Counterexample to regularity in average-distance problem

Dejan Slepcev
Carnegie Mellon University, slepcev@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/math
Part of the Mathematics Commons

Published In
Annales de l'Institut Henri Poincare (C) Non Linear Analysis, 31, 1, 169-184.
ABSTRACT. The average-distance problem is to find the best way to approximate (or represent) a given measure \( \mu \) on \( \mathbb{R}^d \) by a one-dimensional object. In the penalized form the problem can be stated as follows: given a finite, compactly supported, positive Borel measure \( \mu \), minimize

\[
E(\Sigma) = \int_{\mathbb{R}^d} d(x, \Sigma) d\mu(x) + \lambda \mathcal{H}^1(\Sigma)
\]

over the set of connected closed sets, \( \Sigma \), where \( \lambda > 0 \), \( d(x, \Sigma) \) is the distance from \( x \) to the set \( \Sigma \), and \( \mathcal{H}^1 \) is the one-dimensional Hausdorff measure. Here we provide, for any \( d \geq 2 \), an example of a measure \( \mu \) with smooth density, and convex, compact support, such that the global minimizer of the functional is a rectifiable curve which is not \( C^1 \). We also provide a similar example for the constrained form of the average-distance problem.

1. INTRODUCTION

Given a positive, compactly supported, Borel measure \( \mu \) on \( \mathbb{R}^d \), \( d \geq 2 \), \( \lambda > 0 \), and \( \Sigma \) a nonempty subset of \( \mathbb{R}^d \) consider

\[
E(\Sigma) = \int_{\mathbb{R}^d} d(x, \Sigma) d\mu(x) + \lambda \mathcal{H}^1(\Sigma)
\]

The average-distance problem is to minimize the functional over \( \mathcal{A} = \{ \Sigma \subset \mathbb{R}^d : \Sigma \text{ connected and compact} \} \).

The problem was introduced by Buttazzo, Oudet, and Stepanov [3] and Buttazzo and Stepanov [4]. They studied the problem in the constrained form, where instead of \( \mathcal{H}^1 \) penalization one minimizes

\[
F(\Sigma) = \int_{\mathbb{R}^d} d(x, \Sigma) d\mu(x) \quad \text{over } \mathcal{A}_1 := \{ \Sigma \in \mathcal{A} : \mathcal{H}^1(\Sigma) \leq \ell \}.
\]

Over the past few years there has been a significant progress on understanding of the functional; some of which we outline below. An excellent overview article has recently been written by Lemenant [8].

The problem has wide ranging applications. When interpreted as a simplified description of designing the optimal public transportation network then \( \mu \) represents the distribution of passengers, and \( \Sigma \) is the network. The desire is to design the
network that minimizes the total distance of passengers to the network. Another related problem which can be reduced to the average-distance problem, studied in [3], is when we think of passengers as workers that need to get to their workplace. Then two measures are initially given, the distribution of where workers reside and where they work. Again the goal is to find the optimal network that minimizes the total transportation cost (traveling along the network is for free).

A related interpretation is that of finding the optimal irrigation network (the irrigation problem).

Another interpretation, whose application in a related setting is presently investigated by Laurent and the author, is to find a good one dimensional representation to a data cloud. Here $\mu$ represents the distribution of data points. One wishes to approximate the cloud by a one-dimensional object. The first term in (1) then charges the errors in the approximation, while the second one penalizes the complexity of the representation.

The existence of minimizers of $E$ follows from the theorems of Blaschke and Gołąb [4]. In this paper we investigate their regularity. It was shown in [4] that, at least for $d = 2$, the minimizer is topologically a tree made of finitely many simple rectifiable curves which meet at triple junctions (no more that three branches can meet at one point). The authors also show that the minimizer is Ahlfors regular (which was extended to higher dimensions by Paolini and Stepanov, [10]), but further regularity of branches remained open. Recently Tilli [12] showed that every compact simple $C^{1,1}$ curve is a minimizer of the average-distance problem (in the constrained form (27)) where $\mu$ is the characteristic function of a small tubular neighborhood of the curve. This suggests that $C^{1,1}$ is the best regularity for minimizers one can expect (even if $\mu$ were smooth). Further criteria for regularity were established by Lemenant [7].

Due to the presence of the $H^1$ term one might expect that, if $\mu$ is a measure with smooth density, $\Sigma$ is at least $C^1$. A recent paper by Buttazzo, Mainini, and Stepanov [2] suggests that this may not be the case, and exhibits a measure $\mu$ which is a characteristic function of a set in $\mathbb{R}^2$, for which there exists a stationary point of $E$ which has a corner. Furthermore the results on the blow-up of the problem by Santambrogio and Tilli [11] support the possibility of corners. Here we prove that minimizers which are not $C^1$ are indeed possible. That is for any $d \geq 2$ provide an example of a measure $\mu$ with smooth density for which we prove that the minimizer is a curve which has a corner, and is thus not $C^1$. One of the difficulties in dealing with global energy minimizers is that the functional is not convex. To be able to treat them we introduce constructions and an approximation technique that may be of independent interest.

Our approach is based on approximating the measure $\mu$ of our interest by particle measures $\mu_n$ (i.e. the ones that have only atoms). For particle measures $\mu_n$ the average-distance problem (1) has a discrete formulation that can be carefully analyzed. In particular the minimizers are trees with piecewise linear branches. Our
starting point is the construction of a particle measure with three particles, \( \bar{\mu} \), for which we can show that the minimizer is a wedge (curve with exactly two line segments), see Figure 1. We then show that if \( \bar{\mu} \) is smoothed out a bit then the minimizer will still have a corner (even if we also add a smooth background measure of small total mass, \( q \), that makes the support of the perturbed measure convex). We denote the smooth perturbed measure by \( \mu_{q,\delta} \) were \( \delta \) is the smoothing parameter. To show that a minimizer of \( E \) for \( \mu_{q,\delta} \) has a corner when \( \delta \) and \( q \) are small, we consider discrete approximations \( \mu_{q,\delta,n} \) of \( \mu_{q,\delta} \). We show that the minimizers \( \Sigma_{q,\delta,n} \) of \( E \) corresponding to \( \mu_{q,\delta,n} \) have a corner whose opening is bounded from above independent of \( n \). We furthermore obtain appropriate estimates on the minimizers which guarantee convergence as \( n \to \infty \) to a minimizer \( \Sigma_{q,\delta} \) of \( E \) corresponding to \( \mu_{q,\delta} \) and insure that the corner remains in the limit.

1.1. Outline. In Section 2 we list some of the basic properties of the functional \( E \) given in (1), in particular its continuity properties with respect to parameters and scaling with respect to dilation of \( \mu \). In Section 3 we consider the energy (1) with \( \mu \) being a particle measure. We obtain conditions for criticality, information of the projection of the measure \( \mu \) onto the minimizer \( \Sigma \), and appriory estimates on the curvature (turning angle). The basic three-particle configuration, \( \bar{\mu} \), for which the minimizer is a wedge is also introduced. The construction of the counterexample is carried out in Section 4. We introduce the perturbation and use elementary geometry to obtain various geometric facts about the minimizer of the average-distance problem corresponding to the discrete approximation of the perturbed measure: \( \mu_{q,\delta,n} \). The result of these efforts is that the minimizer must have a corner with a large turning angle (jump in the tangent direction). Finally we take the limit \( n \to \infty \) to obtain that the minimizer for \( \mu_{q,\delta} \) has a corner too. In Section 5 we use a scaling argument to show that the minimizers of the constrained problem (2) can have corners too.

2. Properties of the functional

Let \( \mathcal{A} \) be the set of compact connected subsets of \( \mathbb{R}^d \). Given \( \Sigma \in \mathcal{A} \), for \( y \in \Sigma \) we define the region of influence of \( y \),

\[
R(y) = \{ x \in \mathbb{R}^d : (\forall z \in \Sigma) \ d(x, z) \geq d(x, y) \}.
\]

In the next two lemmas we study continuity properties of \( E \) with respect to dependence on \( \Sigma \) and \( \mu \).

Lemma 1. For any \( \mu \in \mathcal{P} \) and any \( \lambda > 0 \), the functional \( E_\mu : \mathcal{A} \to \mathbb{R} \) is lower semicontinuous with respect to Hausdorff convergence.

