A sub-Riemannian modular framework for diffeomorphism based analysis of shape ensembles

Barbara Gris, Stanley Durrleman, Alain Trouvé

A SUB-RIEMANNIAN MODULAR FRAMEWORK FOR DIFFEOMORPHISM BASED ANALYSIS OF SHAPE ENSEMBLES

BARBARA GRIS *, STANLEY DURRLEMAN†, AND ALAIN TROUVÉ‡

Abstract. Deformations, and diffeormophisms in particular, have played a tremendous role in the field of statistical shape analysis, as a proxy to measure and interpret differences between similar objects but with different shapes. Diffeomorphisms usually result from the integration of a flow of regular velocity fields, whose parameters have not enabled so far a full control of the local behaviour of the deformation.

In this work, we propose a new mathematical and computational framework, in which diffeomorphisms are built on the combination of local deformation modules with few degrees of freedom. Deformation modules contribute to a global velocity field, and interact with it during integration so that the local modules are transported by the global diffeomorphic deformation under construction. Such modular diffeomorphisms are used to deform shapes and to provide the shape space with a sub-Riemannian metric.

We then derive a method to estimate a Fréchet mean from a series of observations, and to decompose the variations in shape observed in the training samples into a set of elementary deformation modules encoding distinctive and interpretable aspects of the shape variability. We show how this approach brings new solutions to long lasting problems in the fields of computer vision and medical image analysis. For instance, the easy implementation of priors in the type of deformations offers a direct control to favor one solution over another in situations where multiple solutions may fit the observations equally well. It allows also the joint optimisation of a linear and a non-linear deformation between shapes, the linear transform simply being a particular type of modules.

The proposed approach generalizes previous methods for constructing diffeomorphisms and opens up new perspectives in the field of statistical shape analysis.

1. Introduction. Shape data play a very peculiar role within the blooming field of data science. Unlike many other kinds of data, the "right" representation of shape data remains largely an open problem despite a long lasting research effort, for instance in the fields of computer vision and medical image analysis. Another specificity of shape data is their ability to be easily visualized and interpreted. This fact often demands the derivation of learning methods that yield statistical estimate also in the form of shapes that may be visualized and interpreted. The definition of generative statistical models is notably difficult due to the inherent structure of shape data, and the absence of obvious mathematical framework to define a shape space that accounts for this structure.

A category of methods builds on the idea to deform one shape onto another to quantify their differences. The deformation is used here to position one shape with respect to another in a high-dimensional manifold. This line of works, which follows the pioneering work of Grenander, grounded itself in the seminal vision of D'Arcy Thompson [13, 32], has played an important role in the fields of medical image analysis, computational analysis, or biological shape analysis. In these fields, shape data are derived from 2D or 3D imaging systems, and collections of such data form a set of observations of the same organ, tissue or structure that vary in shape across different individuals. Since such samples are usually observed in full (e.g. without occlusion) and share a similar topology, it makes sense to try to estimate a

*CMLA, ENS Cachan, CNRS, Université Paris-Saclay, France AND Inria Paris-Rocquencourt, Sorbonne Universités, UPMC Univ Paris 06 UMR S1127, Inserm U1127, CNRS UMR 7225, ICM, F-75013, Paris, France

†Inria Paris-Rocquencourt, Sorbonne Universités, UPMC Univ Paris 06 UMR S1127, Inserm U1127, CNRS UMR 7225, ICM, F-75013, Paris, France

‡CMLA, ENS Cachan, CNRS, Université Paris-Saclay, France

diffeomorphic deformation, namely a smooth and invertible space deformation with smooth inverse, which warps one shape onto another.

Various mathematical constructions have been derived from this idea [7, 24, 34]. Such constructions usually include the definition of a particular group of diffeomorphisms and an action of the deformation on the shape data, which together yields then a shape space provided with a metric. These methods may be pooled under the umbrella of the generic concept of "shape spaces" as recently formalised in [4]. A group of diffeomorphisms G is built by integrating trajectories of vector fields belonging a fixed Hilbert space V, and the inner product of this space allows the definition of a Riemannian structure on the group G. Then a shape space is defined as a manifold M, on which this group G may act. The Riemannian structure on G enables the definition of a distance on the shape space M.

These geometrical constructions have been then included into generative statistical models to estimate an "average shape", usually called "template", and the deformations warping the template to each observation [2, 9, 21, 22, 33]. The parameters of these deformations are used to give an estimate of the variance of the observations in the tangent-space of the manifold M at the template point. Template and deformations together are referred to as an "atlas".

Key to these statistical shape analysis methods is the definition of the deformation group and its metric. In the vast majority of cases, the metric structure of the deformation group is inherited from the Hilbert space V of driving velocity fields and therefore offers little control over the local behaviour of the deformations. Resulting deformations show complex non-linear patterns, which are difficult to isolate and interpret. Deformation parameters usually give only a vague description of these patterns. Coming back to Grenander's vision, as detailed in the introduction of [14], shape analysis methods should learn from the observations the typical local deformation patterns, which are indicative of the biological processes that have led to the observed variability. By nature, these patterns are constrained to be local and have very few degrees of freedom. Typical types of such local patterns may include local scaling, torque, or fanning pattern for instance. In the Grow Random Iterated Diffeomorphisms (GRID) model [15, 25, 31], deformation patterns are built thanks to active points, called seeds, whose placements are modelled as a Poisson process and which generate local particular deformation. The global diffeomorphism results in the iterative composition of such deformations. Locations of seeds and types of deformations are estimated at each time so that they maximize the likelihood and then are not transported by the diffeomorphism between iterations.

It is relatively easy to define vector fields as the superimposition of local infinitesimal deformations of a given type. What is difficult is to know how to update this vector field during integration to generate large diffeomorphic deformations. An idea may be to keep the vector field stationary as in [6, 27, 36], but then the local deformation generators are attached to a fixed background, and not to the object that is deformed. One needs not only to combine local deformation generators into a single vector field, but also to transport the local generators with the deformation that results from the integration of the velocity fields.

In the framework of Large Diffeomorphic Deformation Mapping (LDDMM) [24], the structure of the velocity fields depends on the data, although it is possible to constrain the velocity field to be generated by a set of local translations acting at a specific scale [10]. Other attempts to gain control over the local structure of the deformation is to replace the local translation, which is a singular momentum, by a

higher-order momentum encoding local affine transform [17, 29]. The scale at which translations act may be also adjusted by using a superimposition of windowing functions with various scales and to select the ones that are the most relevant to describe shape variability [26, 28]. In [35], local constraints depending on the shape are introduced, and allow a particular control on the structure of vector field. This structure evolves along the resulting flow and induces a sub-Riemannian structure on the shape space. But in [35] the idea is more to provide good finite dimensional approximations to the full group of deformations through the introduction of constraints than consider the introduction of more prior-based constraints.

The main contribution of this work is to define "modular" diffeomorphisms, which are defined by a set of local infinitesimal deformation generators called modules, extending the work of [16]. These modules combine to generate velocity fields embedded in a common Hilbert space V. Deformations resulting from the integration of trajectories of such velocity fields are used to update the parameters of the modules. This construction forms a dynamical system with a feedback loop, in which modules interact in a dynamical way. Shapes can be deformed by these modular diffeomorphisms, and a concept of cost associated to each module provides the shape space with a sub-Riemannian metric. This structure in general is not Riemannian because of the modularity constraints on the velocity field. We define modules and costs, so that elementary modules may be combined into more complex compound modules, thus allowing the definition of a tree structure for modules. The deformation is determined by the root module, which may be decomposed in a hierarchy of sub-modules.

Modular diffeomorphisms allow an explicit control of the local structure of the diffeomorphisms, in the sense that one may choose beforehand the type and scale of the local deformation in each region of the space. Their parameters have a clear and intuitive meaning. They depict the dynamical modular structure of the deformation. An interesting feature of this construction is that it encompasses previous models of diffeomorphic deformations, such as those in [10, 24].

In the following, we present the geometrical construction of the modular diffeomorphisms in Section 2 and 3). In Section 4, we include such deformation model in a method to build an atlas from a series of shape data. We show then in Section 5 how this approach may bring new solutions to simple but yet challenging problems in shape analysis.

2. Modular large deformation.

2.1. Deformation Module. We recall first the mathematical framework which defines deformations as the result of the integration of time-varying velocity fields. Let d and ℓ be non-zero integers, we define $C^\ell_0(\mathbb{R}^d)$ the space of vector fields of class C^ℓ on $\mathbb{R}^d$ whose derivatives of order less than or equal to ℓ converge to zero at infinity and we equip it with norm $|v|_\ell = \sup\{|\frac{\partial^{\ell_1+\cdots+\ell_d} v(x)}{\partial x_1^{\ell_1}\cdots x_d^{\ell_d}}| \mid x \in \mathbb{R}^d, (\ell_1,\cdots,\ell_d) \in \mathbb{N}^d, \ell_1+\cdot+\ell_d \leq \ell\}$ such that it is a Banach space. We define $\mathrm{Diff}^\ell_0(\mathbb{R}^d)$ the space of $C^\ell-$diffeomorphisms of $\mathbb{R}^d$ that converge to identity at infinity, that is an open set of the affine Banach space $Id+C^\ell_0(\mathbb{R}^d)$ and as such equipped with a natural smooth differential structure. We will consider particular trajectories of $\mathrm{Diff}^\ell_0(\mathbb{R}^d)$ defined as flow of particular trajectories of $C^\ell_0(\mathbb{R}^d)$.

Proposition 1. *[5] Let v be an element of $L^1([0,1], C^\ell_0(\mathbb{R}^d))$, ie a time-dependent vector field such that $t \in [0,1] \mapsto |v(t)|_\ell$ is integrable. Then there exists a unique*

absolutely continuous solution φ^v, called ***the flow*** *of v, to the system*

$$\left\{ \begin{array}{rcl} \dot{\varphi}^v(t) & = & v(t) \circ \varphi^v(t) \\ \varphi(0) & = & Id \end{array} \right.$$

where $\phi^v(t) \in \mathrm{Diff}_0^\ell(\mathbb{R}^d)$ for any $t \in [0,1]$.

This proposition allows us to consider flows of time-dependent vector fields, where "time" t refers here to a variable of integration. As we are interested in how these flows can deform a certain shape, we need to precise what is a shape, and how a diffeomorphism of $\mathrm{Diff}_0^\ell(\mathbb{R}^d)$ can act on it. We will use the notion of shape space defined by S. Arguillère in [5], which we recall here.

Definition 1. *Let $\mathcal{O}$ be a manifold of finite dimension and $k \in \mathbb{N}^*$. Assume that the group $\mathrm{Diff}_0^\ell(\mathbb{R}^d)$ continuously acts on $\mathcal{O}$, according to the action*

$$\begin{array}{rcl} \mathrm{Diff}_0^\ell(\mathbb{R}^d) \times \mathcal{O} & \to & \mathcal{O} \\ (\varphi, o) & \mapsto & \varphi \cdot o \end{array} \tag{2.1}$$

*We say that $\mathcal{O}$ is a C^k-**shape space of order** ℓ on $\mathbb{R}^d$ if the following conditions are satisfied:*

1. *for each $o \in \mathcal{O}$, $\phi \in \mathrm{Diff}_0^\ell(\mathbb{R}^d) \mapsto \phi \cdot o$ is Lipschitz with respect to the norm $|\cdot|_\ell$ and is differentiable at $Id_{\mathbb{R}^d}$. This differential is denoted ξ_o and is called the **infinitesimal action** of $C_0^\ell(\mathbb{R}^d)$.*
2. *The mapping $\xi : (o, v) \in \mathcal{O} \times C_0^\ell(\mathbb{R}^d) \mapsto \xi_o v$ is continuous and its restriction to $\mathcal{O} \times C_0^{\ell+k}(\mathbb{R}^d)$ is of class C^k.*

An element o of $\mathcal{O}$ is called a ***shape****, and $\mathbb{R}^d$ will be referred to as the* ***ambient space****.*

The simplest example of shape space is the one of landmarks, where shapes are given as a collection of a fixed number of points: $\mathcal{O} = \{o = (x_1, \cdots, x_n) \in (\mathbb{R}^d)^n \mid x_i \neq x_j \texttt{ if } i \neq j\}$. The action of $\mathrm{Diff}_0^\ell(\mathbb{R}^d)$ is given by $\phi \cdot o = (\phi(x_1), \cdots, \phi(x_n))$, namely the deformed shape, and the infinitesimal action for $v \in C_0^\ell(\mathbb{R}^d)$ is $\xi_o(v) = (v(x_1), \cdots, v(x_n))$ for $o = (x_1, \cdots, x_n) \in \mathcal{O}$, namely the velocity of the trajectories of shape points. These actions make $\mathcal{O}$ a shape space of order ℓ for any $\ell \geq 1$.

Another example is a set of N Delta Dirac 1-currents encoding tangents of a 2D or 3D curve [11]. Shapes in this shape space are represented by base points c_k and tangent vectors τ_k. The action of $\mathrm{Diff}_0^\ell(\mathbb{R}^d)$ (where $d = 2$ or 3) is given by $\phi.((c_1, \tau_1), \ldots, (c_N, \tau_N)) = ((\phi(c_1), d_{c_1}\phi(\tau_1)), \ldots, (\phi(c_N), d_{c_N}\phi(\tau_N)))$, and the infinitesimal action as $((v(c_1), dv_{c_1}(\tau_1)), \ldots, (v(c_N), dv_{c_1}(\tau_N)))$.

Remark 1. *If $\mathcal{O}_1$ and $\mathcal{O}_2$ are two shape spaces of order ℓ on $\mathbb{R}^d$, then $\mathcal{O} = \mathcal{O}_1 \times \mathcal{O}_2$ is also a shape space of order ℓ on $\mathbb{R}^d$.*

The following proposition was proved in [5] and shows that a shape transported by a flow of diffeomorphisms (see Proposition 1) satisfies a differential equation.

Proposition 2. *For every $o \in \mathcal{O}$, the mapping $\phi \mapsto \phi \cdot o$ is of class C^1. Moreover, if $v \in L^1([0,1], C_0^\ell(\mathbb{R}^d))$, for $a \in \mathcal{O}$, the curve $o : t \in [0,1] \mapsto o(t) = \varphi^v(t) \cdot a$ is absolutely continuous and satisfies for almost every $t \in [0,1]$, $\dot{o} = \xi_o(v)$.*

2.1.1. Definition and first examples. We will now define the core concept of our approach, the concept of deformation module. Intuitively, a deformation module is a structure that embeds a vector field generation mechanism. This mechanism depends on some geometrical descriptors, which specify the local structure of the induced vector fields (think about the positioning of some actuators), and a finite numbers of control parameters commanding the actuators. For a given positioning, a cost is specified for any possible values of the control parameters, so that optimal

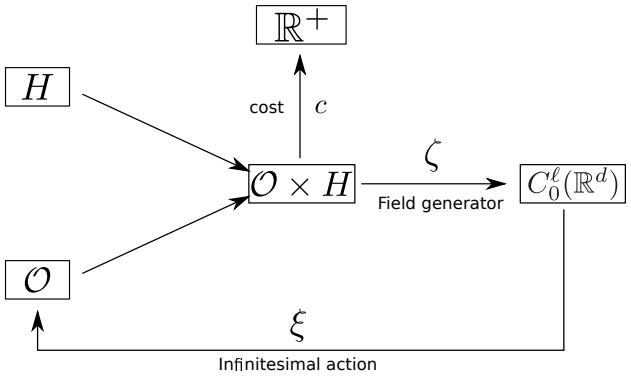

FIG. 2.1. *Schematic view of a deformation module.*

policies can be defined. For each resulting vector fields, a feedback mechanism is defined to update the positioning of the geometrical descriptors. Desirable consistency properties lead to consider the geometrical descriptors as defining a shape on which one has a diffeomorphic action. Hence following [35], the positioning of the geometrical descriptors will be represented as a shape in a shape space.

In the following we set $d \in \mathbb{N}^*$ the dimension of the ambient space.

Definition 2. *Let $k, \ell \in \mathbb{N}^*$. We define $M = (\mathcal{O}, H, \zeta, \xi, c)$ as a C^k-**deformation module of order** ℓ with **geometrical descriptors** in $\mathcal{O}$, **controls** in H, **infinitesimal action** ξ, **field generator** ζ and **cost** c, if*

- *$\mathcal{O}$ is a C^k-shape space of $\mathbb{R}^d$ of order ℓ with infinitesimal action $\xi : C_0^\ell(\mathbb{R}^d) \times \mathcal{O} \longrightarrow T\mathcal{O}$,*
- *H is a finite dimensional Euclidean space,*
- *$\zeta : (o, h) \in \mathcal{O} \times H \to (o, \zeta_o(h)) \in \mathcal{O} \times C_0^\ell(\mathbb{R}^d)$ is continuous, with $h \mapsto \zeta_o(h)$ linear and $o \mapsto \zeta_o$ of class C^k,*
- *$c : (o, h) \in \mathcal{O} \times H \to c_o(h) \in \mathbb{R}^+$ is a continuous mapping such that $o \mapsto c_o$ is smooth and for all $o \in \mathcal{O}$, $h \mapsto c_o(h)$ is a positive quadratic form on H, thus defining smooth metric on $\mathcal{O} \times H$.*

A deformation module is defined by the way it can generate a vector field, which is given by the field generator ζ, and by the feedback action on vector fields, which is given by the infinitesimal action ξ. A schematic view of the construction of a module can be seen in figure 2.1.

