projected on the true control points in a pre-processing phase. This error metric captures the mismatch between the sets of space-shifts that can be generated from the pair of parameters (c_0, A_0).

Evaluation scenarii and results. In addition to the previously introduced scenarii, configurations with varying allowed number of geometrical components $q \in \{2, 4, 6\}$ are also evaluated. We call the configuration with $\sigma_\epsilon = 0.02$, $q = 4$, $n = 100$, and $E(n_i) = 7$ the reference one. Augmented with the 11 configurations differing from this reference by a single parameter, a total of 12 calibration problems are solved 10 times each by running the stochastic MCMC-

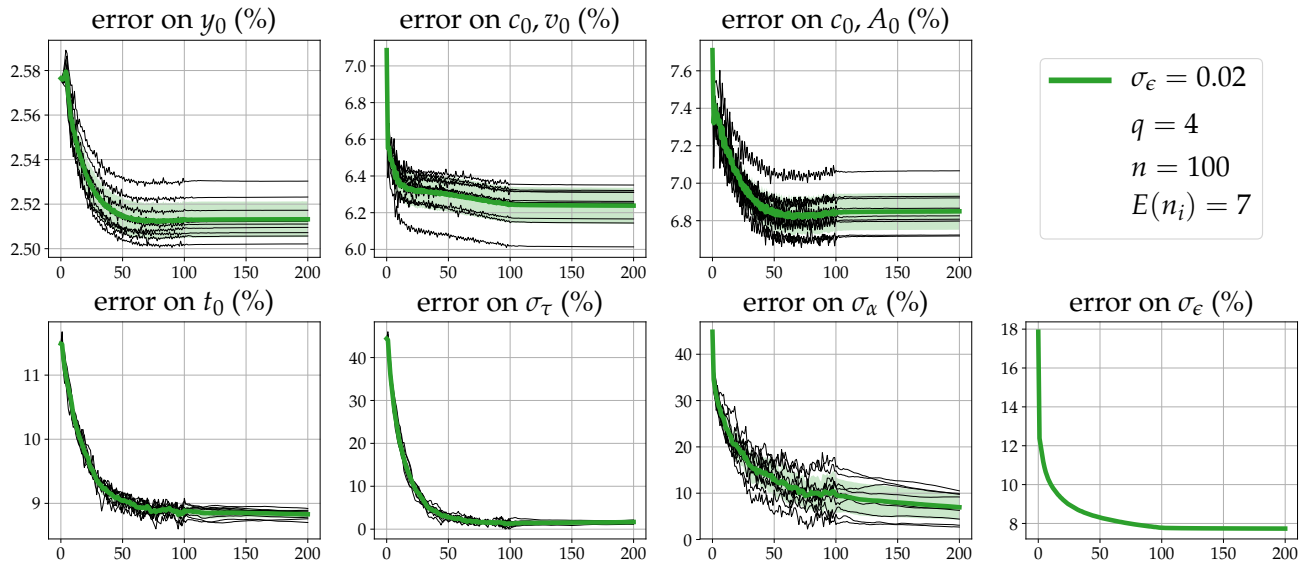


Fig. 5: Evolution of the error metrics across the 200 allowed iterations of the MCMC-SAEM algorithm, for the reference configuration: noise standard deviation $\sigma_\epsilon = 0.02$, $q = 4$ estimated components of geometrical variability, learning on a data set composed of $n = 100$ with on average $E(n_i) = 7$ visits per subject, spanning 5 time units. The 10 solid black curves correspond to 10 independent runs of the same – stochastic – MCMC-SAEM algorithm; the bold green curve is their average and the light green region indicates the associated standard deviation. The algorithm consistently converges towards very similar parameters at each run, and those estimated parameters are satisfyingly close to the MAP estimate.

	y_0 (%)	c_0, m_0 (%)	c_0, A_0 (%)	t_0 (%)	σ_τ (%)	σ_α (%)	σ_ϵ (%)
reference	2.5 ± 0.01	6.2 ± 0.10	6.9 ± 0.10	8.8 ± 0.06	1.7 ± 0.28	7.0 ± 2.73	7.7 ± 0.01
$\sigma_\epsilon = 0.00$	2.7 ± 0.01	5.9 ± 0.17	7.3 ± 0.19	7.4 ± 0.24	2.6 ± 0.60	8.1 ± 5.24	36.4 ± 0.03
$\sigma_\epsilon = 0.01$	6.0 ± 0.01	2.5 ± 0.09	4.3 ± 0.05	2.7 ± 0.07	0.7 ± 0.21	0.8 ± 0.44	15.1 ± 0.03
$\sigma_\epsilon = 0.02$	2.5 ± 0.01	6.2 ± 0.10	6.9 ± 0.10	8.8 ± 0.06	1.7 ± 0.28	7.0 ± 2.73	7.7 ± 0.01
$\sigma_\epsilon = 0.03$	1.4 ± 0.01	1.9 ± 0.09	3.7 ± 0.12	4.8 ± 0.08	2.3 ± 0.38	3.8 ± 1.23	5.1 ± 0.01
$\sigma_\epsilon = 0.05$	1.8 ± 0.02	4.1 ± 0.21	5.0 ± 0.16	1.5 ± 0.10	1.1 ± 0.34	0.9 ± 0.33	1.7 ± 0.01
$E(n_i) = 3$	1.6 ± 0.02	5.7 ± 0.29	6.2 ± 0.20	5.2 ± 0.22	3.9 ± 1.12	7.1 ± 2.03	6.5 ± 0.03
$E(n_i) = 5$	4.4 ± 0.01	5.9 ± 0.32	8.2 ± 0.13	1.1 ± 0.15	5.8 ± 0.56	0.7 ± 0.36	6.7 ± 0.02
$E(n_i) = 7$	2.5 ± 0.01	6.2 ± 0.10	6.9 ± 0.10	8.8 ± 0.06	1.7 ± 0.28	7.0 ± 2.73	7.7 ± 0.01
$E(n_i) = 9$	2.4 ± 0.01	3.2 ± 0.05	5.1 ± 0.13	7.2 ± 0.03	1.8 ± 0.11	2.7 ± 0.17	6.3 ± 0.01
$q = 2$	2.5 ± 0.04	18.0 ± 0.28	100.0 ± 0.00	10.6 ± 0.16	9.3 ± 0.52	3.1 ± 0.84	172.4 ± 0.07
$q = 4$	2.5 ± 0.01	6.2 ± 0.10	6.9 ± 0.10	8.8 ± 0.06	1.7 ± 0.28	7.0 ± 2.73	7.7 ± 0.01
$q = 6$	2.5 ± 0.01	6.8 ± 0.10	5.8 ± 0.14	8.9 ± 0.07	1.7 ± 0.20	9.3 ± 1.80	7.0 ± 0.02
$n = 50$	3.4 ± 0.01	3.6 ± 0.16	6.7 ± 0.14	4.7 ± 0.04	0.2 ± 0.16	2.0 ± 0.17	7.6 ± 0.02
$n = 100$	2.5 ± 0.01	6.2 ± 0.10	6.9 ± 0.10	8.8 ± 0.06	1.7 ± 0.28	7.0 ± 2.73	7.7 ± 0.01
$n = 200$	1.8 ± 0.00	3.9 ± 0.08	4.2 ± 0.08	2.6 ± 0.04	0.8 ± 0.12	3.7 ± 0.19	6.9 ± 0.01

Table 1: Final average normalized performance metrics and associated standard deviations, obtained after 10 independent runs of the MCMC-SAEM algorithm in varied configurations. The algorithm is run for 200 iterations in all configurations. The reference configuration corresponds to a noise level $\sigma_\epsilon = 0.02$, an average number of visits per subject $E(n_i) = 7$, $q = 4$ allowed components of geometrical variability, and $n = 100$ input subjects.