Proof. Assume that \( \Sigma_n \in \mathcal{A} \) converge to \( \Sigma \) in Hausdorff metric, \( d_H \). Gołąb’s theorem (see [1]) gives the lower semicontinuity of the \( H^1 \) measure. Thus it is enough to prove the continuity of the first term of the energy. Note that for any \( x \in \mathbb{R}^d \)

\[
|d(x, \Sigma_n) - d(x, \Sigma)| \leq d_H(\Sigma_n, \Sigma).
\]
To illustrate why, assume that for some \( x \)
\[
d(x, \Sigma_n) > d(x, \Sigma) + d_H(\Sigma_n, \Sigma).
\]
Considering \( y \) to be the closest point to \( x \) on \( \Sigma \) gives
\[
d(x, \Sigma_n) > d(x, y) + \inf_{z \in \Sigma_n} d(y, z) \geq d(x, \Sigma_n)
\]
which is a contradiction. Thus
\[
\left| \int_{\mathbb{R}^d} d(x, \Sigma_n) d\mu(x) - \int_{\mathbb{R}^d} d(x, \Sigma) d\mu(x) \right| \leq d_H(\Sigma_n, \Sigma) \mu(\mathbb{R}^d)
\]
which implies the claim. \( \square \)

**Lemma 2.** Consider \( \Sigma \in \mathcal{A} \) and \( \lambda > 0 \). The mapping \( \mu \mapsto E_{\mu}(\Sigma) \) is continuous with respect to weak-* convergence of measures in \( \mathcal{P}_R \).

**Proof.** We recall that the Wasserstein metric, \( d_W \), metrizes the weak-* convergence of measures on the set of measures supported in \( B(0, R) \). Therefore if \( \mu_n \rightharpoonup \mu \) in \( \mathcal{P}_R \) then \( d_W(\mu_n, \mu) \to 0 \) as \( n \to \infty \). Hence there exists a coupling (i.e. a transportation plan), \( \Pi_n \) between \( \mu \) and \( \mu_n \) such that
\[
\int_{B(0,R) \times B(0,R)} |x-y|^2 d\Pi_n(x,y) \to 0 \text{ as } n \to \infty.
\]
Therefore
\[
\left| \int_{\mathbb{R}^d} d(x, \Sigma) d\mu_n(x) - \int_{\mathbb{R}^d} d(x, \Sigma) d\mu(x) \right| = \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} d(x, \Sigma) - d(y, \Sigma) d\Pi_n(x,y) \right|
\]
\[
\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y| d\Pi_n(x,y)
\]
\[
\leq \sqrt{R} d_W(\mu_n, \mu) \to 0 \text{ as } n \to \infty.
\]

(4)
\( \square \)

**Lemma 3.** If \( \mu_n \rightharpoonup \mu \) in the weak topology of measures in \( \mathcal{P}_R \) then \( E_{\mu_n} \rightharpoonup E_\mu \) with respect to Hausdorff convergence of sets on \( \mathcal{A} \).

**Proof.** To prove the \( \Gamma \)-convergence we need to show the following

- Lower-semicontinuity. Assume \( \mu_n \rightharpoonup \mu \) and \( \Sigma_n \to \Sigma \) in Hausdorff metric as \( n \to \infty \). Then
\[
\liminf_{n \to \infty} E_{\mu_n}(\Sigma_n) \geq E_\mu(\Sigma).
\]

- Construction. Assume \( \mu_n \rightharpoonup \mu \). For any \( \Sigma \in \mathcal{A} \) there exists a sequence \( \Sigma_n \in \mathcal{A} \), such that \( \Sigma_n \to \Sigma \) in Hausdorff metric and
\[
\lim_{n \to \infty} E_{\mu_n}(\Sigma_n) = E_\mu(\Sigma).
\]
The construction claim follows from Lemma 2 by taking $\Sigma_n = \Sigma$.

Let us consider the lower-semicontinuity. As before, $\mu_n \rightharpoonup^* \mu$ implies $d_W(\mu_n, \mu) \to 0$ as $n \to \infty$. As in the estimate (4)

$$ \left| \int_{\mathbb{R}^d} d(x, \Sigma_n)d\mu_n(x) - \int_{\mathbb{R}^d} d(x, \Sigma_n)d\mu(x) \right| \leq \sqrt{R}d_W(\mu_n, \mu) \to 0 \text{ as } n \to \infty.$$

We note that the bound does not depend on $\Sigma_n$. Therefore, using Lemma 1

$$ \lim_{n \to \infty} E_{\mu_n}(\Sigma_n) = \lim_{n \to \infty} E_{\mu}(\Sigma_n) \geq E_{\mu}(\Sigma). $$

\[ \Box \]

**Corollary 4.** Assume that $\mu_n \rightharpoonup^* \mu$ in $\mathcal{P}_R$ and that $\Sigma_n$ is a minimizer of $E_{\mu_n}$. Then along a subsequence $\Sigma_n \rightharpoonup \Sigma$ where $\Sigma$ is a minimizer of $E_{\mu}$.

**Proof.** Since by Blaschke’s theorem (see [1]) the sequence $\Sigma_n$ has a subsequence which converges in Hausdorff metric the claim follows from the $\Gamma$ convergence. \[ \Box \]

**Corollary 5.** Assume $E_{\mu}$ has a unique minimizer $\Sigma$ and that $\mu_n \rightharpoonup^* \mu$ in $\mathcal{P}_R$. Then for every $\varepsilon > 0$ there exists $n_0$ such that for all $n \geq n_0$ any minimizer $\Sigma_n$ of $E_{\mu_n}$ satisfies $d_H(\Sigma, \Sigma_n) < \varepsilon$.

**Proof.** Assume that the claim does not hold. Than there exists $\varepsilon > 0$ and a sequence $\Sigma_{n_k}$ of minimizers of $E_{\mu_{n_k}}$ such that for each $k$, $d_H(\Sigma, \Sigma_{n_k}) \geq \varepsilon$. By relabeling, we can assume $n_k = k$ for all $k$. By Blaschke’s theorem, there exists $\bar{\Sigma} \in \mathcal{A}$ such that, along a subsequence, $\Sigma_n \rightharpoonup \bar{\Sigma}$ as $n \to \infty$ in Hausdorff metric. We can again assume that the subsequence is the whole sequence. Furthermore $\bar{\Sigma}$ is connected and thus belongs to $\mathcal{A}$. We note that $d_H(\Sigma, \bar{\Sigma}) \geq \varepsilon$. By the lower-semicontinuity part in the $\Gamma$-convergence, $\bar{\Sigma}$ is a minimizer of $E_{\mu}$, which contradicts the uniqueness assumption. \[ \Box \]

**Lemma 6.** Let $R > 0$. Let $\gamma_n : [0, 1] \to \overline{B}(0, R)$ be a sequence of Lipschitz curves with constant-speed parameterization (i.e. $|\gamma_n'(s)| = \text{length}(\gamma_n)$ for a.e. $s \in [0, 1]$). Assume that $\sup_n \text{length}(\gamma_n)$ and $\sup_n \|\gamma_n'\|_{BV}$ are finite. Then along a subsequence $\gamma_n$ converges to a Lipschitz curve $\gamma$ in the sense that

- $\gamma_n \to \gamma$ in $C^\alpha$ as $n \to \infty$, for any $\alpha \in [0, 1)$,
- $\gamma_n' \to \gamma'$ in $L^p$ as $n \to \infty$, for any $p \in [1, \infty)$, and
- $\gamma_n'' \rightharpoonup^* \gamma''$ in the space of finite signed Borel measures as $n \to \infty$.

**Proof.** The constant-speed assumption and the uniform bound on the lengths imply that $\|\gamma_n'\|_{L^\infty}$ are uniformly bounded and thus there is a uniform bound on the Lipschitz norm for the curves. The fact that $\gamma_n$ converges along a subsequence in $C^\alpha$ for any $\alpha \in [0, 1)$ follows since the set of Lipschitz functions with values in $\overline{B}(0, R)$ is compactly embedded in $C^\alpha$. To obtain the convergence that holds for all $\alpha$ at the same time one also uses a diagonalization argument.
From the embedding theorem of $BV$ spaces (see [1, 9]), it follows that for some $g \in BV([0, 1], \mathbb{R}^d)$, along a further subsequence, $\gamma'_n \to g$ in $L^1$ as $n \to \infty$ and $\gamma''_n \rightharpoonup g'$ in the space of signed measures as $n \to \infty$. Using the definition of the weak derivative it follows that $g = \gamma'$. Since $\|\gamma'_n\|_{L^\infty}$ are uniformly bound by interpolation it follows that for all $p \in [1, \infty)$, $\gamma'_n \to \gamma'$ in $L^p$ as $n \to \infty$. Furthermore $|\gamma'_n| \to |\gamma'|$ in $L^1$ as $n \to \infty$ and constant-speed assumption imply that and moreover $|\gamma'(s)| = \text{length}(\gamma)$ for a.e. $s \in [0, 1]$ and in particular $\gamma$ is a Lipschitz curve. □

2.1. Scaling of $E_\mu$ with respect to dilations of $\mu$. Given a set $A \subset \mathbb{R}^d$ and $r > 0$ we define $A^r = \{x : rx \in A\}$.

