Calculus of Variations



Weight metamorphosis of varifolds and the LDDMM-Fisher-Rao metric

Hsi-Wei Hsieh 1 · Nicolas Charon 2 [D

Received: 10 December 2021 / Accepted: 7 June 2022 / Published online: 22 June 2022 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

This paper introduces and studies a metamorphosis framework for geometric measures known as varifolds, which extends the diffeomorphic registration model for objects such as curves, surfaces and measures by complementing diffeomorphic deformations with a transformation process on the varifold weights. We consider two classes of cost functionals to penalize those combined transformations, in particular the LDDMM-Fisher-Rao energy which, as we show, leads to a well-defined Riemannian metric on the space of varifolds with existence of corresponding geodesics. We further introduce relaxed formulations of the respective optimal control problems, study their well-posedness and derive optimality conditions for the solutions. From these, we propose a numerical approach to compute optimal metamorphoses between discrete varifolds and illustrate the interest of this model in the situation of partially missing data.

Mathematics Subject Classification $49J15 \cdot 49Q20 \cdot 68T10$

1 Introduction

Diffeomorphic shape analysis has come a long way since its origins in the 90s and some seminal works such as [1, 2]. Besides being at the origin of a constant development of various new mathematical models and numerical methods, it has further shown its wide potential for applications most notably to domains such as computational anatomy or computer vision. Shape analysis is typically concerned with the usual issues of standard statistics, for instance the quantitative comparison of objects, the estimation of the mean of a population or of

Communicated by J. Jost.

 Nicolas Charon charon@cis.jhu.edu
 Hsi-Wei Hsieh hsiwei.hsieh@austin.utexas.edu

- Oden Institute for Computational Engineering and Sciences, The University of Texas, Austin, United States
- Center of Imaging Sciences, Johns Hopkins University, Baltimore, United States



the directions of principal morphological variability. Yet it remains such an active field primarily because of the very intricate mathematical structure of shape spaces that makes the generalization of those notions particularly delicate and still a largely open problem. Taking for example the case of shapes such as curves, surfaces or submanifolds in an Euclidean space, one generally needs to consider those objects as elements of a quotient of an infinite dimensional space by an infinite dimensional group, specifically as the equivalence classes of parametrization functions modulo all their reparametrizations (and in some cases other additional invariance groups such as rigid motions). Thus, even the definition of an adequate notion of metric is highly non trivial and much work has been conducted towards the construction and study of so called *intrinsic* Riemannian metrics on such quotient shape spaces, see e.g. [3, 4].

The construction of metrics and by extension the statistical analysis of shapes can be addressed through a different approach however, which was pioneered by the works of Grenander [1]. In Grenander's shape space framework, one views shapes as objects being acted on by a certain (potentially infinite dimensional) group G of extrinsic deformations. For the above situation of submanifolds embedded in \mathbb{R}^n , it could be for instance the group of diffeomorphisms $Diff(\mathbb{R}^n)$ or a subgroup such as rigid, affine or projective transformations that acts by transporting submanifolds. Then the distance between two shapes can be technically induced from a right-invariant distance on the deformation group itself by looking for a minimal deformation to transform one shape to the other. It is thus quite naturally that this line of work triggered the study of Riemannian metrics on diffeomorphism groups, among which the model coined Large Deformation Diffeomorphic Metric Mapping (LDDMM) in [5] proved particularly prolific. This in good part because it allows to operate with large deformations of the space and provides a principled approach to deal with a variety of geometric shapes including landmarks, images, curves and surfaces or even tensor fields. Yet, this metric formulation involves solving a registration problem i.e. finding an optimal deformation in the group between two given shapes, which a fortiori assumes that those two shapes belong to the same orbit for the action of G. As this is often not a realistic setting when dealing with real data or because of the difficulty of actually solving such boundary value problems, it is very common to relax the exact matching constraint and only enforce that the deformation maps the two shapes approximately as measured by some data attachment (or fidelity) term. For shapes such as landmarks or images, this measure of similarity can be simply taken as the sum of squared differences between the landmark positions or pixel values [5, 6]. But the case of curves and surfaces is typically more elusive because of the aforementioned reparametrization invariance that needs to be embedded within the fidelity term. From a discrete perspective, this means a notion of discrepancy between two curves or two surfaces that does not assume predefined point correspondences and that is robust to differences of sampling and mesh structure.

This precise issue motivated, in particular, the adaptation of ideas from the field of *geometric measure theory* in order to obtain convenient representations for the design of adequate data attachment metrics in the case of submanifold data. The underlying principle is to map shapes into certain spaces of generalized measures and compare them through those measure representations. This was first proposed based on the framework of mathematical currents in [7, 8] and later extended to the representations of *varifolds* [9] and *oriented varifolds* [10]. In each case, the construction of kernel metrics on the corresponding measure spaces lead to



simple fidelity terms that can be effectively used in the above inexact diffeomorphic registration framework. But interestingly, this approach does not need to limit itself to submanifolds, as varifolds in fact encompass a much wider category of geometric objects which can be loosely described as spatial distributions of local orientation planes. Thus, in their recent works [11, 12], the authors suggested to formulate a more general diffeomorphic registration problem directly on varifolds themselves by considering a proper notion of group action of $Diff(\mathbb{R}^n)$ on the varifold space.

Nevertheless, even formulated in the more general setting of varifolds, diffeomorphic models can remain insufficient in adequately dealing with some of the geometric variability encountered in data. Indeed, it is common for two given shapes to exhibit differences which cannot be entirely represented by a diffeomorphism. An obvious situation is the presence of topological changes between the two shapes. It can be also the result of imbalances, such as different fiber densities when comparing white matter fiber tracts obtained from diffusion MRI [13]. The attempt to complement diffeomorphic deformations with transformations of a different nature was first formalized through the fundamental concept of *metamorphoses* in [14]. With subsequent works that include [15–19], it appeared that metamorphoses can provide an effective framework to extend the Riemannian metric setting of LDDMM by incorporating a richer class of shape transformations. Yet metamorphoses have so far been primarily studied and implemented for images and landmarks. With the exception of the measure metamorphosis model of [16], there has been very little to no work done on trying to adapt this framework to submanifolds and let alone to varifolds.

The main goal of this paper is precisely to address that issue by defining and studying a new Riemannian metric on varifolds based on a specific model of varifold metamorphosis that generalizes the purely diffeomorphic approach of [12]. Our model essentially augments the diffeomorphic transport of a varifold with a dynamical change of its weight (or mass) at each point. In order to associate a metric to the varifold metamorphosis, we take inspiration from a related class of models in the context of optimal transport and specifically the unbalanced optimal transport framework that was introduced independently in [20] and [21]. In its dynamic formulation, the Wasserstein-Fisher-Rao (or Hellinger-Kantorovich) metric between measures of \mathbb{R}^n (i.e. 0-dimensional varifolds) that is defined in those works combine usual optimal transport with a Fisher-Rao metric to penalize weight changes in the measures. By analogy, we introduce and study the LDDMM-Fisher-Rao metric between varifolds where the optimal transport component is here replaced by a metric induced from the right-invariant metric on diffeomorphisms of the LDDMM model through its action on varifolds. Besides the mathematical analysis of this novel varifold metamorphosis model, we also tackle its numerical implementation for which we focus on discrete varifolds, namely finite sums of generalized Dirac masses, and again introduce a relaxation of the matching problem based on the aforementioned kernel fidelity terms. The practical interest of this numerical framework for data applications is multifold. Through some of the presented simulations, we will show that it can provide robustness to different types of density imbalances in structured and unstructured geometric data. But we shall also illustrate its potential to deal with partially observed data, which has been a recurrent and challenging issue for many different shape analysis models [22–27].

Relationship to other works. The approach we introduce in this paper relates but differs from several previous works in the following way. In the special case of 0-dimensional varifolds, it leads to a metamorphosis metric between classical measures of \mathbb{R}^n which shares the same diffeomorphic component as the measure metamorphosis model of [16] but combined with



165 Page 4 of 41 H.-W. Hsieh, N. Charon

the Fisher-Rao weight transformation metric of unbalanced optimal transport [20, 21]. Unlike these two models, the LDDMM-Fisher-Rao metric is only well-defined between measures belonging to the same orbit under the combined action of diffeomorphisms and reweighting functions. Compared to [16] however, our use of a more constrained model and metric for the non-diffeomorphic part of the metamorphosis allows us to circumvent the theoretical issues that were uncovered by the authors of [16]; in particular we recover the existence of geodesics that remain in the space of positive measures. In contrast to unbalanced optimal transport on the other hand, the diffeomorphic component of our model guarantees smooth bijective geometric transformations which is often desirable in registration problems. Moreover, as we shall see, the generalization from measures to higher-dimensional varifolds and the corresponding change in the transport action induces further significant differences with these two models. Finally, related to the aforementioned challenge of partial data registration, we should mention the two recent works [25] and [26] which both examine alternative approaches that also rely, to some degree, on the varifold representation. The key difference consists in the fact that [25] rather modifies the fidelity term used in the registration problem into a pseudo-distance that allows for partial overlap of the matched shape and the target while our approach actively models and estimate local varifold weight change itself. On the other hand, the method of [26], although it is also based on the estimation of weight changes, focuses on the space of curves equipped with intrinsic Sobolev metrics and does not fall in the setting of Riemannian metrics between varifolds that we follow in the present work.

Structure of the paper. The paper is organized as follows. In Sect. 2, we start by reviewing basic definitions and properties of varifolds, their relationship to submanifolds and the action of the diffeomorphism group. We then introduce a generalized action involving varifold weight changes and proceed in the definition of induced metrics from the resulting transformation model. Although we are primarily interested in the LDDMM-FR metric mentioned earlier (and introduced in Sect. 2.4), we first discuss in Sect. 2.3 a simplified model involving a static L^2 penalty on the weight change. This approach was only briefly considered in our preliminary work [28] that in fact mainly focused on the special case of constant weight change functions, and we shall expand its analysis in this paper. In the case of LDDMM-FR, we prove that it leads to a well-defined distance between varifolds in a given orbit and show the existence of geodesics. We further derive, in Sect. 2.5, the exact expression of those geodesics between two single Dirac 0- or 1- varifolds. In Sect. 3, we introduce relaxed versions of the matching problems that rely on fidelity metrics derived from reproducing kernels on varifolds; this allows to extend the two models to the comparison of varifolds in different orbits. We also obtain the existence of solutions to the corresponding optimal control problems under adequate assumptions and derive optimality conditions for those solutions. Based on these, in Sect. 4, we propose a numerical approach to estimate the optimal matching between discrete varifolds. It is illustrated with several simple examples, in particular for the registration of shapes with partially missing data, which are meant to serve as proof-of-concept of the validity of the model. Our Python implementation is also made openly available on Github ¹. For the purpose of readability and concision of the main text, we have grouped all proofs of the theorems and propositions in the Appendix.

https://github.com/charoncode/Var_metamorph



2 Weight metamorphoses on varifolds

2.1 Diffeomorphic varifold transformation

The model we shall study in this paper builds on the varifold diffeomorphic registration approach that the authors had introduced in their previous works [11, 12]. In the following paragraphs, we give a brief summary of the general framework of those papers and thereby introduce some notations and definitions that will be necessary for the upcoming sections.

In all the paper, we shall consider the Euclidean space \mathbb{R}^n with n > 2 to be the ambient space in which our "shapes" of interest live. The Euclidean inner product between two vectors $a, b \in \mathbb{R}^n$ will be written $a \cdot b$ or $a^T b$. For any integer $1 \le r \le \infty$, a C^r diffeomorphism of \mathbb{R}^n is a bijective map $\mathbb{R}^n \to \mathbb{R}^n$ of class C^r such that the inverse is also C^r . These will constitute our set of geometric deformations as we shall detail further below. In addition to the Euclidean space, another important set for the rest of the paper is the d-dimensional oriented Grassmannian for $0 \le d \le n$ which we denote by \widetilde{G}_d^n and which is defined as the set of all d-dimensional oriented linear subspaces of \mathbb{R}^n . The oriented Grassmannian carries a natural manifold structure as it can be identified with the quotient of groups $SO(n)/(SO(d)\times SO(n-d))$ d)). Alternatively, one can think of an element $U \in \widetilde{G}_d^n$ as the equivalence class of all the oriented frames $(u_1, \ldots, u_d) \in (\mathbb{R}^n)^d$ that span U with the correct orientation. In particular, for d = 1 or d = n - 1, we can identify \widetilde{G}_d^n with the unit sphere \mathbb{S}^{n-1} . Furthermore, there is a natural metric on \widetilde{G}^n_d that is inherited from this frame representation. Given $U,U'\in \widetilde{G}^n_d$ and (u_1,\ldots,u_d) and (u'_1,\ldots,u'_d) representative frames of those oriented spaces, it is defined

$$\langle U, U' \rangle = \det(u_i \cdot u'_i)_{i, j=1, \dots, d} \tag{1}$$

We can now define the central mathematical object of this paper, namely oriented varifolds which correspond to an oriented version of the classical notion of varifold introduced within the field of geometric measure theory in the seminal works of [29, 30]. Specifically,

Definition 1 An oriented d-varifold μ on \mathbb{R}^n is a nonnegative finite Radon measure on the space $\mathbb{R}^n \times \widetilde{G}_d^n$. We denote by \mathcal{V}_d the space of all oriented d-varifolds.

In the rest of the paper, with a slight abuse of vocabulary, we will use the word varifold instead of oriented varifold to keep the denomination short but note that usual varifolds in the sense of [30, 31] simply result in replacing the oriented Grassmannian in the above definition by its unoriented counterpart. From the Riesz representation theorem, we can equivalently view any varifold μ as a distribution, i.e. an element of the dual space $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)^*$, where $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$ denotes the set of continuous functions vanishing at infinity on $\mathbb{R}^n \times \widetilde{G}_d^n$. It is defined for any test function $\omega \in C_0(\mathbb{R}^n \times G_d^n)$ by:

$$(\mu|\omega) \doteq \int_{\mathbb{R}^n \times \widetilde{G}_{s}^n} \omega(x, U) d\mu(x, U). \tag{2}$$

The weight of a varifold $\mu \in \mathcal{V}_d$ is the finite Radon measure $|\mu|$ on \mathbb{R}^n defined by $|\mu|(A) :=$ $\mu(A \times \widetilde{G}_{d}^{n})$ for all Borel subset A of \mathbb{R}^{n} . As a consequence of the disintegration theorem for measures on product spaces (c.f. [32] Chap. 2), any varifold μ can be decomposed into its weight measure and a family of probability measures on the oriented Grassmannian, namely:

Proposition 1 Let $\mu \in \mathcal{V}_d$. For $|\mu|$ -almost every x in \mathbb{R}^n , there exists a probability measure v_x on \widetilde{G}_d^n such that $x \mapsto v_x$ is $|\mu|$ -measurable and $\mu = |\mu| \otimes v_x$ meaning that for all $\omega \in C_0(\overset{a}{\mathbb{R}}^n \times \widetilde{G}^n_d)$



$$(\mu|\omega) = \int_{\mathbb{R}^n} \int_{\widetilde{G}_d^n} \omega(x, U) d\nu_x(U) d|\mu|(x). \tag{3}$$

Varifolds can be transformed by the action of diffeomorphisms in particular via the notion of pushforward. If $\mu \in \mathcal{V}_d$ and ϕ is a C^1 -diffeomorphism, the *pushforward* of μ by ϕ is the varifold $\phi_{\sharp}\mu \in \mathcal{V}_d$ such that for all $\omega \in C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$:

$$(\phi_{\#}\mu|\omega) \doteq \int_{\mathbb{R}^d \times \widetilde{G}_d^n} \omega(\phi(x), d_x \phi \cdot U) J_U \phi(x) d\mu(x, U). \tag{4}$$

In the above, $d_x \phi$ is the Jacobian of ϕ at x and $d_x \phi \cdot U$ denotes the oriented subspace obtained by transporting U by the linear map $d_x \phi$ i.e. if (u_1, \ldots, u_d) is an oriented frame spanning $U, d_x \phi \cdot U$ is the oriented subspace spanned by the frame $(d_x \phi(u_1), \ldots, d_x \phi(u_d))$. Finally $J_U \phi(x)$ is the Jacobian determinant of ϕ at x along the subspace U i.e. the change of d-volume induced by ϕ along U which is given precisely by $J_U \phi(x) = \sqrt{\det(d_x \phi(u_i) \cdot d_x \phi(u_j))_{i,j}}$ if (u_1, \ldots, u_d) is any orthonormal frame of U. As we shall explain just below, this seemingly convoluted definition of varifold pushforward extends the classical diffeomorphic transformation of submanifolds.

Varifolds provide a representation that embeds a very wide class of mathematical structures, among which usual densities and discrete measures. In \mathcal{V}_d , we will write $\delta_{(x,U)}$ a Dirac mass located at $x \in \mathbb{R}^n$ with attached oriented subspace $U \in \widetilde{G}_d^n$. For such a Dirac varifold, we have as a particular case of the above definitions $|\delta_{(x,U)}| = \delta_x$ and for any diffeomorphism ϕ , $\phi_\# \delta_{(x,U)} = J_U \phi(x) \delta_{(\phi(x),d_x\phi\cdot U)}$. But beyond densities and Dirac masses, varifolds further encompass geometric structures such as submanifolds or rectifiable subsets of \mathbb{R}^n . Indeed, let X be a d-dimensional oriented rectifiable subset of \mathbb{R}^n . As this will not be of critical importance for the rest of the paper, we refer the reader to [31] or [12] for precise definitions of rectifiable sets; otherwise the reader unfamiliar with those notions can instead restrict X to be an oriented d-dimensional smooth submanifold of \mathbb{R}^n . Then X can be naturally represented by a varifold $\mu_X \in \mathcal{V}_d$ defined by:

$$(\mu_X|\omega) \doteq \int_X \omega(x, T_x X) \, d\text{vol}_X(x) \tag{5}$$

where $T_x X \in \widetilde{G}_d^n$ denotes the oriented tangent space to X at x and vol_X the d-volume measure of X. We will refer to such varifolds associated to rectifiable subsets through (5) as *rectifiable varifolds*. In this case, the disintegration given by Proposition 1 is given more specifically by $|\mu_X| = \operatorname{vol}_X$ and $\nu_x = \delta_{T_x X}$. In addition, one can check by direct application of the area formula that for any diffeomorphism ϕ , the pushforward varifold $\phi_\# \mu_X$ is nothing but the varifold $\mu_{\phi(X)}$ associated to the deformed set $\phi(X)$. Lastly, we conclude this brief review by pointing out that, beyond its interest for the shape analysis problems we consider here, this representation of rectifiable sets as varifolds can be also very useful in computational geometry, for example in the estimation of discrete curvatures [33] and curvature flows [34].

2.2 Weight change model

Although the model discussed above does involve transformation of mass through Jacobians of deformations, in many situations, considerable inconsistencies or density variations between measures cannot be fully described by diffeomorphic transformations. We will thus extend the diffeomorphic varifold transformation model by augmenting the pushforward action (4) with a weight or density changing process. To be concrete, we consider rescaling functions living in $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$, the space of positive Borel measurable functions defined on



 \mathbb{R}^n . $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ is a group under pointwise multiplication, and each element α can be applied to varifolds $\mu \in \mathcal{V}_d$ via the action defined by:

$$(\alpha \mu | \omega) = (\mu | \alpha \omega) = \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \alpha(x) \omega(x, T) d\mu(x, T). \tag{6}$$

Informally speaking, this action modifies the density of a varifold at each point in the space \mathbb{R}^n .

We can combine the applications of the diffeomorphism group $\mathrm{Diff}(\mathbb{R}^n)$ and of $\mathcal{B}(\mathbb{R}^n,\mathbb{R}_+)$ on varifolds by semi-direct product. We define a homomorphism Ψ : Diff(\mathbb{R}^n) \mapsto $\operatorname{Aut}(\mathcal{B}(\mathbb{R}^n,\mathbb{R}_+))$, which maps elements φ to automorphisms $\Psi_{\varphi}(\alpha) \doteq \alpha \circ \varphi$ of $\mathcal{B}(\mathbb{R}^n,\mathbb{R}_+)$. The group $\mathrm{Diff}(\mathbb{R}^n) \ltimes_{\Psi} \mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ of semi-direct product between $\mathrm{Diff}(\mathbb{R}^n)$ and $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ can be defined by the following group law:

$$(\varphi_1, \alpha_1) \cdot (\varphi_2, \alpha_2) \doteq (\varphi_1 \circ \varphi_2, (\alpha_1 \circ \varphi_2)\alpha_2).$$

It is straightforward to verify that the identity is (id, 1) and the inverse of the element (φ, α) is $(\varphi^{-1}, \frac{1}{\alpha \circ \varphi^{-1}})$. A natural left action of the group $\mathrm{Diff}(\mathbb{R}^n) \ltimes_{\Psi} \mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ on the space of varifolds \mathcal{V}_d can be then defined as follows:

$$(\varphi, \alpha) \cdot \mu \doteq \varphi_{\#}(\alpha \mu) = (\alpha \circ \varphi^{-1})\varphi_{\#}\mu, \tag{7}$$

which corresponds to first rescaling μ by α then deforming $\alpha\mu$ by φ , or equivalently first deforming μ by φ then rescaling the deformed varifold by the weight function α composed with the coordinate change φ^{-1} .

