

AVERAGE-DISTANCE PROBLEM WITH CURVATURE PENALIZATION FOR DATA PARAMETERIZATION: REGULARITY OF MINIMIZERS

XIN YANG LU^{1,2,*} AND DEJAN SLEPČEV³

Abstract. We propose a model for finding one-dimensional structure in a given measure. Our approach is based on minimizing an objective functional which combines the average-distance functional to measure the quality of the approximation and penalizes the curvature, similarly to the elastica functional. Introducing the curvature penalization overcomes some of the shortcomings of the average-distance functional, in particular the lack of regularity of minimizers. We establish existence, uniqueness and regularity of minimizers of the proposed functional. In particular we establish $C^{1,1}$ estimates on the minimizers.

Mathematics Subject Classification. 49Q20, 49Q10, 35B65.

Received November 30, 2019. Accepted December 28, 2020.

1. INTRODUCTION

The average-distance problem was introduced by Buttazzo, Oudet and Stepanov in [5]. A number of properties of the solutions were established by Buttazzo and Stepanov in [10, 11], and by Paolini and Stepanov in [32]. An alternative variant, often referred to as “*penalized formulation*”, was introduced by Buttazzo, Mainini and Stepanov in [4]. Here we consider a more general form where we consider the general power of the distance $p \in [0, \infty)$, while the works above focused on $p = 1$:

Problem 1.1. Given $p \in [0, \infty)$, $d \geq 2$, a finite measure $\mu \geq 0$ with compact support in \mathbb{R}^d , and $\lambda > 0$, minimize

$$G_\mu^\lambda(\Sigma) := \int_{\mathbb{R}^d} d(x, \Sigma)^p d\mu(x) + \lambda \mathcal{H}^1(\Sigma), \quad d(x, \Sigma) := \inf_{y \in \Sigma} |x - y|,$$

with the unknown Σ varying in the family

$$\mathcal{A} := \{\Sigma \subseteq \mathbb{R}^d : \Sigma \text{ compact, path-wise connected and } \mathcal{H}^1(\Sigma) < +\infty\}.$$

Keywords and phrases: Average-distance problem, elastica functional, Willmore energy, curve fitting.

¹ Department of Mathematics and Statistics, Lakehead University, Thunder Bay, ON P7B5E1, Canada.

² Department of Mathematical Sciences, McGill University, Montréal, QC H3A0B9, Canada.

³ Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

* Corresponding author: xlu8@lakeheadu.ca

Problem 1.1 finds applications in urban planning (see for instance Buttazzo, Pratelli and Stepanov [7], Buttazzo and Santambrogio [8]; a review is available in Buttazzo and Santambrogio [9], and in the monograph by Buttazzo *et al.* [6]) and is closely related to several functionals in data analysis which look for low dimensional structures in data by minimizing the loss combined with a regularization term (see for instance Smola *et al.* [35]). In particular it is closely related to some functionals that arise as regularizations of principal curves, as we discuss below.

To consider the application of the average-distance problem in data parameterization consider a compactly supported measure μ describing the data distribution. Denote by Σ (the unknown) an one-dimensional object which parameterizes the, potentially noisy, data distribution. Then $\int_{\mathbb{R}^d} d(x, \Sigma)^p d\mu$ represents the approximation error, while $\lambda \mathcal{H}^1(\Sigma)$ is a cost associated to the complexity of such representation. Minimizing G_μ^λ is thus determining the “best” parameterization, balancing approximation error and complexity. In practice one would often deal with the empirical measure μ_n of a sample of μ , rather than μ itself, but note that in our setup μ can be a general measure, and thus one can consider μ_n in the place of μ when considering the basic properties of the minimizers when the empirical loss is used.

It is often of interest to use even simpler objects to approximate measures, namely parameterized curves. Let

$$\mathcal{C}^* := \{\varphi^* : [0, 1] \rightarrow \mathbb{R}^d : \varphi^* \text{ Lipschitz and } |(\varphi^*)'| \text{ is constant } \mathcal{L}^1\text{-a.e.}\}.$$

For curves $\varphi^* \in \mathcal{C}^*$, we define its “length” $L(\varphi^*)$ as its total variation:

$$L(\varphi^*) := \|\varphi^*\|_{TV([0,1])} = \sup_n \left(\sup_{0 \leq t_0 < \dots < t_{n-1} < t_n = 1} \sum_{j=1}^n |\varphi^*(t_j) - \varphi^*(t_{j-1})| \right).$$

For the sake of simplicity, in the following we will work with arc-length (instead of constant speed) parameterizations. Let

$$\mathcal{C} := \{\varphi : [0, a] \rightarrow \mathbb{R}^d : a \geq 0, \varphi \text{ Lipschitz and } |\varphi'| = 1 \text{ } \mathcal{L}^1\text{-a.e.}\}.$$

Similarly, given $\varphi \in \mathcal{C}$, $\varphi : [0, a] \rightarrow \mathbb{R}^d$, we define its “length” $L(\varphi)$ as its total variation. The map “reparameterization by constant speed”

$$\Phi : \mathcal{C} \rightarrow \mathcal{C}^*, \quad \Phi(\varphi) = \varphi^*, \quad \varphi^* : [0, 1] \rightarrow \mathbb{R}^d, \quad \varphi^*(t) := \varphi(ta) \quad (1.1)$$

will be frequently used. Thus, by construction, the domain of φ is $[0, L(\varphi)]$. More details about the space \mathcal{C} , including its topology, will be presented in Section 2. For any given curve ψ , we denote by Γ_ψ its image (which is independent of the parameterization).

The average-distance problem for parameterized curves is the following:

Problem 1.2. Given $p \in [1, \infty)$, $d \geq 2$, a finite measure $\mu \geq 0$ with compact support in \mathbb{R}^d , and $\lambda > 0$, minimize

$$E_\mu^\lambda(\varphi) := \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu + \lambda L(\varphi).$$

over all $\varphi \in \mathcal{C}$.

In mathematical literature on average-distance problem [4, 5, 10, 11, 32] the power $p = 1$ is considered, while in applications to machine learning $p = 2$ is the most common.

Problem 1.2 is an alternative to the classic “principal curve” introduced by Hastie and Stuetzle [22] and further studied by Duchamp and Stuetzle [17], for discovery and parameterization of one dimensional structures in data. Tibshirani [36] introduced principal curves with curvature penalty. In Kégl, Krzyzak, Linder, Zeger [23], the authors studied the principal curves with length constraint which is posed as a minimization of the mean squared distance and is hence similar to Problem 1.2. They also introduced an iterative algorithm to find the optimal curves. Subsequently, Biau and Fisher [2] proposed an equivalent formulation of principal curves, with the advantage of the self-consistency condition being explicit, and studied the problem with either length or curvature bounds. More recently, Delattre and Fischer [13] proved several theoretical properties, such as lack of self consistency, bounded curvature as a measure and absence of multiple points, of minimizers of the principal curves problem with length constraint. In terms of the regularity of curvature these results are similar (while the techniques are different) to results in [28, 29] where the penalized problem was studied. Kirov and one of authors [24] relaxed the Problem 1.2 to allow for multiple curves, and developed an efficient algorithm for both Problem 1.2 and the relaxed problem. Delicado [14] introduced a new notion of principal curves based on the so-called “principal oriented points”. Gerber and Whitaker [20] studied a relaxed notion of principal curves, (*i.e.* “weak principal curves”) and have shown it to be equal to the conditional expectation curve of a projection distance functional. A manifold learning algorithm based on the gradient descent of the projection distance functional (see *e.g.* [19], (5)) was then formulated in Gerber, Tasdizen, Whitaker [19]. Another notion of principal curves, defined based only on local quantities such as the gradient and Hessian matrix of an estimate of the probability density was proposed in Ozertem and Erdogmus [31]. Furthermore, Problem 1.2 is related to the lazy traveling salesman problem, see for instance Polak and Wolanski [33].

Let us recall what is known about the obstacles to regularity of minimizers of Problem 1.2:

- (*) In [34] it has been proven that, even if the reference measure μ satisfies $\mu \ll \mathcal{L}^d$ and $d\mu/d\mathcal{L}^d$ is C^∞ regular, Problem 1.1 may still admit minimizers which are simple curves failing to be C^1 regular. Moreover for any corner (*i.e.*, point where C^1 regularity fails), a positive amount of mass is projected on it (see [28], Lemma 2.1 for a detailed discussion about “projections”).
- (**) In [27] it has been proven that in general, even when the reference measure μ has the form $\mu = \sum_{j=1}^{\infty} a_j \chi_{A_j}$ (here χ denotes the characteristic function of the subscripted set), for suitable choices of parameters λ , $\{a_j\}$ and subsets $\{A_j\}$ ($j = 1, 2, \dots$), Problem 1.1 can admit minimizers which are simple curves for which C^1 regularity fails on a non closed set.

Noting that if a minimizer Σ of Problem 1.1 is a simple curve, then it admits a parameterization $\varphi \in \mathcal{C}$ minimizing Problem 1.2, it follows that the formulation of Problem 1.2 presents some drawbacks when used in data parameterization:

- fact (*) implies that a positive fraction of the data could be projected onto a single point. This is undesirable in data parameterization, since it results in a loss of information.
- In [28], Lemma 3.1 however it has been proven that the aforementioned issue is somewhat inevitable, since for any *endpoint*, at least mass $\lambda > 0$ is projected on it. Nevertheless, this implies that there are at most $\lfloor 1/\lambda \rfloor$ endpoints. Thus, from a practical point of view, endpoints can be “singled out” and analyzed with ad hoc arguments, mitigating the aforementioned issue. However fact (**) proves that the set of corners can be quite irregular, which makes “singling out” corners much more difficult than “singling out” endpoints.

It has been proven, by Lemenant in [25] and by the authors in [28], as well as by Delattre and Fischer [13] for the constrained problem, that regularity and curvature of any given subset of a minimizer is related to the amount for mass when $p = 1$ and otherwise the $p - 1$ -moment of mass that projects of the subset considered. We propose a modification of Problem 1.2, by adding a term penalizing the integrated squared curvature. We will consider the general case, where the power p belongs to $[1, \infty)$.

Problem 1.3. Given $d \geq 2$, a finite measure $\mu \geq 0$ compactly supported in \mathbb{R}^d , parameters $\lambda, \varepsilon > 0$, $p \geq 1$, minimize

$$\mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi) := \begin{cases} \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu + \lambda L(\varphi) + \varepsilon \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1, & \text{if } \varphi \in H^2([0, L(\varphi)]; \mathbb{R}^d) \cap \mathcal{C}, \\ +\infty & \text{otherwise,} \end{cases} \quad (1.2)$$

with $\kappa_\varphi := \frac{d}{dt} \frac{\varphi'}{\|\varphi'\|} : [0, L(\varphi)] \rightarrow \mathbb{R}$ denoting the curvature.

Note also

$$(\forall \varphi \in \mathcal{C}, \varepsilon > 0) \quad \mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi) < +\infty \text{ implies that } \varphi \text{ is } C^1 \text{ regular,}$$

since by Sobolev inequality in one dimension H^2 embeds in C^1 . The term $\int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1$ will be often referred to as “curvature term”. For future reference, given a curve $\psi \in \mathcal{C}$, the notation κ_ψ will denote the curvature of ψ . Existence of minimizers will be proven in Theorem 2.4.

The integrated squared curvature penalization is related to the Willmore energy, introduced in differential geometry and widely used in image analysis. Among the vast literature regarding Willmore energy, we cite Bretin, Lachaud and Oudet [3], where, the authors used the Willmore energy to develop an algorithm to reconstruct the contours of digital data, Du *et al.* [16], Dondl *et al.* [15], which are dedicated to formulating an approximation of the Willmore energy via the more easily treated phase field functions.

We note that curvature penalization related to principal curve problem has been considered by Tibshirani [36]. Indeed the functional studied in [36] has strong connections to ours. The focus of that paper however was on modeling and numerical algorithms.

