The Fshape Framework for the Variability Analysis of Functional Shapes

B. Charlier¹ · N. Charon² · A. Trouvé³

Received: 2 December 2014 / Revised: 18 April 2015 / Accepted: 20 August 2015 / Published online: 19 November 2015 © SFoCM 2015

Abstract This article introduces a full mathematical and numerical framework for treating functional shapes (or fshapes) following the landmarks of shape spaces and shape analysis. Functional shapes can be described as signal functions supported on varying geometrical supports. Analyzing variability of fshapes’ ensembles requires the modeling and quantification of joint variations in geometry and signal, which have been treated separately in previous approaches. Instead, building on the ideas of shape spaces for purely geometrical objects, we propose the extended concept of fshape bundles and define Riemannian metrics for fshape metamorphoses to model geometric-functional transformations within these bundles. We also generalize previous works on data attachment terms based on the notion of varifolds and demonstrate the utility of these distances. Based on these, we propose variational formulations of the atlas estimation problem on populations of fshapes and prove existence of solutions for the different models. The second part of the article examines thoroughly the numerical implementation of the tangential simplified metamorphosis model by detailing discrete expressions for the metrics and gradients and proposing an optimization scheme for the atlas estimation problem. We present a few results of the methodology on a synthetic dataset as well as on a population of retinal membranes with thickness maps.

Keywords Shape analysis · Signals on manifolds · Large deformation models · Metamorphoses · Varifolds · Atlas estimation algorithms

Communicated by Peter J. Olver.

N. Charon charon@cis.jhu.edu

¹ IMAG, UMR 5149, Université de Montpellier, Montpellier, France
² CIS, Johns Hopkins University, Baltimore, MD, USA
³ CMLA, UMR 8536, École normale supérieure de Cachan, Cachan, France
Mathematics Subject Classification 49M25 · 49Q20 · 58B32 · 58E50 · 68U05 · 68U10

1 Introduction

Shape spaces have emerged as a natural mathematical setting to think about shapes as a structured space, usually a differential manifold or even more a Riemannian manifold. In that setting, group actions of diffeomorphisms [6,13,25–27,34,37,38] that should be rooted to the seminal work of Arnold [4] are powerful vehicles to build a full theoretical as well as computational framework for a comprehensive quantitative analysis of shape variability recently coined as diffeomorphometry and successfully applied in computational anatomy [28].

The main purpose of the paper is to develop a framework embedding the situation of geometrical shapes carrying functional information that we call here fshapes for functional shapes. The idea of functional shape is quite natural since in many scientific settings, a geometrical shape \( X \) (basically a submanifold) has associated with it a scalar field \( f : X \to \mathbb{R} \) attached to every geometrical point of the support \( X \). What is much more unusual however is to consider a fshape, i.e., the pair \( (X, f) \) as a single object that should live in some well-defined and well-structured ensemble \( \mathcal{F} \) of fshapes from which further processing can be derived. We believe this approach of putting the effort on a well-grounded definition of shape spaces that play a similar role to functional spaces in modern analysis should be pushed forward to the setting of fshapes. This global point of view leading to the notion of geometry of shape spaces is now well established with fruitful development in many regards [24] for purely geometrical shapes, but as far as we know the extension of such approach to functional shapes has only started very recently in [10].

A core issue in working with fshapes is their mixed geometrical and functional nature so that smooth infinitesimal transformations of a given fshape \( (X, f) \) should combine a geometrical and functional component simultaneously. The geometrical part should transport the supporting manifold \( X \), and the functional one should modify the functional signal \( f \). This notion of combined geometrical and functional infinitesimal transformation has been introduced previously in [33] in the setting of images on a fixed support \( X \) (where \( X \) is the unit square or a flat torus). In that case the infinitesimal signal evolution \( \delta f \) is the combination of two factors: on the one hand, the variation of the signal due to the geometric transport of pixel values, and on the other hand, a purely additive perturbation of the signal. This has been further studied in [35] and conceptualized as the metamorphosis framework in [36]. The situation we are looking at here is more general since now the support is a submanifold that can freely evolve during the metamorphosis (as in the example of Fig. 1). A first global outcome is that the associated space of fshapes is not only an orbit \( G \cdot X_0 \) under the action of a subgroup of smooth diffeomorphisms \( G \) of the ambient space on a geometrical template \( X_0 \): The orbit \( G \cdot X_0 \) is now the base space of a vector bundle, and above each manifold \( X \) in the orbit we consider the full vector space \( L^2(X) \) of square integrable signals so that we end up with a vector bundle \( \mathcal{F} \) over the orbit (see Fig. 2). The precise construction of the fshape bundle \( \mathcal{F} \) is described in Sect. 2 as well as...
the associated metamorphosis-based Riemannian metric for which we can prove two key results: the existence of geodesic between any two fshapes (Theorem 2) and the existence of Karcher means for a population of such fshapes in $\mathcal{F}$ (Theorem 3).

There is no hope to build a fshape bundle $\mathcal{F}$ that can contain any possible pair $(X, f)$ in particular since any two fshapes within $\mathcal{F}$ will have diffeomorphic supports. Every fshape in $\mathcal{F}$ should be understood as an ideal model for truly noisy observed fshapes that generically do not belong to $\mathcal{F}$. As a consequence, a second core issue is to build proper smooth data attachment terms or dissimilarity measures that can be defined between any arbitrary pair of fshapes with possibly non-diffeomorphic supports. The powerful setting of mathematical currents equipped with dual norms for purely geometrical shapes [18] has been successfully extended to the situation of fshapes in [11]. However, we develop here (Sect. 4) a different approach encoding non-oriented tangential space information that is based on the concept of varifold along the lines of [12] and extended here to the new situation of functional varifolds. We
believe that this new type of dissimilarity measure, which is able to compare fshapes on smooth manifolds as well as on polyhedral meshes, is well suited in combination with metamorphoses distances to define a theoretical and computational framework for fshapes. In particular, we establish that useful functional varifolds metrics can be induced by smooth embeddings of functional varifolds into reproducible kernel Hilbert spaces (RKHSs) associated with computable kernels. We then establish that the resulting metric has several key regularity properties with respect to smooth geometrical and functional variations of fshapes (Theorem 5). Such regularity results open the way to various smooth and computationally tractable variational problems on fshapes involving observed fshapes.

We concentrate on the central problem of template (or atlas in the context of computational anatomy) estimation which is the extension of the Karcher mean problem to the much broader situation where the observed fshapes do not belong to any a priori defined fshape bundle $\mathcal{F}$. A much more studied situation is the case of purely image data that have been pushed forward on practical ground in the framework of actions of diffeomorphisms in [19]. Here the goal is to obtain, via a coherent fshape framework, a mean template $(X, f)$ encoding simultaneously the mean geometrical and functional information from a dataset $(X^i, f^i)_{1 \leq i \leq N}$ of fshapes as the minimizer of a well-posed variational problem. This variational problem is basically the sum of the square of the metamorphosis distances between the template $(X, f)$ and the approximating models $(\tilde{X}^i, \tilde{f}^i)_i$ of the observations $(X^i, f^i)_i$ within the fshape bundle $\mathcal{F}$ [associated with $(X, f)$] with the addition of the square of the fvarifold discrepancy measure for the difference between the models $(\tilde{X}^i, \tilde{f}^i)_i$ and the true observation $(X^i, f^i)_i$. We prove that introducing a hypertemplate in the spirit of [22, 23] and restricting the template to belong to the fshape bundle $\mathcal{F}_0$ of a hypertemplate $(X_0, f_0)$ are sufficient to establish a well-posed variational problem (Sect. 4). The proof of existence of an optimal template $(X, f)$ as the minimizer within the fshape bundle $\mathcal{F}_0$ is established on fairly general assumptions that cover most of the practical situations (Theorem 6). Note that even restricted to the situation of purely geometrical shapes, as far as we know, there are no available results in the literature providing a rigorous proof of existence of an average shape from a population of noisy geometrical shapes. Available rigorous results have usually focused on the situation where the dataset is sampled from a Riemannian manifold (usually finite dimensional). When the functional part is also involved, the situation is more complex due to the interactions between function and geometry inside the Riemannian metamorphosis distance but also the fvarifold metric defining the dissimilarity term. We establish similar existence results in a slightly simpler situation, called the tangential model for the metamorphosis metric, where the infinitesimal signal variations along a path are computed with respect to a frozen $L^2(X)$ metric on the geometrical support $X$ of the initial fshape.

Since the paper aims at presenting both a theoretical and a computational frameworks, we extensively explore the material needed to bridge the gap between the theoretical framework and its derivation into a computational and algorithmic one in part II of this paper. We believe that this part will give also more practical insights into what goals can be achieved with the fshape framework. Moreover, we provide the details necessary to conform to the minimal standard of reproducible research. Just as the absence of a well-posed theoretical framework may often produce inconsistent
algorithms, we believe that the lack of precise descriptions of the underlying numerical schemes and hidden “tricks” may also hinder the development of incrementally better solutions and fair comparison by other researchers. We start in Sect. 6 with the derivation of a discrete framework based on polyhedral representation of fshapes leading to particle-based approximations of the fvarifold dissimilarity metric and of the Hamiltonian underlying the optimal control formulation of the geodesic trajectories. The actual implemented framework is based on the tangential model for the metamorphosis metric that leads to a slightly more straightforward variational problem. The optimization scheme is carefully described in Sect. 7, and numerical illustrations are provided in Sect. 8 on real and synthetic datasets. A special attention is given in Sect. 9 on various numerical issues and pitfalls that may affect the computation process. In particular, we discuss the very important problem of open surfaces that introduce free boundaries evolutions during the optimization process. The tuning of free parameters are also discussed. The code called fshapetk is made available in the GIT repository https://github.com/fshapes/fshapesTk to reproduce all the experiments presented here that are based on public data.

As discussed in this introduction, this paper describes in part I a new theoretical shape space framework to work with functional shapes with an emphasis on precise mathematical statements and proofs. The more technical aspects of the novel framework have been collected and presented in “Appendix” to ease the reading of the paper. The reader more interested in applying the framework could read Sect. 2 together with Sect. 3.2, skip Sects. 4 and 5 and jump over directly to Sects. 6–9 dealing with the numerical part of the paper.

Part I Theoretical Foundations

2 Riemannian Metamorphosis Framework for Fshapes

In the classical Grenander’s setting, shape spaces are modeled as sets of shapes homogeneous under the action of a group of space transformations. Metrics between shapes are then induced from right-invariant Riemannian metrics on the group itself. The goal of this section is to propose a similar but extended framework for the case of functional shapes. We show, in the first place, that sets of functional shapes can be structured naturally as vector bundles and we then define a Riemannian setting to model and quantify transformations within those spaces.

2.1 Fshape Bundles

Let’s introduce a finite dimensional vector space $E$ that shall be the embedding space of all shapes. In the large deformations’ model (LDDMM, cf [6]), diffeomorphisms are constructed as flows of time-varying vector fields. The basic ingredient is a reproducing kernel Hilbert space (RKHS) of vector fields on $E$ that is denoted $V$ and which is continuously embedded in $C^0_1(E, E)$ the space of continuously differentiable functions from $E$ to $E$ vanishing at the infinity. Let $G = G_V$ be the associated group of diffeomorphisms obtained by flowing vector fields in $L^2([0, 1], V)$ (see [3,37,38]).
Let $X_0$ be a homogeneous space generated by a finite volume $d$-dimensional rectifiable compact subset $X_0$ (see [29]); i.e., $X_0 = G \cdot X_0 = \{ \phi(X_0) \mid \phi \in G \}$.

Now we consider the space

$$\mathcal{F} \doteq \{(X, f) \mid X \in X_0 \text{ and } f \in L^2(X)\},$$

where $L^2(X)$ is the space of square integrable functions on $X$, i.e., the set of functions $f : X \to \mathbb{R}$ such that

$$\int_X f^2(x) dH^d(x) < +\infty$$

for $H^d$ the $d$-dimensional Hausdorff (or volume) measure. $\mathcal{F}$ can be considered as a vector bundle with fiber $L^2(X_0)$ (here, we will not try to define any explicit differentiable structure on it so that the last statement should be considered as formal). Each element of $\mathcal{F}$ is thus given as a couple of a geometrical shape $X$ and a signal function living on $X$ that we call a functional shape (see [11]) or fshape.

2.2 Metamorphosis Distance on Fshapes Bundles

For a purely geometrical deformation $\phi \in GV$, a natural transport of a fshape $(X, f)$ would be given by $\phi \cdot (X, f) = (\phi(X), f \circ \phi^{-1})$, which corresponds to deforming the support $X$ and transporting the values of the signal $f$ onto the deformed shape, which is the exact generalization of image deformation. Such a model, originally presented in [11], remains insufficient to account for variations of signals within fibers themselves. This justifies the following efforts to propose a model of joint geometric and functional transformations.

Let’s consider $(X, f) \in \mathcal{F}$ and instantaneous velocities $(v, h) \in L^2([0, 1], V \times L^2(X))$. We can define from $(v, h)$ a path $(\phi^v_t, \zeta^h_t)_{t \in [0,1]}$ where $\phi^v_t$ is the usual flow of $v$ starting from the identity and $t \mapsto \zeta^h_t = \int_0^t h_s ds$ is the path in $L^2(X)$ with instantaneous speed given by $h$. From that path, we get a path $t \mapsto (X_t, f_t)$ in $\mathcal{F}$ defined by

$$(X_t, f_t) \doteq (\phi^v_t \cdot X, (f + \zeta^h_t) \circ (\phi^v_t)^{-1}).$$

(2)

where $\phi^v_t \cdot X \doteq \phi^v_t(X)$ is the natural diffeomorphisms transport action on rectifiable subsets of $E$. We denote $(\phi^v_1, \zeta^h_1)$ the end point value of the path. Now we denote for $\gamma_V, \gamma_f > 0$

$$E_X(v, h) \doteq \frac{\gamma_V}{2} \int_0^1 |v_t|^2_V dt + \frac{\gamma_f}{2} \int_0^1 \int_X |h_t|^2(x) \left|d_x \phi^v_t \mid_{T_x X}\right| dH^d(x) dt$$

(3)

where $T_x X$ is the tangent space to $X$ at point $x$ (defined $H^d$-almost everywhere on $X$ if $X$ is rectifiable, cf [15] 3.2.19) and, for any vector space $U$, $|d_x \phi^v_t \mid_{U_x}|$ denotes the Jacobian of $\phi^v_t$ at $x$ restricted to $U$, which equals $|d_x \phi^v_t(u_1) \wedge \cdots \wedge d_x \phi^v_t(u_d)| = \mathcal{F}$.
\((\det((d_x\phi^\nu_i(u_{i}), \ d_x\phi^\nu_i(u_j)))_{i,j})^{1/2}\) if \((u_1, \ldots, u_d)\) is an orthonormal basis of \(U\). Note that we shall sometimes omit the writing of \(dH^d\) for integral on \(X\) when the context is clear.

It is clear that \(E_X(v, h) < \infty\) since by definition \(\int_0^1 |v_t|^2 dt < \infty\) and since the usual controls on \(\phi_t\) along finite energy paths give a uniform control in \((x, t)\) of \(d_x\phi_t\) and the existence of an increasing function \(C : \mathbb{R}_+ \to \mathbb{R}_+\) (independent of \(X\)) such that

\[
E_X(v, h) \leq \frac{\gamma}{2} \int_0^1 |v_t|^2 dt + C \left( \int_0^1 |v_t|^2 dt \right) (\gamma/2) \int_0^1 |h_t|^2_{L^2(X)} dt < \infty. \tag{4}
\]

Now for any \((X, f)\) and \((X', f')\) \(\in \mathcal{F}\) we define

\[
d_{\mathcal{F}}((X, f), (X', f')) \hat{=} \left( \inf \left\{ E_X(v, h) \mid \phi^\nu_v \cdot X = X', \ f' = \left( f + \zeta_t^h \right) \circ (\phi^\nu_1)^{-1} \right\} \right)^{1/2} \tag{5}
\]

which is a true distance on \(\mathcal{F}\) as stated in the following theorem:

**Theorem 1** The function \(d_{\mathcal{F}}\) given by (5) defines a distance on \(\mathcal{F}\), i.e., is symmetric, satisfies the triangle inequality and \(d_{\mathcal{F}}((X, f), (X', f')) = 0\) iff \(X = X'\) and \(f = f' \in L^2(X)\).

Moreover, the distance is finite everywhere on \(\mathcal{F} \times \mathcal{F}\).

**Proof** The symmetry comes from a usual process of time reversal.

Let \((X, f), (X', f') \in \mathcal{F}\) and \((v, h) \in L^2([0, 1], V \times L^2(X))\) such that

\[
X' = \phi^\nu_1 \cdot X, \ f' = \left( f + \zeta_t^h \right) \circ (\phi^\nu_1)^{-1}. \tag{6}
\]

If we define

\[
\tilde{v}_t \hat{=} -v_{1-t} \text{ and } \tilde{h}_t \hat{=} -h_{1-t} \circ (\phi^\nu_1)^{-1}
\]

for any \(t \in [0, 1]\), then for \(X_t \hat{=} \phi^\nu_v \cdot X\) (so that \(X_0 = X\) and \(X_1 = X'\)) we have

\[
\phi^\nu_v \cdot X_0 = X_t = \phi^\nu_{1-t} \cdot X_1
\]

and \((\tilde{v}, \tilde{h}) \in L^2([0, 1], V \times L^2(X'))\) with

\[
E_X(v, h) = E_{X'}(\tilde{v}, \tilde{h}). \tag{7}
\]

Since one easily checks that \(\zeta_t^h = -\zeta_t^h \circ (\phi^\nu_1)^{-1}\) so that if \(f' = (f + \zeta_t^h) \circ (\phi^\nu_1)^{-1}\), we have \(f = (f' + \zeta_t^h) \circ (\phi^\nu_1)^{-1}\) with \(X = \phi^\nu_1 \cdot X'\). This gives immediately the symmetry.

Concerning the triangle inequality, it comes from a usual process of path concatenation. Let \((X, f), (X', f')\) and \((X'', f'')\) be three fshapes in \(\mathcal{F}\) such that

\[
\mathcal{F} \ni \begin{matrix}
\circlette\bigcirc\bigcirc
\end{matrix}
\]

\[
\mathcal{F} \ni \begin{matrix}
\circlette\bigcirc\bigcirc
\end{matrix}
\]
$d_{\mathcal{F}}((X, f), (X', f')) > 0$ and $d_{\mathcal{F}}((X', f'), (X'', f'')) > 0$. One easily checks that for $(v, h) \in L^2([0, 1], V \times L^2(X))$ and $(v', h') \in L^2([0, 1], V \times L^2(X'))$ with

$$(X', f') = \left( \phi_1^v \cdot X, \left( f + \zeta_1^h \right) \circ (\phi_1)^{-1} \right)$$

and

$$(X'', f'') = \left( \phi_1^{v'} \cdot X', \left( f' + \zeta_1^h \right) \circ (\phi_1^{v'})^{-1} \right)$$

then denoting for $\alpha, \beta > 1$ such that $1/\alpha + 1/\beta = 1$ and $s \in [0, 1]$

$$\text{Cat}_\alpha \left( (v', h'), (v, h) \right) = \beta \left( v'_\beta(s-1/\alpha), h'_\beta(s-1/\alpha) \circ \phi \right) 1_{s \geq 1/\alpha} + \alpha \left( v_{as}, h_{as} \right) 1_{0 \leq s < 1/\alpha}$$

with $\phi = \phi_1^v$, we have for $(\tilde{v}, \tilde{h}) \doteq \text{Cat}_\alpha ((v', h'), (v, h))$ that $(\tilde{v}, \tilde{h}) \in L^2([0, 1], V \times L^2(X))$ and

$$(X'', f'') = \left( \phi_1^{\tilde{v}} \cdot X, \left( f + \zeta_1^{\tilde{h}} \right) \circ (\phi_1)^{-1} \right)$$

so that

$$d_{\mathcal{F}} ((X, f), (X'', f'')) \leq E_X \left( \tilde{v}, \tilde{h} \right)^{1/2}.$$

However, for $\alpha_* = (E_X (v, h)^{1/2} + E_X (v', h')^{1/2}) / E_X (v, h)^{1/2}$, we check easily that

$$E_X \left( \tilde{v}, \tilde{h} \right)^{1/2} = E_X (v, h)^{1/2} + E (v', h')^{1/2}$$

and the triangle inequality follows immediately.

Now, if $d_{\mathcal{F}}((X, f), (X', f')) = 0$ then there exists a sequence $\phi_n$ such that $\phi_n \cdot X = \phi_n(X) = X'$ and $\phi_n \to \text{Id}$ on $X$. In particular, $X'$ is dense in $X$. As $X'$ is compact (since $X$ is), $X \subset X'$. By symmetry we get $X = X'$. Moreover, there exists also a sequence $\zeta_n \in L^2(X)$ such that $\zeta \to 0$ in $L^2(X)$ and $f' = (f + \zeta_n) \circ \phi_n^{-1}$ with $\phi_n(X) = X$. We get

$$\int_X (f' - f)^2 d\mathcal{H}^d (x) \leq 2 \int_X \left( f \circ \phi_n^{-1} - f \right)^2 d\mathcal{H}^d (x) + 2 \int_X \left( \zeta_n \circ \phi_n^{-1} \right)^2 d\mathcal{H}^d (x) \to 0$$

and the result is proved.

A last point to be verified is that the distance $d_{\mathcal{F}}$ is finite between any pair of $f$-shapes. Indeed, by definition there exists between any $X$ and $X' \in X_0$ a path $i \mapsto \phi_i^{v_i} \cdot X$ with $v_i \in L^2([0, 1], V)$. Now, if $h_i = (f' \circ \phi_i^{v_i} - f)$, we get $(v, h) \in L^2([0, 1], V \times L^2(X))$ and $f' = (f + \zeta_i^{v_i}) \circ (\phi_i^{v_i})^{-1}$ so that $d_{\mathcal{F}}((X, f), (X', f')) \leq E_X (v, h)^{1/2} < \infty$.
As stated in the introduction, the framework extends metamorphoses on images [36] which corresponds to the case where \( X_0 \) is \( E \) itself or is the unit cube of \( E \) (with the extra condition that \( \phi \cdot X_0 = X_0 \) for any \( \phi \in G \)). Here the support can be \( d \)-dimensional and is not fixed.

### 2.3 Existence of Geodesics

A natural question is the existence of a minimizing geodesic between fshapes in \( \mathcal{F} \). We have the following theorem:

**Theorem 2** For any \((X, f)\), and \((X', f')\) in \( \mathcal{F} \), there exists \((v, h) \in L^2([0, 1], V \times L^2(X))\) such that \( E_X(v, h)^{1/2} = d_\mathcal{F}((X, f), (X', f')) \).

In particular, if

\[
(X_t, f_t) \doteq (\phi^{v}_t \cdot X, (f + \int_0^t h_s ds) \circ (\phi^{v}_t)^{-1})
\]

the path \( t \mapsto (X_t, f_t) \) can be considered as a minimizing geodesic between \((X, f)\) and \((X', f')\).

**Proof** The proof extends the results for metamorphoses. It is sufficient to show that \((v, h) \mapsto E_X(v, h)\) is lower semicontinuous for the weak convergence of the space \( L^2([0, 1], V \times L^2(X)) \).

Indeed, if this is the case, then from any minimizing sequence \((v_n, h_n)\) such that \( E_X(v_n, h_n) \to d_\mathcal{F}((X, f), (X', f'))^2 \) we deduce that \( v_n \) is bounded on \( L^2([0, 1], V) \) and using the inequality

\[
\int_0^1 \int_X |h_{n,t}(x)|^2 d\mathcal{H}^d(x) dt \\
\leq \int_0^1 \left( \sup_{x \in X} |d_x \phi_t^{v_n} \mid_{\mathcal{T}_X}^{-1} \right) dr \int_X |h_{n,t}(x)|^2 \left( d_x \phi_t^{v_n} \mid_{\mathcal{T}_X} \right) d\mathcal{H}^d(x) \\
\leq C \left( \int_0^1 |v_{n,t}|^2 \right) E_X(v_n, h_n)
\]

where \( C(\cdot) \) is an increasing function depending only on \( V \), we get that the sequence \((h_n)\) is bounded in \( L^2([0, 1], L^2(X)) \). Hence, by weak compactness of strong balls in \( L^2([0, 1], V \times L^2(X)) \), we can assume that, up to the extraction of a subsequence, \((v_n, h_n)\) weakly converges toward \((v_\infty, h_\infty) \in L^2([0, 1], V \times L^2(X)) \) and by lower semicontinuity of \( E_X \), we deduce that \( E_X(v_\infty, h_\infty) \leq d_\mathcal{F}((X, f), (X', f'))^2 \). We only need to check that \((X', f') = (\phi^{v_\infty}_1 \cdot X, (f + \xi^{h_\infty}_1) \circ (\phi^{v_\infty}_1)^{-1}) \). This last result follows from the fact that if \((v_n, h_n)\) weakly converges to \((v_\infty, h_\infty)\), then \( \phi^{v_n}_1 \to \phi^{v_\infty}_1 \) uniformly on any compact sets (which is a well-known result) and \((f + \xi^{h_n}_1) \circ (\phi^{v_n}_1)^{-1} \) weakly converge to \((f + \xi^{h_\infty}_1) \circ (\phi^{v_\infty}_1)^{-1} \). The last weak convergence is straightforward.
The proof of the lower semicontinuity itself is done now. We know from a classical result on weak convergence that \( \int_0^1 |v_{x,t}|^2 \, dt \leq \liminf_{n \to \infty} \int_0^1 |v_{x,t}|^2 \, dt \). Moreover, we have

\[
\int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{x}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x) \, dt \\
= \lim_{n \to \infty} \int_0^1 \int_X h_{x,t}(x) h_{x,t}(x) \left( \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \right)^{1/2} \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X}^{1/2} \, d\mathcal{H}^d(x) \, dt \\
\leq \liminf_{n \to \infty} \left( \int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x) \, dx \right)^{1/2} \\
\times \left( \int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x) \, dt \right)^{1/2}
\]

by definition of the weak convergence of \((h_n)\). Now, since \(d_x \phi_{t}^{v_{n}} \to d_x \phi_{t}^{v_{n}}\) uniformly on \(t \in [0, 1]\) and \(x \in X\), we get

\[
\int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x) \, dt \\
\leq \lim_{n \to \infty} \int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x) \, dx \\
\leq \liminf_{n \to \infty} \int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x).
\]

by definition of the weak convergence of \((h_n)\). Now, since \(d_x \phi_{t}^{v_{n}} \to d_x \phi_{t}^{v_{n}}\) uniformly on \(t \in [0, 1]\) and \(x \in X\), we get

\[
\int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x) \, dt \\
\leq \lim_{n \to \infty} \int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x) \, dx \\
\leq \liminf_{n \to \infty} \int_0^1 \int_X |h_{x,t}(x)|^2 \left| d_x \phi_{t}^{v_{n}} \right|_{\mathcal{T}_X} \, d\mathcal{H}^d(x).
\]

\[\square\]

### 2.4 Karcher Means on Fshape Bundles

We consider here the existence problem of Karcher mean on the fshape bundle \(\mathcal{F}\). The problem can be stated as follows: Let \((X^i, f^i)_{1 \leq i \leq N}\) be a family of fshapes in \(\mathcal{F}\). Does there exist \((X^*, f^*) \in \mathcal{F}\) minimizing the sum of the square distances to each \((X^i, f^i)\), i.e., solution of the minimizing problem:

\[
\min_{(X, f) \in \mathcal{F}} \sum_{i=1}^{N} d_{\mathcal{F}} \left( (X, f), (X^i, f^i) \right)^2.
\]

Note that this problem does not make sense on practical ground since usually there is no reason for a family of observed fshapes to belong to a predefined fshape bundle \(\mathcal{F}\). However, on a theoretical perspective the existence of such Karcher mean is a quite
important point in addition to the existence of geodesic between any two fshapes in \( \mathcal{F} \). Such a result shows that \( \mathcal{F} \) has basic important properties for further statistical analysis.