This extended group action leads, as one can expect, to richer orbits than the sole diffeomorphic pushforward. For a given varifold μ , one can see for instance that the reweighing function α may set some parts of μ to zero mass, which we will leverage for applications to partial matching or for the modelling of topological changes.

2.3 A static L² energy model

In the spirit of Grenander's approach to construct metrics on shape spaces [1], we may attempt to recover a notion of distance on the space of varifolds by introducing right-invariant Riemannian metrics on the transformation group and relying on the previous group action. For purely diffeomorphic transformations, this was addressed in particular through the Large Diffeomorphic Deformation Metric Mapping (LDDMM) model introduced in [5] which we will recap briefly.

Its key principle is to consider diffeomorphisms which are obtained as the flow of a vector field on \mathbb{R}^n . Specifically, let V be a predefined Hilbert space of vector fields $\mathbb{R}^n \to \mathbb{R}^n$ with metric written $\|\cdot\|_V$. All throughout this paper, we will make the assumption that V is continuously embedded into $C_0^2(\mathbb{R}^n,\mathbb{R}^n)$, the space of C^2 vector fields of \mathbb{R}^n that vanish at infinity as well as all first and second order derivatives. In other words, there exists a constant $c_V > 0$ such that for any $u \in V$, we have $\|u\|_{2,\infty} \le c_V \|u\|_V$ where $\|\cdot\|_{2,\infty}$ denotes the supnorm of u and all derivatives up to order 2. Under such assumptions, V can be further shown (c.f. [35] Chap. 8) to be a reproducing kernel Hilbert space (RKHS) associated to a positive vector kernel which we shall write $K_V: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$. The Hilbert norm of V can be also expressed based on this kernel operator, which we shall used later on. Now, considering the space $L^2([0, 1], V)$ of time-dependent vector fields in V, for any $v \in L^2([0, 1], V)$, the flow map $t \in [0, 1] \mapsto \varphi_t^v$ is defined for any $t \in [0, 1]$ and $x \in \mathbb{R}^n$ by the integral equation:



$$\varphi_t^v(x) = x + \int_0^t v_s \circ \varphi_s^v(x) ds$$

or equivalently by $\varphi_0^v = \operatorname{Id}$ and $\partial_t \varphi_t^v = v_t \circ \varphi_t^v$. The results of [5] (see also [35] Chap. 7) guarantee that $G_V \doteq \{\varphi_1^v \mid v \in L^2([0,1],V)\}$ is a subgroup of the group of C^1 -diffeomorphisms of \mathbb{R}^n which can be equipped with the right-invariant distance such that:

$$d_{G_V}(\mathrm{Id},\phi) = \inf \left\{ \left(\int_0^1 \|v_t\|_V^2 dt \right)^{1/2} \mid v \in L^2([0,1],V), \ \varphi_1^v = \phi \right\}.$$

Intuitively, $\|v_t\|_V^2$ represents the instantaneous cost of the deformation at time t and the total cost of a deformation $\phi \in G_V$ is obtained by minimizing the full path energy $\int_0^1 \|v_t\|_V^2 dt$. This metric on deformations can in turn be used to measure the distance between two varifolds μ and μ' (provided they belong to the same orbit under the action of G_V) by finding an optimal deformation field v minimizing $\int_0^1 \|v_t\|_V^2 dt$ under the constraint that $(\varphi_1^v)_\# \mu = \mu'$. This is the idea underlying varifold diffeomorphic registration models considered in several earlier works such as [9, 10, 12].

Now, for the weight change component, we shall first discuss a simple L^2 penalty on α which is consistent with the approach introduced in our previous work [28] in the more restrictive setting of discrete varifolds. As we shall see however, this energy will not allow us to recover a real notion of metric on varifolds. Given a reference varifold μ , we define $L^2(|\mu|)$ to be the space of real-valued functions on \mathbb{R}^n which are square integrable with respect to the weight measure $|\mu|$. If μ' is in the orbit of μ for the action of $G_V \times L^2(|\mu|)$, we can define a deformation- L^2 discrepancy between μ and μ' as:

$$D_{V-\mathcal{L}^2}(\mu, \mu')^2 = \inf\left\{\frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_{\mathbb{R}^n} (\alpha(x) - 1)^2 d|\mu|(x)\right\}$$
(8)

where the infimum is taken over all time dependent deformation fields $v \in L^2([0,1],V)$ and all weight rescaling function $\alpha \in L^2(|\mu|)$ under the constraint that $(\varphi,\alpha) \cdot \mu = \mu'$. The energy that is being minimized in (8) combines the previous deformation cost with a second term measuring how α deviates from a base value of 1 on $\sup(|\mu|)$ with $\gamma > 0$ being a balancing parameter between those two terms. $D_{\mathcal{D}-\mathcal{L}^2}$ provides a relatively natural notion of discrepancy between the two varifolds this formulation allows for simple mathematical analysis and implementation as evidenced previously by [28] and developed further in Sect. 3.2. However, it is quite clear that it does not define a real metric on the orbit of μ . This is due to the energy in (8) being fundamentally non symmetric since the second integral involves only the first of the two varifold and would thus differ if one instead goes from μ' to μ .

Note that the existence of optimal v and α achieving the infimum in (8) can be shown relatively easily but, since these are not linked to any notion of geodesic for an actual metric, we will skip that discussion and postpone the study of existence to the relaxed version of (8) in Sect. 3.2.

2.4 The LDDMM-Fisher-Rao metric

In order to recover a true notion of Riemannian metric on each orbit, we actually need to also introduce a dynamical model for the weight change α as opposed to the static model of the previous section. This is also the idea behind the general concept of metamorphoses in shape spaces [14]. From there on, similar to the deformation φ_t^v , we will thus consider a time-varying weight change function $t \mapsto \alpha_t \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ which is governed by the



simple ODE $\partial_t \alpha_t = g_t \alpha_t$. Here g_t can be interpreted as a growth factor and used as the control for the weight change function. The transformed varifold at $t \in [0, 1]$ is now $\mu_t =$ $(\alpha_t \circ (\varphi_t^v)^{-1})(\varphi_t^v)_{\#}\mu_0.$

To define an adequate energy on $\tilde{\eta}$, we can consider a classical and relatively natural Riemannian metric on spaces of densities: the Fisher-Rao metric studied e.g. in [36]. Given a reference measure λ_0 on \mathbb{R}^n , $\lambda = \rho \lambda_0$ a measure with positive density ρ with respect to λ_0 and $h: \mathbb{R}^n \to \mathbb{R}$, the Fisher-Rao metric at λ is defined as:

$$G_{\lambda}^{FR}(h,h) = \int_{\mathbb{R}^n} \frac{h}{\rho} \cdot \frac{h}{\rho} d\lambda = \int_{\mathbb{R}^n} \frac{h^2}{\rho} d\lambda_0.$$

The Fisher-Rao metric has been shown to satisfy very specific properties in particular when it comes to diffeomorphism invariance [37] and is also used as a penalty in several versions of unbalanced optimal transport [20, 21]. Inspired from those works, we propose to use a similar approach to measure and penalize the cost of the weight change process. Specifically, in our context, the infinitesimal variation of α_t can be quantified as $G_{|u_t|}^{FR}(\partial_t \alpha_t \circ (\varphi_t^v)^{-1}, \partial_t \alpha_t \circ (\varphi_t^v)^{-1})$ $(\varphi_t^v)^{-1}$) and the total energy of the path $t \mapsto \alpha_t$ is then:

$$\int_0^1 G^{FR}_{|\mu_t|}(\partial_t \alpha_t \circ (\varphi^v_t)^{-1}, \partial_t \alpha_t \circ (\varphi^v_t)^{-1}) dt = \int_0^1 \int_{\mathbb{R}^n} g_t^2 \circ (\varphi^v_t)^{-1}(x) \ d|\mu_t|(x).$$

Combined with the LDDMM deformation energy already described in Sect. 2.3, we could then define the squared distance between two varifolds μ_0 and μ' as the minimizer of the total cost:

$$\frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_0^1 \int_{\mathbb{R}^n} g_t^2 \circ (\varphi_t^v)^{-1}(x) \ d|\mu_t|(x). \tag{9}$$

subject to $\mu' = \mu_1 = (\alpha_1 \circ (\varphi_1^v)^{-1})(\varphi_1^v)_{\#}\mu_0$. The key difference with the static L^2 metric model of sect. 2.3 is that the energy of an instantaneous change in weight now evolves alongside the current transformed measure μ_t as opposed to fixing it to the initial μ_0 .

We shall in fact consider a formulation of the above distance that is formally equivalent but simpler to handle technically in view of the derivations that will follow. It is obtained by introducing the change of variable $\tilde{\alpha}_t = \sqrt{\alpha_t}$ which gives after differentiation $\partial_t \tilde{\alpha}_t = \frac{1}{2} \eta_t$ with $\eta_t = g_t \sqrt{\alpha_t}$ and (9) can be rewritten with respect to η as:

$$C_{\mu_0}(v,\eta) = \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_0^1 \int_{\mathbb{R}^n} \eta_t^2 \circ (\varphi_t^v)^{-1}(x) \ d|(\varphi_t^v) + \mu_0|(x) dt. \tag{10}$$

Using the definition of $(\varphi_t^v)_{\#}\mu_0$ and the change of variable $x \mapsto \varphi_t^v(x)$ in the above integral, one also obtains the following equivalent expression of the cost:

$$C_{\mu_0}(v,\eta) = \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \eta_t^2(x) J_U \varphi_t^v(x) d\mu_0(x,U) dt.$$
 (11)

Let us now write the precise definition of the LDDMM-Fisher-Rao (LDDMM-FR) metric between varifolds. For any $\mu \in \mathcal{V}_d$, define the *orbit* of μ as:

$$\Theta(\mu) = \left\{ ((\tilde{\alpha}_1^{\eta})^2 \circ (\varphi_1^v)^{-1}) (\varphi_1^v)_{\#} \mu \text{ s.t } (v, \eta) \in L^2([0, 1], V \times L^2(|\mu|) \\
\tilde{\alpha}_t^{\eta}(x) > 0 \text{ for all } t \in [0, 1] \text{ and } |\mu| - a.e \ x \in \mathbb{R}^n \right\}$$
(12)

in which $\tilde{\alpha}_t^{\eta}$ is by definition the solution of $\partial_t \tilde{\alpha}_t^{\eta} = \frac{1}{2} \eta_t$ with $\tilde{\alpha}_0^{\eta} = 1$ on supp($|\mu|$) i.e. $\tilde{\alpha}_t^{\eta}(x) = 1 + \frac{1}{2} \int_0^t \eta_s(x) ds$ for all $x \in \text{supp}(|\mu|)$. Note that in the definition of $\Theta(\mu)$, we



165 Page 10 of 41 H.-W. Hsieh, N. Charon

impose the constraint $\tilde{\alpha}_t^{\eta}(x) > 0$ meaning that we do not allow cancellation of mass for varifolds in $\Theta(\mu)$. In contrast, we also define the *extended orbit* of μ as:

$$\bar{\Theta}(\mu) = \left\{ ((\tilde{\alpha}_1^{\eta})^2 \circ (\varphi_1^v)^{-1}) (\varphi_1^v)_{\#} \mu \text{ s.t } (v, \eta) \in L^2([0, 1], V \times L^2(|\mu|) \right. \\ \tilde{\alpha}_t^{\eta}(x) \ge 0 \text{ for all } t \in [0, 1], \ |\mu| - a.e \ x \in \mathbb{R}^n \right\}.$$
 (13)

Now, let $\mu_0 \in \mathcal{V}_d$ and $\mu_1 \in \bar{\Theta}(\mu)$. We define the LDDMM-FR distance between μ_0 and μ_1 through the following optimal control problem on (v, η) :

$$d_{V-\mathcal{FR}}(\mu_0, \mu_1)^2 = \inf_{(v,\eta) \in L^2([0,1], V \times L^2([\mu_0])} C_{\mu_0}(v,\eta)$$
(14)

where the cost $C_{\mu_0}(v,\eta)$ is given by (10), $\mu_t = (\tilde{\alpha}_t^{\eta})^2 \circ (\varphi_t^v)^{-1})(\varphi_t^v)_{\#}\mu$ and the control system is:

$$\begin{cases} \partial_t \varphi_t^v = v_t \circ \varphi_t^v \\ \partial_t \tilde{\alpha}_t^{\eta} = \frac{1}{2} \eta_t \end{cases}$$

In what follows, we shall call a path μ_t (or $(\varphi_t^v, \tilde{\alpha}_t^{\eta})$) obtained from the above system of ODEs an **admissible** path.

Despite the fact that $d_{V-\mathcal{FR}}$ is built by combining two Riemannian metrics on the geometric deformation and the weight change function respectively, we emphasize that the two are coupled in particular through the presence of the Jacobian change in the second term of C_{μ_0} . Therefore it is not yet clear that one recovers an actual distance between varifolds nor that optimal controls v and η (and thus geodesics) exist. Those questions are addressed by the following theorems.

Theorem 1 For any $\mu_0 \in \mathcal{V}_d$, the function $d_{\mathcal{D}-\mathcal{FR}}$ given by (14) induces a distance on $\Theta(\mu_0)$, i.e. it is symmetric, satisfies the triangle inequality and $d_{\mathcal{D}-\mathcal{FR}}(\mu', \mu'') = 0$ if and only if $\mu' = \mu''$.

The proof is provided in Appendix. We point out that it is here essential to restrict to $\Theta(\mu_0)$ since the symmetry and triangle inequality would not hold anymore in the extended orbit $\bar{\Theta}(\mu_0)$. The second question is whether one can recover an optimal control and thereby a geodesic path $t \mapsto \mu_t$ for the LDDMM-FR distance between two varifolds. As a preliminary step, let us consider the problem of minimizing the cost with respect to η when the deformation field v is fixed which is the object of the following Lemma.

Lemma 1 Let $\mu' \in \bar{\Theta}(\mu)$ and $v \in L^2([0,1],V)$ a fixed time-dependent vector field such that there exists $\eta_0 \in L^2([0,1],L^2(|\mu|))$ for which $\tilde{\alpha}_t^{\eta_0} \geq 0$ and $(\varphi_1^v,(\tilde{\alpha}_1^{\eta_0})^2) \cdot \mu = \mu'$. Then the variational problem:

$$\inf_{\eta \in L^2([0,1],L^2(|\mu|))} \int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \eta_t^2(x) J_U \varphi_t^v(x) \, d\mu(x,U) dt$$

subject to $\tilde{\alpha}_t^{\eta}(x) \geq 0$ and $(\varphi_1^{v}, (\tilde{\alpha}_1^{\eta})^2) \cdot \mu = \mu'$ has a unique solution given by:

$$\bar{\eta}_t(x) = 2 \frac{\tilde{\alpha}_1^{\eta_0}(x) - 1}{h_t(x) \int_0^1 1/h_s(x) ds}$$

in which $h_t(x) = \int_{\widetilde{G}_d^n} J_U \varphi_t^v(x) dv_x(U)$ where $\mu = |\mu| \otimes v_x$ denotes the disintegration of μ over $\mathbb{R}^n \times \widetilde{G}_d^n$.



Interestingly, we see that for fixed v, one obtains a unique minimizer with respect to η which can be expressed in closed form with respect to the resulting deformation φ_{ν}^{v} . This can be used to prove the existence of solutions to the full optimal control problem.

Theorem 2 Let $\mu \in \mathcal{V}_d$ a varifold with compact support and $\mu' \in \bar{\Theta}(\mu)$. Then there exists $(v^*, \eta^*) \in L^2([0, 1], V \times L^2(|\mu|))$ such that $d_{\mathcal{D}-\mathcal{FR}}(\mu, \mu') = \sqrt{C_{\mu}(v^*, \eta^*)}$.

The proofs of Lemma 1 and Theorem 2 are provided in Appendix. Unlike the restricted problem of Lemma 1, the uniqueness of the minimizer for the joint optimization problem does not necessarily hold anymore.

2.5 Explicit form of the geodesics between single Diracs

Before we introduce a relaxed registration formulation in the next section that will be more amenable to the numerical estimation of the LDDMM-FR metric and its geodesics, it is fairly enlightening to first specify the above in the simplest situation of two single Dirac varifolds. We will focus specifically on 0- and 1-varifolds since, under certain assumptions, we will be able to recover explicit expression for the distance and geodesics and draw interesting comparisons between the different models.

0-varifolds. Let us fix two single Dirac 0-varifolds (i.e. usual measures) of \mathbb{R}^n , $\mu_0 = r_0 \delta_{x_0}$ and $\mu_1 = r_1 \delta_{x_1}$ with $r_0 > 0$ and $r_1 \ge 0$. In such a case, the action of the deformation and reweighting process are essentially acting independently on x and r respectively and it is quite easy to compute explicitly the optimal v and η together with the distance and the geodesic. We have specifically:

Proposition 2 The LDDMM-FR distance between the two Dirac measures $r_0\delta_{x_0}$ and $r_1\delta_{x_1}$ is:

$$d_{V-\mathcal{FR}}(r_0\delta_{x_0},r_1\delta_{x_1}) = \frac{1}{2}d_{K_V}(x_0,x_1)^2 + 2\gamma(\sqrt{r_1} - \sqrt{r_0})^2.$$

where d_{K_V} is the distance associated to the Riemannian metric on \mathbb{R}^n given by $G_x(h, h) = h^T K_V(x, x)^{-1}h$. The geodesic is $r(t)\delta_{x(t)}$ with x(t) being the geodesic between x_0 and x_1 for the metric G and $r(t) = ((1-t)\sqrt{r_0} + t\sqrt{r_1})^2$

The proof is given in the appendix. Note that in the particular case of a radial scalar kernel of the form $K_V(x, y) = \rho\left(\frac{|x-y|^2}{\sigma^2}\right) \operatorname{Id}_{n \times n}$, the above reduces simply to $d_{\mathcal{D}-\mathcal{FR}}(r_0 \delta_{x_0}, r_1 \delta_{x_1}) =$ $\frac{1}{2}|x_1-x_0|^2+2\gamma(\sqrt{r_1}-\sqrt{r_0})^2$ and $x(t)=(1-t)x_0+tx_1$. In other words, the Dirac travels in straight line between its initial and final positions while the weight evolution follows the typical geodesic of the Fisher-Rao metric.