Remark 2. *By definition, the geometrical descriptors of a deformation module are considered as a "shape" in the shape space $\mathcal{O}$. In the following examples, these "shapes" are the centers of the scaling, of the rotation or the base points of translation vectors. The fact that they are "shapes" in the sense of Definition 1 allows us to use the global deformation under construction to deform or transport them via the infinitesimal action ξ. These "shapes" may or may not coincide with points from the input shape data set. For example, it is sensible not to locate the center of a scaling on the boundaries of an object. We will see in the following that the input shape data may be considered as geometrical descriptors of a particular module, called a silent module, which is deformed thanks to its infinitesimal action, but which does not contribute to the global velocity field.*

We will now present two examples of simple deformation modules, a richer pre-

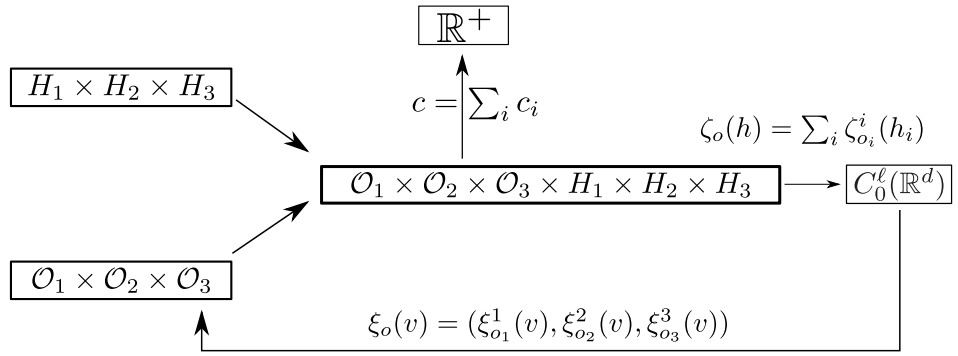

FIG. 2.2. *Schematic view of a combination of three deformation modules.*

sentation of different deformation modules will be done in Section 2.3

First example: sum of local translations. This first example explains how the construction of [10] can be seen as a C^k-deformation module of order ℓ for any $k, \ell \in \mathbb{N}^*$. We want to build a deformation module M that generates vector fields that are a sum of D local translations acting at scale σ in the ambient space $\mathbb{R}^d$ ($d \in \mathbb{N}^*$). We set V_σ the scalar Gaussian Reproducing Kernel Hilbert Space (RKHS) of scale σ (its kernel will be denoted $K_\sigma : (x, y) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto \exp(-|x-y|^2/2\sigma^2)$), $\mathcal{O} \doteq (\mathbb{R}^d)^D$ the shape space of D landmarks with infinitesimal action $\xi_o : v \in C^\ell_0(\mathbb{R}^d) \mapsto (v(z_i))_i$, where $o = (z_i) \in \mathcal{O}$ (application of the vector field at each point), and $H \doteq (\mathbb{R}^d)^D$ (families of D vectors). For $o = (z_i) \in \mathcal{O}$, we define $\zeta_o : h = (\alpha_i) \in H \mapsto \sum_{i=1}^D K_\sigma(z_i, \cdot)\alpha_i$ and $c_o : h = (\alpha_i) \in H \mapsto |\sum_i K_\sigma(z_i, \cdot)\alpha_i|^2_{V_\sigma}$. It is easy to show that $M = (\mathcal{O}, H, \zeta, \xi, c)$ defines a deformation module of order ℓ.

Second example: local scaling or rotations. Let $\ell, k \in \mathbb{N}^*$. The second example is a C^k-deformation module of order ℓ which generates vector fields that are local scalings of fixed scale $\sigma \in \mathbb{R}^+$ in the ambient space $\mathbb{R}^2$. A local scaling is parametrized by its center $o \in \mathbb{R}^2$ and its scale factor $h \in \mathbb{R}$. The point o plays the role of the geometrical descriptor, and the factor h the control parameter. From o, we build 3 points $z_j(o)$ and 3 unit vectors d_j as described in Figure 2.3 (they also depend on the fixed parameter σ). The idea is to build the vector field generated by the geometrical descriptor o and the control h as an interpolation of the values at these points $z_j(o)$ thanks to vectors d_j: $\zeta_o(h) \doteq h \sum_{j=1}^3 K_\sigma(z_j(o), \cdot)d_j$, where K_σ is the kernel of the scalar gaussian RKHS of scale σ. We emphasize here that points $z_j(o)$ and vectors d_j are *intermediate tools* used to build the vector field but that the latter is only parametrized by o and h. We then define the deformation module M by the following spaces : $\mathcal{O} \doteq \mathbb{R}^2$, $H = \mathbb{R}$ and the following mappings : for $o \in \mathcal{O}$, $\zeta_o : h \in H \mapsto \zeta_o(h)$ as given above, $\xi_o : v \in C^\ell_0(\mathbb{R}^2) \mapsto v(o)$, the velocity field at the scaling center, and $c_o : h \in H \mapsto |\zeta_o(h)|^2_{V_\sigma} = h^2 \sum_{j,j'} K_\sigma(z_j, z_{j'}) d_j^T d_{j'}$, the squared norm of the generated velocity field in the RKHS V_σ. Then $M = (\mathcal{O}, H, \zeta, \xi, c)$ defines a C^k-deformation module of order ℓ. Indeed space $\mathcal{O}$ is a C^k-shape space (of one landmark) of order ℓ and $\zeta : o \in \mathcal{O} \mapsto \zeta_o \in L(H, C^\ell_0(\mathbb{R}^2))$ is locally Lipschitz because distances between $z_i(o)$ and $z_i(o')$ ($1 \leq i \leq 3$) are the same as the distance between o and o'. For the same reason ζ is globally continuous. Other properties can be easily verified. This construction can be generalised to any other affine deformation by changing the rule to build vectors d_j, as for example in Figure 2.4 where the generated vector field is a local rotation.

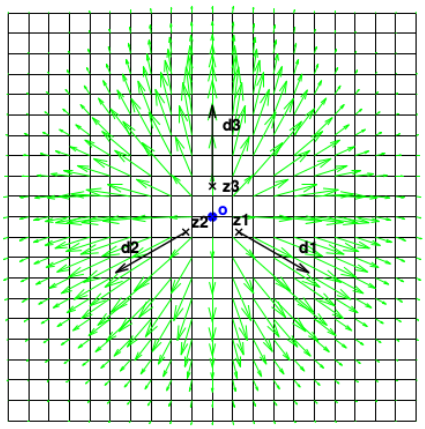

FIG. 2.3. *Local scaling. Geometrical descriptor o (in blue) and intermediate tools (in black). Plot of the resulting vector field in green.*

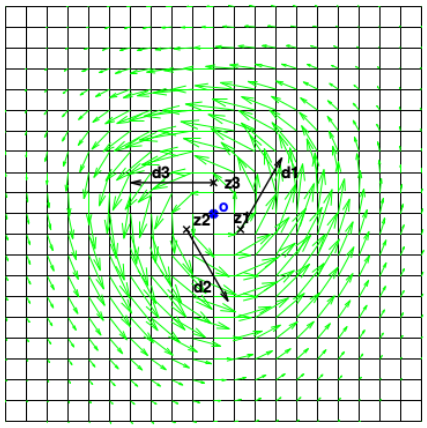

FIG. 2.4. *Local rotation. Geometrical descriptor o (in blue) and intermediate tools (in black). Plot of the resulting vector field in green.*

A key point in the design of a consistent approach for deformation modules is the possibility to deduce existence of optimal deformations between shapes as solutions of the optimal control problem associated with the choice of a cost for the control parameters. This requirement imposes constraints on the choice of costs. For instance, it seems important for the cost to be related with some metric of the induced vector field. This is the spirit of the following embedding condition:

Definition 3. *Let $M = (\mathcal{O}, H, \zeta, \xi, c)$ be a C^k-deformation module of order ℓ. We say that M satisfies the **Uniform Embedding Condition (UEC)** if there exists a Hilbert space of vector fields V continuously embedded in $C_0^{\ell+k}(\mathbb{R}^d)$ and a constant $C > 0$ such that for all $o \in \mathcal{O}$ and for all $h \in H$, $\zeta_o(h) \in V$ and*

$$|\zeta_o(h)|_V^2 \leq C c_o(h)$$

Remark 3. *Examples of deformation modules presented previously satisfy this UEC. Indeed in these two examples the infinitesimal action ξ takes values in the Hilbert space V_σ which is continuously embedded in $C_0^{\ell+k}(\mathbb{R}^d)$ (see [11]), and the cost is defined by the squared-norm of the generated vector field in V_σ.*

2.1.2. Combination. We have presented in the previous section examples of simple, base deformation modules generating *simple* vectors fields. We want now to build multi-modular vector fields by summing vector fields generated by different deformation modules. In order to do so, we define the combination of modules as follows (a schematic view of this combination can be seen in Figure 2.2).

Definition 4. *Let $M^l = (\mathcal{O}^l, H^l, \zeta^l, \xi^l, c^l)$, $l = 1 \cdots L$, be L C^k-deformation modules of order ℓ. We define the **compound module** of modules M^l by $\mathcal{C}(M^l, l = 1 \cdots L) = (\mathcal{O}, H, \zeta, \xi, c)$ where $\mathcal{O} \doteq \prod_l \mathcal{O}^l$, $H \doteq \prod_l H^l$ and for $o = (o^l)_l \in \mathcal{O}$, $\zeta_o : h = (h^l) \in H \mapsto \sum_l \zeta^l_{o^l}(h^l)$, $\xi_o : v \in C_0^\ell(\mathbb{R}^d) \mapsto (\xi^l_{o^l}(v))_l \in T_o\mathcal{O}$ and $c_o : h = (h^l) \in H \mapsto \sum_l c^l_{o^l}(h^l)$.*

A key point of our framework is the following stability result under combination:

Proposition 3. *If $M^l = (\mathcal{O}^l, H^l, \zeta^l, \xi^l, c^l)$, $l = 1 \cdots L$, are C^k-deformation modules of order ℓ, then $\mathcal{C}(M^l, l = 1 \cdots L)$ is a C^k-deformation module of order ℓ. Furthermore, if each M^l satisfies UEC, then $\mathcal{C}(M^l, l = 1 \cdots L)$ also satisfies UEC.*

Proof. It is clear that $\mathcal{C}(M^l, l = 1 \cdots L)$ is a C^k-deformation module of order ℓ. Let suppose that each M^l satisfies UEC. We define

$$\pi : w = (w_1, ..., w_L) \in W \doteq \prod_l V^l \mapsto \sum_i w_i \in C_0^\ell(\mathbb{R}^d).$$

Then space $V \doteq \pi(W)$ can be equipped with the following norm: for $v \in V$, $|v|_V^2 = \inf\{\sum_l |v_l|_{V^l}^2 \mid \pi((v_l)_l) = v\}$, such that it is a Hilbert space continuously embedded in $C_0^\ell(\mathbb{R}^d)$. For any $o = (o^l) \in \mathcal{O}$ and $h = (h^l) \in H$ we have

$$|\zeta_o(h)|_V^2 \leq \sum_{l=1}^L |\zeta_{o^l}^l(h^l)|_{V_l}^2 \leq \sum_{l=1}^L C_l c_{o^l}^l(h^l) \leq (\max_{1 \leq l \leq L} C_l) c_o(h)$$

and then $\mathcal{C}(M^l, l = 1 \cdots L)$ satisfies UEC. □

Remark 4. *Note that even if costs of elementary modules M^l are given by $c_{o^l}^l(h^l) = |\zeta_{o^l}^l(h^l)|_{V^l}^2$ as in our previous examples, in general (when π is not one to one) the cost of the compound module is not the squared norm of the compound velocity field $\zeta_o(h) = \sum_l \zeta_{o^l}^l$ in the global embedding RKHS V, i.e. $c_o(h) = \sum_l |\zeta_{o^l}^l(h^l)|_{V^l}^2 \neq |\zeta_o(h)|_V^2$. Then in general $C > 1$ and c is not the pullback metric on $\mathcal{O} \times H$ of the metric on $\mathcal{O} \times V$. The cost $c_o(h)$ does not directly depend on the norm of the generated vector field $\zeta_o(h)$ but on its specific decomposition as a sum of elementary vector fields $\zeta_{o^l}^l(H^l)$. As we will consider configurations (o, h) that minimize the cost $c_o(h)$ (for a given action $\xi_o(\zeta_o(h))$ on the geometrical descriptor), different choices of cost c^l can favour different decompositions for the same resulting vector field. Moreover in practice one can compute easily the cost of the compound module c from the elementary costs c^l.*

2.2. Large deformations. In this section, we show how large deformations can be generated from a given deformation module $M = (\mathcal{O}, H, \zeta, \xi, c)$ satisfying UEC. These large deformations are obtained as the integration of a trajectory of vector fields $v : t \in [0, 1] \mapsto v_t \in V$ that are modular, meaning that at each time t one can write $v_t = \zeta_{o_t}(h_t)$ with $(o_t, h_t) \in \mathcal{O} \times H$. During the integration of the trajectory we want the geometrical descriptor of the module to be transported by the flow and therefore, denoting φ^v the flow of v (see Proposition 1), that v_t belongs to $\zeta_{o_t}(H)$, with $o_t = \varphi_t^v(o_{t=0})$.

Definition 5. *Let $a, b \in \mathcal{O}$. We denote $\Omega_{a,b}$ the set of mesurable curves $t \mapsto (o_t, h_t) \in \mathcal{O} \times H$ where o_t is absolutely continuous (a.c.), starting from a and ending at b, such that, for almost every $t \in [0, 1]$, $\dot{o}_t = \xi_{o_t}(v_t)$, where $v_t \doteq \zeta_{o_t}(h_t)$, and*

$$E(o, h) \doteq \int_0^1 c_{o_t}(h_t) dt < \infty.$$

*The quantity $E(o, h)$ is called the **energy** of (o, h) and $\Omega_{a,b}$ is the set of **controlled paths of finite energy** starting at a and ending at b.*

If UEC is satisfied, we can build large deformations from these trajectories :

Proposition 4. *Let us suppose that M satisfies UEC. Let $(o, h) \in \Omega_{a,b}$ and for each t, $v_t = \zeta_{o_t}(h_t)$. Then $v \in L^2([0, 1], V) \subset L^1$, the flow φ^v exists, $h \in L^2([0, 1], H)$ and for each $t \in [0, 1]$, $o_t = \phi_t^v.o_0$. The final diffeomorphism $\varphi_{t=1}^v$ is called a **modular large deformation generated by** a.*

Proof. Let $(o,h) \in \Omega_{a,b}$ and $v : t \in [0,1] \mapsto v_t \doteq \zeta_{o_t}(h_t)$. From the UEC we get : $\int_0^1 |v(t)|_V^2 dt \leq C \int_0^1 c_{o_t}(h_t)dt < \infty$ since $(o,h) \in \Omega_{a,b}$. Then $v \in L^2([0,1],V) \subset L^1([0,1],V) \subset L^1([0,1],C_0^\ell(\mathbb{R}^d))$ and its flow φ^v can be defined. As explained in Proposition 2 we can deduce that $o(t) = \varphi_t^v \cdot a$ for each t.
Let us now define, for $o \in \mathcal{O}$, $||c_o^{-1}|| := \sup_{|h|_H=1} c_o(h)^{-1}$ (well defined as H is of finite dimension). As $o \mapsto c_o$ and $t \mapsto o(t) = \varphi_t^v \cdot a$ are continuous, $\sup_t ||c_{o_t}^{-1}||$ is finite and then $\int_0^1 |h_t|_H^2 dt \leq (\sup_t ||c_{o_t}^{-1}||) \int_0^1 c_{o_t}(h_t)dt < \infty$ so $h \in L^2([0,1],H)$.

□

Remark 5. *A modular large deformation is parametrized by an initial value of geometrical descriptor $o_{t=0}$ and a trajectory of control $h \in L^2([0,1],H)$.*

2.3. Examples. We present here a list of possible deformation modules, which comes for the first case with an illustration of a modular large deformation. Most of these modules will be used in the numerical experiments in Section 5. Some modules will be presented only in dimension 2, but may be easily generalized in higher dimensions.