Our main result is the following:

**Theorem 3** For any family \((X^i, f^i)_{1 \leq i \leq N}\) of fshapes in \( \mathcal{F} \), there exists at least a Karcher mean \((X_*, f_*)\) in \( \mathcal{F} \), i.e., a solution of problem (17).

**Proof** The first step of the proof is to reparametrize problem (17) in the Hilbert space

\[
W \doteq \prod_{i=1}^{N} L^2 \left( \left[ 0, 1 \right], V \times L^2 \left( X^i \right) \right)
\]

Indeed, using the symmetry of the problem, we can exchange the role of \((X, f)\) and \((X^i, f^i)\) and look for a family of minimizing paths \( t \mapsto (X^i_t, f^i_t) \), starting from every fshape \((X^i, f^i)\) and ending on a common fshape \((X, f) \in \mathcal{F}\), each path being parametrized by \((v^i, h^i) \in L^2 \left( \left[ 0, 1 \right], V \times L^2 \left( X^i \right) \right)\). Hence, introducing

\[
W_0 \doteq \left\{ \alpha = \left( v^i, h^i \right)_{1 \leq i \leq N} \in W \mid \left( X^i_1, f^i_1 \right) = \left( \tilde{X}^i_1, \tilde{f}^i_1 \right), \ \forall 1 \leq i \leq N \right\}
\]

where \((X^i_1, f^i_1) = (\phi_1^{v^i} \cdot X^i, (f^i + \xi_1^{h^i}) \circ (\phi_1^{v^i})^{-1})\) we need to prove that if we define

\[
J(\alpha) \doteq \sum_{i=1}^{N} E_{X^i} (v^i, h^i) \quad (18)
\]

then \(W_0\) is closed and \(J\) is lower semicontinuous (l.s.c.) for the weak topology on \(W\).

Since the fact that \(J\) is l.s.c. was established inside the proof of Theorem 2, we just need to check that \(W_0\) is closed for the weak topology. So, let \((\alpha_n)_{n \geq 0}\) be a sequence in \(W_0\) weakly converging to \(\alpha_\infty \doteq (v_\infty^i, h_\infty^i)_{1 \leq i \leq N} \in W\), which we will denote by \(\alpha_n \rightharpoonup \alpha_\infty\). We get that \(\phi_1^{v^i_n} \to \phi_1^{v^i_\infty}\) and \((\phi_1^{v^i_n})^{-1} \to (\phi_1^{v^i_\infty})^{-1}\) for the uniform \(C^1\) topology on compact sets so that in particular, if \(X_* = \phi_1^{v^i_\infty} \cdot X^i\) then since by construction \(\phi_1^{v^i_n} \cdot X^i = \phi_1^{v^i_1} \cdot X^i\) we get that \(d_H(\phi_1^{v^i_n} \cdot X^i, X_*) \to 0\) where \(d_H\) denotes the Hausdorff distance\(^1\) and

\[
\phi_1^{v^i_\infty} \cdot X^i = X_* \quad (19)
\]

Now, for any \(1 \leq i \leq N\), if we denote \(f^i_{1,n} \doteq (f^i + \xi_1^{h^i_n})\) and \(\psi_n^i \doteq (\phi_1^{v^i_n})^{-1} \circ \phi_1^{v^i_1} : X^1 \to X^i\), we have (since \(\alpha_n \in W_0\) and \(\alpha_n \rightharpoonup \alpha_\infty\))

\[
f^i_{1,n} \circ \psi_n^i = f^i_{1,n} \text{ and } f^i_{1,n} \rightharpoonup f^i_{1,\infty} \quad (20)
\]

\(^1\) We are using here that for \(X\) compact, the mapping \(v \mapsto \phi_1^{v^i} \cdot X\) is continuous for the weak convergence on \(v \in L^2 \left( \left[ 0, 1 \right], V \right)\) and the convergence for the Hausdorff metric on the set of all compact subsets of \(E\).
where the last statement is straightforward. Let us check that \( \tilde{f}_{1,n}^i \circ \psi_n \to \tilde{f}_{1,\infty}^i \circ \psi_\infty \) so that we will get \( \tilde{f}_{1,\infty}^i \circ \psi_\infty = \tilde{f}_{1,\infty}^i \) and since we have (19) we will get \( \alpha_\infty \in W_0 \) and \( W_0 \) weakly closed.

For that, let \( g : X^1 \to \mathbb{R} \) be a Lipschitz mapping on \( X^1 \) and denote now \( \langle f, f' \rangle_{X^1} = \int_{X^1} f(x) f'(x) d\mathcal{H}^d(x) \) the usual dot product on \( L^2(X^1) \). Since \( X^1 \) is compact and \( \mathcal{H}^d(X^1) < \infty \), we have that \( \mathcal{H}^d \big|_{X^1} \) is a Radon measure on the compact metric space \( X^1 \) so that such Lipschitz mapping \( g \) are dense in \( L^2(X^1) \) and we just need to check that \( \langle \tilde{f}_{1,n}^i \circ \psi_n - \tilde{f}_{1,\infty}^i \circ \psi_\infty, g \rangle_{X^1} \to 0 \) to prove the weak convergence. We have

\[
\begin{align*}
\langle \tilde{f}_{1,n}^i \circ \psi_n - \tilde{f}_{1,\infty}^i \circ \psi_\infty, g \rangle_{X^1} &= \left\langle \tilde{f}_{1,n}^i \circ \psi_n - \tilde{f}_{1,n}^i \circ \psi_\infty, g \right\rangle_{X^1} + \left\langle \tilde{f}_{1,n}^i \circ \psi_\infty - \tilde{f}_{1,\infty}^i \circ \psi_\infty, g \right\rangle_{X^1}.
\end{align*}
\]

(21)

Concerning the \( B \) term, by change of variable and using the fact that the Jacobian \( d(\psi_\infty)^{-1} \) is bounded on \( X^1 \) we get

\[
B_n = \left\langle \tilde{f}_{1,n}^i - \tilde{f}_{1,\infty}^i, g \circ (\psi_\infty)^{-1} \left| d \left( \psi_\infty^i \right)^{-1} . \xi \right| \right\rangle_{X^1} \to 0,
\]

(22)

where \( \xi \) is the unit \( d \)-vector representing the tangent space defined \( \mathcal{H}^d \big|_{X^1} \)-a.e. on \( X^i \). Concerning the \( A \) term, we have again by change of variable

\[
A_n = \left\langle \tilde{f}_{1,n}^i, g \circ (\psi_n)^{-1} \left| d \left( \psi_n^i \right)^{-1} . \xi \right| - d \left( \psi_\infty^i \right)^{-1} . \xi \right\rangle_{X^i} + \left\langle \tilde{f}_{1,n}^i, \left( g \circ (\psi_n)^{-1} - g \circ (\psi_\infty)^{-1} \right) \left| d \left( \psi_\infty^i \right)^{-1} . \xi \right| \right\rangle_{X^i}.
\]

(23)

The first term of (23) is bounded by

\[
\left| \tilde{f}_{1,n}^i \right|_{L^2(X^i)} |g|_{L^\infty(X^1)} \left| d \left( \psi_n^i \right)^{-1} . \xi \right| - d \left( \psi_\infty^i \right)^{-1} . \xi \right|_{L^\infty(X^i)} \to 0
\]

(24)

since the weak convergence of \( \tilde{f}_{1,n}^i \) implies that \( \left| \tilde{f}_{1,n}^i \right|_{L^2(X^i)} \) stays bounded. The second term of (23) is bounded by

\[
|g|_{\text{Lip}} \left| \tilde{f}_{1,n}^i \right|_{L^2(X^i)} \left| (\psi_n^i)^{-1} - (\psi_\infty^i)^{-1} \right|_{L^\infty(X^i)} \left| d \left( \psi_n^i \right)^{-1} . \xi \right|_{L^\infty(X^i)} \to 0
\]

(25)

where \( |g|_{\text{Lip}} \) is the Lipschitz constant of \( g \) on \( X^i \). \( \square \)
2.5 The Tangential Model

The previous metamorphosis metrics can be also approximated in a simpler setting, which becomes closer to a natural extension of the classical image framework in the context of fshapes. We shall call it the **tangential model**. As we shall see in Sect. 6.5, it also has the important advantage of leading to much simpler and more convenient equations for the signal dynamics simplifying the implementation. Instead of computing the cost of the functional evolution \( t \mapsto \zeta_t \) along a time dependent \( L^2(X_t) \) metric

\[
\int_0^1 \int_X |\dot{\zeta}_t(x)|^2 \, dx \, |\xi_t \circ (\phi_t^y)^{-1}|^2(y) \, dH^d(y) \, dt
\]

one can freeze the metric to its initial value \( X_0 = X \) neglecting the change of measure weight:

\[
\int_0^1 \int_X |\dot{\zeta}_t(x)|^2 \, dx \, |\xi_t^y||\phi_t^y|^2 \, dH^d(x) \, dt \approx \int_0^1 \int_X |\dot{\zeta}_t(x)|^2 \, dH^d(x) \, dt
\]

which gives after optimization with fixed end points condition \( \zeta_1 \) the usual \( L^2(X) \) cost

\[
|\zeta_1|^2_X = \int_X |\zeta_1(x)|^2 \, dH^d(x)
\]

which can be interpreted naturally as the log-likelihood of a Gaussian noise in the statistical framework.

Note as we say in the introduction of this section that this is the usual noise term that appears in the classical situation of image matching and this can be considered as some kind of tangential metric for the functional part. Concerning the geometrical part, we can keep the usual \( L^2([0, 1], V) \) penalization to define

\[
\tilde{E}_X(v, \zeta) = \gamma V \int_0^1 |v_t|_V^2 \, dt + \gamma f \int X |\zeta|^2_X
\]

and consider

\[
\tilde{d}_F((X, f), (X', f')) = \left( \inf \left\{ \tilde{E}_X(v, \zeta) \mid X' = \phi_t^v \cdot X, \quad f' = (f + \zeta) \circ (\phi_t^v)^{-1} \right\} \right)^{1/2}.
\]

With that definition, \( \tilde{d}_F \) is no more a distance satisfying the symmetry and triangle inequality as previously, but the minimization of \( \tilde{E}_X \) can be seen as the minimization of the log-likelihood of \((v, \zeta)\) given \((X, f)\) and \((X', f')\) for a quite natural statistical model.
In this framework, one can still consider the problem of existence for two given \( f \text{shapes} \) \((X, f)\) and \((X', f')\) in \( \mathcal{F} \) of an optimal \((v_\ast, \zeta_\ast)\) such that

\[
\begin{align*}
(X', f') &= (\phi_1^{v_\ast} \cdot X, (f + \zeta_\ast) \circ (\phi_1^{v_\ast})^{-1}) \\
\tilde{E}_X(v_\ast, \zeta_\ast) &= \min \left\{ \tilde{E}_X(v, \zeta) \mid (v, \zeta) \in F(X', f') \right\} \\
\text{where} \quad F(X', f') &= \left\{ (v, \zeta) \in L^2([0, 1], V) \times L^2(X) \mid (X', f') = (\phi_1^{v} \cdot X, (f + \zeta) \circ (\phi_1^{v})^{-1}) \right\}
\end{align*}
\]

The proof existence of \((v_\ast, \zeta_\ast)\) can be done along the very same lines that the proof of existence of geodesics and is omitted. The proof is even simpler since the functional metric is frozen and does not change along the path. We get eventually the following result:

**Theorem 4** For \((X, f)\) and \((X', f')\) in \( \mathcal{F} \), there exists an optimal solution of the exact matching problem between \((X, f)\) and \((X', f')\); i.e., there exists \((v_\ast, \zeta_\ast)\) in \( L^2([0, 1], V) \times L^2(X) \) such that (31) holds.

### 3 Dissimilarity Measure Between Fshapes

#### 3.1 Dissimilarity Measures: State of the Art

The previous framework has been focusing so far on the comparison of fshapes within one given common bundle \( \mathcal{F} \). It is clear that it becomes inoperant in most practical situations where the elements of a dataset cannot be assumed to belong to the same bundle. The reason is that the geometrical supports of two subjects need not be obtained from one the other by a deformation belonging to the group \( G \). In some more usual cases of shape spaces such as sets of images, landmarks, curves or surfaces, the general approach has been to introduce dissimilarity measures between the subjects, which provides data attachment terms to perform inexact registration and atlas estimation in practice.

For images and landmarks, \( L^2 \) distances are natural and have been commonly used to compare such objects. The case of curves, surfaces and submanifolds is theoretically more involved and has drawn consistent attention in several fields of mathematics. In computational anatomy, one possible and very powerful setting was proposed through the adaptation of the concept of currents [14, 16]. The dissimilarity measures between oriented submanifolds of given dimension defined in these frameworks derive from Hilbert metrics on spaces of differential forms and combine the advantages of not relying on parametrizations or point to point matchings, of being easily computable for discrete shapes and robust to shape sampling. One potential drawback is the issue of consistent orientation that is required in the currents’ representations. More recently, an alternative methodology based on varifolds was introduced and implemented in [12], which has the interest of being close to the idea of currents while overcoming the problem of orientation.
Now, fspheres do pose additional difficulties with respect to the definition of dissimilarity measures. This problem has been only addressed very recently in [11] where the authors define the extended notion of functional currents that, similarly to usual currents, embeds fspheres in some dual of spaces of differential forms. The Hilbert metrics that are defined on functional currents can be then used again as dissimilarity measures to perform inexact registration between fspheres, as exposed in the article. Functional currents could be thus very well used in the generalized problem of atlas estimation that we focus on in this paper. This has been touched upon, essentially from the numerical point of view, in [10] (chapter 4). Although the results and algorithms presented in this paper could be transposed quite easily to fshape dissimilarities provided by functional currents, we shall work instead with the slightly modified approach of varifolds. As aforementioned, this is a way to get rid of shape orientation which is usually quite desirable in applied situations. In the following section, we briefly present the functional varifolds’ mathematical setting that stands for a natural generalization of varifolds to fspheres. The more technical and discrete computations shall be detailed in Sect. 6 dedicated to the algorithmic part itself.

3.2 Functional Varifolds

The mathematical concept of varifold goes back to geometric measure theory and the original exposition of Almgren [2] which was further developed considerably by Allard in [1] subsequently. The link to computational anatomy’s problems is fairly more recent and is in most part presented in [12] or [10]. We shall frequently refer the reader to these references for additional mathematical details. What we present in this section is an extension of varifolds to represent functional shapes, which is much related to the functional current idea of [11].

3.2.1 Representation of Fshapes as Functional Varifolds

As previously, we call $E$ the $n$-dimensional vector space embedding all geometrical supports of fspheres. For any integer $1 \leq d \leq n$, we will write $G_d(E)$ for the Grassmann manifold of all $d$-dimensional (non-oriented) subspaces of $E$. $G_d(E)$ is a compact manifold that can be embedded trivially in the space $\mathcal{L}(E)$ of linear maps from $E$ to $E$, by identifying any $V \in G_d(E)$ with the orthogonal projector $p_V$ on $V$. The usual space of $d$-dimensional varifolds is defined intuitively as the space of spatially spread distributions of Grassmannians or, more rigorously, as finite Borel measures on the product $E \times G_d(E)$ (cf [12]). Now, to account for the existence of signals at each point, one can simply augment the varifold representation with a signal component, which we formalize by the following definition:

**Definition 1** We say that $\mu$ is a $d$-dimensional functional varifold (fvarifold in short) if $\mu$ is a Borel finite measure on the space $E \times G_d(E) \times \mathbb{R}$ or equivalently if $\mu \in C_0(E \times G_d(E) \times \mathbb{R})'$, the dual of the space of continuous functions vanishing at infinity.
Note that we consider only the case of real-valued signals here, but this could be extended to different situations, as vector or tensor-valued signals (cf [10]). Now, any functional shape \((X, f)\), with \(X\) a \(d\)-dimensional rectifiable subset, can be represented by a functional varifold \(\mu_{(X, f)}\) defined by:

\[
\mu_{(X, f)}(\omega) = \int_X \omega(x, T_x X, f(x))d\mathcal{H}^d(x)
\]

where \(\mathcal{H}^d\) is the \(d\)-dimensional Hausdorff measure on \(E\).

Some particularly simple functional varifolds are the Dirac delta distributions that, in this context, can be written in the form \(\delta_{(x, V, f)}\) with \(x \in E, V \in G_d(E)\) and \(f \in \mathbb{R}\), and act on any function \(\omega \in C_0(E \times G_d(E) \times \mathbb{R})\) by the relation:

\[
\delta_{(x, V, f)}(\omega) = \omega(x, V, f).
\]

We shall detail, in Sect. 6, how to approximate polyhedral fshapes by a finite sums of Dirac in order to perform practical computations.

Now, an important point is to express the way that geometric-functional transformations act on functional varifolds, in such a way that this action is consistent with the transport of fshapes that we have been considering in Sect. 2.2. In a very similar fashion as with regular varifolds (cf [12]), one can express it by usual pull-back and push-forward operations. Let \(\phi \in \text{Diff}(E)\) and \(\zeta\) a measurable function on \(E\), this is given by the following set of equations:

\[
\begin{cases}
\forall \omega \in C_0(E \times G_d(E) \times \mathbb{R}), \ ((\phi, \zeta)_*\mu) (\omega) = \mu ((\phi, \zeta)^*\omega) \\
\text{where} \\
((\phi, \zeta)^*\omega) (x, V, f) = |d_x \phi|_V \omega(\phi(x), d_x \phi(V), f + \zeta)
\end{cases}
\]

where for \(V \in G_d(E), |d_x \phi|_V\) denotes as previously the Jacobian of \(\phi\) along subspace \(V\) (i.e., the volume change along \(V\) at point \(x\)) and \(d_x \phi(V)\) is the image of \(V\) by the invertible linear application \(d_x \phi\). It is then a simple verification that one has the following consistency property:

**Proposition 1** Let \((X, f)\) be a fshape where \(X\) is a \(d\)-dimensional rectifiable subset and \(f\) a \(L^2\) function on \(X\). If \(\phi \in \text{Diff}(E)\) and \(\zeta \in L^2(X)\) then:

\[
(\phi, \zeta)_*\mu_{(X, f)} = \mu_{(\phi(X), (f + \zeta) \circ \phi^{-1})}.
\]

The proof does not involve any additional difficulty than in the case of usual varifolds, for which we refer the reader to [12].

Thus, functional varifolds enable all possible fshapes of given dimension to be embedded into a common space of distributions. The next step then is to equip such a space with a metric that shall induce a dissimilarity measure on the set of fshapes. This can be efficiently addressed by introducing reproducing kernels on the product \(E \times G_d(E) \times \mathbb{R}\), as we explain in the following text.
3.2.2 RKHS of Functional Varifolds

The use of reproducing kernel in the context of currents and varifold spaces has been argued and implemented many times, for instance in [10, 14, 16]. The general scheme is to consider Hilbert spaces embedded in $C_0(E \times G_d(E) \times \mathbb{R})$ which are built as the reproducing kernel Hilbert space (RKHS) associated with a certain positive kernel on $E \times G_d(E) \times \mathbb{R}$. A natural and convenient (but not exhaustive) way to define kernels on product spaces is to consider tensor products of kernels, the classical result from kernel theory [5] being that:

**Lemma 1** Let $A$ and $B$ be two sets and $k_A, k_B$ positive kernels, respectively, on $A$ and $B$. Then, $k_A \otimes k_B$ defined for all $a_1, a_2 \in A$ and $b_1, b_2 \in B$ by:

$$k_A \otimes k_B((a_1, b_1), (a_2, b_2)) = k_A(a_1, a_2)k_B(b_1, b_2)$$

is a positive kernel on $A \times B$.

Now going back to the case of functional varifolds itself, one has the following property:

**Proposition 2** Let $k_e$ be a positive kernel on the space $E$ such that $k_e$ is continuous, bounded and for all $x \in E$, the function $k_e(x, \cdot)$ vanishes at infinity, $k_t$ a kernel on the manifold $G_d(E)$ that is also continuous, and $k_f$ a kernel on $\mathbb{R}$ continuous, bounded and such that $k_f(f, \cdot)$ vanishes at infinity for all $f \in \mathbb{R}$. Then, $W$, the RKHS associated with the positive kernel $k = k_e \otimes k_t \otimes k_f$, is continuously embedded into the space $C_0(E \times G_d(E) \times \mathbb{R})$.

**Proof** The proof is very similar to the one of [12] in the case of varifolds. By Lemma 1, $k$ is indeed a positive kernel on $E \times G_d(E) \times \mathbb{R}$ and by definition:

$$k((x, V, f), (\tilde{x}, \tilde{V}, \tilde{f})) = k_e(x, \tilde{x})k_t(V, \tilde{V})k_f(f, \tilde{f})$$  \hspace{1cm} (35)

and so, thanks to the assumptions on the kernels, $k((x, V, f), \cdot)$ is continuous on $E \times G_d(E) \times \mathbb{R}$ and belongs to $C_0(E \times G_d(E) \times \mathbb{R})$. The vector space $W_0$ generated by these functions is thus included in $C_0(E \times G_d(E) \times \mathbb{R})$. Moreover, if $\omega \in W_0$, by the reproducing kernel property, we have that:

$$\omega(x, V, f) = \delta_{(x, V, f)}(\omega) = \langle k((x, V, f), \cdot), \omega \rangle_W.$$

With Cauchy–Schwarz inequality: $|\omega(x, V, f)| \leq \|k((x, V, f), \cdot)\|_W \|\omega\|_W$. In addition,

$$\|k((x, V, f), \cdot)\|_W = \sqrt{k((x, V, f), (x, V, f))}$$

and all three kernels $k_e, k_t$ and $k_f$ are bounded so that $k$ is also bounded. We conclude that $|\omega|_\infty \leq \sqrt{k(\cdot, \cdot)}_\infty \|\omega\|_W$. Thus, Cauchy sequences in $W_0$ for the $W$-norm are also Cauchy sequences for the infinite norm. It results that all their limits belong to $C_0(E \times$
$G_d(E \times \mathbb{R})$ and therefore $W$ is included in $C_0(E \times G_d(E) \times \mathbb{R})$. The previous inequality then holds for all $\omega \in W$, which shows that the inclusion embedding $i: W \hookrightarrow C_0(E \times G_d(E))$ is indeed continuous.

Consequently, there exists a continuous mapping $i^*$ of the space of fvarifolds $C_0(E \times G_d(E) \times \mathbb{R})'$ into the dual of $W$. This induces a pseudo-distance on fvarifolds resulting from the Hilbert structure of the RKHS. If we introduce the isometry $K_W: W' \rightarrow W$ defined by $(K_W \mu, \omega)_W = \mu(\omega)$ for all $\mu \in W'$ and $\omega \in W$, by the reproducing kernel property, we know that $K_W \delta_{(x,V,f)} = k((x, V, f), \cdot)$. Then, for all $x_1, x_2 \in E$, $V_1, V_2 \in G_d(E)$ and $f_1, f_2 \in \mathbb{R}$,

$$\langle \delta_{(x_1,V_1,f_1)}, \delta_{(x_2,V_2,f_2)} \rangle_W = \langle K_W \delta_{(x_1,V_1,f_1)}, K_W \delta_{(x_2,V_2,f_2)} \rangle_W = K_W \delta_{(x_2,V_2,f_2)}(x_1, V_1, f_1)$$

and thus, we have the following expression for the inner product between two Diracs:

$$\langle \delta_{(x_1,V_1,f_1)}, \delta_{(x_2,V_2,f_2)} \rangle_W = k_e(x_1, x_2) k_t(V_1, V_2) k_f(f_1, f_2). \quad (36)$$

Now, if $(X, f)$ and $(Y, g)$ are two fshapes (of dimension $d$) and $\mu_{(X,f)}$ and $\mu_{(Y,g)}$ the associated fvarifolds (as defined previously), one can prove easily the following formula:

$$\langle \mu_{(X,f)}, \mu_{(Y,g)} \rangle_W = \int_X \int_Y k_e(x, y) k_t(T_x X, T_y Y) k_f(f(x), g(y)) d\mathcal{H}^d(x) d\mathcal{H}^d(y). \quad (37)$$

One can thus propose a dissimilarity measure between any two fshapes that is simply the norm of the difference in the space of fvarifolds, i.e., $\|\mu_{(X,f)} - \mu_{(Y,g)}\|_W^2$, which can be computed easily using (37). We shall detail more specifically the discrete expressions in Sect. 6.

