



**HAL**  
open science

# One-dimensional approximation of measures in Wasserstein distance

Antonin Chambolle, Vincent Duval, Joao Miguel Machado

► **To cite this version:**

Antonin Chambolle, Vincent Duval, Joao Miguel Machado. One-dimensional approximation of measures in Wasserstein distance. 2024. hal-04082932v3

**HAL Id: hal-04082932**

**<https://hal.science/hal-04082932v3>**

Preprint submitted on 15 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ONE-DIMENSIONAL APPROXIMATION OF MEASURES IN WASSERSTEIN DISTANCE

ANTONIN CHAMBOLLE, VINCENT DUVAL, AND JOÃO MIGUEL MACHADO

ABSTRACT. We propose a variational approach to approximate measures with measures uniformly distributed over a 1-dimensional set. The problem consists in minimizing a Wasserstein distance as a data term with a regularization given by the length of the support. As it is challenging to prove existence of solutions to this problem, we propose a relaxed formulation, which always admits a solution. In the sequel we show that, under some assumption on the original measure, a solution to the relaxed problem is solution to the original one. Finally we prove that, whenever the original measure has a density in  $L^{\frac{d}{d-1}}(\mathbb{R}^d)$ , any optimal solution is supported by an Ahlfors regular set.

## 1. INTRODUCTION

In this paper we study the following 1-dimensional (1D) shape optimization problem: given a reference probability measure  $\rho_0 \in \mathcal{P}_p(\mathbb{R}^d)$  (the set of Borel probability measures  $\rho$  with  $\int_{\mathbb{R}^d} |x|^p d\rho < +\infty$ ,  $p \geq 1$ ), we seek to approximate  $\rho_0$  with measures uniformly supported on a one-dimensional connected subset of  $\mathbb{R}^d$ . This approximation is done by means of the following variational problem

$$(P_\Lambda) \quad \inf_{\Sigma \in \mathcal{A}} W_p^p(\rho_0, \nu_\Sigma) + \Lambda \mathcal{H}^1(\Sigma),$$

where the measure  $\nu_\Sigma$  is defined as

$$(1.1) \quad \nu_\Sigma \stackrel{\text{def.}}{=} \frac{1}{\mathcal{H}^1(\Sigma)} \mathcal{H}^1 \llcorner \Sigma, \text{ for } \Sigma \in \mathcal{A} \stackrel{\text{def.}}{=} \left\{ \Sigma \subset \mathbb{R}^d : \begin{array}{l} 0 < \mathcal{H}^1(\Sigma) < +\infty \\ \text{compact, connected.} \end{array} \right\},$$

and  $\mathcal{H}^1$  denotes the 1-dimensional Hausdorff measure in  $\mathbb{R}^d$ . The term  $W_p$  denotes the usual Wasserstein distance on the space of probability measures (see [30, 32] and Section 2.1.3).

One can trace the idea of approximating a probability measure by a 1D set back to the concept of *principal curves* from the seminal paper [16], which extends linear regression to regression using general curves, and introduces a variational problem to define such curves. In this variational sense, a principal curve minimizes the expectation of the distance to the curve, w.r.t. a probability measure describing a data set (with some regularization to ensure existence). As proposed in [17], a length constraint is a simple and intrinsic way to ensure existence. The properties of such minimizers have been studied in detail in e.g. [20, 12].

---

*Date:* October 15, 2024.

The authors are members of the Inria MOKAPLAN team. V.D. gratefully acknowledges support from the Agence Nationale de la Recherche (CIPRESSI, ANR-19-CE48-0017-01).

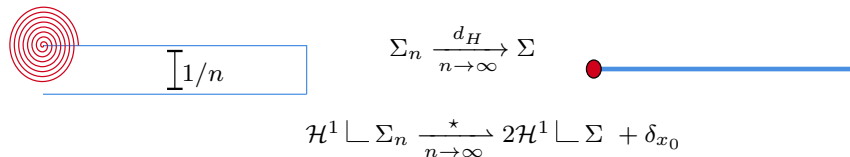


FIGURE 1. Concentration effects for the weak convergence of measures. Here  $\Sigma_n$  is made of two lines getting closer and a spiral converging to a point. In the Hausdorff limit we obtain a segment with shorter length, and a singular limiting measure.

A further generalization consists in replacing the curve with a more general one-dimensional compact and connected set, yielding the *average distance minimizer problem* introduced in [7], and its dual counterpart *maximum distance minimizer problem* [27, 19]. Such problems were conceived for applications in urban planning, where one seeks to minimize the average distance to a transportation network, giving rise to the need for a larger class of 1D sets allowing for bifurcations.

While the above-mentioned problems only focus on some geometric approximation of the support of the measure, approximating a measure in the sense of weak convergence is sometimes more desirable. In [18, 8], the authors have proposed optimal transport based methods for the projection of probability measures onto classes of measures supported on simple curves, using the Wasserstein distance as a data term. Potential applications range from 3D printing to image compression and reconstruction. In [13], the data fidelity term is chosen to be a discrepancy, see also [26]. The advantage of using discrepancies is that approximation rates can be given independently from the dimension, being therefore a good alternative to overcome the curse of dimensionality. The problem we study is an attempt to generalize this class of problems to the approximation with one-dimensional connected sets.

One difficulty when studying  $(P_\Lambda)$  is that the class of measures  $\nu_\Sigma$  is not closed in the usual weak topologies considered for the space of probability measures. While a sequence of sets  $(\Sigma_n)_{n \in \mathbb{N}}$  in  $\mathcal{A}$  with uniformly bounded length will have subsequences converging (in the Hausdorff sense) either to a point or a set in  $\mathcal{A}$ , the corresponding measures  $\nu_{\Sigma_n}$  might converge to a measure which is not necessarily uniform on that set: longer parts of  $\Sigma_n$  might concentrate in the limit on shorter parts of  $\Sigma$ , see Figure 1.

Hence, minimizing sequences converge in general to a measure which is not of the form  $\nu_\Sigma$ , and we need to determine a relaxation of our energy in a topology for which the Wasserstein distance is lower semi-continuous, such as the narrow convergence. The relaxed energy takes the form

$$(\overline{P}_\Lambda) \quad \inf_{\nu \in \mathcal{P}_p(\mathbb{R}^d)} W_p^p(\rho_0, \nu) + \Lambda \mathcal{L}(\nu),$$

where the length functional  $\mathcal{L}$ , defined in Section 3.1, generalizes the notion of length of the support of a measure, see for instance Example 3.6. We will show later on, in Proposition 3.8, that  $\mathcal{L}$  is the lower semi-continuous relaxation, for the narrow topology, of the functional  $\ell$  given by  $\mathcal{H}^1(\Sigma)$  for measures of the form  $\nu_\Sigma$ ,

and  $+\infty$  else, see (3.1). We also find that  $\mathcal{L}(\nu) < \infty$  if and only if  $\text{supp } \nu \in \mathcal{A}$  or  $\nu$  is a Dirac mass. The following Theorem gathers the various results proved throughout this paper.

**Theorem 1.1.** *Let  $\rho_0 \in \mathcal{P}_p(\mathbb{R}^d)$ ,  $\Lambda > 0$ . Then  $(\overline{P}_\Lambda)$  admits a solution  $\nu$ , and there exists  $\Lambda_\star \geq 0$  such that if  $\Lambda > \Lambda_\star$ ,  $\nu$  is a Dirac mass. For  $\Lambda < \Lambda_\star$ ,  $\nu$  is supported by a set  $\Sigma \in \mathcal{A}$  and the following properties hold.*

- (1) *If  $\rho_0$  is absolutely continuous w.r.t.  $\mathcal{H}^1$ , or has a  $L^\infty$  density w.r.t.  $\mathcal{H}^1$ , then so does  $\nu$ .*
- (2) *If  $\rho_0$  does not give mass to 1D-rectifiable sets, then  $\nu = \nu_\Sigma$  and therefore is a solution to the original problem  $(P_\Lambda)$ .*
- (3) *If  $\rho_0 \in L^{\frac{d}{d-1}}(\mathbb{R}^d)$ , then  $\Sigma$  is Ahlfors regular, i.e. there is  $r_0$  depending on  $d, p, \rho_0$  and  $\mathcal{L}(\nu)$  and  $C$  depending only on  $d, p$  such that for any  $x \in \Sigma$  and  $r \leq r_0$  it holds that*

$$r \leq \mathcal{H}^1(\Sigma \cap B_r(x)) \leq Cr.$$

The paper is organized as follows: in Section 2 we recall a few tools from optimal transport and geometric measure theory. Next, in Section 3 we go through the definition of the length functional and its properties as well as the relaxed problem and the existence of a solution. In Section 4 we discuss the existence of  $\Lambda_\star$ . In Section 5 (Theorem 5.4) we prove point (1) from Theorem 1.1, while the existence is proved in Section 6 (Theorem 6.4), and the Ahlfors regularity is studied in Section 7.

## 2. PRELIMINARIES

We start by introducing notions of convergence for sets and measures which will be useful to study problem  $(P_\Lambda)$  as well as the relaxed one  $(\overline{P}_\Lambda)$ . Next we describe some instrumental properties of the objects we shall use throughout the paper, namely the rectifiable sets and measures.

### 2.1. Convergence of sets and measures.

**2.1.1. Hausdorff and Kuratowski convergence.** We recall some useful definitions of convergence for sets, see for instance [29, Chap. 4], [3, Chap. 6].

The Hausdorff distance between two sets  $A, B$  is defined as:

$$(2.1) \quad d_H(A, B) \stackrel{\text{def.}}{=} \max \left\{ \sup_{a \in A} \text{dist}(a, B), \sup_{b \in B} \text{dist}(b, A) \right\},$$

where  $\text{dist}(\cdot, A)$  denotes the distance function to the set  $A$ . A sequence  $(A_n)_{n \in \mathbb{N}}$  of closed subsets of  $\mathbb{R}^d$  converges in the Hausdorff sense to  $A$  if  $\lim_{n \rightarrow \infty} d_H(A_n, A) = 0$ , and we write  $A_n \xrightarrow[n \rightarrow \infty]{d_H} A$ . One can prove that this notion of convergence is equivalent to uniform convergence of the distance functions. Since the latter are all 1-Lipschitz, as a consequence of Arzela-Ascoli's Theorem it follows that if the sequence is contained in a compact set, one can always extract a convergent subsequence. This compactness result is known as *Blaschke's Theorem*, see [3, Theorem 6.1].

A sequence of closed sets  $C_n$  converges in the sense of Kuratowski to  $C$ , and we write  $C_n \xrightarrow[n \rightarrow \infty]{K} C$ , whenever the two properties hold:

- (1) Given a sequence  $x_n \in C_n$ , all its cluster points are contained in  $C$ .
- (2) For all points  $x \in C$  there exists a sequence  $x_n \in C_n$ , converging to  $x$ .

Again, one can show that  $C_n \rightarrow C$  in the sense of Kuratowski if and only if  $\text{dist}(x, C_n) \rightarrow \text{dist}(x, C)$  (possibly infinite if  $C = \emptyset$ ) locally uniformly (see [29, Cor. 4.7]). In addition, the Bolzano-Weierstrass property holds for the Kuratowski convergence as well, *i.e.* any sequence of closed sets has a subsequence which converges, possibly to the empty set.

It is classical that Hausdorff and Kuratowski convergences coincide on sequences on uniformly bounded compact sets, the next lemma describes a more subtle relationship. We prove it in Appendix B.

**Lemma 2.1.** *Let  $(C_n)_{n \in \mathbb{N}}$  be a sequence of closed sets in  $\mathbb{R}^d$ , converging to  $C$  in the sense of Kuratowski. Then, for any  $x \in \mathbb{R}^d$ ,*

$$C_n \cap \overline{B_R(x)} \xrightarrow[n \rightarrow \infty]{d_H} C \cap \overline{B_R(x)},$$

for every radius  $R > 0$  such that  $\overline{C \cap B_R(x)} = C \cap \overline{B_R(x)}$ . Moreover, that condition holds for all  $R > 0$  except in a countable set.