We should point out that the curvature penalty was actually already introduced in the elastica functional, which is one of the earliest functionals of the calculus of variations. The elastica problem was proposed in 1691 by Jacob Bernoulli, who studied “the bendings or curvatures of beams, drawn bows, or of springs of any kind, caused by their own weight or by an attached weight or by any other compressing forces...”, [21], page 17. Indeed, if the measure μ in Problem 1.3 is purely atomic, then the term $\int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu$ is a type of “pulling” force on the minimizer, while length (resp. curvature) term $L(\gamma)$ (resp. $\int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1$) account for the elastic forces. The elastica problem was studied by Bernoulli [1] who provided a partial solution. Subsequently, Euler described in [18] two radically different solutions. Both of these works played a fundamental role in the history of the calculus of variations and their influence extends to this day. For a comprehensive look at the history of elastica problem see the article of Levien [26] and the doctoral thesis of Goss [21].

Problem 1.3 is also related to the notion of splines [12]. In particular the cubic smoothing spline for the given collection of data points $\{(x_i, y_i) : x_i \in [0, 1], y_i \in \mathbb{R}\}$ where $x_i < x_{i+1}$ for all $i = 1, \dots, n-1$ is the function $f : [0, 1] \rightarrow \mathbb{R}$ minimizing

$$\frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2 + \varepsilon \int_0^1 f''(x) dx.$$

Between the points x_i and x_{i+1} the function f is a cubic polynomial, while at the points x_i , f is differentiable, but the second derivative typically has a jump. As in our problem ε controls the level of smoothness. In the $\varepsilon \rightarrow \infty$ limit, the problem converges to the best linear fit, while in the $\varepsilon \rightarrow 0+$ limit the solutions converge to the interpolating cubic spline.

The main difference compared to our problem is that if one considers splines as curves in 2D, $x \mapsto (x, f(x))$ then one coordinate, x is given a-priori, and it already specifies the order in which the points are visited. The lack of information/requirement about the order in which the data points are to be approximated is a major source of complication for our problem. Among other things it is the reason why one needs to add the length penalty, as otherwise the problem does not have a minimizer.

The connection between Problems 1.3 and 1.2 will be discussed in Lemma 2.6. The main goal is to prove that minimizers are $C^{1,1}$ regular. The main result is:

Theorem 1.4. *Given $d \geq 2$, a finite measure μ with compact support in \mathbb{R}^d , and parameters $\lambda, \varepsilon > 0$, $p \geq 1$, any minimizer φ of $\mathcal{E}_\mu^{\lambda, \varepsilon, p}$ is $C^{1,1}$ regular and φ' is Lipschitz continuous, with Lipschitz constant not greater than*

$$Y(\text{diam supp}(\mu), \lambda, p, \varepsilon) := \sqrt{\frac{2(2\vartheta + \lambda)}{\varepsilon}}, \quad (1.3)$$

where

$$\vartheta := pD^{p-1}, \quad D := 2\text{diam supp}(\mu). \quad (1.4)$$

Note that $Y(\text{diam supp}(\mu), \lambda, p, \varepsilon)$ diverges as $\varepsilon \rightarrow 0$: this is indeed necessary, in view of Lemmas 2.6 and 2.7 below. This paper is structured as follows. In Section 2 we present some preliminary definitions and results, and prove existence of minimizers of Problem 1.3. In Section 3 we prove that minimizers of Problem 1.3 are $C^{1,1}$ regular. Moreover, we obtain estimates on the mass projecting onto a subset in relation to its length.

2. PRELIMINARIES

The aim of this section is to present preliminary notions and results, which will be useful in Section 3. In particular we will prove existence of minimizers for Problem 1.3 in Theorem 2.4. The first step is to endow the space \mathcal{C} with a suitable topology.

We define the following metric on \mathcal{C} :

$$m(\varphi, \psi) = \min \{ \|\Phi(\varphi) - \Phi(\psi)\|_{L^\infty}, \|\Phi(\varphi(L(\varphi) - \cdot)) - \Phi(\psi)\|_{L^\infty} \} + |L(\varphi) - L(\psi)|. \quad (2.1)$$

It is straightforward to check that m is a metric on \mathcal{C} . Intuitively, the convergence with respect to m is equivalent to orientation invariant uniform convergence, plus convergence of length. In proofs it is sometimes useful to identify the preferred orientation along a sequence. In particular given a sequence $\{\varphi_n\} \subseteq \mathcal{C}$ and $\varphi \in \mathcal{C}$ we let

$$\zeta_n := \begin{cases} \Phi(\varphi_n) & \text{if } \|\Phi(\varphi_n) - \Phi(\varphi)\|_{L^\infty(0,1)} \leq \|\Phi(\varphi_n(L(\varphi_n) - \cdot)) - \Phi(\varphi)\|_{L^\infty(0,1)}, \\ \Phi(\varphi_n(L(\varphi_n) - \cdot)) & \text{if } \|\Phi(\varphi_n) - \Phi(\varphi)\|_{L^\infty(0,1)} \geq \|\Phi(\varphi_n(L(\varphi_n) - \cdot)) - \Phi(\varphi)\|_{L^\infty(0,1)}, \end{cases}$$

where Φ is the function defined in (1.1). We note that the sequence $\{\varphi_n\}$, converges to φ with respect to metric m and we write $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi$ if:

- (i) $L(\varphi_n) \rightarrow L(\varphi)$ as $n \rightarrow \infty$ and
- (ii) the sequence $\{\zeta_n\}$, converges uniformly to $\Phi(\varphi)$ as $n \rightarrow \infty$.

We note that the notion of convergence implies that if $\varphi_n \xrightarrow{\mathcal{C}} \varphi$, then for arbitrary times $t_1 < t_2$ it holds $L(\Phi^{-1}(\zeta_n)|_{[t_1, t_2]}) \rightarrow L(\varphi|_{[t_1, t_2]})$ as $n \rightarrow \infty$, (i.e., the lengths of restrictions converge).

We turn to establishing some technical lemmas. Note that for any measure μ , $\lambda > 0$, $p \geq 1$, it holds

$$\inf_{\mathcal{C}} E_\mu^{\lambda, p} \leq \int_{\text{supp}(\mu)} |x - z|^p d\mu(x) \leq \text{diam}(\text{supp}(\mu))^p < +\infty, \quad (2.2)$$

where

$$E_\mu^{\lambda,p}(\varphi) := \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu(x) + \lambda L(\varphi),$$

and $z \in \text{supp}(\mu)$ is an arbitrarily chosen point, since clearly $\psi : \{0\} \rightarrow \{z\}$ belongs to \mathcal{C} . In particular, it can be assumed that any minimizing sequence $\{\gamma_n\}$ satisfies

$$\sup_n L(\gamma_n) < +\infty, \quad \sup_n \int_0^{L(\gamma_n)} \kappa_{\gamma_n}^2 ds < +\infty.$$

Lemma 2.1. *Given $d \geq 2$, a compact set $K \subseteq \mathbb{R}^d$, let $\{\gamma_n\} : [0, 1] \rightarrow K$ be a sequence of constant speed parameterized curves satisfying*

$$\sup_n \|\gamma_n'\|_{TV} < +\infty, \quad \sup_n L(\gamma_n) < +\infty,$$

with $\|\cdot\|_{TV}$ denoting the total variation semi-norm. Then there exists a subsequence of $\{\gamma_n\}$ which we do not relabel, and a curve $\gamma : [0, 1] \rightarrow K$ such that

$$\begin{aligned} \{\gamma_n\} &\rightarrow \gamma && \text{in } C^\alpha \text{ for any } \alpha \in [0, 1), \\ \{\gamma_n'\} &\rightarrow \gamma' && \text{in } L^p \text{ for any } p \in [1, +\infty), \\ \{\kappa_{\gamma_n}\} &\xrightarrow{*} \kappa_\gamma && \text{in the space of signed Borel measures.} \end{aligned}$$

It is worth noting that in our case, since we work with curves γ with uniformly bounded energy, and hence uniformly bounded $\int_0^1 \kappa_\gamma^2 ds$, the weak-* convergence of the curvature measures implies the weak L^2 convergence.

Proof. For the proof we refer to [34], Lemma 6. □

Lemma 2.2. *Given a sequence of constant (positive) speed curves $\{\gamma_n\} : [0, 1] \rightarrow \mathbb{R}^d$, converging uniformly to $\gamma : [0, 1] \rightarrow \mathbb{R}^d$, such that*

$$\sup_n |\gamma_n'| < +\infty, \quad \sup_n \int_0^1 \kappa_{\gamma_n}^2 ds < +\infty, \tag{2.3}$$

then it holds

$$L(\gamma) = \lim_{n \rightarrow +\infty} L(\gamma_n).$$

Note that this is a much stronger result than the general lower semicontinuity of length. In particular, due to the curvature penalization, it states that any minimizing sequence $\{\gamma_n\}$ (which surely satisfies (2.3) which admits a uniform limit γ , satisfies also $\{L(\gamma_n)\} \rightarrow L(\gamma)$. This will be crucial for the proof of Lemma 2.3.

For brevity, given two vectors v_1, v_2 , the notation $\angle v_1 v_2$ will denote the angle between v_1 and v_2 . That is,

$$\angle v_1 v_2 \in [0, \pi], \quad \angle v_1 v_2 := \arccos \frac{\langle v_1, v_2 \rangle}{|v_1| |v_2|},$$

where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean scalar product of \mathbb{R}^d .

Proof. Let $\{\gamma_n\}$ be a sequence satisfying (2.3).

Case 1^o: Γ_γ is a singleton. We show that since $\{\gamma_n\} \rightarrow \gamma$ uniformly, the uniform bound on the integrated squared curvature gives $\{L(\gamma_n)\} \rightarrow 0$. Without loss of generality, let $\Gamma_\gamma = \{0\}$, and let $\{\varepsilon_n\} \downarrow 0$ be a sequence such that $\Gamma_{\gamma_n} \subseteq B(0, \varepsilon_n)$. Let $\tilde{\gamma}_n$ be the arc-length reparameterizations of γ_n . We first claim that:

- for any n , and a.e. $t \leq L(\gamma_n) - 3\sqrt{2}\varepsilon_n$, there exists $s \in [t, t + 3\sqrt{2}\varepsilon_n]$ such that

$$\langle \tilde{\gamma}'_n(t), \tilde{\gamma}'_n(s) \rangle \leq \frac{1}{\sqrt{2}}.$$

To show this claim, it suffices to note that if this were not the case, *i.e.* $\langle \tilde{\gamma}'_n(t), \tilde{\gamma}'_n(s) \rangle \geq \frac{1}{\sqrt{2}}$ for all $s \in [t, t + 3\sqrt{2}\varepsilon_n]$, then

$$\langle \tilde{\gamma}_n(t + 3\sqrt{2}\varepsilon_n) - \tilde{\gamma}_n(t), \tilde{\gamma}'_n(t) \rangle \geq \int_t^{t+3\sqrt{2}\varepsilon_n} \langle \tilde{\gamma}_n(s), \tilde{\gamma}'_n(t) \rangle ds \geq 3\varepsilon_n,$$

which contradicts the fact that $\Gamma_{\gamma_n} \subseteq B(0, \varepsilon_n)$. The claim is thus proven.