2.3.1. Constrained local transformations. In Section 2.1.1 we presented the example of a deformation module generating vector fields that are always a local scaling at a fixed scale σ. In a more general setting, we can design deformation modules generating a particular type of local transformation by choosing different vectors d_j. Let us fix these vectors d_j and therefore the corresponding local transformation A. Let us now build the deformation module that generates vector fields that are a sum of P replications of A at P different locations. We set $\mathcal{O} \doteq (\mathbb{R}^2)^P$, $H \doteq \mathbb{R}^P$ and, for $o = (o_i)_i \in \mathcal{O}$, $\zeta_o : h = (h_i) \in H \mapsto \sum_{i=1}^P h_i \sum_{j=1}^3 K_\sigma(z_j(o_i), \cdot)d_j$ (with points $z_j(o_i)$ defined in Section 2.1.1, Example 2), $\xi_o : v \in C_0^\ell(\mathbb{R}^d) \mapsto (v(o_i))_i \in T\mathcal{O}$ (application of the vector field to each point), $c_o : h = (h_i)_i \in H \mapsto |\zeta_o(h)|_{V_\sigma}^2$. As for the deformation module generating one local scaling, it can be shown that $M = (\mathcal{O}, H, \zeta, \xi, c)$ is a deformation module satisfying UEC.

In Figure 2.5 we present an example of large deformation (see Proposition 4) generated by the combination of two deformation modules generating constrained local transforms: one generating a scaling at scale σ_1 and the second generating a rotation at scale $\sigma_2 = \sigma_1/3$. The deformation at each time is represented by the deformation of the grid. We can see that geometrical descriptors are transported by the global flow created by the two deformation modules. As its scale is smaller, the area of action of the local rotation is smaller than the one of the scaling. Then the area of influence of the local rotation is transported by the vector field created by the local scaling, while the geometrical descriptor of the local scaling is almost constant. We represent also in this figure intermediate tools d_j, which are recomputed at each time from the geometrical descriptors (centres of the scaling and rotation) and therefore are not transported by the flow. This example shows how complex modular large deformations can be naturally built from simple base modules, and how their mutual interaction during the integration of the trajectory is encoded in the compound deformation module.

2.3.2. Unconstrained local affine transformations. We present here the deformation module generating vector fields that are locally an affine deformation (at a fixed scale σ), without any other prior on the local deformation pattern. We explain here the construction in dimension 2 without loss of generality. For $o \in \mathbb{R}^2$ we define

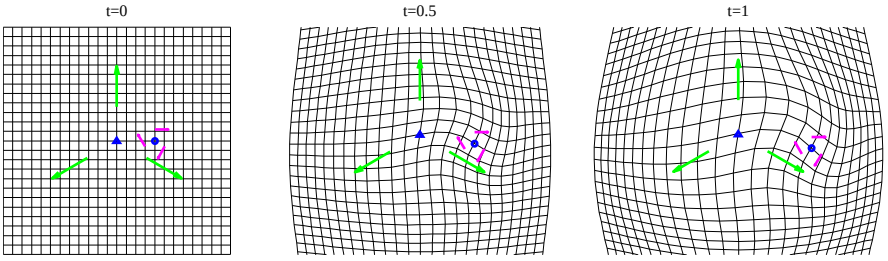

FIG. 2.5. *Example of a large deformation. Geometrical descriptors are in blue (triangle for scaling, circle for rotation). Vectors are intermediate tools used to build vector fields (magenta for rotation and green for scaling)*

points $(z_j(o))_j$ as in the previous case. We define spaces $\mathcal{O} \doteq \mathbb{R}^2$, $H \doteq (\mathbb{R}^2)^3$ (groups of 3 vectors of $\mathbb{R}^2$) and applications $\zeta : (o, h = (\alpha_j)) \in \mathcal{O} \times H \mapsto \sum_j K_\sigma(z_j(o), \cdot)\alpha_j$, $\xi_o : v \in C_0^\ell(\mathbb{R}^2) \mapsto v(o)$ and $c_o : h = (\alpha_j) \in H \mapsto |\zeta_o(h)|^2_{V_\sigma}$. Then $M = (\mathcal{O}, H, \zeta, \xi, c)$ is a C^k-deformation module that satisfies UEC and generates vector fields that are unconstrained local transformations at scale σ. This example differs from the sum of translation in Section 2.1.1 (example 1) as the 3 centres of translations here are glued together. This example differs also from the previous one (Section 2.3.1), as here the directions d_j are free to generate any local transform. These directions become then control parameters, whereas they were fixed in the previous example. Similarly to previous examples local affine transforms may be combined into a deformation module creating a superimposition of P unconstrained local transforms at different locations. This module class differs from the poly-affine framework [27] in that the neighbourhood which is affected by the local affine transformation is transported by the global deformation via the infinitesimal feedback action ξ.

2.3.3. Deformation module generating a sum of local translations. This deformation module enables us to see the construction of [10] as an instance of a deformation module and was detailed in Section 2.1.1, Example 1. The cost of this deformation module is $c : (o = (z_i), h = (\alpha_i)) \mapsto |\zeta_o(h)|^2_{V_\sigma} = \sum_{i,j} K_\sigma(z_i, z_j)\alpha_i^T\alpha_j$. Another possibility to generate sum of translations would be to consider the combination of modules creating one translation, such that the cost would be $\sum_i |K_\sigma(\cdot, z_i)\alpha_i|^2_{V_\sigma} = \sum_i |\alpha_i|^2$. The first case is more interesting because minimizing the cost tends to separate centres z_i and then forces local translations of the same scale to act everywhere it is needed, while in the second case several could converge to the same location, making it redundant.

2.3.4. Example of combination: a multi-scale sum of local translations. Let us fix P scales σ_l and for each l a number of translations D_l. We present here the construction of the deformation module that generates vector fields that are a sum of multi-scale local translations, with D_l translations for each scale σ_l. For each l can be built, as defined in Section 2.1.1 (first example), a module M^l generating vector fields that are a sum of D_l local translations acting at scale σ_l. The multi-scale module M is then the combination of these modules M^l (see Section 2.1.2). In particular the vector field created by geometrical descriptor $o = (z_j^l) \in \mathcal{O} = \Pi_l \left((\mathbb{R}^d)^{D_l}\right)$ and control $h = (\alpha_j^l) \in H = \Pi_l \left((\mathbb{R}^d)^{D_l}\right)$ is $\zeta_o(h) = \sum_l \sum_j K_{\sigma_l}(z_j^l, \cdot)\alpha_j^l$. It is clear here that, unlike in the framework of [26] where vector fields are generated thanks

to a sum of gaussian kernels, centres of local translations are different for each scale. This property is kept for optimal trajectories (minimising the cost) as we force these points to be different at initial time. The cost is, for $o = (z_j^l) \in \mathcal{O}$ and $h = (\alpha_j^l) \in H$, $c_o(h) = \sum_l \sum_j K_{\sigma_l}(z_j^l, z_{j'}^l)\alpha_j^{lT}\alpha_{j'}^l$. Even though this cost is similar to the framework of vector bundle presented in [30], the control variables which are optimised in our framework are only the vectors of local translations $h = (\alpha_j^l) \in H$ and not the global vector field. Therefore, optimal trajectories are different from the vector bundle framework and the decomposition of the vector field in a sum of local translations acting at different scales and centres at different points is preserved.

2.3.5. Deformation module generating a sum of local translations with a prior on the direction. In the previous examples of deformation modules based on local translations (Sections 2.3.3 and 2.3.4), the direction and magnitude of the translation vector were considered as control parameters. Therefore, during the integration of the flow, the direction of the translation at each time needs to be determined as a optimal solution for a given criterion. By contrast, we may want to update the direction of the translation during deformation by using a prior rule, thus considering the direction no more as a control parameter but as a geometrical descriptor instead. In this case, only the magnitude of the translation vector is considered as a control parameter. We give two examples of such modules for translations acting at scale σ.

Translation with constant direction. In this example, the translation vector is kept constant during integration, meaning that the direction of the translation is linked to a fixed background and not to the shape that is deformed. We denote $u \in \mathbb{R}^d$ this direction, and set $\mathcal{O} \doteq \mathbb{R}^d$, $H \doteq \mathbb{R}$ (the control is scalar) and for $(z, u) \in \mathbb{R}^d \times \mathbb{R}^d$, $h \in \mathbb{R}$, $v \in C_0^1(\mathbb{R}^d)$, $\zeta_o(h) \doteq K_\sigma(z, \cdot)hu$, $\xi_o(v) = (v(z), 0)$ and $c_o(h) = |\zeta_o(h)|_{V_\sigma}^2 = h^2$. The geometrical descriptors are the base point of the translation vector which is transported by the deformation flow, and the direction u which is kept constant. This deformation module can be generalized to the deformation module generating vector fields that are a sum of P local translations at scale σ with fixed directions u_i by setting $\mathcal{O} \doteq (\mathbb{R}^d)^P \times (\mathbb{R}^d)^P$, $H \doteq \mathbb{R}^P$ (the control is made of P scalars), $V \doteq V_\sigma$ and for $o = (z_i, u_i) \in \mathcal{O}$, $h = (h_i) \in H$, $v \in C_0^1(\mathbb{R}^d)$, $\zeta_o(h) \doteq \sum_i K_\sigma(z_i, \cdot)h_i u_i$, $\xi_o(v) = (v(z_i), u_i)$ and $c_o(h) = |\zeta_o(h)|_{V_\sigma}^2$. This deformation module satisfies also UEC and will be used in examples of Section 5.1.

Translation with direction updated by adjoint action. This example is the same as the previous one, except that we change the update rule for the direction u of the translation. The infinitesimal action on geometrical descriptors writes now $\xi_o(v) = (v(z), -dv_z^T u)$ (with $o = (z, u)$). The integration of this infinitesimal action gives the action of the deformation ϕ on the geometrical descriptors o as $(\phi, o = (z, u)) \mapsto (\phi(z), (\mathrm{d}\phi_{\phi(x)}^{-1})^T u)$, where the second part is the so-called adjoint action. If one considers the direction u as the normal to an infinitesimal surface patch, then $(\mathrm{d}\phi_{\phi(x)}^{-1})^T u)$ is the direction of the normal of the deformed surface patch. If we consider a compound module combining this deformation module with others, the direction of the translation at integration time t will only depend of the generated global diffeomorphism ϕ and the direction u at time $t = 0$. As above, the combination of P of these modules also satisfy UEC. They will be used in examples in Section 5.2 and 5.3.

2.3.6. Silent modules. We present here a last example of deformation modules: modules generating a vector field that is always null. For a choice of $\mathcal{O}$, a C^k-shape space of order ℓ with its infinitesimal action ξ, we set $H \doteq \{0\}$ (null space for the

controls) and for $o \in \mathcal{O}$, $\zeta_o : h \in H \mapsto 0$, $c_o : h \in H \mapsto 0$. Then $M = (\mathcal{O}, H, \zeta, \xi, c)$ defines a C^k-deformation module of order ℓ satisfying UEC, which will be referred to as **the silent deformation module induced by shape space** $\mathcal{O}$. For instance, if $\mathcal{O}$ is the shape space made of a collection of points as in the examples of landmarks (Section 2.1.1), then these points, considered as a silent module, will feel the velocity field generated by active modules located around them, and move accordingly, but will not contribute to this velocity field. The introduction of silent modules is necessary if one wants to use active deformation modules whose geometric descriptors do not derive from the input shape data, for instance for scaling module whose center is not forced to be a vertex or point of the input shape data, or for the direction of a translation which is not forced to be normal to a surface mesh.

3. Sub-Riemannian setting. We explain here how modular large deformations may be used to provide shape spaces with a sub-Riemannian metric, and therefore define a distance between shapes.

We consider a C^k-deformation module $M = (\mathcal{O}, H, \zeta, \xi, c)$ of order $\ell \in \mathbb{N}^*$. A geometrical descriptor of the module is a "shape" in the shape space $\mathcal{O}$ and, intuitively, the orbit of this geometrical descriptor under the action of a regular group of diffeomorphisms forms a Riemannian manifold. If one considers only the diffeomorphisms which result from the integration of a modular velocity field (generated by the module M), then one provides this Riemannian manifold with a sub-Riemannian structure (Section 3.1). This construction allows the definition of a sub-Riemannian distance and of optimal trajectories between two geometrical descriptors in the shape space $\mathcal{O}$ (Section 3.2). In general, M is a compound module built from a silent module induced by the shape space of input shape data and user-defined deformation modules. Therefore, the sub-Riemannian distance is defined for the augmented data set containing the input data and the geometrical descriptors of the user-defined modules. In this case, this distance cannot be used to compare directly the input shape with another input shape data, since first, one may not assume that the second shape derives from the input shape by a modular deformation, and second, one does not know what would be the corresponding active modules in the second shape. This practical case will be detailed in Section 3.2.2.

3.1. A sub-Riemannian structure on $\mathcal{O}$. We suppose that M satisfies UEC, and we fix a space of vector field V and a constant $C > 0$ such that V is continuously embedded in $C_0^{\ell+k}(\mathbb{R}^d)$ and for all $o \in \mathcal{O}$ and for all $h \in H$, $\zeta_o(h) \in V$ and $|\zeta_o(h)|_V^2 \leq C c_o(h)$. We use now the notion of continuous sub-Riemannian structure on a manifold, following the definition given in [4].

Definition 6. *Let $\mathcal{M}$ be a manifold of finite dimension. We define a continuous sub-Riemannian structure on $\mathcal{M}$ as a triple $(\mathcal{E}, g, \rho)$, where : $\mathcal{E} \to \mathcal{M}$ is a smooth vector bundle on $\mathcal{M}$ endowed with a smooth, Riemannian metric g, and $\rho : \mathcal{E} \to T\mathcal{M}$ a continuous vector bundle morphism.*

The composition of the field generator ζ with the infinitesimal action ξ yields a continuous vector bundle morphism $\rho : (o, h) \in \mathcal{O} \times H \mapsto (o, \rho_o(h) = \xi_o \circ \zeta_o(h)) \in T\mathcal{O}$. Moreover, the cost c induces a smooth Riemannian metric g on the vector bundle $\mathcal{O} \times H$. Then $(\mathcal{O} \times H, g, \rho)$ defines a sub-Riemannian manifold, which we will denote $\mathcal{O}_H$. This structure is the key to define trajectories of modular deformations. At each shape $o \in \mathcal{O}$, the space $\zeta_o(H)$ is the space of vector fields that can be generated by o. The **horizontal space** $\rho_o(H) = \xi_o(\zeta_o(H))$ is the set of tangent vectors of $T_o\mathcal{O}$ that can be obtained by the action of the geometrical descriptor o on itself.

Remark 6. *Note that the dimension of the horizontal space at $o \in \mathcal{O}$, namely the*

rank of the sub-Riemannian structure at $o \in \mathcal{O}$, may depend on o. For instance, let us consider the deformation module $M = (\mathcal{O}, H, \zeta, \xi, c)$ obtained by combining a module $M^1 = (\mathcal{O}^1, H^1, \zeta^1, \xi^1, c^1)$ generating a local scaling in $\mathbb{R}^2$ at scale σ (see Section 2.1.1, second example) and a module $M^2 = (\mathcal{O}^2, H^2, \zeta^2, \xi^2, c^2)$ generating a sum of three local translations in $\mathbb{R}^2$ at scale σ (see Section 2.1.1, first example). For each geometrical descriptor $o = (o^1, o^2 = (z_i)_{1\leq i\leq 3}) \in \mathcal{O}^1 \times \mathcal{O}^2 = \mathbb{R}^2 \times (\mathbb{R}^2)^3$ and each control $h = (h^1, h^2 = (\alpha_i)_{1\leq i\leq 3}) \in H^1 \times H^2 = \mathbb{R} \times (\mathbb{R}^2)^3$, the generated vector field is $\zeta(o,h) = h^1 \sum_j K(z_j(o^1), \cdot) d_j(o^1) + \sum_i K(z_i, \cdot)\alpha_i$ with $z_j(o^1)$ and $d_j(o^1)$ defined in Section 2.1.1, Example 2. Then in the particular case where o^1 and $o^2 = (z_i)_i$ are such that $z_i = z_i(o^1)$ for $i \in \{1,2,3\}$, $\dim(\xi_o(\zeta_o(H))) = 2 \times 3 = 6$ while in other cases $\dim(\xi_o(\zeta_o(H))) = 2 \times 3 + 1 = 7$.

The cost c equips the vector bundle $\mathcal{O} \times H$ with a metric g, which may be used to derive a **metric $g^{\mathcal{O}}$ on $\mathcal{O}_H$**: if δo, $\delta o' \in \rho_o(H)$, let h, $h' \in Ker(\rho_o)^{\perp}$ (orthogonal for the metric g on $\mathcal{O} \times H$) such that $\rho_o(h) = \delta o$ and $\rho_o(h') = \delta o'$, then $g^{\mathcal{O}}_o(\delta o, \delta o') \doteq g_o(h, h')$. Note that for $\delta o \in \rho_o(H)$, $|\delta o|^2_o \doteq g^{\mathcal{O}}_o(\delta o, \delta o) = \inf\{c_o(h) \mid \rho_o(h) = \delta o,\ h \in H\}$. This sub-Riemannian metric $g^{\mathcal{O}}$ will be the one used to build the new distance on $\mathcal{O}$. f H such that $\rho_o(h) = \delta o$ and $c_o(h) = |\rho_o(h)|^2_o$.