Given a measure $\mu$ and $r > 0$ we define the dilation of $\mu$ to scale $r$ to be the measure $D_r\mu$ such that for any $\mu$-measurable set $A$

$$D_r\mu(A) = \mu\left(\frac{A}{r}\right).$$

We note that since both terms of $E$ are scale linearly with respect to length

$$E_\mu(\Sigma) = \frac{1}{r} E_{D_r\mu}(r\Sigma).$$

Therefore if $\Sigma$ is a minimizer of $E_\mu$ then $r\Sigma$ is a minimizer of $E_{D_r\mu}$.

3. Discrete data

In this section we consider the case that $\mu$ is a discrete (or particle) measure:

$$\mu = \sum_{i=1}^{n} m_i \delta_{x_i}$$

where $m_i > 0$ and $x_i \in \mathbb{R}^d$. The measures $m_i \delta_{x_i}$ are called particles. We denote the support of $\mu$ by $X = \{x_1, \ldots, x_n\}$.

**Lemma 7.** If $\mu$ is discrete then every minimizer $\Sigma$ is graph with straight edges.

**Proof.** Let $\varphi : X \to \Sigma$ be the mapping that assigns to each $x_i \in X$ a point on $\Sigma$ which is the closest to $x_i$ (if the closest point is nonunique, an arbitrary one is chosen). Let $y_i = \varphi(x_i)$ and $Y = \{y_1, \ldots, y_n\}$ (the points are not necessarily distinct). Let $A$ be the Steiner tree containing the set $Y$. The Steiner tree is the connected graph with minimal total length of edges containing the vertices in $Y$ (it can have other vertices as well). We note that it is also the connected set of minimal $H^1$ measure containing $Y$. For further information on Steiner trees we refer to [5] and [6].

Furthermore note that $E(A) \leq E(\Sigma)$ with equality holding only if $\Sigma$ is also a Steiner tree, which proves the claim. We remark that $\Sigma$ may be different than $A$ since Steiner trees are not necessarily unique. □
Now that we know that $\Sigma$ has straight edges, we can study it more carefully. We define the vertices, $V$ as the set of those points $v$ in $\Sigma$ for which there exists a point $x$ in $X$ such that $v$ is the closest point to $x$ in $\Sigma$:

$$V = \{ v \in \Sigma : (\exists x \in X)(\forall z \in \Sigma) \ d(x, v) \leq d(x, z) \}.$$  

Note that $Y \subset V$ and that it is possible that at a vertex of degree two the angle is $180^\circ$. Since, by above the segments of $\Sigma$ connecting the vertices must be line segments, we define edges, $S$, as follows: for $v, w \in V$, $\{v, w\} \in S$ is an edge if the line segment $[v, w] \subseteq \Sigma$. We note that since $\Sigma$ is made of finitely many line segments, $V$ must be finite. Thus we can write $V = \{ v_1, \ldots, v_m \}$. Since

$$\int_{\mathbb{R}^d} d(x, \Sigma) d\mu(x) = \sum_{i=1}^{n} m_i d(x_i, \Sigma) = \sum_{i=1}^{n} m_i d(x_i, V)$$

$\Sigma$ must be connected the graph of minimal total length containing the vertices $V$. That is $\Sigma$ is a Steiner tree [5] for the set $V$ too.

The following facts on Steiner trees are available in the classical paper by Gilbert and Pollak [5].

**Proposition 8.** Let $G = (V, S)$ be as above. Then

(i) $G$ is a tree, that is it does not contain a closed loop.

(ii) If $\{u, v\}$ and $\{v, w\}$ are edges then the angle $\angle uvw \geq 120^\circ$.

(iii) The maximal degree of a vertex is three.

(iv) If $v$ is a vertex of degree three then the angles between edges at $v$ are $120^\circ$, and thus all three edges belong to a 2-dimensional plane.

We call the vertices of degree one the **endpoints**, the ones of degree two **corners**, and the ones of degree three **triple junctions**. Given $j = 1, \ldots, m$ let $I_j$ be the set of indices of points in $X$ for which $v_j$ is the closest point in $V$

$$I_j = \{ i \in \{1, \ldots, n\} : (\forall k = 1, \ldots, m) \ d(x_i, v_j) \leq d(x_i, v_k) \}$$

$$= \{ i \in \{1, \ldots, n\} : (\forall y \in \Sigma) \ d(x_i, v_j) \leq d(x_i, y) \}.$$

If $i \in I_j$ then we say that $x_i$ **talks to** $v_j$. We say that a vertex $v_j$ is **tied down** if for some $i$, $v_j = x_i$. We then say that $v_j$ is **tied to** $x_i$. Note that if $v_j$ is tied to $x_i$ then $i \in I_j$ and $T_{ij} = m_i$. A vertex which is not tied down is called **free**. We show below that if $x_i$ talks to $v_j$ and $v_j$ is free then $x_i$ cannot talk to any other vertex.

Consider an $n$ by $m$ matrix $T$ such that

$$T_{ij} \geq 0, \sum_{j=1}^{m} T_{ij} = m_i, \text{ and } T_{ij} > 0 \implies i \in I_j$$
Note that $\mu = \sum_{i=1}^{n} \sum_{j=1}^{m} T_{ij} \delta_{x_i}$. We define $\sigma$ to be the projection of $\mu$ onto the set $\Sigma$, in the sense that the mass from $\mu$ is transported to a closest point on $\Sigma$. That is

$$\sigma = \sum_{j=1}^{m} \sum_{i=1}^{n} T_{ij} \delta_{v_j}.$$  

We note that the matrix $T$ describes an optimal transportation plan between $\mu$ and $\sigma$ with respect to any of the transportation costs $c(x, y) = |x - y|^p$, for $p \geq 1$. We claim that such matrix $T$ exists. It is enough to consider mapping $\varphi$ from the proof of Lemma 7 and then define

$$T_{ij} = \begin{cases} m_i & \text{if } \varphi(x_i) = v_j \\ 0 & \text{otherwise}. \end{cases}$$

We note that in this discrete setting

$$E(\Sigma) = \sum_{i=1}^{n} m_i d(x_i, \Sigma) + \lambda \sum_{\{v, w\} \in S} |v - w|$$

$$= \sum_{j=1}^{m} \sum_{i \in I_j} T_{ij} |x_i - v_j| + \lambda \sum_{\{v_j, v_k\} \in S} |v_j - v_k|$$

Lemma 9. Assume that $\Sigma$ minimizes $E$ for discrete $\mu$ defined in (5). Let $V$ be the set of vertices as defined in (7) and $T$ be any matrix (transportation plan) satisfying (8). For any vertex $v_j$:

(i) If $v_j$ is an endpoint then let $w$ be the vertex such that $\{v_j, w\}$ is an edge. If $v_j$ is free then

$$\sum_{i \in I_j} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} + \lambda \frac{w - v_j}{|w - v_j|} = 0$$

If $v_j$ is tied to $x_k$ then

$$\left| \sum_{i \in I_j, i \neq k} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} + \lambda \frac{w - v_j}{|w - v_j|} \right| \leq m_k$$

(ii) If $v_j$ is a corner then let $\{w_1, v_j\}$ and $\{v_j, w_2\}$ be the edges at the corner. If $v_j$ is free then

$$\sum_{i \in I_j} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} + \lambda \left( \frac{w_1 - v_j}{|w_1 - v_j|} + \frac{w_2 - v_j}{|w_2 - v_j|} \right) = 0$$

If $v_j$ is tied to $x_k$ then

$$\left| \sum_{i \in I_j, i \neq k} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} + \lambda \left( \frac{w_1 - v_j}{|w_1 - v_j|} + \frac{w_2 - v_j}{|w_2 - v_j|} \right) \right| \leq m_k.$$
(iii) If \( v_j \) is a triple junction and if \( v_j \) is free then

\[
\sum_{i \in I_j} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} = 0.
\]

If \( v_j \) is tied to \( x_k \) then

\[
\left| \sum_{i \in I_j} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} \right| \leq m_k.
\]

Proof. To prove (i) and (ii), consider now the configuration \( \Sigma_v \) which is obtained from \( \Sigma \) just by changing the location of \( v_j \) to \( v_j \). Let \( S_v \) be the set of edges of the new graph. Formulation (10) provides

\[
E(\Sigma_v) \leq \sum_{i \in I_j} T_{ij} |x_i - v| + \lambda \sum_{\{v,v_l\} \in S_v} |v - v_l|
\]

\[
+ \sum_{l=1, l \neq j}^m \sum_{i \in I_l} T_{il} |x_i - v_l| + \lambda \sum_{\{v_l,v_l\} \in S_v} |v_k - v_l| =: F(v)
\]

Note that \( F \) defined above maps \( \mathbb{R}^d \rightarrow \mathbb{R} \) and that \( E(\Sigma) = F(v_j) \).