SAEM-GD algorithm. Figure 5 plots the evolution of the error metrics across the allowed 200 iterations for the reference configuration: the black lines correspond to the 10 different runs, and in green is represented their mean and standard deviation. The algorithm is stable, and converges to very similar results at each run. Those results are close to the MAP estimate. Table 1 gives the average final error metrics and the associated standard deviation for varying true noise level σ_ϵ configurations. A first comment is again that the algorithm is very stable, as demonstrated by the relatively small standard deviations. Interestingly, a higher noise levels does not necessarily correlate with a degraded estimation of the parameters for this calibration task: the data noise seems to actually help the algorithm to explore more of the parameters domain. The unique clear pattern is that the estimated noise variance σ_ϵ gets relatively closer to the true value when this one increases. In all configurations, the MAP parameters are satisfyingly estimated. Table 1 studies as well the impact of the window of observation length $E(n_i)$. Long periods of observation are favorable to estimate the longitudinal-related parameters, i.e. v_0 which encodes the population progression, σ_τ and σ_α which capture the dynamical variability. The same table compares as well the estimation quality when the true number of sources is underestimated ($q=2$), perfectly chosen ($q=4$) or overestimated ($q=6$). The reconstruction ability of our model, measured by σ_ϵ, increases with q, and seems to saturate once the optimal number of sources has been reached. Parameters are in general less well estimated in the $q=2$ configuration, and of similar quality in the two remaining ones. Finally, Table 1 shows that the number of subjects over which the model is learned n has a major influence over the quality of the estimation. Almost all metrics are improved in the configuration with $n=200$ subjects.

In conclusion, the proposed MCMC-SAEM-GD algorithm successfully solves our model calibration problem in varied configurations. The stochastic procedure is stable across independent repetitions, converging to always very similar points. The presence of noise in the training data is well-handled, and actually seems to act as a good regularizer for the estimation procedure. An underestimated number of sources does not harm the convergence of the procedure, but mostly impairs the reconstruction ability of the learned model. This number should therefore be gradually increased to meet the reconstruction goals of the experimenter, keeping in mind that an intrinsic optimal performance will be reached when q is large enough. Finally, as far the data set size is concerned, increases in the number of subjects or in the number of visits both seem to be beneficial for model calibration, in comparable proportions.

5.1.3 Personalization after calibration

Once calibrated, the longitudinal shape models are personalized to the training data. The estimated individual parameters α_i, τ_i, s_i are compared to their true value. In order to be comparable with the true sources, the estimated sources are

first brought back to the cotangent space defined by the true control points c_0.

Figure 6 plots the estimated z_i against the true corresponding values. The acceleration factors are almost perfectly estimated, when the onset ages and sources are also estimated with a low variance, but at the same time with a non-negligible bias. In the case of the onset ages, this individual parameter compensates by this bias the error committed on the reference time t_0 during calibration. More generally, all three individual parameters might present biases when estimated after calibrating the model, because they compensate residual errors committed by the calibration operation: α_i compensates for an error in the norm of v_0, and s_i for an error in the norm of the columns of A_0. We argue that those biases are less critical indicators of performance than the associated variances, which capture the relative uncertainty on the individual parameters. Finally, two outliers can be noticed in Figure 2 for the pace of progression α_i, as well as for the onset age τ_i, although in reduced proportions. These outliers correspond to extremely reduced windows of observation $t_{i,n_i} - t_{i,1}$, respectively equal to 0.07 and 0.20, when the theoretical mean is equal to 5.

Table 2 summarizes the results in all configurations, giving for each of the twelve considered scenarii the median error and associated median absolute deviation times 1.4826 (which is a robust estimator of the standard deviation for Gaussian distributions) on $z_i = (\alpha_i, \tau_i, s_i)$ when personalizing the average calibrated model. Focusing on the estimation variability, it appears that the sources s_i are the best estimated parameters, followed by the onset ages τ_i

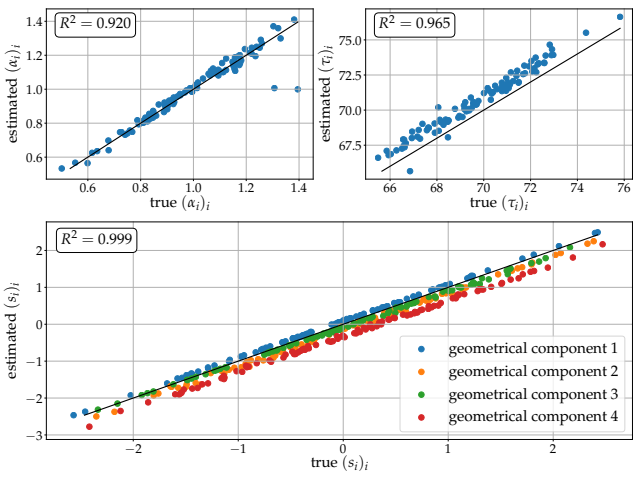


Fig. 6: Comparison of the estimated individual parameters $z_i = (\alpha_i, \tau_i, s_i)$ after personalization of the mean calibrated model to the simulated observations, in the reference scenario. In each scatter plot, the identity is represented by the solid black line. The R^2 value for the sources is an average over the four geometrical components.

	$\Delta\alpha_i$ (%)	$\Delta\tau_i$ (%)	Δs_i (%)
reference	4.4 ± 11.9	44.2 ± 15.3	-11.7 ± 3.3
$\sigma_\epsilon = 0.00$	-3.8 ± 4.3	42.4 ± 24.8	-5.4 ± 6.8
$\sigma_\epsilon = 0.01$	-2.1 ± 6.4	11.8 ± 9.4	-11.1 ± 12.0
$\sigma_\epsilon = 0.02$	4.4 ± 11.9	44.2 ± 15.3	-11.7 ± 3.3
$\sigma_\epsilon = 0.03$	-7.6 ± 15.7	25.5 ± 15.0	-6.7 ± 2.5
$\sigma_\epsilon = 0.05$	-21.8 ± 23.4	4.4 ± 14.7	-9.1 ± 3.0
$E(n_i) = 3$	9.3 ± 32.7	-12.1 ± 19.3	-8.8 ± 3.2
$E(n_i) = 5$	2.9 ± 13.0	3.2 ± 20.8	-3.6 ± 7.5
$E(n_i) = 7$	4.4 ± 11.9	44.2 ± 15.3	-11.7 ± 3.3
$E(n_i) = 9$	-8.2 ± 7.3	46.0 ± 16.1	-7.6 ± 2.3
$q = 2$	-1.6 ± 18.0	48.3 ± 50.7	-2.4 ± 69.4
$q = 4$	4.4 ± 11.9	44.2 ± 15.3	-11.7 ± 3.3
$q = 6$	9.8 ± 12.3	45.3 ± 15.8	-11.7 ± 3.3
$n = 50$	-12.9 ± 8.1	24.8 ± 18.3	-6.9 ± 6.1
$n = 100$	4.4 ± 11.9	44.2 ± 15.3	-11.7 ± 3.3
$n = 200$	9.8 ± 10.5	13.2 ± 10.9	-13.1 ± 2.4

Table 2: Median of the residual errors and associated median absolute deviation times 1.4826 for the estimated individual parameters, expressed in percentage of the corresponding ground-truth standard deviations $\sigma_\alpha = 0.2$, $\sigma_\alpha = 2$ and $\sigma_s = 1$. The results are given for the reference scenario plus eleven perturbed scenarii, where either the noise level σ_ϵ, the average number of visits per subject $E(n_i)$, the allowed number of geometrical components q or the number of subjects n is varied.

and pace of progression α_i. The estimation of the pace of progression α_i quickly deteriorates with increasing levels of noise σ_ϵ, reaching almost 25% of the true standard deviation $\sigma_\alpha = 0.2$ in the most noisy configuration. The estimation of the onset ages τ_i and sources seems more robust, with no clear tendency. An increased number of visits per subject $E(n_i)$ seems to help frankly the estimation of the pace α_i, as well as τ_i in lesser proportions. The sources s_i remain well-estimated in all scenarii. No clear difference can be noticed between the reference $q = 4$ and the over-estimated number of geometrical components case $q = 6$, suggesting that adding components does not hampers the personalization of a calibrated model. However, underestimating this number of components with $q = 2$ largely deteriorates the estimation of the sources s_i as it could be expected, and even of the dynamical parameters α_i and τ_i in reduced proportions. Finally, an increased number of subjects n favours a better post-calibration personalization, even if this conclusion is less clear for the pace parameter α_i.