If $L(\gamma_n) > 3\sqrt{2}\varepsilon_n$ were to hold, then there exist $s_n > t_n$ such that

$$s_n - t_n \leq 3\sqrt{2}\varepsilon_n, \quad \langle \tilde{\gamma}'_n(t_n), \tilde{\gamma}'_n(s_n) \rangle \leq \frac{1}{\sqrt{2}}.$$

This in turn gives

$$\int_0^1 \kappa_{\gamma_n}^2 ds \geq \int_{t_n}^{s_n} |\tilde{\gamma}''_n|^2 ds \geq \frac{1}{s_n - t_n} \left| \int_{t_n}^{s_n} \tilde{\gamma}''_n ds \right| \geq \frac{|\tilde{\gamma}'_n(t_n) - \tilde{\gamma}'_n(s_n)|^2}{3\sqrt{2}\varepsilon_n} \geq \frac{|1 - \frac{1}{\sqrt{2}}|^2}{3\sqrt{2}\varepsilon_n},$$

where the last inequality follows from

$$|\tilde{\gamma}'_n(t_n) - \tilde{\gamma}'_n(s_n)| \geq \langle \tilde{\gamma}'_n(t_n) - \tilde{\gamma}'_n(s_n), \tilde{\gamma}'_n(t_n) \rangle = 1 - \underbrace{\langle \tilde{\gamma}'_n(s_n), \tilde{\gamma}'_n(t_n) \rangle}_{\leq 1/\sqrt{2}}.$$

This clearly contradicts (2.3). Thus the only possibility is $L(\gamma_n) \leq 3\sqrt{2}\varepsilon_n$ for all sufficiently large n , which clearly implies $L(\gamma_n) \rightarrow 0 = L(\gamma)$.

Case 2^o Γ_γ is not a singleton. Without loss of generality, we can assume $\inf_n |\gamma'_n| > 0$. This because by assumption we have $\gamma_n \rightarrow \gamma$ uniformly. As γ_n are arc-length parameterized, $|\gamma'_n|$ is equal to its length $L(\gamma_n) = \mathcal{H}^1(\gamma_n)$, with \mathcal{H}^1 denoting the Hausdorff-1 measure. By the lower semicontinuity of \mathcal{H}^1 (Golab's theorem) we then have $\liminf_{n \rightarrow +\infty} \mathcal{H}^1(\gamma_n) \geq \mathcal{H}^1(\gamma) > 0$.

Note:

1. $\{\gamma_n\} \rightarrow \gamma$ uniformly, *i.e.* $\{\gamma_n\}$ is bounded in $L^\infty([0, 1]; \mathbb{R}^d)$;
2. since the speed of $\{\gamma_n\}$ is equal to $|\gamma'_n|$ a.e., hypothesis $\sup_n |\gamma'_n| < +\infty$ gives that $\{\gamma_n\}$ is bounded in $W^{1, \infty}([0, 1]; \mathbb{R}^d)$;
3. the curvature κ_{γ_n} describes the variation of the unit tangent derivative, and absolute continuity of κ_{γ_n} with respect to the Lebesgue measure (which follows from (2.3)) gives

$$\kappa_{\gamma_n} = |\gamma'_n|^{-1} \frac{d}{ds} \left(\frac{\gamma'_n}{|\gamma'_n|} \right) = |\gamma'_n|^{-1} \frac{\gamma''_n}{|\gamma'_n|}.$$

Since $\inf_n |\gamma'_n| > 0$, (2.3) implies that $\{\gamma_n\}$ is bounded in $W^{2,2}([0, 1]; \mathbb{R}^d)$.

Thus there exists $\tilde{\gamma} \in W^{2,2}([0, 1]; \mathbb{R}^d)$ and a subsequence $\{\gamma_{n_k}\}$ converging to $\tilde{\gamma}$ in the weak topology of $W^{2,2}([0, 1]; \mathbb{R}^d)$. Since the embedding $W^{2,2}([0, 1]; \mathbb{R}^d) \hookrightarrow W^{1,2}([0, 1]; \mathbb{R}^d)$ is compact, $\{\gamma_{n_k}\}$ converges to $\tilde{\gamma}$ strongly in $W^{1,2}([0, 1]; \mathbb{R}^d)$, hence $\{\gamma_{n_k}\}$ converges to $\tilde{\gamma}$ strongly in $L^2([0, 1]; \mathbb{R}^d)$. Since by hypothesis $\{\gamma_{n_k}\}$ converges to γ strongly in $L^\infty([0, 1]; \mathbb{R}^d)$, it follows $\tilde{\gamma} = \gamma$. Strong convergence of $\{\gamma_{n_k}\}$ to γ in $W^{1,2}([0, 1]; \mathbb{R}^d)$ implies

$$L(\gamma_{n_k}) = \int_0^1 |\gamma'_{n_k}| ds \rightarrow \int_0^1 |\gamma'| ds = L(\gamma).$$

To achieve convergence for the whole sequence, note that if there exists a subsequence $\{\gamma_{n_h}\}$ such that $\lim_{h \rightarrow +\infty} L(\gamma_{n_h}) \neq L(\gamma)$, then the above construction allows to further extract a sub-subsequence $\{\gamma_{n_{h(j)}}\}$ converging to γ in the strong topology of $W^{1,2}([0, 1]; \mathbb{R}^d)$, which in turn implies $\lim_{j \rightarrow +\infty} L(\gamma_{n_{h(j)}}) = L(\gamma)$. Contradiction. \square

The next result proves compactness of sub-levels for $\mathcal{E}_\mu^{\lambda,p}$.

Lemma 2.3. *Given $d \geq 2$, finite measure $\mu \geq 0$ compactly supported in \mathbb{R}^d , parameters $\lambda, \varepsilon > 0$, $p \geq 1$, $M \geq \inf_{\mathcal{C}} \mathcal{E}_\mu^{\lambda,\varepsilon,p}$, a sequence $\{\varphi_n\} \subseteq \mathcal{C} \cap \{\mathcal{E}_\mu^{\lambda,\varepsilon,p} \leq M\}$, there exists $\varphi_\infty \in \mathcal{C}$ such that (upon subsequence, which will not be relabeled) $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi_\infty$.*

Proof. Consider an arbitrary sequence $\{\varphi_n\} \subseteq \mathcal{C} \cap \{\mathcal{E}_\mu^{\lambda,\varepsilon,p} \leq M\}$. If there exists a subsequence $\{n_k\}$ such that $\{L(\varphi_{n_k})\} \rightarrow 0$, then $\{\Gamma_{\varphi_{n_k}}\} \xrightarrow{d_{\mathbb{H}}} \{z\}$ for some $z \in \mathbb{R}^d$, and letting $\varphi_\infty : \{0\} \rightarrow \{z\}$ concludes the proof. Now assume

$$\liminf_{n \rightarrow +\infty} L(\varphi_n) > 0. \quad (2.4)$$

Let $\varphi_n^* := \Phi(\varphi_n)$, $n \in \mathbb{N}$. We show that $\{\varphi_n^*\}$ is bounded in $W^{2,2}([0, 1]; \mathbb{R}^d)$.

Claim 1. $\sup_n L(\varphi_n^*) < +\infty$.

Consider an arbitrary index n . Since $\mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi_n) \leq M$, it follows

$$\lambda L(\varphi_n^*) = \lambda L(\varphi_n) \leq \mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi_n) \leq M \implies L(\varphi_n^*) \leq M/\lambda.$$

Claim 2. $\sup_n \int_0^1 |\kappa_{\varphi_n^*}|^2 d\mathcal{L}^1 < +\infty$.

Consider an arbitrary index n . Note that, for any $\varphi \in \mathcal{C} \cap \{\mathcal{E}_\mu^{\lambda,\varepsilon,p} \leq M\}$ with $L(\varphi) > 0$, it holds

$$\begin{aligned} \int_0^1 |\kappa_{\varphi^*}|^2 d\mathcal{L}^1 &= \int_0^1 \left| \frac{d}{ds} \left(\frac{\varphi'(sL(\varphi))}{|(\varphi^*)'|} \right) \right|^2 ds = \int_0^1 |\varphi''(sL(\varphi))|^2 ds \\ &\stackrel{s' := sL(\varphi)}{=} \frac{1}{L(\varphi)} \int_0^{L(\varphi)} |\varphi''(s')|^2 ds' = \frac{1}{L(\varphi)} \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1, \end{aligned} \quad (2.5)$$

where $\varphi^* := \Phi(\varphi)$. Condition $\mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi_n) \leq M$ then gives

$$\varepsilon L(\varphi_n) \int_0^1 |\kappa_{\varphi_n^*}|^2 d\mathcal{L}^1 = \varepsilon \int_0^{L(\varphi_n)} |\kappa_\varphi|^2 d\mathcal{L}^1 \leq \mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi_n) \leq M.$$

Claim 3. $\bigcup_n \Gamma_{\varphi_n} \subseteq K$ for some compact set K .

Consider an arbitrary index n . Given r , if $\Gamma_{\varphi_n} \cap (\text{supp}(\mu))_r = \emptyset$, where

$$(\text{supp}(\mu))_r := \{x \in \mathbb{R}^d : d(x, \text{supp}(\mu)) \leq r\},$$

then

$$\int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu \geq r^p.$$

Claim 1 proves $\sup_n L(\varphi_n) \leq M/\lambda$, thus if $\Gamma_{\varphi_n} \cap (\mathbb{R}^d \setminus (\text{supp}(\mu))_{(2M)^{1/p} + M/\lambda})$ is non empty, then

$$\Gamma_{\varphi_n} \cap (\text{supp}(\mu))_{(1.5M)^{1/p}} = \emptyset,$$

hence

$$M \geq \mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi_n) \geq \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu \geq 1.5M.$$

Contradiction. Letting $K := (\text{supp}(\mu))_{(2M)^{1/p} + M/\lambda}$ proves the claim.

Claims 1, 2 and 3 prove that $\{\varphi_n^*\}$ is bounded in $W^{2,2}([0, 1]; \mathbb{R}^d)$, hence (upon subsequence) it converges to φ_∞^* in the strong topology of $W^{1,2}([0, 1]; \mathbb{R}^d)$. Therefore, denoting by φ_∞ the arc-length reparameterization of φ_∞^* , we get $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi_\infty$, concluding the proof. \square

Now we can prove existence of minimizers.

Theorem 2.4. *For $d \geq 2$, given finite measure $\mu \geq 0$ compactly supported in \mathbb{R}^d , and parameters $\lambda, \varepsilon > 0$, $p \geq 1$, the functional $\mathcal{E}_\mu^{\lambda, \varepsilon, p}$ admits a minimizer in \mathcal{C} .*

Proof. Consider an arbitrary minimizing sequence $\{\varphi_n\} \subseteq \mathcal{C}$. In view of (2.2), assume

$$\mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi_n) \leq (\text{diam supp}(\mu))^p + \xi \quad \text{for some } 0 < \xi \ll 1.$$

Lemma 2.3 gives the existence of a limit curve φ_∞ . Convergence $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi_\infty$ gives

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu = \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_\infty})^p d\mu, \quad \lim_{n \rightarrow +\infty} L(\varphi_n) = L(\varphi_\infty). \quad (2.6)$$

It remains to prove

$$\liminf_{n \rightarrow +\infty} \int_0^{L(\varphi_n)} |\kappa_{\varphi_n}|^2 d\mathcal{L}^1 \geq \int_0^{L(\varphi_\infty)} |\kappa_{\varphi_\infty}|^2 d\mathcal{L}^1. \quad (2.7)$$

Let $\varphi_n^* := \Phi(\varphi_n)$, $n \in \mathbb{N}$, and $\varphi_\infty^* := \Phi(\varphi_\infty)$. Note that

$$\sup_n \int_0^{L(\varphi_n)} |\kappa_{\varphi_n}|^2 d\mathcal{L}^1 \leq \sup_n \frac{\mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi_n)}{\varepsilon} \leq \frac{(\text{diam supp}(\mu))^p + \xi}{\varepsilon} < +\infty,$$

thus, in view of (2.5), $\sup_n \int_0^{L(\varphi_n)} |\kappa_{\varphi_n^*}|^2 d\mathcal{L}^1 < +\infty$, and the sequence $\{d\kappa_{\varphi_n^*}/d\mathcal{L}^1\}$ is bounded in $L^2([0, 1])$. Therefore (upon subsequence, which will not be relabeled) there exist $g \in L^2([0, 1])$ such that $\{d\kappa_{\varphi_n^*}/d\mathcal{L}^1\} \rightharpoonup g$, hence

$$\left\{ \frac{d\kappa_{\varphi_n^*}}{d\mathcal{L}^1} \cdot \mathcal{L}^1_{[0,1]} \right\} \rightharpoonup g \cdot \mathcal{L}^1_{[0,1]} \quad \text{weakly in } L^2([0, 1]), \quad \|g\|_{L^2([0,1])} \leq \liminf_{n \rightarrow +\infty} \left\| \frac{d\kappa_{\varphi_n^*}}{d\mathcal{L}^1} \right\|_{L^2([0,1])}. \quad (2.8)$$