Remark 1 In comparison, the Wasserstein-Fisher-Rao metric of unbalanced optimal transport which was introduced independently in [20] and [21] involves a more intricate interaction between position and weight. Indeed, it is shown that the distance is given by:

$$d_{WFR}(r_0\delta_{x_0}, r_1\delta_{x_1}) = 2\gamma^2 \left[r_0 + r_1 - 2\sqrt{r_0r_1\cos\left(\frac{|x_1 - x_0|}{2\gamma}\right)} \right]$$

where $\overline{\cos}(z) = \cos(|z| \wedge \frac{\pi}{2})$ and that there are in fact two types of geodesics: when $|x_1 - x_0| >$ $\pi\delta$, one obtains the geodesic $[t^2r_1\delta_{x_1}+(1-t)^2r_0\delta_{x_0}]$ in which no transport occur while for $|x_1 - x_0| < \pi \delta$ the geodesic is a travelling Dirac $r(t) \delta_{x(t)}$ where r(t) and x(t) are determined by $r(t) = At^2 - 2Bt + r_0$, $\dot{x}(t) = \omega_0/r(t)$ with



165 Page 12 of 41 H.-W. Hsieh, N. Charon

$$\begin{split} \omega_0 &\doteq 2\gamma \tau \sqrt{\frac{r_0 r_1}{1+\tau^2}}, \quad \tau \doteq \tan\left(\frac{x_1-x_0}{2\gamma}\right), \quad A = r_1 + r_0 - 2\sqrt{\frac{r_0 r_1}{1+\tau^2}}, \\ B &= r_0 - \sqrt{\frac{r_0 r_1}{1+\tau^2}}. \end{split}$$

1-varifolds. As we shall see, the situation becomes quite a bit more complex in the case of 1-varifolds since unlike the 0-dimensional case, there is now a coupling between the diffeomorphic transformation and weight change through the Jacobian factor. In the following, we will directly restrict to a radial kernel $K_V(x,y) = \rho\left(\frac{|x-y|^2}{\sigma^2}\right) \operatorname{Id}_{n\times n}$ for simplicity for which we assume without loss of generality that $\rho(0) = 1$. We will further assume for technical reasons that ρ is C^1 and $\rho'(0) < 0$. Let us again consider two single Dirac 1-varifolds $\mu_0 = r_0 \delta_{(x_0,u_0)}$ and $\mu_1 = r_1 \delta_{(x_1,u_1)}$ where $u_0, u_1 \in \widetilde{G}_1^n \sim \mathbb{S}^{n-1}$ are two unit vectors giving the directions of the lines attached to the two Diracs and $r_0 > 0$, $r_1 \ge 0$. We have the following result which proof can be found in the appendix:

Proposition 3 The LDDMM-FR distance between the Dirac 1-varifolds $r_0\delta_{(x_0,u_0)}$ and $r_1\delta_{(x_1,u_1)}$ is:

$$d_{V-\mathcal{FR}}(r_0\delta_{(x_0,u_0)},r_1\delta_{(x_1,u_1)}) = \frac{\|x_1 - x_0\|^2}{2} + \frac{\tau}{2}\arccos(u_0^T u_1)^2 + 2\tau v^2.$$

where $\tau = -\sigma^2/(2\rho'(0)) > 0$ and ν is given by:

$$v = \begin{cases} -\frac{1}{2} \ln \left(\frac{\chi - 1}{\chi + 1} \right), & \text{if } r_1 = 0\\ \ln \left[-\sqrt{\frac{r_1}{r_0}} \frac{1 - \sqrt{1 + \frac{r_0}{r_1}}(\chi^2 - 1)}{\chi - 1} \right], & \text{if } 0 < r_1 \le r_0\\ \ln \left[\sqrt{\frac{r_1}{r_0}} \frac{1 + \sqrt{1 + \frac{r_0}{r_1}}(\chi^2 - 1)}{\chi + 1} \right], & \text{if } 0 < r_0 \le r_1 \end{cases}$$

with $\chi = \sqrt{1 + \frac{\tau}{\gamma r_0}}$. Whenever $u_1 \neq -u_0$, the geodesic is unique and of the form $r(t)\delta_{(x(t),u(t))}$ where:

$$\begin{cases} x(t) = (1-t)x_0 + tx_1 \\ u(t) = \frac{\sin((1-t)\theta)u_0 + \sin(t\theta)u_1}{\sin(\theta)} & \text{if } \theta = \arccos(u_0^T u_1) \neq 0, \quad u(t) = u_0 \text{ otherwise} \\ r(t) = \left(\frac{\sqrt{r_0}\sinh((1-t)\nu) + \sqrt{r_1}\sinh(\nu t)}{\sinh(\nu)}\right)^2 \end{cases}$$

The above geodesic equations still have a fairly natural interpretation. The Dirac is again being transported along the straight line joining x_0 to x_1 , its direction vector u(t) rotates along the great circle between u_0 and u_1 at constant angular speed. However the dynamics of the weight r(t) becomes more complex than with 0-varifolds. As noted earlier, the parameter γ essentially controls how much of the weight change is performed by the diffeomorphism itself versus the weight rescaling function. In that regard, it is interesting to look at the two limit cases $\gamma \to 0$ and $\gamma \to +\infty$. In the former case, we see that $\chi \to +\infty$ and $\nu \to 0$ which leads to the distance $d_{V-\mathcal{FR}}(r_0\delta_{(x_0,u_0)},r_1\delta_{(x_1,u_1)}) = \frac{|x_1-x_0|^2}{2} + \frac{\tau}{2}\arccos(u_0^Tu_1)^2$.



This is as expected the same distance as for two Dirac masses with the same weight $r_0 = r_1$. On the other hand, when $\gamma \to +\infty$, the model reduces to the pure diffeomorphic matching of varifolds introduced in [12]. In this case, we see that $\chi \to 1^+$ which gives the distance $d_{V-\mathcal{FR}}(r_0\delta_{(x_0,u_0)}, r_1\delta_{(x_1,u_1)}) = \frac{|x_1-x_0|^2}{2} + \frac{\tau}{2}\arccos(u_0^Tu_1)^2 + \frac{\tau}{2}\ln\left(\frac{r_1}{r_0}\right)^2$. Unlike with LDDMM-FR, it becomes in that case impossible to reach the zero mass $(r_1 = 0)$ in finite distance. We further illustrate the effect of γ on the geodesic and optimal transformation with the numerical simulations of Fig. 1.

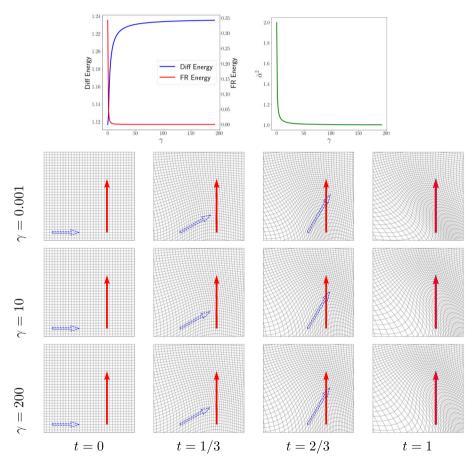


Fig. 1 Exact matching between the two Diracs $\mu = r_0 \delta_{(x_0, u_0)}$ and $\mu' = r_1 \delta_{(x_1, u_1)}$ with $r_0 = 0.5, r_1 = 1$, $x_0 = (0,0), \ x_1 = (1,0) \in \mathbb{R}^2$ and $u_0 = (1,0)$ and $u_1 = (0,1)$. The figure illustrates the effect of γ on the geodesic and optimal transformation calculated from the expression of Proposition 3. The first row shows the relative contributions of the deformation and Fisher-Rao energies as well as the estimated $(\tilde{\alpha}_1^{\eta})^2$ as functions of γ . The next rows showcase the geodesics obtained for three specific values $\gamma = 0.001$, $\gamma = 10$ and $\gamma = 200$. A Dirac mass $r\delta_{(x,u)}$ is here represented as the vector ru which foot point is at the position x. The transformed source varifold is plotted in blue, the target in red



165 Page 14 of 41 H.-W. Hsieh, N. Charon

3 Relaxed matching problem

As follows from their respective variational formulations, the estimation of D_{D-L^2} or $d_{D-\mathcal{FR}}$ and the associated optimal paths require solving optimal control problems with a prescribed terminal condition. This is often not directly tractable except for simple cases and not even necessarily desirable in practice. One of the reason may be that μ' does not belong to the orbit of μ under the joint action of deformations and reweighing functions or, even if it does, imperfections such as noise or segmentation inconsistencies in the data can lead to unnatural behaviour of the distances and of their geodesics if one enforces this terminal matching constraint exactly. Thus, as for many other problems in variational image and shape analysis, it is common to only enforce such a constraint through the addition of a fidelity term to the energy i.e. consider a relaxed (or inexact) matching problem. Yet, unlike images or landmarks, for more complex objects such as measures and varifolds, deriving adequate fidelity metrics that can be nicely embedded within the type of variational problems considered here is not straightforward. In fact, this issue has been the object of several different works in the past such as [9, 10, 38–42]. In this paper, we shall rely on fidelity terms obtained from reproducing kernel metrics on the space of varifolds which have proved successful for that purpose in diffeomorphic registration problems. We thus give a succinct overview of their construction and properties in Sect. 3.1 before focusing on the relaxed versions of the optimal control problems of Sects. 2.3 and 2.4 for which we again study the existence of solutions.

3.1 Kernel metrics on varifolds

Metrics on measure spaces derived from Reproducing Kernel Hilbert Spaces (RKHS), which are also referred to as maximum mean discrepancy in the field of statistics, provide a convenient class of fidelity terms that lead to fully explicit formulas for discrete measures. They essentially rely on the representation of measures as objects of a dual functional space. In a nutshell, to a (scalar) positive definite kernel on $\mathbb{R}^n \times \widetilde{G}_d^n$ with adequate regularity properties is associated, thanks to Aronszajn theorem [43], a reproducing kernel Hilbert space of continuous functions on $\mathbb{R}^n \times \widetilde{G}_d^n$ which dual Hilbert metric will induce a (pseudo-)metric between varifolds. Specifically, we will consider tensor product kernels by relying on the following (Proposition 2 in [11]):

Proposition 4 Let k^{pos} and k^G be continuous positive definite kernels on \mathbb{R}^n and \widetilde{G}_d^n respectively. Assume in addition that for any $x \in \mathbb{R}^n$, $k^{pos}(x,\cdot) \in C_0(\mathbb{R}^n)$. Then $k(x,U,x',U') := k^{pos}(x,x')k^G(U,U')$ for all $x,x' \in \mathbb{R}^n$ and $U,U' \in \widetilde{G}_d^n$ defines a positive definite kernel on $\mathbb{R}^n \times \widetilde{G}_d^n$ and the RKHS W associated to k is continuously embedded into $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$.

Note that we use lower case characters k here to refer to scalar kernels in contrast with the matrix-valued kernel K_V of the deformation field space introduced in Sect. 2.3. Now, as any varifold can be viewed as a linear form on $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$ through (2), the Hilbert norm $\|\cdot\|_W$ of W induces the dual Hilbert metric:

$$d_{W^*}(\mu, \mu') = \|\mu' - \mu\|_{W^*} = \sup_{\|\omega\|_W \le 1} (\mu' - \mu|\omega). \tag{15}$$

The reason why (15) may only give a pseudo-distance rather than a distance between varifolds is because the RKHS W may fail to be dense in $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$. Kernels for which this additional property does hold are called C^0 -universal and many examples, characterizations and constructions of C^0 -universal kernels have been proposed, we shall refer the reader to the discussion in Section 3.2 of [11] as well as [44] for more details on this. Under this condition



on k, we obtain a distance on \mathcal{V}_d which can be shown to metrize the weak-* convergence of varifolds when restricted to specific subsets of \mathcal{V}_d , meaning that $d_{W^*}(\mu_n, \mu) \to 0$ as $n \to +\infty$ if and only if $\mu_n \stackrel{*}{\rightharpoonup} \mu$ i.e. for any continuous compactly supported function $\omega \in C_c(\mathbb{R}^n \times \widetilde{G}_d^n)$, $(\mu_n|\omega) \to (\mu|\omega)$. The precise result, which we shall need later, is the following (Corollary 1 in [12]):

Proposition 5 For M > 0 and K a compact subset of $\mathbb{R}^n \times \widetilde{G}_d^n$, define $\mathcal{V}_{d,M,K} = \{\mu \in \mathcal{V}_d \mid \mu(\mathbb{R}^n \times \widetilde{G}_d^n) \leq M \text{ and supp}(\mu) \subset K\}$. If the kernel k is C^0 -universal then the dual RKHS distance d_{W^*} metrizes the weak-* convergence of varifolds on $\mathcal{V}_{d,M,K}$.

Yet the kernel metric d_{W^*} , as a dual Hilbert metric, remains essentially flat and, unlike the LDDMM-FR distance considered here, does not yield any relevant notion of geodesics. However, its key advantages is that it can be evaluated between any two varifolds μ and μ' in \mathcal{V}_d and in closed form for given kernel k. Indeed, it follows from the Hilbert structure and reproducing kernel property that $d_{W^*}(\mu, \mu')^2 = \|\mu\|_{W^*}^2 - 2\langle \mu, \mu' \rangle_{W^*} + \|\mu'\|_{W^*}^2$ with

$$\langle \mu, \mu' \rangle_{W^*} = \iint_{(\mathbb{R}^n \times \widetilde{G}^n)^2} k(x, U, x', U') d\mu(x, U) d\mu'(x', U').$$

In particular for discrete varifolds $\mu = \sum_{i=1}^{N} r_i \delta_{x_i, U_i}$ and $\mu = \sum_{j=1}^{N} N' r'_j \delta_{x'_j, U'_j}$, the above simply becomes:

$$\langle \mu, \mu' \rangle_{W^*} = \sum_{i=1}^{N} \sum_{j=1}^{N'} k(x_i, U_i, x'_j, U'_j) r_i r'_j.$$
 (16)

Furthermore, by selecting kernels of higher order regularity, we see that the above expression depends smoothly on the positions, directions and weights of the respective varifolds. All these characteristics make kernel metrics d_{W^*} well suited as relaxation terms for our variational problems.

We shall not discuss in many more details the different families of kernels k^{pos} and k^G that could be selected and the corresponding properties they induce on the distance: these questions have been examined quite thoroughly in previous publications notably [10] and [42]. For some of the upcoming mathematical results and in all numerical applications, we shall specifically restrict k^{pos} to a radial kernel $k^{pos}(x, x') = h^{pos}(|x - x'|)$ and k^G to a zonal kernel on \widetilde{G}_d^n namely a function of the form $k^G(U, U') = h^G(\langle U, U' \rangle)$ for the inner product given by (1).

3.2 Relaxed static L² metric problem

We shall first focus on the static model of Sect. 2.3. Given a varifold kernel k and its RKHS W as above, we will replace problem (8) with the following relaxed version:

$$\inf \left\{ \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_{\mathbb{R}^n} (\alpha(x) - 1)^2 d|\mu|(x) + \frac{\lambda}{2} \|\mu_1 - \mu'\|_{W^*}^2 \right\}$$
 (17)

where the minimization is again over $v \in L^2([0,1],V)$ and $\alpha \in L^2(|\mu|)$, and with $\mu_1 = (\varphi_1^v)_\#(\alpha\mu_0)$ although we now only impose that μ_1 approximately matches μ' as measured by their kernel distance $\|\mu_1 - \mu'\|_{W^*}$. The parameter $\lambda > 0$ essentially controls how small this distance should be. We have the following result of well-posedness for this problem which proof follows the standard approach of calculus of variations and is given in Appendix 1:



Theorem 3 Assume that the kernel k is C_0 -universal, that W is continuously embedded in $C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$ and that $supp(\mu) \subset K$ for some compact subset K of $\mathbb{R}^n \times \widetilde{G}_d^n$. Then there exist $(v, \alpha) \in L^2([0, 1], V) \times L^2(|\mu|)$ achieving the infimum in (17).

Note that the above embedding assumption of the RKHS W into $C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$ can be recovered quite simply by imposing some adequate regularity assumptions on the kernels k, as follows from e.g. the results of [45].

In view of the implementation of this model, we shall now specify problem (17) to the particular situation of a discrete source varifold μ and derive the optimality conditions for the resulting finite-dimensional optimal control problem. Thus let us now assume that $\mu = \sum_{i=1}^N r_i \delta_{(x_i,U_i)}$ for some $N \in \mathbb{N}$, $x_i \in \mathbb{R}^n$, $U_i \in \widetilde{G}_d^n$ and $r_i > 0$. Similarly, we shall assume that μ' is of the form $\mu' = \sum_{i=1}^{N'} r_i' \delta_{(x_i',U_i')}$. Following a similar approach as in [12] and [28], we will represent each $U_i \in \widetilde{G}_d^n$ by an ordered frame of d vectors $\{u_i^{(1)}, \ldots, u_i^{(d)}\}$ in \mathbb{R}^n such that U_i is the oriented d-dimensional space spanned by this frame and those vectors are furthermore chosen so that the d-volume of the corresponding parallelotope $|u_i^{(1)} \wedge \ldots \wedge u_i^{(d)}| = \sqrt{\det(u_i^{(l)} \cdot u_i^{(l')})_{l,l'=1,\ldots,d}}$ is equal to r_i . Note that the choice of frame vectors satisfying those conditions is a priori not unique but this is not an issue here since each different term in the energy functional are independent of this choice. This allows to consider the state of the control system as the finite-dimensional variable $q = ((x_i, u_i^{(k)})_{1 \leq i \leq N, 1 \leq k \leq d})$. In addition, the control $\alpha \in L^2(|\mu|)$ can be now represented more simply as the vector $\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}_+^N$ where each α_i stands as the reweighting factor for the i-th Dirac mass. The optimal control problem can be then rewritten as follows:

$$\min_{v,\alpha} C_{V-L^{2}}(v,\alpha) \doteq \frac{1}{2} \int_{0}^{1} \|v_{t}\|_{V}^{2} dt + \frac{\gamma}{2} \sum_{i=1}^{N} r_{i}(\alpha_{i} - 1)^{2}
+ \frac{\lambda}{2} \left\| \left(\sum_{i=1}^{N} \alpha_{i} r_{i} |u_{i}^{(1)}(1) \wedge \ldots \wedge u_{i}^{(d)}(1)| .\delta_{x_{i}(1),U_{i}(1)} \right) - \mu' \right\|_{W^{*}}^{2}$$

where $U_i(1)$ is the oriented d-plane spanned by $\{u_i^{(1)}(1), \ldots, u_i^{(d)}(1)\}$ and the evolution of the state $q(t) = ((x_i(t), u_i^{(k)}(t))_{1 \le i \le N, 1 \le k \le d})$ at $t \in [0, 1]$ is governed by the control system:

$$\begin{cases} \dot{x}_i(t) = v_t(x_i(t)) \\ \dot{u}_i^{(k)}(t) = d_{x_i(t)} v_t(u_i^{(k)}(t)) \end{cases}$$

First, one can notice that for v being fixed, the minimization with respect to α is a quadratic program with non-negativity constraints. However, the problem is non-convex in the deformation field v. Still, necessary conditions satisfied by an optimal v can be obtained from the Pontryagin maximum principle (PMP) of optimal control [46], specifically by using the extension derived in [47]. Their derivation is very similar to that of previous related models [12, 28] and we will skip some of the details for concision. Introducing the costate variables $p = (p_i^x, p_i^{u_k})$, the problem's Hamiltonian is:

$$H(p,q,v) \doteq \sum_{i=1}^{N} \langle p_i^x, v(x_i) \rangle + \sum_{i=1}^{N} \sum_{k=1}^{d} \langle p_i^{u_k}, d_{x_i} v(u_i^{(k)}) \rangle - \frac{1}{2} \|v\|_V^2.$$



If (v, α) is a minimizer of (17), the PMP leads to the existence of a costate function $t \mapsto p(t) \in H^1([0, 1], \mathbb{R}^{nN(1+d)})$ such that $\partial_v H(p(t), q(t), v_t) = 0$ i.e.:

$$v_t(\cdot) = \sum_{i=1}^{N} K_V(x_i(t), \cdot) p_i^x(t) + \sum_{k=1}^{d} \partial_1 K_V(x_i(t), \cdot) (u_i^{(k)}(t), p_i^{u_k}(t))$$
 (18)

where K_V is the kernel of the Hilbert space V and $\partial_1 K_V$ denotes the differential of K_V with respect to his first argument, and p(t) is governed by the adjoint equations $\dot{p}(t) = -\partial_a H(p(t), q(t), v_t)$ which correspond to:

$$\begin{cases} \dot{p}_{i}^{x}(t) = -d_{x_{i}(t)}v_{t}^{T} p_{i}^{x}(t) - \sum_{k=1}^{d} d_{x_{i}(t)}^{(2)} v_{t}(\cdot, u_{i}^{(k)}(t))^{T} p_{i}^{u_{k}}(t) \\ \dot{p}_{i}^{u_{k}}(t) = -d_{x_{i}(t)}v_{t}^{T} p_{i}^{u_{k}}(t) \end{cases}$$
(19)

with the terminal condition that $p_i^x(1)$ and $p_i^{uk}(1)$ are given by minus the derivative of the varifold term with respect to $x_i(1)$ and $u_i^{(k)}(1)$ respectively. To be more explicit, based on the expression of the kernel metric (16), we have $p_i^x(1) = -\partial_{x_i}g(q(1))$, $p_i^{u(k)}(1) = -\partial_{u(k)}g(q(1))$ where:

$$g(q) = \frac{\lambda}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{i} \alpha_{j} r_{i} r_{j} |u_{i}^{(1)} \wedge \dots \wedge u_{i}^{(d)}| . |u_{j}^{(1)} \wedge \dots \wedge u_{j}^{(d)}| k(x_{i}, U_{i}, x_{j}, U_{j})$$

$$- \lambda \sum_{i=1}^{N} \sum_{j=1}^{N'} \alpha_{i} r_{i} r_{j}' |u_{i}^{(1)} \wedge \dots \wedge u_{i}^{(d)}| k(x_{i}, U_{i}, x_{j}', U_{j}') + ||\mu'||_{W^{*}}^{2}.$$
(20)

Note that we did not expand explicitly the last term $\|\mu'\|_{W^*}^2$ since it is here only a constant with respect to q.