We will now present some definitions and results given in [1], which allow the definition of a sub-Riemannian distance on $\mathcal{O}$.

Definition 7. *[1] Let $\mathcal{M}$ be a smooth manifold of finite dimension equipped with a continuous sub-Riemannian structure $(\mathcal{E}, g, \rho)$. A **horizontal system** is a curve $t \in [0,1] \mapsto (q(t), u(t)) \in \mathcal{E}$ such that $t \in [0,1] \mapsto u(t) \in \mathcal{E}_{q(t)}$ is of class L^2 (ie $\int_0^1 g_{q(t)}(u(t), u(t))\mathrm{d}t < \infty$), and its projection : $t \in [0,1] \mapsto q(t)$ is absolutely continuous and satisfies for almost every $t \in [0,1]$, $\dot{q}(t) = \rho_{q(t)}u(t)$. A **horizontal curve** is the projection $q(\cdot)$ to $\mathcal{M}$ of a horizontal system.*

Remark 7. *For $a, b \in \mathcal{O}$, space $\Omega_{a,b}$ (see Definition 5) is exactly the set of horizontal system connecting a and b.*

Definition 8. *[1] Let $o : [0,1] \longrightarrow \mathcal{O}$ be a horizontal curve, we define its **sub-Riemannian length**:*

$$l(o) = \int_0^1 |\dot{o}_t|_{o_t}\mathrm{d}t = \int_0^1 \sqrt{g^{\mathcal{O}}_{o_t}(\dot{o_t}, \dot{o_t})}\mathrm{d}t.$$

*For $a, b \in \mathcal{O}$ we can then define the **sub-Riemannian distance***

$$D(a,b) = inf\{l(o) \mid \exists h : (o,h) \in \Omega_{a,b}\}.$$

Lemma 1. *[1] Let o be a horizontal curve. There exists a horizontal curve $\tilde{o}$ and a Lipschitz bijective reparametrization $\gamma : [0,1] \longrightarrow [0,1]$ such that $\tilde{o} = o \circ \gamma$ and for a.e t, $|\dot{\tilde{o}}(t)|_{\tilde{o}(t)} = l(o)$.*

Definition 9. *[1] Let $a, b \in \mathcal{O}$ and $(o,h) \in \Omega_{a,b}$. We define the **length of** (o,h) by*

$$l(o,h) = \int_0^1 \sqrt{c_{o_t}(h_t)}\mathrm{d}t$$

and then

$$D_H(a,b) = \inf\{l(o,h) \mid (o,h) \in \Omega_{a,b}\}.$$

Remark 8. *If $\Omega_{a,b}$ is empty, both $D(a,b)$ and $D_H(a,b)$ have an infinite value.*

We will now show that studying D and D_H amounts to consider same trajectories, and that D is a real distance on $\mathcal{O}$.

Definition 10. *[1] Let $o : [0,1] \longrightarrow \mathcal{O}$ be a horizontal curve, for each $t \in [0,1]$, let $h^\star(t) \in H$ be the only element of H such that $c_o(h^\star(t)) = |\dot{o}_t|_o^2$. We say that $h^\star : t \mapsto h^\star(t)$ is the* ***minimal control associated with*** *o.*

Lemma 2. *[1] Let $o : [0,1] \longrightarrow \mathcal{O}$ a horizontal curve, its minimal control $h^\star$ is measurable and of class L^2.*

Thanks to the previous lemma, we can deduce the following proposition:

Proposition 5. *[1] Let $a, b \in \mathcal{O}$, $D(a,b) = D_H(a,b)$.*

We can now prove that D defines a distance on $\mathcal{O}$.

Proposition 6. *As M satisfies UEC, the sub-Riemanian distance D is a true distance on $\mathcal{O}$.*

Proof. It is clear that D is a pseudo-distance, we need to show that if $D(a,b) = 0$ then $a = b$. We will use a result proved in [5]: if we set for $\phi \in \mathrm{Diff}_0^\ell(\mathbb{R}^d)$, $d(Id, \phi) \doteq \inf\{\int_0^1 |v_t|_V \mathrm{d}t \mid \forall t\ v_t \in V \text{ and } \varphi_{t=1}^v = \phi\}$ (we remind that V is the Hilbert space of vector fields defined UEC) and for $a, b \in \mathcal{O}$, $dist(a,b) = \inf\{d(Id,\phi) \mid \phi \cdot a = b\}$ then $dist$ defines a distance (taking its value in $[0, +\infty]$).
Let $a, b \in \mathcal{O}$ such that $D(a,b) = 0$, then there exists $(o^n)_{n\in\mathbb{N}}$ such that for each n there exists h^n such that $(o^n, h^n) \in \Omega_{a,b}$ and $l(o^n) \longrightarrow 0$. By choosing h^n the minimal control of o^n, we also have $l(o^n, h^n) \longrightarrow 0$. Yet, thanks to the UEC, for each n: $dist(a,b) \leq \int_0^1 |\zeta_{o^n}(h^n)|_V \leq \sqrt{C} \int_0^1 \sqrt{c_{o^n}(h^n)} = \sqrt{C} l(o^n, h^n) \longrightarrow 0$. Then $a = b$. □

3.2. Optimal trajectories. Thanks to Proposition 5, minimizers of l in $\Omega_{a,b}$ (for $a, b \in \mathcal{O}$) characterise the distance between a and b. However the quantity $l(o,h) = \int_0^1 \sqrt{c_o(h)}$ is hard to study and then it is necessary to link it with the energy $E(o,h) = \int_0^1 c_o(h)$.

Proposition 7. *[1] Let $a, b \in \mathcal{O}$ and let (o,h) be in $\Omega_{a,b}$. Then (o,h) minimizes E in $\Omega_{a,b}$ if and only if it minimizes l in $\Omega_{a,b}$ and its cost $c_o(h)$ is constant.*

Remark 9. *We deduce that along minimizers (if they exist), $D_H(a,b)^2 = l(o,h)^2 = E(o,h)$.*

3.2.1. Existence of optimal trajectories. Propositions 5 and 7 show that calculating the distance $D(a,b)$ between two elements a, b of $\mathcal{O}$ amounts to looking for horizontal system minimizing the energy E, which is easiest to study. Therefore, in the next paragraph we characterize horizontal system minimizing E. We need here to restrain ourselves to a certain type of shapes, obtained through an adaptation of a definition given by S. Arguillère in [5].

Definition 11. *An element o of $\mathcal{O}$, is said to be of* ***compact support*** *if there exists a compact set K of $\mathbb{R}^d$ such that for all $\phi \in D_0^\ell(\mathbb{R}^d)$, $\phi \cdot o$ only depends on $\phi_{|K}$ and $\phi \in D_K^\ell \mapsto \phi \cdot o$ is continuous with $D_K^\ell := \{\phi_{|K} \mid \phi \in D_0^\ell(\mathbb{R}^d)\}$ equipped with the distance deduced from the norm on $\{v_{|K} \mid v \in C_0^\ell(\mathbb{R}^d)\}$: $|v|_{\ell,K} = \sup\{|\frac{\partial^{\ell_1+\cdots+\ell_d} v(x)}{\partial x_1^{\ell_1} \cdots x_d^{\ell_d}}| \mid x \in K, (\ell_1, \cdots, \ell_d) \in \mathbb{N}^d, \ell_1 + \cdot + \ell_d \leq \ell\}$.*

Remark 10. *Examples presented in Section 2.3 correspond to geometrical descriptors of compact support.*

Lemma 3. *Let $o \in \mathcal{O}$ be of compact support and let K be a compact set of $\mathbb{R}^d$ such that for all ϕ of $D_0^\ell(\mathbb{R}^d)$ $\phi \cdot o$ only depends on K and $\phi \in D_K^\ell \mapsto \phi \cdot o$ is continuous. Then for each $\phi \in D_0^\ell(\mathbb{R}^d)$, $\phi \cdot o$ is of compact support and $\psi \in D_0^\ell(\mathbb{R}^d) \mapsto \psi \cdot (\phi \cdot o)$ only depends on the compact set $\phi(K)$.*

Proof. Let ϕ, ψ and ψ' be elements of $D_0^\ell(\mathbb{R}^d)$ such that $\psi_{|\phi(K)} = \psi'_{|\phi(K)}$. Then $\psi\cdot(\phi\cdot o) = (\psi\circ\phi)\cdot o = (\psi'\circ\phi)\cdot o$ because $\psi\circ\phi_{|K} = \psi'\circ\phi_{|K}$. Then $\psi\cdot(\phi\cdot o) = \psi'\cdot(\phi\cdot o)$ and we conclude that for all $\psi \in D_0^\ell(\mathbb{R}^d)$, $\psi\cdot(\phi\cdot o)$ only depends on $\phi(K)$. Besides from Faà di Bruno's formula one gets that for ϕ an element of $D_0^\ell(\mathbb{R}^d)$, there exists a constant $C_{|\phi|_{\ell,K}}$ such that for each ψ, ψ' in $D_0^\ell(\mathbb{R}^d)$, $|\psi\circ\phi - \psi'\circ\phi|_{\ell,K} \leq C_{|\phi|_{\ell,K}}|\psi - \psi'|_{\ell,\phi(K)}$. Then for $\phi D_0^\ell(\mathbb{R}^d)$, $\psi \in D_0^\ell(\mathbb{R}^d) \mapsto \psi\cdot(\phi\cdot o)$ is continuous and therefore $\phi\cdot o$ is of compact support.

□

Let a and b be two geometrical descriptors of compact support, we prove here the existence of minimizers of the energy (if $\Omega_{a,b}$ is non-empty), and therefore of trajectories reaching sub-Riemannian distance D.

Theorem 1. *We recall that the deformation module M satisfies UEC. If $\Omega_{a,b}$ is non-empty, the energy E (see Definition 5) reaches its minimum on $\Omega_{a,b}$.*

Proof. Let (o^n, h^n) be a minimizing sequence of E in $\Omega_{a,b}$ and let, for each n, φ^n be the flow associated to (o^n, h^n) as defined in Proposition 4: $\varphi^n = \varphi^{v_n}$ with $v_n = \zeta_{o^n}(h^n)$. Since $\int_0^1 |v_t^n|_V^2 \mathrm{d}t \leq CE(o^n, h^n)$ (from UEC), the sequence $(v_n)_n$ is bounded in $L^2([0,1], \mathbb{R}^d)$ so up to the extraction of a subsequence we can assume that v^n converges weakly to $v^\infty \in L^2([0,1], \mathbb{R}^d)$. Let us define φ^∞ the flow of v^∞. As a is of compact support, there exists K compact of $\mathbb{R}^d$ such that for all $\phi \in D_0^\ell(\mathbb{R}^d)$, $\phi\cdot a$ only depends on $\phi_{|K}$. Moreover, as K is compact, [11] shows that $\sup_{(t,x)\in[0,1]\times K} |\phi_t^n(x) - \phi_t^\infty(x)| \longrightarrow 0$ so that, as $\phi \in D_K^\ell \mapsto \phi\cdot a$ is continuous, $o_\cdot^n = \phi_\cdot^n \cdot a$ converges to $o_\cdot^\infty = \phi_\cdot^\infty \cdot a$ uniformly on $[0,1]$. Therefore there exists a compact set L of $\mathcal{O}$ such that for all t, $o_t^\infty \in L$ and for all n, $o_t^n \in L$. Then $\sup_{o\in L} ||c_o^{-1}||$ (see proof of Proposition 4) is finite and for each n: $\int_0^1 |h^n(t)|_H^2 \mathrm{d}t \leq \sup_{o\in L} ||c_o^{-1}|| E(o^n, h^n)$. Therefore h^n is bounded in $L^2([0,1], H)$ so up to the extraction of a subsequence we can assume that h^n converges weakly to $h^\infty \in L^2([0,1], H)$. Let us show that $(o^\infty, h^\infty) \in \Omega_{a,b}$. Let $w \in L^2([0,1], V)$, we have

$$\begin{aligned} |\int_0^1 \langle v_t^\infty - \zeta_{o_t^\infty}(h_t^\infty), w_t\rangle_V \mathrm{d}t| \quad \leq \quad & |\int_0 \langle v_t^\infty - \zeta_{o_t^n}(h_t^n), w_t\rangle_V \mathrm{d}t| \\ & + |\int_0^1 \langle \zeta_{o_t^n}(h_t^n) - \zeta_{o_t^\infty}(h_t^n), w_t\rangle_V \mathrm{d}t| \\ & + |\int_0^1 \langle \zeta_{o_t^\infty}(h_t^n) - \zeta_{o_t^\infty}(h_t^\infty), w_t\rangle_V \mathrm{d}t. \end{aligned}$$

As $\zeta_{o^n}(h^n)$ converges weakly to v^∞ the first term tends to 0. In the same way, $h \in L^2([0,1], H) \mapsto \int_0^1 \langle \zeta_{o_t^\infty}(h_t), w_t\rangle_V \mathrm{d}t$ is continuous since, as ζ is of class C^1 with respect to o, $o \in L \mapsto \zeta_o \in L(H, V)$ is bounded on L (which contains o_t^∞ for all t). Then as h^n weakly converges to h^∞, $\int_0^1 \langle \zeta_{o_t^\infty}(h_t^n) - \zeta_{o_t^\infty}(h_t^\infty), w_t\rangle_V \mathrm{d}t$ tends to 0. Therefore

$$\begin{aligned} |\int_0^1 \langle v_t^\infty - \zeta_{o_t^\infty}(h_t^\infty), w_t\rangle_V \mathrm{d}t| \quad & \leq \quad \limsup |\int_0^1 \langle (\zeta_{o_t^n} - \zeta_{o_t^\infty})(h_t^n), w_t\rangle \mathrm{d}t| \\ & \leq \quad \limsup (\int_0 |w_t|_V^2 \mathrm{d}t \int_0^1 ||\zeta_{o_t^n} - \zeta_{o_t^\infty}||^2 |h_t^n|^H \mathrm{d}t)^{1/2} \\ & = \quad 0 \end{aligned}$$

since h^n is bounded in $L^2([0,1], H)$ and, as o^n uniformly converges to o^∞, $||\zeta_{o_t^n} - \zeta_{o_t^\infty}|| \longrightarrow 0$. Since w is arbitrary, $v^\infty = \zeta_{o^\infty}(h^\infty)$ and $\dot{o}_t^\infty = \xi_{o_t^\infty}(v_t^\infty)$ so that $(o^\infty, h^\infty) \in \Omega_{a,b}$. We now need to show that $E(o^\infty, h^\infty) = \lim E(o^n, h^n)$. Since $h \mapsto \int_0^1 c_{o_t^\infty}(h)\mathrm{d}t$ is continuous and convex $\int_0^1 c_{o_t^\infty}(h_t^\infty)\mathrm{d}t \leq \liminf \int_0^1 c_{o_t^\infty}(h_t^n)\mathrm{d}t$. Moreover since c is a continuous metric, there exists $C : o \mapsto C_o$ continuous such that $c_o(h) = (C_o h|h)_{H^\star}$, so that $|\int_0^1 (c_{o_t^\infty}(h_t^n) - c_{o_t^n}(h_t^n))\mathrm{d}t| \leq (\sup_t ||C_{o_t^\infty} - C_{o_t^n}||) \int_0^1 |h_t^n|_H^2 \mathrm{d}t \longrightarrow 0$.

Then we obtain $\int_0^1 c_{o_t^\infty}(h_t^\infty)\mathrm{d}t \leq \liminf \int_0^1 c_{o_t^n}(h_t^n)\mathrm{d}t$ and $E(o^\infty, h^\infty) = \lim E(o^n, h^n)$. □

3.2.2. Extended optimal trajectories and their equations. We will now explain how optimal trajectories can be computed. The following construction is valid for generic deformation modules satisfying UEC, but we will detail here the practical case where one wants to study the differences between shapes belonging to a common shape space $\mathcal{F}$ thanks to a user-defined deformation module $M^1 = (\mathcal{O}^1, H^1, \zeta^1, \xi^1, c^1)$ independent from the shape space $\mathcal{F}$ (and satisfying UEC). We then consider the silent deformation module induced by shape space $\mathcal{F}$ $M^2 = (\mathcal{F}, H^2, \zeta^2, \xi^2, c^2)$ (see Section 2.3.6) and its combination $M = (\mathcal{O}, H, \zeta, \xi, c)$ with M^1 (see Section 2.1.2). M is a deformation module satisfying UEC and for $a = (o, f) \in \mathcal{O} = \mathcal{O}^1 \times \mathcal{F}$, $h = (h^1, 0) \in H$, the application to a of the vector field generated by a and the control h is

$$\xi_o \circ \zeta_o(h) = \xi_o \circ \zeta^1_{o^1}(h^1) = (\xi^1_{o^1} \circ \zeta_{o^1}(h^1), \xi^2_f \circ \zeta_{o^1}(h^1)).$$

Let fix a and b two shapes of $\mathcal{O} = \mathcal{O}^1 \times \mathcal{F}$ of compact support. We distinguish in these shapes the input data (component belonging to $\mathcal{F}$) and the geometrical descriptors of the user-defined modules (component belonging to $\mathcal{O}$): $a = (o^1, f^1)$ and $b = (o^2, f^2)$. We consider now the practical case where the target shape b does not derive from the source shape a by the action of a modular diffeomorphism. Instead, we propose to find the "optimal" shape $\hat{b} = (\hat{o}^2, \hat{f}^2)$ in the orbit of a, i.e. such that $\hat{b} = \phi.a$, so that the deformed shape $\hat{b} = \phi.a$ falls as close as possible to b in the sense of a measure μ. As in practice we only know the component f^2 of b, $\mu(\hat{b}, b)$ will only depend on $\hat{f}^2$ and f^2. Simultaneously we derive the differential equations to compute the optimal trajectory between a and $\hat{b}$.