5.1.4 Simulation after calibration and personalization

After calibration and peronalization, the learned model and empirical distribution of the random effects can be leveraged to simulate new shape trajectories. Figure 4 gives some ran-

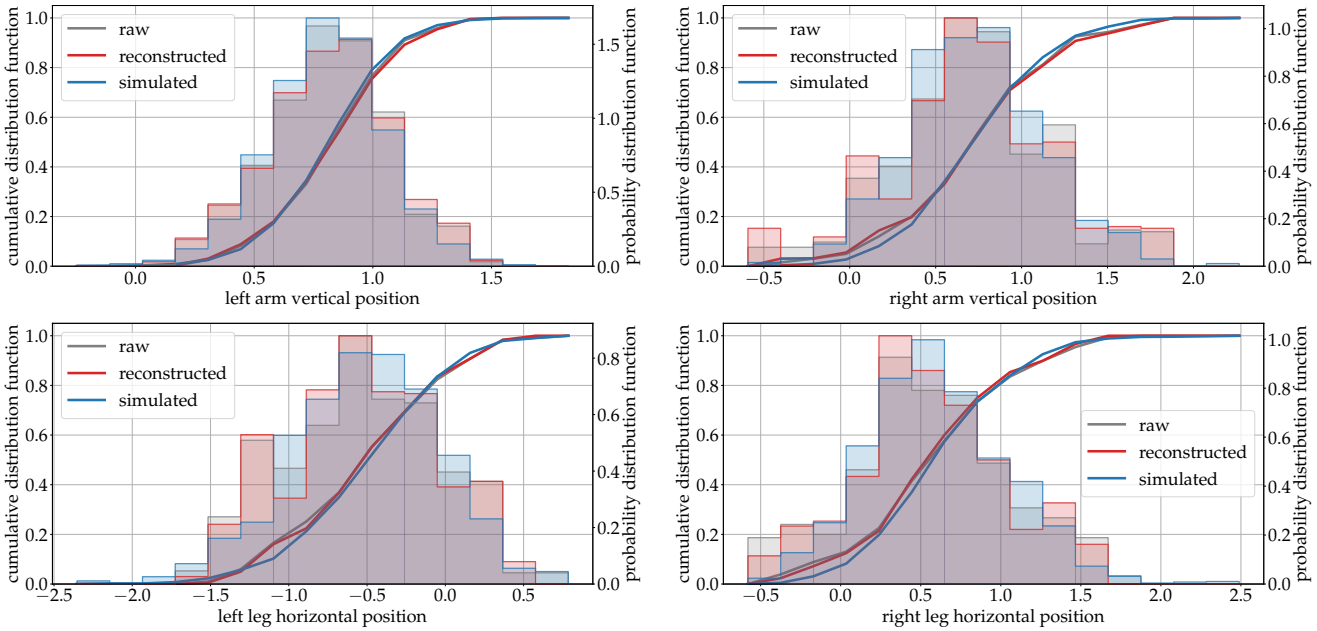


Fig. 7: Distribution of the position of landmarks of interest in the raw (i.e. original), reconstructed (by personalization of the calibrated model) and simulated data sets, for the reference configuration. Those landmarks of interest are indicated by green dots on Figure **??**. The simulated distributions are similar to the corresponding raw ones, suggesting that the spatio-temporal variability of the original data set has been successfully captured.

domly selected samples from such simulated trajectories for the reference scenario, where (see equation 19):

- the fixed effects θ^{m} are averages over the 10 calibrations;
- the random effects z^{s} are drawn according to independent normal distributions with mean and standard deviations equal to the values given by Table 2;
- the visit ages t^{s} are drawn according to the true procedure based the average calibrated values for t_0, σ_τ, and the empirical average $\langle n_i \rangle_i$ for $E(n_i)$ (see Section 5.1.1).

Figure 7 compares the distribution of vertical or horizontal positions of the tips of the raw (see Section 5.1.1), reconstructed (see Section 5.1.3) and simulated dancing ghosts. Those landmarks of interest are indicated by green dots and arrow on Figure **??**, and form the statistic ζ introduced in equation (19). A total of 1000 subjects are simulated, when only 100 were available for model calibration. The three distributions are very similar, indicating that the learned distribution of shape trajectories seems to correctly capture the true distribution, or at least some relevant modes of geometrical variability present in the initial data set. Note that the left arm vertical position is mainly time-related, when the positions of the remaining limbs are mainly related to static geometrical variability.

5.2 Dynamic facial expressiveness under emotions

5.2.1 Data and preprocessing

The Birmingham University 3D dynamic facial expression database [54] gathers short videos sequences from 101 subjects (of which 58 female, 43 male). Each subject mimics in 6 distincts sequences the basic emotions which are Anger, Disgust, Fear, Happiness, Sadness and Surprise. For each of those 606 sequences we uniformly extract 8 frames spanning from the first to the 36-th one, which correspond to a subsampling of the first 1.4 seconds of each video. We do not work directly with the texture video, but with a set of 75 semi-automatically extracted landmarks, which were readily available along with this data set. Every frame is registered to a reference one by Procrustes alignment.

5.2.2 Model calibration: learned emotion models

We learn 6 distinct longitudinal atlas models: one per emotion, calibrated on the $n = 101$ sequences of $n_i = 8$ frames for all subjects i. We choose $q = 10$ number of sources. Figure 8 plots in green the learned mean scenario for each emotion. Qualitatively, those mean scenarii seem to reasonably capture relevant geometrical patterns for each emotion: the Disgust, Fear, Happiness and Surprise models feature large displacements in the area of the mouth in particular. The Sadness expression is more mute, with a subtle displace-