Yet, as we mentioned earlier, this does not necessarily result in a real distance on the space of varifolds because the dual application $i^*$ does not need to be an embedding. This is actually the case if and only if the RKHS $W$ is dense in $C_0(E \times G_d(E) \times \mathbb{R})$, in which case the kernel $k$ is said to be $c_0$-universal. Characterizations and properties of $c_0$-universal kernels are thoroughly studied and presented in [8] and [31]. It is shown among other interesting results that products of $c_0$-universal kernels are also $c_0$-universal. Thus, in our construction, we can obtain a $c_0$-universal kernel on $E \times G_d(E) \times \mathbb{R}$ if all kernels $k_e, k_t$ and $k_f$ are themselves $c_0$-universal. It is also an easy thing to check that the restriction of universal kernels to any subset preserves this property. Now, on a generic vector space $\mathbb{R}^m$, if one considers more specifically translation-invariant kernels that write $k_m(x, y) = \psi(x - y)$ with $\psi \in C_0(\mathbb{R}^m)$, the following general characterization is proved in [31] Proposition 8:

**Proposition 3** $k_m$ is a positive definite $c_0$-universal kernel on $\mathbb{R}^m$ if an only if $\psi$ is the Fourier transform of a finite positive Borel measure $\Lambda$ on $\mathbb{R}^m$ such that $\text{supp}(\Lambda) = \mathbb{R}^m$. 
This is for instance the case with all Gaussian kernels which are the Fourier transforms of other Gaussians, thus having support on all $\mathbb{R}^m$. The proposition also shows that kernels with compactly supported $\psi$ are $c_0$-universal. It gives anyhow a practical criterium to construct kernels which provide in the end a real dual distance on varifolds. Only the construction of kernel $k_t$ on the Grassmann manifold is not totally immediate but can be done through the aforementioned embedding of $G_d(E)$ in $L(E)$ (we refer to [12] for more details).

To conclude this section, let us mention a generalization of Proposition 2 to higher-order regularities that shall be useful in the rest of the paper.

**Proposition 4** Let $k$ be a positive kernel on the product space $E \times G_d(E) \times \mathbb{R}$ such that $k$ is continuously differentiable of order $2p$ and such that all the derivatives of $k$ up to the order $p$ are bounded. Assume in addition that for any $(x, V, f) \in E \times G_d(E) \times \mathbb{R}$, the function $k((x, V, f), \cdot)$ and all its derivatives up to order $p$ vanish at infinity. Then, the RKHS associated with $k$ is continuously embedded into $C^p_0(E \times G_d(E) \times \mathbb{R})$.

The proof follows the same pattern as the one of Proposition 2 and can be also adapted from the one given in [16] Chapter 2.

### 3.3 Properties of the Metrics

#### 3.3.1 Control Results

We now show a few control results on fvarifolds norm that shall be useful for the following. As a measure on $E \times G_d(E) \times \mathbb{R}$, we will call the total variation norm of a given fvarifold $\mu$ the quantity $\mu(E \times G_d(E) \times \mathbb{R})$. In particular, one can easily check that the total variation norm of a rectifiable fvarifold $\mu(X, f)$ equals $\mathcal{H}^d(X)$. We have the following control of $W'$-norms:

**Proposition 5** For a RKHS $W$ continuously embedded into $C_0(E \times G_d(E) \times \mathbb{R})$, there exists a constant $C_W > 0$ such that for all $\mu \in C_0(E \times G_d(E) \times \mathbb{R})$:

$$
\| \mu \|_{W'} \leq C_W \mu(E \times G_d(E) \times \mathbb{R})
$$

**Proof** From the continuous embedding property, we have the existence of $C_W > 0$ such that for all $\omega \in W'$, $|\omega|_\infty \leq C_W \| \omega \|_W$. Now, by definition of the kernel metric $\| \mu \|_{W'}^2 = \mu(K_W \mu)$ for $K_W$ the Riesz isometry between $W'$ and $W$ and thus:

$$
\| \mu \|_{W'}^2 = \mu(K_W \mu) = \int K_W \mu(x, V, F)d\mu(x, V, F) \leq |K_W \mu|_\infty \mu(E \times G_d(E) \times \mathbb{R}) \\
\leq C_W |K_W \mu|_\infty \mu(E \times G_d(E) \times \mathbb{R}) \\
\leq C_W \| \mu \|_{W'} \mu(E \times G_d(E) \times \mathbb{R})
$$

so that dividing on both sides by $\| \mu \|_{W'}$, we obtain the result. \qed
A direct corollary is that the RKHS norm of a rectifiable fvarifold is controlled by the $d$-volume of its geometrical support.

A second important property to establish is the continuity of RKHS norms with respect to deformations of geometrical supports. We first consider general $C^1$-diffeomorphisms of $E$ such that $\phi$ and $d\phi$ tends to $Id$ at infinity, and we will denote $\|\phi-Id\|_{1,\infty} = |\phi-Id|_{\infty} + |d\phi-Id|_{\infty}$. The result we show is the following:

**Proposition 6** Let $X$ be a $d$-dimensional rectifiable subset of $E$ of finite volume and $f \in L^2(X)$. Assume that $W$ is constructed as in Proposition 2 and is continuously embedded into $C^1_0(E \times G_d(E) \times \mathbb{R})$. Then:

$$\|\mu(X,f) - \mu(\phi(X), f \circ \phi^{-1})\|_{W'} \to 0$$

as $\|\phi-Id\|_{1,\infty} \to 0$.

**Proof** We start by writing:

$$\mu(X,f) = \int_X \delta(x,T_x X, f(x)) d\mathcal{H}^d(x)$$

and, in the same way,

$$\mu(\phi(X), f \circ \phi^{-1}) = \int_{\phi(X)} \delta(y,T_{\phi(X)}(\phi(X)), f(\phi^{-1}(y))) d\mathcal{H}^d(x).$$

Now, applying the area formula (Corollary 3.2.20 in [15]), we have:

$$\mu(\phi(X), f \circ \phi^{-1}) = \int_X \delta(\phi(x), d_x \phi(T_x X), f(x)) |d_x \phi|_{T_x X} |d\mathcal{H}^d(x).$$

where $|d_x \phi|_{T_x X}$ is the local volume change at $x$ along $T_x X$. We obtain therefore:

$$\|\mu(X,f) - \mu(\phi(X), f \circ \phi^{-1})\|_{W'} \leq \int_X \|d_x \phi|_{T_x X} \| \delta(\phi(x), d_x \phi(T_x X), f(x)) - \delta(x,T_x X, f(x)) \|_{W'} d\mathcal{H}^d(x). \quad (38)$$

Now, we assumed that the kernel $k$ of $W$ is built as a tensor product $k = k_e \otimes k_t \otimes k_f$. We can then write $W = W_e \otimes W_f$ where $W_e$ is the RKHS of kernel $k_e \otimes k_t$ and $W_f$ the one of kernel $k_f$. It is then straightforward to check that:

$$\|d_x \phi|_{T_x X} \| \delta(\phi(x), d_x \phi(T_x X), f(x)) - \delta(x,T_x X, f(x)) \|_{W'} = \|\delta f(x)\|_{W'_f} \|d_x \phi|_{T_x X} \| \delta(\phi(x), d_x \phi(T_x X)) - \delta(x,T_x X) \|_{W'_e}.$$
Since \( \|\delta_{f(x)}\|_{W'_g} = \sqrt{k_f(\varepsilon)} \) and \( k_f \) is bounded, we have \( \|\delta_{f(x)}\|_{W'_g} \leq C \) for some constant \( C > 0 \). Moreover:

\[
\left\| \int_{\varepsilon} \phi(x) dx \right\|_{W'_g} - \left\| \int_{\varepsilon} \phi(x) dx \right\|_{W'_g} \\
\leq \left\| \phi(x) \right\|_{W'_g} - \left\| \phi(x) \right\|_{W'_g} + \left\| \phi(x) \right\|_{W'_g} - \left\| \phi(x) \right\|_{W'_g}.
\]

Focusing on the first term, we have again, since \( k_e \) and \( k_t \) are bounded, that \( \|\delta_{\phi(x)}d_{\phi(T_x)} X)\|_{W'_g} \) is uniformly bounded on \( X \). In addition, the volume variation \( |d_x \phi|_{T_x X} - 1 \) converges to 0 uniformly on \( X \) whenever \( \|\phi - I|_{1,\infty} \rightarrow 0 \). As for the second term in the sum of (39), we know that \( \|\delta_{\phi(x)}d_{\phi(T_x)} X) - \delta_{\phi(x)}d_{\phi(T_x)} X)\|_{W'_g} = \sup_{\|\omega\|_{W'_g}} 1 \left| (\delta_{\phi(x)}d_{\phi(T_x)} X) - (\delta_{\phi(x)}d_{\phi(T_x)} X) (\omega) \right| \) and for any \( \omega \in W'_g \), we have:

\[
\left| (\delta_{\phi(x)}d_{\phi(T_x)} X) - (\delta_{\phi(x)}d_{\phi(T_x)} X) \right| = |\omega(x) - d_x \phi|_{T_x X} - \omega(x, T_x X)|
\]

and \( d_{E \times G_d} ((\phi(x), d_{\phi(T_x)} X), (x, T_x X)) \) is uniformly upper bounded for \( x \in X \) by \( C \|\phi - I|_{1,\infty} \) for some \( C > 0 \). In addition, with the assumptions on the kernels, we also have that \( |d\omega|_{H} \leq C \|\omega\|_{W'_g} \) for some \( C > 0 \). It results eventually that \( \|\delta_{\phi(x)}d_{\phi(T_x)} X) - \delta_{\phi(x)}d_{\phi(T_x)} X)\|_{W'_g} \) is bounded above by \( C \|\phi - I|_{1,\infty} \) for some constant \( C > 0 \). Now, since both terms in (39) converge to 0 uniformly for \( x \in X \), it results the same behavior for the integral of (38), which completes the proof.

Now, going back to the previous model of deformations, we consider diffeomorphisms \( \phi = \phi_1^\nu \) obtained as the flow at time 1 of a time-varying vector fields \( v \in L^2([0, 1], V) \), where \( V \) is a given Hilbert space of vector fields on \( E \). The result of Proposition 6 implies the following important corollary:

**Corollary 1** If \( V \) is continuously embedded into \( C_0^2(E, E) \) then, for any fixed and bounded rectifiable set \( X \) of finite volume and \( L^2 \) signal \( f \) on \( X \), the application \( v \mapsto \mu_{(\phi_1^\nu(X)), f \circ (\phi_1^\nu)^{-1}} \) is weakly continuous from \( L^2([0, 1], V) \) to \( W' \).

**Proof** This essentially relies on classical results on differential equations and flows that can be found in [16, 38]. It is shown in particular that, since \( V \hookrightarrow \mu_{(\phi_1^\nu(X)), f \circ (\phi_1^\nu)^{-1}} \), diffeomorphisms \( \phi_1^\nu \) and their differentials tend to \( I| \) at infinity. In addition, if \( v_n \) is a sequence that weakly converges to \( v \) in \( L^2([0, 1], V) \), then \( \phi_1^{v_n} \) and \( d\phi_1^{v_n} \) converge, respectively, to \( \phi_1^v \) and \( d\phi_1^{v_n} \) uniformly on every compact subset of \( E \). Then, thanks to proposition 6, we are allowed to conclude that \( \mu_{(\phi_1^{v_n}(X)), f \circ (\phi_1^{v_n})^{-1}} \rightharpoonup \mu_{(\phi_1^v(X)), f \circ (\phi_1^v)^{-1}} \), which proves the weak continuity.

### 3.3.2 Variation Formula for Fvarifold Metrics

We now extend the variation formula shown in [12] to the present setting of functional varifolds. The aim is to have a theoretical description of the variations of fvarifold metrics with respect to variations of a functional shape in both its geometrical support...
and signal. As we shall see, the behavior is qualitatively similar, except for the orientation, to the setting of currents that was examined in [10]. Fixing a functional shape \((X, f)\) with \(X\) a compact submanifold and \(f : X \to \mathbb{R}\) a \(C^1\) signal on \(X\), we wish to compute variations of terms like \(\langle \mu(X, f), \mu \rangle \) for any \(\mu \in W'\). Using the previous isometry \(K_W\), we know that \(\langle \mu(X, f), \mu \rangle W' = \langle K_W \mu(X, f), K_W \mu \rangle W = \mu(X, f) (\omega)\) with \(\omega = K_W \mu\) thanks to the reproducing kernel property, and thus, one is led to consider variation of terms \(\mu(X, f)(\omega)\) for \(\omega \in W\). Rigorously, such variations can be expressed with respect to the action of infinitesimal deformations in the geometric-functional domain. One can define such infinitesimal geometric-functional deformations by considering, exactly as in the previous tangential model, a \(C^1\) compactly supported vector field \(v\) on \(E\) and a \(L^2\) function \(h\) on \(X\). Now we shall denote by \(\psi_t\) the flow of \(v\) at time \(t\) and by \(\psi_t\) the application defined on \(X \times \mathbb{R}\) by \(\psi_t(x, m) = (\phi_t(x), m + th(x))\). In addition, \(\psi\) can be extended straightforwardly to the whole space \(E \times \mathbb{R}\) by setting \(\psi_t(x, m) = (\phi_t(x), m + th(x))\) where \(h|_X = h\) and \(h = 0\) elsewhere. The function \(h\) is then measurable on \(E\) and the push-forward action \(\psi^*\omega\) is well defined by equation (34). The variation we wish to compute can be expressed by:

\[
\frac{d}{dt} \bigg|_{t=0} \mu(\phi_t(X), (f + th) \circ \phi^{-1})(\omega) = \frac{d}{dt} \bigg|_{t=0} \int_X \psi_t^* \omega(x, T_x X, f(x)) d\mathcal{H}^d(x)
= \int_X \frac{d}{dt} \bigg|_{t=0} \psi_t^* \omega(x, T_x X, f(x)) d\mathcal{H}^d(x)
= \int_X (\xi(v, h) \omega)(x, T_x X, f(x)) d\mathcal{H}^d(x)
\]

where \(\xi(v, h) \omega \equiv \frac{d}{dt} \bigg|_{t=0} \psi_t^* \omega\) is a notation for the derivative of function \(\omega\) in the direction of \((v, h)\). Now, the result is the following:

**Theorem 5** Let \(X\) be a compact submanifold and \(f\) a \(C^1\) function on \(X\). Let \(v\) be any \(C^1\) vector field with compact support defined on \(E\) and \(h\) a \(L^2\) function on \(X\). We denote by \(\psi_t(x, m) = (\phi_t(x), m + th(x))\). Then, we have for any \(C^1\) function \((x, V, m) \mapsto \omega(x, V, m)\) on \(E \times G_d(E) \times \mathbb{R}\):

\[
\int_X (\xi(v, h) \omega)(x, T_x X, f(x)) d\mathcal{H}^d(x) = \int_X \left( \frac{\partial \omega}{\partial x} - \text{div}_X \left( \frac{\partial \omega}{\partial V} \right) - \omega H_X \left| v^\perp \right) + \frac{\partial \omega}{\partial m} (h - \langle \nabla f, v^\top \rangle) d\mathcal{H}^d
+ \int_{\partial X} \left( v, \omega v^\top + \left( \frac{\partial \omega}{\partial V} \right| v^\perp \right) d\mathcal{H}^{d-1}
\]

where \(v^\top\) and \(v^\perp\) denote the tangential and normal part of \(v\) along \(X\), \(v\) is the unit outward normal along \(\partial X\), and \(H_X\) the mean curvature vector to \(X\).

The proof is given in “Appendix 1.” This formula is interesting at several levels because it gives qualitative information on the gradients of the kernel metrics with respect to points and signal values of the fshape \((X, f)\). In particular, as already

\[\text{ Springer}\]
pointed out in [12], for a constant signal $f$, we see that, for points in the interior of $X$, the variation with respect to the geometrical support only involves the orthogonal component of $v$ which means that the gradient of the metric is orthogonal to the shape. This is somehow quite natural since tangential components of $v$ do not change the shape itself in that situation. However, we see that for non-constant signals, a term involving the tangential component $v^\top$ appears in the variation and is concentrated on regions of important gradient of the signal $f$. The second significant consequence to mention is the presence of special terms on the boundary of $X$. On points located on $\partial X$, the gradient is orthogonal to the boundary (but not necessarily to $X$). The presence of these boundary singularities have important consequences on the numerical behavior of the gradients that we shall address with more details in Sect. 9.

4 Mathematical Formulation of Atlas Estimation

To go beyond Karcher means within a single bundle (as exposed in Sect. 2.4) and address the problem of atlas computation, we have to consider now a family of (possibly noisy) observed fshapes $(X^i, f^i)_{1 \leq i \leq N}$. The main point here is the optimization of a template $(X, f)$ given the observations. We propose, in this section, a way to formalize the atlas estimation as a variational problem. The existence of solutions to such problems will be examined in the next section.

4.1 Template Space

Ideally, the template should be optimized among all the possible templates to avoid any bias effect toward a particular configuration but this looks as a quite badly posed optimization problem. We follow here a simpler and secure route by introducing a priori a large but restricted space of possible templates as another vector bundle $\mathcal{F}_0$ generated by a hypertemplate $(X_0, f_0)$. As presented previously, $\mathcal{F}_0$ is defined from the $(X_0, f_0)$ as the set of every fshape defined from the mapping $(\phi, \zeta) \mapsto (\phi \cdot X_0, (f_0 + \zeta) \circ \phi^{-1})$ for $(\phi, \zeta) \in G_0 \times L^2(X_0)$ with $G_0 \cong \mathbb{GV}_0$ is the group of diffeomorphisms generated by an additional RKHS space $V_0$ of vector fields on $E$:

$$\mathcal{F}_0 \doteq \{ (X, f) \mid (X, f) = (\phi \cdot X_0, (f_0 + \zeta) \circ \phi^{-1}), (\phi, \zeta) \in G_0 \times L^2(X_0) \}. \hspace{1cm} (41)$$

One could take $V_0 = V$, but we may want less regularity on $V_0$ than on $V$ to generate a larger space and get closer to the ideal case.

Now, given any $(X, f) \in \mathcal{F}_0$, we should consider the associated fshape bundle $\mathcal{F}_X$

$$\mathcal{F}_X \doteq \left\{ (X', f') \mid (X', f') = (\phi \cdot X, (f + \zeta) \circ \phi^{-1}), (\phi, \zeta) \in G \times L^2(X) \right\} \hspace{1cm} (42)$$

corresponding to the previously introduced space $\mathcal{F}$. 
4.2 Variational Formulation (Metamorphosis Riemannian Setting)

In the Riemannian setting, the problem of atlas estimation can be addressed as the minimization of

\[
J_{\text{riem}}^1 \left( (X, f), \left( X^i, f^i \right)_i \right) = d_{\mathcal{F}_0} ((X_0, f_0), (X, f))^2 + \sum_{i=1}^{N} \left( d_{{\mathcal{F}_X}} \left( (X, f), \left( \tilde{X}^i, \tilde{f}^i \right) \right) \right)^2 + g \left( \left( \tilde{X}^i, \tilde{f}^i \right), \left( X^i, f^i \right) \right) \quad (43)
\]

on

\[
(X, f) \in \mathcal{F}_0 \text{ and } \left( \tilde{X}^i, \tilde{f}^i \right)_{1 \leq i \leq N} \in \mathcal{F}_X^N \quad (44)
\]

where \( g \) is the “out-of-orbit” or previous data attachment term, which, in that case, can be defined, as we explained in Sect. 3.2 by fvarifold Hilbertian metrics.

The problem can be then reformulated equivalently as a minimization problem on

\[
\begin{cases}
(v^0, h^0) \in L^2 ([0, 1], V_0 \times L^2(X_0)) \\
(v^i, h^i) \in L^2 ([0, 1], V \times L^2(X)) , \quad 1 \leq i \leq N,
\end{cases}
\]

with

\[
X = \phi_1^0 \cdot X_0
\]

of

\[
J_{\text{riem}}^2 ((v^0, h^0), (v^i, h^i)_i) = \frac{\gamma V_0}{2} \int_0^1 \left| v^0_t \right|_{V_0}^2 \, dt + \frac{\gamma f}{2} \int_0^1 \int_{X_0} \left| h^0_t \right|_{V}^2 \, dx \, dt + \frac{\gamma W}{2} g \left( \left( \tilde{X}^i, \tilde{f}^i \right), \left( X^i, f^i \right) \right) \quad (46)
\]

where

\[
\begin{cases}
\tilde{f}^i = \left( f + \xi_1^{h^i} \right) \circ \left( \phi_1^{v^i} \right)^{-1} , \quad 1 \leq i \leq N, \\
\text{and}
\end{cases}
\]

\[
f = \left( f_0 + \xi_1^{h^0} \right) \circ \left( \phi_1^{v^0} \right)^{-1}.
\]

(47)
4.3 Variational Formulation (Tangential Setting)

The atlas estimation problem can also be formulated using the tangential metric and with a statistical flavor. It seems quite natural to choose the template as the pivotal point for the linearization toward the hypertemplate \((X_0, f_0)\) and the observations \((X^i, f^i)_{1 \leq i \leq N}\) that is to freeze the functional metric on the \(L^2(X)\) space.

The derivation is then quite straightforward and we jump directly to the associated minimization problem which is the minimization of the functional

\[
J^{\text{tan}}((v^0, \zeta^0), (v^i, \zeta^i)_{i}) = \frac{\gamma V_0}{2} \int_0^1 |v^0_t|^2_{V_0} \, dt + \frac{\gamma f}{2} |\zeta^0|^2_X + \sum_{i=1}^N \left( \frac{\gamma V}{2} \int_0^1 |v^i_t|^2_{V} \, dt + \frac{\gamma f}{2} |\zeta^i|^2_X + \frac{\gamma g}{2} \left( \tilde{X}^i, \tilde{f}^i \right), \left( X^i, f^i \right) \right) \tag{48}
\]

in the variables

\[
\begin{align*}
(v^0, \zeta^0) & \in L^2([0, 1], V_0) \times L^2(X) \\
(v^i, \zeta^i) & \in L^2([0, 1], V) \times L^2(X), \quad 1 \leq i \leq N, \\
\text{with} & \\
X &= \phi_{v^0}^{-1} \cdot X_0
\end{align*}
\tag{49}
\]

where we have

\[
\begin{align*}
\tilde{f}^i &= (f + \zeta^i) \circ \left( \phi_{v^i}^{-1} \right), \quad 1 \leq i \leq N, \\
\text{and} & \\
f &= f_0 \circ \left( \phi_{v^0}^{-1} \right) + \zeta^0. \tag{50}
\end{align*}
\]

The signals \((\zeta^i)_{1 \leq i \leq N}\) will be named the residual functions of the atlas.

5 Existence Results for Fshape Atlases

5.1 Introduction

The existence of an atlas for a population of fshapes with fcurrents or fvarifold-based data attachment term does not follow from the same arguments than in the more well-known pure geometrical case. The optimization of a signal \(f\) on the geometrical template and for the more sophisticated model of additional residuals \(\zeta^i\) to match each observation \((X^i, f^i), 1 \leq i \leq N\), introduces new difficulties.

\[
\text{ Springer}
\]
The main one is that the fvarifold attachment term \( \| \mu(X, f_i) - \mu(X, f) \|_W^2 \) is not lower semicontinuous for the weak convergence in \( L^2(X) \) as a function of \( f \) for \( X \) fixed. This comes from the nonlinearities appearing in the various kernels defining the RKHS norm on \( W \). Yet, interestingly, it remains continuous for the weak convergence on the space of measures. We prove that minimizing measure sequences are tight and that existence results can be established under quite general assumption for the extended functional.

However, the existence of strong solution (i.e., a true fshape template \((X, f)\)) does not seem to be guaranteed in full generality but is true as soon as the regularization on the \( L^2 \) penalization on \( f \) [and on the residuals \((\zeta^i)_i\), when introduced] is strong enough. This condition, although imposed theoretically for the existence of a fshape solution, is not clearly required for practical fshape atlases that we have computed so far. We show also that the computed template function and residuals are smooth proving a regularization effect of fvarifold attachment terms. The result holds in both the tangential and metamorphosis setting for fshapes, which is summed up by Theorem 6 (tangential model) and Theorem 7 (metamorphosis model), whose proofs are the core issues of this section.

The most technical part of the proof is the existence of the fshape template \((X, f)\) on which we will focus in the first place. The proof we present follows basically the so-called direct method of geometric measure theory where we first show the existence of a minimizer in a larger space of varifolds and then show that this solution does indeed result from a true fshape.

### 5.2 Existence of the Template Fshape

We shall focus first on the tangential setting of 4.3. In this part, we will also consider the simpler situation where all variables \((v^i)_i\) and \((\zeta^i)_i\) in Eq. (48) are frozen and show that a minimum of the functional exists in the variables \(X\) and \(f\). We will start by assuming that, in addition, \(X\) is also a fixed bounded and finite volume \(d\)-dimensional rectifiable subset of \(E\).

#### 5.2.1 Existence with \(X\) Fixed

In this subsection, we sketch out the proof of the Proposition 7 below. All the technical details are given in “Appendix 2.”

**Proposition 7** Assume that \(W\) is continuously embedded in \(C^2_0(E \times G_d(E) \times \mathbb{R})\) that \(X\) and \((X^1)_{1 \leq i \leq N}\) are finite volume bounded \(d\)-dimensional rectifiable subsets and that \(f^i \in L^2(X^i)\) for \(1 \leq i \leq N\). Assume that \(\gamma_f/\gamma_W\) is large enough with \(\gamma_f, \gamma_W \geq 0\). Then,

\[
J_X(f) = \frac{\gamma_f}{2} \int_X |f(x)|^2 d\mathcal{H}^d(x) + \frac{\gamma_W}{2} \sum_{i=1}^N \| \mu(X^i, f^i) - \mu(X, f) \|_W^2
\]
achieves its minimum on $L^2(X)$ and any minimizer $f_*$ is such that $f_* \in L^\infty(X)$. Moreover, if $X$ is also a $C^p$ submanifold and $W \hookrightarrow C^0_0(E \times G_d(E) \times \mathbb{R})$ with $m \geq \max \{p, 2\}$, then $f_* \in C^{p-1}(X)$.

The proof proceeds in two steps. The first one is to show the existence of a minimizer in the larger space of $f$-varifolds: Namely, for $X$ fixed as previously, we consider the set $\mathcal{M}_X$ of all Borel finite measures $\nu$ on $E \times G_d(E) \times \mathbb{R}$ such that:

$$\int h(x, V) d\nu(x, V, f) = \int_X h(x, T_x X) dH^d(x). \quad (51)$$

for all continuous and bounded applications $h$ on $E \times G_d(E)$. These are $f$-varifolds that have a particular marginal on $E \times G_d(E)$. Note that any varifold $\mu(X, f)$ for $f$ a $L^2$ function on $X$ belongs to $\mathcal{M}_X$. In addition, $\mathcal{M}_X$ is a closed subset for the usual weak convergence of measures defined by $\nu_n \rightharpoonup \nu_\infty$ if for any $\omega \in C_b(E \times G_d(E) \times \mathbb{R})$, the space of continuous bounded functions, we have

$$\nu_n(\omega) \to \nu_\infty(\omega). \quad (52)$$

Then, we consider the extended functional $\tilde{J}$ defined on measures $\nu \in \mathcal{M}_X$ by

$$\tilde{J}(\nu) = \frac{\gamma_f}{2} \nu(|f|^2) + \frac{\gamma_W}{2} \sum_{i=1}^N \|\nu - \mu(X, f_i)\|_{W'}^2 \quad (53)$$

where $\nu(|f|^2)$ is the notation we shall use for $\int |f|^2 d\nu$. Then,

**Lemma 2** There exists $\nu_* \in \mathcal{M}_X$ that minimizes the functional $\tilde{J}$.

The second step consists in proving that $\nu_*$ can be actually expressed as a $f$-varifold associated with a true $f$-shape $(X, f_*)$ with $f_* \in L^\infty(X)$. We also have detailed the full proofs of Lemma 2 in “Appendix 2.”