**2.1.2. Convergence of measures.** Given a Borel set  $X \subset \mathbb{R}^d$ , we denote by  $\mathcal{M}(X)$  (resp.  $\mathcal{M}_+(X)$ ) the collection of the finite (resp. finite positive) Radon measures on  $X$ . The space of Borel probability measures on  $X$  is denoted by  $\mathcal{P}(X)$ , and  $\mathcal{P}_p(X)$  refers to its subset of probability measures with finite  $p$ -moment ( $p \geq 1$ ).

Following [3], we say that a sequence  $(\mu_n)_{n \in \mathbb{N}}$  of Radon measures on  $\mathbb{R}^d$  locally weakly- $\star$  converges to some Radon measure  $\mu$ , if, for every continuous function with compact support  $\phi \in C_c(\mathbb{R}^d)$ ,

$$(2.2) \quad \int_{\mathbb{R}^d} \phi d\mu_n \xrightarrow[n \rightarrow \infty]{} \int_{\mathbb{R}^d} \phi d\mu.$$

Any sequence  $(\mu_n)_{n \in \mathbb{N}}$  of Radon measures such that  $\sup_{n \in \mathbb{N}} |\mu_n|(K) < +\infty$  for every compact set  $K \subset \mathbb{R}^d$  has a locally weakly- $\star$  convergent subsequence.

If  $\{\mu_n\}_{n \in \mathbb{N}} \subset \mathcal{M}(\mathbb{R}^d)$  is a sequence of *finite* Radon measures and (2.2) holds for every bounded continuous function  $\phi \in C_b(\mathbb{R}^d)$ , we say that  $\mu_n$  narrowly converges to  $\mu$ , and we write  $\mu_n \xrightarrow[n \rightarrow \infty]{} \mu$ . When  $(\mu_n)_{n \in \mathbb{N}}$  is a sequence of probability measures, that convergence is often referred to as the *weak convergence* of probability measures.

If  $X$  is compact, any sequence of probability measures  $\{\mu_n\}_{n \in \mathbb{N}} \subset \mathcal{P}(X)$  has a weakly convergent subsequence. More generally, if  $X$  is not compact, compactness for the narrow convergence requires the assumptions of Prokhorov's Theorem, see [2, Theorem 2.8].

**2.1.3. Optimal transport and the Wasserstein distance.** The Wasserstein distances  $W_p$  are defined through the value function of an optimal transport problem, see

[4, 30, 32] for details. Given two probability measures  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ , we set

$$(2.3) \quad W_p^p(\mu, \nu) \stackrel{\text{def.}}{=} \min_{\gamma \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma,$$

where  $\Pi(\mu, \nu) \stackrel{\text{def.}}{=} \{\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : \pi_{0\sharp}\gamma = \mu, \pi_{1\sharp}\gamma = \nu\}$  is the space of transport couplings, and  $\pi_i$  denote the projections, *i.e.*  $\pi_0(x, y) = x$  and  $\pi_1(x, y) = y$ . Whenever  $\mu$  does not have atoms, the value of (2.3) coincides with

$$(2.4) \quad \inf_{T\sharp\mu = \nu} \int_{\mathbb{R}^d} |x - T(x)|^p d\mu,$$

where the inf is taken over all measurable maps  $T$  such that  $T\sharp\mu(A) = \nu(A) = \mu(T^{-1}(A))$  for any Borel set  $A$ .

The optimal transport problem can be analogously defined for any pair of positive  $\mu, \nu$  in  $\mathcal{M}_+(\mathbb{R}^d)$ . In this case, the Wasserstein distance becomes a 1-homogeneous functional and is finite if and only if the measures have finite  $p$ -moments and the same total mass  $\mu(\mathbb{R}^d) = \nu(\mathbb{R}^d)$ .

The Wasserstein distance is l.s.c. with respect to the narrow convergence, and continuous in a compact domain, [32, Lemma 4.3].

**2.2. Gołab's Theorem.** We now study the lower semicontinuity of  $(P_\Lambda)$ . “Gołab's Theorem” [15] shows that along sequences of connected sets, the length is lower semi-continuous with respect to the Hausdorff convergence [24, Chapter 10]. It is of course also true if the sequence has a uniformly bounded number of connected components.

The issue is that the compactness of Hausdorff convergence is not transferred to the weak convergence of measures of the form  $\mathcal{H}^1 \llcorner \Sigma$  which may concentrate in the limit. In general, one can prove the following:

**Theorem 2.2** (Density version of Gołab's Theorem). *Let  $(\Sigma_n)_{n \in \mathbb{N}}$  be a sequence of closed and connected subsets of  $\mathbb{R}^d$  converging in the sense of Kuratowski to some closed set  $\Sigma$  and having locally uniformly finite length, *i.e.* for all  $R > 0$*

$$\sup_{n \in \mathbb{N}} \mathcal{H}^1(\Sigma_n \cap B_R(x_0)) < +\infty.$$

*Define the measures  $\mu_n \stackrel{\text{def.}}{=} \mathcal{H}^1 \llcorner \Sigma_n$ , and let  $\mu$  be a local weak- $\star$  cluster point of this sequence. Then  $\text{supp } \mu \subset \Sigma$  and it holds that*

$$\mu \geq \mathcal{H}^1 \llcorner \Sigma,$$

*in the sense of measures.*

Such a result is hidden in the proof in [5] of the usual thesis of Gołab's Theorem, see also [28]. Yet in this variant, we consider a Kuratowski convergence and do not restrict the sets to be uniformly bounded, or have bounded lengths. This is useful for the proof of Theorem 6.1 where we consider sequences of blow-ups of sets. The proof of Theorem 2.2 is given in Appendix B.

**2.3. Rectifiable sets and measures.** We now introduce the notions of *rectifiable sets* and *rectifiable measure*, which will be crucial for understanding the fine properties of the elements of  $\mathcal{A}$ .

**Definition 2.3.** Let  $M \subset \mathbb{R}^d$  be a Borel set and  $k \in \mathbb{N}$ , we say that  $M$  is countably  $\mathcal{H}^k$ -rectifiable, or shortly  $k$  rectifiable, if there are countably many Lipschitz functions  $f_i : \mathbb{R}^k \rightarrow \mathbb{R}^d$  such that

$$\mathcal{H}^k \left( M \setminus \bigcup_{i \in \mathbb{N}} f_i(\mathbb{R}^k) \right) = 0.$$

A Radon measure  $\mu$  is said to be  $k$ -rectifiable if it is supported over a  $k$ -rectifiable set and  $\mu \ll \mathcal{H}^k$ .

In the simple case  $M = f(E)$ , for  $E \subset \mathbb{R}^k$ , one can define the tangent space at a point of differentiability of  $f$  as

$$\nabla f(z)(\mathbb{R}^k), \text{ for } x = f(z).$$

This is a parametric definition that can be extended to  $k$ -rectifiable sets. It turns out the parametric notion of tangentiability can be expressed in terms of measure theory. Given a Borel set  $M$ , we set the measure  $\mu = \mathcal{H}^k \llcorner M$ , and we consider the family of blow-up measures

$$(2.5) \quad \mu_r \stackrel{\text{def.}}{=} r^{-k} \Phi_{\sharp}^{x,r} \mu = \mathcal{H}^k \llcorner \left( \frac{M-x}{r} \right), \text{ for } \Phi^{x,r} \stackrel{\text{def.}}{=} \frac{\text{id} - x}{r}.$$

If  $M$  is countably  $\mathcal{H}^k$ -rectifiable, and  $\mathcal{H}^k(M \cap K) < +\infty$  for every compact set  $K$ , we say that  $M$  is locally  $\mathcal{H}^k$ -rectifiable, and then the blow-up Theorem, see [23, Theorem 10.2], states that for  $\mathcal{H}^k$ -a.e.  $x \in M$  this family of measures converges in the weak- $\star$  topology to a measure of the form  $\mathcal{H}^k \llcorner \pi_x$ , for a unique  $k$ -plane  $\pi_x \in G(k, d)$ , the Grassmannian of  $k$ -planes of  $\mathbb{R}^d$ .

More generally define the  $k$ -density, whenever the limit exists, of a Radon measure  $\mu$  as

$$(2.6) \quad \theta_k(\mu, x) \stackrel{\text{def.}}{=} \lim_{r \rightarrow 0^+} \frac{\mu(B_r(x))}{\omega_k r^k} \text{ and } \theta_k(M, x) \stackrel{\text{def.}}{=} \theta_k(\mathcal{H}^k \llcorner M, x),$$

where  $\omega_k$  is the volume of the unit  $k$ -dimensional ball, see [3, 23]. A direct consequence of the blow-up Theorem is that  $\mathcal{H}^k$ -a.e. point of a  $k$ -rectifiable set has  $k$ -density 1. Analogously for a  $k$ -rectifiable measure  $\mu$  it holds that  $\mu = \theta_k(\mu, x) \mathcal{H}^k \llcorner M$ .

The equivalence between all notions was completed with the work of Preiss and the notion of a tangent space to a measure, see for instance the monograph [11]. If a measure (*resp.* a set) has a finite  $k$ -density, *i.e.* the limit in (2.6) exists and is finite  $\mathcal{H}^k$ -a.e., then this measure (*resp.* set) is  $k$ -rectifiable. The previous discussion is summarized in the following theorem.

**Theorem 2.4.** *Let  $\mu$  be a Radon measure over  $\mathbb{R}^d$ , the following are equivalent.*

- (i)  $\mu$  is  $k$ -rectifiable

(ii) For  $\mathcal{H}^k$ -a.e.  $x \in \text{supp } \mu$ , it holds that

$$r^{-k} \Phi_{\sharp}^{x,r} \mu \xrightarrow[r \rightarrow 0]{*} \theta_k(\mu, x) \mathcal{H}^k \llcorner \pi_x,$$

for a unique  $k$ -plane  $\pi_x \in G(k, d)$ .

(iii) For  $\mathcal{H}^k$ -a.e.  $x \in \text{supp } \mu$ , the  $k$ -density of  $\mu$  in (2.6) exists and is finite.

In the previous Theorem, if we take  $\mu = \mathcal{H}^k \llcorner M$  where  $M$  is a countably  $\mathcal{H}^k$ -rectifiable set we define the *approximate tangent space* of  $M$  at  $x$  as  $T_x M \stackrel{\text{def.}}{=} \pi_x$ , where  $\pi_x$  is the unique  $k$ -plane from point (ii).

**Definition 2.5.** Let  $M \subset \mathbb{R}^d$  be a  $k$ -rectifiable set. We say that  $x \in M$  is a *rectifiability point* when the locally weakly- $\star$  convergence of point (ii) from Theorem 2.4 holds, with  $\mu = \mathcal{H}^k \llcorner M$  and  $\theta_k(\mu, x) = 1$ .

Now we pass to our case of interest, of compact connected sets  $\Sigma$  with finite length,  $\mathcal{H}^1(\Sigma) < +\infty$ , in view of (1.1). From [10, Prop. 30.1 and Cor. 30.2] or [1, Thm. 4.4], any compact connected set with finite length is also path-connected, *i.e.* any pair of its points can be joined by a continuous arc.

Such sets are also known to be 1-rectifiable, see [5, Thm. 4.4.8], and hence they enjoy the properties of Theorem 2.4. In the next Lemma, we show that the blow-up of some  $\Sigma \in \mathcal{A}$  around a rectifiability point is precisely its approximate tangent space.

**Lemma 2.6.** *Given  $\Sigma \in \mathcal{A}$ , then for  $\mathcal{H}^1$ -a.e.  $y \in \Sigma$ , it holds that*

$$\frac{\Sigma - y}{r} \xrightarrow[r \rightarrow 0^+]{K} T_y \Sigma \quad \text{and} \quad \frac{\Sigma - y}{r} \cap \overline{B_R(0)} \xrightarrow[r \rightarrow 0^+]{d_H} T_y \Sigma \cap \overline{B_R(0)}, \quad \text{for all } R > 0.$$

*Proof.* First we take a rectifiability point  $y \in \Sigma$  with tangent space  $T_y \Sigma$ , by Theorem 2.4 such points cover  $\mathcal{H}^1$ -a.a. of  $\Sigma$ . In particular, point (ii) of the Theorem shows that  $\mathcal{H}^1 \llcorner ((\Sigma - y)/r) \xrightarrow[r \rightarrow 0]{*} \mathcal{H}^1 \llcorner T_y \Sigma$ . Let  $T$  be the (Kuratowski) limit of a subsequence  $(\Sigma - y)/r_k$ . Clearly, the limit measure  $\mathcal{H}^1 \llcorner T_y \Sigma$  is supported by  $T$ , hence  $T_y \Sigma \subset T$ . Thanks to Lemma 2.1 and Theorem 2.2, for almost all  $R > 0$ ,

$$(2.7) \quad \mathcal{H}^1(T \cap B_R) \leq \liminf_{k \rightarrow \infty} \mathcal{H}^1 \left( \frac{\Sigma - y}{r_k} \cap B_R \right) = \mathcal{H}^1(T_y \Sigma \cap B_R),$$

which shows that up to a  $\mathcal{H}^1$ -negligible set,  $T = T_y \Sigma$ .