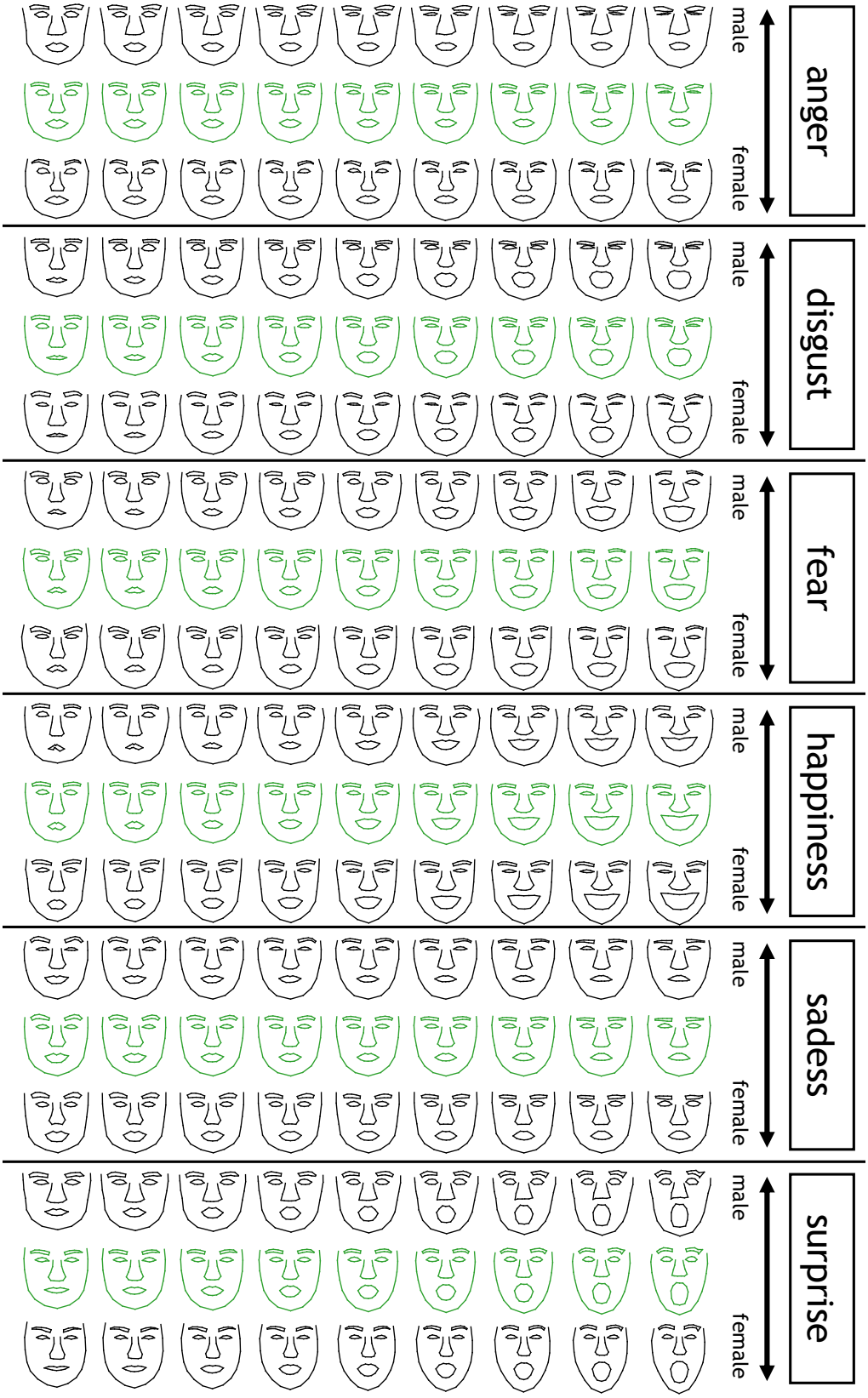


Fig. 8: Learned emotion spatio-temporal models. The population geodesic is plotted in green, and the shifted progressions along the gender mode of geometrical variability are plotted in black.

ment of the eyebrows. The Anger population model shows a combined displacement of both the eyes and the eyebrows.

5.2.3 Gender-specific emotion patterns

The learned model are personalized to the corresponding training data sets, giving for each sequence an optimal $z_i = (\alpha_i, \tau_i, s_i)$. We only focus on exploiting the individual source parameters $s_i \in \mathbb{R}^q = \mathbb{R}^{10}$ in this section: emotion-wise 1D partial least square regression models predicting the gender from s_i are fitted with the `sklearn` library. Those models project the explanatory variable (here the sources s_i) in a 1-dimensional latent space and regress the response variable (here the gender) against those projections [1]. The regression and projection parameters are jointly learned, and we finally obtain for each emotion an optimal linear combination operator on $\mathbb{R}^{10}$ that best explains the gender differences from the projected sources. Those differences are significant for all emotion models, with p-values always below 10^{-5} when testing for projected sources group differences in mean with a Student t-test.

Figure 8 gives the corresponding temporal progressions (in black). For all emotion models, the male subjects tend to have wider faces than females, as it can very clearly be seen in the area of the cheeks or of the nose for the Anger and Surprise models. An interesting difference in the position of the eyebrows can be noticed in the case of the Disgust model, and similarly in the opening of the eyes for the Sadness model.

5.2.4 Application to classification

We propose to automatically recognize the emotion that a sequence is representing based on the dynamic coefficients α_i, τ_i that are learned by personalizing an emotion model to the sequence. We perform a 5-fold cross-validation: the 101 subjects are split into 5 groups, balanced for gender. For each split:

- a total of 6 longitudinal shape models are learned on the training sequences for each emotion;
- those models are personalized to all the 606 sequences: for each sequence a total of 6 z_i vectors are therefore estimated;
- for each sequence, the learned temporal parameters α_i, τ_i are stacked into vectors of $6 \times 2 = 12$ scalars;
- those feature vectors are then leveraged to train and test a simple linear discriminant classifier on the corresponding train and test sequences.

Table 3 gives the confusion matrix obtained with this procedure, averaged over the 5 folds. The average classification accuracy is $67.08\,\%$, above the chance level which amounts to $16.67\,\%$. For comparison, [5] reported an average accuracy of almost $100\,\%$, [51] of $90.44\,\%$, and [19] of $74.63\,\%$. We emphasize however that our performance is achieved:

- using the default linear discriminant analysis from the `sklearn` library, without any hyperparameter tuning as in [5] with random forest, in [51] with hidden markov model or in [19] with radial support vector machine;
- on all the 606 available sequences, without any manual selection of a subset of 60/101 subjects as it is done in [5, 51] or of 507/606 sequences as done in [19];
- based only on 12 intuitive scalar features per sequence.

Qualitatively, we conclude that the relative dynamics of facial emotions in terms of time-lag and pace of expression extracted by calibrating and personalizing longitudinal shape models contain highly discriminative information for emotion recognition.

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	64.3	7.0	8.1	4.0	16.6	-
Disgust	13.7	55.1	12.4	14.8	1.9	2.0
Fear	1.0	16.6	58.6	13.9	7.0	3.0
Happy	1.9	6.0	13.0	79.1	-	-
Sad	16.5	2.0	14.2	1.1	66.2	-
Surprise	1.0	3.0	16.0	-	1.0	79.1

Table 3: Average confusion matrix across 5-fold linear discriminant classification. The sequence features consist in a 12-scalar vector that stacks the 6 pairs of dynamical parameters α_i, τ_i obtained by personalizing the 6 emotion models. The average accuracy is $67.08\,\%$.

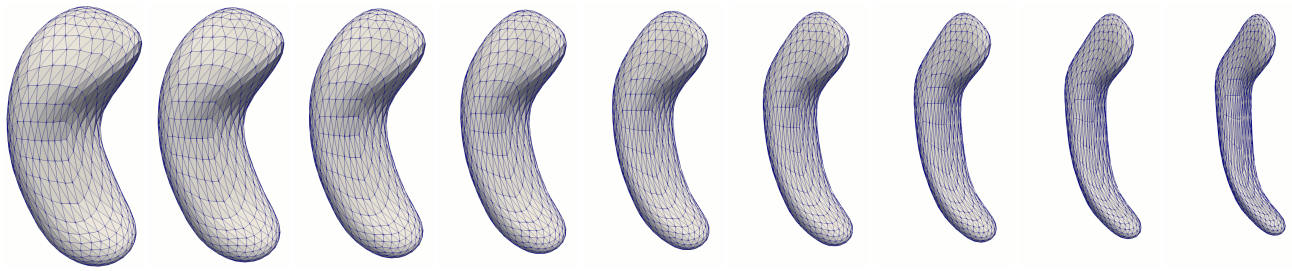
5.3 Hippocampal atrophy in Alzheimer's disease

5.3.1 Data and preprocessing

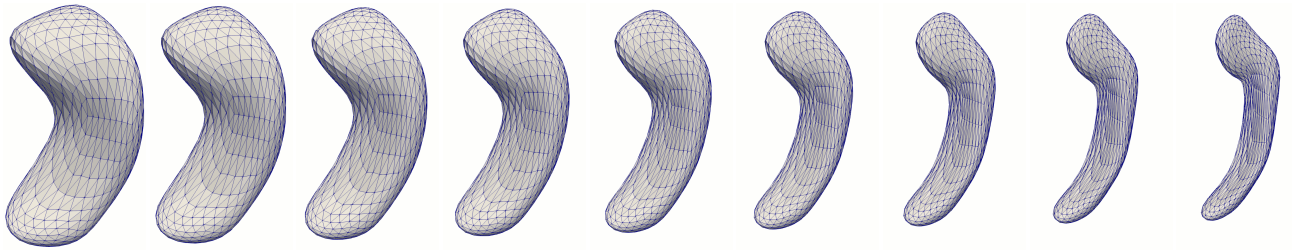
Data used in the preparation of this section were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003

Number of subjects	322
Number of visits	1993
Average number of visits per subject ($\pm$ std)	5.8 ($\pm$ 2.4)
Average age ($\pm$ std)	74.0 ($\pm$ 6.7)
Sex ratio (F/M in %)	41.2 / 58.8
Amyloid status (+/-/unknown in %)	73.2 / 7.1 / 19.7
APOE carriership (%)	65.2
Education (mean $\pm$ std, in years)	15.9 ($\pm$ 2.8)
Marital status (married/not married in %)	80.9 / 19.1

Table 4: Summary statistics of the medical data set of Alzheimer's disease patients.