Lemma 2.1 gives (upon subsequence, which will not be relabeled) $\{\kappa_{\varphi_n^*}\} \xrightarrow{*} \kappa_{\varphi_\infty^*}$, hence $\kappa_{\varphi_\infty^*} = g \cdot \mathcal{L}^1_{[0,1]}$. Combining (2.8), observation (2.5), and $\{L(\varphi_n)\} \rightarrow L(\varphi_\infty)$ proves (2.7). Combining (2.6) and (2.7) gives $\mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi_\infty) \leq \liminf_{n \rightarrow +\infty} \mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi_n)$, concluding the proof. \square

Lemma 2.5. *Given $d \geq 2$, a finite measure $\mu \geq 0$ compactly supported in \mathbb{R}^d , parameters $\lambda, \varepsilon > 0$, $p \geq 1$ and a minimizer $\varphi \in \operatorname{argmin} \mathcal{E}_\mu^{\lambda, \varepsilon}$, it holds:*

1. *length estimate:*

$$L(\varphi) \leq \frac{(\operatorname{diam} \operatorname{supp}(\mu))^p}{\lambda}, \quad (2.9)$$

2. *curvature term estimate:*

$$\int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 \leq \frac{(\operatorname{diam} \operatorname{supp}(\mu))^p}{\varepsilon}, \quad (2.10)$$

3. *confinement condition:* $\Gamma_\varphi \subseteq (\operatorname{supp}(\mu))_{\operatorname{diam} \operatorname{supp}(\mu) + (\operatorname{diam} \operatorname{supp}(\mu))^p/\lambda}$, where for given $r \geq 0$,

$$(\operatorname{supp}(\mu))_r := \{x \in \mathbb{R}^d : d(x, \operatorname{supp}(\mu)) \leq r\}.$$

Proof. Estimates (2.9) follows from

$$\lambda L(\varphi) \leq \mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi) = \inf_{\mathcal{C}} \mathcal{E}_\mu^{\lambda, \varepsilon, p} \stackrel{(2.2)}{\leq} (\operatorname{diam} \operatorname{supp}(\mu))^p,$$

while (2.10) follows from

$$\varepsilon \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 \leq \mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi) = \inf_{\mathcal{C}} \mathcal{E}_\mu^{\lambda, \varepsilon, p} \stackrel{(2.2)}{\leq} (\operatorname{diam} \operatorname{supp}(\mu))^p.$$

To prove the confinement condition, note that for any $\psi \in \mathcal{C}$, if $\Gamma_\psi \cap (\operatorname{supp}(\mu))_r = \emptyset$, then it follows $\int_{\mathbb{R}^d} d(x, \Gamma_\psi)^p d\mu \geq r^p$. Inequality (2.2) gives $\Gamma_\psi \cap (\operatorname{supp}(\mu))_{\operatorname{diam} \operatorname{supp}(\mu)} \neq \emptyset$. Combining with length estimate (2.9) concludes the proof. \square

The next result proves a connection between Problems 1.3 and 1.2. Let $E_\mu^{\lambda, p} : \mathcal{C} \rightarrow [0, +\infty)$ be defined by

$$E_\mu^{\lambda, p}(\psi) := \int_{\mathbb{R}^d} d(x, \Gamma_\psi)^p d\mu + \lambda L(\psi).$$

Lemma 2.6. *Given $d \geq 2$, a finite measure $\mu \geq 0$ compactly supported in \mathbb{R}^d , parameters $\lambda > 0$, $p \geq 1$, and a sequence $\{\varepsilon_n\} \rightarrow 0$, then $\{\mathcal{E}_\mu^{\lambda, \varepsilon_n, p}\}$ Γ -converges to $E_\mu^{\lambda, p}$ in the topology of \mathcal{C} as $n \rightarrow +\infty$.*

Proof. The proof of Γ -convergence is split into two steps.

Step 1. Γ -lim inf inequality: for any sequence $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi$ it holds

$$\liminf_{n \rightarrow +\infty} \mathcal{E}_\mu^{\lambda, \varepsilon_n, p}(\varphi_n) \geq E_\mu^{\lambda, p}(\varphi).$$

Since $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi$, it follows $\{\Gamma_{\varphi_n}\} \xrightarrow{d_{\mathcal{H}}} \Gamma_\varphi$ and $\{L(\varphi_n)\} \rightarrow L(\varphi)$. As proven in [5], the functional

$$\mathcal{A} \ni \Sigma \mapsto \int_{\mathbb{R}^d} d(x, \Sigma)^p d\mu$$

is continuous with respect to Hausdorff distance, hence

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu = \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu.$$

Thus we have

$$\begin{aligned} \liminf_{n \rightarrow +\infty} \mathcal{E}_\mu^{\lambda, \varepsilon_n, p}(\varphi_n) &\geq \liminf_{n \rightarrow +\infty} \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu + \lambda L(\varphi_n) \\ &= \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu + \lambda L(\varphi) \\ &= E_\mu^{\lambda, p}(\varphi). \end{aligned}$$

Step 2. Γ -lim sup inequality: for any φ , there exists a sequence $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi$ such that

$$\limsup_{n \rightarrow +\infty} \mathcal{E}_\mu^{\lambda, \varepsilon_n, p}(\varphi_n) \leq E_\mu^{\lambda, p}(\varphi).$$

Let $\{\tilde{\varphi}_k\}$ be a sequence of smooth functions such that $\{\tilde{\varphi}_k\} \rightarrow \varphi$ uniformly. By relabeling the sequence, using $k(n) = \max \left\{ 1, \sup \left\{ k : \int_0^{L(\tilde{\varphi}_k)} |\kappa_{\tilde{\varphi}_k}|^2 d\mathcal{L}^1 \leq \varepsilon_n^{-1/2} \right\} \right\}$, and defining $\varphi_n = \tilde{\varphi}_{k(n)}$ we have that for all n large enough

$$\int_0^{L(\varphi_n)} |\kappa_{\varphi_n}|^2 d\mathcal{L}^1 \leq \varepsilon_n^{-1/2}, \quad \text{and} \quad \{\varphi_n\} \rightarrow \varphi \quad \text{uniformly.}$$

Thus

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu = \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu, \quad \lim_{n \rightarrow +\infty} \varepsilon_n \int_0^{L(\varphi_n)} |\kappa_{\varphi_n}|^2 d\mathcal{L}^1 = 0,$$

which gives

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathcal{E}_\mu^{\lambda, \varepsilon_n, p}(\varphi_n) &= \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu + \lambda L(\varphi_n) + \varepsilon_n \int_0^{L(\varphi_n)} |\kappa_{\varphi_n}|^2 d\mathcal{L}^1 \\ &= \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu + \lambda L(\varphi) \\ &= E_\mu^{\lambda, p}(\varphi), \end{aligned}$$

concluding the proof. \square

Note that the upper bound in (2.10) diverges as $\varepsilon \rightarrow 0$: this is indeed necessary, in view of Lemmas 2.6, and 2.7 below:

Lemma 2.7. ([34]) *There exist μ and λ such that $E_\mu^{\lambda,1}$ admits a unique minimizer which is not C^1 regular.*

Lemma 2.8. *Given $d \geq 2$, a sequence of measures $\{\mu_n\}$ compactly supported in \mathbb{R}^d , a finite measure $\mu \geq 0$ compactly supported in \mathbb{R}^d , such that $\{\mu_n\} \xrightarrow{*} \mu$, and parameters $\lambda, \varepsilon > 0$, $p \geq 1$, then $\{\mathcal{E}_{\mu_n}^{\lambda,\varepsilon,p}\}$ Γ -converges to $\mathcal{E}_\mu^{\lambda,\varepsilon,p}$ in the topology of \mathcal{C} .*

Proof. The proof will be split in two steps.

Γ -lim inf inequality: for any sequence $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi$, it holds $\liminf_{n \rightarrow +\infty} \mathcal{E}_{\mu_n}^{\lambda,\varepsilon,p}(\varphi_n) \geq \mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi)$.

If $\liminf_{n \rightarrow +\infty} \mathcal{E}_{\mu_n}^{\lambda,\varepsilon,p}(\varphi_n) = +\infty$, then the Γ -lim inf inequality follows. Thus assume (upon subsequence, which will not be relabeled) $\liminf_{n \rightarrow +\infty} \mathcal{E}_{\mu_n}^{\lambda,\varepsilon,p} = \lim_{n \rightarrow +\infty} \mathcal{E}_{\mu_n}^{\lambda,\varepsilon,p}(\varphi_n) < +\infty$. Convergences $\{\mu_n\} \xrightarrow{*} \mu$ and $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi$ give $\{L(\varphi_n)\} \rightarrow L(\varphi)$, and

$$\begin{aligned} \lim_{n \rightarrow +\infty} \left| \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu_n - \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu \right| &= \lim_{n \rightarrow +\infty} \left| \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu_n - \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu \right| \\ &\quad + \left| \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu - \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu \right| = 0. \end{aligned} \quad (2.11)$$

Then the same arguments from the proof of Theorem 2.4 prove lower semicontinuity for the curvature term, hence $\liminf_{n \rightarrow +\infty} \mathcal{E}_{\mu_n}^{\lambda,\varepsilon}(\varphi_n) \geq \mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi)$.

Γ -lim sup inequality: for any $\varphi \in \mathcal{C}$, there exists a sequence $\{\varphi_n\} \xrightarrow{\mathcal{C}} \varphi$ such that

$$\limsup_{n \rightarrow +\infty} \mathcal{E}_{\mu_n}^{\lambda,\varepsilon,p}(\varphi_n) \leq \mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi).$$

Given $\varphi \in \mathcal{C}$, case $\mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi) = +\infty$ is trivial. Assume $\mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi) < +\infty$, and let $\varphi_n := \varphi$ for any n . Clearly

$$(\forall n) \quad L(\varphi_n) = L(\varphi), \quad \int_0^{L(\varphi_n)} |\kappa_{\varphi_n}|^2 d\mathcal{L}^1 = \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1.$$

Since $\{\mu_n\} \xrightarrow{*} \mu$, it follows

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_n})^p d\mu_n = \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu,$$

concluding the proof. \square

3. REGULARITY

The aim of this section is to prove Theorem 1.4. As a consequence, we have Corollaries 3.3 and 3.4, which estimate (for minimizers) the mass projecting on a subset in relation to its length. This proves that Problem 1.3 is effectively a better candidate for application in data parameterization than Problem 1.2, since it does not exhibit the undesirable properties described in (*) and (**). The results are proven for generic (finite) dimension. However while Theorem 1.4 is proven for generic measures, Corollaries 3.3 and 3.4 are restricted to absolutely continuous measures.

We first prove a weaker regularity result.

Lemma 3.1. *Given $d \geq 2$, a finite measure $\mu \geq 0$ compactly supported in \mathbb{R}^d , and parameters $\lambda, \varepsilon > 0$, $p \geq 1$, any minimizer $\varphi \in \operatorname{argmin}_{\mathcal{C}} \mathcal{E}_{\mu}^{\lambda, \varepsilon, p}$ is $C^{1,1/2}$ regular with*

$$(\forall t_0, t_1 \in [0, L(\varphi)]) \quad |\varphi'(t_1) - \varphi'(t_0)| \leq \left(\frac{(\operatorname{diam} \operatorname{supp}(\mu))^p}{\varepsilon} |t_1 - t_0| \right)^{1/2}.$$

Proof. Let φ be an arbitrary minimizer. Definition (1.2) requires that $\varphi \in H^2([0, L(\varphi)]; \mathbb{R}^d)$, which, by Sobolev embedding, implies that φ' is continuous.