3.3 Relaxed LDDMM-FR problem

Let us now introduce a corresponding inexact formulation for the estimation of the LDDMM-FR metric defined by (14) and (11). Relying again on the varifold kernel metric as data fidelity term, we thus consider the minimization:

$$\inf \left\{ \frac{1}{2} \int_{0}^{1} \|v_{t}\|_{V}^{2} dt + \frac{\gamma}{2} \int_{0}^{1} \int_{\mathbb{R}^{n} \times \widetilde{G}_{d}^{n}} \eta_{t}^{2}(x) J_{U} \varphi_{t}^{v}(x) d\mu(x, U) dt + \frac{\lambda}{2} \|\mu_{1} - \mu'\|_{W^{*}}^{2} \right\}$$
(21)

over $v \in L^2([0,1],V)$ and $\eta \in L^2([0,1],L^2(|\mu|))$, where we have $\mu_t = (\tilde{\alpha}_t^\eta)^2 \circ (\varphi_t^v)^{-1})(\varphi_t^v)_\#\mu$. The study of existence of solutions for (21) is however more delicate than with the previous model since the Fisher-Rao penalty does not lead to the necessary weak compactness properties needed in the direct method of calculus of variations. Using a slightly different argument, one is able to prove existence in the particular case of a discrete varifold μ and under specific technical assumptions on the varifold kernel k.

Theorem 4 Assume that the kernel k satisfies $k(x, U, x', U') \ge 0$ and k(x, U, x, U) > 0 for all $x, x' \in \mathbb{R}^n$ and $U, U' \in \widetilde{G}^n_d$ and that W is continuously embedded in $C^1_0(\mathbb{R}^n \times \widetilde{G}^n_d)$. Let $\mu = \sum_{i=1}^N r_i \delta_{(x_i, U_i)}$ be a discrete varifold and $\mu^{tar} \in \mathcal{V}_d$. Then the relaxed problem (21) has a solution $(v, \eta) \in L^2([0, 1], V) \times L^2([0, 1], L^2(|\mu|))$.



165 Page 18 of 41 H.-W. Hsieh, N. Charon

The proof is detailed in Appendix 1. We point out that the above assumptions on the positivity of k are satisfied by most of the kernels that we typically consider in numerical simulations (with the notable exception of the kernel norms related to the model of currents [48]). Moreover, although in practice we are primarily interested in this relaxed approach for discrete varifolds from a numerical point of view, we leave it to future investigations to extend the result of Theorem 4 to more general varifolds μ and kernel families.

Now, as in the previous section, we shall derive optimality conditions for minimizers of this relaxed problem. We again assume that $\mu = \sum_{i=1}^N r_i \delta_{(x_i,U_i)}$ and $\mu' = \sum_{i=1}^{N'} r_i' \delta_{(x_i',U_i')}$. Similar to the above, we can represent $\eta \in L^2([0,1],L^2(|\mu|))$ more simply as a function $\eta \in L^2([0,1],\mathbb{R}^N)$ for each $t \in [0,1]$, we shall rewrite for simplicity $\tilde{\alpha}_t^{\eta} = (\tilde{\alpha}_i(t))_{i=1,\dots,N} \in \mathbb{R}^N$, each component of the vectors being associated to one of the Dirac of μ . We can then describe the state of the optimal control problem by the variable $q = ((x_i, u_i^{(k)}, \tilde{\alpha}_i)_{1 \le i \le N, 1 \le k \le d})$ and the control system is:

$$\begin{cases} \dot{x}_{i}(t) = v_{t}(x_{i}(t)) \\ \dot{u}_{i}^{(k)}(t) = d_{x_{i}(t)}v_{t}(u_{i}^{(k)}(t)) \\ \dot{\tilde{\alpha}}_{i}(t) = \frac{1}{2}\eta_{t,i} \end{cases}$$

with the initial conditions $x_i(0) = x_i$, $(u_i^{(k)}(0))_{k=1,\dots,d} = (u_i^{(k)})$ any frame spanning U_i with $|u_i^{(1)} \wedge \dots \wedge u_i^{(d)}| = r_i$ and $\tilde{\alpha}_i(0) = 1$. Furthermore, for the costate variable $p = (p_i^x, p_i^{u_k}, p_i^{\tilde{\alpha}})$, the Hamiltonian is now given by:

$$H(q, p, v, \eta) = \sum_{i=1}^{N} \langle p_i^x, v(x_i) \rangle + \sum_{i=1}^{N} \sum_{k=1}^{d} \langle p_i^{u_k}, d_{x_i} v(u_i^{(k)}) \rangle + \frac{1}{2} \sum_{i=1}^{N} p_i^{\tilde{\alpha}} \eta_i$$
$$- \frac{1}{2} \|v\|_V^2 - \frac{\gamma}{2} \sum_{i=1}^{N} |u_i^{(1)} \wedge \dots \wedge u_i^{(d)}| \eta_i^2.$$

From this expression of the Hamiltonian and the PMP, we deduce that if (v, η) is a minimizer then:

$$v_{t}(\cdot) = \sum_{i=1}^{N} K_{V}(x_{i}(t), \cdot) p_{i}^{x}(t) + \sum_{k=1}^{d} \partial_{1} K_{V}(x_{i}(t), \cdot) (u_{i}^{(k)}(t), p_{i}^{u_{k}}(t))$$

$$\eta_{t,i} = \frac{p_{i}^{\tilde{\alpha}}}{2\gamma |u_{i}^{(1)}(t) \wedge \cdots \wedge u_{i}^{(d)}(t)|}$$

with the costate function p(t) satisfying the following adjoint equations:

$$\begin{cases} \dot{p}_{i}^{x}(t) = -d_{x_{i}(t)}v_{t}^{T}p_{i}^{x}(t) - \sum_{k=1}^{d}d_{x_{i}(t)}^{2}v_{t}(\cdot, u_{i}^{(k)}(t))p_{i}^{u_{k}}(t) \\ \dot{p}_{i}^{u_{k}}(t) = -d_{x_{i}(t)}v_{t}^{T}p_{i}^{u_{k}}(t) + \frac{\gamma\eta_{i,t}^{2}}{2|u_{i}^{(1)}(t)\wedge\cdots\wedge u_{i}^{(d)}(t)|}\sum_{\ell=1}^{d}C_{i}^{\ell k}(t)u_{i}^{(k)}(t) \\ \dot{p}_{i}^{\tilde{\alpha}}(t) = 0 \Rightarrow p_{i}^{\tilde{\alpha}}(t) = p_{i}^{\tilde{\alpha}} = \text{Cte.} \end{cases}$$

in which the $(C_i^{\ell k}(t))_{1 \leq \ell, k \leq d}$ denote the coefficients of the cofactor matrix of the Gramian $(\langle u_i^{(\ell)}(t), u_i^{(k)}(t) \rangle)_{1 \leq \ell, k \leq d}$. Moreover, one has again the terminal conditions on the costates $p_i^x(1) = -\partial_{x_i} g(q(1)), \ p_i^{u^{(k)}}(1) = -\partial_{u_i^{(k)}} g(q(1))$ and $p_i^{\tilde{\alpha}}(1) = -\partial_{\tilde{\alpha}_i} g(q(1))$ where g is still given by the expression (20) with α_i replaced by $\tilde{\alpha}_i^2$.



Finally, a last condition resulting from the PMP applied to this problem is the conservation of the Hamiltonian over time, namely for all $t \in [0,1], H(q(t), p(t), v_t, \eta_t) =$ $H(q(0), p(0), v_0, \eta_0)$. Furthermore, based on the above expressions of the optimal v_t and η_t and using the reproducing kernel property for the kernel K_V and its derivatives (c.f. [45]), one can show that:

$$H(q(t), p(t), v_t, \eta_t)$$

$$= \frac{1}{2} \|v_t\|_V^2 + \frac{\gamma}{2} \sum_{i=1}^N |u_i^{(1)}(t) \wedge \cdots \wedge u_i^{(d)}(t)| \eta_{t,i}^2$$

$$= \frac{1}{2} \sum_{i=1}^N \langle p_i^x(t), v_t(x_i(t)) \rangle + \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^d \langle p_i^{u_k}(t), d_{x_i(t)} v_t(u_i^{(k)}(t)) \rangle + \frac{1}{4} \sum_{i=1}^N p_i^{\tilde{\alpha}} \eta_{t,i}.$$

From there, we get that the optimal value of the full transformation energy can be expressed as the following function of the initial state and costate variables:

$$\frac{1}{2} \int_{0}^{1} \|v_{t}\|_{V}^{2} dt + \frac{\gamma}{2} \sum_{i=1}^{N} \int_{0}^{1} |u_{i}^{(1)}(t) \wedge \cdots \wedge u_{i}^{(d)}(t)| \eta_{t,i}^{2} dt$$

$$= H(q(0), p(0), v_{0}, \eta_{0})$$

$$= \frac{1}{2} \sum_{i,j=1}^{N} \left\langle p_{i}^{x}(0), K_{V}(x_{j}, x_{i}) p_{j}^{x}(0) + \sum_{k=1}^{d} \partial_{1} K_{V}(x_{j}, x_{i}) (u_{j}^{(k)}, p_{j}^{u_{k}}(0)) \right\rangle$$

$$+ \frac{1}{2} \sum_{i,j=1}^{N} \sum_{k=1}^{d} \left\langle p_{i}^{u_{k}}(0), \partial_{2} K_{V}(x_{j}, x_{i}) (u_{i}^{(k)}, p_{j}^{x}(0)) \right\rangle$$

$$+ \frac{1}{2} \sum_{i,j=1}^{N} \sum_{k=1}^{d} \left\langle p_{i}^{u_{k}}(0), \partial_{1,2}^{2} K_{V}(x_{j}, x_{i}) (u_{i}^{(k)}, u_{j}^{(l)}) p_{j}^{u_{l}}(0) \right\rangle + \frac{1}{8\gamma} \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{(p_{i}^{\tilde{\alpha}})^{2}}{r_{i}} \qquad (22)$$

4 Implementation and results

4.1 Registration algorithms

In order to numerically approximate solutions of the optimal control problems (17) and (21) in the case of discrete Dirac varifolds, we extend the method proposed in [12], namely we solve those problems using a shooting scheme based on the state and costate equations derived in Sects. 3.2 and 3.3. Our implementation in Python further leverages some recently developed libraries in order to efficiently evaluate the different functionals involved and their gradients. It is openly accessible on Github². In the next paragraphs, we detail a little more specifically the different components of our approach.

Optimization scheme. In both models, we perform joint optimization over the deformation i.e., thanks to the above Hamiltonian equations, over the initial costate variables $p_i^x(0)$ and $p_i^{u^{(k)}}(0)$ and the weight changes, namely the α_i 's in the static model and the $p_i^{\tilde{\alpha}}$'s in the LDDMM-FR case. This is done using the limited memory BFGS (L-BFGS) scheme of the



² https://github.com/charoncode/Var_metamorph

SciPy library with the extra non-negativity constraints on the weight change variables. The actual computation of the cost functions and their gradients is explained below.

Numerical integration. Given the initial conditions $(x_i(0))$ and $(u_i^{(k)}(0))$ together with some values for the initial costates $(p_i^x(0))$ and $(p_i^{u(k)}(0))$ (as well as $(p_i^{\tilde{u}})$ for the LDDMM-FR model), we compute the approximate evolution of those variables for $t \in [0, 1]$ by numerical integration of the coupled system of state and costate nonlinear ODEs. In our implementation, we fix a certain number T of discrete time points (typically we take T = 15 in our experiments) and use a Runge-Kutta scheme of order 4 as numerical integrator. It is implemented via the PyTorch library in order to take advantage of the built-in back propagation pipelines and GPU computations.

Cost function evaluation and gradient computation. The cost functions in (17) and (21) are made, on the one hand, of the transformation energies which are directly functions of the initial costates $(p_i^x(0))$, $(p_i^{u^{(k)}}(0))$ and α_i (or $(p_i^{\tilde{\alpha}})$) as shown in (22) and, on the other, of the fidelity term (20) which only depends indirectly on those costates via the terminal state q(1). In both cases, our implementation simply relies on PyTorch and its CUDA bindings to compute those functions on GPU and automatically differentiate them. For the fidelity term, the gradient with respect to the final state q(1) is then back propagated through the RK4 scheme as explained above to automatically recover its gradient with respect to the initial costates. We also point out that the most recurrent and numerically intensive operation in both the integration of the Hamiltonian equations as well as the computation of the varifold data fidelity term consists in the evaluation of kernel convolutions over all Diracs of the source shape. This is typically not handled efficiently in PyTorch itself as it operates by building large kernel matrices, which poses memory and time issues in the case of large sets of Diracs. We remedy this particular problem by taking advantage of the recently developed PyKeops library [49] that provides specialized implementations of kernel operations that remain compatible with PyTorch.

Parameters. The fundamental parameters in both registration models are the choice of deformation kernel K_V , of the varifold kernel k defining the fidelity term and the coefficients γ and λ that weigh the relative importance of the different terms in the cost functional. In our implementation and the simulations of the next section, we use a Gaussian kernel $K_V(x,y) = \exp\left(-|x-y|^2/\sigma_V^2\right) \operatorname{Id}_{n\times n}$ where $\sigma_V > 0$ represents the deformation scale. As for k, we restrict to separable kernel of the form described at the end of Sect. 3.1 and chosen among the different specific classes considered and discussed extensively in [10, 42]. As in most those prior works, the scales of the different kernels are selected manually according to the typical size of the considered shapes and desired level of accuracy of the matching. Our implementation further allows for multistep and multiscale strategies to be used in particular as ways to improve the robustness of the minimization procedure. We point out that developing more data-driven and automatic selection methods for these various parameters is an important ongoing research topic.

4.2 Numerical results

We conclude this paper by presenting a few simulations based on the inexact registration algorithms presented above for 1- and 2- varifolds in \mathbb{R}^n (n=2 or 3). We stress that, in all the examples that involve curves of surfaces, these objects are first converted into corresponding 1- or 2- discrete varifolds to be processed by the algorithm. Specifically, we use the same discretization scheme for polygonal curves and triangulated surfaces that was



introduced in earlier publications [10, 12, 42] in which each facet is approximated by a single Dirac $r_i \delta_{(x_i, U_i)}$ located at the center of mass x_i , with weight r_i given by the area (or length) of the facet and the oriented space U_i being represented by the 1 or 2 frame vectors spanned by the edges. The optimal costates estimated by the approach in turn allow to reconstruct the optimal deformation path φ_i^v and transformed weights $\alpha_i r_i$. For better visualization, rather than showing the transformation of their associated varifolds, we typically plot the corresponding transformation induced on the original curve or surface with the weights being represented as changing colors over the respective facets.

In Fig. 2, we show the result for a simple simulated example in which the unit circle with uniform weight of 1 is matched to an ellipse with piecewise constant weights (displayed by the different colors in the upper left image). On the first row is shown the obtained optimal deformations φ_1^{υ} via deformation grids as well as several intermediate time steps of the shape's evolution for both the LDDMM-L² and LDDMM-FR models (note that for the static L^2 model, the weights being displayed correspond to the linear interpolation $((1-t)+t\alpha_i)\,r_i)$. Both models lead to a close geometric match with similar diffeomorphic deformations as well as a relatively good agreement with the weights of the target shape. We observe nevertheless a sharper concentration of weights around the ground truth values in the case of LDDMM-FR as evidenced by the histogram plots although this comes at the expense of small oscillations around those values.

In Fig. 3, we consider the registration of two surfaces of amygdala each with uniform weights equal to 1, the target surface being obtained by artificially removing portions of the

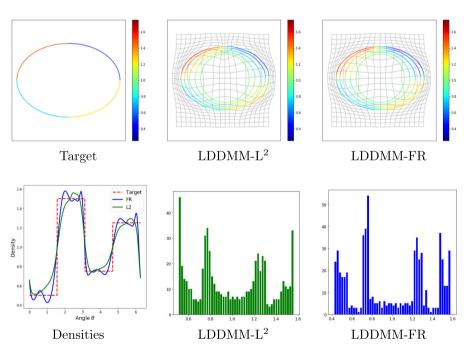


Fig. 2 Registration of a circle with uniform weight density 1 to an ellipse with piecewise constant weights, .5, .75, 1.25 and 1.75. The first column shows the target ellipse with colors corresponding the weight at each location and the plot of those weights as functions of the angle coordinate along the curve. The second column displays the evolution of the transformed curve as well as the histogram of final weight values for the LDDMM-L² model. The corresponding plots for the LDDMM-FR model are on the third column



165 Page 22 of 41 H.-W. Hsieh, N. Charon

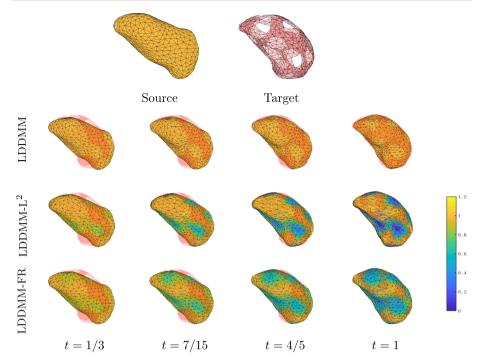


Fig. 3 Registration of two amygdala surfaces (data originating from the BIOCARD dataset [50]) with missing subregions computed with the LDDMM-L², LDDMM-FR, and standard LDDMM approach. Colors on the second and third rows correspond to the weight values at each location. The Chamfer distance to the ground truth is 0.7476 (standard LDDMM), 0.6252 (LDDMM-L²), and 0.5618 (LDDMM-FR.)

ground truth target. We again compared the results of LDDMM-L² and LDDMM-FR as well as the standard LDDMM registration model (computed with the approach of [12] which is equivalent to selecting a very large γ). In this experiment, we took $\lambda=10$ in all models and $\gamma=0.1$ for LDDMM-L² and LDDMM-FR. Such missing parts in the target typically induce, to different degrees, loss of precision in the registration a pure deformation model is being used due to the mismatches between the mass of the source and target at the locations of those missing regions. In contrast, the joint estimation of a weight function in the LDDMM-L² and LDDMM-FR models, as evidenced by the plots in the last two rows of Fig. 3, allows to automatically identify the corresponding missing portions of the source shape and set their weights closer to 0. This leads in turn to a more accurate match. To evaluate it more precisely, we also computed the Chamfer distance between the deformed source surface and the original (complete) ground truth target, showing that the LDDMM-FR model achieves the lowest value in this example.

This is even more clearly exemplified by the result of Fig. 4 that shows the registration of femur surfaces in which the target shape only corresponds to a relatively small part of the source. In such a case, standard diffeomorphic registration of the two varifolds shown on the second row leads to severe shrinking of the bottom section of the source femur as an attempt to geometrically eliminate this extra mass as well as important residual mismatch of



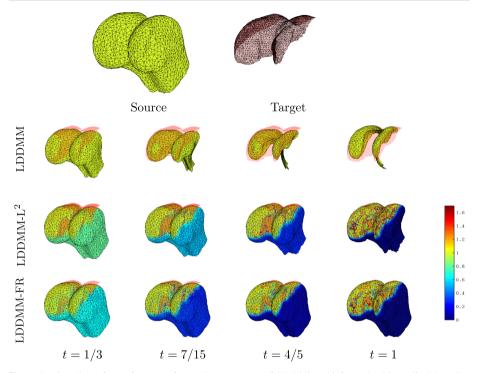


Fig. 4 Registration of two femur surfaces (data courtesy of W. Zbijewski from the Biomedical Imaging department at the Johns Hopkins University) with missing subregions obtained with the standard LDDMM (corresponding to the limit case $\gamma \to +\infty$ of our models), the LDDMM-L² and the LDDMM-FR approach. The colors on the second and third rows correspond to the values of the estimated weight density at each point

the rest of the bone. This is mostly alleviated by the two models LDDMM-L² and LDDMM-FR which instead successfully erase the unmatching part by setting weights to 0 at those locations while resulting in much more natural deformations.

Lastly, with Fig. 5, we show an example of registration involving non-manifold synthetic data in which we compute the registration of a pair of two 1-varifold Diracs in \mathbb{R}^2 (shown as red arrows in the figure with length corresponding to their respective weight) onto the crossing distribution of Diracs shown in blue. For both the standard LDDMM and LDDMM-FR models (here with $\lambda=10$), we obtain a transformed Dirac pair in which each Dirac aligns with one of the two bundles while their respective weights (i.e. length of the arrows in the plot) match the relative densities of the corresponding bundles. In the case of the pure LDDMM model however, this is done via different local expansion of the diffeomorphic mapping along the two bundle directions. In contrast, with the LDDMM-FR model (for which we chose $\gamma=0.01$), the deformation mainly restricts to aligning the Dirac's directions whereas the change of mass is the result of the weight change function. Although only a very simple case, this example suggests the possible interest of these types of models to register or process, e.g., white matter fiber tracts with varying fiber densities as was already pointed out in the authors' previous work [28].