This construction may be seen from a statistical point of view using the following generative model : we suppose that the probability density function (with respect to a given reference measure) of the transformation $\hat{b}$ of a through a modular large deformation is proportional to $\exp(-D(a, \hat{b})^2/\sigma_0^2)$ where D is the sub-Riemannian distance built on $\mathcal{O}$ (see Section 3). We also suppose that, the density of the conditionnal distribution of b given $\hat{b}$ is proportional to $\exp -(\mu(\hat{b}, b)/\sigma_1^2)$. Eventually, we want to estimate the most probable $\hat{b}$, given a and b: the likelihood of $\hat{b}$ knowing a and b is proportional to $\exp(-D(a, \hat{b})^2/\sigma_0^2) \times \exp(-\mu(\hat{b}, b)/\sigma_1^2)$ and its maximisation amounts to the minimisation of:

$$\mu(\hat{b}, b)/\sigma_1^2 + D(a, \hat{b})^2/\sigma_0^2 \tag{3.1}$$

As shown in Theorem 1, for each $\hat{b}$ such that $D(a, \hat{b})$ is finite, there exists a trajectory $h^\star$ of controls such that $D(a, \hat{b}) = \int \sqrt{c_o(h^\star)}$, with o satisfying $o_{t=0} = a$ and $\dot{o} = \xi_o \circ \zeta_o(h^\star)$. Then minimizing (3.1) with respect to $\hat{b}$ amounts to minimizing the following quantity with respect to h

$$\frac{1}{\sigma_1^2}\mu(o_{t=1}, b) + \frac{1}{\sigma_0^2}\Big(\int \sqrt{c_{o_t}(h_t)}dt\Big)^2 \tag{3.2}$$

where $o_{t=0} = a$ and $\dot{o} = \xi_o \circ \zeta_o(h)$.

Besides, thanks to Proposition 7 we know that along minimizing trajectories $\left(\int \sqrt{c_{o_t}(h_t)}dt\right)^2 = \int c_{o_t}(h_t)dt$. Then maximizing (3.2) amounts to minimizing the following quantity with respect to h

$$J_{a,b}(h) = \frac{1}{\sigma^2}\mu(o_{t=1}, b) + \int c_{o_t}(h_t)dt \tag{3.3}$$

with $\sigma = \frac{\sigma_1}{\sigma_0}$, $o_{t=0} = a$ and $\dot{o} = \xi_o \circ \zeta_o(h)$. Note that, as $M = \mathcal{C}(M^1, M^2)$ (with M^2 silent) and μ only depends on the component belonging to $\mathcal{F}$, $J_{a,b}$ can be written :

$$J_{a,b}(h) = \frac{1}{\sigma^2}\mu(f_{t=1}, f^2) + \int c^1_{o^1}(h^1)$$

with $o_{t=0} = a$, $b = (o^2, f^2)$, $o = (o^1, f)$, $h = (h^1, 0)$ and $\dot{o} = (\dot{o}^1, \dot{f}) = \left(\xi^1_{o^1} \circ \zeta^1_{o^1}(h^1), \xi^2_f \circ \zeta^1_{o^1}(h^1)\right)$.

A trajectory o of $\mathcal{O}$ starting at a such that there exists $h \in L^2([0,1], H)$ so that (o, h) is a horizontal system (see Definition 7) and h minimizes $J_{a,b}$, will be called an **optimal trajectory starting at** a. These trajectories can be well characterized thanks to the next result, which we prove following the idea of the proof of [4].

Theorem 2. *We recall that $M = (\mathcal{O}, H, \zeta, \xi, c)$ is a C^k-deformation module of order C^ℓ satisfying UEC with $k, l \geq 1$. We suppose that μ is C^1. If $h \in L^2([0,1], H)$ minimizes functional $J_{a,b}$ then there exists a path $\eta : t \in [0,1] \longrightarrow \eta_t \in T^*_{o_t}\mathcal{O}$ such that with*

$$\mathcal{H} : (o, \eta, h) \in T^*\mathcal{O} \times H \mapsto \Big(\eta | \xi_o(\zeta_o(h))\Big) - \frac{1}{2}c_o(h)$$

*the **Hamiltonian** of the system, $\eta_{t=1} = -\partial_1\mu(o_{t=1}, b)$ and (in a local chart)*

$$\left\{\begin{array}{rcl} \frac{\mathrm{d}o}{\mathrm{d}t} & = & \xi_o \circ \zeta_o(h) \\ \frac{\mathrm{d}\eta}{\mathrm{d}t} & = & -\frac{\partial \mathcal{H}}{\partial o} \\ \frac{\partial \mathcal{H}}{\partial h} & = & 0 \end{array}\right. \tag{3.4}$$

Proof. In this proof we suppose that $\mathcal{O}$ is an open subset of $\mathbb{R}^N$. As previously we associate to each $h \in L^2([0,1], H)$ the trajectory o^h of $\mathcal{O}$ such that $o_{t=0} = a$ and $\dot{o}^h_t = \xi_{o_t} \circ \zeta_{o_t}(h_t)$. For each $h \in L^2([0,1], H)$, o^h is absolutely continuous on $[0,1]$ (see Proposition 2) and then belongs to $H^1_a([0,1], \mathcal{O})$ (elements in $H^1([0,1], \mathcal{O})$ starting at a). We define the new functional $\tilde{J} : (o, h) \in H^1([0,1], \mathcal{O}) \times L^2([0,1], H) \mapsto \int_0^1 c_o(h) + \mu(o_{t=1}, b)$. Then $h \in L^2([0,1], H)$ minimizes J if and only if (o^h, h) minimizes $\tilde{J}$ under the constraint $0 = \Gamma(o, h) \doteq \dot{o} - \xi_o \circ \zeta_o(h)$. Functions $\tilde{J}$ and Γ are of class C^1 and $\partial_o\Gamma$ is an is an isomorphism for each $o \in H^1([0,1], \mathcal{O})$. Indeed let $o \in H^1([0,1], \mathcal{O})$ and $\alpha \in L^2([0,1], \mathbb{R}^N)$, we can define $\delta o \in H^1_0([0,1], \mathbb{R}^N)$ by $\delta o(t=0) = 0$ and $\dot{\delta o} = \partial_o(\xi_o \circ \zeta_o(h)) + \alpha$ (the solution is well defined). Then $\alpha = \partial_o\Gamma(o, h) \cdot \delta o$ and therefore $\partial_o\Gamma(o, h)$ is surjective. Moreover $\partial_o\Gamma(o, h)$ in injective as for $\delta o \in H^1_0([0,1], \mathbb{R}^N)$, $\partial_o\Gamma(o, h)\delta o = 0$ implies $\delta o = 0$ by Cauchy uniqueness. We therefore deduce (thanks to the implicit function theorem) that $\Gamma^{-1}(\{0\})$ is a manifold.

Let (o, h) be a minimizer of $\tilde{J}$ over the set $\Gamma^{-1}(\{0\})$, from [20] (Theorem 4.1) can be shown that there exists a non trivial Lagrange multiplier $\eta \in L^2(0, 1, \mathbb{R}^N)^\star = L^2(0, 1, \mathbb{R}^{N,\star})$ such that $\mathrm{d}\tilde{J}_{(o,h)} + (d\Gamma_{(o,h)})^\star(\eta) = 0$. It is shown in [5] that $\dot{\eta} = -\partial_o\mathcal{H}$ and $\partial_h\mathcal{H} = 0$. □

Remark 11. *As c_o is positive definite, there exists an invertible symmetric operator $C : o \in \mathcal{O} \mapsto C_o \in L(H, H^\star)$ such that for all $(o, h) \in \mathcal{O} \times H$, $c_o(h) = (C_o h|h)$. Then the third equality in (3.4) allows to compute h: $h = C_o^{-1}\rho_o^\star\eta$ with $\rho_o = \xi_o \circ \zeta_o$ and $\rho_o^\star$ such that $(\eta|\rho_o(h))_{T_o\mathcal{O}} = (\rho_o^\star\eta|h)_H$.*

We define the **reduced Hamiltonian**

$$\mathcal{H}_r(o, \eta) \doteq \mathcal{H}(o, \eta, C_o^{-1}\rho_o^\star(\eta)) = \frac{1}{2}(\rho_o^\star(\eta)|C_o^{-1}\rho_o^\star(\eta)) = \frac{1}{2}c_o(C_o^{-1}\rho_o^\star(\eta))$$

and as $\nabla_h\mathcal{H} = 0$, the system of equations (3.4) can be written:

$$\left\{ \begin{array}{rcl} \frac{\mathrm{d}o}{\mathrm{d}t} & = & \frac{\partial \mathcal{H}_r}{\partial \eta} = \rho_o(C_o^{-1}\rho_o^\star(\eta)) \\ \frac{\mathrm{d}\eta}{\mathrm{d}t} & = & -\frac{\partial \mathcal{H}_r}{\partial o} \end{array} \right. \tag{3.5}$$

Proposition 8. *If the module M is C^j of order ℓ with $j, \ell \geq 2$ then solutions of Equation (3.5) exist for any $(o_{t=0}, \eta_{t=0})$ and are totally defined by these initial values.*

Proof. In this case $(o, \eta) \mapsto H_r(o, \eta)$ is of class at least C^2 so $(o, \eta) \mapsto (\frac{\partial \mathcal{H}_r}{\partial \eta}, -\frac{\partial \mathcal{H}_r}{\partial o})$ is at least C^1.

□

Then by choosing an initial momentum $\eta \in T_o^*\mathcal{O}$ one can generate an optimal trajectory starting at a. Optimal trajectories are parametrized by initial values of geometrical descriptor and momentum, so in dimension $2 \times (\dim(\mathcal{O}^1) + \dim(\mathcal{F}))$.

Remark 12. *It is important to note that even though $h = (h^1, 0)$, $\rho_o^\star(\eta) \neq \rho_{o^1}^{1,\star}(\eta^1)$ (with $\eta = (\eta_1, \eta^2) \in T_{o^1}^*\mathcal{O}^1 \times T_f^*\mathcal{F}$ and $\rho_{o^1}^1 = \xi_{o^1}^1 \circ \zeta_{o^1}^1$). Then, even though the component of geometrical descriptor belonging to $\mathcal{F}$ generates only a null vector field, its initial momentum has an influence on the trajectory.*

Remark 13. *From Equation (3.4) we re-deduce that cost $c_o(h)$ is constant along minimizing trajectories. Indeed let (o, h) be such an optimal trajectory and let η be the trajectory of momenta, such that (o, η, h) satisfies Equation (3.4). Then, as $\rho_o(h) = \xi_o \circ \zeta_o(h) = \nabla_\eta\mathcal{H}$, one gets: $\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = \partial_o\mathcal{H}\frac{\mathrm{d}o}{\mathrm{d}t} + \partial_\eta\mathcal{H}\frac{\mathrm{d}\eta}{\mathrm{d}t} + \partial_h\mathcal{H}\frac{\mathrm{d}h}{\mathrm{d}t} = (\nabla_o\mathcal{H}, \nabla_\eta\mathcal{H}) - (\nabla_\eta\mathcal{H}, \nabla_o\mathcal{H}) = 0$. So the Hamiltonian is constant along optimal trajectories and as it is equal to half of the cost, $\int_0^1 c_o(h) = c_{o_{t=0}}(h_{t=0})$ along minimizing trajectories.*

4. Modular analysis of shape variability. The goal of this section is to show how the previous geometrical construction may be used to infer statistical properties from a series of shape data belonging to a common shape space $\mathcal{F}$. The idea is that these properties may be studied thanks to one deformation module obtained through the combination of the silent deformation module induced by $\mathcal{F}$ and a user-defined active deformation module. Computing an atlas of the data set will correspond to computing one initial value of the geometrical descriptor, and modular large deformations generated by it, bringing its silent component as close as possible to target shapes. Geometrical descriptors of the silent deformation module will be referred to as the template shape and plays the role of an "average" of the data set. The active deformation module is a dictionary, in which the shape variations seen in the data may be decomposed. The choice of the types of modules is left to the user, whereas parameters such as the initial position of their geometrical descriptors are to be optimised for a given training shape data set. In the spirit of [9], the user provides an example of template shape with the desired topology (i.e. number of points, edges between points, number of connected components, etc..), and the atlas construction method will optimize its shape (i.e. the position of the points or vertices) to be at the "center"

of the training samples. The method needs to optimize at the same time the other geometrical descriptors and find the P modular large deformations ϕ^k, so that each warped template resembles as much as possible to one of the target shapes f^k_{target}. Given the initial position of geometrical descriptors, modular large deformations that need to be estimated are parametrized by initial momenta. So finally the method will estimate one initial position of geometrical descriptors and P initial momenta. These momenta may be used subsequently for other statistical tasks, such as clustering, classification or regression with covariates. We will now detail this construction, and then present the algorithmic method.

4.1. Building an atlas of shapes. Let $\mathcal{F}$ be a shape space of order $\ell \geq 2$, let f^k_{target} be a series of P different shapes of $\mathcal{F}$, and let $\tilde{M} = (\tilde{\mathcal{O}}, \tilde{H}, \tilde{\zeta}, \tilde{\xi}, \tilde{c})$ be the deformation module thanks to which we want to study this series of shapes. As previously we build $M = (\mathcal{O}, H, \zeta, \xi, c)$ the compound module of $\tilde{M}$ and the silent deformation module induced from $\mathcal{F}$. We suppose that M is C^j of order ℓ and satisfies UEC. The estimation of the atlas may be done in a coherent Bayesian framework. However, the derivation of a Bayesian approach as in [3] is out of the scope of this paper, and we propose here a more straightforward extension of the geometrical construction of Section 3.2.2. It is the analogue of the concept of Fréchet mean, which has been used intensively in the field of Computational Anatomy [21, 18, 9]. This approach amounts here to minimising the following quantity with respect to the geometrical descriptors of the template module o_{temp} and P trajectories $h^k \in L^2([0,1], H)$:

$$\frac{1}{\sigma^2} \sum_k \mu(f^k_{t=1}, f^k_{target}) + \int c_{o^k}(h^k) \tag{4.1}$$

with for each k, $o^k_{t=0} = o_{temp}$ and $\dot{o}^k = \xi_{o^k} \circ \zeta_{o^k}(h^k)$.

Like in Section 3.2.2, can be shown that for a fixed value of o_{temp}, the minimiser $(h^k)_k$ of (4.1) is such that there exist P trajectories $\eta^k : t \in [0,1] \mapsto \eta^k_t \in T^*_{o^k_t}\mathcal{O}$ so that, with $\mathcal{H} : (o, \eta, h) \in T^*\mathcal{O} \times H \mapsto (\eta | \xi_o(\zeta_o(h))) - \frac{1}{2} c_o(h)$,

$$\left\{ \begin{array}{rcl} h^k & = & h^{o^k, \eta^k} \doteq C^{-1}_{o^k} \rho^\star_{o^k} \eta^k \\ \frac{\mathrm{d}o^k}{\mathrm{d}t} & = & \xi_{o^k}(\zeta_{o^k}(h^k) \\ \frac{\mathrm{d}\eta^k}{\mathrm{d}t} & = & -\frac{\partial \mathcal{H}}{\partial o}(o^k, \eta^k, h^k) \end{array} \right. \tag{4.2}$$

where $C : o \in \mathcal{O} \mapsto C_o$ is smooth and satisfies for each $h \in H$, $c_o(h) = (C_o h | h)_H$, and $\rho^\star_o$ is such that, for all $h \in H$, $\eta \in T^*_o\mathcal{O}$, $(\eta | \rho_o(h)) = (\rho^\star_o(\eta), h)$. Then whole trajectories (o^k, η^k, h^k) are defined by initial values of o^k and η^k. Moreover, thanks to Remark 13, the cost is constant along optimal trajectories and so for each k: $c_{o_{temp}}(h^{o_{temp}, \eta^k_0}) = \int_0^1 c_{o^k}(h^{o^k, \eta^k})$. Therefore computing the atlas of shapes f^k_{target}, $k = 1 \cdots P$ thanks to the deformation module M^1 amounts to minimizing:

$$E(o_{temp}, (\eta^k_0)_k) = \sum_{k=1}^{P} c_{o_{temp}}(h^{o_{temp}, \eta^k_0}) + \frac{1}{\sigma^2} \mu\Big(\phi^{o^k, \eta^k} \cdot f_{temp}, f^k_{target}\Big) \tag{4.3}$$

where for each k, (o^k, η^k) starts at (o_{temp}, η^k_0) and satisfies Equation (4.2), and $\phi^{o^k, h^k} = \varphi^{v^k}_{t=1}$ with $v^k = \zeta_{o^k}(h^k)$. The first term of this sum will be referred to

as the **Regularity term** while the second one will be called the **Data term**.

