

HV GEOMETRY FOR SIGNAL COMPARISON

RUIYU HAN, DEJAN SLEPČEV, AND YUNAN YANG

Dedicated to Bob Pego whose knowledge and love of mathematics greatly inspire us.

ABSTRACT. In order to compare and interpolate signals, we investigate a Riemannian geometry on the space of signals. The metric allows discontinuous signals and measures both horizontal (thus providing many benefits of the Wasserstein metric) and vertical deformations. Moreover, it allows for signed signals, which overcomes the main deficiency of optimal transportation-based metrics in signal processing. We characterize the metric properties of the space of signals and establish the regularity and stability of geodesics. Furthermore, we introduce an efficient numerical scheme to compute the geodesics and present several experiments which highlight the nature of the metric.

1. INTRODUCTION

In [30], Miller and Younes introduced a deformation-based geometry on the space of L^2 functions in arbitrary dimension. It allows for and measures both horizontal deformations (as does the Wasserstein metric) and vertical deformations (as does the L^2 norm). Unlike the Wasserstein distance, it allows for signals that change signs. We will refer to the resulting geometry as the HV geometry throughout the paper.

Trounev and Younes [36] provided some of the foundational results, and subsequent works generalizing the approach to a variety of distances and settings. However, there have been few works carefully studying the associated geometry of the signals and theoretically establishing its properties. This is in stark contrast with the abundance of developments of variants of optimal transportation. In the hope of fostering further development of this and related geometries, we provide a largely self-contained introduction in perhaps the simplest setting, namely one-dimensional signals. We then prove new results characterizing the resulting metric, providing various a priori estimates, and establishing the regularity and stability of minimizing geodesics. Furthermore, we introduce a simple and efficient scheme to compute the metric.

While this framework of [36] has been used in applications, primarily in image processing, we are unaware that it has been applied to the setting of one-dimensional signals. Here we argue that when considered in the space of signals (one-dimensional functions), it provides a viable metric that has desirable features and can be effectively computed. This opens the door for a variety of applications.

Given a finite interval, which, for simplicity, we set to be $[0, 1]$, we consider the space of signals to be the L^2 functions. The paths on the space of signals are described by (weak) solutions of the transport equation with a source. In particular, a path connecting $f_0, f_1 \in L^2(0, 1)$ is described by

$$\begin{aligned} \partial_t f &= -\partial_x f \cdot v + z \quad \text{on } [0, 1] \times [0, 1], \\ v(0, \cdot) &= v(1, \cdot) = 0, \\ f(\cdot, 0) &= f_0, \quad f(\cdot, 1) = f_1. \end{aligned} \tag{1}$$

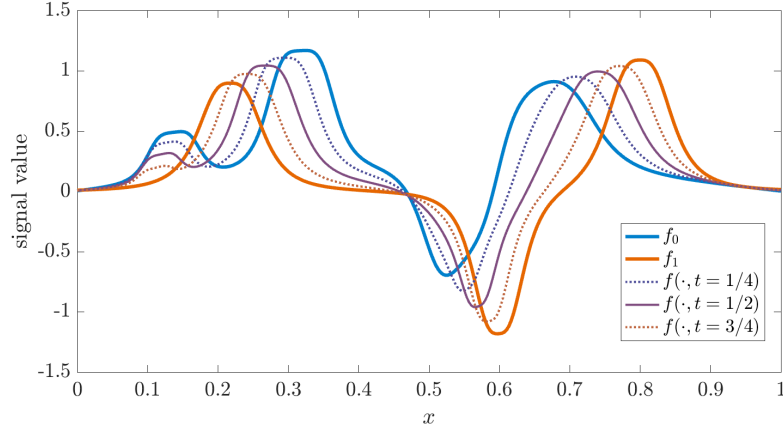


FIGURE 1. An example of the computed geodesic in the space of signals according to the HV geometry for $\kappa = 0.1$, $\lambda = 0.01$, and $\varepsilon = 0.0005$. We note that for some features, horizontal transform dominates, while some parts are matched by vertically moving the signals.

The action measuring the effort of deforming f_0 to f_1 along the particular path is

$$(2) \quad A_{\kappa, \lambda, \varepsilon}(f, v, z) = \frac{1}{2} \int_0^1 \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 \, dx \, dt$$

where $\kappa > 0$, $\lambda \geq 0$, and $\varepsilon > 0$.

The v^2 term measures the horizontal movement of the signal, while the z^2 measures the vertical change. The parameter κ is for modeling flexibility, as it accounts for horizontal and vertical variations differently. We note that the horizontal transport is modeled by the transport equation term, $\partial_x f \cdot v$, instead of the continuity equation $\text{div}(fv)$. Thus, even if $z \equiv 0$, the integral of f is not preserved unless $\text{div } v \equiv 0$. The λv_x^2 term measures how far from being conservative the transport is. It is helpful to mention that in a higher dimensional analog of the metric v_x term would be replaced by $\text{div } v$. In essence, it describes how much space needs to be created/removed to expand/contract the region where f takes a certain value. The εv_{xx}^2 term is necessary for regularity. In particular, it ensures that very small regions cannot be inflated at an arbitrarily small cost, which would create an undesirable shortcut that could become the dominant transport mechanism. Indeed we discuss in Section 2.1 that the geometry can degenerate if $\varepsilon = 0$.

We define the distance between functions f_0 and f_1 as the infimum of the action among all admissible paths connecting them

$$d_{HV(\kappa, \lambda, \varepsilon)} := \inf_{(f, v, z)} \int_0^1 \sqrt{\frac{1}{2} \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 \, dx} \, dt = \inf_{(f, v, z)} \sqrt{A_{\kappa, \lambda, \varepsilon}(f, v, z)}.$$

The second equality is proved using reparameterization. When parameters κ , λ , and ε are clear in the context, we use the short-hand notation d_{HV} . See Figure 1 for an example of the computed geodesic between two signed signals based on the HV geometry.

The metric we describe above is not new; modulo some minor technical details, it is the metric introduced by Miller and Younes [30] and studied by Trounev and Younes [37] and belongs to the family of metamorphoses studied by Trounev, Younes, and Holm [22]. It is also closely related to the metrics studied by Charlier, Charon, and Trounev [10, 11]. In particular, [37] and [10, Theorem 2] show the existence of geodesics, and [11, Property 1] shows the completeness of the metric. The works [22, 10, 11] display

very general frameworks and large families of possible deformation-based ways to compare functions on both fixed and changing sets and manifolds.

Here we present a simple and self-contained description of the particular metric of interest in the space of signals. Our problem description is primarily Eulerian instead of the approach based on groups of diffeomorphisms. The Eulerian point of view, allows us to consider paths between signals with less regularity. This allows us to prove that metrics that only penalize one derivative do not lead to a geodesic space, as minimizing geodesics may not exist; see Section 2.1. We prove the existence of geodesics and the completeness of the resulting metric space of signals, which are analogous to those in [36] and [11], respectively. Our approach allows us to characterize the convergence in the resulting space of signals as the L^2 convergence; see Theorem 2.16. We are not aware of such a result in the other frameworks. In particular, knowing that the resulting metric is not weaker than the L^2 distance provides clarity about the distance that deformations induce. Furthermore, we prove a priori estimates on the minimizers, allowing us to exhibit that the minimizers preserve the regularity of the signals. In particular, Proposition 3.1 shows the result for signals in H^1 and Remark 3.2 shows it for smooth signals.

Furthermore, we introduce a numerical algorithm, which iterates between optimizing the action over v and z with f fixed and optimizing over z and f with v fixed. Each of the subproblems amounts to minimizing a quadratic functional under a linear constraint (unlike the full problem where the constraint is nonlinear) and thus can be efficiently performed. We observe that the algorithm quickly (within a few iterations) converges to a local minimum of the action. By iterating between solving these two sub-problems, we can show that the action functional (2) monotonically decreases on the continuous level. We employ a Lagrangian approach when optimizing over z and f with v fixed, which involves numerical integration and numerical interpolation. When optimizing the action over v and z with f fixed, we solve a fourth-order partial differential equation (PDE) through the finite-difference method. Since the geodesic obtained depends on the initial data, we also propose a velocity initialization for which matches the k most prominent peaks of each signal. To mitigate the issue of only finding a local minimizer, we optimize over $k \in \{0, 1, \dots, k_{max}\}$ to select the initialization that leads to the smallest action value, where k_{max} is the number we pick, denoting the maximum number of prominent peaks we aim to match. In particular, $k = 0$ corresponds to zero-velocity initialization.

1.1. Literature review. Designing and choosing the right way to compare the functions considered are important in a number of contexts. In inverse problems, such as denoising, deblurring, or wave-form inversion, choosing an appropriate way to measure loss is crucial to the outcome. If the signals represent the distribution of mass where the location of the mass is more important than the density, optimal transportation metrics perform especially well [15, 17]. In the setting of optimal transportation with quadratic cost, the equation in (1) is replaced by the continuity equation:

$$\partial_t f = -\operatorname{div}(fv)$$

and the action is

$$A_W(f, v) = \int_0^1 \int_0^1 v^2 f \, dx dt.$$

The optimal transportation metrics have the requirement that the functions compared are nonnegative with the same total mass. The mass requirement was relaxed by the unbalanced optimal transport [24, 13], which allows for the change of the mass, but still requires non-negativity. For unbalanced transport, the continuity equation can have a multiplicative source term, and the action also penalizes signal amplification:

$$\begin{aligned} \partial_t f &= -\operatorname{div}(fv) + zf, \\ \text{and } A_{UB}(f, v) &= \int_0^1 \int_0^1 (v^2 + z^2) f \, dx dt. \end{aligned}$$

There is a need to compare general signals with both positive and negative values in many applications, such as full-waveform inversion [17]. It poses a limitation for applications of optimal transportation. Several ideas for adapting optimal transportation approaches to signed signals have been introduced [17]. They include adding a constant, exponentiating the signal, comparing positive and negative parts separately, and considering the transportation between the graphs of functions on the product space (e.g., TL^p metric [34]). Comparing the signals on graphs has been shown as effective in full-waveform inversion and related problems [27]. However, all of these approaches lack one of the key elements, namely the Riemannian nature that allows for interpolation or the desired invariance $d(f_0, f_1) = d(-f_0, -f_1)$.

Geometric ways for comparing shapes and submanifolds of $\mathbb{R}^d$ have received a lot of attention in computer vision, computational anatomy and geometry processing. Riemannian geometries on the space of curves have been much studied. Mumford and Michor [28] showed that if only the L^2 norm of the velocity is considered in the action, then the metric degenerates. Bruveris, Michor and Mumford [9, 8] and Nardi, Peyre, and Vialard [31] established that the space of immersed plane curves is geodesically complete when the Sobolev or, respectively, BV -based metric involves the L^2 norm of two or more derivatives. Early works on computational anatomy and nonrigid registration have led to studies of geometries of the shapes of surfaces, as well as higher dimensional manifolds, many of which are surveyed in the book by Younes [41]. Recent results [3, 4] established the local well-posedness of geodesic equations when the metrics penalize at least one derivative of the deformation field.

A different line of work investigates the spaces of shapes considered as interiors of sets. In some sense, this takes into account the mass contained in the set, as does the Wasserstein distance. Liu, Pego and one of the authors, [25], showed that restricting the Wasserstein geometry to the space of characteristic functions is not viable, as the geodesics do not exist. In fact, minimizing curves converge weakly to the Wasserstein geodesics in the unconstrained space. Wirth, Bar, Rumpf, and Sapiro [40] considered actions that penalize the gradient of the velocity field and showed that, in this case, the geometry is viable. The framework was refined by Rumpf and Wirth [32, 33], who also provided a numerical method for finding geodesics.

In parallel with the studies of the geometry of the spaces of curves and shapes, researchers considered the metrics which allow for comparing signals, such as gray-level images, while allowing for both deformations of the domain and intensity variations. Trouné [35] introduced the basic description based on group actions. This was refined into the description of deformable templates [36] and metamorphoses [37] by Trouné and Younes. As we indicated at the beginning, the geometry we study belongs to this family. Eulerian description of metamorphosis was carried out in [22]. Charlier, Charon, and Trouné [10, 11] considered spaces where one compares manifolds together with functions on them. The authors of [10, 11] showed that under sufficient regularity assumptions, in the space of manifold, function pairs are geodesically complete. Many of the numerical approaches (e.g. [29, 20]) to computing the minimizing geodesics between the given images relied on the so called shooting methods where one iterates computing the forward geodesic for the given initial conditions and adjusts the initial conditions. Berkels, Effland, and Rumpf [6] took a different approach and gave a variational formulation based on discrete-in-time paths where one minimizes the deformation between consecutive images in a way that is consistent with the action integrated in time. This is a promising and well founded approach. One difference to our approach is that their sub-problems remain nonconvex.

1.2. Outline. The rest of the paper is organized as follows. In Section 2, we define the HV geometry and the associated distance and establish their properties. In particular, we present the scaling invariances of the metric in Proposition 2.1. In Section 2.2, we rigorously identify the tangent space. The identification is analogous but slightly different from the one [36] in that instead of equivalence classes, we identify the representatives that achieve the minimal length (just as the gradient vector fields minimize the action for Wasserstein geometry). In Section 2.3, we provide a number of a priori estimates, prove the existence of geodesics (in a slightly different way from that in [36]), and obtain representation formulas satisfied by

geodesics. In Section 2.4, we prove the completeness of the d_{HV} metric on $L^2(0, 1)$ and that the topology induced by d_{HV} on $L^2(0, 1)$ is the same as the one induced by the L^2 norm. Moreover, we show in Proposition 2.3 that without involving the second-order derivative of the velocity, action minimizers may not exist. In Section 3, we further study several properties of the geodesics. In Proposition 3.1, we show that if the starting and the ending signals are in H^1 , then the signals along the minimizing geodesic remain in H^1 . Furthermore, Remark 3.2 indicates that if the signals are smooth, so is the geodesic connecting them. Proposition 3.3 establishes the L^2 stability of minimizing geodesics with respect to perturbations of the endpoints. We note that this is not a straightforward compactness result since, initially, one only has control over the L^2 norms of the signals. Section 3.1 is devoted to establishing first-order optimality conditions of the action minimizers. We first obtain these conditions for general L^2 signals. If the signals are further in H^1 , we show that the Euler–Lagrange equations have the form of differential equations. In Section 4, we introduce our numerical method. We first present the iterative minimization scheme that uses two sub-problems which are based on Euler–Lagrange equations and representation formulas for the geodesics. The discretization we use based on finite difference is included in Section 4.2. We describe our approach to the path initialization in Section 4.3. As the energy often has local minimizers, it is important to use initialization informed by the signals. In Section 5, we display several illustrative examples that highlight the properties of the HV signal geometry, as well as the numerical approach we take. We also apply our scheme to signals from the ECG datasets and seismology. Finally we discuss the parameter selection in Section 5.1.

2. HV GEOMETRY AND ITS PROPERTIES.

To rigorously define the HV distance we start by precisely defining the set of admissible paths. Let $\mathcal{V} := L^2((0, 1), H^2(0, 1) \cap H_0^1(0, 1))$. Given $f_0, f_1 \in L^2(0, 1)$ we define the set of admissible paths to be

$$(3) \quad \begin{aligned} \mathcal{A}(f_0, f_1) := & \{ (f, v, z) \mid f \in L^2((0, 1), L^2(0, 1)) \cap C([0, 1], H^{-1}(0, 1)), \\ & v \in \mathcal{V}, z \in L^2((0, 1), L^2(0, 1)), \\ & \partial_t f + \partial_x f \cdot v = z \text{ weakly, } f(\cdot, 0) = f_0, f(\cdot, 1) = f_1 \}. \end{aligned}$$

By weak solutions of $\partial_t f + \partial_x f \cdot v = z$ above we mean that for any test function $\phi \in C^\infty([0, 1]^2)$,

$$(4) \quad \begin{aligned} & - \int_0^1 \int_0^1 f \partial_t \phi \, dx \, dt + \int_0^1 f_1(x) \phi(x, 1) \, dx - \int_0^1 f_0(x) \phi(x, 0) \, dx \\ & = \int_0^1 \int_0^1 f \phi \partial_x v \, dx \, dt + \int_0^1 \int_0^1 f \partial_x \phi v \, dx \, dt + \int_0^1 \int_0^1 \phi z \, dx \, dt. \end{aligned}$$

The HV distance is then defined by

$$(5) \quad d_{HV(\kappa, \lambda, \varepsilon)} := \inf_{(f, v, z) \in \mathcal{A}(f_0, f_1)} \int_0^1 \sqrt{\frac{1}{2} \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 \, dx} \, dt$$

The definition of the distance implies the following simple, but useful properties.

Proposition 2.1. *Consider $f_0, f_1 \in L^2(0, 1)$. Let $c > 0$. Then*

- (i) $d_{HV}(f_0, f_1) \leq \|f_0 - f_1\|_{L^2}$.
- (ii) $d_{HV}(-f_0, -f_1) = d_{HV}(f_0, f_1)$
- (iii) $d_{HV}(f_0 + c, f_1 + c) = d_{HV}(f_0, f_1)$
- (iv) $d_{HV(c^2 \kappa, c^2 \lambda, c^2 \varepsilon)}(cf_0, cf_1) = cd_{HV(\kappa, \lambda, \varepsilon)}(f_0, f_1)$

To indicate the behavior of the action with respect to rescaling the space extend f_0 and f_1 periodically to $\mathbb{R}$. Likewise, given a path (f, v, z) consider it extended periodically to $\mathbb{R}$. Then for $L \in \mathbb{N}$

(v) $A_{L^2\kappa,\lambda,\varepsilon/L^2}(f(L\cdot, \cdot), v(L\cdot, \cdot), z(L\cdot, \cdot)) = A_{\kappa,\lambda,\varepsilon}(f, v, z)$, where the action is considered only on $[0, 1]$, as usual.

Property (i) is proved by computing the action of the linear interpolation while the remaining properties are proved by a straightforward application of the definition of the action.

2.1. Degeneracy without the second derivatives. Before proving rigorous results about d_V , we show that if $\varepsilon = 0$ the geometry of signals would not have all of the desirable properties. In particular we show that there exist functions f_0 and f_1 such that there does not exist any minimizers for the action.

To be more precise, consider the set of admissible paths to be as in (3), but with $\mathcal{V}$ replaced by $L^2((0, 1), H_0^1(0, 1))$. Consider the action

$$(6) \quad A(f, v, z) = \int_0^1 \int_0^1 v^2 + v_x^2 + z^2 dx dt.$$

We start by noting that linear interpolation is not optimal when f_0 and f_1 are constant functions that differ sufficiently.