Notice that, if there is some  $x \in T \setminus T_y \Sigma$ , we may consider some ball  $B_s(x)$  which does not intersect  $T_y \Sigma$ . Since  $T$  is the limit of connected sets,  $x$  must be path-connected in  $T$  to some point in  $(B_s(x))^c$ , so that  $\mathcal{H}^1(T \cap B_s(x)) \geq s$ . This contradicts (2.7). Hence  $T = T_y \Sigma$ , and is independent on the subsequence, and we deduce that  $(\Sigma - y)/r \xrightarrow{K} T_y \Sigma$ . The convergence in the Hausdorff distance follows from Lemma 2.1.  $\square$

### 3. THE LENGTH FUNCTIONAL AND THE RELAXED PROBLEM

If a minimizing sequence  $\Sigma_n$  converges to some set  $\Sigma$ , we cannot expect weak cluster points of the measures  $\nu_{\Sigma_n}$  to have the form  $\nu_{\Sigma}$ , see Figure 1. Hence the objective of  $(P_{\Lambda})$  is not lower semi-continuous for the narrow convergence, and, in



this section, we introduce a relaxation for  $(P_\Lambda)$ . First, we define a functional which extends the length of the support and we discuss some of its properties, then we use it to define the relaxed problem.

**3.1. Definition and elementary properties.** Recalling that  $\mathcal{A}$  is the collection of the compact connected sets  $\Sigma \subset \mathbb{R}^d$  with  $0 < \mathcal{H}^1(\Sigma) < +\infty$ , we consider

$$(3.1) \quad \ell : \mathcal{P}(\mathbb{R}^d) \ni \nu \mapsto \begin{cases} \mathcal{H}^1(\Sigma), & \text{if } \nu = \frac{1}{\mathcal{H}^1(\Sigma)} \mathcal{H}^1 \llcorner \Sigma \text{ for some } \Sigma \in \mathcal{A}, \\ +\infty, & \text{otherwise,} \end{cases}$$

so that  $(P_\Lambda)$  becomes  $\inf_\nu W_p^p(\rho_0, \nu) + \Lambda \ell(\nu)$ . As discussed above,  $\ell$  is not l.s.c., hence we introduce the following relaxation, which we call the *length functional*. For any  $\nu \in \mathcal{P}(\mathbb{R}^d)$ , we define

$$(3.2) \quad \mathcal{L}(\nu) \stackrel{\text{def.}}{=} \begin{cases} \inf \{ \alpha \geq 0 \mid \alpha \nu \geq \mathcal{H}^1 \llcorner \text{supp } \nu \}, & \text{if supp } \nu \text{ is connected,} \\ +\infty, & \text{otherwise,} \end{cases}$$

with the convention that  $\inf \emptyset \stackrel{\text{def.}}{=} +\infty$ . Notice that, since  $\nu$  is a probability measure,  $\mathcal{L}(\nu) \geq \mathcal{H}^1(\text{supp } \nu)$ , and that  $\mathcal{L}(\nu) = 0$  if and only if  $\nu = \delta_x$  for some  $x \in \mathbb{R}^d$ . As a result,  $0 < \mathcal{L}(\nu) < \infty$  if and only if  $\text{supp } \nu \in \mathcal{A}$ . Moreover, for any  $\Sigma \in \mathcal{A}$  and  $\nu_\Sigma \stackrel{\text{def.}}{=} \frac{1}{\mathcal{H}^1(\Sigma)} \mathcal{H}^1 \llcorner \Sigma$ , we have  $\mathcal{L}(\nu_\Sigma) = \mathcal{H}^1(\Sigma) = \ell(\nu_\Sigma)$ .

*Remark 3.1.* Definition (3.2) also makes sense for any positive measure  $\mu \in \mathcal{M}_+(\mathbb{R}^d)$ . In that case, thanks to Theorem 2.2, it may be easily shown to be lower semi-continuous with respect to the weak convergence, defining  $\mathcal{L}(0) = 0$  (see also Section 3.2). Yet then, of course, even for uniformly distributed measures such as  $\nu = \theta \mathcal{H}^1 \llcorner \Sigma$  for some  $\theta > 0$ , its value does not coincide with the length of the support anymore (it rather is  $\mathcal{H}^1(\Sigma)/\nu(\mathbb{R}^d)$ ).

In Section 3.2 below, we prove that  $\mathcal{L}$  is the lower semi-continuous envelope of  $\ell$  for the narrow topology of probability measures. Before that, let us discuss some alternative formulations for  $\mathcal{L}$ . Following [3, Sec. 2.4], we consider the upper derivative,

$$(3.3) \quad \forall x \in \text{supp } \nu, \quad D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)(x) \stackrel{\text{def.}}{=} \limsup_{r \rightarrow 0^+} \frac{\mathcal{H}^1(B_r(x) \cap \text{supp } \nu)}{\nu(B_r(x))}.$$

**Proposition 3.2** (Alternative definitions of  $\mathcal{L}$ ). *Let  $\nu \in \mathcal{P}(\mathbb{R}^d)$  be such that  $\text{supp } \nu$  is connected. Then*

$$(3.4) \quad \mathcal{L}(\nu) = \sup \left\{ \frac{\mathcal{H}^1(U \cap \text{supp } \nu)}{\nu(U)} \mid U \text{ open, } U \cap \text{supp } \nu \neq \emptyset \right\}$$

$$(3.5) \quad = \sup \left\{ \frac{\mathcal{H}^1(B_r(x) \cap \text{supp } \nu)}{\nu(B_r(x))} \mid r > 0, x \in \text{supp } \nu \right\}$$

$$(3.6) \quad = \|D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)\|_\infty,$$

where  $\|\cdot\|_\infty$  denotes the supremum norm over  $\text{supp } \nu$ .

*Proof.* It is immediate that

$$(\text{R.H.S. of (3.2)}) \geq (\text{R.H.S. of (3.4)}) \geq (\text{R.H.S. of (3.5)}) \geq (\text{R.H.S. of (3.6)}).$$

Now, assume that  $\|D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)\|_\infty < +\infty$  and let  $\alpha > \|D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)\|_\infty$ . For every compact set  $K \subset \mathbb{R}^d$  and every  $x \in K \cap (\text{supp } \nu)$ , there is some  $r(x) > 0$  such that  $\mathcal{H}^1(B_r(x) \cap (\text{supp } \nu)) \leq \alpha \nu(B_r(x))$ . We may extract from the covering  $(B_{r(x)}(x))_{x \in K \cap (\text{supp } \nu)}$  with open sets a finite covering  $(B_{r_i}(x_i))_{i=1}^N$  of  $K \cap (\text{supp } \nu)$ . As a result

$$\mathcal{H}^1(K \cap (\text{supp } \nu)) \leq \sum_{i=1}^N \alpha \nu(B_{r_i}(x_i)) \leq N\alpha < +\infty,$$

so that  $\mathcal{H}^1 \llcorner (\text{supp } \nu)$  is a Radon measure. We may thus apply [3, Prop. 2.21] to deduce

$$(\text{R.H.S. of (3.6)}) \geq (\text{R.H.S. of (3.2)}).$$

If  $\|D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)\|_\infty = +\infty$ , the inequality holds trivially, which completes the proof.  $\square$

The length functional inherits some of the properties of the  $\mathcal{H}^1$  measure.

**Proposition 3.3.** *Let  $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$ , be a  $k$ -Lipschitz function, with  $k > 0$ . Then*

$$(3.7) \quad \forall \nu \in \mathcal{P}(\mathbb{R}^d), \quad \mathcal{L}(f_\# \nu) \leq k \mathcal{L}(\nu).$$

*Proof.* If  $\mathcal{L}(\nu) = +\infty$ , there is nothing to prove. Otherwise,  $\text{supp } \nu$  is compact, and  $\text{supp}(f_\# \nu) = f(\text{supp } \nu)$ . Moreover, for any open set  $U \subset \mathbb{R}^d$ , since  $f^{-1}(U)$  is open,

$$U \cap (\text{supp } f_\# \nu) \neq \emptyset \iff \nu(f^{-1}(U)) > 0 \iff f^{-1}(U) \cap (\text{supp } \nu) \neq \emptyset.$$

Now, let  $U$  be an open set which intersects  $\text{supp}(f_\# \nu)$ . Using that

$$U \cap f(\text{supp } \nu) \subset f(f^{-1}(U) \cap \text{supp } \nu)$$

we get

$$\begin{aligned} \frac{\mathcal{H}^1(U \cap \text{supp}(f_\# \nu))}{f_\# \nu(U)} &= \frac{\mathcal{H}^1(U \cap f(\text{supp } \nu))}{\nu(f^{-1}(U))} \leq \frac{\mathcal{H}^1(f(f^{-1}(U) \cap \text{supp } \nu))}{\nu(f^{-1}(U))} \\ &\leq k \frac{\mathcal{H}^1(f^{-1}(U) \cap \text{supp } \nu)}{\nu(f^{-1}(U))} \\ &\leq k \mathcal{L}(\nu) \end{aligned}$$

since  $f^{-1}(U)$  is an open set which intersects  $\text{supp } \nu$ . Taking the supremum over all  $U$  yields the claimed inequality.  $\square$

It is also possible to express the length-functional using the Besicovitch differentiation theorem [3, Thm. 2.22]. Assume that  $\mathcal{H}^1(\text{supp } \nu) < +\infty$  (otherwise  $\mathcal{L}(\nu) = +\infty$ ). Then, the measure  $\mathcal{H}^1 \llcorner \text{supp } \nu$  is Radon, and the limit

$$(3.8) \quad D_\nu(\mathcal{H}^1 \llcorner \text{supp } \nu)(x) \stackrel{\text{def.}}{=} \lim_{r \rightarrow 0^+} \frac{\mathcal{H}^1(B_r(x) \cap \text{supp } \nu)}{\nu(B_r(x))}$$

$$(3.9) \quad \left( \text{resp. } D_{\mathcal{H}^1 \llcorner \text{supp } \nu}(\nu)(x) \stackrel{\text{def.}}{=} \lim_{r \rightarrow 0^+} \frac{\nu(B_r(x))}{\mathcal{H}^1(B_r(x) \cap \text{supp } \nu)} \right)$$

exists for  $\nu$ -a.e.  $x$  (resp.  $\mathcal{H}^1 \llcorner \text{supp } \nu$ -a.e.  $x$ ).

**Proposition 3.4** (Alternative definitions, II). *Let  $\nu \in \mathcal{P}(\mathbb{R}^d)$  such that  $\text{supp } \nu$  is connected and  $\mathcal{H}^1(\text{supp } \nu) < +\infty$ . Then*

$$(3.10) \quad \mathcal{L}(\nu) = \begin{cases} \left\| \frac{d(\mathcal{H}^1 \llcorner \text{supp } \nu)}{d\nu} \right\|_{L^\infty} & \text{if } (\mathcal{H}^1 \llcorner \text{supp } \nu) \ll \nu, \\ +\infty & \text{otherwise.} \end{cases}$$

$$(3.11) \quad = \begin{cases} 0 & \text{if } \text{supp } \nu \text{ is a singleton,} \\ \left\| \left( \frac{d\nu}{d(\mathcal{H}^1 \llcorner \text{supp } \nu)} \right)^{-1} \right\|_{L^\infty_{\mathcal{H}^1 \llcorner \text{supp } \nu}} & \text{otherwise.} \end{cases}$$

Notice that in Proposition 3.4, both “norms” may take the value  $+\infty$ , and in (3.11), we adopt the convention that  $1/0 = +\infty$ .