### 5.2.2 Existence with Non-fixed $X$

We now consider the existence of a template when $X$ is no more fixed and is estimated as well. In this case, as we mentioned earlier, it is enough to introduce a RKHS Hilbert space $V_0$ continuously embedded in $C^2_0(E, E)$, an initial hypertemplate $X_0$ and consider also an optimization of the template $X$ in the orbit of $X_0$ under the action of $\phi^0 \in G_0$, the group of diffeomorphisms associated with $V_0$, i.e., $X = \phi^0 \cdot X_0$. To prove existence result, we will need to introduce a penalization depending on the distance between $X_0$ and $X$, i.e., on $d_{G_0}(\text{Id}, \phi^0)$. A typical functional would be, if $X_0$ is the orbit of $X_0$ under $G_0$ and $d_{X_0}(X, X') = \inf_{\phi \in G_0, \phi \cdot X = X'} d_{G_0}(\text{Id}, \phi)$ is the induced distance between two templates in the orbit,
\[ J_1(X, f) = \frac{\gamma V_0}{2} d\chi_0(X_0, X)^2 + \frac{\gamma f}{2} \int_X |f(x)|^2 d\mathcal{H}^d(x) + \frac{\gamma W}{2} \sum_{i=1}^{N} \|\mu(X_i, f_i) - \mu(X, f)\|_{W'}^2 \]  

(54)

Since a diffeomorphism \( \phi^0 \in G_0 \) is obtained as the flow of a time-varying vector field \( v^0 \in L^2([0, 1], V_0) \), we consider the minimization of the following functional:

\[ J_2(v^0, f) = \frac{\gamma V_0}{2} \int_0^1 |v_{1\mid V_0}^0|^2 dt + \frac{\gamma f}{2} \int_X |f(x)|^2 d\mathcal{H}^d(x) + \frac{\gamma W}{2} \sum_{i=1}^{N} \|\mu(X_i, f_i) - \mu(X, f)\|_{W'}^2 \]

for \( v^0 \in L^2([0, 1], V_0) \), \( f \in L^2(X) \) with \( X = \phi^0 \cdot X_0 \). However, it is more convenient for \( X = \phi^0 \cdot X_0 \) to consider the change of variable \( f \mapsto f_0 = f \circ \phi^0 \) from \( L^2(X) \to L^2(X_0) \) so that we keep working in a fixed space \( L^2(X_0) \). Hence, we end up with functional

\[ J_3(v^0, f_0) = \frac{\gamma V_0}{2} \int_0^1 |v_{1\mid V_0}^0|^2 dt + \frac{\gamma f}{2} \int_X |f_0(x)|^2 |d_x \phi^0|_{\gamma} |d\mathcal{H}^d(x)_{\gamma} = f_0 |f(x)|^2 d\mathcal{H}^d(x) \text{ for } f = f_0 \circ (\phi^0)^{-1} \in L^2(X) \]

\[ + \frac{\gamma W}{2} \sum_{i=1}^{N} \|\mu(X_i, f_i) - \phi^0 \cdot \mu(X_0, f_0)\|_{W'}^2 \]  

(55)

with \( f_0 \in L^2(X_0) \) and \( \phi \cdot \mu \in W' \) denoting for a varifold \( \mu \) its diffeomorphic transport by a \( C^1 \) diffeomorphism \( \phi \) defined by \( (\phi \cdot \mu)(\omega) = \int |d_x \phi|^\omega(\phi(x), d_x \phi(V), f) d\mu(x, V, f) \) for \( \omega \in W \). The existence of a minimizer \( (v_{*}^0, f_{*}^0) \) for \( J_3 \) gives immediately the existence of a minimizer \( (X_*, f_*) \) for \( J_1 \) with \( X_* = \phi_1^0 \cdot X_0 \), \( f_* = f_0 \circ (\phi_1^0)^{-1} \in L^2(X_*) \).

The existence result then becomes:

**Proposition 8** Assume that \( W \) is continuously embedded in \( C^2_0(E \times G_d(E) \times \mathbb{R}) \) that \( X_0 \) and \( (X^i)_{1 \leq i \leq N} \) are finite volume and bounded \( d \)-dimensional rectifiable subsets and that \( f^i \in L^2(X^i) \) for \( 1 \leq i \leq N \). Assume \( \gamma V_0 > 0 \) and \( \gamma_f / \gamma W \) is large enough with \( \gamma_f, \gamma_W \geq 0 \). Then,

- \( J_1 \) given by Eq. (54) achieves its minimum on \( \{(X, f) \mid X \in X_0, f \in L^2(X)\} \);
- any minimizer \( (X_*, f_*) \) is such that \( f_* \in L^\infty(X_*) \);
- if \( X_0 \) is also a \( C^p \) submanifold and \( W \hookrightarrow C^m_0(E \times G_d(E) \times \mathbb{R}) \) with \( m \geq \max\{p, 2\} \), \( f_* \in C^{p-1}(X_*) \).

The proof of Proposition 7 can in fact be easily adapted to this new situation. We can consider the formulation of Eq. (55) with \( J_3 \) a functional on the vector field \( v^0 \) and
the function $f \in L^2(X_0)$. With respect to $v^0$, thanks to the penalization in (55), we can restrict the search of a minimum on a closed ball $B_0$ of given radius $b$ in $L^2([0,1], V_0)$, which guarantees at the same time that the Jacobians $|d_x\phi^0_i|_{TV}$ are uniformly lower bounded. This closed ball is also compact for the weak topology in $L^2([0,1], V_0)$. In addition, it follows from Corollary 1 that $v^0 \mapsto \sum_{i=1}^N \|\mu_{(X^i,f^i)} - \phi^0_i\|_{W_{v^0}}^2$ is weakly continuous on $L^2([0,1], V_0)$ and it is also classical that $v^0 \mapsto \|v^0\|_{L^2([0,1], V_0)}$ is lower semicontinuous for the weak convergence topology. Therefore, for all fixed $f_0 \in L^2(X_0)$, $v^0 \mapsto J_3(v^0, f_0)$ is weakly lower semicontinuous on $L^2([0,1], V_0)$. It results that we obtain existence of a minimizing vector field $v^0$ and, reasoning as in the previous subsection and “Appendix 2,” one deduces easily the claim of Proposition 8.

5.3 Existence of Full Fshape Atlases (Tangential Setting)

We now generalize the previous results to the existence of complete atlases of fshepae’ datasets. In addition to the template $(X, f)$, one wants to simultaneously estimate transformations from the template to each subject. In the tangential model of Sect. 4.3, these consist in deformations $(\phi^i)_{1 \leq i \leq N}$ obtained as flows of time-varying vector fields $(v^i)_t$ and residual signals $(\zeta^i)_t$ that are $L^2$ functions on $X$. In that situation, following Sect. 4.3, the optimization functional for atlas estimation writes:

$$J(X, f, (\zeta^i)_i, (v^i)_i) = \frac{\gamma V}{2} d_X(X_0, X)^2 + \frac{\gamma f}{2} \int_X |f(x)|^2 d\mathcal{H}^d(x)$$

$$+ \sum_{i=1}^N \left( \frac{\gamma V}{2} \int_0^1 \left| v^i_t \right|^2 dt + \frac{\gamma \zeta}{2} \int_X \left| \zeta^i(x) \right|^2 d\mathcal{H}^d(x) \right)$$

$$+ \frac{\gamma W}{2} \left\| \mu_{(X^i,f^i)} - \mu_{(\phi^i_t(X),(f+\zeta)_{\circ(\phi^i)^{-1}})} \right\|_{W_{v^0}}^2$$

(56)

The main existence result is the following:

**Theorem 6** Assume that $W$ is continuously embedded in $C^0_0(E \times G_d(E) \times \mathbb{R})$, that $X_0$ and $(X_i)_{1 \leq i \leq N}$ are finite volume bounded $d$-dimensional rectifiable subsets and that $f^i \in L^2(X^i)$ for $1 \leq i \leq N$. Assume $\gamma V, \gamma V > 0$ and $\gamma f / \gamma W$ and $\gamma \zeta / \gamma W$ are large enough with $\gamma f, \gamma W, \gamma \zeta \geq 0$. Then

- $J$ given by Eq. (56) achieves its minimum on $\{(X, f, (\zeta^i)_i, (v^i)_i) \mid X \in X_0, f \in L^2(X), \zeta \in L^2(X)^N, (v^i)_t \in L^2([0,1], V)^N\}$;
- any minimizer $(X_*, f_*, (\zeta^*_i)_i, (v^*_i)_i)$ is such that $f_*$ and $\zeta^i$ for $1 \leq i \leq N$ are in $L^\infty(X_*)$;
- if $X_0$ is also a $C^p$ submanifold and $W \hookrightarrow C^m_0(E \times G_d(E) \times \mathbb{R})$ with $m \geq \max\{p,2\}$, $f_*$ and the $\zeta_*$’s are in $C^{p-1}(X_*)$.

The proof leans essentially on the same arguments as detailed in the previous subsections: We refer the reader to “Appendix 3.” This result is of fundamental importance for the rest of the paper since it ensures that the atlas estimation problems that we shall
study numerically in the next sections do have at least a solution, and we see that this holds with only $L^2$ regularity assumptions on the signals.

5.4 Existence in the Metamorphosis Framework

Interestingly, the previous existence of solutions in the simplified tangential setting can be also used to show existence of solutions to the corresponding variational problem in the metamorphosis framework given by (43). In that model, the subjects are obtained approximately as metamorphoses of the template $(X, f)$, which is itself a metamorphosis of an hypertemplate $(X_0, f_0)$ where $X_0$ is a bounded finite volume rectifiable subset of $E$ and $f_0 \in L^2(X_0)$. The important lemma that bridges both approaches is the following:

Lemma 3 For $v$ fixed and $\zeta_1 \in L^2(X)$ fixed, the infimum over $h \in L^2([0, 1], L^2(X))$ of

$$\gamma f \int_0^1 \int_X |h_t(x)|^2 |dx| \phi_t^v \mid_{\Gamma x} \mid d\mathcal{H}^d(x) dt$$

under the constraint that $\int_0^1 h_t \mid dt = \zeta_1$ is reached on a unique point $h^* \in L^2([0, 1], L^2(X))$ given by

$$h_t^* \mid x = C(x) \frac{\zeta_1(x)}{|dx| \phi_t^v \mid_{\Gamma x}}$$

where $C(x) = (\int_0^1 \frac{1}{|dx| \phi_t^v \mid_{\Gamma x}} ds)^{-1}$.

Proof The proof could be deduced from general optimization results, but for sake of completeness we give here a short proof. Let us first notice that over $(x, t) \in X \times [0, 1]$ we have $0 < \inf |dx| \phi_t^v \mid_{\Gamma x} \leq \sup |dx| \phi_t^v \mid_{\Gamma x} < +\infty$.

Now, if we denote $\alpha_t(x) = |dx| \phi_t^v \mid_{\Gamma x}^{-1/2}$ and consider the change of variable $g_t(x) = h_t(x)/\alpha_t(x)$, the problem is equivalent (after the use of Fubini-Tonelli Theorem) to the minimization on $g \in L^2([0, 1], L^2(X))$ of

$$\gamma f \int_X \left(\int_0^1 |g_t(x)|^2 dt\right) d\mathcal{H}^d(x)$$

under the constraint $\int_0^1 \alpha_t(x) g_t(x) ds = \zeta_1(x)$ for a.e. $x \in X$.

However, this is a separable problem, and for any fixed $x \in X$, we have to consider on $H \doteq L^2([0, 1])$ the very simple quadratic minimization problem: $\inf |\mu|_H^2$ under the constraint $(\alpha(x), u)_H = \zeta_1(x)$ that has the unique solution $u = \lambda(x) \alpha(x)$ with

$$\lambda(x) = \zeta_1(x)/|\alpha(x)|_H^2.$$

Hence, denoting $h_t^*(x) = \zeta_1(x) |\alpha_t(x)|^2 |\alpha(x)|_H^2$, we get the result. □
Now, introducing, as in 2.2, the Riemannian energies

\[ E_{X_0}(v^0, h^0) = \frac{\gamma v_0}{2} \int_0^1 |v_r|^2 v_0 \, dt + \frac{\gamma f_0}{2} \int_0^1 \int_X |h^0_r|^2 (x) |d_x \phi^0_r|_{T_x X} \, dH^d(x) \, dt \]

and

\[ E_X(v, h) = \frac{\gamma v}{2} \int_0^1 |v_r|^2 \, dt + \frac{\gamma f}{2} \int_0^1 \int_X |h_r|^2 (x) |d_x \phi_r|_{T_x X} \, dH^d(x) \, dt \]

the atlas estimation functional becomes:

\[
J \left( \left( v^0, h^0 \right), \left( v^i, h^i \right) \right) = E_{X_0} \left( v^0, h^0 \right) + \sum_{i=1}^N \left( E_X \left( v^i, h^i \right) + \frac{\gamma_W}{2} \left\| \mu(X^i, f^i) - \mu \left( \phi^0_i(X), \left( f + \xi^0_i \right) \circ (\phi^0_1)^{-1} \right) \right\|_{W^i}^2 \right)
\]

(58)

where \((X, f) = (\phi^0_1(X_0), (f_0 + \xi^0_i) \circ (\phi^0_1)^{-1})\).

**Theorem 7** Let \(W\) be continuously embedded in \(C^2_0(E \times G_d(E) \times \mathbb{R})\), \(X_0\) and \((X^i)_{1 \leq i \leq N}\) finite volume bounded \(d\)-dimensional rectifiable subsets and \(f^i \in L^2(X^i)\) for \(1 \leq i \leq N\). Assume that \(\gamma v_0, \gamma_V > 0\) and that \(\gamma_f / \gamma_W\) and \(\gamma_\xi / \gamma_W\) are large enough with \(\gamma_f, \gamma_W, \gamma_\xi \geq 0\). Under these assumptions,

- \(J\) defined by Eq. (58) achieves its minimum on \(v^0 \in L^2([0, 1], V_0), \ h^0 \in L^2([0, 1], L^2(X_0)), \ (v^i)_i \in L^2([0, 1], V)^N\) and \((h^i)_i \in L^2([0, 1], L^2(X))^N\);
- any minimizer \((X_*, f_*, (\xi^i)_i), (v^i)_i)\) with \(f_* = (f_0 + \xi^0_1) \circ (\phi^0_1)^{-1}\) and \(\xi^i = \xi^i_1 \circ (\phi^0_1)^{-1}\) is such that \(f_*\) and \(\xi^i\) for \(1 \leq i \leq N\) are in \(L^2(X_*)\);
- if \(X_0\) is also a \(C^p\) submanifold, \(f_0\) a \(C^{p-1}\) function on \(X\) and \(W \rightarrow C^m_0(E \times G_d(E) \times \mathbb{R})\) with \(m \geq \max\{p, 2\}\), then \(f_*\) and the \(\xi^i\)'s are in \(C^{p-1}(X_*)\).

The proof can be found in “Appendix 4.”

**Part II Numerical Implementation**

**6 A Discrete Framework for Fshape**

**6.1 Introduction**

The previous sections mainly dealt with the setting of a well-posed theoretical framework to work with fshapes. Here, we would like to tackle the problem of the actual processing of fshapes. We believe that the description of an explicit numerical scheme is of the utmost importance for the fshape setting to be of practical utility. Our main...
motivations come from medical imaging where acquisition is done in digital forms. It means that data are discrete and in very high dimensions (tens of thousands of points). Our goal is to demonstrate that the fshape framework may be used to handle real data and not just low dimensional examples. We provide an implementation of our atlas estimation method called fshapetk. It is written in matlab, C, and Cuda, and it is available at the GIT repository https://github.com/fshapes/fshapesTk with a documentation [9] and examples scripts.

6.2 Fvarifold Norm of Polyhedral Meshes

We assume hereafter that data at hand are finite functional polyhedral meshes composed by $T$ cells of dimension $d$ immersed in $E = \mathbb{R}^n$. Most common examples are piecewise linear planar curves corresponding to the case $n = 2$ and $d = 1$, piecewise linear curves in space corresponding to $n = 3$ and $d = 1$, and piecewise triangular surfaces in the space corresponding to $n = 3$ and $d = 2$. A finite polyhedral fshape $(X, f)$ is fully described by three matrices: A first $P \times n$ matrix $x$ contains the coordinates of the $P$ vertices $x_k$ in $E$ (i.e., $x = (x_k)_{1 \leq k \leq P}$), a second $P \times 1$ column vector $f$ contains the $P$ values $f_k = f(x_k) \in \mathbb{R}$ of the signal (i.e., $f = (f_k)_{1 \leq k \leq P}$), and a third $T \times (d + 1)$ matrix contains the list of edges for each cell so that each row contains $(d + 1)$ integers corresponding to the indices of the vertices of the cell.

In the sequel, it will be convenient to introduce the corresponding geometrical diffeomorphic action in the discrete setting: For a diffeomorphism $\phi \in G$ and a discretized fshape $(x, f) = (x_k, f_k)_{1 \leq k \leq P}$, we denote

$$\phi \cdot (x, f) = (\phi \cdot x, \phi \cdot f)$$

where $\phi \cdot x = \phi(x_k)$ (Lagrangian point of view). Note also that the connectivity matrix (i.e., list of edges) remains unchanged through the deformation process.

There is a rather natural way to approximately represent polyhedral meshes in the fvarifold space. The method is very similar to the one described in [12] in the case of fcurrents space. Using formula (33), each cell is coded by a Dirac $\delta(\hat{x}_\ell, V, \hat{f}_\ell)$ so that a discrete polyhedral fshape is viewed as a distribution of Dirac spread in the space $\mathbb{R}^n$. Therefore, the measure $\mu(X, f)$ of Eq. (32) which is associated with $(X, f)$ is approximated by the following finite sum of Dirac

$$\mu(X, f) \approx \mu(x, f) = \sum_{\ell=1}^{T} r_\ell \delta(\hat{x}_\ell, V_\ell, \hat{f}_\ell)$$

where $\hat{x}_\ell \in \mathbb{R}^n$ is the center of the $\ell$-th polyhedral cell, $V_\ell \in G_d(\mathbb{R}^n)$ gives the direction of the cell, $\hat{f}_\ell \in \mathbb{R}$ is the average value of the signal on this cell, and finally, $r_\ell \in \mathbb{R}$ is equal to the $d$-volume of the cell. Note that for the case of surfaces (resp. curves), $V_\ell$ will be simply a unit normal vector (resp. unit tangent) and the orientation will be removed directly “by the kernel”: The formula of the kernel $k_t$ used to compute
the scalar product in the fvarifold space will be invariant with respect to changes in
the orientation as in Eq. (62).

As an illustration, let us precisely describe how to get the representation of the \( \ell \)-th
triangle \( T_\ell \) belonging to a polyhedral functional surface \( (X, f) \). If we assume that
the vertices of \( T_\ell \) (read in the connectivity matrix) are \((x_{k_1}, x_{k_2}, x_{k_3}) \in (\mathbb{R}^3)^3 \) for some
\( k_1, k_2, k_3 = 1, \ldots, P \), we have

\[
\begin{align*}
\hat{x}_\ell &= \frac{x_{k_1} + x_{k_2} + x_{k_3}}{3}, \\
V_\ell &= \frac{(x_{k_2} - x_{k_1}) \wedge (x_{k_3} - x_{k_1})}{\| (x_{k_2} - x_{k_1}) \wedge (x_{k_3} - x_{k_1}) \|}, \\
\hat{f}_\ell &= \frac{f_{k_1} + f_{k_2} + f_{k_3}}{3}, \\
r_\ell &= \frac{\| (x_{k_2} - x_{k_1}) \wedge (x_{k_3} - x_{k_1}) \|}{3},
\end{align*}
\]

(61)

where \( \| \cdot \| \) is the standard Euclidean norm in \( \mathbb{R}^3 \) and \( \wedge \) corresponds to the cross
product. To compute the fvarifold norm of a functional surface, we use in our numerical
experiments Gaussian kernels satisfying assumptions of Proposition 2. Let us define:

\[
\begin{align*}
k_e(\hat{x}_1, \hat{x}_2) &= \exp \left( -\frac{\| \hat{x}_1 - \hat{x}_2 \|^2}{\sigma_e^2} \right), \\
k_t(V_1, V_2) &= \exp \left( -\frac{2}{\sigma_t^2} \left( 1 - \langle V_1, V_2 \rangle \right) \right), \\
k_f(\hat{f}_1, \hat{f}_2) &= \exp \left( -\frac{\| \hat{f}_1 - \hat{f}_2 \|^2}{\sigma_f^2} \right).
\end{align*}
\]

(62)

where \( \sigma_e, \sigma_t, \sigma_f > 0 \) are scale parameters fixed by the practitioner. Thence the scalar
product of two fvarifolds \( \mu_{(X, f)} \) and \( \mu_{(Y, g)} \) associated with two polyhedral
triangular functional surfaces \( (X, f) \) and \( (Y, g) \) and approximated, respectively, by
\( \sum_{\ell=1}^{T_x} r_\ell \delta_{\hat{x}_\ell, V_\ell, \hat{f}_\ell} \) and \( \sum_{\ell'=1}^{T_y} s_{\ell'} \delta_{\hat{y}_{\ell'}, W_{\ell'}, \hat{g}_{\ell'}} \), may be written:

\[
\left[ \mu_{(X, f)}, \mu_{(Y, g)} \right]_{W'} \approx \left[ \mu_{(X, f)}, \mu_{(Y, g)} \right]_{W'} = \sum_{\ell=1}^{T_x} \sum_{\ell'=1}^{T_y} r_\ell s_{\ell'} k_e(\hat{x}_\ell, \hat{y}_{\ell'}) k_t(V_\ell, W_{\ell'}) k_f(\hat{f}_\ell, \hat{g}_{\ell'}). \quad (63)
\]

We discuss in Sect. 9.2 some numerical aspects of the fvarifolds.

6.3 Discrete Approximations of Continuous Fshapes

A continuous fshape \((X_c, f_c)\) can be approximated by a finite polyhedral fshape in the
fvarifold space as follows. First, extract \( P \) points and their corresponding signal from
\((X_c, f_c)\). Then, compute a mesh with the extracted points to define a polyhedral fshape
\((X, f)\). Finally, use (60) to approximate \( \mu_{(X, f)} \) by a finite sum \( \mu_{(X, f)} \) of Diracs.

Although this method of approximation seems reasonable in many practical cases,
we do not provide explicit conditions to ensure the convergence of \( \mu_{(X, f)} \) toward
\( \mu_{(X_c, f_c)} \) when the number \( P \) of extracted vertices tends to infinity. To the best of
our knowledge, this is still an open problem and a famous illustration of pathological
cases is the Schwarz polyhedron. Nevertheless, there exist various theoretical results concerning the polyhedral approximation of continuous surfaces for the currents’ norms and the varifold norms; see, e.g., [30]. Moreover, in [32], the author studies sufficient conditions to ensure the convergence of the area of triangular meshes toward the area of the continuous surface. A general result extending the aforementioned works to our geometric-functional framework involves some materials beyond the scope of this paper.

6.4 Shooting Equations for the Metamorphosis Riemannian Framework

In the large deformation setting for usual shapes, it was shown (see, for instance, [27]) that the dynamics of optimal vector fields giving geodesics in groups of diffeomorphisms can be described through Hamiltonian systems of equations, called the forward equations. In the discrete situation of a finite set of points, these actually reduce to a coupled evolution of the position of the particles and extra variables called the momenta attached to every particle. In addition, the Hamiltonian structure implies that all geodesic trajectories are eventually parametrized only by the initial positions and momenta, which is the principle underlying geodesic shooting algorithms for shape matching (see [3] for an extensive presentation of the Hamiltonian setting for shape deformation analysis).

We now describe how to obtain similar shooting frameworks in our more general situation of fshapes, first for the fshape metamorphosis model. As discussed above, any element \((X, f)\) in \(\mathcal{F}\) will be discretized as a family \((x, f) = (x_k, f_k)_{1 \leq k \leq P}\) of points \(x\) with signal value \(f\). In that setting, for a proper weighting matrix (the precise definition of this matrix depends on actual choices for the approximation of the \(L^2\) norm), we can discretize the \(L^2(x)\) dot product as \((D(x)h)\cdot h)\) for any \(h = (h_k) \in \mathbb{R}^P\), where we use the notation \((D(x)h)\cdot h)\) to denote the matrix product \(h^T D(x)h\). The continuous problem (3) is then approximated by a simple discrete control problem

\[
\min \frac{\gamma_N}{2} \int_0^1 |v(t)|^2 \, dt + \frac{\gamma_f}{2} \int_0^1 (D(x(t))h(t)\cdot h(t)) \, dt
\]

for fixed end point conditions \((x(t), f(t))|_{t=0} = (x_0, f_0)\), \((x(t), f(t))|_{t=1} = (x_1, f_1)\) and controlled dynamics in \(V \times \mathbb{R}^P\) given by

\[
\begin{cases}
\dot{x}_t = v_t \cdot x_t \\
\dot{f}_t = h_t
\end{cases}
\]

where \(v \cdot x = (v(x_k))_{1 \leq k \leq P}\).

Remark 1 An important point is to note that the discrete evolution is based on a Lagrangian particle-based representation and to understand where Eq. 65) is coming from. In the discrete setting, \(t \mapsto x_t = (x_{t,k})_{1 \leq k \leq P}\) is the evolution the vertices of a polyhedral mesh; i.e., \(x_{t,k} = \phi^t_v(x_k)\), where \(\phi^t_v\) is the flow of the time dependent vector field \(t \mapsto v_t\) and \(x = (x_k)\) are the vertices on the initial mesh \(X\).
so that \( \dot{x}_{t,k} = v_t(\phi_t^v(x_k)) = v_t(x_{t,k}) \) which gives the first equation of (65). Moreover, \( f_i = (f_{i,k})_{1 \leq k \leq P} \) are the signal values attached to \( X_t \) at the vertices positions \( x_t = (x_{t,k})_{1 \leq k \leq P} \) so that \( f_{i,k} \) is the discretization of the continuous signal \( f_i : X_t \to \mathbb{R} \) in Lagrangian coordinates; i.e., \( f_{i,k} = f_t(\phi_t^v(x_k)) = f_t(x_{t,k}) \). In particular, one has \( \dot{f}_{i,k} = \frac{d f_t(\phi_t^v(x_k))}{dt} = \frac{d f_t(h_t(x_k))}{dt} = h_t(x_k) = h_{t,k} \) where \( h_t = (h_{t,k}) = (h_t(x_k)) \) is the discretization of \( h_t : X \to \mathbb{R} \) on the vertices \( (x_k) \) which gives the second equation of (65).