(a) Left hippocampus mean progression.



(b) Right hippocampus mean progression.

Fig. 9: Learned mean progression of the left and right hippocampi shapes after calibration of the shape models to the data extracted from 1993 visits of $n = 322$ patients that eventually converted to Alzheimer's disease after a phase of mild cognitive impairments. Ages (from left to right, in years): 58.6, 63.0, 67.4, 71.8, 76.2, 80.7, 85.1, 89.5, 93.9.

as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).

We select all the T1-weighted MRIs of subjects that were diagnosed as presenting mild cognitive impairements at some visit, and diagnosed as converted to Alzheimer's disease at some later visit. See Table 4 for summary statistics. This data set amounts to a total of 1993 visits from $n = 322$ subjects. Second-take "re-test" MR image are available for 1838 of those visits and will be used to measure the intrinsic data noise. All those $1993 + 1838 = 3831$ images are pre-processed exactly in the same manner, starting with the longitudinal pipeline of FreeSurfer[1] (version 5.3.0) [20, 21]. The skull-stripped brains are then aligned with an affine 12-degrees-of-freedom transformation onto the Colin27 average brain[2], with FSL 5.0[3] [53]. Mesh representations of the geometry of the left and right hippocampi are obtained from the original images as follow:

- the volumetric segmentations of the hippocampi computed with FreeSurfer are transformed into meshes using the aseg2srf script of July 2009[4],
- the resulting meshes are decimated by a 88% factor using Paraview 5.4.1[5] [2],
- then aligned using the previously-computed global affine transformation estimated with the FSL software,
- residual pose differences among subjects are then removed by rigidly aligning the meshes from the baseline image of each subject to the corresponding hippocampus mesh in the Colin27 atlas image, this transformation with 6 degrees of freedom being computed with the GMMReg script of June 2008[6] [28],
- the same transformation is finally used to align the meshes from the follow-up images of the same subject.

5.3.2 Calibrated hippocampal atrophy models

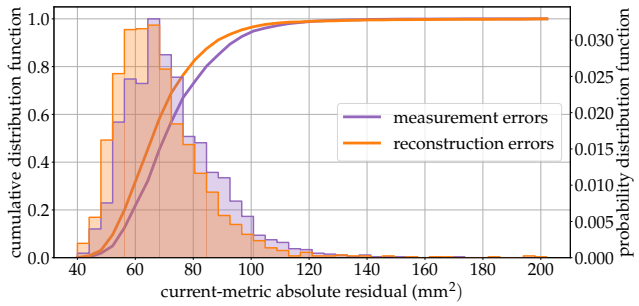
We calibrate two longitudinal shape models on all the 1993 meshes of the left and right hippocampi respectively, choosing in both cases $q = 8$ components of geometrical variability. The chosen deformation kernel width is $\sigma = 10\,mm$. The current noise model [52] is leveraged to compute distances between meshes without point correspondence, with a kernel width of $\sigma_{\mathcal{E}} = 5\,mm$.

Figure 9 plots the estimated average progression. The learned mean progression consists in an atrophy of both the left and right hippocampi, in line with current medical knowledge about this subcortical structure, which is known to be highly involved into memory cognitive tasks.

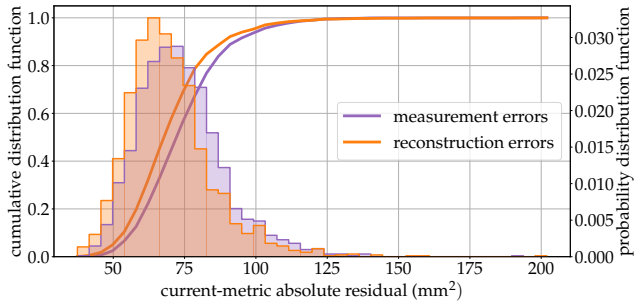
[1] available at: https://surfer.nmr.mgh.harvard.edu
[2] available at: http://www.bic.mni.mcgill.ca/ServicesAtlases/Colin27
[3] available at: https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/
[4] available at: https://brainder.org
[5] available at: www.paraview.org
[6] available at: https://github.com/bing-jian/gmmreg

5.3.3 Personalization to unseen data

We assess the reconstruction ability of the calibrated models in a 5-fold cross-validation fashion. The $n = 322$ subjects are split into 5 groups; 2×5 distinct shape models are calibrated on the so-defined training sets for the left and right hippocampi structures. Those models are then personalized to the unseen test subjects. The distribution of the residual discrepancies, measured thanks to the learning current metric with $\sigma_{\mathcal{E}} = 5\,mm$, is compared to the intrinsic noise in the learning data set, assessed by comparing with the same current metric the regular and "re-test" meshes. Figure 10 plots the obtained distributions of the current-metric absolute residuals. We read that the reconstruction errors are on average inferior to the intrinsic uncertainty on the data, and with a lower variance as well. We argue therefore that our modeling approach allow to correctly represent the progression of the hippocampi of an individual developing Alzheimer's disease with a reduced set of 2×10 scalars, which are for each hippocampus the pace of progression α_i, the onset age τ_i, and the eight geometrical sources s_i.



(a) Left hippocampus. The mean error is $68.5 \pm 15.9\,mm^2$ for the shape model, and $83.2 \pm 36.0\,mm^2$ for the re-test measurement.



(b) Right hippocampus. The mean error is $69.8 \pm 15.0\,mm^2$ for the shape model, and $85.2 \pm 40.1\,mm^2$ for the re-test measurement.

Fig. 10: Comparison of the generalization error to unseen data of the learned shape models and the intrinsic measurement error. The discrepancies between meshes are computed with the current metric with $\sigma_{\mathcal{E}} = 5\,mm$, without assuming any point-to-point correspondence.

			left hippocampus		right hippocampus	
genetic	gender (female vs. male)	α_i	$\times 1.23$	[**]	$\times 1.21$	[**]
		τ_i	-12.4 months	[**]	-8.7 months	[*]
		s_i	± 0.54	[***]	± 0.57	[****]
	APOE-ϵ4 (carrier vs. non-carrier)	α_i	$\times 1.22$	[*]		
		τ_i	-35.8 months	[***]	-32.5 months	[**]
		s_i				
biological	amyloid (positive vs. negative)	α_i	$\times 1.52$	[**]	$\times 1.67$	[*]
		τ_i				
		s_i				
environmental	marital (married vs. non-married)	α_i	$\times 1.14$	[*]		
		τ_i	-42.5 months	[***]	-36.3 months	[**]
		s_i				
	education (nb. of years of education)	α_i				
		τ_i	-3.7 months/y	[**]	-5.1 months/y	[***]
		s_i				

Table 5: Significant associations of individual parameters with genetic, biological and environmental factors: effect sizes, confidence intervals at 95%, and significance levels of the adjusted p-values (thresholds 5%, 1%, 0.1%, 0.01%). Time-shifts τ_i are in months, others have no units. Directions of space-shifts are not signed. The 23 subjects (out of $n = 322$) without amyloid information have been discarded.