Thus φ is C^1 regular. Hölder inequality gives

$$\frac{|\varphi'(t_1) - \varphi'(t_0)|^2}{|t_1 - t_0|} \leq \frac{1}{|t_1 - t_0|} \left(\int_{t_0}^{t_1} |\kappa_{\varphi}| d\mathcal{L}^1 \right)^2 \leq \int_{t_0}^{t_1} |\kappa_{\varphi}|^2 d\mathcal{L}^1 \leq \frac{\mathcal{E}_{\mu}^{\lambda, \varepsilon, p}(\varphi)}{\varepsilon} \leq \frac{(\operatorname{diam} \operatorname{supp}(\mu))^p}{\varepsilon},$$

with the last inequality due to the fact that φ has no more energy compared to any singleton $\psi : \{0\} \rightarrow \{z\}$, $z \in \operatorname{supp}(\mu)$, *i.e.*

$$\mathcal{E}_{\mu}^{\lambda, \varepsilon, p}(\varphi) \leq \mathcal{E}_{\mu}^{\lambda, \varepsilon, p}(\psi) \leq \int_{\operatorname{supp}(\mu)} |x - z|^p d\mu(x) \leq \operatorname{diam}(\operatorname{supp}(\mu))^p.$$

The proof is thus complete. □

Note that the proof of Lemma 3.1 is quite simple, and consists of comparing the curvature term with condition (2.2), without any kind of construction. As one may expect, this kind of proof can be refined, and stronger regularity results can be established.

3.1. Lipschitz regularity of derivatives

Note that the mean value theorem implies that for any $0 < C < +\infty$

$$(\forall p \geq 1, 0 \leq a, b \leq C) \quad |a^p - b^p| \leq p|a - b|C^{p-1}. \quad (3.1)$$

Now we are ready to prove Theorem 1.4. Note that, for an vectors $v_1, v_2 \in S^{d-1}$ (unit ball of \mathbb{R}^d) it holds that

$$\frac{1}{2} \angle v_1 v_2 \leq |v_1 - v_2| \leq \angle v_1 v_2. \quad (3.2)$$

This follows from the fact that, $|v_1 - v_2| = 2 \sin \frac{\angle v_1 v_2}{2} \geq \angle v_1 v_2$, since $0 \leq \angle v_1 v_2 \leq \pi$. For a schematic representation, see Figure 1.

Proof. (of Thm. 1.4) Let φ be an arbitrary minimizer. Since Lemma 3.1 proved that φ' is continuous,

$$\lim_{|s-t| \rightarrow 0} |\varphi'(s) - \varphi'(t)| = \lim_{|s-t| \rightarrow 0} \angle \varphi'(s) \varphi'(t) = 0. \quad (3.3)$$

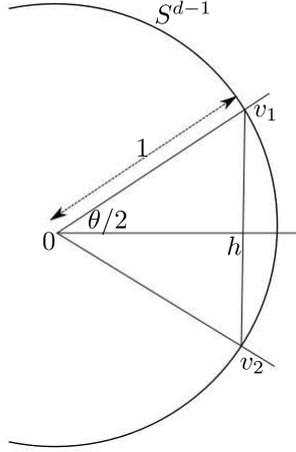


FIGURE 1. Schematic representation of $\theta := \angle v_1 v_2$. Elementary geometric considerations give $|v_1 - v_2| = 2|v_1 - h| = 2 \sin(\theta/2)$.

To show that φ' is Lipschitz continuous it suffices to show that

$$\sup_{0 \leq t_0 < L(\varphi)} \limsup_{t_1 \rightarrow t_0^+} \frac{|\varphi'(t_0) - \varphi'(t_1)|}{|t_1 - t_0|} < \infty.$$

Consider any $0 \leq t_0 < L(\varphi)$ and $M < \limsup_{t_1 \rightarrow t_0^+} \frac{|\varphi'(t_0) - \varphi'(t_1)|}{|t_1 - t_0|}$.

The main ideas are:

1. Using the original curve φ , we construct the competitor ψ in (3.4) below. Intuitively (see Fig. 2), ψ is obtained by first expanding by a factor of 2 a piece of the original curve φ with very high curvature (potentially losing connectedness in the process), and then translating in a suitable way the remaining part to regain connectedness (see (3.4)).
2. We estimate the difference $E_\mu^\lambda(\psi) - E_\mu^\lambda(\varphi)$, with E_μ^λ defined in Problem 1.2. Since the translations in the previous Step are by vectors with norm at most ξ , it follows that

$$\int_{\mathbb{R}^d} [d(x, \Gamma_\psi)^p - d(x, \Gamma_\varphi)^p] d\mu(x) \leq C_1 \xi,$$

for some geometric constant $C_1 > 0$ (Claim 1 below). Similarly, since we will add an arc of circle (to regain connectedness for ψ), the difference of lengths $L(\psi) - L(\varphi)$ is exactly ξ . Finally, we estimate the gain for the curvature term (Claim 2).

Let h be the homothety of center $\varphi(t_0)$ and ratio 2. That is, for any $x \in \mathbb{R}^d$, $h(x)$ is defined by $h(x) = \varphi(t_0) + 2(x - \varphi(t_0))$. For $t_1 > t_0$ consider the curve

$$\psi : [0, L(\varphi) + \xi] \longrightarrow \mathbb{R}^d, \quad \psi(t) := \begin{cases} \varphi(t) & \text{if } t \leq t_0, \\ h(\varphi((t_0 + t)/2)) & \text{if } t_0 \leq t \leq t_0 + 2\xi, \\ \varphi(t - \xi) + (h(\varphi(t_1)) - \varphi(t_1)) & \text{if } t \geq t_0 + 2\xi, \end{cases} \quad (3.4)$$

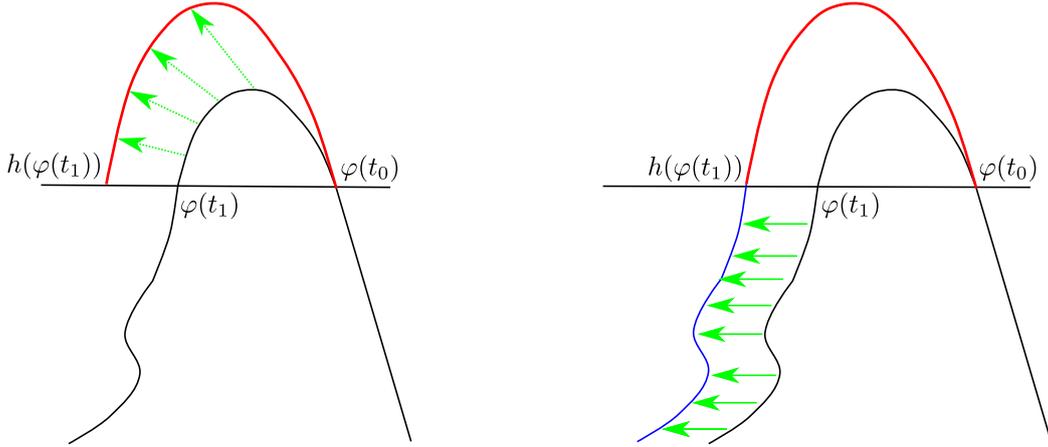


FIGURE 2. *Left*: first, scale the part $\varphi([t_0, t_1])$ using the homothety h with center $\varphi(t_0)$ and ratio equal to 2. *Right*: translate the part $\varphi([t_1, L(\varphi)])$ by the vector $h(\varphi(t_1)) - \varphi(t_1)$ so to regain connectedness. The competitor ψ is made of $\varphi([0, t_0])$ from the original curve, and the red and blue pieces.

where,

$$\xi := |\varphi(t_1) - \varphi(t_0)|.$$

As $|\varphi'| \equiv 1$, \mathcal{L}^1 -a.e., and $\psi(t_0) = \varphi(t_0)$, it follows that

$$|h(\varphi(t_1)) - \varphi(t_1)| = |\varphi(t_1) - \varphi(t_0)| = \xi. \quad (3.5)$$

By construction ψ is parameterized by arc-length, and $L(\psi) = L(\varphi) + \xi$.

Claim 1. $E_\mu^\lambda(\psi) - E_\mu^\lambda(\varphi) \leq (2\vartheta + \lambda)\xi$, with ϑ defined in (1.4).

Since $\psi|_{[t_1+\xi, L(\varphi)+\xi]}$ differs from $\varphi|_{[t_1, L(\varphi)]}$ by a translation of the vector $h(\varphi(t_1)) - \varphi(t_1)$, for any point

$$x \in \mathcal{K}_1 := \{z : d(z, \Gamma_\varphi) = |z - \varphi(s)| \text{ for some } s \in [t_1, L(\varphi)]\}$$

it holds

$$\begin{aligned} d(x, \Gamma_\psi) &\leq d(x, \Gamma_{\psi|_{[t_1+\xi, L(\varphi)+\xi]}}) \leq d(x, \Gamma_{\varphi|_{[t_1, L(\varphi)]}}) + |h(\varphi(t_1)) - \varphi(t_1)| \\ &\stackrel{(3.5)}{\leq} d(x, \Gamma_{\varphi|_{[t_1, L(\varphi)]}}) + \xi. \end{aligned}$$

Integrating with respect to μ gives, in view of (3.1),

$$\begin{aligned} \int_{\mathcal{K}_1} d(x, \Gamma_\psi)^p d\mu - \int_{\mathcal{K}_1} d(x, \Gamma_\varphi)^p d\mu &\leq \int_{\mathcal{K}_1} |d(x, \Gamma_\psi)^p - d(x, \Gamma_\varphi)^p| d\mu \\ &\leq \xi p (D + \xi)^{p-1} \mu(\mathcal{K}_1). \end{aligned} \quad (3.6)$$

Since $\psi|_{[0, t_0]} = \varphi|_{[0, t_0]}$, for any point

$$x \in \mathcal{K}_2 := \{z : d(z, \Gamma_\varphi) = |z - \varphi(s)| \text{ for some } s \in [0, t_0]\}$$

it holds that

$$d(x, \Gamma_\psi) \leq d(x, \Gamma_{\psi|_{[0, t_0]}}) = d(x, \Gamma_{\varphi|_{[0, t_0]}}),$$

thus

$$\int_{\mathcal{K}_2} d(x, \Gamma_\psi)^p d\mu \leq \int_{\mathcal{K}_2} d(x, \Gamma_\varphi)^p d\mu. \quad (3.7)$$

Given that for any point

$$x \in \mathcal{K}_3 := \{z : d(z, \Gamma_\varphi) = |z - \varphi(s)| \text{ for some } s \in [t_0, t_1]\}$$

it holds that

$$d(x, \Gamma_\psi) \leq |x - \psi(t_0)| = |x - \varphi(t_0)| \stackrel{|\varphi'| \equiv 1}{\leq} \stackrel{\mathcal{L}^1\text{-a.e.}}{\leq} d(x, \Gamma_\varphi) + \xi,$$

integrating with respect to μ implies, in the view of (3.1),

$$\int_{\mathcal{K}_3} d(x, \Gamma_\psi)^p d\mu - \int_{\mathcal{K}_3} d(x, \Gamma_\varphi)^p d\mu \leq \xi p (D + \xi)^{p-1} \mu(\mathcal{K}_3). \quad (3.8)$$

Since $\mathbb{R}^d = \mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3$, combining (3.6), (3.7) and (3.8) gives

$$\int_{\mathbb{R}^d} d(x, \Gamma_\psi)^p d\mu - \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu \leq 2p(D + \xi)^{p-1} \xi. \quad (3.9)$$

By construction $L(\psi) = L(\varphi) + \xi$. Since ξ can be chosen arbitrarily small, it follows

$$E_\mu^\lambda(\psi) - E_\mu^\lambda(\varphi) \leq (2\vartheta + \lambda)\xi,$$

proving the claim.