Lemma 2.2. *If $\varepsilon = 0$, then for all $\lambda \in [0, \infty)$ there exists $H \in \mathbb{R}$ such that the linear interpolation path between $f_0 \equiv 0$ and $f_1 \equiv H$ is not optimal.*

Proof. Consider $f_0 \equiv 0$ and $f_1 \equiv H$ for some $H > 0$. One possible path between f_0 and f_1 is by L^2 interpolation. This path is defined by $v \equiv 0$ and $z \equiv H$. Then

$$A(f, v, z) = \int_0^1 \int_0^1 z^2 dx dt = H^2$$

We construct a competitor with nonzero velocity v . In the time interval $t \in [0, 1/3]$, use an L^2 interpolation between f_0 and

$$f_2(x) = \begin{cases} H & x \leq s \\ 0 & x > s \end{cases}$$

for fixed $s \in (0, 1/2)$. In this time interval, $v \equiv 0$, and $\int_0^{1/3} \int_0^s z^2 dx dt = 3sH^2$. On the time interval $t \in [1/3, 2/3]$, the signal moves from f_2 to

$$f_3(x) = \begin{cases} H & x \leq 1-s \\ 0 & x > 1-s \end{cases}$$

by moving the edge of the step function with a constant speed $3(1-2s)$. The velocity parameterized by the initial position is $w(x) = 3(1-2s)\frac{x}{s}$ for $x \in (0, s]$ and $w(x) = 3(1-2s)\frac{1-x}{1-s}$ for $x \in (s, 1)$. Then for all $t \in [1/3, 2/3]$, the velocity is $v(x + w(x)(t - \frac{1}{3}), t) = w(x)$. By chain rule $0 \leq v \leq 3(1-2s)$ and $|v_x| \leq \frac{3(1-2s)}{s}$. So

$$\int_{\frac{1}{3}}^{\frac{2}{3}} \int_0^1 v^2 + \lambda v_x^2 dx dt \leq 3(1-2s)^2 \left(1 + \frac{\lambda}{s^2}\right)$$

The interpolation between f_2 and f_3 is that the signal is constant along the trajectories. Finally use L^2 interpolation between f_3 and f_1 . This is symmetric to the time interval $[0, 1/3]$ and adds $3sH^2$ to the action.

Thus the total action for this path is bounded above by $6sH^2 + 3(1-2s)^2(1 + \frac{\lambda}{s^2})$. By picking $s < \frac{1}{6}$, and then $H > 0$ such that $H^2 > \frac{3}{1-6s}(1-2s)^2(1 + \frac{\lambda}{s^2})$, we have that $H^2 > 6sH^2 + 3(1-2s)^2(1 + \frac{\lambda}{s^2})$. Therefore L^2 interpolation is not the optimal path between $f_0 \equiv 0$ and $f_1 \equiv H$. $\square$

Proposition 2.3. *There exists $H > 0$ such that there is no path between $f_0 \equiv 0$ and $f_1 \equiv H$ minimizing the action (6).*

Proof. Assume H satisfies the condition of Lemma 2.2. Consider any path (f, v, z) from f_0 to f_1 . We split the argument into two cases depending on whether $v \equiv 0$.

Case 1° If $v \equiv 0$. Then $f_t = z$ for all x, t , that is the path is linear interpolation. Hence, be Lemma 2.2 the path does not minimize the action.

Case 2° $v \not\equiv 0$. From the path (f, v, z) we can construct a new path $(\tilde{f}, \tilde{v}, \tilde{z})$ by creating two copies of f shrank to interval $\frac{1}{2}$.

$$\begin{aligned}\tilde{f}(x, t) &= \begin{cases} f(2x, t) & x \leq \frac{1}{2} \\ f(2x - 1, t) & x \geq \frac{1}{2} \end{cases} \\ \tilde{v}(x, t) &= \begin{cases} \frac{1}{2}v(2x, t) & x \leq \frac{1}{2} \\ \frac{1}{2}v(2x - 1, t) & x \geq \frac{1}{2} \end{cases} \\ \tilde{z}(x, t) &= \begin{cases} z(2x, t) & x \leq \frac{1}{2} \\ z(2x - 1, t) & x \geq \frac{1}{2}. \end{cases}\end{aligned}$$

A consequence of this is that

$$\tilde{v}_x(x, t) = \begin{cases} v_x(2x, t) & x \leq \frac{1}{2} \\ v_x(2x - 1, t) & x \geq \frac{1}{2}. \end{cases}$$

For the action of this path,

$$\begin{aligned}A(\tilde{f}, \tilde{v}, \tilde{z}) &= \int_0^1 \int_0^{\frac{1}{2}} \tilde{v}^2 + \tilde{v}_x^2 + \tilde{z}^2 dx dt + \int_0^1 \int_{\frac{1}{2}}^1 \tilde{v}^2 + \tilde{v}_x^2 + \tilde{z}^2 dx dt \\ &= \int_0^1 \int_0^{\frac{1}{2}} \frac{1}{4} v(2x, t)^2 + \tilde{v}_x(2x, t)^2 + \tilde{z}(2x, t)^2 dx dt \\ &\quad + \int_0^1 \int_{\frac{1}{2}}^1 \frac{1}{4} \tilde{v}(2x - 1, t)^2 + \tilde{v}_x(2x - 1, t)^2 + \tilde{z}(2x - 1, t)^2 dx dt \\ &= \int \int \frac{1}{4} v^2 + v_x^2 + z^2 dx.\end{aligned}$$

Since v is non-zero, this is strictly less than the action $A(f, v, z)$. Thus the path (f, v, z) is not minimal.

Therefore so no path between f_0 and f_1 minimizes the action. $\square$

2.2. Identification of the tangent space. We note that d_{HV} , defined in (5), can be seen as the length of a curve in the space of signals. Indeed, as we show below, $(L^2(0, 1), d_{HV})$ is geodesic space. From the definition of the admissible curves and the action we see that in Eulerian description of the tangent space

$$(7) \quad \text{Tan}_{f, Euler} = \{-vf_x + z : v \in H^2(0, 1) \cap H_0^1(0, 1), z \in L^2(0, 1)\} \subseteq H^{-1}$$

where vf_x is the element of H^{-1} defined by $\langle -vf_x, \phi \rangle = \int_0^1 v_x f \phi + v f \phi_x dx$ for all $\phi \in H_0^1$. We show below that in the Lagrangian description the tangent space can be identified with a subspace of $(H^2(0, 1) \cap H_0^1(0, 1)) \times L^2(0, 1)$.

Lemma 2.4. *Given $f \in L^2(0, 1)$, $\bar{v} \in H^2(0, 1) \cap H_0^1(0, 1)$ and $\bar{z} \in L^2(0, 1)$ there exists a unique pair $v \in H^2(0, 1) \cap H_0^1(0, 1)$, $z \in L^2(0, 1)$ minimizing the instantaneous action*

$$(8) \quad Q_f(v, z) = \frac{1}{2} \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 dx$$

under the constraint

$$(9) \quad \forall \phi \in H_0^1(0, 1) \quad \int_0^1 f v_x \phi + f v \phi_x + z \phi \, dx = \int_0^1 f \bar{v}_x \phi + f \bar{v} \phi_x + \bar{z} \phi \, dx.$$

We will denote the solution mapping by $S : H^2(0, 1) \cap H_0^1(0, 1) \times L^2(0, 1) \rightarrow H^2(0, 1) \cap H_0^1(0, 1) \times L^2(0, 1)$, that is $(v, z) = S(\bar{v}, \bar{z})$.

Proof. Let $C(f, \bar{v}, \bar{z})$ be the set of $v \in H^2(0, 1) \cap H_0^1(0, 1)$, $z \in L^2(0, 1)$ satisfying the constraint (9). Note that if $(v_n, z_n) \in C(f, \bar{v}, \bar{z})$, $v_n \rightarrow v$ in H_0^1 , $z_n \rightarrow z$ in L^2 as $n \rightarrow \infty$ and $v \in H^2(0, 1)$ then $(v, z) \in C(f, \bar{v}, \bar{z})$. Namely $f \phi \in L^2$ ensures that $\int f(v_n)_x \phi \, dx \rightarrow \int f v_x \phi \, dx$ and $v_n \rightarrow v$ in L^∞ implies $\int f v_n \phi_x \, dx \rightarrow \int f v \phi_x \, dx$ as $n \rightarrow \infty$.

We note that $(\bar{v}, \bar{z}) \in C(f, \bar{v}, \bar{z})$ and $Q_f(\bar{v}, \bar{z}) < \infty$. Let $(v_n, z_n) \in C(f, \bar{v}, \bar{z})$ be a minimizing sequence of Q . Thus $\{v_n\}_n$ is bounded in H^2 and $\{z_n\}_n$ is bounded in L^2 . Hence there is a subsequence, which by relabeling we can assume to be the whole sequence, $v_n \rightarrow v$ in H_0^1 and $z_n \rightarrow z$ in L^2 to some $v \in H^2 \cap H_0^1$ and $z \in L^2$. By above $(v, z) \in C(f, \bar{v}, \bar{z})$. Since Q_f is sequentially lower-somicontinuous with respect to H^1 convergence in v and weak L^2 convergence in z we conclude that (v, z) minimizes Q_f over $C(f, \bar{v}, \bar{z})$.

The uniqueness follows from the fact that Q_f is strictly convex and that $C(f, \bar{v}, \bar{z})$ is convex. $\square$

We now characterize the solution $(v, z) = S(\bar{v}, \bar{z})$ above via the first variation. We note that condition (9) implies that $(v - \bar{v})f_x$ belongs to L^2 , in the sense that $\langle (v - \bar{v})f_x, \phi \rangle = \int_0^1 (z - \bar{z})\phi \, dx$ for all $\phi \in H_0^1(0, 1)$, where $\langle \cdot, \cdot \rangle$ denotes the dual pairing between H^{-1} and H_0^1 . It is straightforward to show that for $f \in L^2(0, 1)$ and $u \in H^2(0, 1) \cap H_0^1(0, 1)$ the condition $u f_x \in L^2(0, 1)$ is equivalent to $u f \in H^1(0, 1)$. This motivates us to introduce the space

$$(10) \quad R(f) = \{u \in H^2(0, 1) \cap H_0^1(0, 1) : u f \in H^1(0, 1)\}.$$

Note that for $u \in R(f)$ $u f_x = (u f)_x - u_x f \in L^2$. We also remark that if $f \in H^1$ then $R(f) = H^2(0, 1) \cap H_0^1(0, 1)$.

Lemma 2.5. Consider $f \in L^2(0, 1)$, $\bar{v} \in H^2(0, 1) \cap H_0^1(0, 1)$ and $\bar{z} \in L^2(0, 1)$. Then $(v, z) = S(\bar{v}, \bar{z})$ if and only if $v \in H^2(0, 1) \cap H_0^1(0, 1)$, $z \in L^2(0, 1)$ and

$$(11) \quad \forall u \in R(f) \quad \int_0^1 \kappa v u + \lambda v_x u_x + \varepsilon v_{xx} u_{xx} + ((v - \bar{v})f_x + \bar{z})(u f_x) \, dx = 0.$$

Proof. If $(v, z) = S(\bar{v}, \bar{z})$ then $v \in H^2(0, 1) \cap H_0^1(0, 1)$, $z \in L^2(0, 1)$. Taking $u \in R(f)$ and $h = u f_x$, which belongs to L^2 , since $u \in R(f)$ we note that $(v + su, z + sh) \in C(f, \bar{v}, \bar{z})$. First variation in of Q_f at (v, z) gives the condition (11).

Now assume that $v \in H^2(0, 1) \cap H_0^1(0, 1)$, $z \in L^2(0, 1)$ and that (11) holds. Let $(\tilde{v}, \tilde{z}) = S(\bar{v}, \bar{z})$. By above (11) holds for $(\tilde{v}, \tilde{z})$ in place of (v, z) . Furthermore note that $v - \tilde{v} = (v - \bar{v}) - (\tilde{v} - \bar{v}) \in R(f)$. Taking $u = v - \tilde{v}$ and subtracting the two forms of (11) gives

$$\int_0^1 \kappa(v - \tilde{v})^2 + \lambda(v_x - \tilde{v}_x)^2 + \varepsilon(v_{xx} - \tilde{v}_{xx})^2 + ((v - \tilde{v})f_x)^2 \, dx = 0$$

Thus $v = \tilde{v}$ and $z = \tilde{z}$. $\square$

It is useful to note that if $f \in H^1(0, 1)$ and hence $R(f) = H^2(0, 1) \cap H_0^1(0, 1)$, the condition (11) means that for $g = -f_x \bar{v} + \bar{z}$, v is a weak solution of

$$(12) \quad \begin{aligned} \varepsilon v_{xxxx} - \lambda v_{xx} + \kappa v + f_x^2 v &= -f_x g \quad \text{on } (0, 1) \\ v &= 0 \quad \text{at } \{0, 1\} \\ v_{xx} &= 0 \quad \text{at } \{0, 1\}. \end{aligned}$$

The above lemmas allow us to characterize the tangential velocities, namely we define

$$(13) \quad \text{Tan}_f = \{(v, z) \in H^2(0, 1) \cap H_0^1(0, 1) \times L^2(0, 1) : (v, z) = S(v, z)\}.$$

Lemma 2.5 gives that a pair $(v, z) \in H^2(0, 1) \cap H_0^1(0, 1) \times L^2(0, 1)$ belongs to Tan_f if and only if

$$(14) \quad \forall u \in R(f) \quad \int_0^1 \kappa v u + \lambda v_x u_x + \varepsilon v_{xx} u_{xx} + z u f_x dx = 0.$$

In the special case that $f \in H^1(0, 1)$ we can characterize Tan_f as the set of pairs (v, z) where $z \in L^2(0, 1)$ and $v \in H^2(0, 1) \cap H_0^1(0, 1)$ is a weak solution of

$$(15) \quad \begin{aligned} \varepsilon v_{xxxx} - \lambda v_{xx} + \kappa v &= -f_x z \quad \text{on } (0, 1) \\ v &= 0 \quad \text{at } \{0, 1\} \\ v_{xx} &= 0 \quad \text{at } \{0, 1\}. \end{aligned}$$

We furthermore remark that if $u \in R(f)$ then, since H^1 functions in one dimension are continuous we have that both uf and u are continuous. Thus $f = uf/u$ must be continuous on the set where $u \neq 0$. In other words if f does not have a representative that is continuous at some x then all functions $u \in R(f)$ are equal to zero at x ; $u(x) = 0$. In particular if f is nowhere continuous then the only u in $R(f)$ is the zero function. Consequently if f is nowhere continuous then $\text{Tan}_f = H^2(0, 1) \cap H_0^1(0, 1) \times L^2(0, 1)$, that every pair is a tangent pair.

2.3. Existence of minimizing paths. Here we show the existence of minimizers of the action (2)

Theorem 2.6. *Consider $f_0, f_1 \in L^2(0, 1)$. There exists an admissible path $(f, v, z) \in \mathcal{A}(f_0, f_1)$ minimizing the action (2).*

Before proving Theorem 2.6, we first establish several properties of solutions of the transport equation. These results and the techniques are standard, we present them for completeness.

Lemma 2.7. *Given $v \in \mathcal{V}$, $z \in L^2((0, 1)^2)$. Let $f_0 \in L^2(0, 1)$. If $f \in L^2((0, 1)^2)$ is a weak solution to the initial value problem*

$$f(\cdot, 0) = f_0, \quad \partial_t f + \partial_x f \cdot v = z \quad \text{on } [0, 1]^2,$$

in the sense that for every $\phi \in C^\infty([0, 1]^2)$, (4) holds, then f has the following representation: for a.e. x, t in $[0, 1]^2$,

$$(16) \quad f(\Phi(x, t), t) = f_0(x) + \int_0^t z(\Phi(x, s), s) ds,$$

where Φ is the flow of the vector field v :

$$(17) \quad \partial_t \Phi(x, t) = v(\Phi(t, x), t), \quad \Phi(0, x) = x.$$

It follows that the weak solution f is unique.

Proof. Let us extend f_0, z, v, f by zero outside of $[0, 1]$, and denote the extensions by $\tilde{z}, \tilde{f}_0, \tilde{v}, \tilde{f}$. Note that $\tilde{v} \in L^2(0, 1, W^{1,\infty}(\mathbb{R}))$ since $H^2(0, 1)$ embeds in $W^{1,\infty}(0, 1)$, $v(0, \cdot) \equiv 0$, $v(1, \cdot) \equiv 0$, and extending by zero preserves the Lipschitz constant.

Then $\tilde{f} \in L^2(0, 1, L^2(\mathbb{R}))$ satisfies that for all test functions $\phi \in C^\infty(\mathbb{R} \times [0, 1])$ with compact support in $\mathbb{R} \times [0, 1]$

$$-\int_0^1 \int_{\mathbb{R}} \tilde{f} \partial_t \phi dx dt - \int_{\mathbb{R}} \tilde{f}_0 \phi(x, 0) dx = \int_0^1 \int_{\mathbb{R}} \tilde{f} (\tilde{v} \partial_x \phi + \phi \partial_x \tilde{v}) dx dt + \int_0^1 \int_0^1 \phi \tilde{z} dx dt.$$

Since $\tilde{v}$ is Lipschitz in space, the representation of solutions of the transport equation (see Proposition 2.3 in [1]), for a.e. x

$$\tilde{f}(\tilde{\Phi}(x, t), t) = \tilde{f}_0(x) + \int_0^t \tilde{z}(\tilde{\Phi}(x, s), s) ds, \quad t \in [0, 1],$$

where $\tilde{\Phi}$ is the flow of the vector field $\tilde{v}$, which satisfies the differential equation

$$\partial_t \tilde{\Phi}(x, t) = \tilde{v}(\tilde{\Phi}(x, t), t), \quad \tilde{\Phi}(0, x) = x$$

for each t . Notice that since $\tilde{v}(x, t) = 0$ for all $x \in \mathbb{R} \setminus (0, 1)$ and all $t \in [0, 1]$, by the uniqueness of flow map we have that for a.e. $x \in (0, 1)$, $\tilde{\Phi}(x, t) \in (0, 1)$ for all t . Define $\Phi(x, t) := \tilde{\Phi}(x, t) \mathbb{1}_{\{x \in [0, 1]\}}$. Then Φ is the flow of v and for a.e. x in $[0, 1]$,

$$f(\Phi(x, t), t) = f_0(x) + \int_0^t z(\Phi(x, s), s) ds, \quad \forall t \in [0, 1].$$

□

Remark 2.8. [Representation formulas for action minimizing paths.] Note that if $z \in L^2((0, 1)^2)$ is a weak solution of $\partial_t z + \partial_x(zv) = 0$ and v satisfies the conditions of the previous lemma, then it is a weak solution of $\partial_t z + (\partial_x z)v = -z\partial_x v$. Thus

$$z(\Phi(x, t), t) = z(x, 0) - \int_0^t z(\Phi(x, s), s) v_x(\Phi(x, s), s) ds$$

which is an integral form of the ODE whose solution is

$$(18) \quad z(\Phi(x, t), t) = z(x, 0) e^{-\int_0^t v_x(\Phi(x, s), s) ds},$$

where Φ is the flow of v . Here we emphasize that by v_x we always denote the partial derivative of v with respect to the first variable, and not the derivative of the composition.