If \( v_j \) is a free vertex then \( F \) is a smooth function near \( v_j \). The minimality of \( \Sigma \) thus implies that \( DF(v_j) = 0 \). Straightforward computation of \( DF \) gives the conditions (11) and (13).

If on the other hand \( v_j \) is tied to \( x_k \) for some \( k \), \( x_k = v_j \) then recall that \( k \in I_j \) and furthermore \( T_{kj} = m_k \). \( F \) is no longer smooth at \( v_j \) but it still has a minimum at \( v_j \). Therefore zero vector must belong to the subgradient of \( F \) at \( v_j \), that is \( 0 \in \partial F(v_j) \). Using that the subgradient at 0 of \( z \mapsto |z| \) is \( B(0,1) \) we conclude that \( 0 \in \partial F(v_j) \) if the conditions (12) and (14) hold at an endpoint and corner, respectively. More precisely if we define

\[
\tilde{F}(v) = \sum_{i \in I_j, i \neq j} T_{ij} |x_i - v| + \lambda \sum_{\{v,v_l\} \in S_v} |v - v_l|
\]

Then, using \( v_j = x_k \),

\[
F(v) = m_k |v_j - v| + \tilde{F}(v) + \text{const.}
\]

and hence

\[
\partial F(v_j) = B(0, m_k) + D\tilde{F}(v_j).
\]

Therefore \( 0 \in \partial F(v_j) \) if \( |D\tilde{F}(v_j)| \leq m_k \).

Obtaining (15) and (16) is analogous, only that one also needs the fact that the angles at triple junction are 120°, see Proposition 8. \( \square \)

Corollary 10. Assume that conditions of the lemma are satisfied.

(i) In two dimensions, \( d = 2 \), if \( v_j \) is a triple junction and is a free vertex then \( I_j = \emptyset \).
(ii) If \( i \) talks to \( v_j \) and \( v_k \) \((j \neq k)\) then both \( v_j \) and \( v_k \) are tied down. Consequently if \( x_i \) talks to \( v_j \) and \( v_j \) is free then \( T_{ij} = m_i \). Furthermore \( m \leq n \).

(iii) If \( v_j \) is an endpoint then

\[
\sum_{i \in I_j} T_{ij} \geq \lambda.
\]

Hence at every endpoint, the measure \( \sigma \) defined in (9) has an atom of the mass at least equal to \( \lambda \). Note that this gives an upper bound on the number of endpoints. A further consequence is that if \( 2\lambda > \sum_{i=1}^n m_i \) then the minimizer \( \Sigma \) is just a single point (which is then a vertex of degree zero).

**Proof.** To prove (i) assume that \( i \in v_j \). Then \( B(x_i, |v_j - x_i|) \cap \Sigma = \emptyset \). But this contradicts the fact that the angles at triple junction are \( 120^\circ \).

To prove (ii) assume that there exist \( i \in \{1, \ldots, n\} \) and \( j,k \) distinct elements of \( \{1, \ldots, m\} \) such that \( i \in I_j \) and \( i \in I_k \). Let \( T \) be a matrix satisfying the condition (8). For \( s \in (0,1) \) consider the the matrix \( T(s) \) obtained from \( T \) by setting:

\[
T_{ij}(s) = m_i(1-s), \quad T_{ik} = m_is \quad \text{and} \quad T_{il} = 0 \quad \text{if} \quad l \notin \{j,k\}.
\]

Note that \( T(s) \) satisfies the condition (8).

We argue by contradiction: assume that \( v_j \) is free. Let us consider first the case that \( v_j \) is an endpoint. Let \( w \) be the vertex for which \( \{v,w\} \) is an edge. The condition (11) must be satisfied for \( T(s) \) for all \( s \in [0,1] \):

\[
m_i(1-s) \frac{x_i-v_j}{|x_i-v_j|} + \sum_{l \in I_j, l \neq i} T_{lj} \frac{x_l-v_j}{|x_l-v_j|} + \lambda \frac{w-v_j}{|w-v_j|} = 0.
\]

One arrives at a contradiction by taking the derivative in \( s \).

The proofs for a corner and for a triple junction are analogous.

Let us explain why the above implies \( m \leq n \). By definition of \( V \), for each \( v_j \in V \), \( I_j \neq \emptyset \), that there is a particle talking to \( v_j \). We claim that for every \( v_j \) there is a particle talking only to \( v_j \). If \( v_j \) is free then by above this is the case with any particle in \( I_j \). If \( v_j \) is tied to \( x_k \) then \( x_k \) only talks to \( v_j \). Consequently there must be more particles then vertices in \( V \).

Let us now consider the claim of (iii). There are two cases. We first consider the case that \( v_j \) is free. Then (11) holds. Taking the inner product with \( \frac{w-v_j}{|w-v_j|} \) and using (ii) above gives

\[
- \sum_{i \in I_j} m_i \frac{x_i-v_j}{|x_i-v_j|} \cdot \frac{w-v_j}{|w-v_j|} = \lambda.
\]

Thus

\[
\sum_{i \in I_j} m_i \geq \sum_{i \in I_j} m_i \left| \frac{x_i-v_j}{|x_i-v_j|} \cdot \frac{w-v_j}{|w-v_j|} \right| \geq \lambda.
\]
If $v_j$ is tied to $x_k$ for some $k$ then the condition \[ (12) \] gives
\[
\left| \sum_{i \in I_j, i \neq k} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} \cdot \frac{w - v_j}{|w - v_j|} + \lambda \right| \leq m_k.
\]
Thus
\[
\lambda \leq m_k - \sum_{i \in I_j, i \neq k} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} \cdot \frac{w - v_j}{|w - v_j|} \leq \sum_{i \in I_j} T_{ij}.
\]

3.1. Turning angle. If $v$ is a corner with adjacent vertices $w_1$ and $w_2$ then the turning angle at $v$ is $TA(v) = \pi - \angle w_1vw_2$. Basically it describes the curvature at $v$. For $A \subset \Sigma$, the turning angle of $A$, $TA(A) = \sum_{v \in A \cap v} TA(v)$. The total turning angle, $TTA = TA(\Sigma)$.

Lemma 11. If $\Sigma$ is a minimizer of $E$ and $A \subset \Sigma$.
\[
(18) \quad TA(A) \leq \frac{\pi}{2\lambda} \sum_{i \in \cup \{ I_j : v_j \in A \}} m_i.
\]

Proof. Let us first consider the case that $A$ is a single corner, $A = \{v_j\}$. Let $w_1$ and $w_2$ be the adjacent vertices. Let $\alpha$ be the half of the turning angle: $TA(v_j) = 2\alpha$. Then $\angle w_1vw_2 = \pi - 2\alpha$. Let $\theta_1 = \frac{w_j - w_1}{|w_j - w_1|}$ and $\theta_2 = \frac{w_2 - v_j}{|w_2 - v_j|}$. Elementary geometry yields that $|\theta_2 - \theta_1| = 2\sin \alpha$.