Claim 2. There exists $t_1 > t_0$, with $t_1 - t_0$, and hence ξ , arbitrarily small such that

$$\int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 - \int_0^{L(\varphi)+\varepsilon} |\kappa_\psi|^2 d\mathcal{L}^1 \geq \frac{1}{2} M^2 \xi.$$

Note that $\psi|_{[0, t_0]} = \varphi|_{[0, t_0]}$, while $\psi|_{[t_1+\xi, L(\varphi)+\xi]}$ differs from $\varphi|_{[t_1, L(\varphi)]}$ by a translation. Thus

$$\int_0^{t_0} |\kappa_\psi|^2 d\mathcal{L}^1 = \int_0^{t_0} |\kappa_\varphi|^2 d\mathcal{L}^1, \quad \int_{t_1+\xi}^{L(\varphi)+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 = \int_{t_1}^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1. \quad (3.10)$$

Moreover

$$\begin{aligned} \int_0^{L(\varphi)+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 &= \int_0^{t_0} |\kappa_\psi|^2 d\mathcal{L}^1 + \int_{t_0}^{t_1+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 + \int_{t_1+\xi}^{L(\varphi)+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 \\ &\stackrel{(3.10)}{=} \int_0^{t_0} |\kappa_\varphi|^2 d\mathcal{L}^1 + \int_{t_0}^{t_1+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 + \int_{t_1}^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1, \end{aligned}$$

therefore

$$\int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 - \int_0^{L(\varphi)+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 = \int_{t_0}^{t_1} |\kappa_\varphi|^2 d\mathcal{L}^1 - \int_{t_0}^{t_1+\xi} |\kappa_\psi|^2 d\mathcal{L}^1.$$

By construction, it holds

$$\int_{t_0}^{t_1+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 = \frac{1}{2} \int_{t_0}^{t_1} |\kappa_\varphi|^2 d\mathcal{L}^1.$$

Using Hölder inequality, by the definition of M there exists $t_1 > t_0$, with $t_1 - t_0$ arbitrarily small such that

$$\int_{t_0}^{t_1} |\kappa_\varphi|^2 d\mathcal{L}^1 \geq \frac{|\varphi'(t_1) - \varphi'(t_0)|^2}{|t_1 - t_0|} \geq M^2 \xi,$$

and thus

$$\int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 - \int_0^{L(\varphi)+\xi} |\kappa_\psi|^2 d\mathcal{L}^1 \geq \frac{1}{2} M^2 \xi, \quad (3.11)$$

and the claim is proven.

Combining Claims 1 and 2 with the minimality of φ gives

$$0 \geq \mathcal{E}_\mu^{\lambda, \varepsilon, p}(\varphi) - \mathcal{E}_\mu^{\lambda, \varepsilon, p}(\psi) \geq \frac{1}{2} \varepsilon M^2 \xi - (2\vartheta + \lambda) \xi,$$

hence

$$M \leq \sqrt{\frac{2(2\vartheta + \lambda)}{\varepsilon}},$$

and the proof is complete. \square

Now we investigate the behavior of estimate (1.3) under scaling. Let μ , λ , ε , p be given, and let φ be a minimizer of $\mathcal{E}_\mu^{\lambda, \varepsilon, p}$. Endow \mathbb{R}^d with an orthogonal coordinate system, and consider the linear map

$$T : \mathbb{R}^d \longrightarrow \mathbb{R}^d, \quad T(x) := rx,$$

where $r > 0$ is a given homothety ratio. Set

$$\varphi_r : [0, rL(\varphi)] \longrightarrow \mathbb{R}^d, \quad \varphi_r(t) := r\varphi(t/r).$$

Note that

$$\begin{aligned} \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_r})^p dT_{\#}\mu &= r^{p-d} \int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu, \\ L(\varphi_r) &= rL(\varphi), \\ \int_0^{L(\varphi_r)} |\kappa_{\varphi_r}|^2 d\mathcal{L}^1 &= \frac{1}{r} \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1, \end{aligned}$$

hence

$$\begin{aligned}
r^{p-d}\mathcal{E}_\mu^{\lambda,\varepsilon,p}(\varphi) &= r^{p-d} \left(\int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu + \lambda L(\varphi) + \varepsilon \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 \right) \\
&= \int_{\mathbb{R}^d} d(x, \Gamma_{\varphi_r})^p dT_{\#\mu} + r^{p-d-1} \lambda L(\varphi_r) + r^{p-d+1} \varepsilon \int_0^{L(\varphi_r)} |\kappa_{\varphi_r}|^2 d\mathcal{L}^1 \\
&= \mathcal{E}_{T_{\#\mu}}^{\lambda_r, \varepsilon_r, p}(\varphi_r),
\end{aligned}$$

where $\lambda_r := r^{p-d-1}\lambda$, $\varepsilon_r := r^{p-d+1}\varepsilon$. Since φ is a minimizer for $\mathcal{E}_\mu^{\lambda,\varepsilon,p}$, it follows that φ_r is a minimizer for $\mathcal{E}_{T_{\#\mu}}^{\lambda_r, \varepsilon_r, p}$. Note that $\text{diam supp}(T_{\#\mu}) = r \cdot \text{diam supp}(\mu)$. Theorem 1.4 gives that φ'_r is Y_r -Lipschitz continuous, where

$$Y_r := \sqrt{\frac{2 \left(2p(2r \text{diam supp}(\mu))^{p-1} + \lambda_r \right)}{\varepsilon_r}}.$$

Note that

$$Y_r = \frac{1}{r} \sqrt{\frac{2 \left(2p(2 \text{diam supp}(\mu))^{p-1} + \lambda \right)}{\varepsilon}} = \frac{Y}{r}, \quad (3.12)$$

with Y defined in (1.3). By construction, $\varphi'_r(t) = \varphi'_r(t/r)$ for any t , therefore the respective Lipschitz constants satisfy $Y_r = Y/r$, which is compatible with (3.12).

3.2. Heuristic arguments

We present some heuristic arguments about the sharpness of the estimate (1.3), which suggest that such estimate has optimal order in ε , for small values of ε . Choose λ such that $2\pi\lambda < (2\pi + 1)$, and let $p := 1$. For $r > 0$, let $S_r := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = r^2\}$, and $\mu := \mathcal{H}_{S_r}^1$. We compare the following competitors:

1. let $\psi_r \in \mathcal{C}$ be a parameterization of S_r ,
2. let $\psi_r^* \in \mathcal{C}$ be a parameterization of the segment $\{(1-s)(-r, 0) + s(r, 0) : s \in [0, 1]\}$.

Note that

$$\begin{aligned}
\int_{\mathbb{R}^2} d(x, \Gamma_{\psi_r}) d\mu &= 0, & L(\psi_r) &= 2\pi r, & \int_0^{L(\psi_r)} |\kappa_{\psi_r}|^2 d\mathcal{L}^1 &= \frac{2\pi}{r}, \\
\int_{\mathbb{R}^2} d(x, \Gamma_{\psi_r^*}) d\mu &= r, & L(\psi_r^*) &= 2r, & \int_0^{L(\psi_r^*)} |\kappa_{\psi_r^*}|^2 d\mathcal{L}^1 &= 0.
\end{aligned}$$

Thus it follows

$$\mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r) = 2\pi\lambda r + \frac{2\pi\varepsilon}{r}, \quad \mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r^*) = r(1 + 2\lambda).$$

Note that for large r , ψ_r is more convenient than ψ_r^* , since $2\pi\lambda r$ dominates $2\pi\varepsilon/r$, and hypothesis $2\pi\lambda < (2\pi + 1)$ implies $\mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r) < \mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r^*)$. For $r \rightarrow 0^+$, the term $2\pi\varepsilon/r$ diverges. For $r \simeq \sqrt{\varepsilon}$ (i.e. r and $\sqrt{\varepsilon}$ differ by a

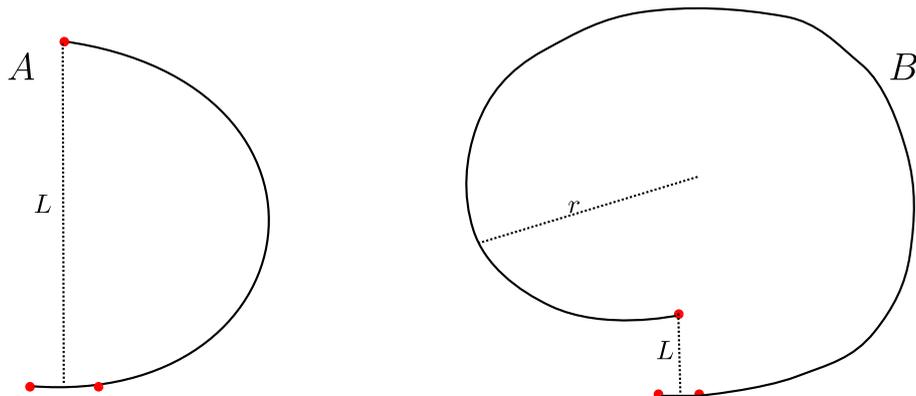


FIGURE 3. A schematic representation of two possible choices of minimizers. Note the different scale in the two figures. The masses are concentrated on red points.

multiplicative constant), it holds

$$\mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r) \simeq 2\pi(1+\lambda)\sqrt{\varepsilon}, \quad \mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r^*) \simeq \sqrt{\varepsilon}(1+2\lambda),$$

hence $\mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r)$ and $\mathcal{E}_\mu^{\lambda,\varepsilon,1}(\psi_r^*)$ have the same order in ε . Thus the maximum radius r_0 such that for $r < r_0$, ψ_r^* is more advantageous than ψ_r , has order $r_0 \simeq \sqrt{\varepsilon}$. The derivative ψ'_{r_0} is approximately (upon constants independent of ε) $\sqrt{1/\varepsilon}$ -Lipschitz continuous.

We present some heuristic arguments about the choice of parameters λ, ε , under given $p \geq 1$. Let μ be a measure concentrated on the red points in Figure 3, and let φ be the minimizer (denoted by black solid lines in Fig. 3). In the following arguments the symbol “ \simeq ” will denote “equal upon universal constants”. For the sake of simplicity we omit any non influential constant.

Configuration A: $\lambda L(\varphi) \simeq \lambda L$, $\varepsilon \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 \simeq \varepsilon/L$, $\int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu$ is infinitesimal compared to the above quantities, since $\Sigma \supseteq \text{supp}(\mu)$. Total energy of configuration A: approximately $\lambda L + \varepsilon/L$.

Configuration B: $r \geq L$, $\lambda L(\varphi) \simeq \lambda r$, $\varepsilon \int_0^{L(\varphi)} |\kappa_\varphi|^2 d\mathcal{L}^1 \simeq \varepsilon/r$, $\int_{\mathbb{R}^d} d(x, \Gamma_\varphi)^p d\mu$ is infinitesimal compared to the above quantities, since $\Sigma \supseteq \text{supp}(\mu)$. Total energy of configuration B: approximately $\lambda r + \varepsilon/r$. Direct computing gives the optimal value $r = \sqrt{\varepsilon/\lambda}$, corresponding to an approximate value of $\sqrt{\lambda\varepsilon}$ for the energy.

Note that configuration B is “less desirable” in data parameterization, since the minimizer contains points further away from $\text{supp}(\mu)$. Moreover, a necessary condition for configuration B to be preferable to configuration A is $\sqrt{\varepsilon/\lambda} > L$, i.e. $\varepsilon > L^2\lambda$. Thus, for $\varepsilon < L^2\lambda$, configuration A is preferable.