165 Page 24 of 41 H.-W. Hsieh, N. Charon

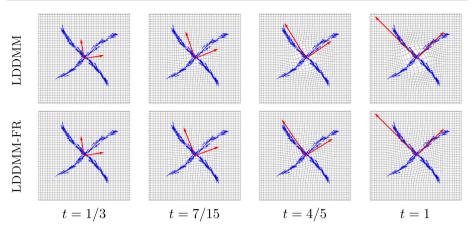


Fig. 5 Registration of a Dirac pair (red) onto crossing bundles of Diracs with distinct densities (120 and 80 Diracs respectively) obtained with the standard LDDMM and the LDDMM-FR approaches

5 Conclusion

In this paper, we proposed extensions of the LDDMM registration model for geometric objects represented as varifolds in which the diffeomorphic deformation is combined with a transformation process of the source varifold weight function. We considered two different classes of cost functionals for these joint transformations, in particular the LDDMM-FR energy in which the weight change is penalized by the Fisher-Rao metric. We showed that the latter is associated to a metamorphosis model on the space of varifolds and induces a well-defined Riemannian metric. We further formulated and studied the corresponding inexact registration problems in which the terminal matching constraint is relaxed through the use of a kernel fidelity metric. Lastly, we derived numerical implementations of those approaches and showcased their potential interest, most notably when it comes to partial geometric data.

Among the persistent shortcomings of the framework developed in this work, we should mention that it remains, to a certain degree, asymmetric in that it allows to modify the weight on the support of the source varifold (and even set it to zero) but cannot a priori "generate" new mass outside of the existing support. In applications to partial data registration, this implies that the model is well suited to deal with missing parts on the target shape but not on the source. Although it is a relatively common in shape analysis that the source shape is taken as a template and thus assumed to be complete, extending this approach to the situation of partial observations for both the source and target is an important and challenging open problem to address in future work.

A second avenue for future research would be to replace the L^2 or Fisher-Rao penalties by different regularization metrics for the weight change function in either a static or metamorphosis setting, with the purpose of imposing spatially smoother weights. In the special case of rectifiable varifolds, one could for instance introduce higher-order Sobolev or total variation norms of the weight change function, by analogy with what has been considered in the context of functional shapes in [19] or elastic shape analysis in [26].



Acknowledgements This work was supported by the National Science Foundation (NSF) under the grant DMS-1945224. The authors would also like to thank Dr. Wojtek Zbijewski and the I-STAR lab for providing the femur data used in Fig. 4 as well as Dr. Sylvain Arguillère for some helpful discussions.

Data availability Some of the datasets analysed in this work (specifically the amygdala and femur data of Fig. 3 and Fig. 4) are not publicly available due to being part of ongoing studies of other research groups but are available from the corresponding author on reasonable request.

Appendix A Proof of Theorem 1

We first prove symmetry by showing that for any μ , $\mu' \in \Theta(\mu_0)$, if there exists an admissible path μ_t from μ to μ' , then there exists a reversed admissible path μ'_t from μ' to μ which consumes the same energy as μ_t . By definition, there exists $(v, \eta) \in L^2([0, 1], V \times L^2(|\mu|))$ such that $\mu_t \doteq ((\tilde{\alpha}_t^{\eta})^2 \circ (\varphi_v^t)^{-1})(\varphi_v^v)_{\#}\mu$. If we define

$$(\bar{v}_t, \bar{\eta}_t) \doteq -(v_{1-t}, \left(\frac{\eta_{1-t}}{\tilde{\alpha}_1^{\eta}}\right) \circ (\varphi_1^v)^{-1}),$$

then it follows from the results of [35] (Chap. 7) that $\varphi_t^{\bar{\nu}} \circ \varphi_1^{\nu} = \varphi_{1-t}^{\nu}$. Furthermore

$$\begin{split} &(\tilde{\alpha}_{t}^{\bar{\eta}} \circ (\varphi_{t}^{\bar{v}})^{-1})(\tilde{\alpha}_{1}^{\eta} \circ (\varphi_{1-t}^{v})^{-1}) \\ &= \left(\left(1 - \frac{1}{2\tilde{\alpha}_{1}^{\eta}(\cdot)} \int_{0}^{t} \eta_{1-s}(\cdot) ds \right) \circ (\varphi_{1}^{v})^{-1} \circ (\varphi_{t}^{\bar{v}})^{-1} \right) \left(\tilde{\alpha}_{1}^{\eta} \circ (\varphi_{1-t}^{v})^{-1} \right) \\ &= \left(\tilde{\alpha}_{1}^{\eta} - \frac{1}{2} \int_{1-t}^{1} \eta_{s}(\cdot) ds \right) \circ (\varphi_{1-t}^{v})^{-1} \\ &= \tilde{\alpha}_{1-t}^{\eta} \circ (\varphi_{1-t}^{v})^{-1}. \end{split}$$

Moreover.

$$\begin{split} \mu_t' &\doteq ((\tilde{\alpha}_t^{\bar{\eta}})^2 \circ (\varphi_t^{\bar{v}})^{-1})(\varphi_t^{\bar{v}})_{\#} \mu' \\ &= ((\tilde{\alpha}_t^{\bar{\eta}})^2 \circ (\varphi_t^{\bar{v}})^{-1})(\varphi_t^{\bar{v}})_{\#} [((\alpha_1^{\eta})^2 \circ (\varphi_1^{v})^{-1})(\varphi_1^{v})_{\#} \mu] \\ &= (\tilde{\alpha}_{1-t}^{\eta})^2 \circ (\varphi_{1-t}^{v})^{-1}(\varphi_{1-t}^{v})_{\#} \mu \\ &= \mu_{1-t}. \end{split}$$

From the equalities above, we can see that

$$\begin{split} \int_{\mathbb{R}^n} \bar{\eta}_t^2 \circ (\varphi_t^{\bar{v}})^{-1} d|(\varphi_t^{\bar{v}})_{\#} \mu'| &= \int_{\mathbb{R}^n} \left(\frac{\eta_{1-t}}{\tilde{\alpha}_1^{\eta}}\right)^2 \circ (\varphi_1^v)^{-1} \circ (\varphi_t^{\bar{v}})^{-1} \frac{1}{(\tilde{\alpha}_t^{\bar{\eta}})^2 \circ (\varphi_t^{\bar{v}})^{-1}} d|\mu_t'| \\ &= \int_{\mathbb{R}^n} \frac{\eta_{1-t}^2 \circ (\varphi_{1-t}^v)^{-1}}{(\tilde{\alpha}_1^{\eta})^2 \circ (\varphi_{1-t}^v)^{-1}} \frac{(\tilde{\alpha}_1^{\eta})^2 \circ (\varphi_{1-t}^v)^{-1}}{(\tilde{\alpha}_{1-t}^{\eta})^2 \circ (\varphi_{1-t}^v)^{-1}} d|\mu_{1-t}| \\ &= \int_{\mathbb{R}^n} \eta_{1-t}^2 \circ (\varphi_{1-t}^v)^{-1} d|(\varphi_{1-t}^v)_{\#} \mu|. \end{split}$$



165 Page 26 of 41 H.-W. Hsieh, N. Charon

Therefore, the cost for the reverse trajectory is the same as the original one:

$$\begin{split} C_{\mu'}(\bar{v},\bar{\eta}) &= \frac{1}{2} \int_0^1 \|v_{1-t}\|_V^2 dt + \frac{\gamma}{2} \int_0^1 \int_{\mathbb{R}^n} \eta_{1-t}^2 \circ (\varphi_{1-t}^v)^{-1} d|(\varphi_{1-t}^v) \# \mu| dt \\ &= \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_0^1 \int_{\mathbb{R}^n} \eta_t^2 \circ (\varphi_t^v)^{-1} d|(\varphi_t^v) \# \mu| dt \\ &= C_{\mu}(v,\eta) < \infty. \end{split}$$

The above also implies that $(\bar{v},\bar{\eta})\in L^2([0,1],V\times L^2(|\mu'|)$. Indeed, as $v\in L^2([0,1],V)$ and $V\hookrightarrow C^1_0(\mathbb{R}^n,\mathbb{R}^n)$, one has from Theorem 7.10 in [35], that $\sup_{t\in[0,1]}\{\|\varphi^v_t-\mathrm{id}\|_{1,\infty},\|(\varphi^v_t)^{-1}-\mathrm{id}\|_{1,\infty}\}<+\infty$ which implies in particular the existence of M>0 such that $1/M\leq J_T\varphi^{\bar{v}}_t(x)\leq M$ for all $t\in[0,1],x\in\mathbb{R}^n$ and $T\in\widetilde{G}^n_d$. Then

$$\int_{0}^{1} \int_{\mathbb{R}^{n}} \bar{\eta}_{t}^{2}(x) d|\mu'|(x) dt \leq M \int_{0}^{1} \int_{\mathbb{R}^{n} \times \widetilde{G}_{d}^{n}} \bar{\eta}_{t}^{2}(x) J_{T} \varphi_{t}^{\bar{v}}(x) d\mu'(x, T) dt$$
$$\leq M C_{\mu'}(\bar{v}, \bar{\eta}) < \infty.$$

where we used (11) for the first inequality. This shows that μ'_t is an admissible path from μ' to μ and that $d_{\mathcal{D}-\mathcal{FR}}$ is symmetric.

Second, we show that if $d_{\mathcal{D}-\mathcal{F}\mathcal{R}}(\mu,\mu')=0$, then $\mu=\mu'$. Now, from the definition of $d_{\mathcal{D}-\mathcal{F}\mathcal{R}}$, there exists a minimizing sequence (v^j,η^j) such that $\mu'=(\varphi_1^{v^j})_\#(\tilde{\alpha}_1^{\eta^j}\mu)$ and $\lim_{j\to\infty}C_\mu(v^j,\eta^j)=0$. This implies in particular that $\int_0^1\|v^j\|_V^2dt$ is uniformly bounded over j and from the same argument as above we can find M>0 such that $1/M\leq J_T\varphi_t^{v^j}(x)\leq M$ for all $j\in\mathbb{N},\,t\in[0,1],\,x\in\mathbb{R}^n$ and $T\in\widetilde{G}_d^n$. We deduce that

$$\begin{split} \int_{\mathbb{R}^{n}} (\tilde{\alpha}_{1}^{\eta^{j}}(x) - 1)^{2} d|\mu|(x) &= \int_{\mathbb{R}^{n}} \left(\int_{0}^{1} \eta_{t}^{j}(x) dt \right)^{2} d|\mu|(x) \\ &\leq \int_{\mathbb{R}^{n}} \int_{0}^{1} (\eta_{t}^{j}(x))^{2} dt \ d|\mu|(x) \\ &\leq M \int_{0}^{1} \int_{\mathbb{R}^{n} \times \tilde{R}_{d}^{n}} (\eta_{t}^{j}(x))^{2} J_{T} \varphi_{t}^{v^{j}}(x) d\mu(x, T) \to 0, \end{split}$$

as $j \to \infty$, the first bound following from Cauchy-Schwarz inequality and the last limit from the fact that the integral is upper bounded by $C_{\mu}(v^j,\eta^j) \to 0$. This also implies that $\int_{\mathbb{R}^n} (\tilde{\alpha}_1^{\eta^j}(x))^2 d|\mu|(x) \to |\mu|(\mathbb{R}^n) < \infty$ and $(\tilde{\alpha}_1^{\eta^j}(x))^2 \to 1 |\mu|$ -a.e. up to a subsequence. From Theorem 4.6.2 in [51], $(\alpha_1^{\eta^j})^2$ converges to 1 in $L^1(|\mu|)$. For any bounded continuous function ω and $j \in \mathbb{N}$

$$\begin{split} (\mu-\mu'|\omega) &= \underbrace{\int_{\mathbb{R}^n \times \tilde{G}_d^n} \omega(x,T) d\mu(x,T) - \int_{\mathbb{R}^n \times \tilde{G}_d^n} \omega(\varphi_1^{v^j}(x),d_x \varphi_1^{v^j} \cdot T) J_T \varphi_1^{v^j}(x) d\mu(x,T)}_{I} \\ &+ \underbrace{\int_{\mathbb{R}^n \times \tilde{G}_d^n} \omega(\varphi_1^{v^j}(x),d_x \varphi_1^{v^j} \cdot T) (1 - (\tilde{\alpha}_1^{\eta^j}(x))^2) J_T \varphi_1^{v^j}(x) d\mu(x,T)}_{I}. \end{split}$$



It is clear that $I \to 0$ from bounded convergence theorem. Also, from the fact that $(\tilde{\alpha}_1^{\eta^j})^2$ converges to 1 in $L^2(|\mu|)$ and the uniform control of Jacobian above, we obtain that

$$|II| \le C \int_{\mathbb{R}^n} |1 - (\tilde{\alpha}^{\eta^j}(x))^2| \, d|\mu|(x) \to 0.$$

This shows that $\mu = \mu'$.

Next, we prove the triangular inequality. Take μ , μ' and μ'' in $\Theta(\mu_0)$ such that $0 < d_{\Theta}(\mu, \mu')$, $d_{\Theta}(\mu', \mu'') < \infty$. Let $(v, \eta) \in L^2([0, 1], V \times L^2(\mathbb{R}^n, |\mu|)$ and $(v', \eta') \in L^2([0, 1], V \times L^2(\mathbb{R}^n, |\mu|))$ $L^2([0,1], V \times L^2(|\mu'|))$ be such that

$$\mu_1 = \mu'$$
 and $\mu'_1 = \mu''$,

where

$$\mu_t \doteq (\varphi_t^v)_{\#}(\tilde{\alpha}_t^{\eta}\mu) \text{ and } \mu_t' \doteq (\varphi_t^{v'})_{\#}(\tilde{\alpha}_t^{\eta'}\mu').$$

Now, let a, b > 1 such that 1/a + 1/b = 1 and denote $u_t \doteq \min\{at, 1\}$ and $u'_t \doteq$ $\max \{b(t-1/a), 0\}$. Defining

$$(\bar{v}_t, \bar{\eta}_t) \doteq a(v_{u_t}, \eta_{u_t}) \mathbf{1}_{[0, 1/a)}(t) + b(v'_{u'_t}, \tilde{\alpha}_1^{\eta} \eta'_{u'_t} \circ \varphi_1^{v}) \mathbf{1}_{[1/a, 1]}(t),$$

then we can easily check that

$$\varphi_t^{\bar{v}} = \varphi_{u_t'}^{v'} \circ \varphi_{u_t}^v,$$

and

$$\begin{split} \tilde{\alpha}_{t}^{\bar{\eta}} &= 1 + \frac{a}{2} \int_{0}^{\min\{t, 1/a\}} \eta_{u_{s}} ds + \frac{\tilde{\alpha}_{1}^{\eta} b}{2} \int_{1/a}^{\max\{t, 1/a\}} \eta'_{u'_{s}} \circ \varphi_{1}^{v} ds \\ &= 1 + \frac{1}{2} \int_{0}^{u_{t}} \eta_{s} ds + \frac{\tilde{\alpha}_{1}^{\eta}}{2} \int_{0}^{u'_{t}} \eta'_{s} \circ \varphi_{1}^{v} ds \\ &= \tilde{\alpha}_{u_{t}}^{\eta} + \tilde{\alpha}_{1}^{\eta} \int_{0}^{u'_{t}} \eta'_{s} \circ \varphi_{1}^{v} ds \\ &= \left(1 + \int_{0}^{u'_{t}} \eta'_{s} \circ \varphi_{1}^{v} ds\right) \tilde{\alpha}_{u_{t}}^{\eta} = (\tilde{\alpha}_{u'_{t}}^{\eta'} \circ \varphi_{1}^{v}) \tilde{\alpha}_{u_{t}}^{\eta}. \end{split}$$

From the above, we can further obtain

$$\begin{split} C_{\mu}(\bar{v},\bar{\eta}) &= \frac{a}{2} \left(\int_{0}^{1/a} \left(\|v_{u_{t}}\|_{V}^{2} + \gamma \int_{\mathbb{R}^{n}} \eta_{u_{t}}^{2} \circ (\varphi_{u_{t}}^{v})^{-1}(x) d |(\varphi_{u_{t}}^{v})_{\#}\mu|(x) \right) \frac{du_{t}}{dt} dt \right) \\ &+ \frac{b}{2} \left(\int_{1/a}^{1} \left(\|v_{u_{t}'}'\|_{V}^{2} + \gamma \int_{\mathbb{R}^{n}} \eta_{u_{t}'}^{2} \circ (\varphi_{u_{t}'}^{v'})^{-1}(x) (\alpha_{1}^{\eta})^{2} \circ (\varphi_{t}^{\bar{v}})^{-1} d |(\varphi_{u_{t}'}^{\bar{v}})_{\#}\mu|(x) \right) \frac{du_{t}'}{dt} dt \right) \\ &= aC_{\mu}(v,\eta) + \frac{b}{2} \left(\int_{1/a}^{1} \left(\|v_{u_{t}'}'\|_{V}^{2} + \gamma \int_{\mathbb{R}^{n}} \eta_{u_{t}'}^{2} \circ (\varphi_{u_{t}'}^{v'})^{-1}(x) d |(\varphi_{u_{t}'}^{v'})_{\#}\mu'|(x) \right) \frac{du_{t}'}{dt} dt \right) \\ &= aC_{\mu}(v,\eta) + bC_{\mu'}(v',\eta') < \infty, \end{split}$$

where the second equality came from the fact that

$$((\tilde{\alpha}_1^{\eta})^2 \circ (\varphi_t^{\bar{v}})^{-1})(\varphi_t^{\bar{v}})_{\#}\mu = (\varphi_t^{\bar{v}})_{\#}(\tilde{\alpha}_1^{\eta})^2\mu = (\varphi_{u'}^{v'})_{\#}(\varphi_1^{v})_{\#}((\alpha_1^{\eta})^2\mu) = (\varphi_{u'}^{v'})_{\#}\mu'.$$



165 Page 28 of 41 H.-W. Hsieh, N. Charon

Again, similar to the argument above, this implies that $(\bar{v}, \bar{\eta}) \in L^2([0, 1], V \times L^2(|\mu|))$. Then, taking $a^* = 1 + \frac{C_{\mu'}(v', \eta')^{1/2}}{C_{\mu}(v, \eta)^{1/2}}$, we obtain that

$$C_{\mu}(\bar{v}, \bar{\eta})^{1/2} = C_{\mu}(v, \eta)^{1/2} + C_{\mu'}(v', \eta')^{1/2}.$$

Now, for any $\varepsilon > 0$, we can find (v, η) and (v', η') satisfying $C_{\mu}(v, \eta)^{1/2} \le d_{\mathcal{D}-\mathcal{FR}}(\mu, \mu') + \varepsilon$ and $C_{\mu}(v', \eta')^{1/2} \le d_{\mathcal{D}-\mathcal{FR}}(\mu', \mu'') + \varepsilon$. Then, taking $(\bar{v}, \bar{\eta})$ as constructed above, we get:

$$d_{\mathcal{D}-\mathcal{F}\mathcal{R}}(\mu,\mu'') \le C_{\mu}(\bar{v},\bar{\eta})^{1/2} \le d_{\mathcal{D}-\mathcal{F}\mathcal{R}}(\mu,\mu') + d_{\mathcal{D}-\mathcal{F}\mathcal{R}}(\mu',\mu'') + 2\varepsilon.$$

for any $\varepsilon > 0$ leading to the triangle inequality.

Finally, we can easily prove that for any μ , $\mu' \in \Theta(\mu_0)$, $d_{\mathcal{D}-\mathcal{FR}}(\mu,\mu') < \infty$. Indeed, from the definition of $\Theta(\mu_0)$, there are admissible paths from μ_0 to μ and μ' separately. We assume that μ and μ' are distinct from μ_0 since this case is trivial. From the arguments in the proof of symmetry, there exists an admissible path from μ to μ_0 . Moreover, we can obtain an admissible path from μ to μ' by concatenating the admissible paths from μ to μ_0 and from μ_0 to μ' with the same argument used above for proving the triangular inequality. Since admissible path have finite energy, we have $d_{\mathcal{D}-\mathcal{FR}}(\mu,\mu') < \infty$.