Note that the initial value of geometrical descriptor o_{temp} is common to all subjects but that the trajectory o^k (starting at o_{temp}) obtained by integrating Equation (4.2) is specific to each subject as it depends on the initial momentum η_0^k, which is specific to each subject.

We emphasize here that the estimated $o_{temp} = (o^1_{temp}, f_{temp})$ has two components. The second one f_{temp} is the template and corresponds to an average of the data set, while the first one o^1_{temp} is a common geometric characterization of the variability among the population of shapes.

4.2. Algorithmic method. In order to minimize Functional (4.3), we use a gradient descent scheme. The first term of the sum only depends on initial values of (o_{temp}, η_0^k) so the gradient can be easily computed. The second term depends on $\phi_{t=1}^{(o^k,\eta^k)} \cdot f_{temp}$, so in order to calculate its gradient we use the following result (see [4]):

Proposition 9. *Let $n \in \mathbb{N}$, let U be an open subset of $\mathbb{R}^n$, $w : U \longrightarrow \mathbb{R}^n$ be a complete C^j vector field on U ($j \geq 1$), G be the function of class C^1 defined on U by $G(q_0) = g(q_{t=1})$ where g is a function of class C^1 and $q : [0,1] \longrightarrow \mathbb{R}^n$ satisfies $\dot{q}_t = w(q_t)$ for almost every $t \in [0,1]$ and $q_{t=0} = q_0$. Then $\nabla G(q_0) = Z(1)$ where $Z : [0,1] \longrightarrow \mathbb{R}^n$ is the solution of $\dot{Z}(t) = dw^T_{q_{(1-t)}} Z_t$ for almost every $t \in [0,1]$ and $Z(0) = \nabla g(q_{t=1})$.*

In our case let us fix k and detail the computation of the gradient of the k-th term of the sum (4.3). We have $U = T^*\mathcal{O} = T^*\mathcal{O}^1 \times T^*F$, $q = (o^k, \eta^k) : t \in [0,1] \longrightarrow (o^k_t, \eta^k_t) \in T^*\mathcal{O}$, $g(o^k_{t=1}, \eta^k_{t=1}) = \mu\Big(f_{k,t=1}, f^k_{target}\Big)$, and $w(o^k, \eta^k) = (\partial_\eta \mathcal{H}, -\partial_o \mathcal{H})(o^k, \eta^k, h^{o^k,\eta^k})$ with H and h^{o^k,η^k} defined in Equation (4.2). Then the gradient of $\mu(\phi_{t=1}^{(o^k,\eta^k)} \cdot f_{temp}, f^k_{target})$ with respect to initial values $(o^k_{t=0}, \eta^k_{t=0})$ can be computed by first integrating Equation 4.2). Then we initialize an adjoint variable $\nu \in T_{(o^k_{t=1},\eta^k_{t=1})}TO$ with initialized with $\nu(t=1) = \nabla_{(o^k_{t=1},\eta^k_{t=1})}\mu(f^k_{t=1}, f^k_{target})$ (where $o^k = (o^{1,k}, f^k)$) such that only the dual variable of $f^k_{t=1}$ is non zero. Finally by integrating $\dot{\nu}_t = dw^T_{(o^k_{1-t},\eta^k_{1-t})}\nu_t$, one obtains $\nabla_{(o^k_{t=0},\eta^k_{t=0})}\mu(\phi_{t=1}^{(o^k,\eta^k)} \cdot f_{temp}, f^k_{target}) = \nu(t=0)$.

The proposed modular deformation framework is very well adapted for an implementation in an object-oriented language. All kinds of modules inherit from a single abstract module class, which contains abstract methods for the implementation of the field generator ζ, the infinitesimal action ξ (which takes another deformation module as argument), the cost c, and their differentials. A meta class contains a list of module, and implements the simple rules presented in Section 2.1.2 for the combination of modules. Now, all the remaining part of the codes depends on this meta class only, so that the definition of new modules is as simple as "plug-and-play". The remaining part of the code encodes the integration of the differential equation to compute large deformations from initial values of geometrical descriptors and momenta for each module, the gradient of a data attachment term, and the optimization method. The code has been implemented in the software Deformetrica [9], so that one may use a large collection of similarity metrics μ, for point clouds, curve and surface meshes in 2D and 3D.

5. Numerical experiments.

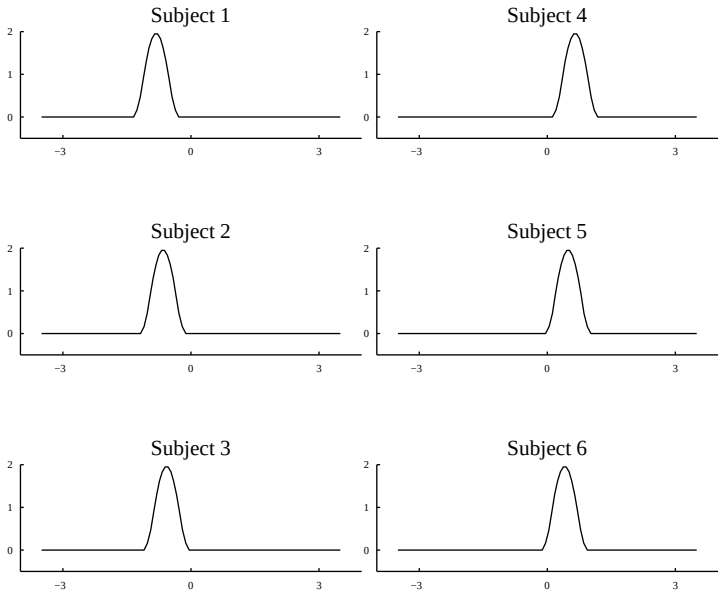

FIG. 5.1. ***Building an atlas using a prior***. *Target shapes*

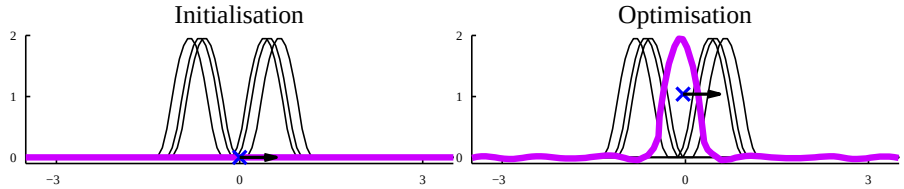

FIG. 5.2. ***Building an atlas using a prior.*** *Prior: horizontal displacement.* ***Template*** *at time* $t = 0$ *(purple curve) and other* ***geometrical descriptor*** *(blue cross).* ***Targets*** *in black. The black vector is the* ***fixed parameter*** *that defines the direction of the translation. Left: Before optimisation. Right: After optimisation.*

5.1. Building an atlas using a prior. In this section, we study the set of shapes presented in Figure 5.1 : each shape has a hump, three of them have the hump rather on their left part, and three other have it on their right part, all at variable locations. Intuitively, there are two possible descriptions of the variability of this collection of shapes. One possibility is to consider that shapes derive from a "template" shape with one central hump by random translations of the hump in either direction. Another possibility is to consider that shapes derive from "template" shape by unfolding the hump in one place and fold a hump at another place. These two models of shape variability would explain the observed samples equally well, and the problem is undecidable without assuming priors on the solution. Determining the template reflecting the structure in the data set, and allowing to study its variability is a well-known problem in computer vision and shape analysis, as described in [19, 23] for instance. If no point-correspondence is assumed, the current implementation of most statistical shape analysis techniques will give one or the other solution in an unpredictable way depending mostly on implementation choices, initialisation,

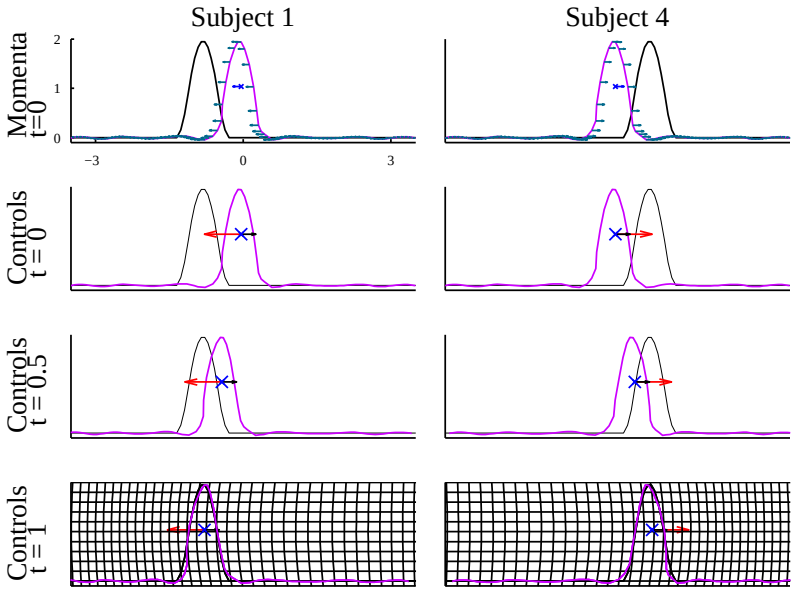

FIG. 5.3. ***Building an atlas using a prior.*** *Prior: horizontal displacement.* ***Template*** *at time t (in purple),* ***target shapes*** *are in black. Other* ***geometrical descriptor*** *at time t (blue cross). First line:* ***momenta*** *at t = 0 (in marine blue are momenta attached to the template's points and in blue the one attached to center of the translation with fixed direction). Three lower lines : the black vector is the fixed parameter that defines the direction of the translation,* ***controls*** *at t are represented by the length of the red arrow.*

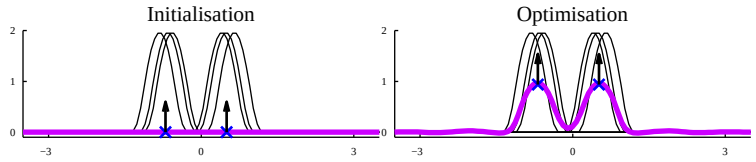

FIG. 5.4. ***Building an atlas using a prior.*** *Prior: vertical displacements.* ***Template*** *at time t = 0 (purple curve) and other* ***geometrical descriptor*** *(blue crosses). The black vector is the fixed parameter that defines the direction of the translation.* ***Targets*** *in black. Left: Before optimisation. Right: After optimisation.*

regularisation techniques, etc..

We show here that our approach based on modular deformations allows the user to decide beforehand which solution he wants to favor. The addition a such a prior on the sought solution is possible by the design of relevant modules.

To obtain a description of shape variability based on horizontal translations of the hump, we use a user-defined deformation module, which generates a vector field that is always a horizontal translation at a fixed large scale σ (see Section 2.3.5, here we use $\sigma = 3$ and a constant direction of translation). We initialise the template curve with a shape with no hump, include it into a silent module and combine it with the translation module. We minimize the Functional (4.3), where μ is the varifold distance between the curves, which measures how well two curves are superimposed without the need to have point correspondence and consistent orientation [8]. The

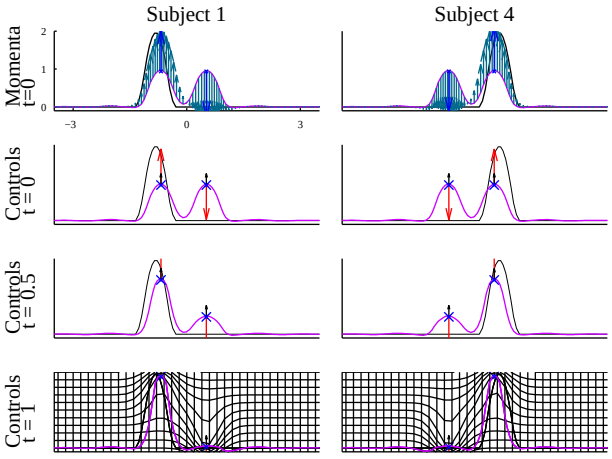

FIG. 5.5. ***Building an atlas using a prior.*** *Prior: vertical displacements.* ***Template*** *at time t (in purple),* ***target shapes*** *are in black. Other* ***geometrical descriptor*** *at time t (blue crosses). First line :* ***momenta*** *at $t = 0$ (in marine blue are momenta attached to the template's points and in blue the ones attached to centers of translations with fixed direction). Three lower lines : black vectors are the fixed parameter that defines the directions of translations,* ***controls*** *at t are represented by the lengths of red arrows.*

minimisation results is the optimal geometrical descriptors of the compound module, here the position of the vertices of the template curve and the base point of the translation, and one initial momentum per target shapes characterising the deformation of the template curve to the corresponding target shape. In Figure 5.2 (left part) we present the initialisation of the template and the geometrical descriptor, momenta η_0^k are initialized at zero. The fixed horizontal vector $u = (1, 0)$ is plotted in black, it is not optimised during the gradient descent and does not evolve during the integration of the trajectory of vector field. In Figure 5.2 (right part) we display the optimized shared parameters of optimal trajectories: the optimized template and position for the geometrical descriptor at time $t = 0$. The diffeomorphic deformation from the template to one target is parametrized by values at time $t = 0$ of the template, the geometrical descriptor and the momentum (dual variable of these quantities). We present on Figure 5.3, for two subjects, this parametrization of trajectories (first row), and also the transport of the template to targets with the geometrical descriptor and the control (last three rows). As the control is scalar, it is plotted as the length of the vector of the translation: the unit horizontal vector (fixed parameter) is plotted in black and in red is plotted this vector multiplied by the scalar control.

By contrast, we may decide to describe the variability in this same data set by using folding/unfolding pattern. The corresponding prior in the deformation model is encoded by one deformation module generating vector fields that are always a sum of two vertical translations at a fixed small scale σ (see Section 2.3.5, here we use $\sigma = 0.4$ and constant directions of translations). In Figure 5.4 can be seen the template and other geometrical descriptors before and after optimisation, as well as fixed vertical vectors . Parametrisation of minimizing trajectories, and trajectories of the template, geometrical descriptors and controls are plotted in Figure 5.5 for two subjects. As previously, controls can be seen in the lengths of vertical vectors, black vectors being fixed vertical unit vectors.

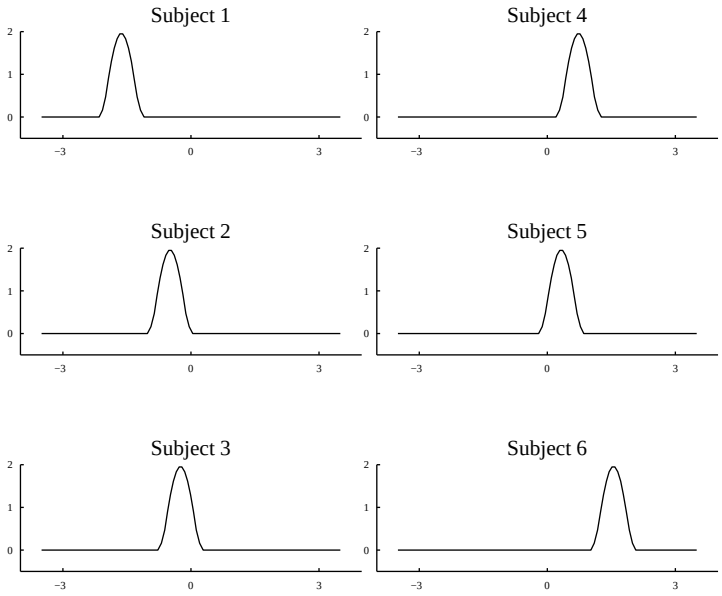

FIG. 5.6. ***Building an atlas using a prior.*** *Different data set.*

In these two experiments the data term $\sum_{k=1}^{P} \frac{1}{\sigma^2} \mu\Big(\phi^{o^k,\eta^k} \cdot f_{temp}, f^k_{target}\Big)$ (see Equation (4.3)) decreases significantly during the optimisation process (divided respectively by 120 and 55): both priors allow to study the variability amongst the population, even though the first one seems to explain a larger part of the variance. These two experiments show that one may obtain different templates and deformations from the same data set by using different prior on shape variability. Our method allows the user to easily encode such prior in an intuitive and controlled manner by the design of suitable modules.

We present now the results obtained for another data set presented in Figure (5.6): here humps are at various locations. In Figure 5.7 is presented the initialisation and optimisation of the template and geometrical descriptor if we incorporate the first kind of prior in the deformation model, *ie* we use the deformation module generating always a horizontal translation. The template has one big centred hump and the data term is divided by 46 during the gradient descent. In Figure 5.8 are presented the same results but if we use the other kind of prior, *ie* the deformation module generating always a sum of two vertical translations. The optimised template seems less appropriate. This is confirmed by the evolution of the data term during the gradient descent : it is only divided by 3. Here by comparing the decrease of the data term we understand that the first prior is more adapted to this data set, and then that the variability among the population is better described by horizontal displacements.