3.3. Consequences of Theorem 1.4

Recall that Problem 1.3 has been introduced to overcome the fact that minimizers of Problem 1.2 can fail to be C^1 regular, which is undesirable for data parameterization. Corollaries 3.3 and 3.4 prove that if $\mu \ll \mathcal{L}^d$, then minimizers of Problem 1.3 do not exhibit such undesirable properties. Some preliminary discussion is required. For given $\Sigma \in \mathcal{A}$ (defined in Problem 1.1), define the set-valued “projection” map as

$$\Pi_\Sigma : \mathbb{R}^d \longrightarrow \mathcal{P}(\Sigma), \quad \Pi(x) = \{y \in \Sigma : |x - y| = d(x, \Sigma)\}, \quad (3.13)$$

with $\mathcal{P}(\Sigma)$ denoting the power set of Σ . We recall that in [30] it has been proven that the ridge

$$\mathcal{R}_\Sigma := \{x : \exists y, z \in \Sigma, y \neq z, |x - y| = |x - z| = d(x, \Sigma)\} = \{x : \#\Pi_\Sigma(x) \geq 2\}$$

is $\mathcal{H}^{d-1,1}$ rectifiable. In particular, if $\mu \ll \mathcal{L}^d$, for any $\Sigma \in \mathcal{A}$ the ridge \mathcal{R}_Σ is μ -negligible, thus $\#\Pi_\Sigma(x) = 1$ μ -a.e. In this case it is possible to define, for μ -a.e. point, the point-valued function

$$\pi_\Sigma : \mathbb{R}^d \longrightarrow \Sigma, \quad \pi_\Sigma(x) := y, \quad y \text{ is the unique point of } \Sigma \text{ satisfying } \{y\} = \Pi_\Sigma(x). \quad (3.14)$$

Definition 3.2. Given $\mu \ll \mathcal{L}^d$, $\Sigma \in \mathcal{A}$ and a subset $B \subseteq \Sigma$, the quantity $\mu(\pi_\Sigma^{-1}(B))$ will be referred to as mass projecting on B in Σ .

For the sake of brevity we will omit writing “in Σ ” if no risk of confusion arises. Note that, for any curve $\psi \in \mathcal{C}$, its image Γ_ψ belongs to \mathcal{A} .

Corollary 3.3. Given $d \geq 2$, a measure $\mu \ll \mathcal{L}^d$, parameters $\lambda, \varepsilon > 0$, $p \geq 1$, a minimizer $\varphi \in \operatorname{argmin}_{\mathcal{C}} \mathcal{E}_\mu^{\lambda, \varepsilon, p}$ and a time $t \in (0, L(\varphi))$, it holds $\mu(\pi_{\Gamma_\varphi}^{-1}(\{\varphi(t)\})) = 0$, i.e. the mass projecting on $\{\varphi(t)\}$ is zero.

Proof. Note that for sufficiently small η , the behavior of the curve $\varphi|_{[t-\eta, t+\eta]}$ is locally approximated (in first order, upon translation and scaling) by the tangent derivative $\varphi'(t)$. Thus the set

$$\{x \in \mathbb{R}^d : d(x, \Gamma_\varphi) = |x - \varphi(t)|\} = \pi_{\Gamma_\varphi}^{-1}(\{\varphi(t)\})$$

is contained in Δ_t , defined as the (unique) $(d-1)$ -hyperplane passing through $\varphi(t)$ and orthogonal to $\varphi'(t)$. Since $\mu \ll \mathcal{L}^d$, it follows $\mu(\Delta_t) = 0$, concluding the proof. \square

Corollary 3.4. Given $d \geq 2$, $q \in [1, +\infty]$, a measure $\mu \ll \mathcal{L}^d$ with Radon-Nikodym derivative $d\mu/d\mathcal{L}^d \in L^q$, parameters $\lambda, \varepsilon > 0$, $p \geq 1$, a minimizer $\varphi \in \operatorname{argmin}_{\mathcal{C}} \mathcal{E}_\mu^{\lambda, \varepsilon, p}$ and a time interval $[t_0, t_1] \subseteq (0, L(\varphi))$, then the mass projecting on $\varphi([t_0, t_1])$ satisfies

$$\mu(\pi_{\Gamma_\varphi}^{-1}(\varphi([t_0, t_1]))) \leq \left\| \frac{d\mu}{d\mathcal{L}^d} \right\|_{L^q} \left(|t_1 - t_0| \omega_{d-1} \left(\frac{D^d}{d} + \frac{D^{d+1}}{d+1} Y \right) \right)^{1-1/q}, \quad (3.15)$$

where

$$Y = Y(\operatorname{diam} \operatorname{supp}(\mu), \lambda, p, \varepsilon) := \sqrt{\frac{2(2\vartheta + \lambda)}{\varepsilon}},$$

$\vartheta := pD^{p-1}$, $D := 2\operatorname{diam} \operatorname{supp}(\mu)$, ω_{d-1} denotes the \mathcal{L}^{d-1} measure of the $d-1$ dimensional unit ball, and $1/q := 0$ if $q = +\infty$.

Proof. Let $\xi := |t_1 - t_0|$. For any t , the set $\pi_{\Gamma_\varphi}^{-1}(\{\varphi(t)\})$ is contained in Δ_t , defined as the $(d-1)$ -hyperplane passing through $\varphi(t)$ and orthogonal to $\varphi'(t)$. Thus it holds

$$\pi_{\Gamma_\varphi}^{-1}(\varphi([t_0, t_1])) \subseteq \bigcup_{t \in [t_0, t_1]} \Delta_t,$$

and it suffices to estimate the μ -measure of the right-hand side term.

Endow \mathbb{R}^d with the standard orthonormal base $\{e_j\}_{j=1}^d$, $e_j := (0, \dots, 0, \underbrace{1}_{j\text{-th position}}, 0, \dots, 0)$. For $t \in [t_0, t_1]$,

define the linear map $X_t : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ as follows:

1. if $e_1 = \varphi'(t)$, then X_t is the identity map,
2. if $e_1 = -\varphi'(t)$, then $X_t(p) := -p$ for any $p \in \mathbb{R}^d$,
3. if $\angle_{e_1} \varphi'(t) \in (0, \pi)$, let X_t be the unique rotation by an angle $\angle_{e_1} \varphi'(t) \in (0, \pi)$, mapping e_1 in $\varphi'(t)$.

Note that for any t , $X_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a conformal isometry. Since Theorem 1.4 proves $C^{1,1}$ regularity for φ , and Y -Lipschitz continuity for φ' , the mapping $t \mapsto X_t e_j$ is C^1 regular for any j . Moreover, it holds

$$\begin{aligned} (\forall t^*)(\forall j) \quad \left| \frac{d}{dt} X_t e_j \Big|_{t=t^*} \right| &\leq \lim_{h \rightarrow 0} \frac{1}{h} |X_{t^*+h} e_j - X_{t^*} e_j| \\ &= \lim_{h \rightarrow 0} \frac{1}{h} |X_{t^*+h} e_1 - X_{t^*} e_1| \\ &= \lim_{h \rightarrow 0} \frac{1}{h} |\varphi'(t^*+h) - \varphi'(t^*)| \leq Y. \end{aligned} \quad (3.16)$$

Consider the moving frame $\{X_t e_1, \dots, X_t e_d\}$. By construction, $\Delta_t = \text{Span}(X_t e_2, \dots, X_t e_d) + \varphi(t)$.

Let $P = \varphi(t_0) + \sum_{j=2}^d P_j X_{t_0} e_j \in \Delta_{t_0}$ be an arbitrarily given point, and define the trajectory

$$\gamma_P : [t_0, t_1] \rightarrow \mathbb{R}^d, \quad \gamma_P(t) := \varphi(t) + \sum_{j=2}^d P_j X_t e_j.$$

Since $t \mapsto X_t e_j$ is C^1 regular for any j , γ_P is C^1 regular for any $P \in \Delta_{t_0}$. Note that, in general, γ_P is not parameterized by arc-length. By construction, it holds

$$(\forall t) \quad \Delta_t = \bigcup_{(P_2, \dots, P_d) \in \mathbb{R}^{d-1}} \left(\varphi(t) + \sum_{j=2}^d P_j X_t e_j \right),$$

and the union of all trajectories covers $\bigcup_{t \in [t_0, t_1]} \Delta_t$, *i.e.*

$$\bigcup_{P \in \Delta_{t_0}} \gamma_P([t_0, t_1]) = \bigcup_{t \in [t_0, t_1]} \Delta_t.$$

Lemma 2.5 gives $\Gamma_\varphi \subseteq (\text{supp}(\mu))_{\text{diam supp}(\mu) + (\text{diam supp}(\mu))^p / \lambda}$, thus

$$\sup_{x \in \text{supp}(\mu), y \in \Gamma_\varphi} |x - y| \leq 2 \text{diam supp}(\mu) + (\text{diam supp}(\mu))^p / \lambda = D.$$

Since X_t is an isometry, it follows

$$\text{supp}(\mu) \cap \bigcup_{P \in \Delta_{t_0}} \gamma_P([t_0, t_1]) = \text{supp}(\mu) \cap \bigcup_{t \in [t_0, t_1]} \Delta_t \subseteq \bigcup_{\substack{P \in \Delta_{t_0}, \\ |P - \varphi(t_0)| \leq D}} \gamma_P([t_0, t_1]). \quad (3.17)$$

To estimate the length of the trajectory $t \mapsto \gamma_P(t)$, direct computation gives

$$\begin{aligned} \int_{t_0}^{t_1} |\gamma'_P(t)| dt &= \int_{t_0}^{t_1} \left| \varphi'(t) + \sum_{j=2}^d P_j \frac{d}{ds} X_s e_j \Big|_{s=t} \right| dt \leq \int_{t_0}^{t_1} |\varphi'(t)| dt + \int_{t_0}^{t_1} \left| \sum_{j=2}^d P_j \frac{d}{ds} X_s e_j \Big|_{s=t} \right| dt \\ &= \xi + \int_{t_0}^{t_1} \sqrt{\sum_{j=2}^d P_j^2 \left| \frac{d}{ds} X_s \Big|_{s=t} \right|^2} dt \stackrel{(3.16)}{\leq} \xi + \int_{t_0}^{t_1} \sqrt{Y^2 \sum_{j=2}^d P_j^2} dt \end{aligned}$$

$$= \xi + \xi|P - \varphi(t_0)|Y. \quad (3.18)$$

This proves that the length of trajectory $t \mapsto \gamma_P(t)$ depends only on $|P - \varphi(t_0)|$. Thus

$$\begin{aligned} \mathcal{L}^d \left(\text{supp}(\mu) \cap \bigcup_{t \in [t_0, t_1]} \Delta_t \right) &\stackrel{(3.17)}{\leq} \mathcal{L}^d \left(\bigcup_{P \in \Delta_{t_0}, |P - \varphi(t_0)| \leq D} \gamma_P([t_0, t_1]) \right) \\ &= \int_{\{P \in \Delta_{t_0} : |P - \varphi(t_0)| \leq D\}} \int_{t_0}^{t_1} |\gamma'_P(t)| dt dP \\ &\stackrel{(3.18)}{\leq} \xi \int_{\{v \in \Delta_{t_0} : |P - \varphi(t_0)| \leq D\}} (1 + |P - \varphi(t_0)|Y) dP \\ &= \xi \int_0^D \int_{\{P \in \Delta_{t_0} : |P - \varphi(t_0)| = r\}} (1 + rY) dP dr \\ &= \xi \int_0^D \omega_{d-1} r^{d-1} (1 + rY) dr \\ &= \xi \omega_{d-1} \left(\frac{D^d}{d} + \frac{D^{d+1}}{d+1} Y \right). \end{aligned}$$

Since by hypothesis $d\mu/d\mathcal{L}^d \in L^q$, Hölder inequality gives

$$\begin{aligned} \mu \left(\text{supp}(\mu) \cap \bigcup_{t \in [t_0, t_1]} \Delta_t \right) &\leq \left\| \frac{d\mu}{d\mathcal{L}^d} \right\|_{L^q} \left(\mathcal{L}^d \left(\text{supp}(\mu) \cap \bigcup_{t \in [t_0, t_1]} \Delta_t \right) \right)^{1-1/q} \\ &\leq \left\| \frac{d\mu}{d\mathcal{L}^d} \right\|_{L^q} \left(\xi \omega_{d-1} \left(\frac{D^d}{d} + \frac{D^{d+1}}{d+1} Y \right) \right)^{1-1/q}, \end{aligned}$$

concluding the proof. \square

APPENDIX A. AN ODE FOR CURVATURE

Here we derive the Euler-Lagrange equation for the energy \mathcal{E} , when the reference measure μ has finitely many atoms. Let $\gamma : [0, L] \rightarrow \mathbb{R}^d$ be an arc-length parameterized minimizer of $\mathcal{E}_\mu^{\lambda, \varepsilon, p}$. Let $\varphi : [0, L] \rightarrow \mathbb{R}^d$ be a C^2 regular perturbation. We show that, in \mathbb{R}^2 , the Euler-Lagrange equation is up to parameters, the same as the Euler-Lagrange equation for elastica functional [26]:

$$\kappa'' + C_1 \kappa^3 + C_2 \kappa = 0.$$

This is expected, as minimizers of \mathcal{E} are “elastica-like”.