*Proof of Proposition 3.4.* First, we prove (3.10). If  $(\mathcal{H}^1 \llcorner \text{supp } \nu) \ll \nu$  then the Lebesgue-Besicovitch differentiation theorem ensures that

$$\mathcal{H}^1 \llcorner \text{supp } \nu = \left( \frac{d(\mathcal{H}^1 \llcorner \text{supp } \nu)}{d\nu} \right) \nu \leq \left\| \frac{d(\mathcal{H}^1 \llcorner \text{supp } \nu)}{d\nu} \right\|_{L^\infty} \nu.$$

Therefore,

$$\mathcal{L}(\nu) \leq \left\| \frac{d(\mathcal{H}^1 \llcorner \text{supp } \nu)}{d\nu} \right\|_{L^\infty} \leq \|D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)\|_\infty = \mathcal{L}(\nu).$$

If  $(\mathcal{H}^1 \llcorner \text{supp } \nu)$  is not absolutely continuous w.r.t.  $\nu$ , there is no  $\alpha > 0$  such that  $\alpha\nu \geq \mathcal{H}^1 \llcorner \text{supp } \nu$ , and  $\mathcal{L}(\nu) = +\infty$ .

Now, we prove (3.11). The case where  $\text{supp } \nu$  is a singleton is already known. We assume now that  $\mathcal{H}^1(\text{supp } \nu) > 0$ , and using the Besicovitch differentiation theorem [3, Thm. 2.22], we decompose

$$(3.12) \quad \nu = \theta \mathcal{H}^1 \llcorner \text{supp } \nu + \nu^s,$$

where

$$\theta(x) \stackrel{\text{def.}}{=} \frac{d\nu}{d(\mathcal{H}^1 \llcorner \text{supp } \nu)}(x) = \lim_{r \rightarrow 0^+} \frac{\nu(B_r(x))}{\mathcal{H}^1(B_r(x) \cap \text{supp } \nu)} = (D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)(x))^{-1}$$

for  $(\mathcal{H}^1 \llcorner \text{supp } \nu)$ -a.e.  $x$ . From the last equality, we get

$$\|\theta^{-1}\|_{L^\infty_{\mathcal{H}^1 \llcorner \text{supp } \nu}} \leq \|D_\nu^+(\mathcal{H}^1 \llcorner \text{supp } \nu)(x)\|_\infty = \mathcal{L}(\nu).$$

To prove the converse inequality, we assume  $\|\theta^{-1}\|_{L^\infty_{\mathcal{H}^1 \llcorner \text{supp } \nu}} < +\infty$  (otherwise there is nothing to prove). Using (3.12), we note that

$$\left( \|\theta^{-1}\|_{L^\infty_{\mathcal{H}^1 \llcorner \text{supp } \nu}} \right) \nu \geq \mathcal{H}^1 \llcorner \text{supp } \nu,$$

so that  $\mathcal{L}(\nu) \leq \|\theta^{-1}\|_{L^\infty_{\mathcal{H}^1 \llcorner \text{supp } \nu}}$ .  $\square$

We may now examine a few examples.

**Example 3.5.** Let  $\nu = \sum_{n=1}^{\infty} 2^{-n} \delta_{q_n}$ , where  $(q_n)_{n \geq 1}$  is a dense sequence in  $[0, 1]$ . The support being the set of points  $x$  such that  $\nu(B_r(x)) > 0$  for all  $r > 0$ , one has  $\text{supp } \nu = [0, 1]$  which is connected. However, using (3.2), we see that  $\mathcal{L}(\nu) = +\infty$ .

**Example 3.6** (Densities on a  $(\mathcal{H}^1, 1)$ -rectifiable set). Let  $\Sigma \subseteq \mathbb{R}^d$  be a closed connected set with  $0 < \mathcal{H}^1(\Sigma) < +\infty$ ,  $\theta: \Sigma \rightarrow \mathbb{R}_+$  a Borel function such that  $\int_{\Sigma} \theta d\mathcal{H}^1 < 1$ , and let  $\nu = \theta \mathcal{H}^1 \llcorner \Sigma + \nu^s$  be a probability measure, where  $\text{supp } \nu^s \subset \Sigma$  and the measures  $\nu^s$  and  $\mathcal{H}^1 \llcorner \Sigma$  are mutually singular. Then  $\mathcal{L}(\nu) = \|1/\theta\|_{L^\infty_{\mathcal{H}^1 \llcorner \Sigma}}$ : the length functional ignores the singular part.

**Example 3.7** (Parametrized Lipschitz curves). Let  $\gamma: [0, 1] \rightarrow \mathbb{R}^d$  be a non-constant Lipschitz curve, and let  $\nu$  such that for all  $f \in C_b(\mathbb{R}^d)$ ,

$$\langle f, \nu \rangle \stackrel{\text{def.}}{=} \frac{1}{\text{len}(\gamma)} \left( \int_0^1 f(\gamma(t)) |\dot{\gamma}(t)| dt \right), \quad \text{where } \text{len}(\gamma) \stackrel{\text{def.}}{=} \int_0^1 |\dot{\gamma}(t)| dt$$

is the length of the curve. By the area formula [14, Thm. 3.2.5],

$$d\nu(y) = \frac{1}{\text{len}(\gamma)} \text{card}(\gamma^{(-1)}(y)) d(\mathcal{H}^1 \llcorner \Sigma)(y)$$

where  $\Sigma = \gamma([0, 1])$ . As a result,

$$(3.13) \quad \mathcal{L}(\nu) = \frac{\text{len}(\gamma)}{\text{ess-min}_{y \in \Sigma} (\text{card}(\gamma^{(-1)}(y)))},$$

where the minimum is an *essential minimum* with respect to  $\mathcal{H}^1 \llcorner \Sigma$ .

**3.2. Lower semi-continuity of the length functional.** Now, we prove that  $\mathcal{L}$  is the lower semi-continuous envelope of  $\ell$  for the narrow convergence.

**Proposition 3.8.** *The functional  $\mathcal{L}$  is the lower semi-continuous envelope of  $\ell$  for the narrow topology. Moreover, for every  $\nu$  such that  $\mathcal{L}(\nu) < +\infty$ ,*

$$(3.14) \quad \mathcal{H}^1(\text{supp } \nu) \leq \mathcal{L}(\nu)$$

with equality if and only if either  $\nu = \delta_x$  for some  $x \in \mathbb{R}^d$  (if  $\mathcal{H}^1(\text{supp } \nu) = 0$ ), or  $\mathcal{H}^1(\text{supp } \nu) > 0$  and  $\nu = \frac{1}{\mathcal{H}^1(\text{supp } \nu)} \mathcal{H}^1 \llcorner \text{supp } \nu$ , i.e.  $\nu = \nu_\Sigma$  for some  $\Sigma \in \mathcal{A}$ , as defined in (1.1).

*Proof of Proposition 3.8:* The inequality (3.14) is clear from the definition of (3.2), so we study the equality case.

If  $\nu = \delta_x$  or  $\nu = \frac{1}{\mathcal{H}^1(\text{supp } \nu)} \mathcal{H}^1 \llcorner \text{supp } \nu$  with  $\mathcal{H}^1(\text{supp } \nu) > 0$ , one readily checks that  $\mathcal{L}(\nu) = \mathcal{H}^1(\text{supp } \nu)$ . Conversely, if (3.14) is an equality, for every Borel set  $B$ ,

$$\begin{aligned} 0 &= \mathcal{L}(\nu) - \mathcal{H}^1(\text{supp } \nu) \\ &= \underbrace{(\mathcal{L}(\nu)\nu(B) - \mathcal{H}^1(B \cap \text{supp } \nu))}_{\geq 0} + \underbrace{(\mathcal{L}(\nu)\nu(B^c) - \mathcal{H}^1(B^c \cap \text{supp } \nu))}_{\geq 0} \end{aligned}$$

so that both terms must be zero. If  $\mathcal{L}(\nu) > 0$ , we deduce

$$\forall B \subset \mathbb{R}^d \text{ Borel, } \nu(B) = \frac{\mathcal{H}^1(B \cap \text{supp } \nu)}{\mathcal{L}(\nu)} = \frac{\mathcal{H}^1(B \cap \text{supp } \nu)}{\mathcal{H}^1(\text{supp } \nu)}.$$

If  $\mathcal{L}(\nu) = 0$ ,  $\mathcal{H}^1(\text{supp } \nu) = 0$  and since  $\text{supp } \nu$  is connected,  $\nu$  is a Dirac mass.

Next we prove that  $\mathcal{L}$  is sequentially lower semi-continuous. We consider  $(\nu_n)_{n \in \mathbb{N}}$  such that  $\nu_n \xrightarrow[n \rightarrow \infty]{} \nu \in \mathcal{P}(\mathbb{R}^d)$  and we show that  $\alpha \stackrel{\text{def.}}{=} \liminf_{n \rightarrow \infty} \mathcal{L}(\nu_n) \geq \mathcal{L}(\nu)$ . If  $\alpha = +\infty$ , we have nothing to prove. Otherwise, up to the extraction of a subsequence, we may assume that  $\lim_{n \rightarrow \infty} \mathcal{L}(\nu_n) = \alpha$  and that  $\mathcal{L}(\nu_n) < +\infty$  for all  $n \in \mathbb{N}$ .

Defining the sequence of compact and connected sets  $\Sigma_n \stackrel{\text{def.}}{=} \text{supp } \nu_n$ , it holds that  $\mathcal{H}^1(\Sigma_n) \leq \mathcal{L}(\nu_n)$ , so that

$$\sup_{n \geq N} \mathcal{H}^1(\Sigma_n) \leq \alpha + 1 < +\infty$$

for  $N$  large enough. Hence, for all  $n \geq N$ ,  $\text{diam}(\Sigma_n) \leq \alpha + 1$ . In addition, let  $x \in \text{supp } \nu$ . Since  $0 < \nu(B_1(x)) \leq \liminf_{n \rightarrow \infty} \nu_n(B_1(x))$ , for all  $n$  large enough  $(\text{supp } \nu_n) \cap B_1(x) \neq \emptyset$ , thus  $\text{supp } \nu_n \subset \overline{B_{\alpha+2}(x)}$ .

Therefore, we may apply Blaschke's Theorem and assume, up to extracting a subsequence, that  $\Sigma_n \xrightarrow[n \rightarrow \infty]{d_H} \Sigma$ . From the weak convergence of measures one has  $\text{supp } \nu \subset \Sigma$ . Let us show that  $\text{supp } \nu = \Sigma$ . If  $\Sigma$  is a singleton  $\{x_0\}$ , we have  $\nu = \delta_{x_0}$ . Otherwise, Theorem 2.2 implies that  $\Sigma \in \mathcal{A}$  and furthermore, as  $\mathcal{L}(\nu_n)\nu_n \geq \mathcal{H}^1 \llcorner \Sigma_n$ , that

$$(3.15) \quad \alpha \nu \geq \mathcal{H}^1 \llcorner \Sigma.$$

Hence, as  $\Sigma$  is connected, for all  $z \in \Sigma$  it holds  $\nu(B_r(z)) > 0$ , confirming that  $\text{supp } \nu = \Sigma$ . Finally from (3.15) we get that

$$\liminf_{n \rightarrow \infty} \mathcal{L}(\nu_n) = \alpha \geq \mathcal{L}(\nu),$$

proving that  $\mathcal{L}$  is l.s.c.

As a result, we have proved that  $\mathcal{L}$  is l.s.c. and that  $\mathcal{L} \equiv \ell$  on the effective domain of  $\ell$ . To show that  $\mathcal{L}$  is the l.s.c. envelope of  $\ell$ , we prove that it is above any l.s.c. functional  $\mathcal{G} \leq \ell$ . Let  $\nu \in \mathcal{P}(\mathbb{R}^d)$ . If  $\mathcal{L}(\nu) = +\infty$ , we have  $\mathcal{G}(\nu) \leq \mathcal{L}(\nu)$ . If  $\mathcal{L}(\nu) < +\infty$ , using Lemma 3.9 below, we can find a sequence  $\nu_{\Sigma_n} \xrightarrow[n \rightarrow \infty]{} \nu$  such that  $\mathcal{H}^1(\Sigma_n) \rightarrow \mathcal{L}(\nu)$ . The lower semi-continuity of  $\mathcal{G}$  yields

$$\mathcal{G}(\nu) \leq \liminf_{n \rightarrow \infty} \mathcal{G}(\nu_{\Sigma_n}) \leq \liminf_{n \rightarrow \infty} \ell(\nu_{\Sigma_n}) = \liminf_{n \rightarrow \infty} \mathcal{H}^1(\Sigma_n) = \mathcal{L}(\nu).$$

□

The proof of Proposition 3.8 relies on the following approximation Lemma.