Let $J(x, t) = e^{-\int_0^t v_x(\Phi(x, s), s) ds}$. We have the following representation:

$$f_1(\Phi(x, 1)) = f_0(x) + \int_0^1 z(\Phi(x, t), t) dt = f_0(x) + z(x, 0) \int_0^1 J(x, t) dt.$$

where the first equality follows from Lemma 2.7. This allows us to determine $z(x, 0)$ and hence

$$(19) \quad z(\Phi(x, t), t) = (f_1(\Phi(x, 1)) - f_0(x)) \frac{J(x, t)}{\int_0^1 J(x, \tau) d\tau}.$$

Hence we have the following formula for $f(\Phi(x, t), t)$

$$(20) \quad \begin{aligned} f(\Phi(x, t), t) &= f_0(x) + \int_0^t z(\Phi(x, s), s) ds = f_0(x) + (f_1(\Phi(x, 1)) - f_0(x)) \frac{\int_0^t J(x, s) ds}{\int_0^1 J(x, s) ds} \\ &= (1 - \eta(x, t)) f_0(x) + \eta(x, t) f_1(\Phi(x, 1)), \quad \text{where } \eta(x, t) = \frac{\int_0^t J(x, s) ds}{\int_0^1 J(x, s) ds}. \end{aligned}$$

Equation (20) implies that at any $t \in [0, 1]$ and a fixed $x \in [0, 1]$, $f(\Phi(x, t), t)$ is an interpolation between the initial condition $f_0(x) = f_0(\Phi(x, 0))$ and the final-time condition $f_1(\Phi(x, 1))$, with a time-dependent weight function $\eta(x, t)$ determined by the velocity v .

We note that if $(f, v, z) \in \mathcal{A}(f_0, f_1)$ then f is a weak solution of $f_t + f_x v = z$. We show in Section 3.1, equation (36) that if the path is a critical point of the action then z solves the continuity equation $z_t + (zv)_x = 0$. The formulas (20) and (19) provide the Lagrangian representation formulas for f and z .

Lemma 2.9. *Let $f_0 \in L^2(0, 1)$, $v \in L^2(0, 1, H^2(0, 1) \cap H_0^1(0, 1))$, $z \in L^2((0, 1)^2)$. If f has the representation (16) then*

$$\max_{0 \leq t \leq 1} \|f(\cdot, t)\|_{L^2(0,1)} \leq \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))})(\|f_0\|_{L^2(0,1)} + \|z\|_{L^2((0,1)^2)}).$$

Proof. Let $D\Phi(\cdot, t)$ be the spatial derivative of $\Phi(\cdot, t)$. Notice that that for every $x \in [0, 1]$

$$\partial_t D(\Phi(x, t)) = D(\partial_t \Phi(x, t)) = D(v(\Phi(x, t), t)) = v_x(\Phi(x, t), t) D\Phi(x, t).$$

Integrating in t , since $D\Phi(x, 0) = 1$

$$(21) \quad D\Phi(x, t) = \exp\left(\int_0^t v_x(\Phi(x, s), s) ds\right) \leq \exp(\sqrt{2t}\|v\|_{L^2(0,1,H^2(0,1))}),$$

where the constant $\sqrt{2}$ comes from the embedding $H_0^1((0, 1)) \hookrightarrow L^\infty((0, 1))$. Therefore

$$(22) \quad \exp(-\sqrt{2t}\|v\|_{L^2(0,1,H^2(0,1))}) \leq D\Phi(x, t) \leq \exp(\sqrt{2t}\|v\|_{L^2(0,1,H^2(0,1))})$$

since for every $t \in [0, 1]$

$$(23) \quad \left| \int_0^t v_x(\Phi(\cdot, s), s) ds \right| \leq \int_0^t \|v_x(\Phi(\cdot, s), s)\|_{L^\infty(0,1)} ds \leq \sqrt{2t}\|v\|_{L^2(0,1,H^2(0,1))}.$$

By Lemma 2.7, for a.e. $y \in [0, 1]$, $t \in [0, 1]$,

$$(24) \quad f(y, t) = f_0(\Phi(\cdot, t)^{-1}(y)) + \int_0^t z(\Phi(\Phi(\cdot, t)^{-1}(y), s), s) ds.$$

By change of variables, for the first term on the right-hand side of (24)

$$\begin{aligned} \int_0^1 f_0^2(\Phi(\cdot, t)^{-1}(y)) dy &= \int_0^1 f_0^2(\Phi(\cdot, t)^{-1}(y)) D\Phi(\cdot, t)^{-1} D\Phi(\cdot, t) dy \\ &\leq \|f_0\|_{L^2(0,1)}^2 \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))}). \end{aligned}$$

For the second term on the right-hand side of (24), using estimate (22) twice and applying Jensen's inequality,

$$\left\| \int_0^t z(\Phi(\Phi(\cdot, t)^{-1}, s), s) ds \right\|_{L^2(0,1)} \leq \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))}) \|z\|_{L^2((0,1)^2)}$$

Plugging the above estimates into (24) we obtain

$$\begin{aligned} \|f(\cdot, t)\|_{L^2} &\leq \|f_0\|_{L^2(0,1)} \exp\left(\frac{\sqrt{2}}{2}\|v\|_{L^2(0,1,H^2(0,1))}\right) + \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))}) \|z\|_{L^2((0,1)^2)} \\ &\leq \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))})(\|f_0\|_{L^2(0,1)} + \|z\|_{L^2((0,1)^2)}). \end{aligned}$$

Therefore, $\max_{0 \leq t \leq 1} \|f(\cdot, t)\|_{L^2(0,1)} \leq \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))})(\|f_0\|_{L^2(0,1)} + \|z\|_{L^2((0,1)^2)})$. □

Remark 2.10. We note that we can combine Remark 2.8 and formula (21) to get another representation of the weak solution of $\partial_t z + \partial_x(vz)$. Namely z is characterized by

$$z(\Phi(x, t), t) D\Phi(x, t) = z(x, 0)$$

for a.e. x, t .

Lemma 2.11. *Given $v \in \mathcal{V}$, $z \in L^2((0, 1)^2)$. Let $f_0 \in L^2(0, 1)$. If $f \in L^2((0, 1)^2)$ is a weak solution to the initial value problem*

$$f(\cdot, 0) = f_0, \quad \partial_t f + \partial_x f \cdot v = z,$$

in the sense that for every $\phi \in C^\infty([0, 1]^2)$,

$$(25) \quad \begin{aligned} & - \int_0^1 \int_0^1 f \partial_t \phi \, dx \, dt - \int_0^1 f_0(x) \phi(x, 0) \, dx \\ & = \int_0^1 \int_0^1 f \phi \partial_x v \, dx \, dt + \int_0^1 \int_0^1 f \partial_x \phi v \, dx \, dt + \int_0^1 \int_0^1 \phi z \, dx \, dt. \end{aligned}$$

Then $f \in C(0, 1, H^{-1})$ and

$$\|\partial_t f\|_{L^2(0,1,H^{-1})} \leq \left[1 + \|v\|_{L^2(0,1,H^2)} \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2)}) \right] (\|f_0\|_{L^2(0,1)} + \|z\|_{L^2((0,1)^2)}).$$

Proof. For any $\phi \in L^2((0, 1), H^1(0, 1))$

$$\begin{aligned} & \left| \int_0^1 \int_0^1 (f \partial_x v + z) \phi \, dx \, dt + \int_0^1 \int_0^1 f v \partial_x \phi \, dx \, dt \right| \\ & \leq \int_0^1 \|f(\cdot, t)\|_{L^2} \|\partial_x v(\cdot, t)\|_{L^2} \|\phi(\cdot, t)\|_{L^\infty} + \|f(\cdot, t)\|_{L^2} \|v(\cdot, t)\|_{L^\infty} \|\partial_x \phi(\cdot, t)\|_{L^2} \, dt \\ & \quad + \|z\|_{L^2((0,1)^2)} \|\phi\|_{L^2((0,1)^2)} \\ & \leq \left(\sqrt{2} \|f\|_{L^\infty(0,1,L^2)} \|v\|_{L^2(0,1,H^1)} + \|z\|_{L^2((0,1)^2)} \right) \|\phi\|_{L^2(0,1,H^1)}, \end{aligned}$$

Thus if we define $\partial_t f(t)$ to be the mapping from H^1 to $\mathbb{R}$ defined by

$$\langle \partial_t f(t), g \rangle = \int_0^1 (f(\cdot, t) \partial_x v(\cdot, t) + z(\cdot, t)) g \, dx + \int_0^1 f(\cdot, t) v(\cdot, t) \partial_x g \, dx$$

we see that $\partial_t f \in L^2(0, 1, H^{-1}(0, 1))$ and its norm is controlled by the right hand side of the inequality above. Furthermore from the definition of weak solution (4) follows that $\partial_t f$ is indeed a weak derivative in time of f in H^{-1} . Therefore $f \in C(0, 1, H^{-1})$.

Therefore, combining Lemma 2.9 and the above estimates, we derive that

$$(26) \quad \begin{aligned} \|\partial_t f\|_{L^2(0,1,H^{-1})} & \leq \|z\|_{L^2((0,1)^2)} + \sqrt{2}\|v\|_{L^2(0,1,H^2)} \|f\|_{L^\infty(0,1,L^2)} \\ & \leq [1 + \sqrt{2}\|v\|_{L^2(0,1,H^2)} \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2)})] \cdot (\|f_0\|_{L^2(0,1)} + \|z\|_{L^2((0,1)^2)}) \end{aligned}$$

□

Now we prove Theorem 2.6.

Proof of Theorem 2.6. To apply the direct method of calculus of variations to show the existence of a minimizing path, we show compactness and lower-semicontinuity.

We recall that by claim (i) of Proposition 2.1 we know that there exists a path of finite action. Thus for any minimizing sequence (f_n, v_n, z_n) , we have that $\{v_n\}_{n \in \mathbb{N}}$ is bounded in $L^2(0, 1, H^2 \cap H_0^1)$ and $\{z_n\}_{n \in \mathbb{N}}$ is bounded in $L^2((0, 1)^2)$. Since $(L^2((0, 1), H^s(0, 1)))^* = L^2((0, 1), H^{-s}(0, 1))$ for any s , Banach-Alaoglu Theorem allows us to extract a subsequence, still denoted by (v_n, z_n) , such that $v_n \rightharpoonup v$ and $z_n \rightharpoonup z$ weakly in $L^2(0, 1, H^2(0, 1) \cap H_0^1(0, 1))$ and $L^2((0, 1)^2)$, respectively.

Since $H^2(0, 1) \hookrightarrow C^1(0, 1)$, $v_n \in L^2(0, 1, C^1(0, 1))$. By Lemma 2.7, f_n has the representation (16) with v, z, Φ replaced by v_n, z_n, Φ_n respectively.

We claim that along a subsequence

$$(27) \quad f_n \rightarrow f \quad \text{in } L^2(0, 1, H^{-\frac{1}{2}}(0, 1)).$$

First notice that by Lemma 2.9, $\{f_n\}_{n \in \mathbb{N}}$ is a bounded set in $L^\infty(0, 1, L^2(0, 1))$. There exists $f \in L^\infty(0, 1, L^2(0, 1))$ such that up to a subsequence, $f_n \xrightarrow{*} f$ in $L^\infty(0, 1, L^2(0, 1))$. By relabeling we assume that the subsequence is the whole sequence. In addition,

$$\|f\|_{L^\infty(0,1,L^2(0,1))} \leq \liminf_{n \rightarrow \infty} \|f_n\|_{L^\infty(0,1,L^2(0,1))}.$$

Applying Lemma 2.11 to (f_n, v_n, z_n) , $\{\partial_t f_n\}_{n \in \mathbb{N}}$ is a bounded set in $L^1(0, 1, H^{-1}(0, 1))$. By Lions-Aubin lemma [12], since $\{f_n\}_{n \in \mathbb{N}}$ is a bounded set in $L^2(0, 1, L^2(0, 1))$, and $H^{-1} \hookrightarrow H^{-\frac{1}{2}} \hookrightarrow L^2$ compactly, there exists $\tilde{f}$ in $L^2(0, 1, H^{-\frac{1}{2}}(0, 1))$, such that up to a subsequence (still denoted by f_n) such that $f_n \rightarrow \tilde{f}$ in $L^2(0, 1, H^{-\frac{1}{2}}(0, 1))$. In fact, $\tilde{f} = f$. Namely f is the weak $*$ limit of f_n in $L^\infty(0, 1, L^2(0, 1))$, and thus the weak limit in $L^2(0, 1, H^{-\frac{1}{2}}(0, 1))$. By the uniqueness of weak limit, $\tilde{f} = f$.

To summarize: there exists $f \in L^2((0, 1)^2)$, $v \in \mathcal{V}$ and $z \in L^2((0, 1)^2)$ such that

$$\begin{aligned} f_n &\xrightarrow{*} f && \text{in } L^\infty(0, 1, L^2(0, 1)) \\ f_n &\rightarrow f && \text{in } L^2(0, 1, H^{-\frac{1}{2}}(0, 1)) \\ z_n &\rightarrow z && \text{in } L^2(0, 1, L^2(0, 1)) \\ v_n &\rightarrow v && \text{in } L^2(0, 1; H^2(0, 1)). \end{aligned}$$

Now pass to the limit in (4). Since $\phi \in C^\infty([0, 1]^2)$. First consider the term $\int_0^1 \int_0^1 f_n \phi \partial_x v_n \, dx \, dt$.

$$\begin{aligned} &\left| \int_0^1 \int_0^1 f_n \phi \partial_x v_n \, dx \, dt - \int_0^1 \int_0^1 f \phi \partial_x v \, dx \, dt \right| \\ &\leq \left| \int_0^1 \int_0^1 f_n \phi \partial_x v_n \, dx \, dt - \int_0^1 \int_0^1 f \phi \partial_x v_n \, dx \, dt \right| \\ &\quad + \left| \int_0^1 \int_0^1 f \phi \partial_x v_n \, dx \, dt - \int_0^1 \int_0^1 f \phi \partial_x v \, dx \, dt \right| \\ &\leq \int_0^1 \|f_n - f\|_{H^{-\frac{1}{2}}(0,1)} \|\phi \partial_x v_n\|_{H^{\frac{1}{2}}(0,1)} \, dt + \left| \int_0^1 \int_0^1 f \phi (\partial_x v_n - \partial_x v) \, dx \, dt \right| \\ &\leq \|f_n - f\|_{L^2((0,1), H^{-\frac{1}{2}}(0,1))} \sup_{n \in \mathbb{N}} \|\phi \partial_x v_n\|_{L^2((0,1), H^{\frac{1}{2}}(0,1))} \\ &\quad + \left| \int_0^1 \int_0^1 f \phi (\partial_x v_n - \partial_x v) \, dx \, dt \right| =: A_1 + B_1, \end{aligned}$$

Consider A_1 . For every t , $\|\phi \partial_x v_n\|_{H^{\frac{1}{2}}} \leq C \|\phi\|_{H^1} \|\partial_x v_n\|_{H^{\frac{1}{2}}} \leq C \|\phi\|_{C^1} \|\partial_x v_n\|_{H^{\frac{1}{2}}}$ by multiplication of Sobolev functions (see Theorem 7.4 of [5]). Then

$$\begin{aligned} \sup_{n \in \mathbb{N}} \|\phi \partial_x v_n\|_{L^2((0,1), H^{\frac{1}{2}}(0,1))} &\leq C \max_{t \in [0,1]} \|\phi(t)\|_{C^1} \sup_{n \in \mathbb{N}} \|\partial_x v_n\|_{L^2((0,1), H^{\frac{1}{2}}(0,1))} \\ &\leq C_\phi \sup_{n \in \mathbb{N}} \|v_n\|_{L^2((0,1), H^2(0,1))}, \end{aligned}$$

where the constant C_ϕ only depends on ϕ . Combining with $f_n \rightarrow f$ in $L^2(0, 1, H^{-\frac{1}{2}}(0, 1))$, we obtain that $A_1 \rightarrow 0$ as $n \rightarrow \infty$. $B_1 \rightarrow 0$ since $\phi f \in L^2(0, 1, L^2(0, 1))$ which could serve as a test function for $\partial_x v_n \rightarrow \partial_x v$ in $L^2(0, 1, H^1(0, 1))$. By analogous argument,

$$\left| \int_0^1 \int_0^1 f_n v_n \partial_x \phi \, dx \, dt - \int_0^1 \int_0^1 f v \partial_x \phi \, dx \, dt \right| \rightarrow 0.$$

For the other terms in (4), passing to the limit is straightforward. Therefore f satisfies (4). By Lemma 2.11, $f \in C(0, 1, H^{-1}(0, 1))$.

The lower semicontinuity of the action function follows directly from the lower-semicontinuity of norms with respect to the weak convergence. Therefore (f, v, z) is a minimizer. $\square$

Showing that d_{HV} is a metric on $L^2(0, 1)$ is straightforward. In particular the triangle inequality is obtained by concatenating minimizers with arc-length parameterization in which

$$d_{HV(\kappa, \lambda, \varepsilon)} = \int_0^1 \sqrt{\frac{1}{2} \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 dx} dt = \sqrt{A_{\kappa, \lambda, \varepsilon}(f, v, z)}.$$

The existence of such reparameterization follows from reparameterization result which states that any admissible path (f, v, z) can be reparameterized in time by any absolutely continuous diffeomorphism of $[0, 1]$. The lemma is a just variation of the lemma on rescaling of distributional solutions to continuity equations (see Lemma 8.1.3 of [2]).

Lemma 2.12. *Let $t := s \in [0, 1] \rightarrow t(s) \in [0, 1]$ be strictly increasing absolutely continuous map with absolutely continuous inverse $s := t^{-1}$. In addition, $t(0) = 0, t(1) = 1$. Then $(f(x, t), v(x, t), z(x, t))$ is a weak solution of*

$$\partial_t f + \partial_x f \cdot v = z \quad \text{in } [0, 1] \times [0, 1]$$

with boundary condition $f(\cdot, 0) = f_0, f(\cdot, 1) = f_1$ if and only if

$$\hat{f}(x, s) =: f(x, t(s)), \quad \hat{v}(x, s) = t'(s)v(x, t(s)), \quad \hat{z}(x, s) := t'(s)z(x, t(s))$$

is the weak solution of solution of

$$\partial_s \hat{f} + \nabla \hat{f} \cdot \hat{v} = \hat{z} \quad \text{in } [0, 1] \times [0, 1]$$

with boundary condition $\hat{f}(\cdot, 0) = f_0, \hat{f}(\cdot, 1) = f_1$.