Analogously to the proof of statement (iii) of Corollary \[ (10) \] that is by using conditions \[ (13) \] and \[ (14) \] and taking inner product with $\frac{\theta_2 - \theta_1}{|\theta_2 - \theta_1|}$, one can show that
\[
(19) \quad 2\lambda \sin \alpha \leq \sum_{i \in I_j} T_{ij}.
\]
Therefore
\[
\alpha \leq \max \left\{ \frac{\pi}{2}, \arcsin \left( \frac{\sum_{i \in I_j} T_{ij}}{2\lambda} \right) \right\} \leq \frac{\pi}{2} \frac{\sum_{i \in I_j} T_{ij}}{2\lambda}
\]
which establishes the claim.

For the general $A \subset \Sigma$ it suffices to sum over the vertices it contains. \hfill \Box

3.2. Region of Influence. Given orthogonal unit vectors $\xi$ and $b$ and an angle $\beta \in [0, \frac{\pi}{2}]$, consider the wedge $W$, with bisector $b$ and opening $2\beta$, defined as follows:
\[
(20) \quad W(\xi, b, \beta) = \{ x \in \mathbb{R}^d : |\xi \cdot x| < b \cdot x \tan \beta \}
\]

By $\ll z_1, \ldots, z_k \gg$ we denote the piecewise linear curve connecting the points $z_1, \ldots, z_k$. Given three points $v_1$, $v_2$, and $v_3$ such that if they are collinear then $v_2$ lies between $v_1$ and $v_3$ consider $\Sigma = \ll v_1, v_2, v_3 \gg$. If the points are collinear the region of influence of $v_2$, defined in \[ (3) \], is the hyperplane passing through $v_2$, orthogonal to $v_3 - v_2$. If points are not collinear then let $\theta_i = \frac{v_{i+1} - v_i}{|v_{i+1} - v_i|}$ for $i = 1, 2$. The region of influence of
Figure 1. The basic configuration.

$v_2$ the translated wedge: $R(v_2) = v_2 + W(\xi, b, \beta)$ where $\xi = \frac{\theta_1 + \theta_2}{|\theta_1 + \theta_2|}$, $b = \frac{\theta_1 - \theta_2}{|\theta_1 - \theta_2|}$, and $\beta = TA(v_2)/2$. We denote the mapping above that outputs the Normal, Bisectris, and Angle given non-collinear points $v_1, v_2, v_3$ by $(\xi, b, \beta) = NBA(v_1, v_2, v_3)$. Then we can write $R(v_2) = v_2 + W(NBA(v_1, v_2, v_3))$.

We note that if $\Sigma = [v_1, \ldots, v_m]$ then for all $i = 2, \ldots, m - 1$, if $TA(v_i) > 0$ $R(v_i) \subseteq v_i + W(NBA(v_{i-1}, v_i, v_{i+1}))$.

3.3. **Basic Configuration.** Here we describe the basic configuration $\tilde{\mu}$, whose perturbation is used in the counterexample to regularity. While the construction is rather flexible we choose a fixed configuration to make the number of parameters of the system easier to manage.

Let $m_1 = m_3 = 0.38$, $m_2 = 0.24$, and $\lambda = 0.36$. Let $\bar{x}_1 = (-1, 0)$, $\bar{x}_2 = (0, 1)$, and $\bar{x}_3 = (1, 0)$.

(21) $\tilde{\mu} = m_1 \delta_{\bar{x}_1} + m_2 \delta_{\bar{x}_2} + m_3 \delta_{\bar{x}_3}$.

Note that the total mass is one.

We now turn to characterizing the minimizer. We note that since $\lambda$ is greater than one third of the total mass the condition (17) implies that the minimizer can have at most two endpoints. Thus the minimizer is either a point or a piecewise linear curve. We reindex the vertices $v_1, \ldots, v_m$ if needed so that the minimizer is $[v_1, \ldots, v_m]$.

Let us note that for any point $a \in \mathbb{R}^2$, $E(\{a\}) > E([\bar{x}_1, a, \bar{x}_3])$ and thus a point cannot be a minimizer. Therefore a minimizer is a piecewise linear curve with either one or two line segments. If it has two line segments then $v_2 \neq \bar{x}_1$ since then it would violate angle condition (ii) of Proposition 8 since we know the minimizer must stay in the convex hull of the support of $\tilde{\mu}$. If $v_1 \neq \bar{x}_1$ and $v_m \neq \bar{x}_3$ then $E([\bar{x}_1, v_1, \ldots, v_m]) < E([v_1, \ldots, v_m])$ since $m_1 > \lambda$. Analogously for $\bar{x}_3$. Therefore $\bar{x}_1$ and $\bar{x}_3$ must be endpoints of any minimizer. So without loss of generality we can set $v_1 = \bar{x}_1$ and $v_m = \bar{x}_3$. If $[v_1, v_2]$ were a minimizer then $\bar{x}_2$ would talk to $(0, 0)$ so $(0, 0)$ would have to be a vertex of the graph (as we defined it (6)). But then the condition (13) cannot hold. So the minimizer must be of the form $[\bar{x}_1, \bar{v}_2, \bar{x}_3]$. 
The criticality condition (13) implies that the only minimizer is the symmetric configuration, \( \Sigma \), presented on Figure 1. Elementary geometry gives:

\[
\bar{v}_2 = (0, H), \quad \text{where } H = \frac{1}{2\sqrt{2}}, \quad L = \frac{3}{2\sqrt{2}}, \quad \text{and } \sin \alpha = \frac{1}{3}.
\]

4. Construction of the counterexample

We start with the basic configuration \( \bar{\mu} \) defined in (21), which we now consider as configuration in \( \mathbb{R}^d \) by taking the values of coordinates from 3 to \( d \) to be zero.

To smooth out the basic configuration, we use a standard mollifier \( \eta \). That is, let \( \eta \) be smooth, radially symmetric, positive on \( B(0, 1) \), equal to zero outside of \( B(0, 1) \), and such that \( \int_{\mathbb{R}^d} \eta(x) dx = 1 \). For \( \delta > 0 \) let \( \eta_\delta(z) = \frac{1}{\delta^d} \eta(\frac{z}{\delta}) \). Let \( \rho_{i, \delta} = m_i \eta_\delta(\cdot - \bar{x}_i) \) for \( i = 1, 2, 3 \) and let \( \mu_\delta \) be the measure with density \( \rho_{1, \delta} + \rho_{2, \delta} + \rho_{3, \delta} \).

To have a measure with connected and convex support we introduce the background measure \( \bar{\mu} \) to be the measure with density \( \eta_{1.5} \). The smooth measure we consider is

\[
\mu_{q, \delta} = (1 - q)\mu_\delta + q\bar{\mu}.
\]

**Theorem 12.** There exists \( \delta > 0 \) and \( 1 > q > 0 \) such that there is a minimizer of \( E \) for \( \lambda = 0.36 \) and \( \mu = \mu_{q, \delta} \) which is a Lipschitz curve such that if one considers its constant-speed parameterization \( \gamma : [0, 1] \rightarrow \mathbb{R}^d \) \( (|\gamma'(s)| = \text{length}(\gamma) \text{ for a.e. } s \in [0, 1]) \) then \( \gamma' \) is an \( \mathbb{R}^d \) valued BV function such that \( \gamma'' \) has an atom of size at least 1 at some \( s \in (0, 1) \). More precisely \( |\gamma''(\{s\})| \geq 1 \).

**Proof.** Assume that \( \varepsilon \) satisfies the condition

\[
(C1) \quad 0 < \varepsilon < \frac{\alpha}{2000}.
\]

Corollary 5 implies that for \( \delta > 0 \) and \( q > 0 \) small enough any minimizer of \( E_{\mu_{q, \delta}} \) lies within \( \varepsilon \) ball of \( \Sigma \) in Hausdorff metric. That is, we can impose

\[
(C2) \quad q > 0 \text{ and } \delta > 0 \text{ are small enough so that any minimizer } \Sigma_{q, \delta} \text{ of } E_{\mu_{q, \delta}} \text{ satisfies } d_H(\Sigma_{q, \delta}, \Sigma) < \frac{\varepsilon}{2}.
\]

We also require:

\[
(C3) \quad q < \frac{2\lambda}{\pi} \frac{\alpha}{20000}.
\]

\[
(C4) \quad \delta < 0.1\varepsilon.
\]

The condition (C3) controls the part of the turning angle which is due to the background measure.