**Lemma 3.9.** *Let  $\nu \in \mathcal{P}(\mathbb{R}^d)$  such that  $\mathcal{L}(\nu) < \infty$ . There exists a sequence  $(\Sigma_n)_{n \in \mathbb{N}} \subset \mathcal{A}$  such that*

- $\Sigma_n \xrightarrow[n \rightarrow \infty]{d_H} \text{supp } \nu$ ,

- $\nu_{\Sigma_n} \xrightarrow{n \rightarrow \infty} \nu$  and  $W_p(\nu_{\Sigma_n}, \nu) \xrightarrow{n \rightarrow \infty} 0$  for any  $p \geq 1$ , where  $\nu_{\Sigma_n}$  is defined as in (1.1).

We also have  $\mathcal{H}^1(\Sigma_n) \xrightarrow{n \rightarrow \infty} \mathcal{L}(\nu)$  and if, in addition  $\mathcal{L}(\nu) > 0$ , we can take  $\mathcal{H}^1(\Sigma_n) = \mathcal{L}(\nu)$  for all  $n \in \mathbb{N}$ .

*Proof.* To simplify the notation, we set  $\alpha = \mathcal{L}(\nu)$  and  $\Sigma = \text{supp } \nu$ . For  $\alpha = 0$  (that is,  $\nu = \delta_{x_0}$  for some  $x_0$ ), we consider

$$\Sigma_n = x_0 + [0, 1/n] \times \{0\}^{d-1}$$

which provides the desired approximation, with  $\mathcal{H}^1(\Sigma_n) = 1/n \rightarrow 0 = \mathcal{L}(\delta_{x_0})$ .

For  $\alpha > 0$ , we start by covering the space with cubes of the form

$$Q_{z,n} \stackrel{\text{def.}}{=} \frac{1}{n} (z + [0, 1]^d), \quad \text{for } z \in \mathbb{Z}^d.$$

For some fixed  $n$ , let  $(Q_{i,n})_{i \in I_n}$  be the collection of the cubes such that  $\nu(Q_{z,n}) > 0$ , since the set  $\Sigma$  is compact,  $I_n$  is finite for a given  $n$ . We define the quantities

$$m_{i,n} \stackrel{\text{def.}}{=} \alpha \nu(Q_{i,n}) - \mathcal{H}^1(\Sigma \cap Q_{i,n}) \leq \alpha,$$

as the excess mass of  $\nu$  in the cube  $Q_{i,n}$  (note that  $m_{i,n} \geq 0$  in view of (3.2)). Our strategy is to modify  $\nu \llcorner Q_{i,n}$  by adding segments with uniform measure inside the cube and having a total length equal to the excess mass  $m_{i,n}$ .

If  $\Sigma \cap \text{int } Q_{i,n} \neq \emptyset$ , take  $x_i$  in this intersection, so that  $B_{\delta_i}(x_i) \subset Q_{i,n}$  for some  $\delta_i > 0$ . Then, set  $N_{i,n} \stackrel{\text{def.}}{=} \left\lceil \frac{m_{i,n}}{\delta_i} \right\rceil$ , and choose  $\delta_{i,j} \geq 0$  for  $j = 1, \dots, N_{i,n}$  such that

$$\sum_{j=1}^{N_{i,n}} \delta_{i,j} = m_{i,n}, \quad \text{and } 0 \leq \delta_{i,j} < \delta_i.$$

Since  $\mathcal{H}^1(\Sigma \cap Q_{i,n}) < +\infty$ , it is possible to choose  $N_{i,n}$  vectors  $v_{i,j} \in \mathbb{S}^{d-1}$  such that the segments  $S_{i,j} \stackrel{\text{def.}}{=} [x_i, x_i + \delta_{i,j} v_{i,j}]$  are contained in  $\text{int } Q_{i,n}$  and satisfy  $\mathcal{H}^1(\Sigma \cap S_{i,j}) = 0$ , for  $j = 1, \dots, N_{i,n}$ .

If  $\Sigma \cap \text{int } Q_{i,n} = \emptyset$ , as the cubes have positive mass, it means that  $\nu$  is concentrated on the boundary of the cube, in which case we take  $x_i \in \Sigma \cap \partial Q_i$  and any family of segments entering the cube will suffice.

Next, we define the measures

$$\nu_{\Sigma_n} \stackrel{\text{def.}}{=} \frac{1}{\mathcal{H}^1(\Sigma_n)} \mathcal{H}^1 \llcorner \Sigma_n \text{ for } \Sigma_n \stackrel{\text{def.}}{=} \Sigma \cup \bigcup_{i \in I_n} \bigcup_{j=1}^{N_{i,n}} S_{i,j}.$$

From the construction, the Hausdorff distance between  $\Sigma$  and  $\Sigma_n$  is at most the diagonal of the cube  $[0, 1/n]^d$ , so that

$$d_H(\Sigma, \Sigma_n) \leq \frac{\sqrt{d}}{n} \xrightarrow{n \rightarrow \infty} 0,$$

and the total length of  $\Sigma_n$  is given by

$$\begin{aligned} \mathcal{H}^1(\Sigma_n) &= \sum_{i \in I_n} \mathcal{H}^1(\Sigma \cap Q_{i,n}) + \sum_{i \in I_n} \sum_{j=1}^{N_{i,n}} \mathcal{H}^1(S_{i,j}) \\ &= \sum_{i \in I_n} \mathcal{H}^1(\Sigma \cap Q_{i,n}) + m_{i,n} = \alpha \sum_{i \in I_n} \nu(Q_{i,n}) = \alpha. \end{aligned}$$

Each  $\Sigma_n \in \mathcal{A}$  since it is connected and compact (as a finite union of compact sets).

To finish the proof, it remains to show that  $\nu_{\Sigma_n} \xrightarrow[n \rightarrow \infty]{} \nu$ . By construction, there exists a compact set  $K \subset \mathbb{R}^d$  such that  $(\text{supp } \nu) \cup \bigcup_{n \geq 1} (\text{supp } \nu_{\Sigma_n}) \subset K$ . Then any function  $\phi \in C_b(\mathbb{R}^d)$  is uniformly continuous on  $K$ , and we denote by  $\omega$  its modulus of continuity. Observing that  $\nu_{\Sigma_n}(Q_{i,n}) = \nu(Q_{i,n})$ , we note that

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \phi d\nu_{\Sigma_n} - \int_{\mathbb{R}^d} \phi d\nu \right| &\leq \sum_{i \in I_n} \left| \int_{Q_{i,n}} \phi d\nu_{\Sigma_n} - \int_{Q_{i,n}} \phi d\nu \right| \\ &\leq \sum_{i \in I_n} \omega(\text{diam } Q_{i,n}) \nu(Q_{i,n}) \leq \omega(\sqrt{d}/n) \xrightarrow[n \rightarrow \infty]{} 0. \end{aligned}$$

Hence  $\nu_{\Sigma_n} \xrightarrow[n \rightarrow \infty]{} \nu$ . But as the support of all such measures is contained in the compact  $K$  and the Wasserstein distance metrizes the weak convergence in  $\mathcal{P}_p(K)$ , see [30, Thm. 5.10], it holds that  $W_p(\nu_{\Sigma_n}, \nu) \xrightarrow[n \rightarrow \infty]{} 0$ .  $\square$

*Remark 3.10.* The conclusions of Proposition 3.8 and Lemma 3.9 still hold when replacing the narrow topology with the local weak- $\star$  topology.

**3.3. A relaxed problem with existence of solutions.** The relaxed problem  $(\overline{P}_\Lambda)$  introduced on page 2 is defined by replacing  $\ell$  in the original problem with its l.s.c. envelope  $\mathcal{L}$ . We define the energy  $\mathcal{E}(\nu) \stackrel{\text{def.}}{=} W_p^p(\rho_0, \nu) + \Lambda \mathcal{L}(\nu)$ , and with a slight abuse of notation, we sometimes write  $\mathcal{E}(\Sigma) = \mathcal{E}(\nu_\Sigma)$  for  $\Sigma \in \mathcal{A}$ . The main point of considering this relaxed problem is that the existence of solutions for  $(\overline{P}_\Lambda)$  follows from the direct method of the calculus of variations.

**Theorem 3.11.** *The relaxed problem  $(\overline{P}_\Lambda)$  admits a solution. In addition,  $\mathcal{E}$  is the l.s.c. envelope of  $W_p^p(\rho_0, \cdot) + \Lambda \ell$ , and:*

$$\inf (P_\Lambda) = \min (\overline{P}_\Lambda).$$

*Proof.* Let  $(\nu_n)_{n \in \mathbb{N}}$  be a minimizing sequence for  $\mathcal{E}$ . Since  $(\sup_n W_p^p(\rho_0, \nu_n)) < +\infty$ , the moments of order  $p$  of  $\nu_n$  are uniformly bounded (see for instance [30, Thm. 5.11]), and we may then extract a (not relabeled) subsequence converging to some  $\nu \in \mathcal{P}(\mathbb{R}^d)$  in the narrow topology (by Prokhorov's theorem). From Proposition 3.8 and the fact that the Wasserstein distance is lower semi-continuous, the functional  $\mathcal{E}$  is l.s.c. and we have that

$$\mathcal{E}(\nu) \leq \liminf_{n \rightarrow \infty} \mathcal{E}(\nu_n) = \inf (\overline{P}_\Lambda).$$

The measure  $\nu$  is a minimizer of  $(\overline{P}_\Lambda)$ .

To show that  $\mathcal{E}$  is the l.s.c. envelope of the original energy one may argue as in the proof of Proposition 3.8. Consider any l.s.c. functional  $\mathcal{G}$  such that

$$\forall \nu \in \mathcal{P}(\mathbb{R}^d), \quad \mathcal{G}(\nu) \leq W_p^p(\rho_0, \nu) + \Lambda \ell(\nu).$$

For every  $\nu$  with  $\mathcal{L}(\nu) < +\infty$ , we use Lemma 3.9 to build a sequence  $(\nu_n)_{n \in \mathbb{N}}$  such that  $W_p^p(\rho_0, \nu_{\Sigma_n}) \rightarrow W_p^p(\rho_0, \nu)$ . Indeed, as  $\nu_{\Sigma_n}$  converges to  $\nu$  for the Wasserstein distance, the triangle inequality gives

$$|W_p(\rho_0, \nu_{\Sigma_n}) - W_p(\rho_0, \nu)| \leq W_p(\nu_{\Sigma_n}, \nu) \xrightarrow[n \rightarrow \infty]{} 0.$$

Hence for any  $\nu \in \mathcal{P}_p(\mathbb{R}^d)$  it holds that

$$\mathcal{G}(\nu) \leq \liminf_{n \rightarrow \infty} (W_p^p(\rho_0, \nu_{\Sigma_n}) + \Lambda \ell(\nu_{\Sigma_n})) = W_p^p(\rho_0, \nu) + \Lambda \mathcal{L}(\nu) = \mathcal{E}(\nu),$$

and we conclude that  $\mathcal{E}$  is the l.s.c. envelope.  $\square$

#### 4. ON THE SUPPORT OF OPTIMAL MEASURES

Our goal for this section is to answer to the question of “how small”  $\Lambda$  must be in Theorem 1.1. For this, in Theorem 4.1 we study when solutions of the relaxed problem  $(\overline{\mathcal{P}}_\Lambda)$  are Dirac masses. Keeping this in mind, the rest of this section can be skipped and the reader can move on to the main results of the paper.