In that situation, introducing co-states or *momenta* \((p, p^f) = (p_k, p_k^f)_{1 \leq k \leq P}\) we get the associated Hamiltonian

\[
H \left( (x, f), \left( p, p^f \right), (v, h) \right) = (p|v \cdot x) + \left( p^f | h \right) - \frac{\gamma V}{2} |v|^2 - \frac{\gamma f}{2} \left( D(x) h | h \right)
\]

so that we deduce from the Pontryagin’s maximum principle that the optimal controls satisfy

\[
v(\cdot) = \frac{1}{\gamma V} \sum_{k=1}^{P} K_V(\cdot, x_k) p_k \quad \text{and} \quad h = \frac{1}{\gamma f} D^{-1}(x) p^f
\]

where \( K_V \) is the kernel associated with the RKHS space \( V \). Then, plugging the optimal control into the Hamiltonian \( H \) we get the reduced Hamiltonian given as

\[
H_f \left( (x, f), \left( p, p^f \right) \right) = \frac{1}{2 \gamma V} (K_{x,x} p | p) + \frac{1}{2 \gamma f} \left( D^{-1}(x) p^f | p^f \right)
\]

with \( K_{x,x} = (K_V(x_k, x_{k'}))_{1 \leq k, k' \leq P} \). Note that the weighting matrix \( D(x) \) may be chosen to be diagonal (“mass lumping”) so that the computation of \( D^{-1}(x) \) is straightforward. From \( H_f \), we can derive the forward equation given by the Hamiltonian dynamics:

\[
\begin{cases}
\dot{x} = \partial_p H_f(x, p) \\
\dot{p} = -\partial_x H_f(x, p).
\end{cases}
\]

As usual, since the Hamiltonian is not depending on time, it is a conserved quantity during the geodesic evolution so that we get from (66) and (67) that

\[
\frac{\gamma V}{2} \int_0^1 |v_t|^2 V dt + \frac{\gamma f}{2} \int_0^1 (D(x_t) h_t | h_t) dt = H_f \left( (x, f), \left( p, p^f \right) \right) \big|_{t=0}.
\]

We see also that the reduced Hamiltonian is a perturbation of the more familiar Hamiltonian for the pure geometrical case mentioned earlier, \( H_{f}^{\text{geo}}(x, p) = \frac{1}{2 \gamma V} (K_{x,x} p | p) = \frac{\gamma V}{2} |v|^2 V \) with a geometric-functional term

\[
H_{f}^{\text{geo-fun}}(x, f, p^f) = \frac{1}{2 \gamma f} \left( D^{-1}(x) p^f | p^f \right).
\]
Hence, the shooting equations contain a new source term which is $\partial_x H^\text{geo–fun}_t$ in the dynamics of the momenta $p$:

$$\dot{p} = -\partial_x H^\text{geo}(x, p) - \partial_p H^\text{geo–fun}_t(x, f, p^f)$$

(71)

and since the $H^\text{geo–fun}_t$ does not depend on $f$, we have

$$\dot{p^f} = -\partial_f H^\text{geo–fun}_t(x, f, p^f) = 0.$$  (72)

In particular, the functional speed $f^*_k$ is only modulated by the evolution of the local weight $D(x_0)$ which depends on the continuous limit both on the divergence of the “tangential” part of $v$ to the manifold spanned by the $x_k$’s and on its normal part when the mean curvature is non-vanishing.

### 6.5 Shooting Equations for the Tangential Setting

In the tangential model of Sect. 2.5, equations are simplified once again. Indeed, the $L^2$ metric being frozen to the initial position of $X$, the second term in (64) becomes

$$\frac{\gamma_f}{2} \int_0^1 (D(x_0) h_t | h_t) dt$$

where $D(x_0)$ is now a constant weighting matrix and the optimal functions $h_t$ for fixed time 1 end point condition $f_1$ are simply given by $h_t \equiv \xi = f_1 - f$, i.e., $f_t = f + t\xi = t f_1 + (1-t) f$ where $\xi = (\xi(x_k))_{1 \leq k \leq P}$ and $\xi$ is the residual introduced in Sect. 4.3. In that situation, the new source term disappears in (71) and the dynamics of $(x, p)$ is described by the usual Hamiltonian equations in the purely geometrical case, which gives:

$$\begin{cases}
\dot{x} = \partial_p H^\text{geo}_t(x, p) \\
\dot{p} = -\partial_x H^\text{geo}_t(x, p) \\
\dot{f}_t = f + t\xi
\end{cases}$$

(73)

These are the forward equations in the tangential model. Compared to metamorphoses, the dynamics on the signal part can be expressed in closed form, while the evolution in spatial position is described by the usual Hamiltonian system without extra term. This makes the tangential model particularly simple for implementation as the two dynamics are separated (it shall be exploited in Sect. 7). Optimizing over trajectories can be reduced to optimization with respect to initial momentum $p_0$ and signal $f_1$ (or the residual $\xi$) since again $v(\cdot) = \frac{1}{\gamma_V} \sum_{k=1}^P K_V(\cdot, x_k) p_k$, $h = \xi$ and

$$\frac{\gamma_V}{2} \int_0^1 |v_t|_V^2 dt + \frac{\gamma_f}{2} \int_0^1 (D(x_0) h_t | h_t) dt = H^\text{geo}_t((x, p))_{t=0} + \frac{\gamma_f}{2} (D(x_0)\xi | \xi).$$

(74)
Variation of functionals with respect to $p_0$ can be obtained by the usual backward integration of the adjoint Hamiltonian system (cf [10] Chapter 1 or [3] Section 4).

7 Algorithms to Compute Mean Template of Fshapes

We present two different algorithms to compute a mean template from a sample of $N$ discrete fshapes. Both methods consist in solving a variational problem via an adaptive gradient descent algorithm presented in Sect. 7.3. As the optimization method is the same in both cases, we just need to describe how to compute the functional and its gradient. These methods are implemented in MATLAB, Cuda and C, and some numerical experiments are shown Sects. 8 and 9.

7.1 Hypertemplate and Tangential Setting

Algorithms 1, 2 and 5 give a way to implement the atlas estimation in the tangential setting described in Sect. 4.3. The principle of the method is illustrated by Fig. 3. The method consists in computing a reduced version based on geodesic shooting and initial momenta of the functional $J_{\text{tan}}$ of Eq. (48) with Algorithm 1 and its gradient with Algorithm 2 and then plugging this in the optimization box given by Algorithm 5.

More precisely, the inputs are $N$ observations $(x^i, f^i)_i$, a hypertemplate $(x^0, f^0)$, the momenta $p_0$ and $(p^i)_i$ and the functional residuals $\zeta_0$ and $(\zeta^i)_i$. All the fshapes should be provided with their corresponding connectivity matrix. The optimization is conducted on the reduced version of $J_{\text{tan}}$ defined as

\[
J_{\text{HT}}^{\text{tan}}(p_0, \zeta_0, (p^i, \zeta^i)_i) \equiv \gamma V_0 \frac{1}{2} |v_{p_0}|^2 V_0 + \gamma_0 \frac{1}{2} |\zeta_0|^2_x + \sum_{i=1}^N \left( \gamma V_0 \frac{1}{2} |v_{p^i}|^2 V + \gamma f_0 \frac{1}{2} |\zeta^i|^2_x + \gamma f_w \frac{1}{2} g \left( \left( \tilde{x}^i, \tilde{f}^i \right), (x^i, f^i) \right) \right) \]

\[(75)\]

Fig. 3 Hypertemplate method: the mean template $(x, f) = \left( \phi^{x_0, f_0} \cdot x_0, f_0 + \zeta_0 \right)$ is a deformed version of the hypertemplate $(x_0, f_0)$.
in the variables

\[
\begin{align*}
(p_0, \xi_0) & \in \mathbb{R}^P \\
(p^i, \xi^i) & \in \mathbb{R}^P, \quad 1 \leq i \leq N,
\end{align*}
\]

with

\[
(x, f) = \left( \phi_1^{p_0} \cdot x_0, f_0 + \xi_0 \right),
\]

\[
(\tilde{x}^i, \tilde{f}^i) = \left( \phi_1^{p^i} \cdot x + \xi^i \right), \quad 1 \leq i \leq N,
\]

where \( \|\ell\|_X^2 = (D(x)\ell|\ell) \) is the discretization of the \( L^2(X) \) norm for \( \ell = (\ell_k)_{1 \leq k \leq P} \in \mathbb{R}^P \) defined on the vertices \( x \in \mathbb{R}^P \) as introduced in Sect. 6.4.

---

**Algorithm 1** Computation of the energy \( J_{HT}^{\text{tan}} \) with formula (75)

1. **Require:** A hypertemplate \( (x_0, f_0) \) and \( N \) fshapes \( (x^i, f^i) \).
2. **Inputs:** \( N + 1 \) momenta \( p_0, (p^i)_i \) and \( N + 1 \) functional residuals \( \xi_0, (\xi^i)_i \).
3. **Begin**
4. **Deformation:** Compute \( x \) by forward integration of \( (x_0, p_0) \) and let \( f \leftarrow f_0 + \xi_0 \).
5. **for** \( i = 1 \) to \( N \) **do**
6. **Deformation:** Compute \( \tilde{x}^i \) by forward integration of \( (x, p^i) \) and let \( \tilde{f}^i \leftarrow f + \xi^i \).
7. **Fvarifold norm:** Compute the fvarifold representation \( \mu(\tilde{x}^i, \tilde{f}^i) \) and \( \mu(x^i, f^i) \); compute \( g_i \leftarrow ||\mu(\tilde{x}^i, \tilde{f}^i) - \mu(x^i, f^i)||^2_{W^*} \).
8. **Penalty terms:** Compute \( \|v^{p^i}\|^2_V \) and \( ||\xi^i||_X^2 \).
9. **end for**
10. **Penalty terms:** Compute \( \|v^{p_0}\|^2_V \) and \( ||\xi_0||_X^2 \).
11. **End**
12. **Outputs:** \( \sqrt{\gamma_0} \cdot \|v^{p_0}\|^2_V + \sqrt{\gamma_0} \cdot ||\xi_0||_X^2 + \sum_i \sqrt{\gamma^i} \cdot ||\xi^i||_X^2 + \sqrt{\gamma^i} \cdot \|v^{p^i}\|^2_V + \sqrt{\gamma^i} \cdot g_i \).

The hypertemplate may be difficult to define, and the choice of this initial guess may obviously affect the quality of the atlas estimation. Nevertheless, the outputs of the algorithms are found to be stable even for the non-trivial experiments shown in Fig. 8. We recommend generating a simple and smooth fshape with a meshing program and choose \( f_0 \) to be a constant. This hypertemplate should be topologically equivalent to the observations as much as possible. Using one of the observations as the hypertemplate could be tempting, but it may induce some bias in the estimated atlas as the final mean template may contain the same specific features as the chosen observation (recall that the template is a diffeomorphic deformation of the hypertemplate).

A bar sign over a variable means that this is the optimal value found by the algorithm. Then, at the end of the minimization procedure, the outputs are: an estimation of the mean template \( (\bar{x}, \bar{f}) = (\phi_1^{p_0} \cdot x_0, f_0 + \bar{\xi}_0) \), and for any \( 1 \leq i \leq N \) a momentum \( \bar{p}^i \) and a functional residual \( \bar{\xi}^i \) so that \( (\phi_1^{p^i} \cdot \bar{x}, \bar{f} + \bar{\xi}^i) \) is close to \( (x^i, f^i) \).
Algorithm 2 Computation of the gradient $\nabla J_{\text{HT}}^{\tan}$

1: Require: A hypertemplate $(x_0, f_0)$ and $N$ fspheres $(x^i, f^i)_i$.
2: Inputs: $N + 1$ momenta $p_0, (p^i)_i$ and $N + 1$ functional residuals $\xi_0, (\xi^i)_i$.
3: Begin
4: \hspace{1em} Deformation: Compute $x$ by forward integration of $(x_0, p_0)$ and let $f \leftarrow f_0 + \xi_0$.
5: \hspace{1em} for $i = 1$ to $N$ do
6: \hspace{2em} Deformation: Compute $\tilde{x}^i$ by forward integration of $(x, p^i)$ and let $\tilde{f}^i \leftarrow f + \xi^i$.
7: \hspace{2em} Gradient of $g_i$ wrt $\xi_0, \xi^i, p^i$ and $x$: Compute directly $\nabla_{\xi_0} g_i$ and $\nabla_{\xi^i} g_i$; compute $(\nabla_x g_i, \nabla_{p^i} g_i)$ by backward integration of $\nabla_{\xi^i} g_i$.
8: \hspace{1em} end for
9: \hspace{1em} End
10: Outputs: $\nabla_{\xi_0} J_{\text{HT}}^{\tan} = \frac{\gamma}{2} \nabla_{\xi_0} |\xi_0|_2^2 + \frac{\gamma_f}{2} \sum_i \nabla_{\xi_0} g_i$; $\nabla_{\xi^i} J_{\text{HT}}^{\tan} = \frac{\gamma}{2} \nabla_{\xi^i} |\xi^i|_2^2 + \frac{\gamma_f}{2} \nabla_{\xi^i} g_i$ for $1 \leq i \leq N$.

Algorithm 3 Computation of the energy $J_{\text{free}}^{\tan}$ of formula (77)

1: Require: $N$ fspheres $(x^i, f^i)_i$.
2: Inputs: A template fsphere $(x, f)$, $N$ momenta $(p^i)_i$ and $N$ functional residuals $(\xi^i)_i$.
3: Begin
4: \hspace{1em} for $i = 1$ to $N$ do
5: \hspace{2em} Deformation: Compute $\tilde{x}^i$ by forward integration of $(x, p^i)$ and let $\tilde{f}^i \leftarrow f + \xi^i$.
6: \hspace{2em} Fvarifold norm: Compute the fvarifold representation $\mu(\tilde{x}^i, \tilde{f}^i)$ and $\mu(x^i, f^i)$; compute $g_i \leftarrow \|\mu(\tilde{x}^i, \tilde{f}^i) - \mu(x^i, f^i)\|_W^2$.
7: \hspace{2em} Penalty terms: Compute $|v^p|^2_{L^2}$ and $|\xi^i|_2^2$.
8: \hspace{1em} end for
9: \hspace{1em} Penalty term: Compute $|f - f_0|_x^2$.
10: End
11: Outputs: $\frac{\gamma_f}{2} |f - f_0|_x^2 + \sum_i \frac{\gamma_f}{2} |v^p|^2_{L^2} + \frac{\gamma_f}{2} |\xi^i|_2^2 + \frac{\gamma_w}{2} g_i$.

7.2 “Free” Mean Fshape and Tangential Model

Algorithms 3, 4 and 5 use a more direct approach to compute a mean fshape compared to the hypertemplate method of Sect. 7.1. The general framework is similar, but the mean template is no longer defined as a deformed version of an hypertemplate. The coordinates of the points composing the mean template will be directly updated along the gradient descent. Notice that there is no extra cost as, in Algorithm 2, we compute the gradient $\nabla_x g_i$ in the backward integration step 7. Some properties of this gradient are discussed Sect. 9.2. The aim of this method is simply to use this gradient directly.
Algorithm 4 Computation of the gradient $\nabla J_{\text{free}}^{\text{tan}}$

1: **Require:** $N$ fshapes $(x^i, f^i)$.
2: **Inputs:** A template $(x, f)$, $N$ momenta $(p^i)$, and $N$ functional residuals $(\zeta^i)$.
3: **Begin**
4: for $i = 1$ to $N$
5: \hspace{1em} **Deformation:** Compute $\tilde{x}^i$ by forward integration of $(x, p^i)$ and let $\tilde{f}^i \leftarrow f + \zeta^i$.
6: \hspace{1em} **Gradient of $g_i$ wrt $f, x^i, p^i$ and $x$:** Compute directly $\nabla_f g_i$ and $\nabla_{\tilde{x}^i} g_i$; compute $(\nabla_{x^i} g_i, \nabla_p g_i)$ by backward integration of $(\nabla_{\tilde{x}^i} g_i, 0)$.
7: \hspace{1em} **Gradient of penalty terms:** Compute directly $\nabla_x |v| p_i^2 \frac{\partial f}{\partial v} \nabla_{\tilde{x}^i} |\tilde{x}^i|^2$ and $\nabla_{\tilde{x}^i} |\tilde{x}^i|^2$.
8: \hspace{1em} **end for**
9: **Gradient of penalty term:** Compute directly $\nabla_f (f - f_0)^2 + \nabla_x (f - f_0)^2$.
10: **End**
11: **Outputs:** $\nabla_x J_{\text{free}}^{\text{tan}} = \frac{\gamma_f}{2} \nabla f + f_0 \frac{\partial f}{\partial v} + \nabla_x (f - f_0)^2$.
12: \hspace{1em} $\nabla_x J_{\text{free}}^{\text{tan}} = \frac{\gamma_f}{2} \nabla f + f_0 \frac{\partial f}{\partial v} + \nabla_x (f - f_0)^2$.
13: \hspace{1em} for $1 \leq i \leq N$; $\nabla_{\tilde{x}^i} J_{\text{free}}^{\text{tan}} = \frac{\gamma_f}{2} \nabla_{\tilde{x}^i} |\tilde{x}^i|^2 + \nabla_{\tilde{x}^i} \nabla_{\tilde{x}^i} g_i$ for $1 \leq i \leq N$.

Algorithm 5 Adaptive gradient descent

1: **Require:** Coefficients used to adapt the step sizes $0 < s^- < 1 < s^+$, A functional $J$ and its gradient $\nabla J = (\nabla u_1, \cdots, \nabla u_k)$. 
2: **Inputs:** Points $(u_1^{\text{init}}, \cdots, u_k^{\text{init}})$ and step sizes $(\delta_1, \cdots, \delta_k)$. 
3: **Begin**
4: \hspace{1em} **Initialization:** $(u_1, \cdots, u_k) \leftarrow (u_1^{\text{init}}, \cdots, u_k^{\text{init}})$ and $(\delta_1, \cdots, \delta_k) \leftarrow (\delta_1^{\text{init}}, \cdots, \delta_k^{\text{init}})$.
5: **repeat**
6: \hspace{1em} **Evaluations:** Compute $J_{\text{cur}}^{\text{new}} \leftarrow J(u_1, \cdots, u_k)$ and for $1 \leq k \leq K$ compute $\nabla u_k J_{\text{cur}}^{\text{new}}$.
7: \hspace{1em} **Update all the variables simultaneously:** Compute $J_{\text{new}}^{\text{new}} \leftarrow J(u_1^{\text{new}}, \cdots, u_k^{\text{new}})$ where $u_k^{\text{new}} \leftarrow u_k + \delta_k^{\text{new}} \nabla u_k J_{\text{cur}}^{\text{new}}$ for $1 \leq k \leq K$.
8: **Adapt steps:** if $J_{\text{new}} < J_{\text{cur}}^{\text{new}}$ then $s \leftarrow s^+$ else $s \leftarrow s^-$ **end if**
9: **Stopping condition:** Declare a boolean, breakLoop $\leftarrow$ False.
10: **repeat**
11: \hspace{1em} **Update each variable separately:** For $1 \leq k \leq K$ compute $J_{\text{new}}^{\text{new}} \leftarrow J(u_1^{\text{new}}, \cdots, u_k^{\text{new}}, \cdots, u_K^{\text{new}})$ where $u_k^{\text{new}} \leftarrow u_k + s \delta_k \nabla u_k J_{\text{cur}}^{\text{new}}$.
12: \hspace{1em} **Update every variable at the same time:** $J_{\text{new}}^{\text{new}} \leftarrow J(u_1^{\text{new}}, \cdots, u_k^{\text{new}})$.
13: **Keep the best configuration:** Let $(u_1^*, \cdots, u_k^*)$ and $(\delta_1^*, \cdots, \delta_k^*)$ be so that $(u_1^{\text{new}}, \cdots, u_k^{\text{new}}) = (u_1^*, \cdots, u_k^*)$ and $\delta_k \nabla u_k J_{\text{cur}}^{\text{new}}$ satisfies $J(u_1^*, \cdots, u_k^*) = \min(J_{\text{new}}^{\text{new}}, \cdots, J_{\text{new}}^{\text{new}}, J_{\text{new}}^{\text{new}})$.
14: \hspace{1em} if $J(u_1^*, \cdots, u_k^*) < J_{\text{cur}}^{\text{new}}$ then $J_{\text{new}}^{\text{new}} \leftarrow J(u_1^*, \cdots, u_k^*)$.
15: \hspace{1em} **Update points and steps sizes:** $(u_1, \cdots, u_k) \leftarrow (u_1^*, \cdots, u_k^*)$ and $(\delta_1, \cdots, \delta_k) \leftarrow (\delta_1^*, \cdots, \delta_k^*)$.
16: breakLoop $\leftarrow$ True.
17: **else**
18: **Decrease all steps sizes:** $(\delta_1, \cdots, \delta_k) \leftarrow (s \delta_1, \cdots, s \delta_k)$.
19: **end if**
20: **until** “breakLoop = True” or “steps sizes are too small.”
21: **until** “Maximum iteration is reached” or “descent is too small” or “steps sizes are too small.”
22: **End**
23: **Output:** Points $(u_1, \cdots, u_k)$ and steps sizes $(\delta_1, \cdots, \delta_k)$.
Fig. 4 “Free” mean template method: The mean template \((x, f)\) is a gradient descent-based update of an initial fshape \((x_0, f_0)\). The notation \((x_r, f_r)\) denotes the state of the mean template at iteration \(r\) to update the mean fshape, thereby making the evolution “free” along the optimization process. Figure 4 gives an illustration of the method.

The free mean fshape method is a gradient descent on the following functional:

\[
J^\text{tan} \left( (x, f), \left( p^i, \xi^i \right) \right) = \gamma_0^2 \| f - f_0 \|^2_x + \sum_{i=1}^{N} \left( \frac{\gamma_V}{2} \left\| v^i \right\|^2_V + \frac{\gamma_f}{2} \left\| \xi^i \right\|^2_x + \frac{\gamma_W}{2} g \left( (\tilde{x}^i, \tilde{f}^i), (x^i, f^i) \right) \right) \tag{77}
\]

in the variables

\[
\begin{align*}
(x, f) &\in E^P \times \mathbb{R}^P, \\
(p^i, \xi^i) &\in E^P \times \mathbb{R}^P, \quad 1 \leq i \leq N, \\
\text{with} \\
(\tilde{x}^i, \tilde{f}^i) &\equiv \left( \phi^{p^i}_1 \cdot x, f + \xi^i \right), \quad 1 \leq i \leq N.
\end{align*} \tag{78}
\]

This algorithm does not have yet the same theoretical warranties as those given by Theorem 6 for the hypertemplate version. We are not able to show, with the same arguments, that there exists proper minimizers of the continuous counterpart of (77) as the fshape \((X, f)\) is not constrained to belong to a single fshape bundle. However, we use in practice a regularized gradient descent to keep the evolution of the template smooth along the minimization. Therefore, if the size of the updates are small enough, the output of the free mean fshape algorithm may be considered as a discretization of a smooth deformation of \((X_0, f_0)\) as Property 4.2.1 of [10] shows. This trick also prevents odd behaviors of the algorithm as discussed in Sect. 9.2. Compared to the hypertemplate algorithm of previous section, this numerical scheme still reduces the constraint imposed on the template evolution which is likely to provide less dependency in the choice of initialization.
The inputs are \(N\) observations \(((x^i, f^i))_i\), an initial fshape \(((x_0, f_0))\), the momenta \(((p^i))_i\) and the functional residuals \(((\xi^i))_i\). All the fshapes should be provided with their respective connectivity matrix. The momenta \(((p^i))_i\) are usually initialized to 0 and \(((\xi^i))_i\) are usually initialized to a constant.

A bar sign over a variable means that this is the optimal value found by the algorithm. At the end of the minimization procedure, the outputs are: an estimation of the mean template \(((\bar{x}, \bar{f}))\), and for any \(1 \leq i \leq N\) the momentum \(\hat{p}^i\) and the functional residual \(\hat{\xi}^i\) so that \((\phi v^\top \hat{p}^i \cdot \bar{x}, \bar{f} + \hat{\xi}^i)\) is close to \((x^i, f^i)\).

### 7.3 Optimization

An atlas estimation is a smooth, non-convex and high-dimensional global optimization problem. We use an adaptive gradient descent algorithm to solve it numerically. This method is popular in geometric atlas estimation as it is quite robust and performs relatively well on various real examples; see [14,17]. In our geometric-functional framework, one main issue comes from the fact that the functionals to minimize depend on various types of variables living in different spaces: momenta, functional values or points coordinates. Adding the functional part in the optimization process makes it harder, compared to a pure geometrical approach, as some non-desirable phenomena may hold. For a thorough discussion, see Sect. 9.1.

Let us assume hereafter that we want to minimize a functional \(J(u_1, \ldots, u_K)\) depending on \(K \geq 1\) types of variables. For instance, in an atlas estimation of fshapes with a hypertemplate (see Sect. 7.1), we have \(J = J_{\text{tan}}^{\text{HT}}\) and \(K = 4\) (namely, the variables \(\xi_0, p_0, (\xi^i)_i\) and \((p^i)_i\)).

If the \(K\) types of variables are non-homogeneous, we note that in practice, the different types of variables should not be updated at the same speed along the optimization process. It means that the gradient is not the proper descent direction to follow in order to reach a reasonable solution. As Algorithm 5 shows, we use \(K\) different steps sizes denoted \((\delta_1, \ldots, \delta_K)\) and we adapt them separately. The quality of the final result will depend on a fine balance between the \(\delta_1, \ldots, \delta_K\).

Some tricks may be also used to improve the optimization strategy. For instance, it may be convenient to regularize some part of the gradient as discussed in Sect. 9.

Another efficient method is to change the \(f\)varifold kernel widths \(\sigma_e, \sigma_f, \sigma_t\) along the optimization process. In that case, the optimization involves several runs: A first gradient descent is performed at coarse scale by choosing large \(\sigma_e, \sigma_f, \sigma_t\), then a second gradient descent is performed with smaller \(\sigma_e, \sigma_f, \sigma_t\) (with starting point equals to the end point of the first run) and so on. In [9] and in the various commented examples in the \texttt{fshapesTk} software, we provide some practical hints to help people choose the scales \(\sigma_e, \sigma_f, \sigma_t\).