5.3.4 Correlations with cofactors

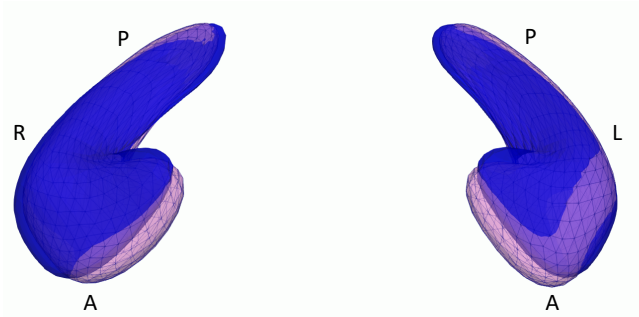
We personalize the models learned on the full 1993 visits to all the $n = 322$ learning subjects, and take an interest in correlating the learned individual parameters $z_i = (\alpha_i, \tau_i, s_i)$ with genetic (gender and APOE-ϵ4 carriership), biological (amyloid status) or environmental (education level and martial status) cofactors.

The parameters α_i and τ_i are regressed against the five considered cofactors, and two-tailed t-tests are performed on the coefficients. A 2-blocks partial least square regression (PLS) model [1] is used to regress the eight sources s_i against the five cofactors in a one-dimensional projection space. A two-tailed t-test is then performed on the coefficients of the multivariate regression of the PLS-learned linear combination of sources against the cofactors. For each case, the obtained five p-values are corrected with the false discovery rate (FDR) method.

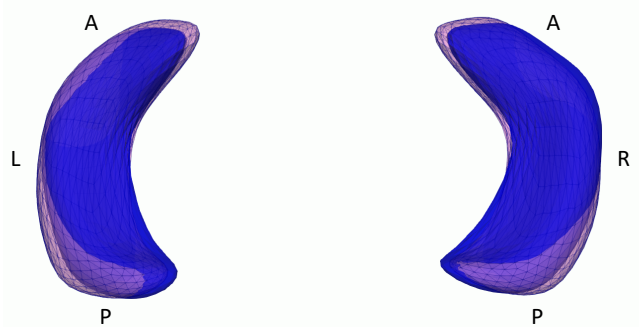
Table 5 gives the obtained results for both left and right hippocampi. We read that the atrophy of the hippocampi develops faster and starts earlier in female subjects. Male and female subjects present significantly different geometries of their hippocampi, after removal of the ageing confounding effect: Figure 11 presents the corresponding mode of geo-

metrical variability. We read that the hippocampal atrophy starts earlier in carriers of at least one APOE-ϵ4 allele, with an effect size of almost three years. The atrophy occurs at an accelerated pace in amyloid-positive subject, as well as for APOE-ϵ4 allele carriers and married subjects but only in a significant manner in the left hemisphere of the brain. Finally, the atrophy occurs earlier in married subjects, as well as in educated subjects.

The obtained results by correlating the learned parameters z_i with the genetic and biological cofactors are in line with current medical knowledge. The results obtained with respect to the marital status are more surprising, and should probably be taken with care as the non-married group, which weights a bit less than 20% of the considered 299 subjects in Table 5, is quite heterogeneous with widowed, divorced, or never married subjects pooled together. Finally, the idea that educated subjects subjects develop an earlier atrophy of their hippocampi can appear surprising at first, but is actually in line with the cognitive reserve theory [48], which supports the idea that education can help to compensate damaged brain anatomy at the clinical level, maintaining unaltered cognitive capacities for a period of time. In other words, cognitive decline would be delayed with respect to the onset of brain atrophy in educated subjects. Making now the reasonable hypothesis that the subjects presenting mild cognitive impairements enroll in the ADNI study at similar cognitive stage explains that the selected subjects present an increased atrophy of their hippocampi for an increased education: they enrolled with a more advanced stage of anatomical pathology, after some years of cognitive compensation.



(a) Coronal view.



(b) Axial view.

Fig. 11: Superposition of the male-like (in blue) and the female-like (in pink) hippocampi geometries, in two standard views. The letters L, R, A, P respectively indicate the left, right, anterior and posterior directions.

5.3.5 Simulation of new hippocampi progressions

The calibrated models and the empirical distribution of random effects z_i learned by their personalization to the training data are leveraged to simulate new progressions of the hippocampi. In order to validate such a simulation method, the simulated trajectories will be sampled at some visit ages, and the empirical distribution of the volumes of the simulated hippocampi will be compared to the distribution of the original hippocampi. This simple volume feature is commonly considered relevant in the context of Alzheimer's disease, and is a commonly-used biomarker of the disease in research studies and clinical trials. Note that in order to validate the simulation method, the distribution of the visit ages must be respected, as well as its joint distribution with the individual parameters z_i. There exist for example a strong correlation between the estimated onset age τ_i and the baseline age of the enrolled subjects $t_{i,1}$: indeed as argued in the previous section, the selected subjects tend to be enrolled in the ADNI study at similar stages of Alzheimer's disease progression. This additional difficulty is addressed by simulating a data set with exactly $n = 322$ subjects and 1993 visits, keeping the same number of visits and same time gaps between visits: only the baseline age is changed. As a preprocessing step:

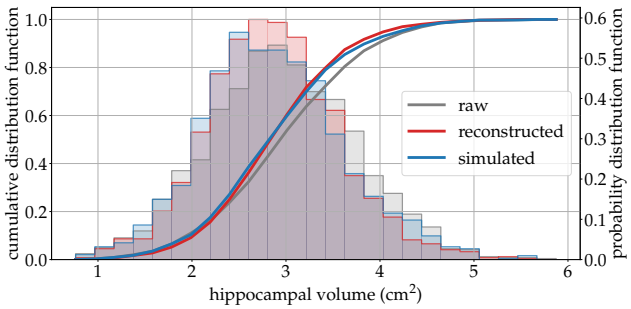
- the empirical joint distribution of the time-related parameters α_i and τ_i augmented with the age at baseline $t_{i,1}$ is carefully modeled with a kernel density;
- the empirical joint distribution of those time-related parameters augmented with the sources s_i is captured by fitting a multivariate Gaussian distribution.

A simulated data set is then created by applying 322 times the following procedure:

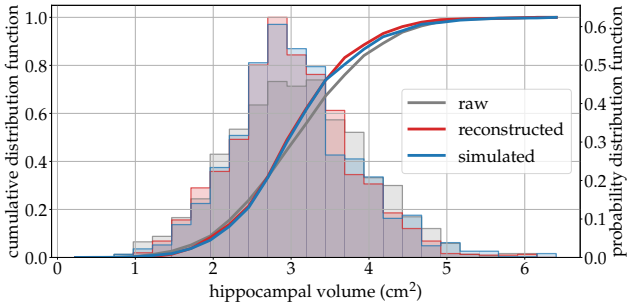
- draw the acceleration factor α_i, the onset age τ_i and the baseline age $t_{i,1}$ from the corresponding kernel density;
- draw the sources s_i from the multivariate Gaussian conditional distribution with respect to its already-drawn time-related parameters;
- draw without replacement the structure of the visits i.e. the number of visits and the time gaps between them from the pool of true visit structures;
- sample the individual hippocampus trajectory defined by $z_i = (\alpha_i, \tau_i, s_i)$ at the baseline age $t_{i,1}$ and the follow-up visits.

This protocol is repeated for both the left and right hippocampi. For each structure, this method is further ramified by applying it separately on the male and female groups: indeed this it is against this cofactor that the individual parameters are the most differently distributed, as it can be read from Table 5. Moreover this cofactor is quite well balanced with 41% of females, thus this additional split still allows to correctly estimate the multivariate Gaussian on the considered eleven parameters $(t_{i,1}, z_i)$.