The following notation will be useful: a point  $x_0$  is said to be a  $p$ -mean of  $\rho_0$  if

$$x_0 \in \arg \min_{y \in \mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^p d\rho_0(x) = \arg \min_{y \in \mathbb{R}^d} W_p(\rho_0, \delta_y).$$

A 2-mean is just the mean of  $\rho_0$ , that is,  $m_{\rho_0} \stackrel{\text{def.}}{=} \int_{\mathbb{R}^d} x d\rho_0(x)$ . For  $p > 1$ , the  $p$ -mean is uniquely defined, but for  $p = 1$  the collection of 1-means is a closed convex set which is not reduced to a singleton in general.

**Theorem 4.1.** *For a fixed measure  $\rho_0 \in \mathcal{P}_p(\mathbb{R}^d)$  there exists a critical parameter  $\Lambda_\star \in [0, \infty)$  such that*

- for  $\Lambda < \Lambda_\star$  no solution of  $(\mathcal{P}_\Lambda)$  is a Dirac measure;
- for  $\Lambda > \Lambda_\star$  it holds that  $\arg \min(\mathcal{P}_\Lambda)$  is the set of  $p$ -means of  $\rho_0$ .

Moreover,  $\Lambda_\star = 0$  if and only if  $\rho_0$  is a Dirac mass.

We start by studying the support of the optimal measure, showing that it is contained in the convex hull of the support of  $\rho_0$ . In the sequel the proof of Theorem 4.1 will be divided in several steps. We end the section with an exemple of  $\rho_0$  composed of 2 Dirac masses.

**4.1. Elementary properties of the support.** Given a set  $A \subset \mathbb{R}^d$  we denote by  $\overline{\text{conv}}A$  its closed convex hull.

**Lemma 4.2.** *Let  $\nu \in \mathcal{P}(\mathbb{R}^d)$  be a solution to  $(\overline{\mathcal{P}}_\Lambda)$ . Then the following properties hold*

- (1)  $\mathcal{H}^1(\text{supp } \nu) \leq \frac{1}{\Lambda} W_p^p(\rho_0, \delta_{m_{\rho_0}})$ , where  $m_{\rho_0}$  is any  $p$ -mean of  $\rho_0$ . In particular,  $\Sigma$  is contained in a ball of diameter  $d_0 \stackrel{\text{def.}}{=} \frac{1}{\Lambda} W_p^p(\rho_0, \delta_{m_{\rho_0}})$ .
- (2)  $\text{supp } \nu \subset \overline{\text{conv}}(\text{supp } \rho_0) \cap B(m_{\rho_0}, 2W_p(\rho_0, \delta_{m_{\rho_0}}) + \frac{2}{\Lambda} W_p^p(\rho_0, \delta_{m_{\rho_0}}))$



*Proof.* For the first point, let  $\Sigma$  denote the support of  $\nu$ . Since  $\nu$  has finite energy we have that  $\mathcal{L}(\nu) \geq \mathcal{H}^1(\Sigma)$ . Thus, since it is also optimal

$$(4.1) \quad \Lambda \mathcal{H}^1(\Sigma) \leq W_p^p(\rho_0, \nu) + \Lambda \mathcal{L}(\nu) \leq W_p^p(\rho_0, \delta_{m_{\rho_0}}) + \Lambda \mathcal{L}(\delta_{m_{\rho_0}}) = W_p^p(\rho_0, \delta_{m_{\rho_0}}).$$

For the second point, let  $C \stackrel{\text{def.}}{=} \overline{\text{conv}}(\text{supp } \rho_0)$ . It is a nonempty closed convex set, therefore the projection onto  $C$  is well-defined and 1-Lipschitz. We denote it by  $f$ . By Proposition 3.3, it holds that  $\mathcal{L}(\nu) \geq \mathcal{L}(f_{\#}\nu)$ . Moreover, for every  $(x, y) \in C \times \mathbb{R}^d$ ,

$$|x - y|^2 = |x - f(y)|^2 + |f(y) - y|^2 + 2 \underbrace{\langle x - f(y), f(y) - y \rangle}_{\geq 0} \geq |x - f(y)|^2$$

with equality if and only if  $y \in C$ . As a result, if  $\gamma$  is an optimal transport plan for  $(\rho_0, \nu)$ ,

$$\begin{aligned} W_p^p(\rho_0, \nu) &= \int |x - y|^p d\gamma(x, y) \geq \int |x - f(y)|^p d\gamma(x, y) \\ &= \int |x - y|^p d((\text{id}, f)_{\#}\gamma)(x, y) \geq W_p^p(\rho_0, f_{\#}\nu) \end{aligned}$$

with strict inequality unless  $y \in C$  for  $\gamma$ -a.e.  $(x, y)$  (hence  $\nu$ -a.e.  $y$ ).

But  $\nu$  is a solution to  $(\overline{\mathcal{P}}_{\Lambda})$ , therefore the inequality

$$W_p^p(\rho_0, \nu) + \Lambda \mathcal{L}(\nu) \geq W_p^p(\rho_0, f_{\#}\nu) + \Lambda \mathcal{L}(f_{\#}\nu)$$

cannot be strict. We deduce that  $y \in C$  for  $\nu$ -a.e.  $y$ , and  $C$  being closed, that  $\Sigma \subset C$ .

Additionally, from (4.1), we have  $W_p^p(\nu, \delta_{m_{\rho_0}}) \leq 2W_p^p(\rho_0, \delta_{m_{\rho_0}})$  and in particular there are points  $y \in \Sigma$  such that  $|y - m_{\rho_0}| \leq 2W_p^p(\rho_0, \delta_{m_{\rho_0}})$ . Combined with the first point, we obtain that  $\Sigma \subset B(m_{\rho_0}, 2W_p^p(\rho_0, \delta_{m_{\rho_0}}) + \frac{2}{\Lambda}W_p^p(\rho_0, \delta_{m_{\rho_0}}))$ .  $\square$

**Example 4.3.** Let  $\rho_0 = \delta_{x_0}$  for some  $x_0 \in \mathbb{R}^d$ . Then both conditions from Lemma 4.2 are sharp and characterize for all  $\Lambda > 0$  the unique solution  $\delta_{x_0}$  of  $(\overline{\mathcal{P}}_{\Lambda})$ .

**4.2. When solutions are Dirac masses.** Now, we discuss whether or not Dirac masses are solutions in the case where  $\rho_0$  is not a Dirac measure. We start with the following Lemma.

**Lemma 4.4.** *Let  $\Lambda > 0$  such that  $\delta_{x_0} \in \arg \min(\mathcal{P}_{\Lambda})$ , for  $\Lambda' > \Lambda$  it holds*

- for  $p > 1$  that  $\delta_{x_0}$  is the unique solution of  $(\mathcal{P}_{\Lambda'})$ ,
- for  $p = 1$  that  $\arg \min(\mathcal{P}_{\Lambda'})$  consists of only Dirac masses.

*Proof.* If  $\delta_{x_0} \in \arg \min(\mathcal{P}_{\Lambda})$ , for any  $p \geq 1$ , and for any measure  $\nu$  with  $\mathcal{L}(\nu) > 0$  it holds that

$$W_p^p(\rho_0, \delta_{x_0}) \leq W_p^p(\rho_0, \nu) + \Lambda \mathcal{L}(\nu) < W_p^p(\rho_0, \nu) + \Lambda' \mathcal{L}(\nu),$$

and hence  $\nu$  cannot be a minimizer of  $(\mathcal{P}_{\Lambda'})$ . Then for any  $p \geq 1$  it holds that  $\arg \min(\mathcal{P}_{\Lambda'})$  consists of Dirac measures. Whenever  $p > 1$ , the function  $y \mapsto W_p^p(\rho_0, \delta_y)$  is strictly convex and hence  $\arg \min(\mathcal{P}_{\Lambda'})$  is a singleton.  $\square$

This simple Lemma allows for the definition of the critical value  $\Lambda_\star$  as follows

$$(4.2) \quad \Lambda_\star \stackrel{\text{def.}}{=} \inf \{ \Lambda \geq 0 : \arg \min (\mathcal{P}_\Lambda) \subset (\delta_x)_{x \in \mathbb{R}^d} \}.$$

As stated in Theorem 4.1,  $\Lambda_\star > 0$  whenever  $\rho_0$  is not a single Dirac mass, which is a direct consequence of the convergence of solutions to  $\rho_0$  when  $\Lambda$  goes to 0.

**Lemma 4.5.** *For every  $\rho_0 \in \mathcal{P}_p(\mathbb{R}^d)$ , and  $\Lambda > 0$ , let  $\nu_\Lambda$  be any solution to  $(\bar{P}_\Lambda)$ . Then*

$$\nu_\Lambda \xrightarrow{\Lambda \rightarrow 0^+} \rho_0.$$

*In particular,  $\Lambda_\star > 0$  unless  $\rho_0$  is a Dirac mass.*

*Proof.* If  $\mathcal{L}(\rho_0) < +\infty$ , it suffices to notice that

$$W_p^p(\rho_0, \nu_\Lambda) \leq W_p^p(\rho_0, \nu_\Lambda) + \Lambda \mathcal{L}(\nu_\Lambda) \leq W_p^p(\rho_0, \rho_0) + \Lambda \mathcal{L}(\rho_0) = \Lambda \mathcal{L}(\rho_0) \xrightarrow{\Lambda \rightarrow 0^+} 0.$$

However, we need to handle the case where  $\mathcal{L}(\rho_0) = +\infty$ .

Let  $\varepsilon > 0$ . By the density of discrete measures in the Wasserstein space, there exists a probability measure of the form  $\mu = \sum_{i=1}^N a_i \delta_{x_i}$  such that  $W_p^p(\rho_0, \mu) \leq \varepsilon$ . We may assume that  $N \geq 2$ . By connecting all the points  $\{x_i\}_{1 \leq i \leq N}$ , we obtain a compact connected set  $\Sigma$  with  $0 < \mathcal{H}^1(\Sigma) < +\infty$ . For every  $\theta \in ]0, 1[$ , we then define

$$\tilde{\rho}_0 \stackrel{\text{def.}}{=} \frac{\theta}{\mathcal{H}^1(\Sigma)} \mathcal{H}^1 \llcorner \Sigma + (1 - \theta)\mu = \theta \nu_\Sigma + (1 - \theta)\mu.$$

and we note that  $\mathcal{L}(\tilde{\rho}_0) \leq \frac{\mathcal{H}^1(\Sigma)}{\theta} < +\infty$ .

By the optimality of  $\nu_\Lambda$ ,

$$W_p^p(\rho_0, \nu_\Lambda) \leq \Lambda \mathcal{L}(\nu_\Lambda) + W_p^p(\rho_0, \nu_\Lambda) \leq \Lambda \mathcal{L}(\tilde{\rho}_0) + W_p^p(\rho_0, \tilde{\rho}_0).$$

Taking the upper limit as  $\Lambda \rightarrow 0^+$ , and using the convexity of the Wasserstein distance yields

$$\limsup_{\Lambda \rightarrow 0^+} W_p^p(\rho_0, \nu_\Lambda) \leq W_p^p(\rho_0, \tilde{\rho}_0) \leq \theta W_p^p(\rho_0, \nu_\Sigma) + (1 - \theta) W_p^p(\rho_0, \mu).$$

Letting  $\theta \rightarrow 0^+$  we obtain  $\limsup_{\Lambda \rightarrow 0^+} W_p^p(\rho_0, \nu_\Lambda) \leq \varepsilon$ . Since  $\varepsilon$  is arbitrary, the claim follows.

For the last statement, we note that  $\text{supp } \rho_0$  must be included in the Kuratowski limits of  $\text{supp } \nu_\Lambda$  as  $\Lambda \rightarrow 0$ , so that if  $\rho_0$  is not a Dirac mass, neither is  $\nu_\Lambda$  for  $\Lambda > 0$  small enough.  $\square$

Next, we show that for  $\Lambda$  large enough, the solution becomes a Dirac measure.

**Proposition 4.6.** *For every  $\rho_0 \in \mathcal{P}_p(\mathbb{R}^d)$ ,  $\Lambda_\star < +\infty$ .*

*Proof.* Choose  $\nu \in \arg \min (\bar{P}_\Lambda)$ , let  $\Sigma \stackrel{\text{def.}}{=} \text{supp } \nu$  and  $y_0 \in \Sigma$ .