### 8 Numerical Experiments

#### 8.1 Synthetic Dataset

We now present a first set of results of the previous atlas estimation algorithms on a synthetic dataset of \(N = 6\) textured statues constructed using \texttt{Sculptris} software.
Fig. 5 Template evolution along the gradient descent steps of the algorithm. On the left is the prototype fshape which serves as the initialization for the template. We use a multiscale approach that decreases, after a certain number of iterations, the characteristic sizes of the kernels on geometry and signal.

The dataset as well as the code to run this experiment (“venus” example) are available in fshapesTk. The set of subjects is shown in the last row of Fig. 7. The atlas is estimated using the free mean fshape tangential evolution algorithm of Sect. 7.2 after the dataset is preregistered with respect to translations. The template is initialized with a prototype fshape with zero-valued signal everywhere in order to avoid as much as possible bias toward one particular individual. Figure 5 shows the current template at several intermediate steps of the estimation. We use all Gaussians for the kernels $k_e$, $k_t$ and $k_f$ defining the data attachment terms as in Eq. (62). The template is obtained by refining twice the scales $\sigma_e$ and $\sigma_f$ throughout the process in order to have a coarse estimation of the atlas in the first place before being able to retrieve finer details from the dataset. The resulting template captures the average shape and signal behavior of the set of subjects: In particular, the algorithm tends to recover the signal patterns that are the most shared among the population, as the uniform coloring of the head for instance.

In addition to the template, we obtain, at convergence, a set of residual signals as well as deformation momenta that map the template on the different subjects both in terms of shape and texture. The kernel for the space $V$ of deformation fields is taken as a sum of two Gaussian kernels, following the approach of [7] which allows to introduce multiscale deformations. We show two of such mappings in Fig. 6. The whole set of shape and texture matchings compared to the original subjects is finally summed up in Fig. 7. In terms of numerics, the subjects of the dataset are unequally sampled between 2800 and 6500 points, while the template of Fig. 5 has 7000 points. Using our GPU implementation for kernel computations, each iteration of the algorithm takes approximately 38 secs for a total atlas estimation time of 3 h (300 iterations) on a server equipped with a Nvidia GTX 555.

The dependency in the initialization of the template is also an important issue. Section 9.3 explains how to make the previous procedures independent of rescaling or sampling changes of the fshapes. But obviously, in both the hypertemplate and “free” template evolution algorithm, the shape of the template is still constrained to live in the diffeomorphic orbit of the initial one. It results that one cannot expect to remove completely the bias resulting from the initialization’s choice. In Fig. 8, we show the template obtained after convergence for different initializations. Although
Fig. 6 Mapping of the template on subjects 2 and 5. From left to right, we display the template and several intermediate time steps of the geometric-functional transformations, i.e., $\phi_t \cdot \bar{x}, \bar{f} + \bar{\xi}_t^i$ where $\bar{\xi}_t^i = t \bar{\xi}^i$ as in formula (73).

Fig. 7 Estimated template and the geometric-functional matchings to each subject
these results do demonstrate some variations for the estimated template, they still show a quite remarkable stability in most of the important geometric and functional features of the dataset.

8.2 OCT Dataset

We now present a template estimation performed on a real dataset. The dataset contains surfaces segmented from volumetric images of the retina acquired by optical coherence tomography (OCT) as described in [20,21]. The aim of these measurements is to detect early glaucoma by analyzing changes of conformations of the inner layers of the retina around the optic disk (where the optic nerve slots into the retina). Data are fshapes: The surfaces represent the lower boundary of the nerves fiber layer (NFL), and the signals represent the thickness of the NFL. A loss of thickness may be an indicator of glaucoma. We depict below two typical observations in two different views: Data 1 is a normal subject in Fig. 10b, c, and data 2 is a glaucoma subject in Fig. 11b, c. The “view 1” has the same scale for the 3 axis, and “view 2” is a flipped version of “view 1” where the depth axis has been scaled (×4) to better represent fine reliefs. The typical size of square boundary ranges from 5 to 7 mm and the signal ranges from 0 to 0.3 mm. The color scale for the signal is the same for all the pictures of this section. The overall geometry of the NFL boundary does not seem to be challenging, but these surfaces contain many boundaries making the problem hard to handle in practice. We discuss in Sect. 9.2 how we manage problems arising from boundary effects. Moreover, the difficulty is increased by the fact that some observations are not
Hypertemplate and the estimated mean (computed with the full dataset containing 51 observations).

(a) Relative positions of the hypertemplate and data 1 and 2 (Figs. 10, 11) (view 1).
(b) Mean template \((\bar{x}, \bar{f})\) (view 1).
(c) Mean template \((\bar{x}, \bar{f})\) (view 2).

Fig. 10 Results for data 1 (control dataset). Figure 10e, f should be compared with Fig. 10b, c, respectively.

(a) Relative position of data 1 and the mean template (view 1).
(b) Data 1 (view 1).
(c) Data 1 (view 2).
(d) Mean template and the residuals \(\bar{x}, \bar{f} + \bar{\xi}^1\) (view 1).
(e) Deformed mean template and residuals \(\phi^1 \cdot \bar{x}, \bar{f} + \bar{\xi}^1\) (view 1).
(f) Deformed mean template and residuals \(\phi^1 \cdot \bar{x}, \bar{f} + \bar{\xi}^1\) (view 2).

centered and the opening may be closed to the boundary of the acquisition area as in Fig. 11.

We ran our code, using the hypertemplate method described Sect. 7.1 on a dataset containing \(N = 51\) observations (19 normals, 25 glaucoma and 7 suspects). The observations were preregistered by hand with respect to translations so that the center of the opening is at the origin. We use downsampled data (raw data contains more than 130,000 points and 270,000 triangles each), and each observation contains about 5000 points and 9850 triangles. The hypertemplate is the flat rectangle with a hole and with a null signal depicted in Fig. 9a. It contains 5700 points and 11,100 triangles. Computations of the mean template \((\bar{x}, \bar{f})\) (Fig. 9b, c) and the deformations took 7 h.
Fig. 11  Results for data 2 (glaucoma dataset). Figure 11e, f should be compared with Fig. 11b, c, respectively. a Relative position of data 2 and the mean template (view 1). b Data 2 (view 1). c Data 2 (view 2). d Mean template and the residuals $\left( \bar{x}, \bar{f} + \xi^2 \right)$ (view 1). e Deformed mean template and residuals $\left( \phi_v \bar{p} \cdot \bar{x}, \bar{f} + \xi^2 \right)$ (view 1). f Deformed mean template and residuals $\left( \phi_v \bar{p} \cdot \bar{x}, \bar{f} + \xi^2 \right)$ (view 2). (120 iterations) using a server equipped with a Nvidia GTX 555 graphical processor unit.

The estimated template is depicted in Fig. 9b, c. The average shape is a bended version of the flat hypertemplate which is a reasonable guess. The baseline signal $f$ captures a light but typical “C”-shaped signal. This blurred signal is due to the rather high variability of signals across the dataset. The results of the deformations and functional residuals estimations for data 1 and 2 are given by Figs. 10 and 11, respectively. The functional part of all the 51 observations is well reconstructed, and we are now able to compare these signals as they are all defined on the mean template; see Figs. 10d, and 11d. The deformations are also satisfying although most of the energy of the deformations is spent to match the (non-informative) outer boundaries as data are misaligned, comparing Figs. 10a and 11a.

9 Numerical Pitfalls

In this section, we discuss some issues affecting the quality of the mean template estimation. We illustrate these problems by numerical experiments and propose some solutions to fix them.
9.1 Mass Cancelation

The varifold norm was first introduced in the pure geometrical setting to avoid mass cancelation appearing with the classical current norms. As orientation matters in the currents’ setting, a surface with positive area sufficiently crumpled may have an arbitrarily small current norm. This phenomenon is common in practical applications and particularly during an atlas estimation. This is described in detail in the introduction of [12] or in the Chapter 3 of [10]. In the (pure geometrical) varifold setting, the norm of a surface cannot decrease too much when two pieces of surface are folded: Theorem 3.4.1 in [10] is in some sense a reciprocal inequality to Proposition 5. Unfortunately, there is no such a control for the varifold norm as one can exhibit a functional surface of positive area and nonzero signal with small varifold norm.

We follow here the notations introduced in Sect. 6. Recall that, in formula (63), the inner product of two varifolds was approximated by a double sum of inner products of Diracs. We assume hereafter that the kernel varifold inner product satisfies the assumptions of Proposition 2. Therefore, by formula (36) we have, for any Diracs $\delta(x_1,V_1,f_1)$ and $\delta(x_2,V_2,f_2)$:

$$\left|\langle\delta(x_1,V_1,f_1),\delta(x_2,V_2,f_2)\rangle_{\mathcal{W}}\right| \leq |k_e(\cdot,\cdot)|_{\infty}|k_t(\cdot,\cdot)|_{\infty}|f_1-f_2|.$$  \hspace{1cm} (79)

The right-hand side of (79) may be arbitrarily small when $|f_1-f_2|$ is large as $k_f(f_1,\cdot)$ is continuous and vanishes at infinity. It means, in particular, that sufficiently high variations in the signal make any two Diracs orthogonal with respect to the varifold inner product. Now, let $(x, f)$ be a finite polygonal mesh approximated by a finite sum of Diracs $\sum_{\ell=1}^{T} r_{\ell}\delta(\hat{x}_{\ell},V_{\ell},\hat{f}_{\ell})$. Equations (37) and (79) yield

$$\|\mu(x,f)\|_{\mathcal{W}}^2 \leq |k_e(\cdot,\cdot)|_{\infty}|k_t(\cdot,\cdot)|_{\infty} \times \left(\sum_{\ell=1}^{T} r_{\ell}^2 k_f\left(\hat{f}_{\ell},\hat{f}_{\ell}\right) + \sum_{\ell,\ell'=1,\ell\neq\ell'}^{T} r_{\ell} r_{\ell'} k_f\left(\hat{f}_{\ell},\hat{f}_{\ell'}\right)\right).$$ \hspace{1cm} (80)

From here on, we consider the case where $k_e$, $k_t$ and $k_f$ are the Gaussian kernels of Eq. (62). Let $0 < \varepsilon < 1$ and assume that $\hat{f}_{\ell} = K\sigma_{f}\ell$ for any $\ell = 1, \ldots, T$ and $K = \sqrt{-\ln \varepsilon}$. Inequality (80) becomes, in that case:

$$\|\mu(x,f)\|_{\mathcal{W}}^2 \leq \sum_{\ell=1}^{T} r_{\ell}^2 + \sum_{\ell,\ell'=1,\ell\neq\ell'}^{T} r_{\ell} r_{\ell'} \exp\left(-K^2|\ell - \ell'|^2\right)$$

$$\leq C_1(T) + \varepsilon C_2(T),$$

where $C_1(T) = \sum_{\ell=1}^{T} r_{\ell}^2$ and $C_2(T) = \sum_{\ell,\ell'=1,\ell\neq\ell'}^{T} r_{\ell} r_{\ell'}$. Thence, if $T$ remains fixed and $K \rightarrow \infty$ the varifold norm of $(x, f)$ is less of equal to $C_1(T)$. Assuming that the fsshape $(x, f)$ is regularly discretized with $P$ points, we may consider that $r_{\ell} \approx P^{-1}$ and $C_1(T) \approx P^{-1}$. When the discretization of a fsshape becomes finer, the number $T$ of triangles increases and the varifold norm of $(x, f)$ may be arbitrarily small.
To illustrate this, we consider the discrete version of the variational problem of Proposition 7 in the simple case where \( N = 1 \) and \( \gamma_W = 2 \). It gives:

\[
\begin{align*}
\min_{f \in \mathbb{R}^{P_x}} J_{\text{fun}}(f) \\
\text{where} \\
J_{\text{fun}}(f) = \frac{\gamma_f}{2} |f|_x^2 + \|\mu(x, f) - \mu(y, g)\|_W^2
\end{align*}
\]

where the source fshape \((x, f) \in E_{P_x} \times \mathbb{R}_{P_x}\) and the target fshape \((y, g) \in E_{P_y} \times \mathbb{R}_{P_y}\) are two flat overlapping squares belonging to the same plane; see Fig. 12. The source \((x, f)\) contains \(P_x = 6400\) vertices and the target \((y, g)\) contains \(P_y = 400\) vertices distinct from the vertices of \(x\). The meaning of (81) is the following: We are trying to register \((x, f)\) onto the fixed target \((y, g)\) by tuning the signal \(f \in \mathbb{R}^{P_x}\) of the source only. This is not obvious to figure out what a good solution of this problem should be. Note that Proposition 7 ensures the existence of a proper solution of the continuous version of (81) if \(\gamma_f\) is large enough.

In Fig. 13, we show three outputs of an adaptive gradient descent on \(J_{\text{fun}}\) with various choices of penalty parameter \(\gamma_f\). We denote \(f^* \in \mathbb{R}^{P_x}\) the solution found after 800 iterations. Figure 13a shows the result with no penalty term; i.e., \(\gamma_f = 0\). The signal \(f^*\) is almost equals to 1 on the overlapping square, whereas \(f^*_s\) oscillates dramatically on the complementary part to decrease the fvarifold norm of the remaining part of \(x\). Figure 13b shows that with a small penalty, we are able to recover a signal close to 1 on the central square but oscillations are still present though weaker compared to Fig. 13a. Finally, Fig. 13c shows the solution \(f^*_s\) found with a larger \(\gamma_f\). The oscillations have almost disappeared, and the signal remains small in the non-overlapping part (in blue). The price to pay is a lower intensity in the central square (around 0.25).

Note that this example is an extreme case, but such a problem arises in real datasets. In medical imaging, it is common to get functional data acquired on a subpart of a surface only. Even with a rigid registration during the preprocessing, subparts with a well-defined signal may be highly non-overlapping on two different subjects. This may yield to oscillating solutions if the gradient descent is used naively. We note that in practice, regularizing the gradient with respect to \(f\) may fix this issue in some cases.
Fig. 13 The fshape $(x, f_{\ast})$ where $f_{\ast}$ is the output of a gradient descent on $J_{\text{fun}}$ of Eq. (81) with various value of $\gamma_f$. a $\gamma_f = 0$. b $\gamma_f = 4$. c $\gamma_f = 20$

Fig. 14 Initial fshape $(x_{\text{init}}, 0)$ is a flat square with a centered opening. The target fshape $(y, 0)$ is a (non smooth) bended version of $(x_{\text{init}}, 0)$. All the signals are zero. a Initialization $(x_{\text{init}}, 0)$. b Target $(y, 0)$

The method is similar to the one described at the end of Sect. 9.2 for the geometric part of the gradient.

9.2 Boundary Problems

In this section, we discuss some problems arising when the “free” mean template method of Sect. 7.2 is used with fshapes containing boundaries. An issue concerns the gradient (with respect to $x$) of the data attachment terms $g_i$ (see Algorithm 3)—note that the following discussion is relevant for both the purely geometric and functional varifolds or currents framework. The values of the signal has no influence here, and we assume without loss of generality that the signals are 0. Typical examples of data with boundaries (inspired by the dataset presented in Sect. 8.2) are depicted in Fig. 14.

To illustrate the problem that one can face, let $E = \mathbb{R}^3$ and $(y, 0) \in E^{P_y} \times \mathbb{R}^{P_y}$ be a fixed discrete target fshape as in Fig. 14b. We consider the following minimization problem,

$$\min_{x \in E^P} J_{\text{geo}}(x)$$

where

$$J_{\text{geo}}(x) = \|\mu(x,0) - \mu(y,0)\|_{W^{1,p,y}}^2,$$

$$\text{(82)}$$
Our goal is to study the behavior of a gradient descent in the variable \( x = (x_k)_{k=1}^P \in E^P \) on \( J_{\geo} \) starting from an initial discrete fshape \((x_{\text{init}}, 0)\). In numerical experiments, the gradient descent on \( J_{\geo} \) does not perform well if the fshape \((x_{\text{init}}, 0)\) contains boundaries as in Fig. 14a. The boundary of a discrete fshape is the set of vertices that are an end of an edge belonging to a single triangle (see Fig. 15).

The main issue on the gradient of \( J_{\geo} \) is the following. The norm of \( \frac{\partial}{\partial x_k} J_{\geo} \) has different orders of magnitude depending on the location of the points \( x_k \) in the fshape: Gradient of boundary points may be much larger than gradient of interior points; see Fig. 16a. These unbalanced values between interior points and boundary points in the gradient induce undesirable effects such as self-crossings or changes in topology during the gradient descent as in Fig. 16b. Note also that the situation becomes worse when the number \( P \) of points in \( x \) increases.

In order to explain the singularities of \( \nabla_x J_{\geo} \) at the boundary points, notice first that this is sufficient to focus on the gradient of the fvarifold inner product as we have \( \nabla_x \| \mu(x, f) - \mu(y, g) \|^2_{W'} = \nabla_x \left( \langle \mu(x, f), \mu(x, f) \rangle_{W'} - 2 \langle \mu(x, f), \mu(y, g) \rangle_{W'} \right) \). Nevertheless, it is hard to clearly understand this phenomenon by looking at the discrete formulas directly. It is then convenient to use the results of Sect. 3.3.2 whose aim is to analyze the variations of the fvarifold inner product in the continuous setting. Thence, we may consider that \((x, 0)\) (resp. \((y, 0)\)) is sampled from a continuous fshape \((X, 0)\) (resp. \((Y, 0)\)) and that \( \langle \mu(x, 0), \mu(y, 0) \rangle_{W'} \) approximates its continuous counter-
Regularized gradient descent on $J_{\text{geo}}$ of Eq. (82) starting from $(x_{\text{init}}, 0)$ of Fig. 14a. a The arrows represent the initial regularized gradient $\tilde{\nabla}_x J_{\text{geo}}(x_{\text{init}})$ on $(x_{\text{init}}, 0)$. The color represents the norm of the gradient. b Output of the regularized gradient descent after few steps (has to be compared with 14b) (Color figure online)

part $\langle \mu(X, 0), \mu(Y, 0) \rangle_{W'}$. We further assume that $(x, 0)$ is regularly discretized, so that the typical size of the interior volume elements $r_\ell$ is of order $P^{-1}$ and the typical size of the boundary volume elements is of order $P^{-d-1}$. Formula (40) gives an expression of the variations of $\langle \mu(X, 0), \mu(Y, 0) \rangle_{W'}$ when $(X, 0)$ moves. The right-hand side of (40) is composed by two separated integral terms whose domains of integration are $X$ and its boundary $\partial X$, respectively. This formula then makes clear the differences between interior points and boundary points. A discrete approximation of the integral over the interior of $X$ should then involve terms of order $P^{-1}$ (i.e., the volume of the surface element), whereas a discrete approximation of the integral over the boundary $\partial X$ should involve terms of order $P^{-d-1}$ (corresponding to the size of the volume element of the boundary). This heuristic explains why there exists a multiplicative factor between the size of the gradient at interior points and boundary points. We also know that this factor is of order $P^{d}$ meaning that the difference increases when the resolution of the discrete fshape $(x, 0)$ increases.

Thus, the gradient $\nabla_x J_{\text{geo}}$ computed thanks to the Euclidean metric involves singular terms on the boundary. A convenient way to address this problem is to compute a gradient via a more regular metric. Let us introduce a (conveniently normalized) Gaussian kernel $K_{\text{reg}}: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ of scale $\sigma_{\text{reg}}$. Let $K_{x,x}^{\text{reg}} = (K_{\text{reg}}(x_k, x_{k'}))_{1 \leq k, k' \leq P} \in \mathbb{R}^{P \times P}$ and define

$$\tilde{\nabla}_x J_{\text{geo}}(x) = K_{x,x}^{\text{reg}} \nabla_x J_{\text{geo}}(x)$$

This new gradient is, in some sense, a convolution of the singular gradient with a Gaussian kernel; see Fig. 17a. As the kernel $K_{\text{reg}}$ is positive definite, $-\tilde{\nabla}_x J_{\text{geo}}$ still defines a descent direction for $J_{\text{geo}}$. Figure 17b shows the output of a regularized gradient descent. This new gradient can be seen as the discretization of a dense vector field belonging to the RKHS $V_{\text{reg}}$ of kernel $K_{\text{reg}}$. The displacement of the template from its initial position in the regularized gradient descent can be then considered.

Fig. 17 Regularized gradient descent on $J_{\text{geo}}$ of Eq. (82) starting from $(x_{\text{init}}, 0)$ of Fig. 14a. a The arrows represent the initial regularized gradient $\tilde{\nabla}_x J_{\text{geo}}(x_{\text{init}})$ on $(x_{\text{init}}, 0)$. The color represents the norm of the gradient. b Output of the regularized gradient descent after few steps (has to be compared with 14b) (Color figure online)
as the integration of a sequence of vector fields living in $V_{\text{reg}}$. It is shown in [10] (Property 4.2.1) that the regularized gradient flow is well defined. Thence, provided that the steps along the gradient descent are sufficiently small, the evolution of the points of $x$ may be considered as a diffeomorphic evolution (belonging to the orbit $G_{V_{\text{reg}}}(x_{\text{init}}, 0)$). Note that, in general, $\sigma_{\text{reg}}$ can be chosen to be different from the scale $\sigma_V$ of the deformations. This allows more flexibility in the template evolution during the gradient descent and justifies the methodology used in the algorithm described in Sect. 7.2.

9.3 Tuning the Parameters

The procedure of atlas estimation for fshapes contains numbers of parameters which have to be tuned by the user. Although there exist some heuristics to choose the values of these parameters, it may be hard to find a good range of settings for a particular dataset. One of our goals is to provide an algorithm behaving in a similar way whatever the scale and the resolution of the data are. In particular, the user should be allowed to tune the parameters at coarse resolution (when computation times are low) and perform a matching or an atlas estimation at high resolution with only few updates for the parameters values. Even if these refinements are not fundamental in theory, it makes the algorithm usable in practice and it has the great advantage of normalizing the choice of the balance parameters $\gamma_0, \gamma_f^0, \gamma_V, \gamma_f, \gamma_W$ between the various penalization terms and the data attachment term, as well as the initial step sizes in Algorithm 5 (these quantities may vary considerably otherwise).

In this section, $J$ denotes either the functional $J_{\text{tan}}^{\text{HT}}$ of Eq. (75) (Algorithms 1, 2) or $J_{\text{tan}}^{\text{free}}$ of Eq. (77) (Algorithms 3, 4).

9.3.1 Scale Invariance

We first describe how we normalize the functional $J$ and its gradient in order to provide an algorithm with outputs that are invariant if a scaling is applied to the geometry or to the signal of the fshapes. To do so, we carefully examine the effect of a scaling (in the geometrical or in the signal space) on each term composing $J$. Without this normalization step, the balance between the different terms of $J$ would be modified by scale which is not desirable in practice.

Let us examine the behavior of the functional when we apply a scaling on the geometrical space $x \mapsto x' = \lambda_e x \in E^P$ and on the signal $f \mapsto f' = \lambda_f f \in \mathbb{R}^P$ for some $\lambda_e, \lambda_f > 0$. We assume that the scale parameters $\sigma_e$ and $\sigma_f$ of the Gaussian kernels $k_e$ and $k_f$ are also rescaled as they measure the size at which the fshapes are compared. These new kernels are denoted $k_{e}'$ and $k_{f}'$ and they verify $k_{e}'(x', y') = k_e(x, y)$ and $k_{f}'(f', g') = k_f(f, g)$. The scaled local volume elements $r_{e}'$ now satisfy $r_{e}' = \lambda_e^d r_e$ and the scaled momenta $p' = \lambda_e p$ so that the displacement generated by $p'$ is also scaled. All the terms composing $J$ are normalized in the same way, and we just detail the case of the fvarifold data attachment term $g_i$. By formula (63), we easily see that the scaled data attachment term is proportional to $\lambda_e^{2d}$ and does not depend on the scale of the signal, namely
\[ g'_i \left( \left( x', f' \right), \left( x^{i'}, f^{i'} \right) \right) = \lambda_c^{2d} g_i \left( \left( x, f \right), \left( x^i, f^i \right) \right). \]

Therefore, we may compute a normalized version \((R_c)^{-2d} g_i\) of \(g_i\) with \(R_c \doteq \max_i \left( \sqrt{\text{trace}(\text{Var}(x^i))} \right)\) where \(\text{Var}(x^i)\) is the covariance matrix of the points cloud \(x^i\). From there on, \(J_{\text{norm}}\) denotes the functional containing the normalized terms of \(J\) (i.e., each term being multiplied by an appropriate power of \(R_c\) and an appropriate power of \(R_f \doteq \max_i \sqrt{\text{Var}(f^i)}\)). By construction, we (formally) have

\[ J_{\text{norm}}(x, p, f) = J'_{\text{norm}}(x', p', f') \]

where \(J'_{\text{norm}}\) is computed for \(\sigma'_e = \lambda_e \sigma_e\) and \(\sigma'_f = \lambda_f \sigma_f\).

We now examine the gradient’s behavior under scaling. The goal is to keep a similar dynamic during the gradient descent whatever the (geometric or functional) scales of the fshapes are. As an illustration, consider the gradient with respect to \(x\): We need \(J_{\text{norm}}(x - \delta \nabla_x J_{\text{norm}}, p, f) = J'_{\text{norm}}(x' - \delta \nabla_x' J'_{\text{norm}}, p', f')\) which yields to

\[ \nabla_x' J'_{\text{norm}} = \lambda_e \nabla_x J_{\text{norm}}. \]

As \(\nabla_x' g'_i = \lambda_c^{-1} \nabla_x g_i\), we use the normalized version \((R_c)^{2-2d} \nabla_x g_i\) of the gradient \(\nabla_x g_i\). Each term composing the gradient \(\nabla J_{\text{norm}}\) is then normalized with an appropriate power of \(R_c\) and \(R_f\), and we denote \(\nabla_n J_{\text{norm}}\) the normalized version \(\nabla J_{\text{norm}}\).

In our implementation of Algorithm 1 and 3 (resp. 2 and 4), we compute \(J_{\text{norm}}\) (resp. \(\nabla_n J_{\text{norm}}\)) rather than \(J\) (resp. \(\nabla J\)). These normalizations then guarantee that using similar but rescaled (or translated and rotated) data will provide comparable energies and energy decreases during gradient descent.

### 9.3.2 Consistency with Respect to P

We now focus on the behavior of the template estimation algorithm when the sampling of the discretized template changes. The methodology and notations will be similar to Sect. 9.3.1. Assume that we are working with surfaces: We wish to understand qualitatively the asymptotic behavior of the expressions of the functional \(J\) and its gradient when the mesh on the template is refined (i.e., when \(P\) and \(T\) increase). The idea is once again to keep a similar behavior in the optimization process at coarse and fine resolutions. We illustrate the invariance with respect to the number of points by a numerical experiment presented in Figs. 18 and 19 where the same geometric-functional matching is performed at various resolutions.