Appendix B Proof of Theorem 2

Proof of the Lemma First, we point out that since $\mu' = (\varphi_1^v, (\tilde{\alpha}_1^{\eta_0})^2) \cdot \mu = (\tilde{\alpha}_1^{\eta_0} \circ (\varphi_1^v)^{-1})^2 (\varphi_1^v)_{\sharp} \mu$ and since we have $(\tilde{\alpha}_1^{\eta_0} \circ (\varphi_1^v)^{-1})^2 \in L^1(|(\varphi_1^v)_{\sharp} \mu|), \mu'$ is absolutely continuous with respect to $|(\varphi_1^v)_{\sharp} \mu|$ and the Radon-Nykodym derivative of $|\mu'|$ with respect to the transported measure $|(\varphi_1^v)_{\sharp} \mu|$ is precisely $(\tilde{\alpha}_1^{\eta_0} \circ (\varphi_1^v)^{-1}(x))^2$ for all $x \in \text{supp}(|(\varphi_1^v)_{\sharp} \mu|) = \text{supp}(|\mu'|)$. Now, denoting $h_t(x) = \int_{\widetilde{G}_d^v} J_U \varphi_t^v(x) dv_x(U)$, we see that the variational problem consists in minimizing:

$$\begin{split} &\int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \eta_t^2(x) J_U \varphi_t^{v}(x) d\mu(x, U) dt \\ &= \int_0^1 \int_{\mathbb{R}^n} \eta_t^2(x) \left(\int_{\widetilde{G}_d^n} J_U \varphi_t^{v}(x) dv_x(U) \right) d|\mu|(x) dt \\ &= \int_0^1 \int_{\mathbb{R}^n} \eta_t^2(x) h_t(x) d|\mu|(x) dt \end{split}$$

over $\eta \in L^2([0,1], L^2(|\mu|))$ subject to $\tilde{\alpha}_t^{\eta}(x) = 1 + \frac{1}{2} \int_0^t \eta_s(x) ds$ with $\tilde{\alpha}_t^{\eta}(x) \geq 0$ and the boundary constraint $\tilde{\alpha}_1^{\eta}(x) = \tilde{\alpha}_1^{\eta_0}(x)$ for $|\mu|$ -a.e $x \in \mathbb{R}^n$. Since this is a linear (although infinite-dimensional) control system with a quadratic cost function, ignoring for now the non-negativity constraint on $\tilde{\alpha}_t^{\eta}$, it has an essentially unique solution which is given by:

$$\bar{\eta}_t(x) = 2 \frac{\tilde{\alpha}_1^{\eta_0}(x) - 1}{h_t(x) \int_0^1 1/h_s(x) ds}$$

and we see that for all $t \in [0, 1]$ and for $|\mu|$ -a.e $x \in \mathbb{R}^n$:

$$\alpha_t^{\bar{\eta}}(x) = 1 + \frac{1}{2} \int_0^t \bar{\eta}_s(x) ds = 1 + (\tilde{\alpha}_1^{\eta_0}(x) - 1) \frac{\int_0^t 1/h_s(x) ds}{\int_0^1 1/h_s(x) ds} \ge 0$$



where the last inequality follows from the fact that $h_t(x) > 0$ for all $x \in \mathbb{R}^n$ and $t \in [0, 1]$ giving $\int_0^t 1/h_s(x)ds \le \int_0^1 1/h_s(x)ds$ as well as $\tilde{\alpha}_1^{\eta_0}(x) \ge 0$ for $|\mu|$ -a.e $x \in \mathbb{R}^n$. Therefore $\bar{\eta}$ is indeed the unique solution of the problem with fixed deformation field v.

Proof of the Theorem Let $(v^m, \tilde{\eta}^m)$ be a minimizing sequence in $L^2([0,1], V \times L^2(|\mu|))$. We have by definition that for all $m \in \mathbb{N}$, $(\varphi_1^{v^m}, (\tilde{\alpha}_1^{\eta^m})^2) \cdot \mu = \mu'$. In addition, since the sequence $C_{\mu}(v^m, \eta^m)$ is bounded, we obtain in particular that (v^m) is a bounded sequence in $L^2([0,1], V)$ from which we deduce that, up to extracting a subsequence, there exists $v^* \in L^2([0,1], V)$ such that v^m converges to v^* weakly in $L^2([0,1], V)$. By the result of Theorem 7.13 in [35], this implies that the mapping $\varphi_t^{v^m}$ converges to the diffeomorphism $\varphi_t^{v^*}$ and $(\varphi_t^{v^m})^{-1}$ converges to $(\varphi_t^{v^*})^{-1}$ in $\|\cdot\|_{1,\infty}$ on any compact subset of \mathbb{R}^n (and thus on $\sup(|\mu|)$), and the convergence is also uniform over $t \in [0,1]$. From this it follows in addition that there exist A > 1 such that $1/A \leq J_T \varphi_t^{v^m}(x) \leq A$ and $1/A \leq J_T (\varphi_t^{v^m})^{-1}(x) \leq A$ for all $t \in [0,1]$, $T \in \widetilde{G}_d^n$, $x \in \sup(|\mu|)$ and $m \in \mathbb{N}$, with the same bounds also holding for $J_T \varphi_t^{v^*}(x)$.

Let us write $\mu = |\mu| \otimes \nu_x$ with ν_x a probability measure on \widetilde{G}_d^n for all $x \in \mathbb{R}^n$ the disintegration of the varifold μ . For each $m \in \mathbb{N}$, we have $\mu' = (\varphi_1^{v^m})_{\#}((\widetilde{\alpha}_1^{\eta^m})^2\mu) = (\widetilde{\alpha}_1^{\eta^m})^2 \circ (\varphi_1^{v^m})^{-1}.(\varphi_1^{v^m})_{\#} \mu$ with $\widetilde{\alpha}_1^{\eta^m}(x) \geq 0$ for $|\mu| - \text{a.e. } x \in \mathbb{R}^n$, which we can equivalently rewrite as $(\varphi_1^{v^m})_{\#}^{-1} \mu' = (\alpha_1^{\widetilde{\eta}^m})^2 \mu$. This shows that the disintegration of $(\varphi_1^{v^m})_{\#}^{-1} \mu'$ must take the form $(\varphi_1^{v^m})_{\#}^{-1} \mu' = |(\varphi_1^{v^m})_{\#}^{-1} \mu'| \otimes \nu_x$. Also, for any measurable subset $B \subseteq \mathbb{R}^n \times \widetilde{G}_d^n$, we have:

$$(\varphi_1^{v^m})_{\#}^{-1}\mu'(B) = \int_{\mathbb{R}^n \times \widetilde{G}_{J}^n} \mathbb{1}|_{B}((\varphi_1^{v^m})^{-1}(x), d_x(\varphi_1^{v^m})^{-1} \cdot T) J_T(\varphi_1^{v^m})^{-1}(x) d\mu'(x, T)$$

and by uniform convergence of $(\varphi_1^{v^m})^{-1}$ to $(\varphi_1^{v^*})^{-1}$ and $d_x(\varphi_1^{v^m})^{-1}$ to $d_x(\varphi_1^{v^*})^{-1}$ on the compact supp($|\mu'|$), we obtain by applying Lebesgue's dominated convergence theorem:

$$\begin{split} &(\varphi_1^{v^m})_{\#}^{-1}\mu'(B) \xrightarrow[m \to \infty]{} \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \mathbb{1}|_B((\varphi_1^{v^*})^{-1}(x), d_x(\varphi_1^{v^*})^{-1} \cdot T) J_T(\varphi_1^{v^*})^{-1}(x) d\mu'(x, T) \\ &= (\varphi_1^{v^*})_{\#}^{-1}\mu'(B) \end{split}$$

On the other hand, let $g(x) \doteq \int_{\widetilde{G}_d^n} \mathbb{1}|_{B}(x,T) d\nu_x(T)$. We obtain from the bounded convergence theorem that

$$\begin{split} (\varphi_1^{v^m})_{\#}^{-1}\mu'(B) &= |(\varphi_1^{v^m})_{\#}^{-1}\mu'| \otimes \nu_x(B) = \int_{\mathbb{R}^n \times \widetilde{G}_d^n} g(x)d(\varphi_1^{v^m})_{\#}^{-1}\mu'(x,T) \\ & \xrightarrow[m \to \infty]{} \int_{\mathbb{R}^n \times \widetilde{G}_d^n} g((\varphi_1^{v^*})^{-1}(x)) J_T(\varphi_1^{v^*})^{-1}(x)d\mu'(x,T) \\ &= \int_{\mathbb{R}^n} g(x)d|(\varphi_1^{v^*})_{\#}^{-1}\mu'|(x) = |(\varphi_1^{v^*})_{\#}^{-1}\mu'| \otimes \nu_x(B). \end{split}$$

It results in particular that $(\varphi_1^{v^*})_\#^{-1}\mu' = |(\varphi_1^{v^*})_\#^{-1}\mu'| \otimes \nu_x$ and that for any measurable $E \subseteq \mathbb{R}^n$, $|(\varphi_1^{v^m})_\#^{-1}\mu'|(E) \to |(\varphi_1^{v^*})_\#^{-1}\mu'|(E)$. Now assume that $E \subseteq \mathbb{R}^n$ is a measurable subset such that $|\mu|(E) = 0$. Then, for all $m \in \mathbb{N}$, using again the equality $(\varphi_1^{v^m})_\#^{-1}\mu' = (\alpha_1^{\bar{\eta}^m})^2\mu$

$$|(\varphi_1^{v^m})_{\#}^{-1}\mu'|(E) = \int_E (\alpha_1^{\tilde{\eta}^m})^2(x)d|\mu|(x) = 0.$$



165 Page 30 of 41 H.-W. Hsieh, N. Charon

Therefore we get $|(\varphi_1^{v^*})_\#^{-1}\mu'|(E)=0$ from which we deduce that $|(\varphi_1^{v^*})_\#^{-1}\mu'|$ is absolutely continuous with respect to $|\mu|$. By Radon-Nykodym's theorem, there exists $\check{\alpha}^*\in L^1(|\mu|)$ such that $|(\varphi_1^{v^*})_\#^{-1}\mu'|=\check{\alpha}^*|\mu|$ and thus $(\varphi_1^{v^*})_\#^{-1}\mu'=\check{\alpha}^*|\mu|\otimes\nu_x=\check{\alpha}^*\mu$ which leads to $\mu'=(\varphi_1^{v^*})_\#(\check{\alpha}^*\mu)$. Setting for all $t\in[0,1],\ \eta_t(x)=2(\sqrt{\check{\alpha}^*}-1),$ we get $\eta\in L^2([0,1],L^2(|\mu|)),\alpha_t^\eta(x)\geq 0$ for $|\mu|$ -a.e. $x\in\mathbb{R}^n$ and the previous equality is equivalent to $\mu'=(\varphi_1^{v^*})_\#((\check{\alpha}_1^\eta)^2\mu)=(\varphi_1^{v^*},(\alpha_1^\eta)^2)\cdot\mu$.

It follows that we can apply Lemma 1 and thus find $\bar{\eta}^* \in L^2([0, 1], L^2(|\mu|))$ such that $\bar{\eta}^*$ minimizes:

$$\int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_t^n} \eta_t^2(x) J_T \varphi_t^{v^*}(x) d\mu(x, T) dt$$

among all $\eta \in L^2(|\mu|)$ such that $\tilde{\alpha}_t^{\eta}(x) \geq 0$ and $(\varphi_1^{v^*}, (\tilde{\alpha}_1^{\eta})^2) \cdot \mu = \mu'$. Therefore we have on the one hand $(\varphi_1^{v^*}, (\tilde{\alpha}_1^{\bar{\eta}^*})^2) \cdot \mu = \mu'$ and by construction, for all $m \in \mathbb{N}$,

$$\int_{0}^{1} \int_{\mathbb{R}^{n} \times \widetilde{G}_{d}^{n}} \bar{\eta}_{t}^{*}(x)^{2} J_{T} \varphi_{t}^{v^{*}}(x) \, d\mu(x, T) dt \leq \int_{0}^{1} \int_{\mathbb{R}^{n} \times \widetilde{G}_{d}^{n}} \eta_{t}^{m}(x)^{2} J_{T} \varphi_{t}^{v^{*}}(x) \, d\mu(x, T) dt$$
(B1)

Moreover since $\varphi_t^{v^m}$ converges to $\varphi_t^{v^*}$ in $\|\cdot\|_{1,\infty}$ uniformly over $t\in[0,1]$ and on a compact set that contains $\operatorname{supp}(|\mu|)$, for any $\varepsilon>0$, there exists $p\in\mathbb{N}$ such that for all $m\geq p$, $t\in[0,1], T\in\widetilde{G}_d^n$ and $x\in\operatorname{supp}(|\mu|)$, we have $|J_T\varphi_t^{v^m}(x)-J_T\varphi_t^{v^*}(x)|\leq\varepsilon$ which also leads to $|J_T\varphi_t^{v^m}(x)-J_T\varphi_t^{v^*}(x)|\leq A\varepsilon J_T\varphi_t^{v^m}(x)$. Going back to (B1), we obtain:

$$\begin{split} &\int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \widetilde{\eta}_t^*(x)^2 J_T \varphi_t^{v^*}(x) \, d\mu(x,T) dt \\ &\leq \int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \eta_t^m(x)^2 J_T \varphi_t^{v^m}(x) \, d\mu(x,T) dt \\ &+ \int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \eta_t^m(x)^2 \left[J_T \varphi_t^{v^*}(x) - J_T \varphi_t^{v^m}(x) \right] d\mu(x,T) dt \\ &\leq (1 + A\varepsilon) \int_0^1 \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \widetilde{\eta}_t^m(x)^2 J_T \varphi_t^{v^m}(x) \, d\mu(x,T) dt \end{split}$$

As this holds for all $\varepsilon > 0$, we obtain that:

$$\int_{0}^{1} \int_{\mathbb{R}^{n} \times \widetilde{G}_{d}^{n}} \bar{\eta}_{t}^{*}(x)^{2} J_{T} \varphi_{t}^{v^{*}}(x) d\mu(x, T) dt$$

$$\leq \lim \inf_{m \to +\infty} \int_{0}^{1} \int_{\mathbb{R}^{n} \times \widetilde{G}_{d}^{n}} \eta_{t}^{m}(x)^{2} J_{T} \varphi_{t}^{v^{m}}(x) d\mu(x, T) dt$$

which, combined with the weak lower semicontinuity of the squared Hilbert norm $v \mapsto \int_0^1 \|v_t\|_V^2 dt$, leads to:

$$C_{\mu}(v^*, \bar{\eta}^*) \leq \lim \inf_{m \to +\infty} C_{\mu}(v^m, \eta^m) = d_{\mathcal{D}-\mathcal{FR}}(\mu, \mu').$$

In conclusion, $(v^*, \bar{\eta}^*)$ is a minimizer of the energy defining the distance between μ and μ' .



Appendix C Proof of Propositions 2 and 3

As we know from Sect. 2.4 that solutions to the geodesic boundary value problem between the two Diracs exist, we will derive necessary conditions satisfied by the solutions summoning the Pontryagin maximum principle (PMP) of optimal control [46] (more specifically the version derived in [47] for the type of infinite-dimensional problems considered here).

Proof of Proposition 2 We start with the 0-varifold case. The optimal control problem here simplifies to the minimization of:

$$C(v, \eta) = \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_0^1 \eta_t^2 r_0 dt$$

over $(v, \eta) \in L^2([0, 1], V \times \mathbb{R})$ subject to the state equations:

$$\begin{cases} \dot{x}(t) = v_t(x(t)) \\ \dot{\tilde{\alpha}}(t) = \frac{1}{2}\eta_t \end{cases}$$

and the boundary conditions $x(0) = x_0$, $\tilde{\alpha}(0) = 1$ and $x(1) = x_1$, $\tilde{\alpha}(1) = \sqrt{r_1/r_0}$. Let (v, η) be a solution. The Pontryagin maximum principle then states the existence of the costate functions $t \mapsto p_x(t) \in \mathbb{R}^n$ and $t \mapsto p_{\tilde{\alpha}}(t) \in \mathbb{R}$ that satisfy the adjoint equations:

$$\begin{cases} \dot{p}_x(t) = -(d_{x(t)}v_t)^T p_x(t) \\ \dot{p}_{\tilde{\alpha}}(t) = 0 \end{cases}$$

from which we immediately deduce that $p_{\tilde{\alpha}}$ is constant. Furthermore, the Hamiltonian of the system is here given by

$$H(x, \tilde{\alpha}, p_x, p_{\tilde{\alpha}}, v, \eta) = p_x^T v(x) + \frac{1}{2} p_{\tilde{\alpha}} \eta - \frac{1}{2} \|v\|_V^2 - \frac{\gamma}{2} r_0 \eta^2$$

and so the optimality conditions of the PMP that write $(v_t, \eta_t) = \operatorname{argmin}_{(v', \eta') \in V \times \mathbb{R}} H(x(t), \tilde{\alpha}(t), p_x(t), p_{\tilde{\alpha}}(t), v', \eta')$ become here $v_t(\cdot) = K_V(x(t), \cdot) p_x(t)$ (c.f. [47] for details on this derivation) and $\eta_t = \frac{p_{\tilde{\alpha}}}{2\gamma r_0}$. From this and the reproducing kernel formula, we first deduce that $\|v_t\|_V^2 = p_x(t)^T K_V(x(t), x(t)) p_x(t)$. Moreover, we get $\dot{x}(t) = K_V(x(t), x(t)) p_x(t)$ and so $p_x(t) = K_V(x(t), x(t))^{-1} \dot{x}(t)$. It follows that the first term in the cost function is equal to

$$\int_0^1 \dot{x}(t)^T G_{x(t)} \dot{x}(t) dt$$

with $G_{x(t)} = K_V(x(t), x(t))^{-1}$. Therefore, as the second term of the cost is independent of v and thus $\dot{x}(t)$, the path $t \mapsto x(t)$ minimizes the above energy subject to the boundary constraints $x(0) = x_0$ and $x(1) = x_1$ which means precisely that x(t) follows the geodesic from x_0 to x_1 for the Riemannian metric on \mathbb{R}^n given by the field of positive definite matrices $G_x = K_V(x, x)^{-1}$. On the other hand, we have:

$$\dot{\tilde{\alpha}}(t) = \frac{1}{2}\eta_t = \frac{p_{\tilde{\alpha}}}{4\gamma r_0}$$

which, together with the boundary conditions $\tilde{\alpha}(0) = 1$ and $\tilde{\alpha}(1) = \sqrt{r_1/r_0}$, leads to $\tilde{\alpha}(t) = (1-t) + t\sqrt{r_1/r_0}$ and so the varifold weight r(t) is:

$$r(t) = \tilde{\alpha}(t)^2 r_0 = ((1-t)\sqrt{r_0} + t\sqrt{r_1})^2.$$



165 Page 32 of 41 H.-W. Hsieh, N. Charon

Proof of Proposition 3 Let us now move on to the 1-varifold case. For concision, we will write again $\tilde{\alpha}(t) = \tilde{\alpha}_t^{\eta}$ and define the auxiliary variable $\tilde{u}(t) = d_{x(t)} \varphi_t^{v}(r_0 u_0)$ so that $u(t) = \frac{\tilde{u}(t)}{|\tilde{u}(t)|}$ and $r(t) = \tilde{\alpha}(t)^2 |\tilde{u}(t)|$. The cost to minimize is then:

$$C(v, \eta) = \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \frac{\gamma}{2} \int_0^1 \eta_t^2 |\tilde{u}(t)| dt$$

with the state equations:

$$\begin{cases} \dot{x}(t) = v_t(x(t)) \\ \dot{\tilde{u}}(t) = d_{x(t)}v_t(\tilde{u}(t)) \\ \dot{\tilde{\alpha}}(t) = \frac{1}{2}\eta_t \end{cases}$$

and the boundary conditions $x(0) = x_0$, $x(1) = x_1$, $\tilde{u}(0) = r_0 u_0$, $\tilde{\alpha}(0) = 1$, $\frac{\tilde{u}(1)}{|\tilde{u}(1)|} = u(1)$ and $\tilde{\alpha}(1)^2 |\tilde{u}(1)| = r_1$. The Hamiltonian for this optimal control problem is now:

$$H(x,\tilde{u},\tilde{\alpha},p_x,p_{\tilde{u}},p_{\tilde{\alpha}},v,\eta) = p_x^T v(x) + p_{\tilde{u}}^T d_x v(\tilde{u}) + \frac{1}{2} p_{\tilde{\alpha}} \eta - \frac{1}{2} \|v\|_V^2 - \frac{\gamma}{2} \eta^2 |\tilde{u}|$$

from which we get the following adjoint equations:

$$\begin{cases} \dot{p}_{x}(t) = -(d_{x(t)}v_{t})^{T} p_{x}(t) - d_{x(t)}^{2} v_{t}(\tilde{u}(t), \cdot)^{T} p_{\tilde{u}}(t) \\ \dot{p}_{\tilde{u}}(t) = -(d_{x(t)}v_{t})^{T} p_{\tilde{u}}(t) + \frac{\gamma}{2} \eta_{t}^{2} \frac{\tilde{u}(t)}{|\tilde{u}(t)|} \\ \dot{p}_{\tilde{u}}(t) = 0 \end{cases}$$
(C1)

where $d_x^2 v(u, \cdot)$ denotes the matrix of the linear mapping $h \in \mathbb{R}^n \mapsto d_x^2 v(u, h) \in \mathbb{R}^n$. Since $p_{\tilde{\alpha}}(t)$ is constant from the last equation above, we will simply write $p_{\tilde{\alpha}}$ in what follows. The optimality conditions of the PMP lead to the following expressions for the optimal controls:

$$\begin{cases} v_t(\cdot) = K_V(x(t), \cdot) p_x(t) + \partial_1 K_V(x(t), \cdot) (\tilde{u}(t), p_{\tilde{u}}(t)) \\ \eta_t = \frac{p_{\tilde{u}}}{2\gamma |\tilde{u}(t)|} \end{cases}$$
(C2)

Now, using the expression of the kernel $K_V(x, y) = \rho\left(\frac{|x-y|^2}{\sigma^2}\right)$, the first equation in (C2) gives for all $x \in \mathbb{R}^n$:

$$v_t(x) = \rho \left(\frac{|x - x(t)|^2}{\sigma^2} \right) p_x(t) - \frac{2}{\sigma^2} \rho' \left(\frac{|x - x(t)|^2}{\sigma^2} \right) [(x - x(t))^T \tilde{u}(t)] p_{\tilde{u}}(t).$$
 (C3)

In particular, we get that $\dot{x}(t) = v_t(x(t)) = p_x(t)$. Furthermore, differentiating the above equation at x = x(t), we find that for all $h \in \mathbb{R}^n$:

$$\begin{aligned} d_{x(t)}v_t(h) &= \frac{1}{\tau} \left[\tilde{u}(t)^T h \right] p_{\tilde{u}}(t) \\ d_{x(t)}v_t^T(h) &= \frac{1}{\tau} \left[p_{\tilde{u}}(t)^T h \right] \tilde{u}(t) \\ d_{x(t)}^2 v_t(\cdot, \bar{u}(t))^T h &= -\frac{1}{\tau} (p_x(t)^T h) \tilde{u}(t) \end{aligned}$$



where we remind the reader that $\tau = -\frac{\sigma^2}{2\rho'(0)} > 0$ and thus

$$\dot{p}_{x}(t) = -\frac{1}{\tau} \left[p_{\tilde{u}}(t)^{T} p_{x}(t) \right] \tilde{u}(t) + \frac{1}{\tau} (p_{x}(t)^{T} p_{\tilde{u}}(t)) \tilde{u}(t) = 0.$$

Therefore $p_x(t)$ is constant and $\ddot{x}(t) = 0$ which, with the two boundary conditions, leads to $p_x = x_1 - x_0$ and $x(t) = (1 - t)x_0 + tx_1$.