Let  $r \stackrel{\text{def.}}{=} \min \{ r' \geq 0 \mid \Sigma \subset B(y_0, r') \}$ : since  $\Sigma$  is connected one has  $r \leq \mathcal{H}^1(\Sigma) < +\infty$ . The convexity of the  $p$ -norm yields

$$\forall x, y \in \mathbb{R}^d, \quad |x - y|^p \geq |x - y_0|^p - p |x - y_0|^{p-1} |y - y_0|.$$

As a result, if  $\gamma$  is an optimal transport plan for  $(\rho_0, \nu)$ ,

$$\begin{aligned} \mathcal{E}(\nu) &= \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \, d\gamma(x, y) + \Lambda \mathcal{L}(\nu) \\ &\geq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y_0|^p \, d\gamma(x, y) - p \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y_0|^{p-1} |y - y_0| \, d\gamma(x, y) + \Lambda \mathcal{H}^1(\Sigma) \\ &\geq \mathcal{E}(\delta_{y_0}) + r \left( \Lambda - p \int_{\mathbb{R}^d} |x - y_0|^{p-1} \, d\rho_0(x) \right). \end{aligned}$$

By optimality of  $\nu$ , we have  $\mathcal{E}(\nu) \leq \mathcal{E}(\delta_{y_0})$ , so that  $r = 0$  and  $\nu$  is a Dirac mass as soon as  $\left( \Lambda - p \int_{\mathbb{R}^d} |x - y_0|^{p-1} \, d\rho_0(x) \right) > 0$ .

On the other hand as soon as  $r > 0$ , this expression must be negative, and it follows that

$$\Lambda_\star \leq p \int_{\mathbb{R}^d} |x - y_0|^{p-1} \, d\rho_0(x).$$

Note that this bound depends on  $\nu$  (through  $\Sigma$ ) and therefore also on  $\Lambda$ . Yet, as observed in the proof of Lemma 4.2, point (2), we can choose  $y_0 \in \Sigma$  with  $|y_0 - m_{\rho_0}| \leq 2W_p(\delta_{m_{\rho_0}}, \rho_0)$ . It follows that

$$\Lambda_\star \leq \max_{y_0 \in B(m_{\rho_0}, 2W_p(\delta_{m_{\rho_0}}, \rho_0))} \left( p \int_{\mathbb{R}^d} |x - y_0|^{p-1} \, d\rho_0(x) \right),$$

which is a (pessimistic) a priori bound depending only on  $\rho_0$ .  $\square$

*Remark 4.7.* In some cases, it is possible to provide sharper bounds on  $\Lambda_\star$ :

- If  $p = 1$ , we see that  $\Lambda_\star \leq 1$ .
- If  $p = 2$ , it can be shown by a simple translation argument that  $\nu$  and  $\rho_0$  have the same barycenter. Then, one may adapt the above argument to get  $\Lambda_\star \leq 2 \int |x - m_{\rho_0}| \, d\rho_0(x)$ , where  $m_{\rho_0} = \int x \, d\rho_0(x)$ .

**4.3. The example of an input with two Dirac masses.** In this subsection we consider the case  $p = 2$ . Let  $x_{-1} = (-1, 0, \dots, 0)$ ,  $x_1 = (1, 0, \dots, 0) \in \mathbb{R}^d$ , and let  $\rho_0 = \frac{1}{2}(\delta_{x_{-1}} + \delta_{x_1})$ . By Lemma 4.2, we know that the solutions to  $(\bar{P}_\Lambda)$  are supported on line segments which are contained in  $[x_{-1}, x_1]$ . We may thus reduce the problem to the one-dimensional setting, with  $x_{-1} = -1$ ,  $x_1 = 1$ . The solution to that problem is given by the following proposition.

**Proposition 4.8.** *For  $p = 2$  and  $\rho_0 = \frac{1}{2}(\delta_{-1} + \delta_1)$ , the unique solution to  $(\bar{P}_\Lambda)$  is given by*

$$(4.3) \quad \nu_\Lambda = \begin{cases} \sqrt{\frac{3\Lambda}{2}} \mathcal{H}^1 \llcorner [-1, 1] + \left(\frac{1}{2} - \sqrt{\frac{3\Lambda}{2}}\right) (\delta_{-1} + \delta_1) & \text{if } 0 < \Lambda < \frac{1}{6}, \\ \frac{1}{3(1-2\Lambda)} \mathcal{H}^1 \llcorner \left[-\frac{3}{2}(1-2\Lambda), \frac{3}{2}(1-2\Lambda)\right] & \text{if } \frac{1}{6} \leq \Lambda < \frac{1}{2}. \\ \delta_0 & \text{if } \Lambda \geq \frac{1}{2}. \end{cases}$$

*Proof.* We fix  $\Lambda > 0$  and denote  $\nu$  a solution. Let  $\alpha = \mathcal{L}(\nu)$ . If  $\alpha = 0$ ,  $\nu$  is a Dirac mass. If  $\alpha > 0$ , we know that the support of  $\nu$  is a connected subset of  $\overline{\text{conv}}\{-1, 1\} = [-1, 1]$ , so that  $\text{supp } \nu = [a, b]$  for  $-1 \leq a < b \leq 1$ . In addition,

letting  $c \in [a, b]$  such that  $\nu([a, c]) \leq 1/2$  and  $\nu([a, c]) \geq 1/2$ , one can check that if some mass is sent from  $\{-1\}$  to  $]c, b]$ , then exchanging it with the same amount of mass sent from  $\{+1\}$  to  $[a, c[$  we reduce the Wasserstein distance. Hence one may assume that the mass coming from  $\{-1\}$  is sent to a measure  $\nu^-$  supported on  $[a, c]$  while the mass from  $\{+1\}$  is sent to a measure  $\nu^+$  supported on  $[c, b]$ , with  $\nu^- + \nu^+ = \nu$ . Observing that  $\nu \geq \frac{1}{\alpha} \mathcal{H}^1 \llcorner [a, b]$  (we are in the case  $\alpha > 0$ ), we introduce the non-negative excess measures:

$$\nu_{\text{exc}}^- = \nu^- - \frac{1}{\alpha} \mathcal{H}^1 \llcorner [a, c], \quad \nu_{\text{exc}}^+ = \nu^+ - \frac{1}{\alpha} \mathcal{H}^1 \llcorner [c, b],$$

and  $\nu_{\text{exc}} = \nu_{\text{exc}}^- + \nu_{\text{exc}}^+$ . Once more, we see that the Wasserstein distance is reduced if all the mass sent from  $\{-1\}$  to  $\nu_{\text{exc}}^-$  is sent to the point  $\{a\}$ , closest to  $\{-1\}$ . Hence, we may assume that  $\nu_{\text{exc}}^- = x\delta_a$ , for  $x \geq 0$ , and similarly,  $\nu_{\text{exc}}^+ = y\delta_b$ , for  $y \geq 0$ . Eventually, we easily see that if  $a > -1$  and  $x > 0$ , then we can extend the segment  $[a, b]$  towards  $\{-1\}$ , adding a small piece  $[a - \delta, \delta]$  for  $\delta \leq \min\{\alpha x, a + 1\}$ , send a fraction  $\delta/\alpha$  of the measure  $x\delta_a$  rather to  $\frac{1}{\alpha} \mathcal{H}^1 \llcorner [a - \delta, a]$ , and reduce again the Wasserstein distance without changing  $\mathcal{L}(\nu)$ . We deduce that  $x = 0$  if  $a > -1$ , similarly  $y = 0$  if  $b < 1$ .

Recalling that for  $p = 2$ ,  $\nu$  must have the same center of mass as  $\rho_0$ , we deduce that  $\nu$  must be equal to

$$\begin{aligned} \nu_{0,0} &\stackrel{\text{def.}}{=} \delta_0, \\ \text{or } \nu_{b,2b} &\stackrel{\text{def.}}{=} \frac{1}{2b} \mathcal{H}^1 \llcorner [-b, b] \quad \text{for some } b \in ]0, 1[ \\ \text{or } \nu_{1,\alpha} &= \frac{1}{\alpha} \mathcal{H}^1 \llcorner [-1, 1] + \left(\frac{1}{2} - \frac{1}{\alpha}\right) (\delta_{-1} + \delta_1) \quad \text{for some } \alpha \geq 2. \end{aligned}$$

Let  $\mathcal{E}(\nu) = \Lambda \mathcal{L}(\nu) + W_2^2(\rho_0, \nu)$  denote the energy to minimize. We have  $\mathcal{E}(\nu_{0,0}) = 1 = \lim_{b \rightarrow 0^+} \mathcal{E}(\nu_{b,2b})$ , and

$$\begin{aligned} \mathcal{E}(\nu_{b,2b}) &= 2\Lambda b + 2 \int_0^b (1-x)^2 \frac{dx}{2b} = \frac{b^2}{3} + (2\Lambda - 1)b + 1 \\ \text{with } \frac{d}{db} \mathcal{E}(\nu_{b,2b}) &= \frac{2b}{3} + 2\Lambda - 1, \\ \mathcal{E}(\nu_{1,\alpha}) &= \Lambda\alpha + 2 \int_0^1 (1-x)^2 \frac{dx}{\alpha} + 0 = \Lambda\alpha + \frac{2}{3\alpha}, \\ \text{with } \frac{d}{d\alpha} \mathcal{E}(\nu_{1,\alpha}) &= \Lambda - \frac{2}{3\alpha^2}. \end{aligned}$$

For  $0 < \Lambda < \frac{1}{6}$ , we check that  $\nu_{1,\alpha^*}$ , for  $\alpha^* \stackrel{\text{def.}}{=} \sqrt{\frac{2}{3\Lambda}}$ , is the unique solution.

For  $\frac{1}{6} \leq \Lambda < \frac{1}{2}$ , we get that  $\nu_{b^*,2b^*}$  is the unique solution, with  $b^* \stackrel{\text{def.}}{=} \frac{3}{2}(1 - 2\Lambda)$ .

For  $\Lambda \geq \frac{1}{2}$ , the functions  $\alpha \mapsto \mathcal{E}(\nu_{1,\alpha})$  and  $b \mapsto \mathcal{E}(\nu_{b,2b})$  are strictly increasing on  $[2, +\infty[$  and  $]0, 1]$  respectively. Therefore  $\nu_{0,0}$  is the unique solution to  $(\overline{P}_\Lambda)$ .  $\square$

## 5. SOLUTIONS ARE RECTIFIABLE MEASURES

Our goal here is to show that whenever  $\rho_0 \ll \mathcal{H}^1$ , any solution  $\nu$  is a rectifiable measure of the form

$$\nu = \theta \mathcal{H}^1 \llcorner \Sigma, \quad \text{for } \theta \in L^1(\Sigma; \mathcal{H}^1).$$

To this end, we introduce the excess measure  $\nu_{\text{exc}}$  as the positive measure given by the mass of  $\nu$  that exceeds the density constraints. We first show that this measure solves a family of localized problems. This is used to prove the absolute continuity w.r.t.  $\mathcal{H}^1 \llcorner \Sigma$ , that is, point (1) of Theorem 1.1.

**5.1. The excess measure.** Let  $\nu$  be a minimizer of  $(\bar{P}_\Lambda)$  with support  $\Sigma$  not reduced to a singleton. From the definition of the length functional we have:

$$\mathcal{L}(\nu) < \infty \text{ if and only if there is } \alpha \geq 0 \text{ such that } \alpha\nu \geq \mathcal{H}^1 \llcorner \Sigma.$$

Setting  $\alpha \stackrel{\text{def.}}{=} \mathcal{L}(\nu) > 0$ , we define the following decomposition

$$(5.1) \quad \nu = \nu_{\mathcal{H}^1} + \nu_{\text{exc}}, \quad \text{where } \nu_{\mathcal{H}^1} \stackrel{\text{def.}}{=} \alpha^{-1} \mathcal{H}^1 \llcorner \Sigma \text{ and } \nu_{\text{exc}} \stackrel{\text{def.}}{=} \nu - \nu_{\mathcal{H}^1}.$$

The part  $\nu_{\mathcal{H}^1}$  is the measure which saturates the density constraint, and the support of the *excess measure*  $\nu_{\text{exc}}$  is where the constraint is inactive.