Let \((x^P, f^P)_P\) be a sequence of regularly discrete fshapes sampled from a continuous fshape \((X, f)\) and containing, respectively, \(P\) points. Let also \((p^P)_P\) be a sequence of discretized momenta (sampled from a continuous vector field \(p\)) attached to the points of \((x^P, f^P)\). If the discretization is sufficiently uniform, we may consider that the typical size of a local volume element \(r_\ell\) is of order \(P^{-1}\). When the number \(P\) of points is large, we then have

\[ \therefore \]
Fig. 18 Geometric-functional registrations of an ellipsoid onto a functional version of the Stanford’s Bunny. The fshapes are depicted in 18a at medium resolution (around 15,000 points). The experiment consists in minimizing the functional $J_{\text{tan}}$ of Eq. (75) in $p^1$ and $\xi^1$ only (here $N = 1$). The results are presented in Fig. 19. a The source $(x, f)$ and the target $(x^1, f^1)$. Signal $f = 0.5$ is constant, and $f^1$ is equal to 1 on the head and on the tail and 0 elsewhere. b Value of $J_{\text{tan}}$ as a function of the step number (x-axis) at three different resolutions.

$$|f^P|_{x_P}^2 = \sum_{\ell} \hat{f}_\ell r_\ell \approx \int_X f^2(x) dH^d(x)$$

where $\hat{f}$ is defined by formula (61). Notice that other discretization methods may be used to approximate the $L^2$ norm on $X$ as discussed in Sect. 6.4. If we now assume that $\|\mu(x_P, f^P) - \mu(x, f)\|_{W'} \to 0$, then formula (63) implies that

$$\|\mu(x_P, f^P)\|_{W'}^2 \approx \|\mu(x, f)\|_{W'}^2$$

when $P$ is large. Finally, if $\|v^{P^P} - v^P\|_V \to 0$ then the initial velocity fields satisfy

$$v^{P^P}(\cdot) = \sum_{k=1} K(\cdot, x_k) p^P_k \approx \int_X K(\cdot, x) p(x) dH^d(x) = v^{P}(\cdot)$$

when $P$ is large. This means that the magnitude of the momenta should be proportional to the inverse of the density of points to generate comparable displacements at various resolution. In particular, the update of the momentum $p^P$ should be of order $P^{-1}$ as Eq. (83) suggests and $\nabla_{x^P} J$ and $\nabla_{f^P} J$ are of order $P^{-1}$. We then multiply these terms by $P$ to ensure the homogeneity of the update and keep a similar dynamics along the optimization procedure even if the resolution...
Fig. 19 Results of registration presented Fig. 18. First row coarse resolution (around 1200 points) computations takes 30 s. Second row high resolution (around 80,000 points) computations take 2.5 h. a t = 0. b t = 1/2. c t = 1. d Target. e t = 0. f t = 1/2. g t = 1. h Target.

of the meshes changes. We provide an example in Fig. 18b where the values of the functional along the gradient descent at various resolutions are plotted.

10 Conclusion

In this article, we have proposed a well-formalized and efficient extension of the ideas of shape spaces for the treatment of functional shapes. To do so, we have introduced the structure of fshapes’ vector bundle and the metamorphosis setting to model and quantify geometric-functional transformations within these bundles. In addition, the concepts of geometric measure theory (varifolds in particular) were generalized to provide dissimilarity metrics between such objects. The combination of these settings enabled us to express atlas estimation on populations of fshapes as a variational problem, for which we were able to prove existence of solutions.

In the second part of the paper, we addressed the issue of providing practical numerical schemes to efficiently perform the functionals’ optimization. In particular, we detailed the discrete expressions corresponding to the fvarifolds’ data attachment terms, proposed a gradient descent algorithm for the estimation of all variables in the atlas and carefully examined some of the important numerical issues related to these algorithms. The resulting codes include atlas estimation on populations of classical curves and surfaces and extend it to fshapes. They are publicly available on the github repository https://github.com/fshapes/fshapesTk. This was implemented under the simplified “tangential model”: The full metamorphosis setting derived in the theoretical part of the paper is likely to be implemented as well in the near future.

Extensions of the present framework are possible in several directions worth exploring as future work. One of them is to consider fshape bundles modeled on more regular spaces than $L^2$ and hopefully recover existence of solutions in a more direct way than...
the proof presented in this paper, at the price of more involved numerical computations. Another track is to generalize such a framework to more general spaces of signals (e.g., vector fields or tensor fields) and more general group actions, which has been considered only partially in [10]. Finally, one important follow-up problem to atlas estimation is the one of statistical analysis and classification based on these fshape atlases: We have deliberately postponed such questions to an upcoming paper.

Acknowledgements We would like to thank Mirza Faisal Beg, Sieun Lee, Evgeniy Lebed, Marinko Sarunic and their collaborators for providing the OCT dataset and for fruitful discussions. We are very grateful to the anonymous reviewers for their detailed and constructive comments that helped us improve the original manuscript. The authors also acknowledge the support of the French Agence Nationale de la Recherche project HM-TC (Number ANR-09-EMER-006).

Appendix 1: Variation Formula for Fvarifolds: Proof of Theorem 5

The proof follows the same steps as the corresponding result for usual varifolds (cf [12]). Given a $C^1$ vector field $v$ on $E$ with compact support, we can consider the 1-parameter group of diffeomorphisms $\phi_t$ with $\phi_0 = \text{Id}$ and $\partial_t |_{t=0} \phi_t = v$. Then, it follows that:

$$
(\mathcal{L}(v, h) \omega)(x, T_x X, f(x)) = \frac{d}{dt} |_{t=0} ((\psi_t^* \omega)(x, T_x X, f(x))
$$

$$
= \frac{d}{dt} |_{t=0} \left| \frac{d}{dt} \phi_t |_{T_x X} \right| \omega(\phi_t(x), d_x \phi_t(T_x X), f(x) + th(x)).
$$

As we see, the previous leads to several terms in the derivative: Differentiate the volume change term $J_t = |\frac{d}{dt} \phi_t |_{T_x X}|$, the function $\omega$ with respect to the position variable, with respect to the tangent space direction and to the signal part. The derivative with respect to point positions and signal values are easy to obtain and equals, respectively, since $\omega$ is assumed to be $C^1$, $(\frac{\partial \omega}{\partial x} v)$ and $(\frac{\partial \omega}{\partial n} h)$. The two other terms require more attention.

Derivative of the Volume Change

For any vector field $u$ defined on $X$, we shall denote by $u^\top$ and $u^\perp$ the tangential and normal components of $u$ with respect to the tangent space of $X$ at each point. We also introduce the connection $\nabla \cdot$ on the ambient space and an orthonormal frame of tangent vector fields $(e_i)_{i=1,...,d}$ on $X$. Now $J_t = \sqrt{\text{det}((d_x \phi_t(e_i), d_x \phi_t(e_j))_{i,j})}$ so a simple calculation shows that:

$$
\frac{d}{dt} |_{t=0} J_t = \sum_{i=1}^d \{e_i, \nabla e_i v\}
$$

Writing $v = v^\top + v^\perp$ provides a first term $\sum_{i=1}^d \{e_i, \nabla e_i v^\top\}$ which is the tangential divergence of the vector field $v^\top$ denoted usually $\text{div}_X(v^\top)$. The second term
becomes \( \sum_{i=1}^{d} (e_i, \nabla e_i v^\perp) \). For all \( i = 1, \ldots, d \), we have \( (e_i, v^\perp) = 0 \) so that after differentiation we find that \( (e_i, \nabla e_i v^\perp) = -\langle \nabla e_i e_i, v^\perp \rangle \). Therefore:

\[
\sum_{i=1}^{d} (e_i, \nabla e_i v^\perp) = - \sum_{i=1}^{d} \langle \nabla e_i e_i, v^\perp \rangle = -\left( \sum_{i=1}^{d} \nabla e_i e_i \right)^\perp, v^\perp.
\]

In this last expression, we recognize the **mean curvature vector** to the submanifold \( X \), which is the trace of the Weingarten map and is denoted \( H_X \). As a result, we find that:

\[
\int_X \omega \frac{d}{dt} \bigg|_{t=0} J_t = \int_X \omega \text{div}_X (v^\perp) - \int_X \omega (H_X, v^\perp).
\]

where, we adopt in this section the shortcut notation \( \int_X g \) to denote the integral \( \int_X g(x) dH^d(x) \) and \( \int_{\partial X} g \) for \( \int_{\partial X} g(x) dH^{d-1}(x) \). Now, the first term can be rewritten as a boundary integral by applying the Divergence Theorem. Indeed, if we denote by \( \tilde{\omega} \) the function defined on \( X \) by \( \tilde{\omega}(x) = \omega(x, T_x X, f(x)) \) which is \( C^1 \), we have \( \text{div}_X (\tilde{\omega} v^\perp) = \tilde{\omega} \text{div}_X (v^\perp) + \nabla v^\perp \tilde{\omega} \). Applying the Divergence Theorem (cf [30] Section 7) on the submanifold \( X \) gives:

\[
\int_X \omega \text{div}_X (v^\perp) = - \int_X \nabla v^\perp \tilde{\omega} + \int_{\partial X} \omega \left( v, v^\perp \right)
\]

where \( v \) is the unit outward normal to the boundary.

**Computation of \( \nabla_{v^\perp} \tilde{\omega} \)**

The previous equation still involve the derivative of the function \( \tilde{\omega} \) along vector field \( v^\perp \). Given the expression of function \( \tilde{\omega} \), this can be expressed as the sum of three terms:

\[
\nabla_{v^\perp} \tilde{\omega} = \left( \frac{\partial \omega}{\partial x} \bigg| v^\perp \right) + \left( \frac{\partial \omega}{\partial V} \bigg| \nabla v^\perp T^X \right) + \frac{\partial \omega}{\partial m} \left( \nabla f, v^\perp \right)
\]

where \( \nabla_{v^\perp} T^X \) is to be understood as the derivative along vector field \( v^\perp \) of the function \( T^X : x \mapsto T_x X \) having values in the Grassmannian. For a given \( x \in X \), we are thus left with deriving the variation of the tangent space \( T_x X \) when moving along a curve \( t \mapsto \gamma(t) \) with \( \gamma(0) = x \) and \( \gamma'(0) = v^\perp(x) \). Let’s call \( V_t \doteq T_{\gamma(t)} X \in G_d(E) \). As already mentioned, it is often more convenient to think of \( V_t \) through the embedding in \( \mathcal{L}(E) \) given by the orthogonal projector \( p_V \). As explained more in [12], the variation of the orthogonal projector \( p_V \) with respect to \( V \) turns out to be the sum of two linear maps transposed to each other, one of which belonging to \( \mathcal{L}(V, V^\perp) \). In this way, one
identifies the tangent space of $G_d(E)$ at $V$ with the space $L(V, V^\perp)$. Let’s consider $(e_1, \ldots, e_d)$ an orthonormal basis of $V_0 = T_x X$ and $(e_1(t), \ldots, e_d(t))$ the parallel transport of this orthonormal frame along the curve $\gamma$. Let’s also take $z$ any vector in $T_x X$ and set $\tilde{z}$ the vector field on $X$ defined by $\tilde{z}(y) = p_{T_x X}(z)$ for all $y \in X$. Then

$$ p_{V_t}(z) = \sum_{i=1}^d (e_i(t), z)e_i(t) $$

and the derivative of (86) writes

$$ \frac{d}{dt} p_{V_t}(z) = \sum_{i=1}^d (\dot{e}_i(t), z)e_i(t) + (e_i(t), z)\dot{e}_i(t). $$

The derivative of $e_i(t)$ can be decomposed into its tangential and normal parts $\dot{e}_i(t) = \nabla_{\dot{\gamma}(t)} e_i(t) + (\nabla_{\dot{\gamma}(t)} e_i)^\perp$. Then, $\nabla_{\dot{\gamma}(t)} e_i(t)$ being the covariant derivative of $e_i$ in $X$ along the curve $\gamma$, this term vanishes since $e_i(t)$ is obtained by parallel transport. Thus, we eventually get that:

$$ \frac{d}{dt} \bigg|_{t=0} p_{V_t}(z) = \sum_{i=1}^d \left( (\nabla_{\dot{\gamma}(0)} e_i)^\perp, z \right) e_i(0) + (e_i(0), z) (\nabla_{\dot{\gamma}(0)} e_i)^\perp. $$

Since $z \in T_x X$, $(\nabla_{\dot{\gamma}(0)} e_i)^\perp, z = 0$ and it results that

$$ \frac{d}{dt} \bigg|_{t=0} p_{V_t}(z) = \sum_{i=1}^d (e_i(0), z) (\nabla_{\dot{\gamma}(0)} e_i)^\perp = (\nabla_{\dot{\gamma}(0)} \tilde{z})^\perp. $$

This is exactly $\Pi(\dot{\gamma}(0), \tilde{z})$ where $\Pi$ is the second fundamental form of $X$. It results that the variation $(\nabla_{v^\top} T^X)_x$ is precisely $\Pi(v^\top, \cdot)_x \in L(T_x X, (T_x X)^\perp)$. By symmetry of the second fundamental form, it also equals to

$$ \nabla_{v^\top} T^X = \Pi(\cdot, v^\top) = \left( \nabla, v^\top \right)^\perp $$

(87)

Now, under the previous identifications of tangent spaces of $G_d(E)$, $\frac{\partial \omega}{\partial V}$ at $V = T_x X$ belongs to $L(T_x X, (T_x X)^\perp)^* \approx L((T_x X)^\perp, T_x X)$ or equivalently is a map from $E$ to $T_x X$ vanishing on $T_x X$ and so the previous term can be written more compactly as

$$ \left( \frac{\partial \omega}{\partial V} \bigg|_{T^X} \right) = \left( \frac{\partial \omega}{\partial V} \bigg|_{\nabla v^\top} \right) $$

and eventually

$$ \nabla_{v^\top} \tilde{\omega} = \left( \frac{\partial \omega}{\partial x} \bigg|_{v^\top} \right) + \left( \frac{\partial \omega}{\partial V} \bigg|_{\nabla v^\top} \right) + \frac{\partial \omega}{\partial m} \left( \nabla f, v^\top \right) $$

(88)
Derivative of Tangent Spaces’ Transport

We now come to the derivative term on the tangent space part in Eq. (84). Again, we identify tangent spaces with their corresponding orthogonal projector. If we now set $V_t = \frac{d}{dt} \phi_t(T_x X)$, one can easily show that [12]:

$$\frac{d}{dt} \big|_{t=0} V_t = p_{T_x X^\perp} \circ \nabla v \big|_{T_x X} \in \mathcal{L} \left( T_x X, (T_x X)^\perp \right).$$

As previously, $\frac{\partial \omega}{\partial V}$ is an element of $\mathcal{L}(T_x X, T_x X^\perp)^* \approx (T_x X^\perp)^* \otimes T_x X$ and which we can write:

$$\frac{\partial \omega}{\partial V} = \sum_{j=d+1}^n \eta_j^* \otimes \alpha_j$$

for $(\eta_{d+1}, \ldots, \eta_n)$ an orthonormal frame of $T_x X^\perp$ and $(\alpha_j)$ some vectors of $T_x X$ (as usual $\eta^*$ denotes the linear form $\langle \eta, \cdot \rangle$). Then, the variation we wish to compute is:

$$\left( \frac{\partial \omega}{\partial V} \big| \nabla v \right) = \sum_{j=d+1}^n \langle \eta_j, \nabla \alpha_j v \rangle.$$

If we introduce $\left( \frac{\partial \omega}{\partial V} \big| v \right) = \sum_{j=d+1}^n \eta_j^*(v) \alpha_j$ which is a tangent vector field on $X$, we have:

$$\text{div}_X \left( \frac{\partial \omega}{\partial V} \big| v \right) = \sum_{i=1}^d \sum_{j=d+1}^n \left( \langle e_i, \nabla_{e_i} \eta \rangle \langle \eta_j, v \rangle + \langle e_i, \nabla_{e_i} \eta_j \rangle \alpha_j \right).$$

The last term in the sum is also $\sum_{j=d+1}^n \langle \eta_j, \nabla \alpha_j v \rangle$, which is nothing else than $\left( \frac{\partial \omega}{\partial V} \big| \nabla v \right)$. As for the two other terms in the sum, it is easy to see that it equals:

$$\left( \sum_{i=1}^d \langle e_i, \nabla_{e_i} \eta \rangle \sum_{j=d+1}^n \eta_j^* \otimes \alpha_j \big| v \right) = \left( \text{div}_X \left( \frac{\partial \omega}{\partial V} \big| v \right) \right).$$

Hence, it follows that:

$$\left( \frac{\partial \omega}{\partial V} \big| \nabla v \right) = \text{div}_X \left( \frac{\partial \omega}{\partial V} \big| v \right) - \left( \text{div}_X \left( \frac{\partial \omega}{\partial V} \big| v \right) \right).$$

Integrating Eq. (89) over the submanifold $X$ and using the Divergence Theorem as before, we find that:

$$\int_X \left( \frac{\partial \omega}{\partial V} \big| \nabla v \right) = \int_{\partial X} v \left( \frac{\partial \omega}{\partial V} \big| \right) v \right) - \int_X \left( \text{div}_X \left( \frac{\partial \omega}{\partial V} \big| v \right) \right).$$
Synthesis

Coming back now to Eq. (84), the sum of all different terms gives:

\[
\int_X (\mathcal{L}(v, h) \omega) = \int_X \left( \frac{\partial \omega}{\partial x} - \text{div}_X \left( \frac{\partial \omega}{\partial V} \right) \right) v - \int_X \left( \frac{\partial \omega}{\partial x} \right) v^\top + \int_X \left( \frac{\partial \omega}{\partial m} \right) (h - \langle \nabla f, v^\top \rangle) \\
+ \int_{\partial X} \left\langle v, \left( \frac{\partial \omega}{\partial V} \right) v \right\rangle + \omega v^\top.
\]

In addition, the integral

\[
\int_X \left( \frac{\partial \omega}{\partial V} \right) \nabla v^\top = \int_{\partial X} \left\langle v, \left( \frac{\partial \omega}{\partial V} \right) v^\top \right\rangle - \int_X \left( \text{div}_X \left( \frac{\partial \omega}{\partial V} \right) \right) v^\top 
\]

thanks once more to the divergence theorem and the fact that \( \left( \frac{\partial \omega}{\partial V} \right) v^\top = 0 \) because \( v^\top \in T_X X \). Since \( v = v^\top + v^\perp \), it follows that several simplifications occur in the previous sum, leading to

\[
\int_X (\mathcal{L}(v, h) \omega) = \int_X \left( \frac{\partial \omega}{\partial x} - \text{div}_X \left( \frac{\partial \omega}{\partial V} \right) - \omega H_X \right) v^\perp \\
+ \int_{\partial X} \left\langle v, \left( \frac{\partial \omega}{\partial V} \right) v \right\rangle + \omega v^\top
\]

which proves the result of Theorem 5.

Appendix 2: Proof of Proposition 7

Perturbation

We introduce now a perturbation process on any measure \( \nu \) on \( E \times G_d(E) \times \mathbb{R} \) that shall be useful for the following. Let \( a > 0 \) to be fixed later and consider for any \( t \in \mathbb{R} \) the function \( \rho_t : \mathbb{R} \to \mathbb{R} \) such that

\[
\rho_t(z) = z + t(\text{sgn}(z) a - z) \mathbf{1}_{|z| > a}
\]

where \( \text{sgn}(z) \) is the sign of \( z \). We have \( \rho_0 = \text{Id}_\mathbb{R} \) and \( \rho_1 \) is a symmetric threshold at level \( a \). Now for any \( t \in \mathbb{R} \), we denote \( \nu_t \) the new measure defined for any \( \omega \in C_b(E \times G_d(E) \times \mathbb{R}) \) as:

\[
\nu_t(\omega) = \int \omega(x, V, \rho_t(f)) \, d\nu(x, V, f).
\]

\[
\text{ Springer}
\]
Obviously, $v_0 = v$ and $v_1$ is such that $v_1(|f| > a) = 0$ so that $t \mapsto v_t$ is an homotopy from $v$ to a measure under which the signal is a.e. bounded by $a$.

**Proof of Lemma 2**

We show the existence of a fvarifold minimizer in $\mathcal{M}^X$ [cf Eq. (51)] for the extended functional $\tilde{J}$. For any $v \in \mathcal{M}^X$ and $t \in \mathbb{R}$, we denote $J_t = \tilde{J}(v_t)$ where $v_t$ is the previously defined perturbation of $v$ (cf Sect. 2.1) and we assume that $J_0 < \infty$ [which is equivalent to say that $v(|f|^2) < \infty$]. We recall that $\|\mu_{(X^i,f^i)} - v_t\|_{W'}^2 = (\mu_{(X^i,f^i)} - v_t)(\omega^i)$ where $\omega^i = K_{W}(\mu_{(X^i,f^i)} - v_t) \in W$. Then, one easily checks that $J_t < \infty$ and, since we assume that $W$ is continuously embedded into $C^2_0(E \times G_d(E) \times \mathbb{R})$, with existing derivative $J_t'$ at any location $t$ given by

$$J_t' = v\left(\frac{d}{dt}(\rho_t(f))\left(\gamma_f \rho_t(f) + \gamma_W \sum_{i=1}^N \partial \omega^i_{f^j}(x, V, \rho_t(f))\right)\right) \quad (93)$$

In the sequel, $C > 0$ denotes a constant, the value of which may vary from one line to another. Using again the continuous embedding of $W$ into $C^2_0(E \times G_d(E) \times \mathbb{R})$, we get that

$$\left|\frac{\partial \omega^i}{\partial f^j}(x, V, \rho_t(f))\right| \leq C \|\omega^i\|_{W'} \leq C \left(\|\mu_{(X^i,f^i)}\|_{W'} + \|v_t\|_{W'}\right). \quad (94)$$

Moreover, as we mentioned after Proposition 5, $\|\mu_{(X^i,f^i)}\|_{W'} \leq C \mathcal{H}^d(X^i)$. Similarly, $\|v_t\|_{W'} \leq C v_t(E \times G_d(E) \times \mathbb{R})$ and, since $v_t \in \mathcal{M}^X$, we have $v_t(E \times G_d(E) \times \mathbb{R}) = \mathcal{H}^d(X)$ and consequently $\|v_t\|_{W'} \leq C \mathcal{H}^d(X)$. Thus, there exists a constant $C' > 0$ such that:

$$\left|\sum_{i=1}^N \frac{\partial \omega^i}{\partial f^j}(x, V, \rho_t(f))\right| \leq C' \sum_{i=1}^N \left(\mathcal{H}^d(X^i) + \mathcal{H}^d(X)\right) \quad (95)$$

Noticing now that $\frac{d}{dt}(\rho_t(f)) \rho_t(f) \leq 0$, that $|\frac{d}{dt}(\rho_t(f))| = 0$ for $|f| \leq a$ and that $|\rho_t(f)| \geq a$ for $|f| \geq a$ and $t \in [0, 1]$, we get for $t \in [0, 1]$

$$J_t' \leq v\left(-\frac{d}{dt}(\rho_t(f))\right) 1_{|f| \geq a} \left(\gamma_f a - \gamma_W C' \sum_{i=1}^N \left(\mathcal{H}^d(X^i) + \mathcal{H}^d(X)\right)\right) \quad (96)$$

so that

$$\tilde{J}(v_1) \leq \tilde{J}(v_0) \text{ if } a \geq C' \gamma_W \sum_{i=1}^N \left(\mathcal{H}^d(X^i) + \mathcal{H}^d(X)\right). \quad (97)$$
An important consequence of (97) is that one can restrict the search of a minimum for \( \tilde{J} \) on fvarifolds \( \nu \) such that

\[
v \left( \mathbf{1}_{|f| > a} \right) = 0
\]

(98)

with \( a = C' \gamma_W \sum_{i=1}^{N} (\mathcal{H}^d (X^i) + \mathcal{H}^d (X)) \). In particular, since \( \nu \in \mathcal{M}^X \), we will have

\[
x \in X \text{ and } |f| \leq a \text{ a.e.}
\]

(99)

Since \( X \) is bounded and \( G_d (E) \) compact, we can restrict the search of a minimum on a measure supported on a compact subset \( K \subset E \times G_d (E) \times \mathbb{R} \) so that we introduce:

\[
\mathcal{M}^{X,K} = \left\{ \nu \in \mathcal{M}^X \mid (x, V, f) \in K \, \nu \text{ a.e.} \right\}.
\]

(100)

An easy check shows that \( \tilde{J} \) is lower semicontinuous on the set \( \mathcal{M}^{X,K} \) for the weak convergence topology. In addition, \( \mathcal{M}^{X,K} \) is sequentially compact. Indeed, if \( \nu_n \) is a sequence in \( \mathcal{M}^{X,K} \) then all \( \nu_n \) are supported by the compact \( K \) and in particular \( (\nu_n) \) is tight. Also, as already noted, there exists a constant \( C > 0 \) independent of \( n \) such that \( \nu_n (E \times G_d (E) \times \mathbb{R}) \leq C \mathcal{H}^d (X) \), and thus, the sequence is uniformly bounded for the total variation norm. It results, thanks to the Prokhorov Theorem, that there exists a subsequence of \( (\nu_n) \) converging for the weak topology. These compactness and lower semicontinuity properties guarantee the existence of a minimizer \( \nu_* \) of \( \tilde{J} \) with \( \nu_* \in \mathcal{M}^{X,K} \) and

\[
\tilde{J}(\nu_*) \leq \inf_{f \in L^2(X)} J_X (f).
\]

(101)

**Proof of the Proposition**

At this point, we do not yet have a minimizer of \( J_X \). The problem is that if the marginal on \( E \times G_d (E) \) of \( \nu_* \) is the transport of \( \mathcal{H}^d_{|X} \) under the application \( x \mapsto (x, T_x X) \), we cannot guarantee that \( \nu_* \) does not weight multiple signal values in the fiber above a location \((x, T_x X)\). We will now show that for large enough \( \gamma_f / \gamma_W \), there exists \( f_* \in L^2(X) \) such that \( \nu_* = \nu_{X,f_*} \) so that we will deduce

\[
J_X (f_*) = \tilde{J}(\nu_*) \leq \inf_{f \in L^2(X)} J_X (f)
\]

(102)

and the existence of a minimizer on \( L^2(X) \).