**Step 1.** *Discrete approximation.* Let \( \mu_{q, \delta, n} \) be an approximation of \( \mu_{q, \delta} \) which is a particle measure such that the Wasserstein distance \( d_W(\mu_{q, \delta}, \mu_{q, \delta, n}) < \frac{1}{n} \) and furthermore that there exists an optimal coupling such that all of the mass in the \((1 - q)\rho_{i, \delta}\) part of \( \mu_{q, \delta} \) is coupled with particles in \( B(\bar{x}_i, \delta) \) for \( i = 1, 2, 3 \). This can be achieved by, say, taking a fine square grid such that \( \bar{x}_1, \bar{x}_2, \) and \( \bar{x}_3 \) are cell centers and then constructing \( \mu_{q, \delta, n} \) by taking the mass of \( \mu_{q, \delta} \) in each grid cell and concentrating it at the center of the grid cell.

Due to Corollary 5, along a subsequence, minimizers, \( \Sigma_{q, \delta, n} \) (if nonunique then any minimizer can be chosen), converge in Hausdorff metric to a minimizer \( \Sigma_{q, \delta} \) of \( E_{\mu_{q, \delta}} \).
We can assume without loss of generality that the whole sequence converges and that $n$ is large enough so that
\[
d_H(\Sigma_{q,\delta,n}, \Sigma_{q,\delta}) < \frac{\varepsilon}{2}
\]
which implies, using (C2), that $d_H(\Sigma_{q,\delta,n}, \Sigma) < \varepsilon$.

**Step 2a Control of the optimal coupling.** We claim that particles of $\mu_{q,\delta,n}$ which lie in $B(\bar{x}_2, \delta)$ can only talk to points on $\Sigma_{q,\delta,n}$ which lie above (in the direction of $e_2$) the hyperplane $P = \{ y : y \cdot e_2 = H - \delta - \frac{\varepsilon}{\cos \alpha} \}$.

To prove the claim consider $x_i \in B(\bar{x}_2, \delta)$. Let $\tilde{x}_i$ be the projection of $x_i$ on the coordinate axis in the direction of the vector $e_2$. That is $\tilde{x}_i = (0, x_{i,2}, 0, \ldots, 0)$. Let $U = \{ z : d(z, \Sigma) < \varepsilon \text{ and } z \cdot e_1 = 0 \}$. We note that $U \cap \Sigma_{q,\delta,n} \neq \emptyset$. An elementary geometry argument shows that the furthest point to $\tilde{x}_2$ on $U$ is $(0, H - \frac{\varepsilon}{\cos \alpha}, 0, \ldots, 0)$. Thus
\[
d(x_i, \Sigma_{q,\delta,n}) \leq d(x_i, \tilde{x}_i) + d(\tilde{x}_i, \Sigma_{q,\delta,n}) < \delta + x_{i,2} - H + \frac{\varepsilon}{\cos \alpha}.
\]
On the other hand
\[
d(x_i, P) = x_{i,2} + \delta - H + \frac{\varepsilon}{\cos \alpha}.
\]

**Step 3a Average tangent direction.** Since $\Sigma_{q,\delta,n}$ has only two endpoints, and is piecewise linear, it can be represented by a constant-speed parameterized curve $\gamma_{q,\delta,n} : [0,1] \to \mathbb{R}^d$. Due to the closeness to $\Sigma$ (by Step 1) there are points on $\Sigma_{q,\delta,n}$ which are within $\varepsilon$ of $\tilde{x}_1$ and $\tilde{x}_3$.

We claim that the endpoints of the curve must lie within $2\varepsilon$ of $\bar{x}_1$ and $\bar{x}_3$. For if that was not the case then all of the mass in, say, $B(\bar{x}_1, \delta)$ would not talk to an endpoint. But then there would not be enough available mass for the lower bound on the mass talking to the endpoints (condition [17]) to be satisfied.
We can require that $\gamma_{q,\delta,n}(0)$ lies close to $\bar{x}_1$. We define

$$\theta(s) = \frac{\gamma'_{q,\delta,n}(s)}{\mid \gamma'_{q,\delta,n}(s) \mid} = \frac{\gamma'_{q,\delta,n}(s)}{\text{length}(\gamma_{q,\delta,n})}.$$ 

We note that $\theta \in BV([0,1],\mathbb{R}^d)$. Let

$$\bar{\theta}_1 = \frac{\bar{v}_2 - \bar{x}_1}{\mid \bar{v}_2 - \bar{x}_1 \mid} \quad \text{and} \quad \bar{\theta}_2 = \frac{\bar{x}_3 - \bar{v}_2}{\mid \bar{x}_3 - \bar{v}_2 \mid}.$$ 

Let $\text{Cyl}_1 = \{ x \in \mathbb{R}^d : d(x,\Sigma) < \varepsilon \text{ and } (x - \bar{x}_1) \cdot \bar{\theta}_1 \in [2\varepsilon, L - 11\varepsilon] \}$ and $\text{Cyl}_2 = \{ x \in \mathbb{R}^d : d(x,\Sigma) < \varepsilon \text{ and } -(x - \bar{v}_2) \cdot \bar{\theta}_2 \in [2\varepsilon, L - 11\varepsilon] \}$, as shown on Figure 2. Let $s_{1,1,n}$ be the first time $\gamma_{q,\delta,n}$ enters $\text{Cyl}_1$ and $s_{1,2,n}$ the largest time $\gamma_{q,\delta,n}$ belongs to $\text{Cyl}_1$. Analogously $s_{2,1,n}$ and $s_{2,2,n}$ are the corresponding times for the domain $\text{Cyl}_2$. Let $\bar{\theta}_{i,\text{avg}} = \frac{\gamma_{q,\delta,n}(s_{i,2,n}) - \gamma_{q,\delta,n}(s_{i,1,n})}{\mid \gamma_{q,\delta,n}(s_{i,2,n}) - \gamma_{q,\delta,n}(s_{i,1,n}) \mid}$ for $i = 1, 2$. We claim that $\angle \bar{\theta}_i \bar{\theta}_{i,\text{avg}} < 0.01\alpha$ for $i = 1, 2$. To see this, note that $\tan \angle \bar{\theta}_i \bar{\theta}_{i,\text{avg}}$ is less than twice the width of the cylinder $\text{Cyl}_i$ divided by its length: $\tan \angle \bar{\theta}_i \bar{\theta}_{i,\text{avg}} < \frac{2\varepsilon}{L - 11\varepsilon} < \frac{4\varepsilon}{L} < 0.005\alpha$ by (C1), which implies the claim.

**Step 4** Tangent at the end of the cylinders. The cylinders $\text{Cyl}_1$ and $\text{Cyl}_2$ were defined so that they lie below the hyperplane $P$, as can be verified by simple trigonometry. Step 2 then implies that no point that belongs to $B(\bar{x}_i, \delta)$ for $i = 1, 2, 3$ talks to any point in the cylinders. Thus, using the turning angle estimate (18) and assumption (C3), $TA(\text{Cyl}_i \cap \Sigma_{q,\delta,n}) \leq \frac{\pi}{20} < \frac{\alpha}{100}$. Therefore the tangent at the any point in $\text{Cyl}_i$ is close to the average tangent. Let

$$\theta_1 = \frac{\gamma'_{q,\delta,n}(s_{1,2,n}^+)}{\mid \gamma'_{q,\delta,n}(s_{1,2,n}) \mid} = \lim_{s \rightarrow s_{1,2,n}} \frac{\gamma'_{q,\delta,n}(s)}{\mid \gamma'_{q,\delta,n}(s) \mid} \quad \text{and} \quad \theta_2 = \frac{\gamma'_{q,\delta,n}(s_{2,1,n}^-)}{\mid \gamma'_{q,\delta,n}(s_{2,1,n}) \mid}$$

be the tangents as the curve exits $\text{Cyl}_1$ and as it enters $\text{Cyl}_2$. Then $\angle \theta_i \theta_{i,\text{avg}} < 0.01\alpha$. Combining with estimates of Step 3 we get

$$\angle \bar{\theta}_i \theta_i \leq \angle \bar{\theta}_i \bar{\theta}_{i,\text{avg}} + \angle \bar{\theta}_{i,\text{avg}} \theta_i < 0.02\alpha \quad \text{for } i = 1, 2.$$

**Step 5** The turning angle at the first contact point. We now show that there exists a vertex at which the turning angle is large (of size at least about $\alpha/2$). This is the key point of the argument. Let us relabel the vertices if necessary, so that their indices are increasing along $\gamma_{q,\delta,n}$ as $s$ increases. Let $v_j$ be the first vertex of $\gamma_{q,\delta,n}$ that talks to any particle of $\mu_{q,\delta,n}$ in $B(\bar{x}_2, \delta)$, as illustrated on Figure 3. The reason that the turning angle has to be "large" is that the tangent cannot turn much prior to $v_j$ (because the vertices talk to very little mass), but for $v_j$ to be able to talk to a point in $B(\bar{x}_2, \delta)$ the tangent must turn by at least by about $\alpha$. Thus it has to turn by that amount precisely at $v_j$. 