Lemma 2.13. *The distance d_{HV} is metric on $L^2(0, 1)$.*

Proof. We claim that the positivity follows from the existence of minimizers in Proposition 2.6. If $f_0 = f_1$ in $L^2(0, 1)$, it is obvious that $d_{HV}(f_0, f_1) = 0$. Consider the case $f_0 \neq f_1$ in $L^2(0, 1)$. Observe that for $(f, v, z) \in \mathcal{A}$, we could reparameterize by arc length in time to make the quantity $\left(\int_0^1 \kappa v^2(x, s) + \lambda v_x^2(x, s) + \varepsilon v_{xx}^2(x, s) + z^2(x, s) dx \right)^{\frac{1}{2}}$ constant in s , with the reparameterization denoted by $(\hat{f}, \hat{v}, \hat{z})$. In particular, the reparameterization and its inverse are almost everywhere differentiable. Applying Lemma 2.12 and changing variables, by Jensen's inequality we have

$$A_{(\kappa, \lambda, \varepsilon)}(\hat{f}, \hat{v}, \hat{z}) \leq A_{(\kappa, \lambda, \varepsilon)}(f, v, z).$$

According to Jensen's inequality, if (f, v, z) is a minimizer of $A(f, v, z)$ then it is simultaneously the minimizer of $\int_0^1 \sqrt{\frac{1}{2} \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 dx} dt$. The existence of minimizer implies that $d_{HV}(f_0, f_1) > 0$, since otherwise the minimizer $f(x, t)$ is constant and $f_0 \equiv f(\cdot, t) \equiv f_1$ which yields contradiction.

The symmetry is direct from the definition. The triangle inequality follows from path concatenation. Consider $f_0, f_1, f_2 \in L^2(0, 1)$. Let (f_{01}, v_{01}, z_{01}) be the minimizing path between f_0 and f_1 and let (f_{12}, v_{12}, z_{12}) be the minimizing path between f_1 and f_2 . Define

$$\tilde{f}(t) = \begin{cases} f_{01}(2t) & \text{for } t \in [0, \frac{1}{2}] \\ f_{12}(2t - 1) & \text{for } t \in (\frac{1}{2}, 1]. \end{cases}$$

$\tilde{v}(t)$ and $\tilde{z}(t)$ are defined analogously. It immediately gives that $d_{HV}(f_0, f_2) \leq d_{HV}(f_0, f_1) + d_{HV}(f_1, f_2)$. $\square$

2.4. Completeness of the d_{HV} metric space and the characterization of its topology. Thus far, we have shown that d_{HV} is a metric on L^2 . We now establish completeness.

Proposition 2.14. *The metric space $(L^2(0, 1), d_{HV})$ is complete.*

Proof. Notice that $d_{HV}(\kappa, \lambda, \varepsilon) \asymp d_{HV(1,1,1)}$, it suffices to consider $d_{HV(1,1,1)}$.

Let $\{g_n\}_{n \in \mathbb{N}} \subset L^2(0, 1)$ be a Cauchy sequence in d_{HV} . Without a loss of generality, we assume that for all $n \in \mathbb{N}$, $d_{HV}(g_1, g_n) \leq 1$. By Proposition 2.6, for $n > 1$, let $(g_{1,n}, v_{1,n}, z_{1,n})$ denotes the minimizer of the action $A(f, v, z)$ with the admissible set $\mathcal{A}(g_1, g_n)$. Then Lemma 2.9 and the definition of d_{HV} give that

$$\begin{aligned} \|g_n\|_{L^2(0,1)} &\leq \exp(\sqrt{2}\|v_{1,n}\|_{L^2(0,1,H^2(0,1))})(\|g_1\|_{L^2(0,1)} + \|z_{1,n}\|_{L^2((0,1)^2)}) \\ &\leq \exp(\sqrt{2}d_{HV}(g_1, g_n))(\|g_1\|_{L^2(0,1)} + d_{HV}(g_1, g_n)) \\ &\leq \exp(\sqrt{2})(\|g_1\|_{L^2(0,1)} + 1). \end{aligned}$$

Then Banach-Alaoglu Theorem gives that there exists a subsequence g_{n_k} converges to some $g_\infty \in L^2(0, 1)$ weakly in $L^2(0, 1)$. We denote the subsequence as $\{\tilde{g}_n\}_{n \in \mathbb{N}}$.

On the other hand, there exists further subsequences (which we relabel to be the original subsequence) $\{v_{1,n}\}, \{z_{1,n}\}$ and $v_{1,\infty}, z_{1,\infty}$ such that

$$\begin{aligned} v_{1,n} &\rightharpoonup v_{1,\infty}, \quad \text{weakly in } L^2(0, 1, H^2(0, 1)), \\ z_{1,n} &\rightharpoonup z_{1,\infty}, \quad \text{weakly in } L^2((0, 1)^2). \end{aligned}$$

Meanwhile, by Lemma 2.11, $\{\partial_t g_{1,n}\}_{n \in \mathbb{N}}$ is a bounded set in $L^2(0, 1, H^{-1}(0, 1))$. Then Lions-Aubin Lemma gives that up to a subsequence $g_{1,n} \rightarrow g_{1,\infty}$ in $L^2(0, 1, H^{-\frac{1}{2}}(0, 1))$. Moreover we have $g_{1,n} \rightharpoonup g_{1,\infty}$ weakly $*$ in $L^\infty(0, 1, L^2(0, 1))$.

Up to a subsequence,

$$\begin{aligned} g_n &\rightarrow g_\infty \quad \text{in } L^2(0, 1) \\ g_{1,n} &\xrightarrow{*} g_{1,\infty} \quad \text{in } L^\infty(0, 1, L^2(0, 1)) \\ g_{1,n} &\rightarrow g_{1,\infty} \quad \text{in } L^2(0, 1, H^{-\frac{1}{2}}(0, 1)) \\ z_{1,n} &\rightharpoonup z_{1,\infty} \quad \text{in } L^2(0, 1, L^2(0, 1)) \\ v_{1,n} &\rightharpoonup v_{1,\infty} \quad \text{in } L^2(0, 1, H^2(0, 1)). \end{aligned}$$

By a similar argument as in the proof of Theorem 2.6, we could verify that $(g_{1,\infty}, v_{1,\infty}, z_{1,\infty})$ is in the admissible set $\mathcal{A}(g_1, g_\infty)$. Thus

$$d_{HV}^2(g_1, g_\infty) \leq \liminf_{n \rightarrow \infty} \|v_{1,n}\|_{L^2(0,1,H^2(0,1))}^2 + \liminf_{n \rightarrow \infty} \|z_{1,n}\|_{L^2((0,1)^2)}^2 \leq \liminf_{n \rightarrow \infty} d_{HV}^2(g_1, g_n).$$

Then without loss of generality, we could take a further sequence in $\{\tilde{g}_n\}_{n \in \mathbb{N}}$ such that for every n , for any $m \geq n$, $d_{HV}(\tilde{g}_m, \tilde{g}_n) < \frac{1}{n}$. We repeat the above arguments with g_1 replaced by $\tilde{g}_n$. Then we have

$$d_{HV}^2(\tilde{g}_n, g_\infty) \leq \liminf_{m \rightarrow \infty} d_{HV}^2(\tilde{g}_m, \tilde{g}_n) < \frac{1}{n^2}.$$

Now we have a subsequence $\{\tilde{g}_n\}_{n \in \mathbb{N}}$ that converges to g_∞ in d_{HV} . For any $\delta > 0$, there exists $K_1 \in \mathbb{N}$ such that $\frac{1}{K_1} < \frac{\delta}{2}$. And there exists $K_2 > 0$ such that for all $m_1, m_2 > K_2$, $d_{HV}(g_{m_1}, g_{m_2}) < \frac{\delta}{2}$. Choose $N = \max\{K_1, K_2\}$,

$$d_{HV}(g_n, g_\infty) \leq d_{HV}(g_n, \tilde{g}_N) + d_{HV}(\tilde{g}_N, g_\infty) < \delta,$$

which gives that $g_n \rightarrow g_\infty$ in d_{HV} . □

Our next goal is to show that convergence in d_{HV} implies convergence in L^2 . Towards that goal we first prove the following estimate.

Lemma 2.15. *For any $f_0, f_1 \in L^2(0, 1)$ and $g_0 \in C^\infty$. Let (f, v, z) be a minimizer of the action with the admissible set $\mathcal{A}(f_0, f_1)$ and $\psi := \Phi(\cdot, 1)$, where $\Phi(x, t)$ is the flow map of $v(x, t)$. Define $g_1 = g_0 \circ \psi^{-1}$.*

Then $g_1 \in C^1$, satisfies $\int_0^1 |g'_1| dx = \int_0^1 |g'_0| dx$ and

$$(28) \quad \|f_1 - g_1\|_{L^2} \leq \exp(\sqrt{2}d_{HV}(f_0, f_1)) [\|f_0 - g_0\|_{L^2} + d_{HV}(f_0, f_1)].$$

Proof. ψ^{-1} is a C^1 diffeomorphism immediately gives that $g_1 \in C^1$ and $\int_0^1 |g'_0| dx = \int_0^1 |g'_1| dx$. Recall that by (22), we have

$$|D\Phi(\cdot, t)| \leq \exp(\sqrt{2}t\|v\|_{L^2(0,1,H^2(0,1))}), \quad \forall t \in [0, 1].$$

In particular,

$$(29) \quad |D\psi| \leq \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))}).$$

By Lemma 2.7, for a.e. $y \in [0, 1]$,

$$f_1(y) - g_1(y) = f_0(\psi^{-1}(y)) + \int_0^1 z(\Phi(\psi^{-1}(y), s), s) ds - g_0(\psi^{-1}(y)).$$

By an analogous estimate as in the proof of Lemma 2.9, we have

$$(30) \quad \begin{aligned} \|f_1 - g_1\|_{L^2} &\leq \left\| \int_0^1 z(\Phi(\psi^{-1}(\cdot), s), s) ds \right\|_{L^2} + \|f_0 \circ \psi^{-1} - g_0 \circ \psi^{-1}\|_{L^2} \\ &\leq \exp(\sqrt{2}\|v\|_{L^2(0,1,H^2(0,1))}) \|z\|_{L^2((0,1)^2)} + \exp\left(\frac{\sqrt{2}}{2}\|v\|_{L^2(0,1,H^2(0,1))}\right) \|f_0 - g_0\|_{L^2} \\ &\leq \exp(\sqrt{2}d_{HV}(f_0, f_1)) [\|f_0 - g_0\|_{L^2} + d_{HV}(f_0, f_1)]. \end{aligned}$$

□

Theorem 2.16. *Let $\{f_n\}_{n \in \mathbb{N}} \subseteq L^2(0, 1)$. If $f_n \rightarrow f$ in d_{HV} , then $f_n \rightarrow f$ in L^2 .*

Proof. It suffices to show the conclusion $d_{HV(1,1,1)}$. We will prove this by showing that for any subsequence of $\{f_n\}_{n \in \mathbb{N}}$, there exists a further subsequence that converges to f in L^2 .

Let $K := 2 \exp(\sqrt{2})$. Let $0 < \delta < 1$ be arbitrarily chosen. By extracting a subsequence, we can assume that for every n , $d_{HV}(f_n, f) \leq \frac{\delta}{3K} < 1$. Since C^∞ is dense in L^2 , there exists $g \in C^\infty$ such that $\|f - g\|_{L^2} \leq \frac{\delta}{3K}$. For any n , let ψ_n be the time-1 flow map of v_n , where (f_n, v_n, z_n) denotes the minimizer of the action $A(f, v, z)$ with the admissible set $\mathcal{A}(f, f_n)$. Take $g_n := g \circ \psi_n^{-1}$ if $f_n \neq f$, and take $g_n = g$ if $f_n = f$.

We will show that $\{g_n\}_{n \in \mathbb{N}}$ is a bounded set in $W^{1,2}(0, 1)$. Thus without loss of generality, we assume $f_n \neq f$. By Lemma 2.15,

$$(31) \quad \|f_n - g_n\|_{L^2} \leq \exp(\sqrt{2}d_{HV}(f, f_n)) [\|f - g\|_{L^2} + d_{HV}(f, f_n)] < \frac{\delta}{3}.$$

On the other hand, by change of variables,

$$\|g_n\|_{L^2}^2 \leq \|g\|_{L^2}^2 \|D\psi_n\|_{L^\infty} \leq \|g\|_{L^2}^2 \exp(\sqrt{2}\|v_n\|_{L^2(0,1,H^2(0,1))}),$$

where we use the estimate (29) with ψ, v replaced by ψ_n, v_n . Then we have

$$\|g_n\|_{L^2} \leq \|g\|_{L^2} \exp\left(\frac{\sqrt{2}}{2}d_{HV}(f, f_n)\right) \leq \|g\|_{L^2} \exp\left(\frac{\sqrt{2}}{2}\right),$$

which implies that $\{g_n\}_{n \in \mathbb{N}}$ is a bounded set in L^2 . Since ψ_n^{-1} is the time-1 flow of $v_n(\cdot, 1-t)$, by (21),

$$\int_0^1 |Dg_n|^2 dx \leq \exp(\sqrt{2}\|v_n\|_{L^2(0,1,H^2(0,1))})\|g'\|_{L^2}^2 \leq \exp(\sqrt{2}d_{HV}(f, f_n))\|g'\|_{L^2}^2$$

which implies that $\{Dg_n\}_{n \in \mathbb{N}}$ is a bounded set in L^2 . Thus $g_n \in W^{1,2}(0,1)$ and

$$\sup_n \|g_n\|_{W^{1,2}} < \exp(\sqrt{2})[\|g\|_{L^2} + \|g'\|_{L^2}^2].$$

Hence $\{g_n\}_{n \in \mathbb{N}}$ is a bounded set in $W^{1,2}(0,1)$. By Morrey's inequality and Arzela-Ascoli compactness criterion, there exists a subsequence g_{n_k} that is Cauchy in L^2 . For sufficiently large N , for any $i, j > N$, we have $\|g_{n_i} - g_{n_j}\|_{L^2} < \frac{\delta}{3}$. Thus

$$\|f_{n_i} - f_{n_j}\|_{L^2} \leq \|f_{n_j} - g_{n_j}\|_{L^2} + \|g_{n_j} - g_{n_i}\|_{L^2} + \|f_{n_i} - g_{n_i}\|_{L^2} < \delta.$$

This implies that up to a subsequence, $\{f_{n_k}\}_{k \in \mathbb{N}}$ is a Cauchy sequence in L^2 . By completeness of L^2 , it converges to some $\tilde{f} \in L^2$. $\tilde{f}$ is also the weak limit of the subsequence in L^2 , by the uniqueness of weak limit, we should have $\tilde{f} = f$ (recall that f is the weak limit of a subsequence of $\{f_{n_k}\}_{k \in \mathbb{N}}$).

Therefore, up to a subsequence, $f_n \rightarrow f$ in L^2 . For any subsequence $f_n \rightarrow f$, it admits a further subsequence that converges to f . This proves that for the whole sequence, $f_n \rightarrow f$ in L^2 . $\square$

3. PROPERTIES OF GEODESICS.

3.1. Euler-Lagrange Equations. Assume $(f, v, z) \in \mathcal{A}(f_0, f_1)$ is a minimizer of the action (2) over the admissible set (3). Below we describe the first-order optimality conditions, first for general paths, (33) and (36), and then under assumption that $f_0, f_1 \in H^1(0,1)$, in which case all conditions are partial differential equations.

To find the first-order optimality conditions, in other words the Euler-Lagrange equations, we first fix f and perform the first variation in v and z . Motivated by a similar reasoning as in Section 2.2, we introduce the space,

$$(32) \quad R_T(f) = \{u \in L^2(0,1, H^2(0,1) \cap H_0^1(0,1)) : uf \in L^2(0,1, H^1(0,1))\}.$$

Note that for $u \in R_T(f)$, for a.e. $t \in [0,1]$, $uf_x = (uf)_x - u_x f \in L^2((0,1)^2)$.

Now we characterize the optimality condition analogous to that in Lemma 2.5. Taking $u \in R_T(f)$ and $h = uf_x$, which belongs to $L^2((0,1)^2)$, we have that, for any $s \in \mathbb{R}$, $(f, v + su, z + sh)$ satisfies the equation (4), and thus $(f, v + su, z + sh) \in \mathcal{A}(f_0, f_1)$. First variation of action (2) gives that (f, v, z) satisfies:

$$(33) \quad \forall u \in R_T(f) \quad \int_0^1 \int_0^1 \kappa v u + \lambda v_x u_x + \varepsilon v_{xx} u_{xx} + z(uf_x) dx dt = 0.$$

We then turn to the optimality of z . To carry out the argument let us denote that minimizer of the action considered by $(\bar{f}, \bar{v}, \bar{z})$. Then $(\bar{f}, \bar{z})$ is a critical point of $\tilde{A}(z) = \int_0^1 \int_0^1 z^2 dx dt$ over the set of $(f, v, z) \in \mathcal{A}(f_0, f_1)$ such that $v = \bar{v}$. This is a convex functional over linear constraint. Thus $(\bar{f}, \bar{z})$ is a global minimizer for fixed $\bar{v}$.

Furthermore by Lemma 2.7, the constraint that $(f, v, z) \in \mathcal{A}(f_0, f_1)$, with $v = \bar{v}$ can be expressed as follows: for a.e. x, t

$$(34) \quad f(\Phi(x, t), t) = f_0(x) + \int_0^t z(\Phi(x, s), s) ds.$$

In other words z needs to satisfy that for a.e. $x \in (0,1)$

$$(35) \quad f_1(\Phi(x, 1)) - f_0(x) = \int_0^1 z(\Phi(x, t), t) dt$$

while $f(\cdot, t)$ for $0 < t < 1$ is defined by (34). Since f does not enter the action directly, we minimize over z alone and define f by (34). After the change of variables $\tilde{\mathcal{A}}(z) = \int_0^1 \int_0^1 z^2(\Phi(x, t), t) D\Phi(x, t) dx dt$. Hence if we define $w(x, t) = z(\Phi(x, t), t)$ the problem is transformed to

$$\begin{aligned} & \text{minimize } \int_0^1 \int_0^1 w^2(x, t) D\Phi(x, t) dx dt \\ & \text{under constraint: } f_1(\Phi(x, 1)) - f_0(x) = \int_0^1 w(x, t) dt \quad \text{for a.e. } x \in (0, 1). \end{aligned}$$

By Cauchy-Schwarz inequality, the minimizer w of this problem should satisfy that $w(x, t) D\Phi(x, t)$ is independent of time. Thus z which minimizes $\tilde{\mathcal{A}}$ satisfies that for a.e. x

$$z(\Phi(x, t), t) D\Phi(x, t) =: z(x, 0).$$

By Remark 2.10, this implies that z is a weak solution of

$$(36) \quad \partial_t z + \partial_x(zv) = 0 \quad \text{on } (0, 1)^2,$$

which is the desired first-order condition on z .