Let \( \delta f \in C_b (E \times G_d (E) \times \mathbb{R}) \) and for any \( t \in \mathbb{R} \) consider the perturbation \( \nu_t \in \mathcal{M}^X \) of any \( \nu \in \mathcal{M}^{X,K} \) such that for any \( g \in C_b (E \times G_d (E) \times \mathbb{R}) \) we have:

\[
\nu_t (g) \doteq \int g(x, V, f + t \delta f (x, V, f)) d\nu(x, V, f).
\]

(103)
Here again, the function $t \mapsto \tilde{f}(v_t)$ is differentiable everywhere and we have for $\omega^i = K_W(\mu(x^i, x^i) - v)$

$$\frac{d}{dt} \tilde{f}(v_t) \bigg|_{t=0} = v \left( \gamma f + \gamma_W \sum_{i=1}^{N} \frac{\partial \omega^i}{\partial f}(x, f) \delta f(x, f) \right),$$

so that when $v = \nu_*$ we get

$$\begin{cases}
\gamma f + \gamma_W A(x, V, f) = 0 \nu_* \text{ a.e.} \\
\text{with}
\end{cases}$$

$$(104)$$

The partial derivative of $f \mapsto \gamma f + \gamma_W A(x, V, f)$ with respect to $f$ equals $\gamma f + \gamma_W \frac{\partial A}{\partial f}(x, V, f)$. As before, using the continuous embedding $W \hookrightarrow C_0^2(E \times G_d(E) \times \mathbb{R})$, we have once again a certain constant $C > 0$ such that

$$\left| \frac{\partial A}{\partial f}(x, V, f) \right| \leq C \sum_{i=1}^{N} \left( \mathcal{H}^d(X^i) + \mathcal{H}^d(X) \right), \text{ for all } (x, V, f) \in E \times G_d(E) \times \mathbb{R}. \hspace{1cm} (105)$$

It results that for $\gamma f/\gamma_W$ large enough and for all $(x, V) \in E \times G_d(E)$, $f \mapsto \gamma f + \gamma_W A(x, V, f)$ is a strictly increasing function going from $-\infty$ at $-\infty$ to $+\infty$ at $+\infty$, and thus, there is a unique solution $\tilde{f}(x, V)$ to (104). Now, since the application $f \mapsto \gamma f + \gamma_W A(x, V, f)$ is also $C^1$ on $E \times G_d(E) \times \mathbb{R}$, we deduce from the Implicit Function Theorem that $\tilde{f}$ is a $C^1$ function on $E \times G_d(E)$. Going back to the solution $\nu_*$, we know that for $\nu_*$ almost every $(x, V, f) \in E \times G_d(E) \times \mathbb{R}$, we have $(x, V, f) \in K$ and $f = \tilde{f}(x, V)$, so that $|\tilde{f}| \leq a$ a.e. For any continuous and bounded function $\omega$:

$$v_* (\omega) = \int \omega (x, V, f) \, dv_* = \int \omega (x, V, \tilde{f}(x, V)) \, dv_*$$

and if we denote by $\tilde{\omega}(x, V) = \omega(x, V, \tilde{f}(x, V))$ which is a continuous and bounded function on $E \times G_d(E)$, we have by definition of the space $\mathcal{M}^\times$ given by (51):

$$v_* (\omega) = v_* (\tilde{\omega}) = \int_X \omega(x, T_x X, \tilde{f}(x, T_x X)) \, d\mathcal{H}^d(x). \hspace{1cm} (106)$$

Therefore, setting $f_*(x) = \tilde{f}(x, T_x X)$ for $x \in X$, we see that $|f_*| \leq a$ so that $f \in L^\infty(X)$ and with (106), we deduce that $v_* = \mu(x, f_*)$ which shows that the solution of the optimization is a varifold associated to a true shape $(X, f_*)$. In addition, if $X$ is a $C^p$ submanifold then $x \mapsto T_x X$ is a $C^{p-1}$ function on $X$ and, if $W \hookrightarrow C_0^m(E \times G_d(E) \times \mathbb{R})$ with $m \geq 2$ and $m \geq p$, $A$ and $\tilde{f}$ are $C^{p-1}$ functions so $f_*$ is also $C^{p-1}$, which concludes the proof of Proposition 7.
Appendix 3: Proof of Theorem 6

We shall basically follow the same steps as in the previous simpler cases. First of all, exactly as in 5.2.2, existence of a template shape $X$ is guaranteed with the same compactness and lower semicontinuity arguments. Thus, we may assume that $X$ is fixed and we only have to show existence of minimizers to the simplified functional:

$$J_X \left( f, (\xi^i)_i, (v^i)_i \right) = \frac{\gamma f}{2} \int_X |f(x)|^2 dH^d(x)$$

$$+ \frac{1}{2} \sum_{i=1}^N \left( \int_0^1 |v^i_t|^2_V dt + \gamma \xi \int_X |\xi^i(x)|^2 dH^d(x) \right)$$

$$+ \gamma W \left\| \mu(\xi^i, f^i) - \mu(\phi^i(f^i(x)), f^i + \xi^i \circ (\phi^i(f^i(x))^{-1}) \right\|^2_W$$

Now, as for $v_0$, due to the presence of the penalizations $\|v_i\|_{L^2([0,1], V)} = \int_0^1 |v^i_t|^2_V dt$, $1 \leq i \leq N$, one can assume that all vector fields $v^i$, $1 \leq i \leq N$, belong to a fixed closed ball $B$ of radius $r > 0$ in $L^2([0,1], V)$. As in the proof of Proposition 7, we first show existence of a minimizer in a space of fvarifolds. Namely, extending the definitions of the previous subsections, we introduce the space $M^X$ of measures $\nu$ on $E \times Gd(E) \times \mathbb{R} \times \mathbb{R}^N$ such that for all continuous and bounded function $h$ on $E \times Gd(E) \times \mathbb{R} \times \mathbb{R}^N$, we have:

$$\nu(h) = \int h(x, V) d\nu(x, V, f, (\xi^i)_i) = \int_X h(x, T_x X) dH^d(x).$$

For a measure $\nu$ on $E \times G_d(E) \times \mathbb{R} \times \mathbb{R}^N$ and a diffeomorphism $\phi$, we denote by $\phi \cdot \nu$ the transport of $\nu$ by $\phi$ defined by:

$$(\phi \cdot \nu)(g) = \int |d_x \phi|_{V} g(\phi(x), d_x \phi(V), f, (\xi^i)_i) d\nu(x, V, f, (\xi^i)_i).$$

We now introduce the extended functional:

$$\tilde{J} \left( (\nu^i)_i \right) = \frac{\gamma f}{2} \nu(|f|^2) + \frac{1}{2} \sum_{i=1}^N \left( \int_0^1 |v^i_t|^2_V dt + \gamma \xi \nu(\xi^i) \right)$$

$$+ \gamma W \left\| \mu(\xi^i, f^i) - (\phi^i(f^i) \cdot \nu^i \right\|^2_W$$

with $(v^i)_i \in (L^2([0,1], V))^N$ and $\nu \in M^X$, $\nu(|f|^2)$ denoting in short the integral of the application $(x, V, f, (\xi^i)_i) \mapsto |f|^2$ with respect to $\nu$. For all $1 \leq i \leq N$, $v^i$ is the fvarifold defined for all $\omega \in \mathbb{W}$ by:
\[ v^i(\omega) = \int \omega(x, V, f + \xi^i)dv(x, V, f, \xi^i). \]

As previously, we can consider the perturbation function \( \rho_t \) acting on signals and the measures
\[
v_t(g) \doteq \int g(x, V, \rho_t(f), (\rho_t(\xi^i))_i)dv(x, V, f, (\xi^i)_i).
\]

Denoting \( J_t = \tilde{J}(v_t, (v^i)_i) \), we have, for \( t \in [0, 1] \),
\[
J'_t = v \left( \frac{d}{dt} (\rho_t(f)) \left( \gamma_f \rho_t(f) + \gamma_W \sum_{i=1}^N \left| d_x \phi^i \right|_V \left| \frac{\partial \omega^j}{\partial f} (\phi^i(x), d_x \phi^i(V), \rho_t(f) + \rho_t(\xi^i)) \right| \right) \right)
+ v \left( \sum_{i=1}^N \frac{d}{dt} (\rho_t(\xi^i)) \left( \gamma_f \rho_t(\xi^i) + \gamma_W \left| d_x \phi^i \right|_V \left| \frac{\partial \omega^j}{\partial f} (\phi^i(x), d_x \phi^i(V), \rho_t(f) + \rho_t(\xi^i)) \right| \right) \right) 
\]
(107)

where, for all \( 1 \leq i \leq N, \omega^i = K_W(\mu(X^i, f^i) - (\phi_1^i) \cdot v^i) \). On the first hand, we know that there exists a constant \( C > 0 \) such that for all \( i, x \in E \) and \( V \in G_d(E) \),
\[
|d_x \phi^i |_V \leq C |d \phi^i |_\infty.
\]

In addition, it is a classical result on flows of differential equations (cf [38]) that there exists a non-decreasing continuous function \( \tau : \mathbb{R}^+ \to \mathbb{R}^+ \) independent of \( v \in L^2([0, 1], V) \) such that \( |d \phi^i |_\infty \leq \tau(\|v\|_{L^2([0, 1], V)}) \). On the other hand, using the same controls as in the previous subsections, we have, for some \( C > 0 \) denoting a constant the value of which may change from one line to another:

\[
\left| \frac{\partial \omega^j}{\partial f} (\phi^i(x), d_x \phi^i(V), \rho_t(f) + \rho_t(\xi^i)) \right| \leq \left| \frac{\partial \omega^j}{\partial f} \right|_\infty \leq C \left\| \omega^i \right\|_W \leq C \left( \left\| \mu(X^i, f^i) \right\|_W + \left\| (\phi^i_1) \cdot v^i \right\|_W \right) \leq C \left( \mathcal{H}^d(X^i) + \left( \phi^i_1 \right) \cdot v^i (E \times G_d(E) \times \mathbb{R}) \right). \]

It is also straightforward that there exists a constant \( C > 0 \) such that \( (\phi^i_1) \cdot v^i (E \times G_d(E) \times \mathbb{R}) \) and, using the fact that \( v \in \mathcal{M}^X \) as already argued in 5.2.1, \( v^i (E \times G_d(E) \times \mathbb{R}) = \mathcal{H}^d(X) \). It results, from all the previous inequalities, the existence of a non-decreasing continuous function that we will still call \( \tau \) such that for all \( i, x, V, f, \xi^i \):

\[
\left| d_x \phi^i |_V \left| \frac{\partial \omega^j}{\partial f} (\phi^i(x), d_x \phi^i(V), \rho_t(f) + \rho_t(\xi^i)) \right| \leq \tau(\|v^i\|_{L^2([0, 1], V)}) (\mathcal{H}^d(X^i) + \mathcal{H}^d(X)) \]
(108)
Following the same path that previously lead to (96)

\[
J'_t \leq \nu \left( -\frac{d}{dt} (\rho_t(f)) \right) 1_{|f_t|>a} \left( \gamma_f a - \gamma_W \sum_{i=1}^N \tau \left( \|v^i\|_{L^2([0,1],V)} \right) (\mathcal{H}^d(X^i) + \mathcal{H}^d(X)) \right) \\
+ \sum_{i=1}^N \nu \left( -\frac{d}{dt} (\rho_t(\xi^i)) \right) 1_{|\xi^i|>a} \left( \gamma_{\xi} a - \gamma_W \tau \left( \|v^i\|_{L^2([0,1],V)} \right) (\mathcal{H}^d(X^i) + \mathcal{H}^d(X)) \right).
\]

(109)

Just as in 5.2.1, this implies that \( \tilde{J}(v_1, (\xi_i^i)_i) \leq \tilde{J}(v_0, (\xi_i^i)_i) \) as soon as:

\[
\begin{cases}
 a \geq \frac{\gamma_W}{\gamma_f} \sum_{i=1}^N \tau \left( \|v^i\|_{L^2([0,1],V)} \right) (\mathcal{H}^d(X^i) + \mathcal{H}^d(X)) \\
 a \geq \max_i \frac{\gamma_W}{\gamma_{\xi}} \tau \left( \|v^i\|_{L^2([0,1],V)} \right) (\mathcal{H}^d(X^i) + \mathcal{H}^d(X))
\end{cases}
\]

Therefore, one may restrict the search of a minimum on a set of measures \( \nu \) that are supported on a compact subset \( K \) of \( E \times G_d \times \mathbb{R} \times \mathbb{R}^N \), which space shall denote again \( \mathcal{M}^{X,K} \). The rest of the proof is now very close to the one of 5.2.1. Due to lower semicontinuity of the functional and the compacity of \( \mathcal{M}^{X,K} \) and \( B \) for the weak convergence topologies (respectively on the space of measures and on \( L^2([0,1],V) \)), we obtain the existence of a minimizer \( (\nu_*, (\xi_i^i)_i) \) for the functional \( \tilde{J} \).

The last step is to prove that \( \nu_* \), which belongs \( a \text{ priori} \) to the measure space \( \mathcal{M}^{X,K} \), can be written under the form \( \nu_* = v_{X,f_*(\xi^i)_i} \), i.e., that there exists functions \( f_* \) and \( \xi^i_* \), \( 1 \leq i \leq N \), on \( X \) such that, for all continuous and bounded function \( g \) on \( E \times G_d(E) \times \mathbb{R} \times \mathbb{R}^N \):

\[
\nu_*(g) = \int_X g(x, T_x X, f_*(x), (\xi^i_*(x))_i) d\mathcal{H}^d(x)
\]

(110)

We then consider variations of the signals \( \delta f, (\delta \xi^i)_i \) all belonging to the space \( C_b(E \times G_d(E) \times \mathbb{R} \times \mathbb{R}^N) \) and the path \( t \mapsto v_t \) defined by:

\[
v_t(g) = \int g(x, V, f + t\delta f (x, V, f, (\xi^i)_i), (\xi^i + t\delta \xi^i (x, V, f, (\xi^i)_i))_i) dv_* (x, V, f, (\xi^i)_i).
\]

Now, if \( J_t = \tilde{J}(v_t, (\nu^*_t)_i) \), expressing that \( J'_t|_{t=0} = 0 \) for all \( \delta f \) and \( (\delta \xi^i)_i \) gives, similarly to 5.2.1, the following set of equations:

\[
(\gamma_f f, (\gamma_{\xi} \xi^i)_i) = -A(x, V, f, (\xi^i)_i) \nu_* - a.e
\]

(111)

with \( A(x, V, f, (\xi^i)_i) \) defined as:

\[
\left( \sum_{i=1}^N \frac{\partial \omega^d}{\partial f} \left( \phi^i_1(x), d_x \phi^i_1(V), f + \xi^i \right), \left( \frac{\partial \omega^d}{\partial f} \left( \phi^i_1(x), d_x \phi^i_1(V), f + \xi^i \right) \right)_i \right).
\]

The derivatives \( \partial_f A \) and \( \partial_{\xi^i} A \) can be shown again to be
uniformly bounded in $x$, $V$, $f$, $\zeta^i$, and a previous argument provides the existence of unique solutions $f = \tilde{f}(x, V)$ and $\zeta^i = \tilde{\zeta}^i(x, V)$, $1 \leq i \leq N$, to (111). The rest of the proof is exactly the same as in the end of “Proof of the Proposition” in “Appendix 2”: We set $f_\ast(x) = \tilde{f}(x, T_x X)$ and $\zeta^\ast_i(x) = \tilde{\zeta}^i(x, T_x X)$, $1 \leq i \leq N$, which are again $L^\infty$ functions on $X$. In addition, one shows easily that the minimizing measure $\nu$ equals $\nu_{x, f\ast, (\zeta^\ast_i)_i}$ in the sense of (110). Finally, the regularity of $f_\ast$ and $\zeta^\ast$ when $X$ is a $C^p$ submanifold is obtained again by applying the Implicit Function Theorem to (111).

**Appendix 4: Proof of Theorem 7**

We shall only sketch the essential steps to adapt the content of Appendix 3. We start by writing (58) in an extensive way. This gives:

$$J \left( \left( v^0, h^0 \right), \left( v^i, h^i \right) \right) = \frac{\gamma V_0}{2} \left\| v^0 \right\|_{L^2([0,1], V_0)}^2 + \frac{\gamma f_0}{2} \int_0^1 \int_{X_0} \left| h^0_0 \right|^2 \left| d_x \phi^0_1 |_{T_x X} \right| d\mathcal{H}^d(x)$$

$$+ \sum_{i=1}^N \left( \frac{\gamma V}{2} \left\| v^i \right\|_{L^2([0,1], V)}^2 + \frac{\gamma f}{2} \int_0^1 \int_X \left| h^i_0 \right|^2 \left| d_x \phi^i_1 |_{T_x X} \right| d\mathcal{H}^d(x) \right.$$ 

$$+ \frac{\gamma W}{2} \left\| \mu(X^i, f^i) - \mu \left( f^i(X), (f+\zeta^i) \circ (\phi^i)^{-1} \right) \right\|_{W'}^2 \right)$$

(112)

Now, with Lemma 3, we know that the optimal functions $h^0_\ast$ and $h^i_\ast$, $1 \leq i \leq N$, are given by (57), and thus, the variational problem of (112) can be replaced by the optimization with respect to residual functions $\zeta^0$ and $\zeta^i$, $1 \leq i \leq N$, living in $L^2(X)$ of the functional:

$$J \left( \left( v^0, \zeta^0 \right), \left( v^i, \zeta^i \right) \right) = \frac{\gamma V_0}{2} \left\| v^0 \right\|_{L^2([0,1], V_0)}^2 + \frac{\gamma f_0}{2} \int_{X_0} C^0(x) \left| \zeta^0(x) \right|^2 d\mathcal{H}^d(x)$$

$$+ \sum_{i=1}^N \left( \frac{\gamma V}{2} \left\| v^i \right\|_{L^2([0,1], V)}^2 + \frac{\gamma f}{2} \int_X C^i(x) \left| \zeta^i(x) \right|^2 d\mathcal{H}^d(x) \right.$$ 

$$+ \frac{\gamma W}{2} \left\| \mu(X^i, f^i) - \mu \left( \phi^i(x), (f+\zeta^i) \circ (\phi^i)^{-1} \right) \right\|_{W'}^2 \right)$$
where \( C^0(x) \doteq \left( \int_0^1 \frac{1}{\vert d_x\phi_s^{0} \circ (\phi_1^{0})^{-1}\vert|_{T_xX}} \right)^{-1} ds \)

and for all \( 1 \leq i \leq N \), \( C^i(x) \doteq \left( \int_0^1 \frac{1}{\vert d_x\phi_s^{i} \circ (\phi_1^{0})^{-1}\vert|_{T_xX}} \right)^{-1} ds \). But we note that the previous, up to the weights in the \( L^2 \) metrics given by functions \( C^i \), becomes now extremely close to the problem examined in Theorem 6. In fact, the proof of Appendix 3 can be adapted almost straightforwardly to this situation. As previously, the essential step is to reformulate the optimization problem in a space of measures. Defining the functions:

\[
\tilde{C}^0(x, H) = \left( \int_0^1 \frac{1}{\vert d_x[\phi_s^{0} \circ (\phi_1^{0})^{-1}]\vert|_H} \right)^{-1} ds
\]

\[
\tilde{C}^i(x, H) = \left( \int_0^1 \frac{1}{\vert d_x\phi_s^{i} \circ (\phi_1^{0})^{-1}\vert|_H} \right)^{-1} ds
\]

for \( 1 \leq i \leq N \) and \( (x, H) \in E \times G_d(E) \), we can set, with the same definitions as in Appendix 3:

\[
\hat{J}(v, (v^i)_i) = \frac{\gamma_{f0}}{2} v (\tilde{C}^0 \mid f)^2 + \sum_{i=1}^N \left( \frac{\gamma}{2} \parallel v^i \parallel^2_{L^2([0,1], V)} + \frac{\gamma_{f}^i}{2} v (\tilde{C}^i \mid \xi_i)^2 \right) + \frac{\gamma_W}{2} \parallel \mu(x^i, f^i) - (\phi^{i}) \cdot v^i \parallel^2_{W^*}
\]

(113)

for \( v \in \mathcal{M}^X \). The rest of the proof follows the same path, relying on the fact that we can assume the vector fields \( v^0 \) and \( v^i \) to be bounded in \( L^2([0,1], V_0) \) and \( L^2([0,1], V) \) as we explained in the beginning of Appendix 3. This implies, as already argued in the same section, that we have uniform lower and upper bounds for \( \vert d_x[\phi_s^{0} \circ (\phi_1^{0})^{-1}] \vert|_H \), \( s \in [0,1] \) and \( 1 \leq i \leq N \) and for the quantities \( \vert d_x[\phi_s^{0} \circ (\phi_1^{0})^{-1}] \vert|_H \). Consequently, we can assume that we have \( \alpha, \beta > 0 \) such that for all \( 1 \leq i \leq N \), \( \alpha \leq \parallel \tilde{C}^i \parallel_{\infty} \leq \beta \). Using these inequalities, one can check that we get equivalent controls as in the proof of Theorem 6 which allows us to conclude the existence of a measure minimizer for the extended functional \( \hat{J} \) and then go back to a $f$shape solution for \( J \) with a similar implicit functions’ argument.

References

1. W. Allard. On the first variation of a varifold. *Annals of mathematics*, 95(3), 1972.
2. F. Almgren. *Plateau’s Problem: An Invitation to Varifold Geometry*. Student Mathematical Library, 1966.
3. S. Arguillere, E. Trélat, A. Trouvé, and L. Younes. Shape deformation analysis from the optimal control viewpoint. *Journal de Mathématiques Pures et Appliquées*, 104(1):139–178, July 2015.
4. V. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l’hydrodynamique des fluides parfaits. *Annales de l’Institut Fourier*, 16(2):319–361, 1966.
5. A. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.
6. M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision*, 61(139–157), 2005.
7. M. Bruveris, L. Risser, and F. Vialard. Mixture of Kernels and Iterated Semidirect Product of Diffeomorphisms Groups. *Multiscale Modeling and Simulation*, 10(4):1344–1368, 2012.
8. C. Carmeli, E. De Vito, A. Toigo, and V. Umanita. Vector valued reproducing kernel hilbert spaces and universality. *Analysis and Applications*, 8(01):19–61, 2010.
9. B. Charlier, N. Charon, and A. Trouvé. A short introduction to the functional shapes toolkit. https://github.com/fshapes/fshapesTk/, 2014–2015.
10. N. Charon. *Analysis of geometric and functional shapes with extensions of currents. Application to registration and atlas estimation*. PhD thesis, ENS Cachan, 2013.
11. N. Charon and A. Trouvé. Functional currents : a new mathematical tool to model and analyse functional shapes. *JMIV*, 48(3):413–431, 2013.
12. N. Charon and A. Trouvé. The varifold representation of non-oriented shapes for diffeomorphic registration. *SIAM journal of Imaging Science*, 6(4):2547–2580, 2013.
13. P. Dupuis, U. Grenander, and M. I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of applied mathematics*, 56(3):587, 1998.
14. S. Durrleman. Statistical models of currents for measuring the variability of anatomical curves, surfaces and their evolution. PhD thesis, Inria Sophia Antipolis, 2009.
15. H. Federer. *Geometric measure theory*. Springer, 1969.
16. J. Glaunès. Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l’anatomie numérique. PhD thesis, Université Paris 13, 2005.
17. J. Glaunès, A. Trouvé, and L. Younes. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:712–718, 2004.
18. J. Glaunès and M. Vaillant. Surface matching via currents. *Proceedings of Information Processing in Medical Imaging (IPMI), Lecture Notes in Computer Science*, 3565(381–392), 2006.
19. S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:S151–S160, 2004.
20. S. Lee, N. Fallah, F. Forooghian, A. Ko, K. Pakzad-Vaezi, A. B. Merkur, A. W. Kirker, D. A. Albiani, M. Young, M. V. Sarunic, and M. F. Beg. Comparative analysis of repeatability of manual and automated choroidal thickness measurements in nonneovascular age-related macular degeneration. *Investigative Ophthalmology and Vision Science*, 53(5):2864–2871, 2013.
21. S. Lee, S. X. Han, M. Young, M. F. Beg, M. V. Sarunic, and P. J. Mackenzie. Optic nerve head and peripapillary morphometrics in myopic glaucoma. preprint, 2014.
22. J. Ma, M. I. Miller, A. Trouvé, and L. Younes. Bayesian template estimation in computational anatomy. *NeuroImage*, 42(1):252 – 261, 2008.
23. J. Ma, M. I. Miller, and L. Younes. A bayesian generative model for surface template estimation. *Journal of Biomedical Imaging*, 2010:16, 2010.
24. M. Micheli, P. W. Michor, and D. Mumford. Sobolev metrics on diffeomorphism groups and the derived geometry of spaces of submanifolds. *Izvestiya: Mathematics*, 77(3):541, 2013.
25. P. W. Michor and D. Mumford. A zoo of diffeomorphism groups on $\mathbb{R}^n$. *Annals of Global Analysis and Geometry*, 44(4):529–540, 2013.
26. M. I. Miller, A. Trouvé, and L. Younes. On the metrics and euler-lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4(1):375–405, 2002.
27. M. I. Miller, A. Trouvé, and L. Younes. Geodesic Shooting for Computational Anatomy. *Journal of Mathematical Imaging and Vision*, 24(2):209–228, 2006.
28. M. I. Miller, L. Younes, and A. Trouvé. Diffeomorphometry and geodesic positioning systems for human anatomy. *TECHNOLOGY*, 2(1):36–43, 2014.
29. F. Morgan. *Geometric measure theory, a beginner’s guide*. Academic Press, 1995.
30. L. Simon. *Lecture notes on geometric measure theory*. Australian national university, 1983.
31. B. K. Sriperumbudur, K. Fukumizu, and G. Lanckriet. On the relation between universality, characteristic kernels and RKHS embedding of measures. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS-10)*, volume 9, pages 773–780, 2010.
32. B. Thibert. *Sur l’approximation géométrique d’une surface lisse. Applications en géologie structurale.* PhD thesis, Université Claude Bernard - Lyon 1, 2003.

33. A. Trouvé. An approach of pattern recognition through infinite dimensional group action. *Rapport de recherche du LMENS*, 1995.

34. A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *Intern. Jour. of Computer Vision*, 28(3):213–221, 1998.

35. A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM Journal of Mathematical Analysis*, 37(1):17–59, 2005.

36. A. Trouvé and L. Younes. Metamorphoses through lie group action. *Foundation of computational mathematics*, 5:173–198, sep 2005.

37. A. Trouvé and L. Younes. *Handbook of Mathematical Imaging*, chapter Shape spaces, pages 1309–1362. Springer, 2011.

38. L. Younes. *Shapes and diffeomorphisms*. Springer, 2010.