Figure 3. Details of the configuration and relevant angles near the tip.

Here is the detailed argument. Let

\[ \theta_- = \frac{v_j - v_{j-1}}{|v_j - v_{j-1}|} \quad \text{and} \quad \theta_+ = \frac{v_{j+1} - v_j}{|v_{j+1} - v_j|}. \]

Let \( s_n \) be the time at which \( \gamma_{q,\delta,n}(s_n) = v_j \). Since points on \( \gamma_{q,\delta,n}|_{[s_1,2,n,s_n]} \) can only talk to the background measure \( q\tilde{\mu} \), it follows from (18) that \( \pi q \frac{\lambda}{2\alpha} < 0.01\alpha \). Combining with Step 4 implies

(23) \[ \angle \bar{\theta}_1 \theta_- < 0.03\alpha. \]

Let \( K = \{ x : d(x, \Sigma) < \varepsilon, \ x \cdot e_2 > H - \delta - \frac{\varepsilon}{\cos \alpha} \} \). From Step 2 follows that \( v_j \in K \). Let \( y \in B(\bar{x}_2, \delta) \). We can decompose vectors in \( \mathbb{R}^d \) into the component in the direction of \( e_2 \) (vertical component) and the one in the orthogonal complement of \( e_2 \) (horizontal component). Elementary geometry implies using the assumptions (C1) and (C4) that the horizontal distance between \( y \) and any point in \( K \) is less than \( 6\varepsilon + 4\delta < 7\varepsilon \), while the vertical distance is greater than \( 1 - H - \varepsilon - \delta > \frac{1}{2} \). Hence

(24) \[ \angle (y - v_j)e_2 \leq \tan \angle (y - v_j)e_2 < 14\varepsilon < 0.01\alpha. \]

Using that \( \angle \bar{\theta}_1 e_2 = \frac{\pi}{2} - \alpha \) and \( \angle \bar{\theta}_1 \theta_- < 0.03\alpha \). We conclude that for all \( y \in B(\bar{x}_2, \delta) \),

\[ \angle (y - v_j)\theta_- \leq \angle (y - v_j)e_2 + \angle e_2 \bar{\theta}_1 + \angle \bar{\theta}_1 \theta_- < \frac{\pi}{2} - 0.96\alpha. \]
Let \( W_j = W(NBA(v_{j-1}, v_j, v_{j+1})) \). Recall that \( R(v_j) \subset W_j \). Note that \( \theta_- \) is orthogonal to one side of the wedge \( W_j \). Also note that by definition of \( v_j \), there exists \( y \in W_j \cap B(\bar{x}_2, \delta) \). Let \( 2\beta \) be the opening of the wedge. Then

\[
TA(v_j) = 2\beta \geq \frac{\pi}{2} - \angle \theta_- (y - v_j) > 0.96\alpha.
\]

**Step 6° Angle bisector estimate.** As a consequence of the estimate we obtain that \( TA(v_j) > 1.88\alpha \).

The idea of this step is as follows: In the previous step we have shown that the turning angle at \( v_j \) is "large". The criticality condition \([13]\) shows that \( v_j \) is thus talking to a large mass. Since the only large mass in the region of influence lies within \( B(\bar{x}_2, \delta) \) this implies that the bisector of the angle \( \angle v_{j-1}v_jv_{j+1} \) passes through, or close to, \( B(\bar{x}_2, \delta) \) which implies that the turning angle at \( v_j \) is about \( 2\alpha \).

More precisely, let \( m_{aux} \) be the mass of the particle of \( \mu_{q, \delta, n} \) at \( v_j \) if it is tied down and zero otherwise. We note that \( m_{aux} \leq q < 0.001\alpha \) by \((C3)\). Therefore using \([13], [14]\), and \( \mu_{q, \delta, n} = \sum_i m_i \delta_{x_i} \) one obtains

\[
\left| \sum_{i \in I_j, x_i \neq v_j} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} \right| \geq \lambda|\theta_+ - \theta_-| - m_{aux} \geq \lambda \left( 2 \sin \frac{TA(v_j)}{2} \right) - 0.001\alpha.
\]

Using that \( TA(v_j) > 0.96\alpha \) and that \( \sin \alpha = \frac{1}{3} \) and \( \lambda = 0.36 \) we conclude that \( \sum_{i \in I_j} T_{ij} > 0.08 \).

Let \( I = \{ i \in I_j, x_i \in B(\bar{x}_2, \delta) \} \), \( w_1 = \sum_{i \in I} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} \), and \( w_2 = \sum_{i \in I \setminus I} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} \). Since for all \( i \in I \), \( \angle (x_i - v_j)(\bar{x}_2 - v_j) < 0.01\alpha \) we conclude that

\[
|w_1| \geq \sum_{i \in I} T_{ij} \frac{x_i - v_j}{|x_i - v_j|} \cdot \frac{\bar{x}_2 - v_j}{|\bar{x}_2 - v_j|} \geq \frac{2}{3} \sum_{i \in I} T_{ij} > 0.05
\]

and \( \angle w_1(\bar{x}_2 - v_j) < 0.01\alpha \), where \( T \) is any matrix satisfying \([8]\). By definition of \( w_2 \):

\[
|w_2| \leq q.
\]

Therefore, using the sine theorem, \( \sin \angle((w_1 + w_2), w_1) < \frac{q}{0.05} \). Conditions \([13], [14]\) give

\[
|w_1 + w_2 + \lambda(\theta_+ - \theta_-)| \leq m_{aux}.
\]

Thus, by sine theorem,

\[
\angle(w_1 + w_2)(-(\theta_+ - \theta_-)) \leq \frac{m_{aux}}{\lambda|\theta_+ - \theta_-|} \leq \frac{3q}{\lambda},
\]

where we used that \( |\bar{\theta}_2 - \bar{\theta}_1| = 2 \sin \alpha = 2/3 \) to obtain lower bound \( |\theta_+ - \theta_-| > 1/3 \).

Hence, using the condition \((C3)\), \( \angle(\theta_- - \theta_+)w_1 \leq \angle(w_1 + w_2)w_1 + \frac{3q}{20.05} + \frac{3q}{\lambda} < \frac{\pi}{2} - \frac{q}{0.05} + \frac{3q}{\lambda} < 1/3 \).
0.01α. Hence for (ν, b, β) = NBA(v_{j-1}, v_j, v_{j+1}) we conclude that \( \angle b(\bar{x}_2 - v_j) \leq \angle b w_1 + \angle w_1(\bar{x}_2 - v_j) < 0.02\alpha \). Therefore, by using the estimates \[23\] and \[24\],
\[
\angle b\theta_- \leq \angle b(\bar{x}_2 - v_j) + \angle (\bar{x}_2 - v_j)e_2 + \angle e_2\theta_1 + \angle \theta_1\theta_- \leq \frac{\pi}{2} - 0.94\alpha.
\]
Therefore
\[ TA(v_j) = \pi - 2\angle b\theta_- \geq 1.88\alpha. \]

**Step 7º Symmetry argument.** In Steps 5 and 6, we considered \( v_j \), the first vertex of \( \gamma_{q,\delta,n} \) which talks to any particle of \( \mu_{q,\delta,n} \) in \( B(\bar{x}_2, \delta) \). The same arguments apply if one considers the last vertex of \( \gamma_{q,\delta,n} \), denote it be \( v_k \) (with \( k \geq j \)), talking to any particle of \( \mu_{q,\delta,n} \) in \( B(\bar{x}_2, \delta) \). Thus \( TA(v_k) \geq 1.88\alpha \). We claim that \( k = j \). For if one assumes that \( k > j \) then
\[ TA(\gamma_{q,\delta,n}([s_{1,2,n}, s_{2,1,n}])) > 3.76\alpha. \]