We next analyze the behaviour of the direction u(t). Rewriting the state and adjoint equations based on (C1) and (C2), we have:

$$\begin{cases} \dot{\tilde{u}}(t) = \frac{1}{\tau} |\tilde{u}(t)|^2 p_{\tilde{u}}(t) \\ \dot{p}_{\tilde{u}}(t) = -\frac{1}{\tau} |p_{\tilde{u}}(t)|^2 + \frac{p_{\tilde{u}}^2}{8\gamma} \frac{\tilde{u}(t)}{|\tilde{u}(t)|^3} \end{cases}$$

Going back to $u(t) = \frac{\tilde{u}(t)}{|\tilde{u}(t)|}$, one has:

$$\begin{split} \dot{u}(t) &= \frac{\dot{\tilde{u}}(t)}{|\tilde{u}(t)|} - \frac{1}{|\tilde{u}(t)|^3} \left[\dot{\tilde{u}}(t)^T \tilde{u}(t) \right] \tilde{u}(t) \\ &= \frac{1}{\tau} |\tilde{u}(t)| \left(p_{\tilde{u}}(t) - \left[u(t)^T p_{\tilde{u}}(t) \right] u(t) \right) \\ &= \frac{1}{\tau} |\tilde{u}(t)| \mathrm{Proj}_{u(t)^{\perp}}(p_{\tilde{u}}(t)) \end{split}$$

with $\operatorname{Proj}_{u(t)^{\perp}}$ denoting the orthogonal projector onto the hyperplane normal to u(t). Letting $p_u(t) \doteq |\tilde{u}(t)| \operatorname{Proj}_{u(t)^{\perp}}(p_{\tilde{u}}(t))$, we find after calculations that $\dot{p}_u(t) = -\frac{1}{\tau} |p_u(t)|^2 u(t)$. Therefore, we have the following coupled system of ODEs on $(u(t), p_u(t))$:

$$\begin{cases} \dot{u}(t) = \frac{1}{\tau} p_u(t) \\ \dot{p}_u(t) = -\frac{1}{\tau} |p_u(t)|^2 u(t) \end{cases}$$
 (C4)

As $p_u(t)$ is orthogonal to u(t) by definition, the above equations imply in particular that $|p_u(t)|^2 = \kappa^2$ is constant. It also follows immediately that for all t, $(u(t), p_u(t))$ stay in the plane spanned by $\{u(0), p_u(0)\}$. Let us therefore identify the plane $\mathrm{Span}\{u(0), p_u(0)\}$ with \mathbb{C} , choosing without loss of generality u(0) = 1. We can then write $u(t) = e^{i\theta(t)}$ with $\theta(0) = 0$, $\theta(1) = \theta_{0,1}$ the angle between u_0 and u_1 and $p_u(t) = \pm i\kappa e^{i\theta(t)}$. With this identification, (C4) leads to $\dot{\theta}(t) = \pm \frac{\kappa}{\tau}$, in other words the direction $\bar{u}(t)$ rotates with constant angular velocity. With the boundary conditions, it shows that $\kappa = \tau \theta_{0,1}$ and thus $\theta(t) = t\theta_{0,1}$. Eventually, assuming $u_1 \neq u_0$, this leads to the expression of u(t):

$$u(t) = \begin{cases} \frac{1}{\sin(\theta_{0,1})} (\sin((1-t)\theta_{0,1})u_0 + \sin(t\theta_{0,1})u_1) & \text{if } \theta_{0,1} \neq 0 \\ u_0(=u_1) & \text{if } \theta_{0,1} = 0 \end{cases}$$
 (C5)

with $\theta_{0,1} = \arccos(u_0^T u_1)$. Note that when $u_1 = -u_0$, the geodesic is not unique as one can rotate from u_0 to u_1 in infinitely many ways.

We are left with determining $r(t) = \tilde{\alpha}(t)^2 |\tilde{u}(t)|$. Let us start by introducing the auxiliary function $m(t) = \sqrt{r(t)} = \tilde{\alpha}(t)\sqrt{|\tilde{u}(t)|}$. Using the state and costate equations, we first see that $\dot{m}(t) = \frac{p_{\tilde{u}}}{4\gamma\sqrt{|\tilde{u}(t)|}} + \frac{1}{2\tau}m(t)(\tilde{u}(t)^Tp_{\tilde{u}}(t))$. Differentiating a second time, we get:



165 Page 34 of 41 H.-W. Hsieh, N. Charon

$$\begin{split} \ddot{m}(t) &= -\frac{p_{\tilde{u}}}{8\tau\gamma\sqrt{|\tilde{u}(t)|}}(\tilde{u}(t)^T p_{\tilde{u}}(t)) + \frac{p_{\tilde{u}}}{8\tau\gamma\sqrt{|\tilde{u}(t)|}}(\tilde{u}(t)^T p_{\tilde{u}}(t)) \\ &+ \frac{1}{4\tau^2}(\tilde{u}(t)^T p_{\tilde{u}}(t))^2 m(t) + \frac{1}{2\tau^2}|\tilde{u}(t)|^2|p_{\tilde{u}}(t)|^2 m(t) \\ &- \frac{1}{2\tau^2}|\tilde{u}(t)|^2|p_{\tilde{u}}(t)|^2 m(t) + \frac{p_{\tilde{u}}^2}{16\tau\gamma}\frac{m(t)}{|\tilde{u}(t)|} \\ &= \frac{1}{4\tau}\left[\frac{1}{\tau}(\tilde{u}(t)^T p_{\tilde{u}}(t))^2 + \frac{p_{\tilde{u}}^2}{4\gamma|\tilde{u}(t)|}\right] m(t). \end{split}$$

Moreover, it turns out that the term inside brackets is constant in time. Indeed, on the one hand we have $\frac{d}{dt}\tilde{u}(t)^T p_{\tilde{u}}(t) = \frac{p_{\tilde{u}}^2}{8\sqrt{|\tilde{u}(t)|}}$ so that:

$$\frac{d}{dt}(\tilde{u}(t)^T p_{\tilde{u}}(t))^2 = \frac{p_{\tilde{u}}^2}{4\gamma |\tilde{u}(t)|} (\tilde{u}(t)^T p_{\tilde{u}}(t))$$

and on the other:

$$\frac{d}{dt}\frac{p_{\tilde{\alpha}}^2}{4\gamma|\tilde{u}(t)|} = -\frac{p_{\tilde{\alpha}}^2}{4\gamma|\tilde{u}(t)|^3}(\tilde{u}(t)^T\dot{\tilde{u}}(t)) = \frac{p_{\tilde{\alpha}}^2}{4\gamma|\tilde{u}(t)|}(\tilde{u}(t)^Tp_{\tilde{u}}(t)).$$

Let us therefore set $v^2 \doteq \frac{1}{4\tau} \left[\frac{1}{\tau} (\tilde{u}(0)^T p_{\tilde{u}}(0))^2 + \frac{p_{\tilde{u}}^2}{4\gamma |\tilde{u}(0)|} \right] > 0$ so that $\ddot{m}(t) = v^2 m(t)$ and therefore $r(t) = m(t)^2$ takes the form $r(t) = C_0 \sinh(C_1 + vt)^2$ for some constants $C_0, C_1 \in \mathbb{R}$. Based on the two boundary conditions for r(t) and hyperbolic trigonometry identities, we find eventually that:

$$r(t) = \left(\frac{\sqrt{r_0}\sinh((1-t)\nu) + \sqrt{r_1}\sinh(\nu t)}{\sinh(\nu)}\right)^2.$$
 (C6)

To recover the explicit expression of ν , let us first express $\dot{r}(0)$ based on (C6). We obtain

$$\dot{r}(0) = 2\nu r_0 \left(\sqrt{\frac{r_1}{r_0}} \frac{1}{\sinh(\nu)} - \coth(\nu) \right). \tag{C7}$$

On the other hand, $\dot{r}(0) = 2m(0)\dot{m}(0)$ and using the expression of $\dot{m}(t)$ obtained earlier, we get

$$\dot{r}(0) = 2\sqrt{r_0} \left(\frac{p_{\tilde{u}}}{4\gamma\sqrt{r_0}} + \frac{1}{2\tau} \sqrt{r_0} (\tilde{u}(0)^T p_{\tilde{u}}(0)) \right)$$
 (C8)

Furthermore, the term $(\tilde{u}(0)^T p_{\tilde{u}}(0))$ may be expressed with respect to $p_{\tilde{\alpha}}$. Indeed, consider the function $h(t) = \frac{\tilde{\alpha}(t)p_{\tilde{\alpha}}}{2} - (\tilde{u}(t)^T p_{\tilde{u}}(t))$. We see that:

$$\dot{h}(t) = \frac{p_{\tilde{\alpha}}^2}{8\gamma |\tilde{u}(t)|} - \frac{p_{\tilde{\alpha}}^2}{8\gamma |\tilde{u}(t)|} = 0$$

and thus h(t) = h(0) = h(1). Now one of the terminal constraint for the optimal control problem is $\tilde{\alpha}(1)^2 |\tilde{u}(1)| = r_1$ and therefore the transversality condition of the PMP yields that the vector $(p_{\tilde{u}}(1), p_{\tilde{\alpha}})^T$ is parallel to $\left(\tilde{\alpha}(1)^2 \frac{\tilde{u}(1)}{|\tilde{u}(1)|}, 2\tilde{\alpha}(1)|\tilde{u}(1)|\right)^T$ from which we deduce that $\tilde{u}(1)^T p_{\tilde{u}}(1) = \frac{\tilde{\alpha}(1)p_{\tilde{u}}}{2}$, in other words h(1) = h(0) = 0. Since $\tilde{\alpha}(0) = 1$, this implies that $\tilde{u}(0)^T p_{\tilde{u}}(0) = \frac{p_{\tilde{u}}}{2}$ from which we can rewrite v^2 as:



$$v^{2} = \left(\frac{p_{\tilde{\alpha}}}{4\tau}\right)^{2} \left[1 + \frac{\tau}{\gamma r_{0}}\right] \implies \frac{p_{\tilde{\alpha}}}{\nu} = \pm \frac{4\tau}{\sqrt{1 + \frac{\tau}{\gamma r_{0}}}}.$$
 (C9)

Moreover, (C8) becomes

$$\dot{r}(0) = \frac{p_{\tilde{\alpha}}}{2\gamma} \left(1 + \frac{\gamma}{\tau} r_0 \right)$$

which combined with (C7) and (C9) leads to the implicit equation on ν :

$$\sqrt{\frac{r_1}{r_0}} \frac{1}{\sinh(\nu)} - \coth(\nu) = \pm \underbrace{\sqrt{1 + \frac{\tau}{\gamma r_0}}}_{=\gamma}.$$
 (C10)

A quick analysis of the function on the left hand side shows that this equation only has a solution with $+\chi$ on the right hand side when $r_1 > r_0$ and with $-\chi$ when $r_1 \le r_0$. Using Mathematica, we find specifically the explicit expressions for ν given in Proposition 3.

Finally, we can express the distance between $r_0\delta_{x_0,u_0}$ and $r_1\delta_{x_1,u_1}$. First, the kernel norm $\|v_t\|_V^2$ can be calculated based on the expressions (C2) and (C3) using the reproducing kernel property for kernel derivatives ([45] Theorem 2.11). Skipping some of the details for brevity, we obtain:

$$\|v_t\|_V^2 = |p_x|^2 + \frac{1}{\tau} |p_{\tilde{u}}(t)|^2 |\tilde{u}(t)|^2$$

$$= |x_1 - x_0|^2 + \frac{1}{\tau} \left(|\text{Proj}_{u(t)^{\perp}}(p_{\tilde{u}}(t))|^2 + (u(t)^T p_{\tilde{u}}(t))^2 \right) |\tilde{u}(t)|^2$$

$$= |x_1 - x_0|^2 + \frac{1}{\tau} |p_u(t)|^2 + \frac{1}{\tau} (\tilde{u}(t)^T p_{\tilde{u}}(t))^2$$

where we have used the fact that $p_x = x_1 - x_0$, $\tilde{u}(t) = u(t)/|u(t)|$ and $p_u(t) = |\tilde{u}(t)| \text{Proj}_{u(t)^{\perp}}(p_{\tilde{u}}(t))$. Furthermore, by the results above, we know that $|p_u(t)|^2 = \kappa^2 = (\tau \arccos(u_0^T u_1))^2$. Therefore:

$$\|v_t\|_V^2 = |x_1 - x_0|^2 + \tau \arccos(u_0^T u_1)^2 + \frac{1}{\tau} (\tilde{u}(t)^T p_{\tilde{u}}(t))^2.$$

Now for the second term in the energy, we have using (C2):

$$\frac{\gamma}{2}\eta_t^2|\tilde{u}(t)| = \frac{p_{\tilde{\alpha}}^2}{8\gamma|\tilde{u}(t)|}$$

leading to:

$$\begin{split} \frac{1}{2} \|v_t\|_V^2 + \frac{\gamma}{2} \eta_t^2 |\tilde{u}(t)| &= \frac{|x_1 - x_0|^2}{2} + \frac{\tau}{2} \arccos(u_0^T u_1)^2 \\ &+ \frac{1}{2} \left(\frac{1}{\tau} (\tilde{u}(t)^T p_{\tilde{u}}(t))^2 + \frac{p_{\tilde{u}}^2}{4\gamma |\tilde{u}(t)|} \right) \end{split}$$

Recalling that the last term is constant in time and equal to $2\tau v^2$ thanks to our earlier derivations, we finally get:



165 Page 36 of 41 H.-W. Hsieh, N. Charon

$$\begin{split} d_{V-\mathcal{FR}}(r_0\delta_{(x_0,u_0)},r_1\delta_{(x_1,u_1)})^2 &= \int_0^1 \left(\frac{1}{2}\|v_t\|_V^2 + \frac{\gamma}{2}\eta_t^2|\tilde{u}(t)|\right) dt \\ &= \frac{|x_1 - x_0|^2}{2} + \frac{\tau}{2}\arccos(u_0^T u_1)^2 + 2\tau v^2. \end{split}$$

Appendix D Proof of Theorem 3

Let us denote by $C(v,\alpha)$ the functional to minimize in (17) and let (v^j,α^j) be a minimizing sequence. Using a similar argument as in the previous existence proofs, since (v^j) is bounded in $L^2([0,1],V)$, we can assume (by extracting a subsequence if necessary) that (v^j) converges weakly to some \bar{v} in $L^2([0,1],V)$ giving that $\|(\varphi_1^{v_j}-\varphi_1^{\bar{v}})|_K\|_{1,\infty}\to 0$ as $j\to +\infty$. Also, since $\gamma>0$, (α^j) is a bounded sequence in $L^2(|\mu|)$ and thus up to extraction of another subsequence, we may assume that we have weak convergence $\alpha^j \to \bar{\alpha}$ to some $\bar{\alpha} \in L^2(|\mu|)$.

From the weak lower semicontinuity of the first two terms in the energy, we deduce:

$$\int_{0}^{1} \|\bar{v}\|_{V}^{2} dt \leq \liminf_{j \to +\infty} \int_{0}^{1} \|v^{j}\|_{V}^{2} dt \tag{D1}$$

$$\int_{\mathbb{R}^n} (\bar{\alpha}(x) - 1)^2 d|\mu|(x) \le \liminf_{j \to +\infty} \int_{\mathbb{R}^n} (\alpha^j(x) - 1)^2 d|\mu|(x) \tag{D2}$$

In addition, since $\|\varphi_1^{v_j}\|_{1,\infty}$ is bounded and μ is supported in the compact set K, there exists another compact subset $K \subset K' \subset \mathbb{R}^n \times \widetilde{G}_d^n$ such that for all $j \in \mathbb{N}$, supp $((\varphi_1^{v_j})_{\#}(\alpha^j \mu)) \subset K'$. Moreover, using the disintegration theorem on the measure μ (c.f. Proposition 1), we see that:

$$\begin{split} |(\varphi_1^{v^j})_{\#}(\alpha^j\mu)|(\mathbb{R}^n) &= \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \alpha^j(x) |J_U \varphi_1^{v^j}(x)| d\mu(x,U) \\ &= \int_{\mathbb{R}^n} \alpha^j(x) \left(\int_{\widetilde{G}_d^n} |J_U \varphi_1^{v^j}(x)| d\nu_x(U) \right) d|\mu|(x) \end{split}$$

For the same reason as above, we have that $|J_U \varphi_1^{v^j}(x)|$ is bounded uniformly over $x \in \text{supp}(|\mu|) \subset K$, $j \in \mathbb{N}$ and $U \in \widetilde{G}_d^n$ from which we get, applying the Cauchy-Schwarz inequality:

$$|(\varphi_1^{v^j})_{\#}(\alpha^j\mu)|(\mathbb{R}^n) \leq C\sqrt{\int_{\mathbb{R}^n}\alpha^j(x)^2d|\mu|(x)} = C\|\alpha^j\|_{L^2(|\mu|)}$$

for some constant C>0. Now since $\|\alpha^j\|_{L^2(|\mu|)}$ is bounded, we deduce that there exists M>0 such that $|\varphi_1^{v^j}|_{\#}(\alpha^j\mu)|(\mathbb{R}^n)\leq M$ for all $j\in\mathbb{N}$. In other words, we have obtained that for all $j,\varphi_1^{v^j}|_{\#}(\alpha^j\mu)$ belong to the space $\mathcal{V}_{d,M,K'}$ defined in Proposition 5. To show the convergence of $(\varphi_1^{v^j})_{\#}(\alpha^j\mu)$ for the W^* metric, we are thus left to show that it converges for the weak-* topology.