If $f_0, f_1 \in H^1(0, 1)$ then we establish in the Proposition 3.1 below that the minimizing path $f \in L^2(0, 1, H^1(0, 1))$. In that case we the resulting Euler–Lagrange equations satisfied by f, v, z , combined with the condition for belonging to $\mathcal{A}(f_0, f_1)$ can be expressed as follows:

$$(37) \quad \varepsilon v_{xxxx} - \lambda v_{xx} + \kappa v + z f_x = 0 \text{ weakly on } (0, 1)^2$$

$$(38) \quad v = 0 \text{ and } v_{xx} = 0 \text{ on } \partial\Omega \times [0, 1]$$

$$(39) \quad z_t + (zv)_x = 0 \text{ weakly on } (0, 1)^2$$

$$(40) \quad f_t + f_x v - z = 0 \text{ weakly on } (0, 1)^2$$

$$(41) \quad f(\cdot, 0) = f_0, \text{ and } f(\cdot, 1) = f_1.$$

The first equations, and boundary conditions, follow directly from (33). The other conditions are identical as before.

3.2. Regularity of geodesics. This subsection presents results regarding the regularity of the geodesics.

Proposition 3.1. *Assume $f_0, f_1 \in H^1$ and let $(f, v, z) \in \mathcal{A}(f_0, f_1)$ be an action minimizing path. Then $f \in L^\infty(0, 1, H^1(0, 1))$, with quantitative estimates on the norm (45), (46).*

Proof. We remark that the existence of action minimizing $(f, v, z) \in \mathcal{A}(f_0, f_1)$ is guaranteed by Theorem 2.6. The argument in Section 3.1 implies that z solves $z_t + (zv)_x = 0$, weakly. As in Remark 2.8, z satisfies the formula (19) and f satisfies (20). By Theorem 2.2.2 in [42], since $\Phi(\cdot, t)$ is bi-Lipschitz in x it suffices to check if $f(\Phi(x, t), t) \in H^1$. We use the composition with Φ in several instances below.

Let $b := e^{\sqrt{2}\|v\|_{L^2(0,1,H^2)}}$. Recall from (22) that for all $x \in [0, 1]$ and $t \in [0, 1]$

$$\frac{1}{b} \leq |D\Phi(x, t)| \leq b.$$

Note that for all t , $\|\int_0^t v_x(\Phi(\cdot, s), s) ds\|_{L^\infty} \leq \|v_x\|_{L^1(0,1,L^\infty)} \leq \sqrt{2}\|v\|_{L^2(0,1,H^2)}$. Thus for all (x, t)

$$(42) \quad \frac{1}{b} \leq J(x, t) \leq e^{\sqrt{2}\|v\|_{L^2(0,1,H^2)}} = b.$$

Change of variables provides that, for a.e. t ,

$$\int_0^1 (\partial_x v_x(\Phi(x, t), t))^2 dx \leq \|D\Phi(\cdot, t)\|_{L^\infty} \|v_{xx}(t)\|_{L^2((0,1))}^2 \leq b \|v\|_{L^2(0,1,H^2)}.$$

Integrating in t and Cauchy-Schwarz inequality imply that for all $t \in [0, 1]$

$$\left\| \partial_x \int_0^t v_x(\Phi(x, s), s) ds \right\|_{L^2(0,1)}^2 \leq tb \|v\|_{L^2(0,1,H^2)}^2.$$

Combining with the chain rule and the L^∞ estimate on J implies that $J \in L^\infty(0, 1, H^1(0, 1))$ and

$$(43) \quad \forall t \in [0, 1] \quad \left\| \int_0^t J(x, s) ds \right\|_{H^1(0,1)} \leq b + b^{3/2} \|v\|_{L^2(0,1,H^2)}$$

Since $J \geq \frac{1}{b}$, $1/\int_0^1 J(x, t) dt$ is in $H^1(0, 1)$. By chain rule [18],

$$(44) \quad \left\| \partial_x \left(1/\int_0^1 J(x, t) dt \right) \right\|_{L^2} \leq b^2 \left\| \partial_x \left(\int_0^1 J(x, t) dt \right) \right\|_{L^2} \leq b^{7/2} \|v\|_{L^2(0,1,H^2)}.$$

Moreover by change of variables and estimate (22), $\|\partial_x f_1(\Phi(x, 1))\|_{L^2} \leq \sqrt{b} \|\partial_x f_1\|_{L^2}$. From (19), (42), (43), and (44), via the product rule and using that $\|gh\|_{L^2} \leq \|g\|_{L^2} \|h\|_{L^\infty}$, we obtain

$$\begin{aligned} \|\partial_x z(\Phi(x, \cdot), \cdot)\|_{L^2((0,1)^2)} &\leq (\|\partial_x f_1(\Phi(\cdot, 1))\|_{L^2} + \|\partial_x f_0\|_{L^2}) b^2 \\ &\quad + (\|f_1(\Phi(\cdot, 1))\|_{L^\infty} + \|f_0\|_{L^\infty}) (b^{5/2} + b^{9/2}) \|v\|_{L^2(0,1,H^2)} \\ &\leq 3(\|f_1\|_{H^1} + \|f_0\|_{H^1}) (b^{5/2} + b^{9/2}) \|v\|_{L^2(0,1,H^2)}. \end{aligned}$$

Note also that, via change of variables,

$$\int_0^1 \int_0^1 z^2(\Phi(x, t), t) dx dt \leq b \|z\|_{L^2((0,1)^2)}^2 \leq b \|f_1 - f_0\|_{L^2}^2.$$

From (20), by a change of variables and estimate (22), for every $t \in [0, 1]$,

$$(45) \quad \|\partial_x f(\cdot, t)\|_{L^2(0,1)} \leq \sqrt{b} \|\partial_x f(\Phi(\cdot, t), t)\|_{L^2} \leq 6(\|f_1\|_{H^1} + \|f_0\|_{H^1}) (1 + b^3 + b^5) \|v\|_{L^2(0,1,H^2)},$$

which using that $\|v\|_{L^2(0,1,H^2)} \leq C_\varepsilon \|f_1 - f_0\|_{L^2}$ can be turned in an estimate where right-hand side depends only on initial and final signal.

The estimate on $\|f(\cdot, t)\|_{L^2}$ from Lemma 2.9 and the estimate on z above provide that

$$(46) \quad \|f(\cdot, t)\|_{L^2(0,1)} \leq b(\|f_0\|_{L^2} + \sqrt{b} \|f_0 - f_1\|_{L^2}),$$

which completes the proof. $\square$

Remark 3.2. The regularity above can be improved to spaces with more regularity. Here we now outline argument that if f_0 and f_1 are smooth then $f(\cdot, t)$ is smooth for all $t \in [0, 1]$. For $f_0, f_1 \in H^1(0, 1)$ Proposition 3.1 gives that the minimizing path $f \in L^\infty(0, 1, H^1(0, 1))$ and $z \in L^\infty(0, 1, H^1(0, 1))$. Thus in equation (37), $zf_x \in L^2(0, 1)$ for every $t \in [0, 1]$. The results on elliptic boundary value problems imply $v(\cdot, t) \in H^4(0, 1)$. Moreover, there exists a constant C such that

$$\|v(\cdot, t)\|_{H^4(0,1)} \leq C(\|zf_x(\cdot, t)\|_{L^2} + \|v(\cdot, t)\|_{H^2}).$$

Thus

$$\|v\|_{L^2(0,1,H^4)} \leq C(\|z\|_{L^2(0,1,H^1)} \|f\|_{L^2(0,1,H^1)} + \|v\|_{L^2(0,1,H^2)}),$$

where the right-hand side, by Proposition 3.1, is a function only depending on f_0 and f_1 .

According the differentiability of ODE solution with respect to parameters, we obtain that for fixed t , the flow map $\Phi(x, t) \in C^3(0, 1)$. In addition, $\Phi(x, t) \in L^\infty(0, 1, C^3)$. Indeed, $D^2\Phi(x, t)$ satisfies the equation:

$$\partial_t D^2\Phi(x, t) = (v_x(\Phi(x, t), t))^2 (D\Phi(x, t))^2 + v_x(\Phi(x, t), t) D^2\Phi(x, t).$$

By Gronwall's inequality and (22) we obtain that $D^2\Phi \in L^\infty((0, 1)^2)$. Inductively, $\Phi \in L^\infty(0, 1, C^3)$.

Now we assume that $f_0, f_1 \in H^4(0, 1)$. The above discussion gives that $v \in L^2(0, 1, H^4)$ and $\Phi \in L^\infty(0, 1, C^3)$. Applying chain rule to (19) and following a similar argument as in the proof of Proposition 3.1, we have that $z \in L^\infty(0, 1, H^3(0, 1))$. Then by (20), $f \in L^\infty(0, 1, H^3(0, 1))$, with $\|f\|_{L^\infty(0, 1, H^3(0, 1))}$ bounded by a function of f_0 and f_1 .

By an iterative argument, we conclude that for any $k \in \mathbb{N}$, $f, z \in L^\infty(0, 1, H^k(0, 1))$ and $v \in L^2(0, 1, H^k(0, 1))$, which implies the smoothness of $f(\cdot, t)$.

3.3. Stability of geodesics. Next, we study the stability of the geodesic in terms of a subsequence approximation.

Proposition 3.3. *Assume $f_0, f_1 \in L^2(0, 1)$, $f_0^n, f_1^n \in L^2(0, 1)$ for all $n \in \mathbb{N}$, and $f_0^n \rightarrow f_0$, $f_1^n \rightarrow f_1$ in $L^2(0, 1)$ as $n \rightarrow \infty$. Let $(f^n, v^n, z^n) \in \mathcal{A}(f_0^n, f_1^n)$ be action minimizing paths. Then there exists $(f, v, z) \in \mathcal{A}(f_0, f_1)$ such that along a subsequence*

$$\begin{aligned} f^n &\xrightarrow{*} f \quad \text{in } L^\infty((0, 1), L^2(0, 1)) \\ f^n &\rightarrow f \quad \text{in } C((0, 1), (L^2(0, 1), d_{HV})) \\ z^n &\rightharpoonup z \quad \text{in } L^2((0, 1), L^2(0, 1)) \\ v^n &\rightharpoonup v \quad \text{in } L^2([0, 1]; H^2(0, 1)). \end{aligned}$$

Furthermore (f, v, z) is an action minimizing path between f_0 and f_1 .

Proof. Since

$$(47) \quad d_{HV}(f_0^n, f_1^n) \leq \|f_0^n - f_0\|_{L^2} + d_{HV}(f_0, f_1) + \|f_1^n - f_1\|_{L^2},$$

we conclude that $A(f^n, v^n, z^n) \leq d_{HV}^2(f_0, f_1) + 1 \leq \|f_0 - f_1\|_{L^2}^2 + 1$ for all n large enough. We can assume, without loss of generality that the inequalities hold for all n . The proof of weak convergence mirrors the proof of Theorem 2.6. To show that (f, v, z) is an action minimizing path, we note that by lower-semicontinuity $A(f, v, z) \leq \liminf_{n \rightarrow \infty} A(f^n, v^n, z^n)$. Since $A(f^n, v^n, z^n) = d_{HV}^2(f_0^n, f_1^n)$, from (47) follows that $A(f, v, z) \leq d_{HV}^2(f_0, f_1)$. Thus $A(f, v, z) = d_{HV}^2(f_0, f_1)$.

We claim that for all t , $f^n(\cdot, t)$ is precompact in L^2 . We note that by Lemma 2.9 the sequence $f^n(\cdot, t)$ is uniformly bounded in L^2 . To show precompactness in L^2 , by [19][Theorem 2.88] we need to show that for all $\delta > 0$ there exists $h_0 > 0$ such that for all $h \in [0, h_0]$ and all n

$$(48) \quad \|f^n(\cdot + h, t) - f^n(\cdot, t)\|_{L^2} < \delta.$$

We proceed with the proof assuming (48), and verify the condition at the end.

Observe that $t \mapsto f^n(\cdot, t)$ is a constant speed curve in $(L^2(0, 1), d_{HV})$, since it is a geodesic. Therefore that family of functions f^n considered as functions between $[0, 1]$ and $(L^2(0, 1), d_{HV})$ are uniformly Lipschitz and thus equicontinuous. Since we know that $\{f^n(\cdot, t)\}_{n \in \mathbb{N}}$ is precompact in $(L^2(0, 1), \|\cdot\|_{L^2})$, it is precompact in L^2 with respect to the d_{HV} metric. Therefore by the Arzela-Ascoli Theorem, for metric-space valued functions, we get that f^n is precompact in $C(0, 1, (L^2(0, 1), d_{HV}))$. We note that this also implies that up to a subsequence, $f^n(\cdot, t)$ converges to $f(\cdot, t)$ in $L^2(0, 1)$ for all t fixed.

We now turn to proving that (48) holds. By (22), for each t , $\Phi_t := \Phi(\cdot, t)$ and Φ_t^{-1} are Lipschitz and the Lipschitz constant for both maps and for all $t \in [0, 1]$ is bounded by $b_n := \exp(\sqrt{2}\|v^n\|_{L^2(0, 1, H^2)})$.

Next we show that the function $\eta^n(x, t) := \frac{\int_0^t J^n(x, s) ds}{\int_0^1 J^n(x, \tau) d\tau}$, where $J^n(x, t) = e^{-\int_0^t v_x^n(\Phi^n(x, s), s) ds}$ is Hölder continuous in x . Since v_x^n is Hölder continuous, we have that for $x, y \in [0, 1]$,

$$|J^n(x, t) - J^n(y, t)| \leq e^{b_n} \int_0^t \|v_{xx}^n(\cdot, s)\|_{L^2} (\text{Lip}(\Phi_s^n))^{\frac{1}{2}} ds |x - y|^{\frac{1}{2}} \leq e^{b_n} b_n^{\frac{1}{2}} \|v^n\|_{L^2(0, 1, H^2)} |x - y|^{\frac{1}{2}},$$

since $|v_x(y) - v_x(y)| \leq |x - y|^{\frac{1}{2}} \|v_{xx}\|_{L^2}$. This implies that for any $0 \leq t \leq 1$, $\int_0^t J^n(x, s) ds$ is in $C^{0, \frac{1}{2}}(0, 1)$ with Hölder constant $e^{b_n} b_n^{\frac{1}{2}} \|v^n\|_{L^2(0,1,H^2)}$. Observe that by (42), $b_n^{-1} \leq J^n(x, t) \leq b_n$. Thus $(\int_0^1 J(x, \tau) d\tau)^{-1}$ is Hölder continuous with constant $e^{b_n} b_n^{\frac{5}{2}} \|v^n\|_{L^2(0,1,H^2)}$. Therefore for any x, y , for any t ,

$$|\eta^n(x, t) - \eta^n(y, t)| \leq (b_n e^{b_n} b_n^{\frac{1}{2}} + b_n e^{b_n} b_n^{\frac{5}{2}}) \|v^n\|_{L^2(0,1,H^2)} |x - y|^{\frac{1}{2}} \leq 2e^{b_n} b_n^{\frac{7}{2}} \|v^n\|_{L^2(0,1,H^2)} |x - y|^{\frac{1}{2}}.$$

By change of variables

$$f^n(y, t) = \eta^n((\Phi_t^n)^{-1}(y), t) f_0^n((\Phi_t^n)^{-1}(y)) + (1 - \eta^n((\Phi_t^n)^{-1}(y), t)) f_1^n(\Phi_1 \Phi_t^{-1}(y))$$

We note that $b := \sup\{b_n : n \in \mathbb{N}\}$ is finite since $\|v^n\|_{L^2(0,1,H^2)}^2 \leq c \|f_0 - f_1\|_{L^2}^2 + 1$ by the observation above, for some $c > 0$ depending on λ, κ , and ε .

We extend f^n, f_0 and f_1 by zero to $\mathbb{R}$. We also extend Φ_t^n as identity ($\Phi_t^n(x) = x$) outside of $[0, 1]$. Choose $g_0, g_1 \in C_c^\infty(\mathbb{R})$ such that $\|f_0 - g_0\|_{L^2(\mathbb{R})} < \frac{\delta}{100b^3}$ and $\|f_1 - g_1\|_{L^2(\mathbb{R})} < \frac{\delta}{100b^3}$. Choose $N(\delta)$ such that, then for any $n > N(\delta)$, $\|f_0^n - f_0\|_{L^2} < \frac{\delta}{100b^3}$, $\|f_1^n - f_1\|_{L^2} < \frac{\delta}{100b^3}$. Then for $n > N(\delta)$,

$$\begin{aligned} & \|f_0^n((\Phi_t^n)^{-1}(y+h)) - f_0^n((\Phi_t^n)^{-1}(y))\|_{L^2(\mathbb{R})} \\ & \leq \|f_0^n((\Phi_t^n)^{-1}(y+h)) - g_0((\Phi_t^n)^{-1}(y+h))\|_{L^2(\mathbb{R})} \\ & \quad + \|g_0((\Phi_t^n)^{-1}(y+h)) - g_0((\Phi_t^n)^{-1}(y))\|_{L^2(\mathbb{R})} \\ & \quad + \|g_0((\Phi_t^n)^{-1}(y)) - f_0^n((\Phi_t^n)^{-1}(y))\|_{L^2(\mathbb{R})} \\ & \leq 2\|D(\Phi_t^n)^{-1}\|_{L^\infty}^{\frac{1}{2}} \|f_0^n - g_0\|_{L^2} + \|g_0((\Phi_t^n)^{-1}(y+h)) - g_0((\Phi_t^n)^{-1}(y))\|_{L^2(\mathbb{R})} \\ & \leq \frac{\delta}{25b^2} + b \text{Lip}(g_0)h, \end{aligned}$$

Similarly,

$$\|f_1^n(\Phi_1(\Phi_t^n)^{-1}(y+h)) - f_1^n(\Phi_1(\Phi_t^n)^{-1}(y))\|_{L^2(\mathbb{R})} \leq \frac{\delta}{25b^2} + b^2 \text{Lip}(g_1)h.$$

Moreover

$$(49) \quad \|\eta^n((\Phi_t^n)^{-1}(y+h), t) - \eta^n((\Phi_t^n)^{-1}(y), t)\|_{L^\infty(0,1)} \leq 2e^{b_n} b_n^4 \|v^n\|_{L^2(0,1,H^2)} h^{\frac{1}{2}} \leq Ch^{\frac{1}{2}},$$

where $C := 2 \exp(b) b^4 (\sqrt{c} \|f_1 - f_0\|_{L^2} + 1)$.