5.2. Performing jointly rigid and non linear registration to study variability among the population. In this section we present the construction of an atlas from a collection of 40 rabbit profiles (see Figure 5.9) with variable ear size and variable positioning of the shape in the ambient space. These shapes are encoded as varifolds [8] so that no point-correspondence is assumed. We propose here to combine rigid registration and local deformation in a single optimisation, whereas in the

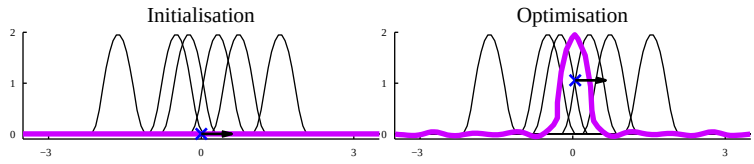

FIG. 5.7. ***Building an atlas using a prior.*** *Prior: horizontal displacement.* ***Template*** *at time $t = 0$ (purple curve) and other* ***geometrical descriptor*** *(blue cross). Black vector is the fixed parameter that defines the direction of translation.* ***Targets*** *in black. Left: Before optimisation. Right: After optimisation.*

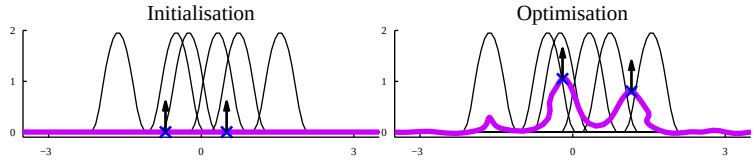

FIG. 5.8. ***Building an atlas using a prior.*** *Prior: vertical displacement.* ***Template*** *at time $t = 0$ (purple curve) and other* ***geometrical descriptors*** *(blue crosses). Black vectors are the fixed parameters that define the direction of translations.* ***Targets*** *in black. Left: Before optimisation. Right: After optimisation.*

vase majority of cases shapes are rigidly registered as a pre-processing step before statistical analysis.

To this end, we propose to combine three deformation modules: the first one generates vector fields that are a rotation at scale 5000 (see Section 2.3.1), the second one generates vector fields that are a translation at scale 5000 (see Section 2.3.3). The last deformation module generates a sum of two translations whose directions are updated by adjoint action at scale 600 (See Section 2.3.5). The first two modules encode a rigid body transformation (at the scale of the shapes), and the third one encodes local non-linear deformation patterns. In Figure 5.10 are plotted template and other geometrical descriptors at $t = 0$, before and after optimisation. Note that the geometrical descriptor of the deformation module generating a sum of two translations with directions updated by adjoint action is composed of two points and two vectors : the initial directions of these two vectors are shared by all subjects. The optimised initial directions are along ears : the variability amongst the population at this scale is in the direction of ears, which was expected. Besides, as a consequence, values of associated controls measure the growth or shrinking of ears and then this particular feature can be studied independently from the rigid registration. In Figure 5.11, on the first columns, are presented parametrisation of minimizing trajectories (template,other geometrical descriptors and momenta) for three subjects. Note that momenta associated to the sum of two translations with updated directions have two components: one associated to base points of translations, and one to vectors of translations. All these vectors are represented attached to the base point of the translation, but in different colors. In the three columns on the right hand side of this figure are represented the trajectory of the template, geometrical descriptors and controls for these three subjects. Controls associated to translations with updated direction are scalar so they are represented as lengths of vectors. The geometrical descriptor of the rotation is a blue circle, circled by a black one. The control associated to the rotation (scalar) is represented by the portion of this black circle which is coloured in red (the

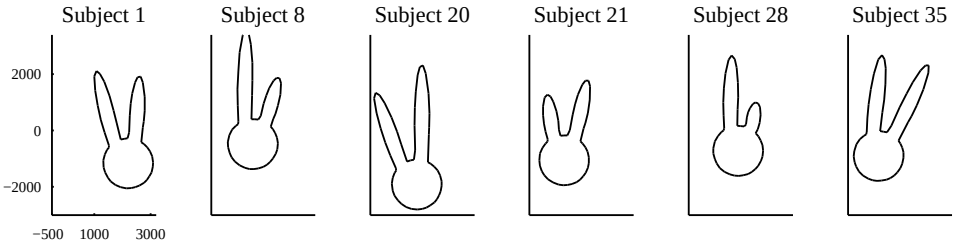

FIG. 5.9. ***Performing jointly rigid and non linear registration.*** *Target shapes*

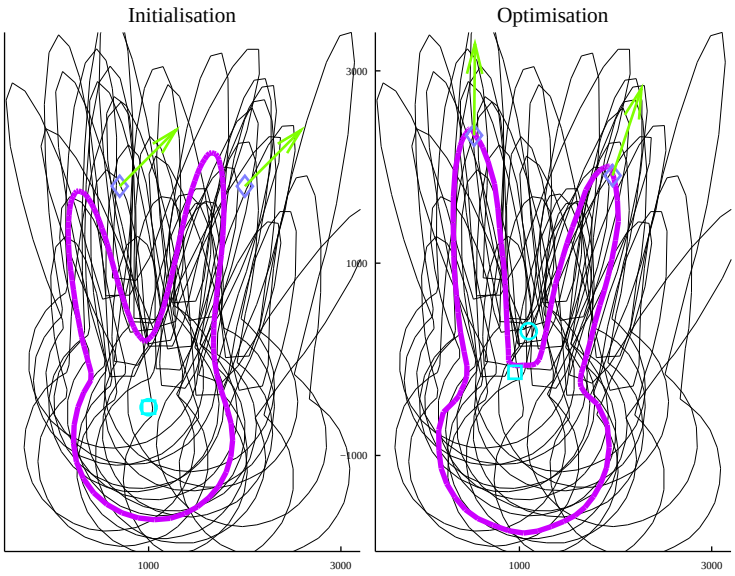

FIG. 5.10. ***Performing jointly rigid and non linear registration. Template*** *at time* $t = 0$ *(purple shape) and other* ***geometrical descriptors****: rotation's one is the blue circle, the translation's ones is the blue square and translations with transported vectors' ones have their dot-component represented by blue diamonds and their vector components represented by green arrows.* ***Targets*** *in black. Left: Before optimisation. Right: After optimisation.*

control is positive if the colouring is anti-clockwise and negative otherwise).

We compare our result with the one obtained by rigidly registering data beforehand and then applying the framework developed in [10]: in this framework the vector field is built as a sum of local translations carried by control points. We use here 50 control points, and we set the scale of translations equal to 500. In Figure 5.12 we present the targets after the rigid registration, as well as the initial position of the template and control points, before and after optimisation. In Figure 5.13 are presented the parametrisation of optimal trajectories for several subjects (in this framework momenta are equal to controls) and the evolution of the template, control points and controls. One can see that non linear deformations are now necessary not only in the area of the ears, but also in the inferior part of the head. This fact is the direct consequence of the rigid-body registration, which does not align the inferior parts of the heads. It shows the sub-optimality of the greedy approach that consists in optimising the rigid part and the non-linear part in two consecutive step. Very few

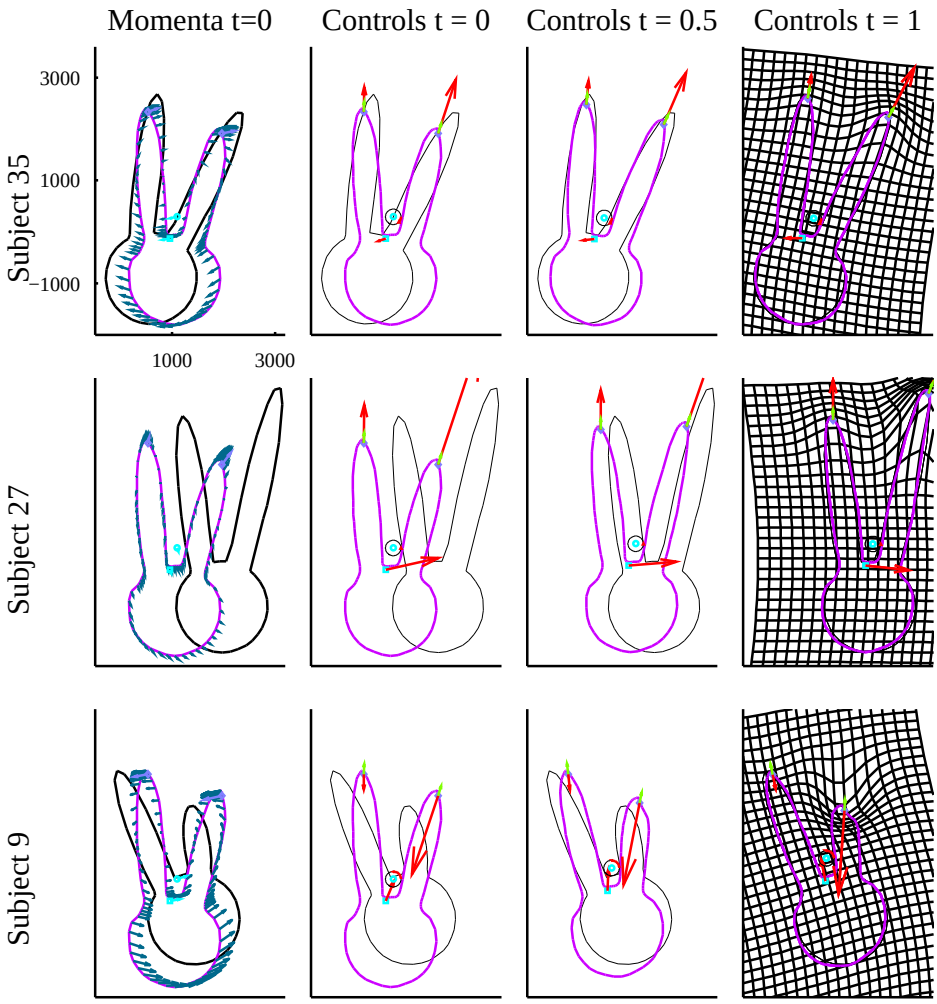

FIG. 5.11. ***Performing jointly rigid and non linear registration. Template*** *at time t (in purple),* ***target shapes*** *are in black. Other* ***geometrical descriptors*** *at time t: rotation's one at the blue circle, the translation's one at the blue square, translations with updated directions ones have their dot-component at blue diamonds and their vector-component represented by green arrows. Left column:* ***Momenta*** *at $t = 0$ in blue, attached to their geometrical descriptor (translations with update directions' ones have their vector-component represented in green, attached to the center of the translation). Three right-hand columns:* ***Controls*** *at time t : the red arrow attached to the blue square is the translation's one, the rotation's one is represented by the portion of the black circle coloured in red, and the one of translations with transported vectors are represented by the lengths and orientations of the red arrows attached to the blue diamond.*

mathematical frameworks may deal with this issue, whereas it is well-known in the statistical shape analysis community. By contrast, the method that we present here allows the joint optimisation of linear and non-linear deformation patterns by the use of combined deformation modules at various scales. The resulting explanation of the variability seen in the data is much more satisfying by displaying a fixed head and ears of variable sizes.

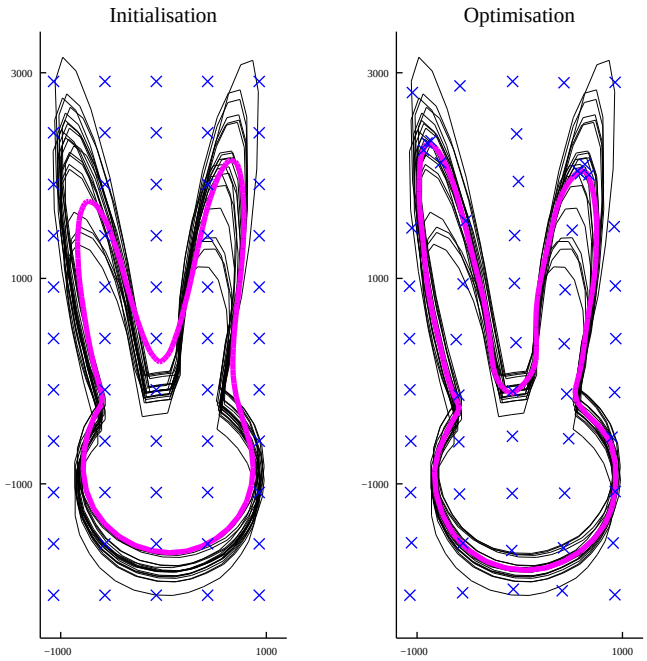

FIG. 5.12. ***Performing rigid before non linear registration. Template*** *at time* $t = 0$ *(purple) and* ***control points*** *(blue crosses).* ***Targets*** *in black. Left: Before optimisation. Right: After optimisation.*

5.3. An example of atlas with a weak prior. In this last example we compute an atlas of the five shapes of skulls presented in 5.14. Similarly to previous examples, shapes are encoded as varifolds [8] so that no point-correspondence is assumed. Here, we do not have clear prior to include in the model. Therefore, we combine 7 deformation modules: one generating a translation at large scale ($\sigma = 200$, see Section 2.3.3), one generating a rotation at large scale ($\sigma = 200$, see Section 2.3.1), one generating a sum of two rotations at middle scale ($\sigma = 100$, see Section 2.3.1), one generating a sum of two scalings at middle scale ($\sigma = 100$, see Section 2.3.1), one generating a sum of four rotations at small scale ($\sigma = 50$, see Section 2.3.1), one generating a sum of four scalings at small scale ($\sigma = 50$, see Section 2.3.1) and one generating a sum of 16 translations of directions updated by adjoint action at small scale ($\sigma = 50$, see Section 2.3.5). The first two deformation modules (those at large scale) enable to perform rigid registration simultaneously with finer deformations. At the smallest scales, translations with transported direction have shared initial positions and directions for all subjects, and their directions are transported by the flow. Using this deformation module instead of one generating translation as in Section 2.3.3, allows to interpret initial vector of these translations as directions of greatest variability among the population (at the small scale). The initial distribution of the geometrical descriptors on a regular lattice and the template before optimisation can be seen in Fig. 5.15. Geometrical descriptors are updated during optimisation to regions of interest. Parametrization of minimizing trajectories (template, initial other geometrical descriptors and momenta) and trajectories of template, geometrical descriptors and controls can be seen in Figure 5.16 for three skulls. Controls associated

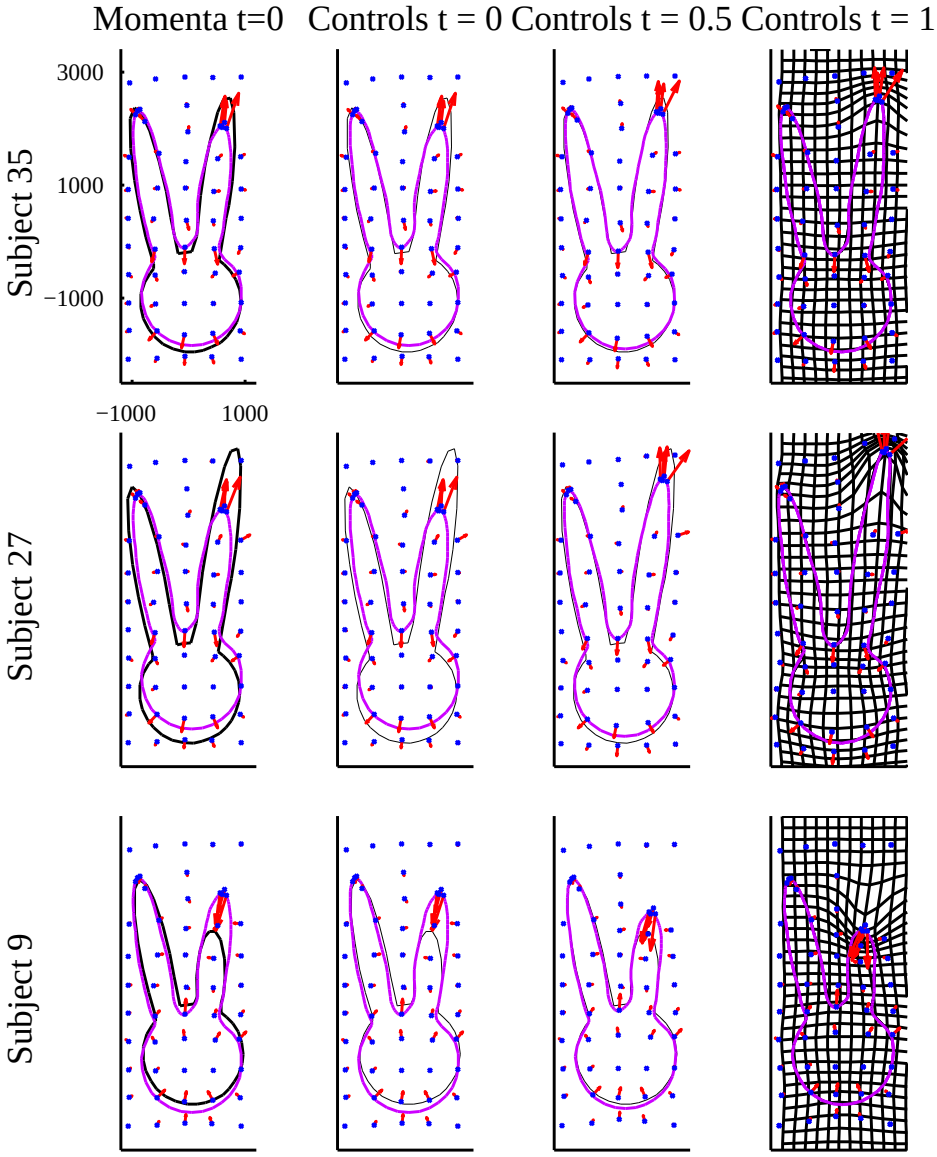

FIG. 5.13. ***Performing rigid before non linear registration.*** *Framework [10].* ***Template*** *at time t (in purple),* ***target shapes*** *are in black.* ***Control points*** *at time t (blue crosses). Left column:* ***Momenta*** *at $t = 0$ (blue arrows). Three right-hand columns:* ***Controls*** *at time t in red.*

to translations with updated directions are scalars and are represented by lengths of vectors attached to blue diamonds. Rotations have their geometrical descriptors represented by blue circles, and their controls (scalars) by the portion of these blue circles that are coloured in red (the control is positive if the colouring is anti-clockwise and negative otherwise). Geometrical descriptors of scalings are represented by blue triangles, and for each one a red triangle corresponding to the image of the blue one by this scaling enables to represent the scalar control. From these results one can see, for