Figure 12 plots the volume distributions of the raw, reconstructed and simulated data sets. The simulated distribution matches almost perfectly the reconstructed one, suggesting that the distribution of the individual parameters z_i has been correctly modeled. The residual mismatch between the raw and reconstructed distributions comes from the smoothing power of the current metric, which ignores the small segmentation artifacts. Figure 13 illustrates this: is displayed an example of data point for which the reconstruction error belongs to the lower percentile of all the 1993 reconstructions, when the absolute volume difference falls into the 79-th percentile of the corresponding distribution. Qualitatively, the small noisy outer protrusions are smoothed out when fitting the shape model using the current metric. Those protrusions are on average oriented towards the exterior of the mesh, leading to a volume bias when comparing the volume distributions of the raw meshes with the reconstructed ones. This volume bias amounts on average to 84.5 mm^2 for the left hippocampus and 67.3 mm^2 for the right one, and explains the mismatch displayed on Figure 12. As a final remark, the slopes of individual linear regressions fitting on the volumes of the raw hippocampi shapes are significantly larger for female subjects than for males, for both hemispheres. This result holds when testing on the reconstructed but also simulated data, indicating that some fundamental population-level results are preserved when generating a synthetic data set with the proposed method.



(a) Left hippocampus: the mean volume is $2958 \pm 779\, mm^3$ for raw data, $2863 \pm 693\, mm^3$ for reconstructed data, and $2865 \pm 746\, mm^3$ for simulated data.



(b) Right hippocampus: the mean volume is $3081 \pm 862\, mm^3$ for raw data, $3014 \pm 754\, mm^3$ for reconstructed data, and $3063 \pm 763\, mm^3$ for simulated data.

Fig. 12: Distribution of the left and right hippocampal volume in the raw, reconstructed and simulated data set. The simulated volume distribution is very close to the volume distribution of the reconstructed data set. The remaining bias between those two distributions and the one corresponding to the raw data comes from the smoothing behavior of the current noise model, used to deal with noisy meshes without point correspondence. See Figure 13.

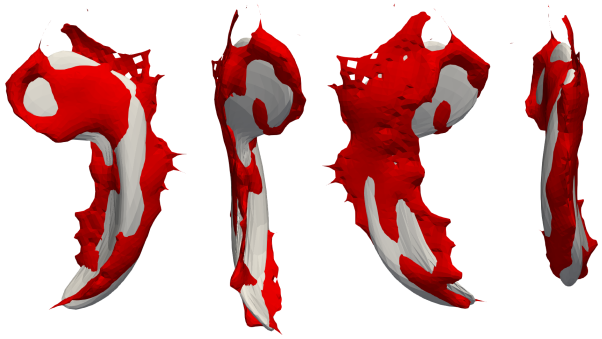


Fig. 13: Several views of a single example of the reconstruction of a right hippocampus structure by the longitudinal shape model. The reconstruction is the smooth white structure, and the raw data point is plotted in red.

Remark. Now validated, the simulation algorithm could be used to synthesize a data set of left and right hippocampi of any number of subjects, with any desired visit sampling. The proposed gender-wise split even allows to achieve any desired male-female balance.

6 Conclusion

Based on the introduced concept of shape spatio-temporal reference frame, we proposed a modeling approach that represents individual data series as continuous trajectories understood as spatio-temporal perturbations of a population average progression. The spatial warp is defined thanks to the exp-parallelization operator on manifolds, when the time warps are simple affine functions. We defined the calibration, personalization and simulation algorithms that take ad-

vantage of this statistical model to meet pragmatic objectives. The calibration algorithm originally combines the MCMC-SAEM stochastic approach with gradient descent to estimate a normative scenario of changes along with the associated spatial and temporal variability from temporally-unaligned individual sequences, which is key for applications where no common time reference frame is available. Personalizing such scenario to a new individual gives quantitative, low-dimensional and intuitive measures of abnormality in terms of an acceleration factor and a time-shift in one hand, and geometrical sources of variability in the other hand. Such individual parameters offer relevant features for classification of correlation tasks, in a post-processing phase. The generative nature of the proposed model naturally offers a simulation algorithm, able to synthesize a new data set resembling the original one, of any desired time sampling or possibly balanced cofactors of interest. We emphasize that the proposed modeling approach is able to deal with meshes without any assumption on their topology, in particular without assuming point-to-point correspondence.

The three proposed algorithms were validated in varied simulated configurations, demonstrating their ability to retrieve the true parameters or reproduce the original data distribution. They were illustrated on a data set of emotive faces, showing the relevance of the learned normative scenarii and their potential for classification. A large medical data set of Alzheimer's disease patient was finally thoroughly analyzed. The learned average population atrophy pattern for the hippocampi subcortical structures are in line with current medical knowledge. Individual sequences are successfully parametrized by 10 scalar spatio-temporal coordinates in the calibrated reference frames. Correlating these coordinates with genetic, biological and environmental cofactors gives valuable insights about atrophy-protective conditions in terms of onset age or pace, as well as typical geometries of the hippocampi after removal of the ageing confounding effect.

As a perspective for this work, ergonomic visualization dashboards could be designed for neurologists and radiologists to enrich the information on which they might base a diagnosis or prognosis. We believe that the estimated individual parameters, and in particular the temporal ones, bear relevant high-level though intuitive information about the current and future state of the considered individual.

Acknowledgements This work has been partly funded by the European Research Council with grant 678304, European Union's Horizon 2020 research and innovation program with grant 666992, and the program Investissements d'avenir ANR-10-IAIHU-06.

References

1. Abdi, H.: Partial least square regression (pls regression). Encyclopedia for research methods for the social sciences **6**(4), 792–795 (2003)
2. Ahrens, J., Geveci, B., Law, C.: Paraview: An end-user tool for large data visualization. The visualization handbook **717** (2005)
3. Allassonnière, S., Durrleman, S., Kuhn, E.: Bayesian mixed effect atlas estimation with a diffeomorphic deformation model. SIAM Journal on Imaging Science **8**, 13671395 (2015)
4. Allassonnière, S., Kuhn, E., Trouvé, A.: Construction of bayesian deformable models via a stochastic approximation algorithm: a convergence study. Bernoulli **16**(3), 641–678 (2010)
5. Amor, B.B., Drira, H., Berretti, S., Daoudi, M., Srivastava, A.: 4-d facial expression recognition by learning geometric deformations. IEEE Trans. Cybernetics **44**(12), 2443–2457 (2014)
6. Atchade, Y.F.: An adaptive version for the metropolis adjusted langevin algorithm with a truncated drift. Methodology and Computing in applied Probability **8**(2), 235–254 (2006)
7. Banerjee, M., Chakraborty, R., Ofori, E., Okun, M.S., Viallancourt, D.E., Vemuri, B.C.: A nonlinear regression technique for manifold valued data with applications to medical image analysis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4424–4432 (2016)
8. Beg, F., Miller, M., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. IJCV (2005)
9. Bône, A., Colliot, O., Durrleman, S.: Learning distributions of shape trajectories from longitudinal datasets: a hierarchical model on a manifold of diffeomorphisms. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9271–9280 (2018)
10. Chakraborty, R., Banerjee, M., Vemuri, B.C.: Statistics on the space of trajectories for longitudinal data analysis. In: Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on, pp. 999–1002. IEEE (2017)
11. Charlier, B., Feydy, J., Glaunès, J.A., Trouvé, A.: An efficient kernel product for automatic differentiation libraries, with applications to measure transport (2017)
12. Charon, N., Trouvé, A.: The varifold representation of nonoriented shapes for diffeomorphic registration. SIAM Journal on Imaging Sciences **6**(4), 2547–2580 (2013)
13. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. IEEE transactions on image processing **5**(10), 1435–1447 (1996)
14. Delyon, B., Lavielle, M., Moulines, E.: Convergence of a stochastic approximation version of the em algorithm. Annals of statistics pp. 94–128 (1999)
15. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. Journal of the royal statistical society. Series B (methodological) pp. 1–38 (1977)
16. Durrleman, S., Allassonnière, S., Joshi, S.: Sparse adaptive parameterization of variability in image ensembles. IJCV **101**(1), 161–183 (2013)
17. Durrleman, S., Pennec, X., Trouvé, A., Braga, J., Gerig, G., Ayache, N.: Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data. International Journal of Computer Vision **103**(1), 22–59 (2013). DOI 10.1007/s11263-012-0592-x. URL https://doi.org/10.1007/s11263-012-0592-x
18. Durrleman, S., Prastawa, M., Charon, N., Korenberg, J.R., Joshi, S., Gerig, G., Trouvé, A.: Morphometry of anatomical shape complexes with dense deformations and sparse parameters. NeuroImage (2014)
19. Fang, T., Zhao, X., Shah, S.K., Kakadiaris, I.A.: 4d facial expression recognition. In: Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, pp. 1594–1601. IEEE (2011)
20. Fischl, B., Dale, A.M.: Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences **97**(20), 11,050–11,055 (2000)

21. Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., et al.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron **33**(3), 341–355 (2002)
22. Fishbaugh, J., Prastawa, M., Gerig, G., Durrleman, S.: Geodesic regression of image and shape data for improved modeling of 4D trajectories. In: ISBI 2014 - 11th International Symposium on Biomedical Imaging, pp. 385 – 388 (2014)
23. Fletcher, T.: Geodesic regression and the theory of least squares on riemannian manifolds. IJCV **105**(2), 171–185 (2013)
24. Gori, P., Colliot, O., Marrakchi-Kacem, L., Worbe, Y., Poupon, C., Hartmann, A., Ayache, N., Durrleman, S.: A Bayesian Framework for Joint Morphometry of Surface and Curve meshes in Multi-Object Complexes. Medical Image Analysis **35**, 458–474 (2017). DOI 10.1016/j.media.2016.08.011. URL https://hal.inria.fr/hal-01359423
25. Hinkle, J., Muralidharan, P., Fletcher, P.T., Joshi, S.: Polynomial Regression on Riemannian Manifolds, pp. 1–14. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
26. Hirsch, M.W.: Differential topology, vol. 33. Springer Science & Business Media (2012)
27. Hyvärinen, A., Karhunen, J., Oja, E.: Independent component analysis, vol. 46. John Wiley & Sons (2004)
28. Jian, B., Vemuri, B.C.: Robust point set registration using gaussian mixture models. IEEE transactions on pattern analysis and machine intelligence **33**(8), 1633–1645 (2011)
29. Joshi, S.C., Miller, M.I.: Landmark matching via large deformation diffeomorphisms. IEEE Transactions on Image Processing **9**(8), 1357–1370 (2000)
30. Kendall, D.G.: Shape manifolds, procrustean metrics, and complex projective spaces. Bulletin of the London Mathematical Society **16**(2), 81–121 (1984)
31. Kim, H.J., Adluru, N., Collins, M.D., Chung, M.K., Bendlin, B.B., Johnson, S.C., Davidson, R.J., Singh, V.: Multivariate general linear models (mglm) on riemannian manifolds with applications to statistical analysis of diffusion weighted images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2705–2712 (2014)
32. Kim, H.J., Adluru, N., Suri, H., Vemuri, B.C., Johnson, S.C., Singh, V.: Riemannian nonlinear mixed effects models: Analyzing longitudinal deformations in neuroimaging. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
33. Koval, I., Schiratti, J.B., Routier, A., Bacci, M., Colliot, O., Allassonnière, S., Durrleman, S., Initiative, A.D.N., et al.: Statistical learning of spatiotemporal patterns from longitudinal manifold-valued networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 451–459. Springer (2017)
34. Liu, D.C., Nocedal, J.: On the limited memory bfgs method for large scale optimization. Mathematical programming **45**(1-3), 503–528 (1989)
35. Lorenzi, M., Ayache, N., Frisoni, G., Pennec, X.: 4D registration of serial brains MR images: a robust measure of changes applied to Alzheimer's disease. Spatio Temporal Image Analysis Workshop (STIA), MICCAI (2010)
36. Lorenzi, M., Ayache, N., Pennec, X.: Schild's ladder for the parallel transport of deformations in time series of images. pp. 463–474. Springer (2011)
37. Louis, M., Bône, A., Charlier, B., Durrleman, S.: Parallel transport in shape analysis: a scalable numerical scheme. In: International Conference on Geometric Science of Information, pp. 29–37. Springer (2017)
38. Louis, M., Charlier, B., Jusselin, P., Susovan, P., Durrleman, S.: A fanning scheme for the parallel transport along geodesics on riemannian manifolds (2017)
39. Manasse, F., Misner, C.W.: Fermi normal coordinates and some basic concepts in differential geometry. Journal of mathematical physics **4**(6), 735–745 (1963)
40. Marin, J.M., Pudlo, P., Robert, C.P., Ryder, R.J.: Approximate bayesian computational methods. Statistics and Computing **22**(6), 1167–1180 (2012)
41. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. Journal of Mathematical Imaging and Vision **24**(2), 209–228 (2006)
42. Muralidharan, P., Fletcher, P.T.: Sasaki metrics for analysis of longitudinal data on manifolds. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pp. 1027–1034. IEEE (2012)
43. Niethammer, M., Huang, Y., Vialard, F.X.: Geodesic regression for image time-series. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 655–662. Springer (2011)
44. Pennec, X.: Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. Journal of Mathematical Imaging and Vision **25**(1), 127–154 (2006)
45. Pennec, X., Fillard, P., Ayache, N.: A riemannian framework for tensor computing. International Journal of Computer Vision **66**(1), 41–66 (2006)
46. Schiratti, J.B., Allassonnière, S., Colliot, O., Durrleman, S.: Learning spatiotemporal trajectories from manifold-valued longitudinal data. In: NIPS 28, pp. 2404–2412 (2015)
47. Singh, N., Hinkle, J., Joshi, S., Fletcher, P.T.: Hierarchical geodesic models in diffeomorphisms. IJCV **117**(1), 70–92 (2016)
48. Stern, Y.: Cognitive reserve and alzheimer disease. Alzheimer Disease & Associated Disorders **20**(2), 112–117 (2006)
49. Su, J., Kurtek, S., Klassen, E., Srivastava, A., et al.: Statistical analysis of trajectories on riemannian manifolds: bird migration, hurricane tracking and video surveillance. The Annals of Applied Statistics **8**(1), 530–552 (2014)
50. Su, J., Srivastava, A., de Souza, F.D., Sarkar, S.: Rate-invariant analysis of trajectories on riemannian manifolds with application in visual speech recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 620–627 (2014)
51. Sun, Y., Yin, L.: Facial expression recognition based on 3d dynamic range model sequences. In: European Conference on Computer Vision, pp. 58–71. Springer (2008)
52. Vaillant, M., Glaunès, J.: Surface matching via currents. In: Information processing in medical imaging, pp. 1–5. Springer (2005)
53. Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M.: Bayesian analysis of neuroimaging data in fsl. Neuroimage **45**(1), S173–S186 (2009)
54. Yin, L., Chen, X., Sun, Y., Worm, T., Reale, M.: A high-resolution 3d dynamic facial expression database. In: Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on, pp. 1–6. IEEE (2008)
55. Younes, L.: Jacobi fields in groups of diffeomorphisms and applications. Quarterly of Applied Mathematics **65**(1), 113–134 (2007)
56. Younes, L.: Shapes and Diffeomorphisms. Applied Mathematical Sciences. Springer Berlin Heidelberg (2010). URL https://books.google.fr/books?id=SdTBtMGgeAUC
57. Zhang, M., Fletcher, P.T.: Finite-dimensional lie algebras for fast diffeomorphic image registration. In: International Conference on Information Processing in Medical Imaging, pp. 249–260. Springer (2015)
58. Zhang, M., Singh, N., Fletcher, P.T.: Bayesian estimation of regularization and atlas building in diffeomorphic image registration. In: IPMI, vol. 23, pp. 37–48 (2013)