Therefore,

$$\begin{aligned} & \|f^n(\cdot, t) - f^n(\cdot + h, t)\|_{L^2(\mathbb{R})} \\ & \leq \|\eta^n((\Phi_t^n)^{-1}(\cdot + h), t) f_0^n((\Phi_t^n)^{-1}(\cdot + h)) - \eta^n((\Phi_t^n)^{-1}(\cdot), t) f_0^n((\Phi_t^n)^{-1}(\cdot))\|_{L^2} \\ & \quad + \|(1 - \eta^n((\Phi_t^n)^{-1}(\cdot + h), t)) f_1^n(\Phi_1 \Phi_t^{-1}(\cdot + h)) - (1 - \eta^n((\Phi_t^n)^{-1}(\cdot), t)) f_1^n(\Phi_1 \Phi_t^{-1}(\cdot))\|_{L^2} \\ & \leq \|\eta^n\|_{L^\infty} \|f_0^n((\Phi_t^n)^{-1}(\cdot + h)) - f_0^n((\Phi_t^n)^{-1}(\cdot))\|_{L^2} \\ & \quad + \|\eta^n((\Phi_t^n)^{-1}(\cdot + h), t) - \eta^n((\Phi_t^n)^{-1}(\cdot), t)\|_{L^\infty} \|f_0^n \circ (\Phi_t^n)^{-1}\|_{L^2} \\ & \quad + \|1 - \eta^n\|_{L^\infty} \|f_1^n(\Phi_1(\Phi_t^n)^{-1}(\cdot + h)) - f_1^n(\Phi_1(\Phi_t^n)^{-1}(\cdot))\|_{L^2} \\ & \quad + \|\eta^n((\Phi_t^n)^{-1}(\cdot + h), t) - \eta^n((\Phi_t^n)^{-1}(\cdot), t)\|_{L^\infty} \|f_1^n \circ \Phi_1 \circ (\Phi_t^n)^{-1}\|_{L^2} \\ & \leq b^2 \left(\frac{\delta}{25b^2} + b \text{Lip}(g_0)h \right) + Ch^{\frac{1}{2}} b \|f_0^n\|_{L^2} + (1 + b^2) \left(\frac{\delta}{25b^2} + b^2 \text{Lip}(g_1)h \right) + Ch^{\frac{1}{2}} b \|f_1^n\|_{L^2} \\ & \leq \frac{3\delta}{25} + 2b^3 (\text{Lip}(g_0) + \text{Lip}(g_1))h + Cb(2 + \|f_0\|_{L^2} + \|f_1\|_{L^2})h^{\frac{1}{2}}. \end{aligned}$$

We pick h_0 , which depend on f_0, f_1, g_0, g_1 , such that for all t , for all $h \in [0, h_0]$, and $n > N(\delta)$, (48) holds. For $n \leq N(\delta)$, there exists h_n such that for all $h \in [0, h_n]$, (48) holds for f^n . Take $\tilde{h} := \min\{h_0, h_1, \dots, h_{N(\delta)}\}$. For all t , for all $h \in [0, \tilde{h}]$, and for all n , (48) holds. $\square$

4. NUMERICAL SCHEME

We propose an iterative minimization scheme to find the minimizers of the action (50) over the admissible paths (51). We first present two convex sub-problems on the continuous level by fixing v and f , respectively. By solving each of the sub-problems, the action functional decays monotonically. We then present our discretization scheme of finding the optimal path based on this problem splitting.

4.1. Two sub-problems. Recall that we are interested in minimizing

$$(50) \quad A_{\kappa, \lambda, \varepsilon}(f, v, z) = \frac{1}{2} \int_0^1 \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 dx dt,$$

over the set of admissible paths defined by

$$(51) \quad \mathcal{A} = \{(f, v, z) : f_t = -f_x v + z, v(0, \cdot) = v(1, \cdot) = 0, f(\cdot, 0) = f_0, f(\cdot, 1) = f_1\}.$$

The Euler–Lagrange equations for this variational problem are given in (37)–(41) under the assumption that $f_0, f_1 \in H^1((0, 1))$. The numerical methods presented in this section and the experiments in Section 5 are based on the assumptions that the signals are $C^2((0, 1))$. Due to the regularity results of Proposition 3.1 and Remark 3.2 as well as the stability result proved in Proposition 3.3, we expect that the geodesics for regular signals can be used to approximate those for general signals.

We point out that among (f, v, z) , the three variables we optimize over, z is determined by f and v due to the constraint set (51). However, it is still impractical to directly minimize (50) over (51) due to the nonlinear constraint. Next, we translate the optimization problem to a fixed-point problem by working with the system of Euler–Lagrange equations (37)–(41).

More specifically, we find a solution to the Euler–Lagrange equations through two convex optimization sub-problems. The method alternates between fixing v while finding the optimal (f, z) and fixing f while searching the optimal (v, z) , both for problem (50)–(51). This method shares similar flavors with many existing optimization algorithms. First, it is related to the so-called block coordinate descent method [38] since we alternately update (f, z) and (v, z) , the coordinate blocks in our problem. Based on [26, P. 266], our algorithm also has local convergence since the action function has a unique minimum in each coordinate block. The fact that our updated new (f, z) or (v, z) is the exact minimizer for each of the sub-problem, sharing similar features with ADMM [39]. The corresponding optimal (f, z) or (v, z) are weighted projections onto the linear constraints determined by (37) – (41) (with the other variable fixed).

Lastly, our method shares the same spirit with the so-called sequential quadratic programming (SQP) [7]. For quadratic programming with nonlinear constraints, SQP solves a sequence of optimization sub-problems using a linearization of the constraints. In our method, we achieve linearization of (51) by fixing v or f . Next, we will discuss in detail the two important sub-problems.

4.1.1. From v to (f, z) . First, for a given v , we consider the sub-problem of finding (f, z) which minimize the action (50) under the constraint that $(f, v, z) \in \mathcal{A}$ for the v fixed. We note that this reduces to minimizing a convex (in fact, quadratic) objective functional under a linear constraint:

$$(52) \quad \min_{f, z} \frac{1}{2} \int_0^1 \int_0^1 z^2 dx dt, \quad \text{s.t.} \quad (f, v, z) \in \mathcal{A}.$$

The first-order optimality conditions are given by equations (39), (40) and (41):

$$(53) \quad \begin{aligned} z_t + (zv)_x &= 0, \\ f_t + f_x v - z &= 0, \\ f(\cdot, 0) &= f_0, f(\cdot, 1) = f_1. \end{aligned}$$

To solve (53), we can use the Lagrangian approach. First, we can obtain a flow map $\Phi(x, t)$ solving (17), and an analytical solution for $z(\Phi(x, t), t)$ presented in (19). Note that we still need to find the quantity J in (19). To do so, we observe that, by Lemma 2.7, $J(x, t) = \exp(-\int_0^t v_x(\Phi(x, s), s)ds)$ is the weak solution of the following auxiliary initial value problem:

$$J_t + (Jv)_x = 0 \quad \text{on } [0, 1]^2, \quad \text{with } J(x, 0) \equiv 1.$$

Hence, by (19), we have an analytic formulation for $z(\Phi(x, t), t)$:

$$(54) \quad z(\Phi(x, t), t) = (f_1(\Phi(x, 1)) - f_0(x)) \frac{J(x, t)}{\int_0^1 J(x, s)ds}.$$

We also have an analytic formulation for $f(\Phi(x, t), t)$ given by (20),

$$(55) \quad f(\Phi(x, t), t) = (1 - \eta(x, t)) f_0(x) + \eta(x, t) f_1(\Phi(x, 1)), \quad \text{where } \eta(x, t) = \frac{\int_0^t J(x, s)ds}{\int_0^1 J(x, s)ds}.$$

As a final step, through a change of coordinate, we obtain functions f and z at the (x, t) coordinate (which are used for v), rather than the flow map coordinate $(\Phi(x, t), t)$. At the discrete level, this will amount to an interpolation step between the points where f and z are computed along the flow, and the desired grid points.

Combining all steps above, we have defined a continuous operator

$$\mathcal{G}_1 : v \mapsto (z, f),$$

which solves (53) and (52). Note that we have $(f, v, z) \in \mathcal{A}$ automatically if $(f, z) = \mathcal{G}_1(v)$. If we denote the (f, v, z) from the previous step by $(f^{\text{old}}, v^{\text{old}}, z^{\text{old}})$, where $(f^{\text{old}}, v^{\text{old}}, z^{\text{old}}) \in \mathcal{A}$, we then have

$$A_{\kappa, \lambda, \varepsilon}(f, v^{\text{old}}, z) \leq A_{\kappa, \lambda, \varepsilon}(f^{\text{old}}, v^{\text{old}}, z^{\text{old}}).$$

The steps above yield functions (f, v^{old}, z) satisfying (38), (39), (40), and (41) for the given v^{old} , but they do not necessarily satisfy (37) in the Euler–Lagrange equations. Otherwise, we have found a set of solutions satisfying the first-order optimality conditions of (50).

4.1.2. From f to (v, z) . Given f from the previous step, we consider the second sub-problem of finding the pair $(v^{\text{new}}, z^{\text{new}})$ that minimizes the action (50) under the constraint that $(f, v, z) \in \mathcal{A}$. This again is a quadratic optimization problem under linear constraint:

$$(56) \quad \min_{v, z} \frac{1}{2} \int_0^1 \int_0^1 \kappa v^2 + \lambda v_x^2 + \varepsilon v_{xx}^2 + z^2 dx dt, \quad \text{s.t. } (f, v, z) \in \mathcal{A}.$$

Using the Euler–Lagrange equation (37) and the constraint $z = f_t + v f_x$ yields the following fourth order boundary value problem for v :

$$(57) \quad \begin{aligned} \varepsilon v_{xxxx} - \lambda v_{xx} + (\kappa + |f_x|^2)v &= -f_t f_x \quad \text{on } (0, 1)^2, \\ v &= 0 \quad \text{on } \{0, 1\} \times (0, 1), \\ v_{xx} &= 0 \quad \text{on } \{0, 1\} \times (0, 1). \end{aligned}$$

Given f and the solution v , we also obtain z . Let us denote the solution operator by $\mathcal{G}_2 : f \mapsto (v^{\text{new}}, z^{\text{new}})$.

We note that

$$(58) \quad A_{\kappa, \lambda, \varepsilon}(f, v^{\text{new}}, z^{\text{new}}) \leq A_{\kappa, \lambda, \varepsilon}(f, v^{\text{old}}, z).$$

Therefore, combining $\mathcal{G}_2$ with the previous step in Section 4.1.1, we have the following inequalities with respect to the action functional:

$$A_{\kappa,\lambda,\varepsilon}(f^{\text{new}}, v^{\text{new}}, z^{\text{new}}) \underbrace{\leq}_{(v^{\text{new}}, z^{\text{new}}) = \mathcal{G}_2(f), f^{\text{new}} = f} A_{\kappa,\lambda,\varepsilon}(f, v^{\text{old}}, z) \underbrace{\leq}_{(f, z) = \mathcal{G}_1(v^{\text{old}})} A_{\kappa,\lambda,\varepsilon}(f^{\text{old}}, v^{\text{old}}, z^{\text{old}}),$$

where $(f^{\text{new}}, v^{\text{new}}, z^{\text{new}}), (f, v^{\text{old}}, z), (f^{\text{old}}, v^{\text{old}}, z^{\text{old}}) \in \mathcal{A}$, all satisfying the constraints. If we define a new operator $\mathcal{G}$ by composing $\mathcal{G}_2$ with $\mathcal{G}_1$, i.e.,

$$(59) \quad \mathcal{G} = \mathcal{G}_2 \circ \mathcal{G}_1 : \mathcal{A} \mapsto \mathcal{A},$$

it gives an update formula after which the action functional decays:

$$(60) \quad A_{\kappa,\lambda,\varepsilon}(\mathcal{G}(f, v, z)) \leq A_{\kappa,\lambda,\varepsilon}(f, v, z).$$

We can then repetitively applying $\mathcal{G}$ until finding a set of solution (f^*, v^*, z^*) where $\mathcal{G}(f^*, v^*, z^*) = (f^*, v^*, z^*)$. That is, (f^*, v^*, z^*) is a fixed point of $\mathcal{G}$, while (60) indicates the contractivity of the fixed-point operator with respect to the action functional. It is easy to verify that (f^*, v^*, z^*) also solves the Euler–Lagrange equations (37)–(41).

Algorithm 1 An iterative scheme for minimizing (50).

- 1: Given an initial guess $(f^{(0)}, v^{(0)}, z^{(0)}) \in \mathcal{A}$, maximum number of iterations N , tolerance $\delta > 0$.
 - 2: **for** $n = 1$ to N **do**
 - 3: Compute $(\tilde{f}, \tilde{z}) = \mathcal{G}_1(v^{(n)})$ with $\mathcal{G}_1$ described in Section 4.1.1 and set $f^{(n+1)} = \tilde{f}$.
 - 4: Set $(v^{(n+1)}, z^{(n+1)}) = \mathcal{G}_2(f^{(n+1)})$ with $\mathcal{G}_2$ described in Section 4.1.2.
 - 5: **if** $|A_{\kappa,\lambda,\varepsilon}(f^{(n+1)}, v^{(n+1)}, z^{(n+1)}) - A_{\kappa,\lambda,\varepsilon}(f^{(n)}, v^{(n)}, z^{(n)})| < \delta$ **then**
 - 6: Return $(f^{(n+1)}, z^{(n+1)}, v^{(n+1)})$ and the minimum action value; **Break**.
 - 7: **end if**
 - 8: **end for**
-

Algorithm 2 An iterative scheme for minimizing (50) with damping.

- 1: Given an initial guess $(f^{(0)}, v^{(0)}, z^{(0)}) \in \mathcal{A}$, maximum number of iterations N , tolerance $\delta > 0$.
 - 2: **for** $n = 1$ to N **do**
 - 3: Compute $(\tilde{f}, \tilde{z}) = \mathcal{G}_1(v^{(n)})$ and set $f^{(n+1)} = \alpha_1 \tilde{f} + (1 - \alpha_1) f^{(n)}$, with α_1 given by Algorithm 3
 - 4: and $\mathcal{G}_1$ described in Section 4.1.1.
 - 5: Compute $(\hat{v}, \hat{z}) = \mathcal{G}_2(f^{(n+1)})$ and set $(v^{(n+1)}, z^{(n+1)}) = \alpha_2(\hat{v}, \hat{z}) + (1 - \alpha_2)(v^{(n)}, z^{(n+\frac{1}{2})})$, with
 - 6: α_2 given by Algorithm 3 and $\mathcal{G}_2$ described in Section 4.1.2.
 - 7: **if** $A_{\kappa,\lambda,\varepsilon}(f^{(n+1)}, v^{(n+1)}, z^{(n+1)}) > A_{\kappa,\lambda,\varepsilon}(f^{(n)}, v^{(n)}, z^{(n)}) - \delta$ **then**
 - 8: Return $(f^{(n+1)}, z^{(n+1)}, v^{(n+1)})$ and the minimum action value; **Break**.
 - 9: **end if**
 - 10: **end for**
-

Algorithm 3 Back-tracking line search for the damping parameter.

```

1: Given the old iterate  $(\bar{f}, \bar{v}, \bar{z})$ , the proposed new iterate  $(\tilde{f}, \tilde{v}, \tilde{z})$ , the objective function  $A_{\kappa, \lambda, \varepsilon}$ , and
   the maximum number of search steps  $\tilde{N}$ . Set  $\alpha = 1$ ,  $\bar{A} = A_{\kappa, \lambda, \varepsilon}(\bar{f}, \bar{v}, \bar{z})$  and FLAG = 0.
2: for  $i = 1$  to  $\tilde{N}$  do
3:   if  $A_{\kappa, \lambda, \varepsilon}(f_\alpha, v_\alpha, z_\alpha) < \bar{A}$  where  $(f_\alpha, v_\alpha, z_\alpha) = (1 - \alpha)(\bar{f}, \bar{v}, \bar{z}) + \alpha(\tilde{f}, \tilde{v}, \tilde{z})$  then
4:     Return  $\alpha$  and set FLAG = 1; Break.
5:   end if
6:    $\alpha \leftarrow \alpha/2$ .
7: end for
8: if FLAG = 0 then
9:   Return  $\alpha = 0$ .
10: end if

```

Algorithm 4 Searching for a path minimizing action (50) via different initializations.

```

1: Given the maximum number of iterations  $N \in \mathbb{N}^+$ ,  $\delta > 0$ ,  $k_{max} \in \mathbb{N}$ .
2: for  $k = 0$  to  $k_{max}$  do
3:   Use the prominence-based matching initialization described in Section 4.3 with parameter  $k$  and
4:   obtain  $(f^{(0)}, v^{(0)}, z^{(0)})$ .
5:   Run Algorithm 2 with  $N$  iterations and tolerance  $\delta$ . Obtain the minimum action value  $\mathcal{J}(k)$ .
6: end for
7: Find  $k^* = \operatorname{argmin} \mathcal{J}(k)$ .
8: Return the optimal path  $(f, v, z)$  and its action value  $\mathcal{J}(k^*)$  for initialization with parameter  $k^*$ .

```

4.2. The discrete scheme. Next, we use a simple first-order numerical scheme to solve the two sub-problems discussed above. If the signals f_0 and f_1 are known to be smooth, higher-order discretization schemes would be recommended for better efficiency and accuracy. If the signals are discontinuous, first-order schemes, on the other hand, are known to mitigate the Gibbs phenomenon [23] that higher-order methods may suffer.

To evaluate $(f, z) = \mathcal{G}_1(v)$, we compute $z(\Phi(x, t), t)$ and $f(\Phi(x, t), t)$ using first-order numerical integration based on (54) and (55). We obtain $z(x, t)$ and $f(x, t)$ from $z(\Phi(x, t), t)$ and $f(\Phi(x, t), t)$ through first-order numerical interpolation in the x variable alone. This step can be implemented in a parallel fashion.

To evaluate $(v, z) = \mathcal{G}_2(f)$, we have a fourth-order PDE for $v(\cdot, t)$ for every fixed t , with $v = 0$ and $v_{xx} = 0$ (if $\varepsilon \neq 0$) on the boundaries $\{0, 1\}$.