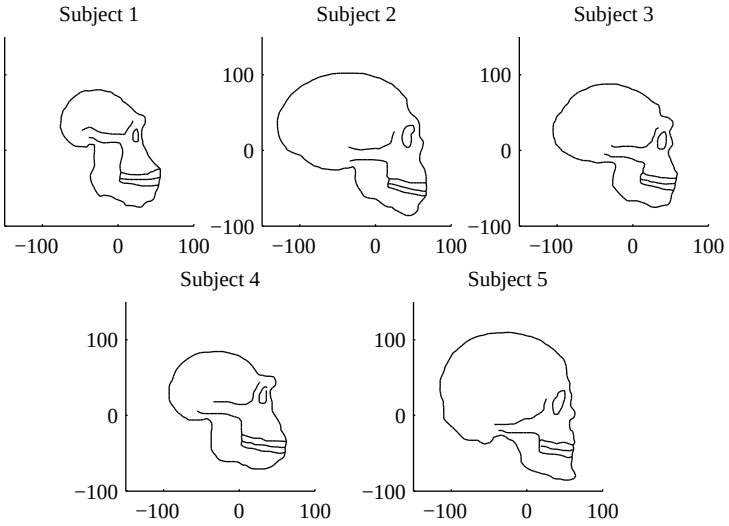

FIG. 5.14. ***Building an atlas with weak priors.*** *Target shapes*

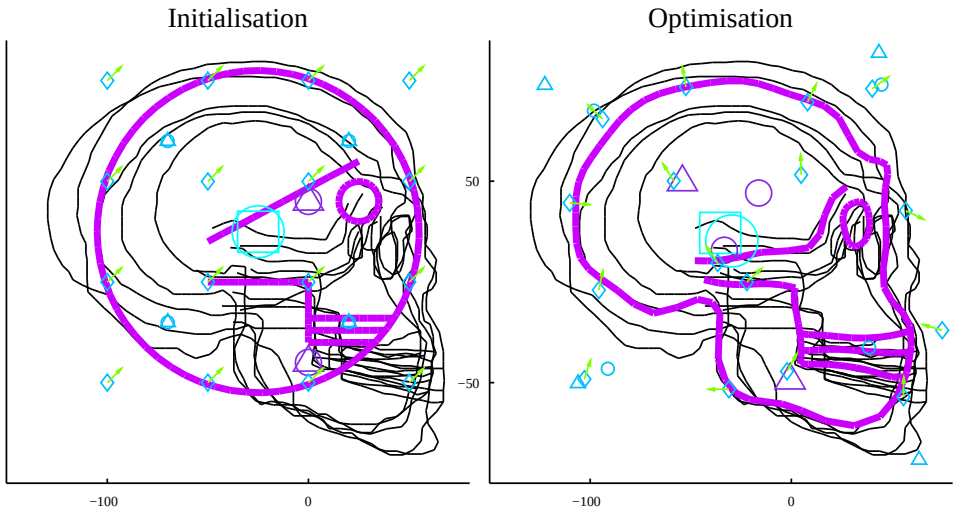

FIG. 5.15. ***Building an atlas without prior. Template*** *at time* $t = 0$ *(purple shape) and* ***geometrical descriptors****:* □ *(translations),* ○ *(rotations),* △ *(scalings) and green vectors attached to* ◊ *(translations with updated directions). The color and size of markers represent the scale (200,100,50). Black shapes are the targets. Left: Before optimisation. Right: After optimisation.*

example, that the size of the cranium is a feature that varies importantly amongst the population as one center of the two local scalings at scale 100 has moved to this area during optimisation, and trajectories of the corresponding control are different for all skulls: negative for skull one (the red triangle is very small so the control is highly negative: his cranium is smaller than the template), close to zero for skull three (the red triangle is almost of the size of the blue one: his cranium is almost of same size as the template) and positive for skull five (the red triangle is bigger than the blue one: his cranium is bigger than the template). Other features of great variabilities can be

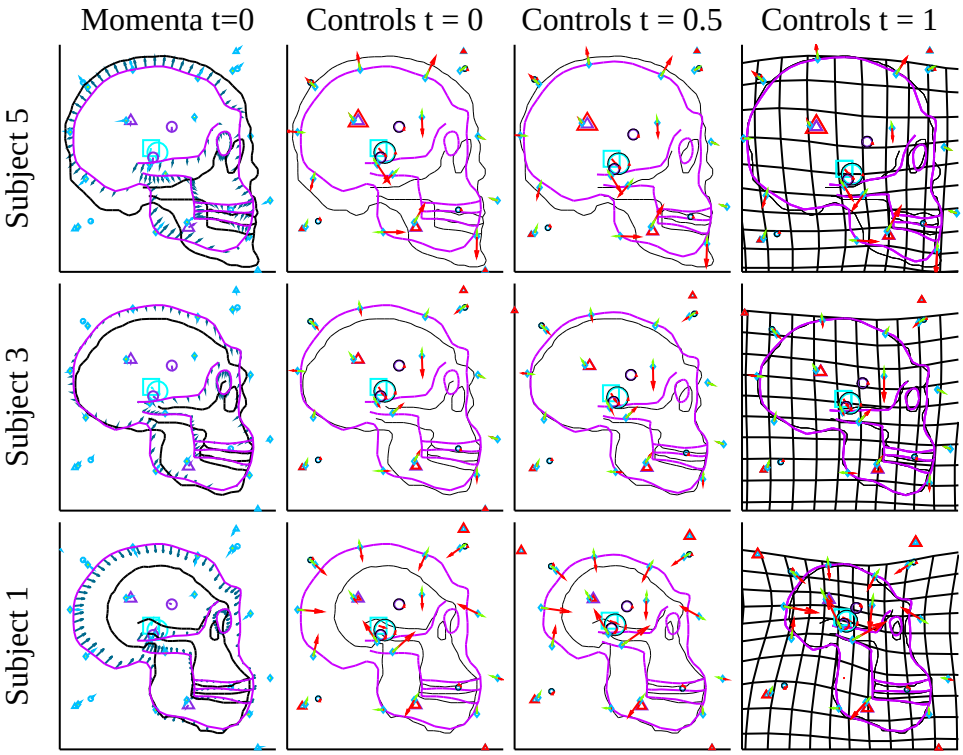

FIG. 5.16. ***Building an atlas with weak prior. Template*** *at time t (in purple),* ***target shapes*** *are in black.* ***Geometrical descriptors*** *at time t:* □ *(translations),* ○ *(rotations),* △ *(scalings) and green vectors attached to* ◊ *(translations with updated directions). The color and size of markers represent the scale (200,100,50). Left column: Momenta at* $t = 0$*. Three right-hand columns:* ***Controls*** *at time t in red (vector for the translation, portion of the blue circle coloured in red for rotations, ratio between the blue and red triangles for scalings and length of the red vectors for translations with updated direction).*

deduced from this results such as, for instance, the shape of the back of the head by inspecting the initial value of vector of the corresponding translation with updated direction.

This experiment shows in a more realistic case how the variability seen in a shape data set may be decomposed in a series of local non-linear and multiscale deformation patterns.

6. Conclusion. In this paper we defined a coherent mathematical framework to build locally constrained diffeomorphic deformations thanks to the introduction of a new concept: deformation modules. The modules constrain the driving velocity field of the deformation to belong to particular subspaces of admissible vector fields. The flow of such sub-Riemannian trajectories of vector fields forms a modular large deformation. These deformations act both on the shapes (namely points, meshes or images) that are embedded into the ambient space, and on the geometrical parameters of the deformation modules themselves, so that they are all transported during the integration of the flow equations. We consider the shapes to be deformed as geometrical parameter of a particular module, called "silent" module, so that shapes and

active geometrical descriptors alike belong to a common "shape space". We show that our construction enables the definition of a sub-Riemmanian distance on this shape space, and that normal geodesics are parameterized in finite dimension thanks to the initial momentum.

We showed how these deformations may be used to summarise the variability observed in a series of example shapes. The proposed method estimates a "template shape", as a Fréchet mean, and decomposes the observed variability into a dictionary of deformation modules encoding specific local deformation patterns. The relative position of each example shape with respect to the template shape on the sub-Riemannian manifold is encoded into the initial momentum. These descriptors of shape variability may be then used for statistical purposes, such as classification, clustering, or regression against covariates.

It is worth noting that the dimension of the initial momentum is larger than the number of control parameters in the modules, since it is the sum of the dimension of geometrical descriptors of the active modules and the number of points in the shapes to be deformed, namely the geometrical descriptor of the silent module. Therefore, the variability is usually encoded in much higher dimension than the number of degrees of freedom of the deformations. There is an interplay between the momentum attached to the active modules and the one attached to the shapes, which enriches the possible dynamics of control parameters in the active modules. In particular, the initial value of the control parameters does not uniquely determine the subsequent deformation, neither do initial momentum of active deformation modules alone. Nonetheless, it is not clear what is the optimal number of parameters to describe the variability of a given data set. One may want to automatically learn this number using sparse penalties in the spirit of [10]. It is also an open question to know whether the same deformations may be obtained in a purely Riemannian framework, where the dimension of the initial momentum exactly matches the number of control parameters, that is the number of degrees of freedom of the deformations.

The presented experiments show different situations where this modular analysis of shape variability may be useful. It allows the user to easily include prior knowledge in the analysis of shape variability, by encoding biological constraints for instance. By combining global rigid-body transformations with a non-linear combination of locally affine transforms, our method allows the joint optimisation of the pose and shape parameters, thus providing a solution to a long-lasting problem in statistical shape analysis.

There are situations where multiple local deformation patterns may explain the shape variability equally well. In such situations, the optimal solution depends on the weights in the combination of costs associated to each module. One may learn such weights in the future by using a Bayesian framework along the lines of [2, 12] for instance.

REFERENCES

[1] A. Agrachev, D. Barilari, and U. Boscain. *Introduction to Riemannian and Sub-Riemannian geometry*. 2014.

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

[3] S. Allassonnière, E. Kuhn, A. Trouvé, et al. Construction of bayesian deformable models via a stochastic approximation algorithm: a convergence study. *Bernoulli*, 16(3):641–678, 2010.

[4] S. Arguillere. *Géométrie sous-riemannienne en dimension infinie et applications à l'analyse mathématique des formes*. PhD thesis, Paris 6, 2014.

[5] S. Arguillere, E. Trélat, A. Trouvé, and L. Younes. Shape deformation analysis from the optimal control viewpoint. *Journal de mathématiques pures et appliquées*, 2015.

[6] V. Arsigny, O. Commowick, N. Ayache, and X. Pennec. A fast and log-euclidean polyaffine framework for locally linear registration. *Journal of Mathematical Imaging and Vision*, 33(2):222–238, 2009.

[7] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008.

[8] N. Charon and A. Trouvé. The varifold representation of non-oriented shapes for diffeomorphic registration. *arXiv preprint arXiv:1304.6108*, 2013.

[9] S. Durrleman, M. Prastawa, N. Charon, J. R. Korenberg, S. Joshi, G. Gerig, and A. Trouvé. Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *NeuroImage*, 101:35–49, 2014.

[10] S. Durrleman, M. Prastawa, G. Gerig, and S. Joshi. Optimal data-driven sparse parameterization of diffeomorphisms for population analysis. In *Information Processing in Medical Imaging*, pages 123–134. Springer, 2011.

[11] J. Glaunes. Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique. *These de sciences, Université Paris*, 13, 2005.

[12] P. Gori, O. Colliot, Y. Worbe, L. Marrakchi-Kacem, S. Lecomte, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman. Bayesian Atlas Estimation for the Variability Analysis of Shape Complexes. In *MICCAI 2013 : Medical Image Computing and Computer Assisted Intervention*, pages 267–274, Nagoya, Japan, Sept. 2013.

[13] U. Grenander. *Elements of pattern theory*. JHU Press, 1996.

[14] U. Grenander, Y.-s. Chow, and D. M. Keenan. *Hands: A pattern theoretic study of biological shapes*, volume 2. Springer Science & Business Media, 2012.

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

[16] B. Gris, S. Durrleman, and A. Trouvé. A sub-riemannian modular approach for diffeomorphic deformations. In *Geometric Science of Information*, pages 39–47. Springer, 2015.

[17] H. Jacobs. Symmetries in lddmm with higher order momentum distributions. *arXiv preprint arXiv:1306.3309*, 2013.

[18] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:S151–S160, 2004.

[19] A. C. Kotcheff and C. J. Taylor. Automatic construction of eigenshape models by genetic algorithm. In *Information Processing in Medical Imaging*, pages 1–14. Springer, 1997.

[20] S. Kurcyusz. On the existence and nonexistence of lagrange multipliers in banach spaces. *Journal of Optimization Theory and Applications*, 20(1):81–110, 1976.

[21] P. Lorenzen, B. C. Davis, and S. Joshi. Unbiased atlas formation via large deformations metric mapping. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pages 411–418. Springer, 2005.

[22] J. Ma, M. I. Miller, A. Trouvé, and L. Younes. Bayesian template estimation in computational anatomy. *NeuroImage*, 42(1):252–261, 2008.

[23] G. McNeill and S. Vijayakumar. Linear and nonlinear generative probabilistic class models for shape contours. In *Proceedings of the 24th international conference on Machine learning*, pages 617–624. ACM, 2007.

[24] M. I. Miller, L. Younes, and A. Trouvé. Diffeomorphometry and geodesic positioning systems for human anatomy. *Technology*, 2(01):36–43, 2014.

[25] N. Portman. The modelling of biological growth: A pattern theoretic approach. 2010.

[26] L. Risser, F. Vialard, R. Wolz, M. Murgasova, D. D. Holm, and D. Rueckert. Simultaneous multi-scale registration using large deformation diffeomorphic metric mapping. *Medical Imaging, IEEE Transactions on*, 30(10):1746–1759, 2011.

[27] C. Seiler, X. Pennec, and M. Reyes. Capturing the multiscale anatomical shape variability with polyaffine transformation trees. *Medical image analysis*, 16(7):1371–1384, 2012.

[28] S. Sommer, F. Lauze, M. Nielsen, and X. Pennec. Sparse multi-scale diffeomorphic registration: the kernel bundle framework. *Journal of mathematical imaging and vision*, 46(3):292–308, 2013.

[29] S. Sommer, M. Nielsen, S. Darkner, and X. Pennec. Higher-order momentum distributions and locally affine lddmm registration. *SIAM Journal on Imaging Sciences*, 6(1):341–367, 2013.

[30] S. Sommer, M. Nielsen, F. Lauze, and X. Pennec. A multi-scale kernel bundle for lddmm:

towards sparse deformation description across space and scales. In *Information Processing in Medical Imaging*, pages 624–635. Springer, 2011.

[31] A. Srivastava, S. Saini, Z. Ding, and U. Grenander. Maximum-likelihood estimation of biological growth variables. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 107–118. Springer, 2005.

[32] D. W. Thompson et al. On growth and form. *On growth and form.*, 1942.

[33] M. Vaillant, M. I. Miller, L. Younes, and A. Trouvé. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23:S161–S169, 2004.

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

[35] L. Younes. Constrained diffeomorphic shape evolution. *Foundations of Computational Mathematics*, 12(3):295–325, 2012.

[36] W. Zhang, J. A. Noble, and J. M. Brady. Adaptive non-rigid registration of real time 3d ultrasound to cardiovascular mr images. In *Information Processing in Medical Imaging*, pages 50–61. Springer, 2007.