We recall that elastica curves are solutions of the so-called “elastica problem”, one of the earliest examples of nonlinear displacement problems, first proposed by Bernoulli in 1691. The integrated squared curvature quantifies the elastic stress, while the average-distance term describes the “pulling force” of the weight hanging on the elastica.

We first estimate the first variation of the curvature term $K(\gamma) = \int_0^L |\gamma''|^2 ds$. Given $h > 0$,

$$K(\gamma + h\varphi) - K(\gamma) = \int_0^L |\gamma' + h\varphi'|^{-1} \left| \frac{d}{ds} \left(\frac{\gamma' + h\varphi'}{|\gamma' + h\varphi'|} \right) \right|^2 ds.$$

Direct computation gives

$$\begin{aligned} \left. \frac{d}{dh} K(\gamma + h\varphi) \right|_{h=0} &= \int_0^L \left[-\langle \gamma', \varphi' \rangle |\gamma''|^2 + 2\langle \gamma'', \varphi'' - \gamma'' \langle \gamma', \varphi' \rangle - \langle \varphi', \gamma'' \rangle \gamma' - \langle \varphi'', \gamma' \rangle \gamma' \rangle \right] ds \\ &= \int_0^L [3(|\gamma''|^2 \gamma')' + 2\gamma'''''] \varphi ds + 2\gamma'' \varphi'' \Big|_0^L, \end{aligned} \quad (\text{A.1})$$

Now we estimate the first variation of the length:

$$\lambda \int_0^L [|\gamma' + h\varphi'| - |\gamma'|] dx = \lambda \int_0^L \frac{h\gamma'\varphi' + O(h^2)}{|\gamma' + h\varphi'| + |\gamma'|} dx.$$

Without loss of generality, we take γ to be arc-length parameterized, *i.e.* $|\gamma'| = 1$ a.e. Dividing by h and then taking the limit $h \rightarrow 0$ gives

$$\lambda \int_0^L \frac{\gamma'\varphi'}{2|\gamma'|} dx = \frac{\lambda}{2} \int_0^L \gamma'' \varphi dx + \frac{\lambda}{2} \gamma'' \varphi \Big|_0^L.$$

Moreover, recall that all the masses projected on the knots, hence the average distance term remains unchanged. Combining with (A.1) and considering perturbations φ which are compactly supported within $(0, L)$ gives

$$\int_0^L \left[\frac{\lambda}{2} \gamma'' + \varepsilon [3(|\gamma''|^2 \gamma')' + 2\gamma'''''] \right] \varphi dx = 0,$$

i.e.,

$$\frac{\lambda}{2} \gamma'' + \varepsilon [3(|\gamma''|^2 \gamma')' + 2\gamma'''''] = \frac{\lambda}{2} \gamma'' + \varepsilon [3(2\langle \gamma'', \gamma'' \rangle \gamma' + |\gamma''|^2 \gamma'') + 2\gamma'''''] = 0. \quad (\text{A.2})$$

In \mathbb{R}^2 , using Frenet-Serret formulas, we have

$$\gamma' = \vec{t}, \quad \gamma'' = \kappa \vec{n}, \quad (\text{A.3})$$

$$\gamma''' = \kappa' \vec{n} - \kappa^2 \vec{t}, \quad \gamma'''' = \kappa'' \vec{n} - 3\kappa \kappa' \vec{t} - \kappa^3 \vec{n}, \quad (\text{A.4})$$

hence (A.2) becomes

$$0 = \frac{\lambda}{2} \kappa \vec{n} + \varepsilon \left[6\langle \kappa \vec{n}, \kappa' \vec{n} - \kappa^2 \vec{t} \rangle \vec{t} + 2(\kappa'' \vec{n} - 3\kappa \kappa' \vec{t} - \kappa^3 \vec{n}) \right] = \left(\frac{\lambda}{2} \kappa + 2\varepsilon [\kappa'' - \kappa^3] \right) \vec{n},$$

i.e.

$$\frac{\lambda}{2} \kappa + 2\varepsilon [\kappa'' - \kappa^3] = 0.$$

Finally, note the segment $\gamma((t_0, t_1))$ minimizes the (sum of) length and curvature terms, *i.e.*

$$\begin{aligned} \gamma|_{(t_0, t_1)} \in \operatorname{argmin} \left\{ \lambda L(\sigma) + \varepsilon \int_0^{L(\sigma)} |\kappa_\sigma|^2 ds : \right. \\ \left. \sigma \text{ is an arc-length parameterized curve with endpoints } \gamma(t_0) \text{ and } \gamma(t_1) \right\}. \end{aligned}$$

Acknowledgements. The authors are grateful to the referees for valuable suggestions. XYL acknowledges the support of NSERC Discovery Grant “Regularity of minimizers and pattern formation in geometric minimization problems”. This research was partly carried out when XYL was affiliated with Instituto Superior Técnico. DS is grateful to the National Science Foundation for support under grants DMS 1211161, CIF 1421502 and DMS 1516677. Both authors are thankful to the Center for Nonlinear Analysis (CNA) for its support.

REFERENCES

- [1] J. Bertoulli, Curvatura laminae elasticae. Ejus identitas cum curvatura lintei a pondere inclusi fluidi expansi. Radii circulorum osculantium in terminis simplicissimis exhibit; una cum novis quibusdarn theorematis huc pertinentibus. *Acta Eruditorum* (1694).
- [2] G. Biau and A. Fischer, Parameter selection for principal curves. *IEEE Trans. Inf. Theory* **58** (2011) 1924–1939.
- [3] E. Bretin, J.-O. Lachaud and É. Oudet, Regularization of discrete contour by Willmore energy. *J. Math. Imag. Vision* **40** (2011) 214–229.
- [4] G. Buttazzo, E. Mainini and E. Stepanov, Stationary configurations for the average distance functional and related problems. *Control Cybernet.* **38** (2009) 1107–1130.
- [5] G. Buttazzo, E. Oudet and E. Stepanov, Optimal transportation problems with free dirichlet regions, in Variational methods for discontinuous structures. Vol. 51 of *Progr. Nonlinear Differential Equations Appl.* Birkhäuser, Basel (2002) 41–65.
- [6] G. Buttazzo, A. Pratelli, S. Solimini and E. Stepanov, Optimal urban networks via mass transportation. Vol. 1961 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin (2009).
- [7] G. Buttazzo, A. Pratelli and E. Stepanov, Optimal pricing policies for public transportation networks. *SIAM J. Optim.* **16** (2006) 826–853.
- [8] G. Buttazzo and F. Santambrogio, A model for the optimal planning of an urban area. *SIAM J. Math. Anal.* **37** (2005) 514–530 (electronic).
- [9] G. Buttazzo and F. Santambrogio, A mass transportation model for the optimal planning of an urban region. *SIAM Rev.* **51** (2009) 593–610.
- [10] G. Buttazzo and E. Stepanov, Optimal transportation networks as free dirichlet regions for the Monge-Kantorovich problem. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2** (2003) 631–678.
- [11] G. Buttazzo and E. Stepanov, Minimization problems for average distance functionals, in *topics from the mathematical heritage of E. de Giorgi. Vol. 14 of Quad. Mat.*, Dept. Math. Seconda Univ. Napoli, Caserta. *Cal. Var. Geom. Measure Theor.* **14** (2004) 47–83.
- [12] C. de Boor, A practical guide to splines. Vol. 27 of *Applied Mathematical Sciences*. Springer-Verlag, New York, revised ed. (2001).
- [13] S. Delattre and A. Fischer, On principal curves with a length constraint. *Ann. Inst. Henri Poincaré Probab. Statist.* **56** (2020) 2108–2140.
- [14] P. Delicado, Another look at principal curves and surfaces. *J. Multivariate Anal.* **77** (2001) 84–116.
- [15] P.W. Dondl, L. Mugnai and M. Röger, A phase field model for the optimization of the Willmore energy in the class of connected surfaces. *SIAM J. Math. Anal.* **46** (2014) 1610–1632.
- [16] Q. Du, C. Liu, R. Ryham and X. Wang, A phase field formulation of the Willmore problem. *Nonlinearity* **18** (2005) 1249–1267.
- [17] T. Duchamp and W. Stuetzle, Geometric properties of principal curves in the plane, in Robust statistics, data analysis, and computer intensive methods (Schloss Thurnau, 1994). Vol. 109 of *Lecture Notes in Statist.* Springer, New York (1996) 135–152.
- [18] L. Euler, Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes sive solutio problematis isoperimetrici latissimo sensu accepti. Lausannæ, Genevæ, apud Marcum-Michaelem Bousquet & socios (1744).
- [19] S. Gerber, T. Tasdizen and R. Whitaker, Dimensionality reduction and principal surfaces via kernel map manifolds, in *2009 IEEE 12th International Conference on Computer Vision. IEEE* (2009) 529–536.
- [20] S. Gerber and R. Whitaker, Regularization-free principal curve estimation. *J. Machine Learning Res.* **14** (2013) 1285–1302.
- [21] V.G.A. Goss, *Snap buckling, writhing and loop formation in twisted rods*. Ph.D. thesis, University College London (2003).
- [22] T. Hastie and W. Stuetzle, Principal curves. *J. Am. Statist. Assoc.* **84** (1989) 502–516.
- [23] B. Kégl, A. Krzyzak, T. Linder and K. Zeger, Learning and design of principal curves. *IEEE Trans. Pattern Anal. Mach. Intell.* **22** (2000) 281–297.
- [24] S. Kirov and D. Slepčev, Multiple penalized principal curves: analysis and computation. *J. Math. Imaging Vision* **59** (2017) 234–256.
- [25] A. Lemenant, About the regularity of average distance minimizers in \mathbb{R}^2 . *J. Convex Anal.* **18** (2011) 949–981.
- [26] R. Levien, *The elastica: a mathematical history* (2008).
- [27] X. Y. Lu, Example of minimizer of the average-distance problem with non closed set of corners. *Rendiconti del Seminario Matematico della Università di Padova* **137** (2017) 19–55.
- [28] X.Y. Lu and D. Slepčev, Properties of minimizers of average-distance problem via discrete approximation of measures. *SIAM J. Math. Anal.* **45** (2013) 3114–3131.
- [29] X.Y. Lu and D. Slepčev, Average-distance problem for parameterized curves. *ESAIM: COCV* **22** (2016) 404–416.
- [30] C. Mantegazza and A.C. Mennucci, Hamilton-Jacobi equations and distance functions on Riemannian manifolds. *Appl. Math. Optim.* **47** (2003) 1–25.

- [31] U. Ozertem and D. Erdogmus, Locally defined principal curves and surfaces. *J. Mach. Learn. Res.* **12** (2011) 1249–1286.
- [32] E. Paolini and E. Stepanov, Qualitative properties of maximum distance minimizers and average distance minimizers in \mathbb{R}^n . *J. Math. Sci. (N. Y.)* **122** (2004) 3290–3309.
- [33] P. Polak and G. Wolansky, The lazy travelling salesman problem in \mathbb{R}^2 . *ESAIM: COCV* **13** (2007) 538–552.
- [34] D. Slepčev, Counterexample to regularity in average-distance problem. *Ann. Inst. Henri Poincaré Anal. Non Linéaire* **31** (2014) 169–184.
- [35] A.J. Smola, S. Mika, B. Schölkopf and R.C. Williamson, Regularized principal manifolds. *J. Mach. Learn. Res.* **1** (2001) 179–209.
- [36] R. Tibshirani, Principal curves revisited. *Stat. Comput.* **2** (1992) 182–190.