Consider a uniform mesh over the spatial domain $[0, 1]$ and the time domain $[0, 1]$. The spatial spacing $\Delta x = 1/N_x$ and the time-domain spacing $\Delta t = 1/N_t$. Let $\mathbf{v}_j \in \mathbb{R}^{N_x+1}$ be a vector approximating $[v(0, j\Delta t), \dots, v(i\Delta x, j\Delta t), \dots, v(1, j\Delta t)]^\top$. The PDE (57) then becomes the following linear system under the finite-difference discretization,

$$(A + \operatorname{diag}(\mathbf{w}_j \odot \mathbf{w}_j)) \mathbf{v}_j = -\boldsymbol{\tau}_j \odot \mathbf{w}_j,$$

where $\odot$ denotes the Hadamard product, $\text{diag}(\mathbf{x})$ denotes a diagonal matrix with elements of vector $\mathbf{x}$ being its diagonal entries, $A \in \mathbb{R}^{(N_x+1) \times (N_x+1)}$ is given by

$$\begin{bmatrix} 1 & & & & & & & & & \\ -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{5\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & & & & & & & \\ \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{6\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & & & & & & \\ & \ddots & & \ddots & & & & & & \\ & & \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{6\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & & & & \\ & & & \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{5\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & & & \\ & & & & \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{5\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & & \\ & & & & & \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{5\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \\ & & & & & & \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{5\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} \\ & & & & & & & \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} & \kappa + \frac{2\lambda}{\Delta x^2} + \frac{5\varepsilon}{\Delta x^4} \\ & & & & & & & & \frac{\varepsilon}{\Delta x^4} & -\frac{4\varepsilon}{\Delta x^4} - \frac{\lambda}{\Delta x^2} \\ & & & & & & & & & 1 \end{bmatrix},$$

and $\mathbf{w}_j, \boldsymbol{\tau}_j \in \mathbb{R}^{N_x+1}$, approximating $f_x(\cdot, t_j)$ and $f_t(\cdot, t_j)$, respectively, are given by

$$\begin{aligned} \mathbf{w}_j &= \frac{1}{\Delta x} \begin{bmatrix} 0 & (f(x_3, t_j) - f(x_2, t_j)) & \dots & (f(x_{N_x+1}, t_j) - f(x_{N_x}, t_j)) & 0 \end{bmatrix}^\top, \\ \boldsymbol{\tau}_j &= \frac{1}{\Delta t} \begin{bmatrix} 0 & (f(x_2, t_{j+1}) - f(x_2, t_j)) & \dots & (f(x_{N_x}, t_{j+1}) - f(x_{N_x}, t_j)) & 0 \end{bmatrix}^\top, \end{aligned}$$

with $f(x_i, t_j) = f(i\Delta x, j\Delta t)$. We remark that the first and last elements of the right-hand side are set to be zero while the first and last rows of the left-hand side are also modified. These two linear equations are to enforce $v = 0$ on the boundary; see (57). The fact that $v_{xx} = 0$ on the boundary implies that the ghost points $v(-\Delta x, \cdot) = -v(\Delta x, \cdot)$, and $v(1 + \Delta x, \cdot) = -v(1 - \Delta x, \cdot)$, which is used in the second and the N_x -th rows of A . The remaining $N_x - 3$ equations in the linear system are to enforce the linear PDE (57). Note that with a given f , we can solve for $\{\mathbf{v}_j\}$ in parallel for all time $\{t_j\}$, or construct a large sparse linear system with respect to $\mathbf{v} = [\mathbf{v}_1^\top \dots \mathbf{v}_{N_t+1}^\top]^\top$ in one single sparse linear solve. When f_0, f_1 are smooth, it is preferable to use the central difference method to obtain $\boldsymbol{\tau}_j$.

Now we have two steps: $\mathcal{G}_1 : v \mapsto (f, z)$ through first-order numerical integration and interpolation, and $\mathcal{G}_2 : f \mapsto (v, z)$ through a linear PDE solver. On the continuous level, we have a monotonic energy decay based on (60). Algorithm 1 combines these two steps to iteratively find approximate local minimizers of (50).

While this algorithm works well for many signals, on the discrete level, due to numerical errors from interpolation, integration, and the PDE solver, (60) may not hold for every iteration of the fixed-point update. To ensure that the *discretized* variational problem still has a monotonically decaying action functional, we can introduce damping parameter α_1 and α_2 , similar to the step size in a gradient descent algorithm; see Algorithm 2 for details. We remark that most of the time, $\alpha_1 = \alpha_2 = 1$ is sufficient to ensure decay of the action value, which is the case on the continuous level. Only when it is close to the target solution, and the numerical errors dominate, one may observe that the action value no longer monotonically decreases when iterating between the two sub-problems. There damping helps obtain more accurate minima. We use back-tracking line search with a shrinking factor $1/2$ to find a proper damping coefficient; see details in Algorithm 3.

4.3. Finding a good initialization. We remark that the fixed points of $\mathcal{G}$ solve the Euler–Lagrange equations (37)–(41), but they may not minimize (50) over (51) due to the possible local minima. In other words, the fixed-point operator $\mathcal{G}$ has multiple fixed points. One can start with different initial guesses and investigate the convergence behavior while choosing the solution with the smallest objective function.

We propose two different types of initial guesses.

- (1) **Zero-velocity initialization.** We set $v^{(0)}(x, t) \equiv 0$, and compute $(f^{(0)}, z^{(0)}) = \mathcal{G}_2(v^{(0)})$. Note that in this case, we have $f^{(0)}(x, t) = (1 - t)f_0(x) + tf_1(x)$ and $z^{(0)}(x, t) = f_1(x) - f_0(x)$.
- (2) **Prominence-matching initialization.** We expect that if there exists a path with action that is substantially smaller than linear interpolation, that path will match nearby large peaks. We observe that matching the tallest peaks is not stable under perturbations as there may be spurious nearby peaks, for example, due to oscillations as in Figure 3. For this reason, we use a more stable notion of how large the peaks are. Namely, we use the notion of *prominence* coming from topography, which describes how large the peaks are compared to their surroundings. The prominence of a signal is defined as the least drop in height necessary to get to another local maximum with a higher value. Consider a positive integer k . For the given f_0 and f_1 , we each select k local maxima with the largest k prominence. The location of the local maxima are denoted by $\{x_i\}$ and $\{y_i\}$, $1 \leq i \leq k$, respectively. We then construct a map $T(x)$ such that $T(x_i) = y_i$ for each i , and $T(0) = 0$, $T(1) = 1$. We use linear interpolation to define its function value for $x \in (0, 1) \setminus \{x_i\}$. We then set the initial velocity to be $v^{(0)}(x, t) = T(x) - x$, which is constant in time, and use $(f^{(0)}, z^{(0)}) = \mathcal{G}_1(v^{(0)})$ as the initial guess for f and z . Also, if the minima in f_0 and f_1 are more significant than the maxima, one can also initialize by matching the prominence of $-f_0$ and $-f_1$.

The zero-velocity initialization is equivalent to a degenerate case of the prominence-matching initialization where $k = 0$. In this scenario, we have $T(x) = x$ through linear interpolation between $T(0) = 0$ and $T(1) = 1$. Later, we will refer to the “zero-velocity initialization” as $k = 0$. We also comment that the initialization is different for various integer k , which may lead to different convergence behavior and local minima of the action functional (50). We suggest trying for a few k values. While there are many variants on how one can incorporate the prominence-matching initialization into the optimization scheme, we outline an Algorithm 4 as an example, which we use to produce the numerical results in Section 5. To be more efficient, one can use fewer iterations when searching for an initialization compared to running the entire Algorithm 1 or Algorithm 2 to find the minimizer of the optimization problem.

5. NUMERICAL EXPERIMENTS

In this section, we present a few examples illustrating the geodesic using the HV geometry¹. Throughout this section, we plot the source signal f_0 in blue and the target signal f_1 in red, while their barycenter under the HV geometry is shown using the color purple. We use dashed lines to indicate the signals at $t = \frac{1}{4}$ and $t = \frac{3}{4}$.

Example 5.1. Nonuniqueness of minimizing geodesics. We consider an example where the source and target signals have two bumps. We use 300 spatial and 290 time intervals to discretize the space-time domain. For all plots in Figure 2, we set the hyperparameters $\kappa = 0.02$, $\lambda = 0.001$, and $\varepsilon = 0.002$. In Figure 2a, each signal has a large bump and a small bump. The large bumps are much bigger. The geodesic is dominated by horizontal transport. In Figure 2b, the two bumps in each signal are of comparable size while their locations remain the same. The geodesics, in this case, is dominated by vertical changes instead of horizontal transport in the previous example. Finally, by adjusting the ratio of the bump heights while fixing their locations, we find a scenario where horizontal transport and vertical changes result in the same action value; see Figure 2c. That is, we found local minimizers, which we believe to be global, where the paths have the same action. Thus, we believe the geodesics are not unique, indicating that the signal space is, at least partly, positively curved.

¹The codes based on the numerical scheme described in Section 4 that reproduce these examples can be found at <https://github.com/yunany/Compute-HV-distance-between-signals.git>.

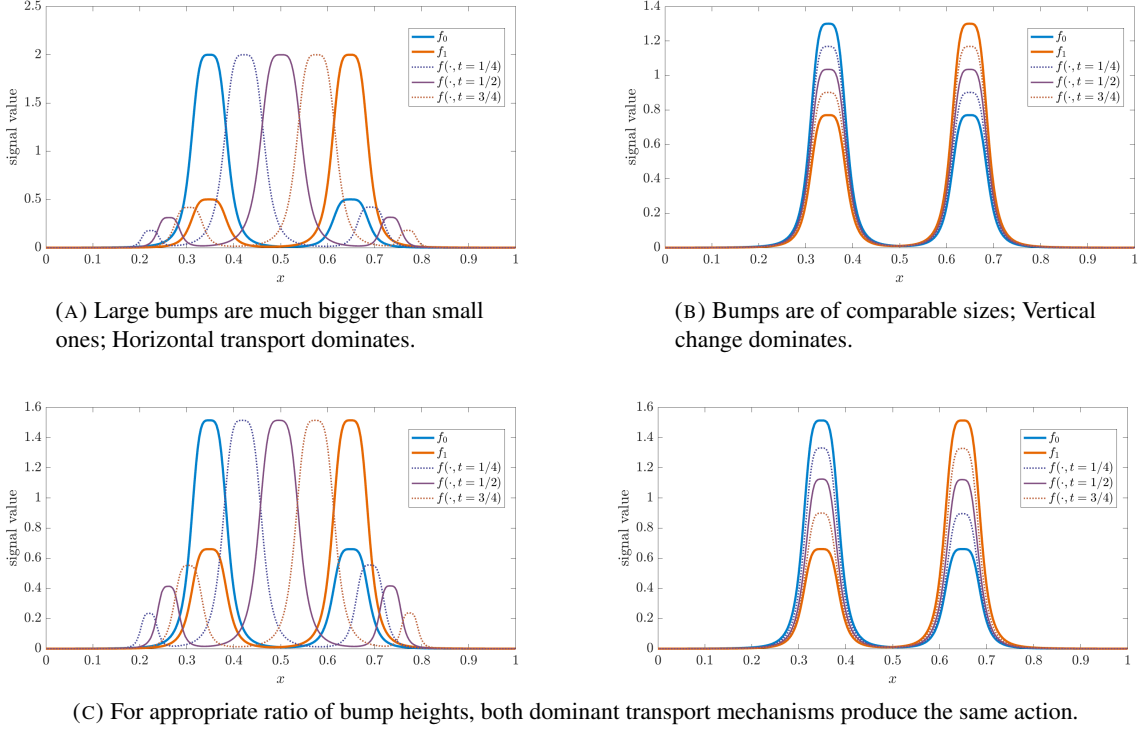
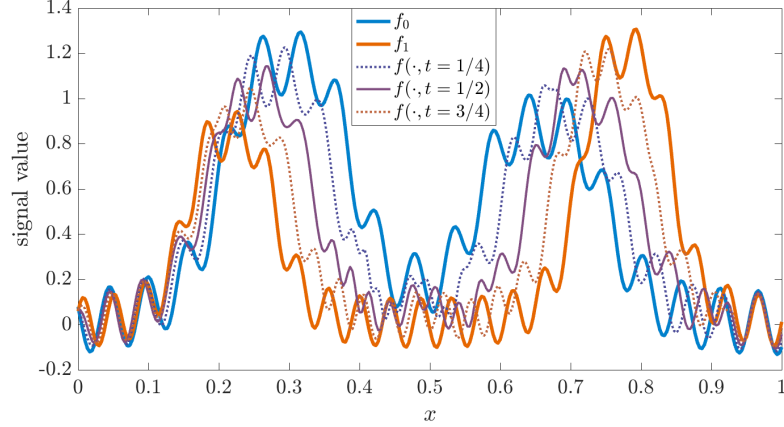


FIGURE 2. Example of the nonuniqueness of the length-minimizing geodesic. We set $\kappa = 0.02$, $\lambda = 0.001$, and $\varepsilon = 0.002$. We use 300 spatial and 290 time intervals.

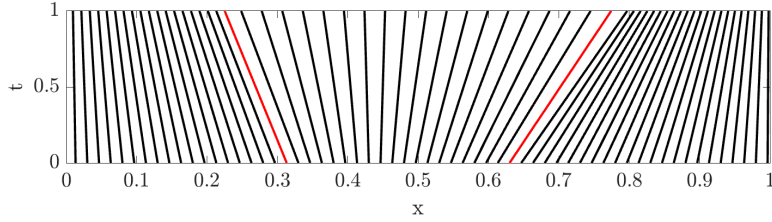
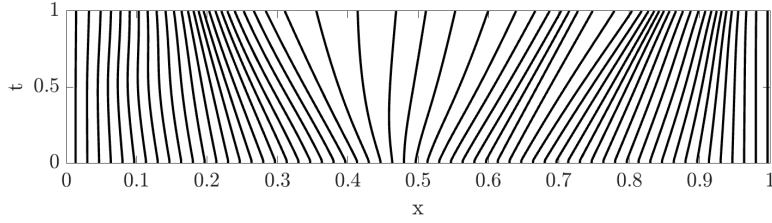
Example 5.2. Bumps with high frequency perturbations. Here we use the prominence-matching initialization introduced in the previous section to find a good starting point for our iterative scheme to converge to the global minimizer of the action function (50). In Figure 3a, we show the geodesics for hyperparameters $\kappa = 10^{-3}$, $\lambda = 5 \times 10^{-5}$, and $\varepsilon = 2.5 \times 10^{-5}$. We used 300 space intervals and 290 time intervals. In Figure 3b, we show the flow map based on the prominence-matching initial velocity, which maps the source signal's two most prominent peaks to the target signal's two most prominent peaks. The trajectories matching the peaks are indicated in red, while the remaining trajectories are obtained by linear interpolation. The flow map with respect to the final converged velocity is shown in Figure 3c, which is relatively close to Figure 3b.

Example 5.3. Signal with discontinuities. In Figure 4, we consider a non-smooth source signal and a smooth target signal. Here, we use the hyperparameters $\kappa = 0.02$, $\lambda = 0.001$, and $\varepsilon = 0.002$. We used 300 space intervals and 290 time intervals. The finite difference discretization on a static mesh implicitly regularizes the signal f_0 . Since we use a first-order numerical scheme, the computed solutions do not suffer from the Gibbs phenomenon as reflected by the geodesic in Figure 4. Moreover, we also observe the gradual transition between the discontinuous feature and the discontinuous feature both in horizontal and vertical directions.

Example 5.4. Growth and expansion. In Figure 5, we consider a “growth” example where the target signal is much bigger in width and height than the source signal. Here, we use the hyperparameters $\kappa = 0.2$, $\lambda = 0.01$, and $\varepsilon = 0.02$. We used 300 space intervals and 290 time intervals. We comment that this example is somewhat sensitive to the choice of hyperparameters, which directly affects the location of the barycenter between f_0 and f_1 under the HV geometry.



(A) geodesics based on the HV geometry

(B) flow map for the initial velocity (red trajectories denote the initial matching of the most prominent two peaks between f_0 and f_1)

(C) flow map for the optimal velocity

FIGURE 3. We use prominence-matching initialization to find the action-minimizing path. We show in (A) the initial signal (blue), final signal (red), and signals at $t = 0.25$, $t = 0.5$, and $t = 0.75$ along the computed geodesic for hyperparameters $\kappa = 10^{-3}$, $\lambda = 5 \times 10^{-5}$, and $\varepsilon = 2.5 \times 10^{-5}$. We used 300 space intervals and 290 time intervals. The flow for the initial velocity is shown in (B), and the flow for the optimal velocity is shown in (C).

Example 5.5. In Figure 6, we compare two signed signals with a single bump where the bumps' widths and locations do not agree. We compute geodesic between them under the HV geometry for the chosen hyperparameters decided by the parameter estimate (61) discussed in Section 5.1. Roughly estimate that $L = 0.3$, $W = 0.4$ and $H = 0.4$, which yields $\kappa = 0.1$, $\lambda = 0.01$ and $\varepsilon = 0.005$. We note that a wide set of parameters would have produced similar geodesics. The path is discretized using 300 spatial and 290 time intervals. The flow map corresponding to the optimal velocity v is plotted at the bottom of Figure 6, where one can observe the transport feature mapping the peak of the source signal (blue) to the peak of the target signal (red).

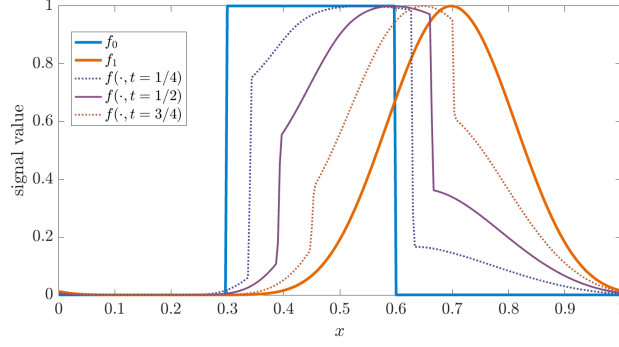


FIGURE 4. Algorithm allows for non-smooth data. Here, we use the hyperparameters $\kappa = 0.02$, $\lambda = 0.001$, $\varepsilon = 0.002$, 300 space intervals, and 290 time intervals.