Thus, let $\omega \in C_c(\mathbb{R}^n \times \widetilde{G}_d^n)$. We have:

$$\left((\varphi_1^{v^j})_\#(\alpha^j\mu)|\omega\right) = \left((\varphi_1^{\bar{v}})_\#(\alpha^j\mu)|\omega\right) + \left((\varphi_1^{v^j})_\#(\alpha^j\mu) - (\varphi_1^{\bar{v}})_\#(\alpha^j\mu)|\omega\right)$$



Looking at the first term, we see that:

$$\begin{split} \left((\varphi_1^{\bar{v}})_{\#}(\alpha^j \mu) | \omega \right) &= \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \alpha^j(x) J_U \varphi_1^{\bar{v}}(x) \omega(\varphi_1^{\bar{v}}(x), d_x \varphi_1^{\bar{v}} \cdot U) d\mu(x, U) \\ &= \int_{\mathbb{R}^n} \alpha^j(x) \left(\int_{\widetilde{G}_d^n} J_U \varphi_1^{\bar{v}}(x) \omega(\varphi_1^{\bar{v}}(x), d_x \varphi_1^{\bar{v}} \cdot U) d\nu_x(U) \right) d|\mu|(x) \\ &\xrightarrow{j \to \infty} \int_{\mathbb{R}^n} \bar{\alpha}(x) \left(\int_{\widetilde{G}_d^n} J_U \varphi_1^{\bar{v}}(x) \omega(\varphi_1^{\bar{v}}(x), d_x \varphi_1^{\bar{v}} \cdot U) d\nu_x(U) \right) d|\mu|(x) \\ &= \left((\varphi_1^{\bar{v}})_{\#}(\bar{\alpha}\mu) | \omega \right) \end{split}$$

where the convergence in the third row follows from the weak convergence of α^j to $\bar{\alpha}$ in $L^2(|\mu|)$ and the fact the function between parentheses is measurable and bounded on \mathbb{R}^n and thus in $L^2(|\mu|)$. As for the second term, we can expand it and see that:

$$\begin{split} & \left| \left((\varphi_{1}^{v^{j}})_{\#}(\alpha^{j}\mu_{0}) - (\varphi_{1}^{\bar{v}})_{\#}(\alpha^{j}\mu)) | \omega \right) \right| \\ & \leq \int_{K} |\alpha^{j}(x)| \left| J_{U}\varphi_{1}^{v^{j}}(x)\omega(\varphi_{1}^{v^{j}}(x), d_{x}\varphi_{1}^{v^{j}} \cdot U) - J_{U}\varphi_{1}^{\bar{v}}(x)\omega(\varphi_{1}^{\bar{v}}(x), d_{x}\varphi_{1}^{\bar{v}} \cdot U) \right| \\ & \times d\mu(x, U) \\ & \leq \|\alpha^{j}\|_{L^{2}(|\mu|)} \cdot \left(\int_{K} \left| J_{U}\varphi_{1}^{v^{j}}(x)\omega(\varphi_{1}^{v^{j}}(x), d_{x}\varphi_{1}^{v^{j}} \cdot U) - J_{U}\varphi_{1}^{\bar{v}}(x)\omega(\varphi_{1}^{\bar{v}}(x), d_{x}\varphi_{1}^{\bar{v}} \cdot S) \right|^{2} d\mu(x, U) \right)^{1/2} \end{split}$$

Now, $\|\alpha^j\|_{L^2(|\mu|)}$ is bounded and by convergence of $\varphi_1^{v^j}$ to $\varphi_1^{\bar{v}}$ in $\|\cdot\|_{1,\infty}$ and uniform continuity of ω , we deduce that

$$\left((\varphi_1^{v^j})_{\#}(\alpha^j\mu_0)-(\varphi_1^{\bar{v}})_{\#}(\alpha^j\mu))|\omega\right)\to 0.$$

Therefore, we have shown that $(\varphi_1^{v^j})_\#(\alpha^j\mu) \stackrel{*}{\rightharpoonup} (\varphi_1^{\bar{v}})_\#(\bar{\alpha}\mu)$. By Proposition 5, this implies that $(\varphi_1^{v^j})_\#(\alpha^j\mu) \stackrel{d_{W^*}}{\longrightarrow} (\varphi_1^{\bar{v}})_\#(\bar{\alpha}\mu)$. As a result,

$$\lim_{j \to \infty} \|(\varphi_1^{v^j})_{\#}(\alpha^j \mu) - \mu'\|_{W^*}^2 = \|(\varphi_1^{\bar{v}})_{\#}(\bar{\alpha}\mu) - \mu'\|_{W^*}^2.$$
 (D3)

Combining (D1), (D2) and (D3), we obtain:

$$C(\bar{v}, \bar{\alpha}) \le \lim \inf_{j \to \infty} C(v^j, \alpha^j)$$

and thus $(\bar{v}, \bar{\alpha})$ is a minimizer of C.

Appendix E Proof of Theorem 4

Let us again denote by $C(v, \eta)$ the functional to minimize in (21) and let (v^j, η^j) be a minimizing sequence. Since $\mu = \sum_{i=1}^N r_i \delta_{(x_i, U_i)}$ is here assumed to be a discrete varifold, we can equivalently view the functions η^j as functions in $L^2([0, 1], \mathbb{R}^N)$ in which each component of the vector $\eta^j(t)$ is associated to the corresponding Dirac in μ . Similarly, $\tilde{\alpha}_1^{\eta^j}$



165 Page 38 of 41 H.-W. Hsieh, N. Charon

can be viewed as a vector in \mathbb{R}^N_+ and to simplify the following derivations, we shall write with a slight abuse of notations $\tilde{\alpha}_1^{\eta^j} = \alpha^j = (\alpha_i^j)_{i=1,\dots,N}$. Then for each $j \in \mathbb{N}$, define $\mu^j = \sum_{i=1}^N r_i^j \delta_{(x_i^j, U_i^j)}$ to be the transformed varifold at t=1 for (v^j, η^j) i.e. $x_i^j = \varphi_1^{v^j}(x_i)$, $U_i^j = d_{x_i} \varphi_1^{v^j} \cdot U_i$ and $r_i^j = (\alpha_i^j)^2 J_{U_i} \varphi_1^{v^j} r_i$. We will first show that, modulo extraction of subsequences, for each $i=1,\dots,N$ the

We will first show that, modulo extraction of subsequences, for each $i=1,\ldots,N$ the sequences (x_i^j) , (U_i^j) and (r_i^j) are converging in \mathbb{R}^n , \widetilde{G}_d^n and \mathbb{R}_+ respectively. Let K be a compact subset of \mathbb{R}^n that contains x_1,\ldots,x_N . Since $C(v^j,\eta^j)$ and thus $\int_0^1 \|v_i^j\|_V^2 dt$ is bounded uniformly in j, using similar arguments as in the previous proofs, we have, up to extraction of a subsequence, that v^j converges weakly to some $v^* \in L^2([0,1],V)$ and thus $\varphi_1^{v^j}$ converges to $\varphi_1^{v*} \in \mathrm{Diff}(\mathbb{R}^n)$ in $\|\cdot\|_{1,\infty}$ on every compact subset of \mathbb{R}^n in particular on K leading to $x_i^j \xrightarrow[j \to \infty]{} \varphi_1^{v*}(x_i) \doteq x_i^*$ and $U_i^j \xrightarrow[j \to \infty]{} d_{x_i} \varphi_1^{v*} \cdot U_i \doteq U_i^*$ for all $i=1,\ldots,N$.

Now we also have that $\|\mu^j - \mu^{tar}\|_{W^*}$ is bounded uniformly in j and furthermore:

$$\begin{split} \|\mu^{j} - \mu^{tar}\|_{W^{*}} &\geq \|\mu^{j}\|_{W^{*}} - \|\mu^{tar}\|_{W^{*}} \\ &= \sqrt{\sum_{i,i'=1}^{N} r_{i}^{j} r_{i'}^{j} k(x_{i}^{j}, U_{i}^{j}, x_{i'}^{j}, U_{i'}^{j})} - \|\mu^{tar}\|_{W^{*}} \\ &\geq \sqrt{\sum_{i=1}^{N} (r_{i}^{j})^{2} k(x_{i}^{j}, U_{i}^{j}, x_{i}^{j}, U_{i}^{j})} - \|\mu^{tar}\|_{W^{*}} \end{split}$$

where the last equality follows from the fact that $k(x_i^j, U_i^j, x_i^j, U_{i'}^j) \geq 0$ by assumption on the kernel of W. Moreover, $(x, U) \mapsto k(x, U, x, U)$ is strictly positive, continuous and we can find a compact subset $K' \subset \mathbb{R}^n$ with (x_i^j, U_i^j) belonging to the compact $K' \times \widetilde{G}_d^n$ for all $i = 1, \ldots, N$ and $j \in \mathbb{N}$. Consequently, taking $\delta = \min\{k(x, U, x, U) \mid (x, U) \in K' \times \widetilde{G}_d^n\} > 0$ we have $k(x_i^j, U_i^j, x_i^j, U_i^j) \geq \delta$ for all i and j and therefore we must have that each sequence $(r_i^j)_{j \in \mathbb{N}}$ is bounded since otherwise we could find a subsequence making the right hand side in the above inequality go to infinity. Thus, up to extracting once again a subsequence, we have $r_i^j \xrightarrow[i \to \infty]{} r_i^*$ with $r_i^* \geq 0$ for all $i = 1, \ldots, N$.

Let us now set $\mu^* = \sum_{i=1}^N r_i^* \delta_{(x_i^*, T_i^*)}$ and define $\eta^* \in L^2([0, 1], \mathbb{R}^N)$ as given by Lemma 1 which means here specifically:

$$\eta_i^*(t) = 2 \frac{\sqrt{\frac{r_i^*}{r_i h_i(1)}} - 1}{h_i(t) \int_0^1 1/h_i(s) ds}$$

where $h_i(t) = J_{U_i} \varphi_t^{v^*}(x_i)$. Then it is easy to check that $\mu^* = (\varphi_1^{v^*}, (\alpha_1^{\eta^*})^2) \cdot \mu$ and thus $\mu^* \in \bar{\Theta}(\mu)$. Then by the same argument as in the proof of Theorem 2 (c.f. Appendix 1), we get that:

$$\frac{1}{2} \int_{0}^{1} \|v_{t}^{*}\|_{V}^{2} dt + \frac{\gamma}{2} \sum_{i=1}^{N} \int_{0}^{1} \eta_{i}^{*}(t)^{2} J_{U_{i}} \varphi_{t}^{v^{*}}(x_{i}) r_{i} dt$$

$$\leq \lim \inf_{j \to \infty} \left(\frac{1}{2} \int_{0}^{1} \|v_{t}^{j}\|_{V}^{2} dt + \frac{\gamma}{2} \sum_{i=1}^{N} \int_{0}^{1} \eta_{i}^{j}(t)^{2} J_{U_{i}} \varphi_{t}^{v^{j}}(x_{i}) r_{i} dt \right).$$



Thus it only remains to examine the convergence of the kernel fidelity term for which it is enough to show that the sequence (μ^j) converges to μ^* for $\|\cdot\|_{W^*}$. Let $\omega \in W$ such that $\|\omega\|_W \leq 1$. Given the continuous embedding assumption, we also have $\|\omega\|_{1,\infty} \leq c_W \|\omega\|_W \leq c_W$ for some constant $c_W > 0$. Then

$$\begin{split} \left| \left(\mu^{j} - \mu^{*} | \omega \right) \right| &= \left| \sum_{i=1}^{N} \left(\omega(x_{i}^{j}, U_{i}^{j}) r_{i}^{j} - \omega(x_{i}^{*}, U_{i}^{*}) r_{i}^{*} \right) \right| \\ &\leq \sum_{i=1}^{N} \left(\left| \omega(x_{i}^{j}, U_{i}^{j}) (r_{i}^{j} - r_{i}^{*}) \right| + \left| (\omega(x_{i}^{j}, U_{i}^{j}) - \omega(x_{i}^{*}, U_{i}^{*})) r_{i}^{*} \right| \right) \\ &\leq \sum_{i=1}^{N} \left(\| \omega \|_{\infty} | r_{i}^{j} - r_{i}^{*}| + r_{i}^{*} \| \omega \|_{1,\infty} \max\{ |x_{i}^{j} - x_{i}^{*}|, |U_{i}^{j} - U_{i}^{*}| \} \right) \\ &\leq 2N \| \omega \|_{1,\infty} \max_{i=1,\dots,N} \{ |x_{i}^{j} - x_{i}^{*}|, |U_{i}^{j} - U_{i}^{*}|, |r_{i}^{j} - r_{i}^{*}| \} \\ &\leq 2N c_{W} \max_{i=1} \sum_{n=1}^{N} \{ |x_{i}^{m} - x_{i}^{*}|, |U_{i}^{j} - U_{i}^{*}|, |r_{i}^{j} - r_{i}^{*}| \} \end{split}$$

Therefore, we see that $\|\mu^j - \mu^*\|_{W^*} = \sup_{\|\omega\|_W \le 1} \left| \left(\mu^j - \mu^* | \omega \right) \right| \xrightarrow[j \to \infty]{} 0$ since $x_i^j \to x_i^*$, $U_i^j \to U_i^*$ and $r_i^j \to r_i^*$.

Finally, combining the previous two estimates, we obtain:

$$C(v^*, \eta^*) \le \lim \inf_{j \to +\infty} C(v^j, \eta^j)$$

which shows that (v^*, η^*) is indeed a minimizer of (21).

References

- Grenander, U.: General Pattern Theory: A Mathematical Study of Regular Structures. Clarendon Press, Oxford (1993)
- Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. IEEE Trans. Image Process. 5(10), 1435–1447 (1996)
- 3. Michor, P., Mumford, D.: An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. Appl. Comput. Harmon. Anal. 23(1), 74–113 (2007)
- 4. Bauer, M., Harms, P., Michor, P.: Sobolev metrics on shape space of surfaces. J. Geom. Mech. 3(4), 389–438 (2011)
- Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int. J. Comput. Vision 61, 139–157 (2005)
- Joshi, S.C., Miller, M.I.: Landmark matching via large deformation diffeomorphisms. Image Processing, IEEE Transactions on 9(8), 1357–1370 (2000)
- Glaunès, J., Vaillant, M.: Surface matching via currents. Proceedings of Information Processing in Medical Imaging (IPMI), Lecture Notes in Computer Science 3565(381-392) (2006)
- Durrleman, S., Pennec, X., Trouvé, A., Ayache, N.: Statistical models of sets of curves and surfaces based on currents. Med. Image Anal. 13(5), 793–808 (2009)
- Charon, N., Trouvé, A.: The varifold representation of non-oriented shapes for diffeomorphic registration. SIAM journal of Imaging Science 6(4), 2547–2580 (2013)
- Kaltenmark, I., Charlier, B., Charon, N.: A general framework for curve and surface comparison and registration with oriented varifolds. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3346–3355 (2017)
- Hsieh, H.-W., Charon, N.: Diffeomorphic registration of discrete geometric distributions. In: Mathematics Of Shapes And Applications, pp. 45–74. World Scientific, (2020)



165 Page 40 of 41 H.-W. Hsieh, N. Charon

12. Hsieh, H.-W., Charon, N.: Metrics, quantization and registration in varifold spaces. Found. Comput. Math. 21, 1317–1361 (2021)

- Gori, P., Colliot, O., Marrakchi-Kacem, L., Worbe, Y., Fallani, F.D.V., Chavez, M., Poupon, C., Hartmann, A., Ayache, N., Durrleman, S.: Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles. IEEE Transactions on Medical Imaging PP(99), 2609–2619 (2016)
- Trouvé, A., Younes, L.: Metamorphoses through lie group action. Foundation of computational mathematics 5, 173–198 (2005)
- Holm, D., Trouvé, A., Younes, L.: The Euler-Poincaré theory of Metamorphosis. Quart. Appl. Math 67(2), 661–685 (2009)
- Richardson, C., Younes, L.: Computing metamorphoses between discrete measures. Journal of Geometric Mechanics 5(1), 131–150 (2013)
- Richardson, C., Younes, L.: Metamorphosis of images in reproducing kernel Hilbert spaces. Advances in Computational Mathematics, 42, 573–603 (2016)
- Berkels, B., Effland, A., Rumpf, M.: Time discrete geodesic paths in the space of images. SIAM J. Imag. Sci. 8(3), 1457–1488 (2015)
- Charon, N., Charlier, B., Trouvé, A.: Metamorphoses of functional shapes in Sobolev spaces. J. Foundations of Comput. Math 18(6), 1535–1596 (2018)
- Liero, M., Mielke, A., Savaré, G.: Optimal Transport in Competition with Reaction: The Hellinger-Kantorovich Distance and Geodesic Curves. SIAM J. Math. Anal. 48(4), 2869–2911 (2016)
- Chizat, L., Peyré, G., Schmitzer, B., Vialard, F.-X.: An interpolating distance between optimal transport and Fisher-Rao metrics. Found. Comput. Math. 18(1), 1–44 (2018)
- 22. Bronstein, A.M., Bronstein, M.M., Bruckstein, A.M., Kimmel, R.: Partial similarity of objects, or how to compare a centaur to a horse. Int. J. Comput. Vision **84**(2), 163 (2009)
- Robinson, D.T.: Functional data analysis and partial shape matching in the square root velocity framework (2012)
- Rodolà, E., Cosmo, L., Bronstein, M.M., Torsello, A., Cremers, D.: Partial functional correspondence. In: Computer Graphics Forum, vol. 36, pp. 222–236 (2017). Wiley Online Library
- Antonsanti, P.-L., Glaunès, J., Benseghir, T., Jugnon, V., Kaltenmark, I.: Partial matching in the space of varifolds. In: International Conference on Information Processing in Medical Imaging, pp. 123–135 (2021). Springer
- Sukurdeep, Y., Bauer, M., Charon, N.: A new variational model for shape graph registration with partial matching constraints. SIAM J. Imag. Sci. 15(1), 261–292 (2022)
- Attaiki, S., Pai, G., Ovsjanikov, M.: Dpfm: Deep partial functional maps. arXiv preprint arXiv:2110.09994 (2021)
- Hsieh, H.-W., Charon, N.: Diffeomorphic registration with density changes for the analysis of imbalanced shapes. In: International Conference on Information Processing in Medical Imaging, pp. 31–42 (2021). Springer
- 29. Almgren, F.: Plateau's Problem: An Invitation to Varifold Geometry. Student Mathematical Library (1966)
- 30. Allard, W.: On the first variation of a varifold. Ann. Math. 95(3), 417–491 (1972)
- 31. Simon, L.: Lecture Notes on Geometric Measure Theory. Australian national university, (1983)
- 32. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Clarendon Press, Oxford (2000)
- Buet, B., Leonardi, G.P., Masnou, S.: Discretization and approximation of surfaces using varifolds. Geometric Flows 3(1), 28–56 (2018)
- Buet, B., Rumpf, M.: Mean curvature motion of point cloud varifolds. arXiv preprint arXiv:2010.09419 (2020)
- 35. Younes, L.: Shapes and Diffeomorphisms. Springer (2019)
- 36. Friedrich, T.: Die fisher-information und symplektische strukturen. Math. Nachr. 153(1), 273-296 (1991)
- Bauer, M., Bruveris, M., Michor, P.W.: Uniqueness of the Fisher-Rao metric on the space of smooth densities. Bull. Lond. Math. Soc. 48(3), 499–506 (2016)
- Glaunès, J., Trouvé, A., Younes, L.: 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)
- Roussillon, P., Glaunès, J.: Kernel Metrics on Normal Cycles and Application to Curve Matching. SIAM J. Imag. Sci. 9(4), 1991–2038 (2016)
- Feydy, J., Charlier, B., Vialard, F.-X., Peyré, G.: Optimal Transport for Diffeomorphic Registration. In: Medical Image Computing and Computer Assisted Intervention, pp. 291–299 (2017)
- Feydy, J., Séjourné, T., Vialard, F.-X., Amari, S.-i., Trouvé, A., Peyré, G.: Interpolating between optimal transport and MMD using Sinkhorn divergences. In: The 22nd International Conference on Artificial Intelligence and Statistics, pp. 2681–2690 (2019). PMLR



- Charon, N., Charlier, B., Glaunès, J., Gori, P., Roussillon, P.: Fidelity metrics between curves and surfaces: currents, varifolds, and normal cycles. In: Riemannian Geometric Statistics in Medical Image Analysis, pp. 441–477. Elsevier, (2020)
- 43. Aronszajn, N.: Theory of reproducing kernels. Trans. Amer. Math. Soc. 68, 337–404 (1950)
- Sriperumbudur, B.K., Fukumizu, K., Lanckriet, G.: 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), vol. 9, pp. 773–780 (2010)
- Glaunès, J., Micheli, M.: Matrix-valued kernels for shape deformation analysis. Imaging and Computing 1(1), 57–139 (2014)
- Pontryagin, L., Boltyanskii, V., Gamkrelidze, R., Mishchenko, E.: The Mathematical Theory of Optimal Processes. John Wiley & Sons (1962)
- Arguillère, S., Trélat, E., Trouvé, A., Younes, L.: Shape deformation analysis from the optimal control viewpoint. Journal de Mathématiques Pures et Appliquées 104(1), 139–178 (2015)
- Glaunès, J., Qiu, A., Miller, M., Younes, L.: Large deformation diffeomorphic metric curve mapping. Int. J. Comput. Vision 80(3), 317–336 (2008)
- Charlier, B., Feydy, J., Glaunès, J., Collin, F.-D., Durif, G.: Kernel operations on the GPU, with autodiff, without memory overflows. J. Mach. Learn. Res. 22(74), 1–6 (2021)
- Miller, M., Younes, L., Ratnanather, J., Brown, T., Trinh, H., Lee, D., Tward, D., Mahon, P., Mori, S., Albert, M.: Amygdalar atrophy in symptomatic Alzheimer's disease based on diffeomorphometry: the BIOCARD cohort. Neurobiol. Aging 36(Supplement 1), 3–10 (2015)
- 51. Durrett, R.: Probability: Theory and Examples, vol. 49. Cambridge University Press (2019)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