However by estimate \[18\]
\[
TA(\gamma_{q,\delta,n}([s_{1,2,n}, s_{2,1,n}])) \leq \frac{\pi}{2\lambda}(m_2 + q) < \frac{\pi}{0.72} \cdot 0.25 < 3.3\alpha
\]
which contradicts the statement above. Therefore \( k \) must equal \( j \). That is \( v_j \) is the only point on \( \gamma_{q,\delta,n} \) talking to particles in \( B(\bar{x}_2, \delta) \). Furthermore analogously to \[23\] it holds that \( \angle \theta_1\theta_+ < 0.03\alpha \). Hence
\[
TA(v_j) = \angle \theta_-\theta_+ \geq \angle \theta_1\theta_2 - \angle \theta_-\theta_1 - \angle \theta_+\theta_2 \geq 2\alpha - 0.03\alpha - 0.03\alpha = 1.94\alpha.
\]

**Step 8º Convergence.** By definition of the turning angle, using that \( \gamma_{q,\delta,n} \) has constant speed parameterization and that \( |\gamma_{q,\delta,n}''| \) is a measure, for any \( k \) it holds that:
\[ |\gamma_{q,\delta,n}''([t_{k,n}])| = |\gamma_{q,\delta,n}(t_{k,n}+) - \gamma_{q,\delta,n}(t_{k,n}-)| = 2\text{length}(\gamma_{q,\delta,n}) \sin \frac{TA(v_k)}{2} \]
where \( t_{k,n} = \gamma_{q,\delta,n}^{-1}(v_k) \). Therefore, using the estimate on the turning angle \[18\],
\[ |\gamma_{q,\delta,n}''([s_{1,2,n}, s_{2,1,n}] \setminus \{s_n\})| \leq \text{length}(\gamma_{q,\delta,n}) TA(\gamma_{q,\delta,n}([s_{1,2,n}, s_{2,1,n}] \setminus \{s_n\})) \leq \text{length}(\gamma_{q,\delta,n}) \frac{\pi q}{2\lambda} < \text{length}(\gamma_{q,\delta,n}) * 0.01\alpha. \]

Given that \( \text{length}(\gamma_{q,\delta,n}) \) is uniformly bounded from above and below in \( n \), that distance between \( |\gamma_{q,\delta,n}(s_{1,2,n}) - \gamma_{q,\delta,n}(s_{2,1,n})| \geq d(\text{Cyl}_1, \text{Cyl}_2) > 0 \) and \( |\gamma_{q,\delta,n}(s_n) - \gamma_{q,\delta,n}(s_{1,2,n})| \geq d(K, \text{Cyl}_1) > 0 \) we conclude that along a subsequence \( s_{1,2,n} \to s_1, s_{2,1,n} \to s_2, \) and \( s_n \to s \) as \( n \to \infty \) with \( 0 < s_1 < s < s_2 < 1 \). By relabeling we can assume that the subsequence is the whole sequence.

Let \( a_n = \gamma_{q,\delta,n}''(\{s_n\}) \) and \( \zeta_n \) be such that \( \gamma_{q,\delta,n}'' = a_n\delta_{s_n} + \zeta_n \). From Step 7 and \[25\] it follows that \( |a_n| > 2 \text{length}(\gamma_{q,\delta,n}) \sin 0.97\alpha. \)
From Lemma 6 it follows that along a subsequence \( \gamma'_{q,\delta,n} \to \gamma'_{q,\delta} \) in \( L^1 \) and \( \gamma''_{q,\delta,n} \rightharpoonup \gamma''_{q,\delta} \) in the space of signed measures as \( n \to \infty \). Along a further subsequence \( a_n \to a \) as \( n \to \infty \). By relabeling we can assume that the subsequence is the whole sequence. The \( L^1 \) convergence of gradients implies that \( \text{length}(\gamma_{q,\delta,n}) \to \text{length}(\gamma_{q,\delta}) \) as \( n \to \infty \), and thus \( |a| \geq 2\text{length}(\gamma_{q,\delta})\sin 0.97\alpha \). Furthermore \( a_n\delta_{\tilde{s}_n} \rightharpoonup a\delta_{\tilde{s}} \) as \( n \to \infty \). Consequently \( \zeta_n \rightharpoonup \zeta \) for some vector of measures \( \zeta \).

Let \( r > 0 \) be small enough so that \( [\tilde{s} - r, \tilde{s} + r] \subset (s_1, s_2) \). Then for all \( n \) large enough, (26) implies \( |\zeta_n|([\tilde{s} - r, \tilde{s} + r]) < \text{length}(\gamma_{q,\delta,n}) \times 0.01\alpha \). Therefore \( |\zeta|([\tilde{s} - r, \tilde{s} + r]) \leq \text{length}(\gamma_{q,\delta}) \times 0.01\alpha \).

Consequently \( |\gamma''_{q,\delta}(\{\tilde{s}\})| \geq |a| - |\zeta(\{\tilde{s}\})| > \text{length}(\gamma_{q,\delta})(2\sin(0.97\alpha) - 0.01\alpha) > 2(L - 3\epsilon)\frac{3}{2}\sin \alpha > 3\sin \alpha = 1 \). Therefore thus \( |\gamma''_{q,\delta}| \) has an atom of size at least 1 at \( \tilde{s} \).

\[ \square \]

### 5. The Constrained Problem

We now consider the original average-distance problem introduced in [3]. The task is to minimize

\[ (27) \quad F(\Sigma) = \int_{\mathbb{R}^d} d(x, \Sigma) d\mu(x) \quad \text{over} \quad \mathcal{A}_1 := \{ \Sigma \in \mathcal{A} : \mathcal{H}^1(\Sigma) \leq 1 \}. \]

5.1. **Construction of the counterexample.** The existence of a measure \( \mu \) for which the minimizer (27) is a Lipschitz curve which has a corner follows from Theorem 12.

**Corollary 13.** There exists \( r > 0, \delta > 0 \) and \( 1 > q > 0 \) such that there is a minimizer of (27) for \( \mu = D_{r,\delta} \) which is a Lipschitz curve such that if one considers its arclength parameterization \( \gamma : [0, 1] \to \mathbb{R}^d \) then \( \gamma' \) is an \( \mathbb{R}^d \) valued BV function such that \( \gamma'' \) has an atom of size at least 1/3 at some \( \tilde{s} \in (0, 1) \). More precisely \( |\gamma''(\{\tilde{s}\})| \geq 1/3 \).

**Proof.** Let \( \lambda = 0.36 \) and let \( \mu_{q,\delta} \) be as in the proof of Theorem 12. Let \( r = \text{length}(\gamma_{q,\delta}) \). Thus \( \mathcal{H}^1(\frac{1}{\epsilon} \Sigma_{q,\delta}) = 1 \) By the scaling discussed in Section 2.1, \( \frac{1}{\epsilon} \Sigma_{q,\delta} \) is a minimizer of \( E_{D_{r,\mu_{q,\delta}}} \). These facts imply that \( \frac{1}{\epsilon} \Sigma_{q,\delta} \) is a minimizer of (27). We claim that \( r < 3 \). The conclusion then follows from properties of \( \Sigma_{q,\delta} \) established in the proof of Theorem 12.

To prove that \( r < 3 \) note that for energy (1) corresponding to \( \mu_{q,\delta} \) we have \( E(\Sigma_{q,\delta}) \leq E(\Sigma) \) since \( \Sigma_{q,\delta} \) is a minimizer. By assumption (C2)

\[ \int d(x, \Sigma_{q,\delta}) d\mu_{q,\delta} \geq \int d(x, \Sigma) - \epsilon d\mu_{q,\delta} = \int d(x, \Sigma) d\mu_{q,\delta} - \epsilon. \]

Using (22), \( E(\Sigma_{q,\delta}) \leq E(\Sigma) \) implies

\[ r = \text{length}(\gamma_{q,\delta}) \leq \mathcal{H}^1(\Sigma) + \frac{\epsilon}{\lambda} = 2 \cdot \frac{3}{2\sqrt{2}} + 0.01 < 3. \]

\[ \square \]
Acknowledgement. The author would like to thank Thomas Laurent, with whom he started studying average-distance type functionals, for many insights and ideas. He is also thankful to Giovanni Leoni for his help and numerous valuable suggestions and observations and to Xin Yang Lu and Edoardo Mainini for careful reading of an early version and many useful comments. Furthermore the author acknowledges the support by NSF grants DMS-0908415 and DMS-1211760. He is also grateful to the Center for Nonlinear Analysis (NSF grant DMS-0635983) for its support. The research was also supported by NSF PIRE grant OISE-0967140.

REFERENCES