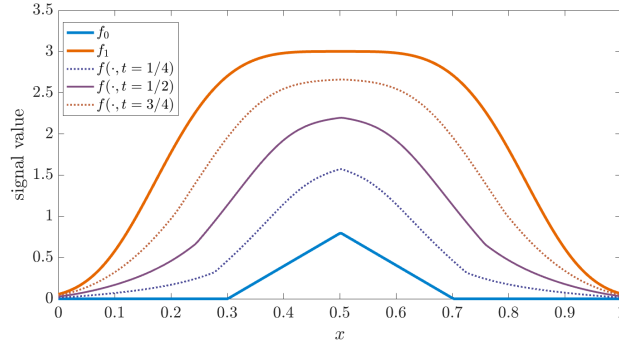


FIGURE 5. The d_{HV} geodesic balances horizontal expansion and vertical growth. We use $\kappa = 0.2$, $\lambda = 0.01$, $\varepsilon = 0.02$, 300 space intervals, and 290 time intervals.

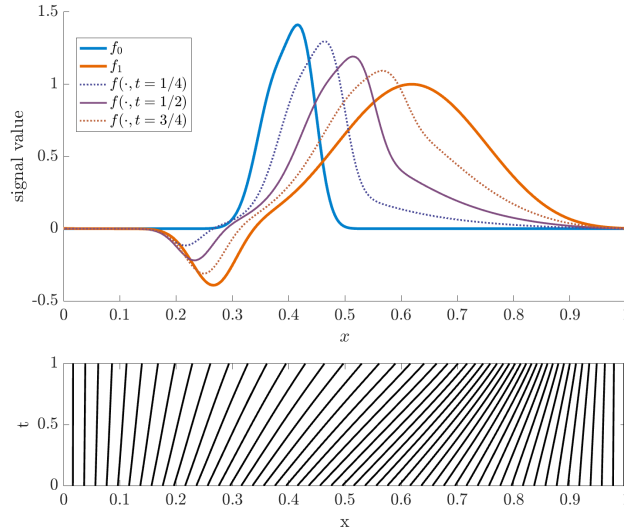


FIGURE 6. We show the initial signals along the computed geodesic for hyperparameters given by formula (61) with $L = 0.3$, $W = 0.4$ and $H = 0.4$. The bottom image shows the flow of the optimal velocity v .

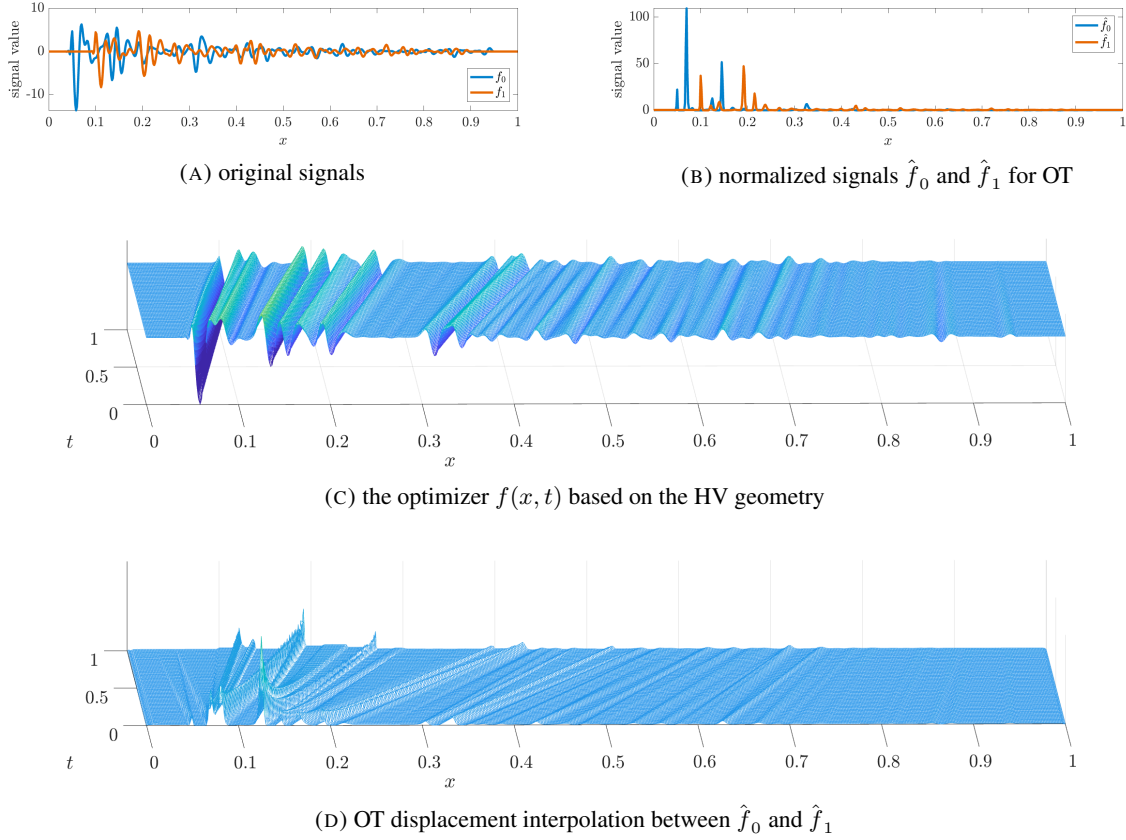


FIGURE 7. We compare the matching of two 1D seismic signals by the flow map induced by the HV geometry and optimal transportation with a quadratic cost (applied to normalized signals). (A): the original signal used in the HV geometry; (B): the normalized signal to satisfy the OT requirements; (C) optimizer $f(x, t)$ from the HV geometry; (D) the displacement interpolation between $\hat{f}_0$ and $\hat{f}_1$ based on the optimal transport map.

Example 5.6. (Seismic signals) Using optimal transportation (OT) for seismic applications has faced difficulties from the constraints that the signals should be nonnegative with equal total mass [16, 17]. In this example, we test the proposed HV geometry for comparing synthetic seismic signals shown in Figure 7a. To compare them using OT, we may normalize the signals first so that they are nonnegative with equal total mass; see Figure 7b for the normalized signals $\hat{f}_0, \hat{f}_1$, squared and then scaled to integrate to one [16]. Figures 7c and 7d show the HV and OT geometry velocity flow maps, respectively. We use a quadratic cost function for OT. For the HV geometry, we set $H = 2$, $L = 0.1$ and $W = 0.02$ for parameters in (61) presented in Section 5.1. Note that in the classic OT, all mass has to be transported through the monotonic map T such that $\int_0^x \hat{f}_0(y) dy = \int_0^{T(x)} \hat{f}_1(y) dy$. This may lead to mass being transported far away and unevenly, as illustrated in Figure 7d. The proposed HV geometry not only can handle signed signals naturally, avoiding the artifacts by preprocessing the signal but also enforces regularity to the velocity.

Example 5.7. Finally, we consider a real-world example. We compare two heartbeats from the ECG database PhysioNet 2017 Challenge [14, 21]; see Figure 8. The geodesic is computed in the space of signals according to the HV geometry for hyperparameters given by (61) below in Section 5.1, with $L = W = 0.1$ and $H = 300$, which one estimates from the given data. We use 600 space intervals and

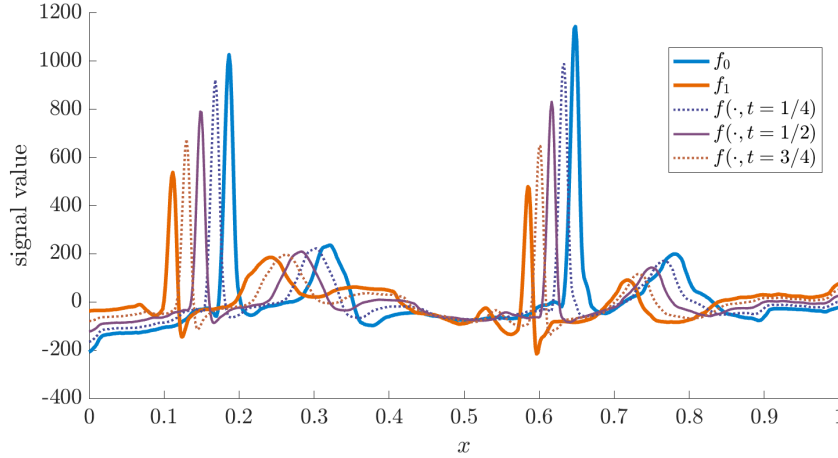


FIGURE 8. Shown is the geodesic between two ECG signals under the HV geometry. We note that the horizontal transform dominates for some features, while some parts are matched by vertically moving the graph. We set κ , λ , and ε based on (61) with $L = W = 0.1$ and $H = 300$. We use 600 space intervals and 150 time intervals.

150 time intervals. We note that large features (R-peaks and T-waves) are matched in a desirable way via horizontal transport, while perturbations of small amplitude are matched via a vertical adjustment. This illustrates the benefits of the HV geometry.

5.1. Parameter selection. An important element in using HV geometry to analyze signals is how to select the parameters κ , λ , and ε . This depends on the length scales present in the data. Here we give a simple rule for selecting the parameters based on the scaling properties of the distance; see Proposition 2.1. Let H be the average vertical variation in the data, W be the typical width of features in the data, and L be the maximum horizontal distance between the features to be matched. Then we suggest using

$$(61) \quad \kappa = 0.01 \frac{H^2}{L^2}, \quad \lambda = 0.02H^2, \quad \text{and} \quad \varepsilon = 0.2H^2W^2.$$

As we mentioned, the scaling of the parameters respects the invariances of the distance. The real number coefficients (0.01, 0.02, 0.2) are based on numerical experiments with different signal types.

We note that given a data set $\mathcal{F} = \{f_1, \dots, f_n\}$ a good suggestion for H would be the typical L^2 distance between the signals:

$$H^2 = \frac{1}{n^2} \sum_i \sum_j \|f_i - f_j\|_{L^2}^2.$$

We also note that H^2 is twice the variance and can thus be computed as a sum over one index:

$$H^2 = \frac{2}{n} \sum_i \|f_i\|_{L^2}^2 - 2 \left\| \frac{1}{n} \sum_i f_i \right\|_{L^2}^2.$$

We remark that, for most signals, the outcome is not very sensitive to the parameters.

ACKNOWLEDGEMENTS

We thank Jianming Wang for providing a key idea for the result in Section 2.1. We are grateful to Katy Craig for stimulating discussions. DS and RH are grateful to NSF for support via grant DMS 2206069. This work was done in part while DS and YY were visiting the Simons Institute for the Theory of

Computing in Fall 2021. YY acknowledges support from Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation.

REFERENCES

- [1] L. AMBROSIO AND G. CRIPPA, *Continuity equations and ode flows with non-smooth velocity*, Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 144 (2014), pp. 1191–1244.
- [2] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.
- [3] M. BAUER, M. BRUVERIS, P. HARMS, AND P. W. MICHOR, *Smooth perturbations of the functional calculus and applications to Riemannian geometry on spaces of metrics*, Comm. Math. Phys., 389 (2022), pp. 899–931.
- [4] M. BAUER, P. HARMS, AND P. W. MICHOR, *Fractional Sobolev metrics on spaces of immersions*, Calc. Var. Partial Differential Equations, 59 (2020), pp. Paper No. 62, 27.
- [5] A. BEHZADAN AND M. HOLST, *Multiplication in sobolev spaces, revisited*, 2015.
- [6] B. BERKELS, A. EFFLAND, AND M. RUMPF, *Time discrete geodesic paths in the space of images*, SIAM J. Imaging Sci., 8 (2015), pp. 1457–1488.
- [7] P. T. BOGGS AND J. W. TOLLE, *Sequential quadratic programming*, Acta numerica, 4 (1995), pp. 1–51.
- [8] M. BRUVERIS, *Completeness properties of Sobolev metrics on the space of curves*, J. Geom. Mech., 7 (2015), pp. 125–150.
- [9] M. BRUVERIS, P. W. MICHOR, AND D. MUMFORD, *Geodesic completeness for Sobolev metrics on the space of immersed plane curves*, Forum Math. Sigma, 2 (2014), pp. Paper No. e19, 38.
- [10] B. CHARLIER, N. CHARON, AND A. TROUVÉ, *The Fshape framework for the variability analysis of functional shapes*, Found. Comput. Math., 17 (2017), pp. 287–357.
- [11] N. CHARON, B. CHARLIER, AND A. TROUVÉ, *Metamorphoses of functional shapes in Sobolev spaces*, Found. Comput. Math., 18 (2018), pp. 1535–1596.
- [12] X. CHEN, A. JÜNGEL, AND J.-G. LIU, *A note on Aubin-Lions-Dubinskiĭ lemmas*, Acta Appl. Math., 133 (2014), pp. 33–43.
- [13] L. CHIZAT, G. PEYRÉ, B. SCHMITZER, AND F.-X. VIALARD, *Unbalanced optimal transport: Dynamic and kantorovich formulation*, 2015.
- [14] G. D. CLIFFORD, C. LIU, B. MOODY, L.-W. H. LEHMAN, I. SILVA, Q. LI, A. E. JOHNSON, AND R. G. MARK, *Af classification from a short single lead ecg recording: The physionet/computing in cardiology challenge 2017*, in 2017 Computing in Cardiology (CinC), 2017, pp. 1–4.
- [15] B. ENGQUIST, B. D. FROESE, AND Y. YANG, *Optimal transport for seismic full waveform inversion*, Commun. Math. Sci., 14 (2016), pp. 2309–2330.
- [16] B. ENGQUIST AND Y. YANG, *Seismic inversion and the data normalization for optimal transport*, Methods and Applications of Analysis, 26 (2019), pp. 133–148.
- [17] ———, *Optimal transport based seismic inversion: Beyond cycle skipping*, Communications on Pure and Applied Mathematics, 75 (2022), pp. 2201–2244.
- [18] L. C. EVANS, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 1998.
- [19] I. FONSECA AND G. LEONI, *Modern methods in the calculus of variations: L^p spaces*, Springer Monographs in Mathematics, Springer, New York, 2007.
- [20] A. FRANÇOIS, P. GORI, AND J. GLAUNÈS, *Metamorphic image registration using a semi-Lagrangian scheme*, in Geometric science of information, vol. 12829 of Lecture Notes in Comput. Sci., Springer, Cham, 2021, pp. 781–788.
- [21] A. L. GOLDBERGER, L. A. N. AMARAL, L. GLASS, J. M. HAUSDORFF, P. C. IVANOV, R. G. MARK, J. E. MIETUS, G. B. MOODY, C.-K. PENG, AND H. E. STANLEY, *Physiobank, physiokit, and physionet*, Circulation, 101 (2000), pp. e215–e220.
- [22] D. D. HOLM, A. TROUVÉ, AND L. YOUNES, *The Euler-Poincaré theory of metamorphosis*, Quart. Appl. Math., 67 (2009), pp. 661–685.
- [23] P. D. LAX, *Gibbs phenomena*, Journal of Scientific Computing, 28 (2006), pp. 445–449.
- [24] M. LIERO, A. MIELKE, AND G. SAVARÉ, *Optimal entropy-transport problems and a new hellinger–kantorovich distance between positive measures*, Inventiones mathematicae, 211 (2017), pp. 969–1117.
- [25] J.-G. LIU, R. L. PEGO, AND D. SLEPČEV, *Least action principles for incompressible flows and geodesics between shapes*, Calc. Var. Partial Differential Equations, 58 (2019), pp. Paper No. 179, 43.
- [26] D. G. LUENBERGER AND Y. YE, *Linear and Nonlinear Programming*, vol. 228, Springer, 2015.
- [27] L. MÉTIVIER, R. BROSSIER, Q. MÉRIGOT, AND E. OUDET, *A graph space optimal transport distance as a generalization of L^p distances: application to a seismic imaging inverse problem*, Inverse Problems, 35 (2019), pp. 085001, 49.
- [28] P. W. MICHOR AND D. MUMFORD, *Riemannian geometries on spaces of plane curves*, J. Eur. Math. Soc. (JEMS), 8 (2006), pp. 1–48.

- [29] M. I. MILLER, A. TROUVÉ, AND L. YOUNES, *Geodesic shooting for computational anatomy*, J. Math. Imaging Vision, 24 (2006), pp. 209–228.
- [30] M. I. MILLER AND L. YOUNES, *Group actions, homeomorphisms, and matching: A general framework*, International Journal of Computer Vision, 41 (2001), pp. 61–84.
- [31] G. NARDI, G. PEYRÉ, AND F.-X. VIALARD, *Geodesics on shape spaces with bounded variation and Sobolev metrics*, SIAM J. Imaging Sci., 9 (2016), pp. 238–274.
- [32] M. RUMPF AND B. WIRTH, *Discrete geodesic calculus in shape space and applications in the space of viscous fluidic objects*, SIAM J. Imaging Sci., 6 (2013), pp. 2581–2602.
- [33] ———, *Variational time discretization of geodesic calculus*, IMA J. Numer. Anal., 35 (2015), pp. 1011–1046.
- [34] M. THORPE, S. PARK, S. KOLOURI, G. K. ROHDE, AND D. SLEPČEV, *A transportation L^p distance for signal analysis*, J. Math. Imaging Vision, 59 (2017), pp. 187–210.
- [35] A. TROUVÉ, *Action de groupe de dimension infinie et reconnaissance de formes*, C. R. Acad. Sci. Paris Sér. I Math., 321 (1995), pp. 1031–1034.
- [36] A. TROUVÉ AND L. YOUNES, *Local geometry of deformable templates*, SIAM J. Math. Anal., 37 (2005), pp. 17–59.
- [37] ———, *Metamorphoses through Lie group action*, Found. Comput. Math., 5 (2005), pp. 173–198.
- [38] P. TSENG, *Convergence of a block coordinate descent method for nondifferentiable minimization*, Journal of optimization theory and applications, 109 (2001), p. 475.
- [39] Y. WANG, W. YIN, AND J. ZENG, *Global convergence of admm in nonconvex nonsmooth optimization*, Journal of Scientific Computing, 78 (2019), pp. 29–63.
- [40] B. WIRTH, L. BAR, M. RUMPF, AND G. SAPIRO, *A continuum mechanical approach to geodesics in shape space*, Int. J. Comput. Vis., 93 (2011), pp. 293–318.
- [41] L. YOUNES, *Shapes and diffeomorphisms*, vol. 171 of Applied Mathematical Sciences, Springer-Verlag, Berlin, 2010.
- [42] W. P. ZIEMER, *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, 1989.

(Ruiyu Han and Dejan Slepčev) DEPARTMENT OF MATHEMATICAL SCIENCES, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PA 15213

Email address: ruiyuh@andrew.cmu.edu, slepcev@math.cmu.edu

(Yunan Yang) INSTITUTE FOR THEORETICAL STUDIES, ETH ZÜRICH, ZÜRICH, SWITZERLAND 8092

Email address: yunan.yang@eth-its.ethz.ch