In the sequel, we fix an optimal transport plan  $\gamma$ , for the problem defining  $W_p^p(\rho_0, \nu)$ , and we define an analogous (non-unique) decomposition of  $\gamma$  and  $\rho_0$  by disintegrating  $\gamma$  w.r.t. the second marginal. From the disintegration theorem [3, Theorem 2.28], there exists a  $\nu$ -measurable family  $\{\gamma_y\}_{y \in \mathbb{R}^d} \subset \mathcal{P}(\mathbb{R}^d)$ , such that  $\gamma = \gamma_y \otimes \nu$ , that is

$$(5.2) \quad \int_{\mathbb{R}^d \times \Sigma} \psi(x, y) d\gamma(x, y) = \int_{\Sigma} \left( \int_{\mathbb{R}^d} \psi(x, y) d\gamma_y(x) \right) d\nu(y), \quad \text{for all } \psi \in L^1(\gamma).$$

We define a decomposition  $\gamma = \gamma_{\mathcal{H}^1} + \gamma_{\text{exc}}$  as

$$(5.3) \quad \gamma_{\mathcal{H}^1}(A \times B) \stackrel{\text{def.}}{=} \int_{\Sigma \cap B} \gamma_y(A) d\nu_{\mathcal{H}^1}(y), \quad \gamma_{\text{exc}}(A \times B) \stackrel{\text{def.}}{=} \int_{\Sigma \cap B} \gamma_y(A) d\nu_{\text{exc}}(y).$$

The decomposition  $\rho_0 = \rho_{\mathcal{H}^1} + \rho_{\text{exc}}$  can be defined as the marginals of  $\gamma_{\mathcal{H}^1}$  and  $\gamma_{\text{exc}}$

$$(5.4) \quad \rho_{\mathcal{H}^1} \stackrel{\text{def.}}{=} (\pi_0)_\# \gamma_{\mathcal{H}^1}, \quad \rho_{\text{exc}} \stackrel{\text{def.}}{=} (\pi_0)_\# \gamma_{\text{exc}}.$$

This way  $\gamma_{\mathcal{H}^1} \in \Pi(\rho_{\mathcal{H}^1}, \nu_{\mathcal{H}^1})$ ,  $\gamma_{\text{exc}} \in \Pi(\rho_{\text{exc}}, \nu_{\text{exc}})$  and they are optimal transport plans between their respective marginals. Indeed if we find a better transport plan for either problem we can construct a better plan for the original problem, contradicting the minimality of  $\gamma$ . We therefore also have a decomposition between the Wasserstein distances

$$(5.5) \quad W_p^p(\rho_0, \nu) = W_p^p(\rho_{\mathcal{H}^1}, \nu_{\mathcal{H}^1}) + W_p^p(\rho_{\text{exc}}, \nu_{\text{exc}}).$$

Let us point out that, although the decomposition of  $\nu$  is natural, there are many ways to decompose  $\gamma$  and  $\rho_0$ , for instance by choosing another disintegration family. In the sequel we show that for any such decomposition the excess must be

concentrated on the graph of the operator given by the (multivalued) projection onto  $\Sigma$

$$(5.6) \quad \Pi_\Sigma(x) \stackrel{\text{def.}}{=} \arg \min_{y \in \Sigma} |x - y|^2.$$

Note that  $\Pi_\Sigma$  is a multivalued operator which is included in the subgradient of the convex conjugate of the function:  $y \mapsto |y|^2/2$  if  $y \in \Sigma$  and  $+\infty$  else.

**Lemma 5.1.** *Let  $\nu$  be a minimizer of  $(\overline{P}_\Lambda)$  and  $\gamma$  an optimal transport plan from  $\rho_0$  to  $\nu$ . Then, for any decomposition  $\gamma = \gamma_{\mathcal{H}^1} + \gamma_{\text{exc}}$ , s.t.  $(\pi_1)_\# \gamma_{\text{exc}} = \nu_{\text{exc}}$ , it holds that*

$$(5.7) \quad \text{supp } \gamma_{\text{exc}} \subset \text{graph}(\Pi_\Sigma).$$

In addition, for any  $\pi_\Sigma$  measurable selection of  $x \mapsto \Pi_\Sigma(x)$ , the measure

$$\nu_{\mathcal{H}^1} + (\pi_\Sigma)_\# \rho_{\text{exc}}$$

is optimal for  $(\overline{P}_\Lambda)$ .

*Proof.* Consider the problem

$$(\overline{Q}_\Lambda) \quad \inf_{\substack{\gamma \in \mathcal{P}_p(\mathbb{R}^d \times \mathbb{R}^d) \\ (\pi_0)_\# \gamma = \rho_0}} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma(x, y) + \Lambda \mathcal{L}((\pi_1)_\# \gamma),$$

which is a reformulation of  $(\overline{P}_\Lambda)$  in terms of the transport plan  $\gamma$  from  $\rho_0$  to  $\nu$ .

Now, let  $(\gamma_{\mathcal{H}^1}, \gamma_{\text{exc}})$  be any suitable decomposition of  $\gamma$  and let  $\pi_\Sigma$  be a measurable selection of  $\Pi_\Sigma$ . We set  $\rho_{\text{exc}} \stackrel{\text{def.}}{=} (\pi_0)_\# \gamma_{\text{exc}}$  and define  $\tilde{\gamma} = \gamma_{\mathcal{H}^1} + (\text{id}, \pi_\Sigma)_\# \rho_{\text{exc}}$ . Then, since still  $\pi_1 \# \tilde{\gamma} \geq \nu_{\mathcal{H}^1}$ , it holds that  $\mathcal{L}(\pi_1 \# \tilde{\gamma}) \leq \mathcal{L}(\nu)$  and

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\tilde{\gamma} &= \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma_{\mathcal{H}^1} + \int_{\mathbb{R}^d} |x - \pi_\Sigma(x)|^p d\rho_{\text{exc}} \\ &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma_{\mathcal{H}^1} + \int_{\mathbb{R}^d \times \Sigma} |x - y|^p d\gamma_{\text{exc}} = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma \end{aligned}$$

Since  $\gamma$  is a minimizer of  $(\overline{Q}_\Lambda)$ , we must have an equality, in particular it holds that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} (|x - y|^p - |x - \pi_\Sigma(x)|^p) d\gamma_{\text{exc}} = 0.$$

Since  $\gamma$ -a.e.  $(x, y)$  is in  $\mathbb{R}^d \times \Sigma$ , the integrand is nonnegative and must vanish  $\gamma_{\text{exc}}$ -a.e. Hence  $(x, y) \in \text{Graph}(\Pi_\Sigma)$  for  $\gamma_{\text{exc}}$ -a.e.  $(x, y)$  and (5.7) follows since  $\text{Graph}(\Pi_\Sigma)$  is closed. As a consequence, the measure  $\nu_{\mathcal{H}^1} + \pi_\Sigma \# \rho_{\text{exc}}$  reaches the minimum for  $(\overline{P}_\Lambda)$  and is optimal.  $\square$

**5.2. Solutions are absolutely continuous.** Now we prove that the solutions to the relaxed problem  $(\overline{P}_\Lambda)$  are absolutely continuous w.r.t.  $\mathcal{H}^1 \llcorner \Sigma$ . The proof is based on the construction of a localized variational problem.

**Lemma 5.2.** *Let  $\nu$  be an optimal solution for the relaxed problem  $(\overline{P}_\Lambda)$  and set  $\alpha = \mathcal{L}(\nu)$ . Let  $\mathcal{S} = \mathcal{S}_0 \times \mathcal{S}_1 \subset \mathbb{R}^d \times \mathbb{R}^d$  be a Borel set and define the transport plan*

$$\gamma_{\mathcal{S}} \stackrel{\text{def.}}{=} \gamma_{\text{exc}} \llcorner \mathcal{S}_0 \times \mathcal{S}_1$$

along with its marginals

$$\rho_{\mathcal{S}} \stackrel{\text{def.}}{=} \pi_{0\sharp} \gamma_{\mathcal{S}} \leq \rho_{\text{exc}} \llcorner \mathcal{S}_0, \quad \nu_{\mathcal{S}} \stackrel{\text{def.}}{=} \pi_{1\sharp} \gamma_{\mathcal{S}}.$$

Then the measure  $\nu_{\mathcal{S}}$  solves the following variational problem

$$(5.8) \quad \inf \left\{ W_p^p(\rho_{\mathcal{S}}, \nu') : \begin{array}{l} \text{there is } \Gamma \text{ such that} \\ \nu' \in \mathcal{M}_+(\Sigma \cup \Gamma), \\ \nu' \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus \Sigma), \\ \Sigma \cup \Gamma \in \mathcal{A}, \nu'(\mathbb{R}^d) = \nu_{\mathcal{S}}(\mathbb{R}^d) \end{array} \right\}$$

More generally, let  $(\sigma_{\mathcal{S},t})_{t \in [0,1]}$  be the constant speed geodesic between  $\rho_{\mathcal{S}}$  and  $\nu_{\mathcal{S}}$  defined through  $\sigma_{\mathcal{S},t} \stackrel{\text{def.}}{=} \pi_{(1-t)\sharp} \gamma_{\mathcal{S}}$ , where  $\pi_t(x, y) \stackrel{\text{def.}}{=} (1-t)x + ty$ . Then for any  $t \in [0, 1]$ , the measure  $\nu_{\mathcal{S}}$  minimizes the variational problem

$$(5.9) \quad \inf \left\{ W_p^p(\sigma_{\mathcal{S},t}, \nu') : \begin{array}{l} \text{there is } \Gamma \text{ such that} \\ \nu' \in \mathcal{M}_+(\Sigma \cup \Gamma), \\ \nu' \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus \Sigma), \\ \Sigma \cup \Gamma \in \mathcal{A}, \nu'(\mathbb{R}^d) = \nu_{\mathcal{S}}(\mathbb{R}^d) \end{array} \right\}.$$

*Proof.* See Appendix A. □

We now craft a specific set  $\mathcal{S}$  to apply the lemma. Given  $\delta > 0$ , we define the set

$$(5.10) \quad D_{\delta} \stackrel{\text{def.}}{=} \{x \in \text{supp } \rho_{\text{exc}} : \delta \leq \text{dist}(x, \Sigma) \leq \delta^{-1}\},$$

And for a fixed point  $y_0 \in \Sigma$ , and  $\delta, r > 0$  consider the new transport plan

$$(5.11) \quad \gamma_{\delta,r} \stackrel{\text{def.}}{=} \gamma_{\text{exc}} \llcorner D_{\delta} \times B_r(y_0)$$

along with its marginals

$$(5.12) \quad \rho_{\delta,r} \stackrel{\text{def.}}{=} \pi_{0\sharp} \gamma_{\delta,r} \leq \rho_{\text{exc}} \llcorner D_{\delta}, \quad \nu_{\delta,r} \stackrel{\text{def.}}{=} \pi_{1\sharp} \gamma_{\delta,r}.$$

From Lemma 5.2 it holds that

$$(5.13) \quad \nu_{\delta,r} \in \arg \min \left\{ W_p^p(\rho_{\delta,r}, \nu') : \begin{array}{l} \text{there is } \Gamma \text{ such that} \\ \nu' \in \mathcal{M}_+(\Sigma \cup \Gamma), \\ \nu' \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus \Sigma), \\ \Sigma \cup \Gamma \in \mathcal{A}, \nu'(\mathbb{R}^d) = \nu_{\delta,r}(\mathbb{R}^d) \end{array} \right\}.$$

We also introduce

$$(5.14) \quad \gamma_{\delta} \stackrel{\text{def.}}{=} \gamma_{\text{exc}} \llcorner D_{\delta} \times \Sigma \text{ and } \nu_{\delta} \stackrel{\text{def.}}{=} \pi_{1\sharp} \gamma_{\delta},$$

so that by definition,  $\nu_{\delta,r} = \nu_{\delta} \llcorner B_r(y_0)$  and  $\nu_{\text{exc}}$  can be further decomposed as  $\nu_{\text{exc}} = \nu_{\delta} + \pi_{1\sharp}(\gamma_{\text{exc}} \llcorner D_{\delta}^c \times \mathbb{R}^d)$ . As  $D_{\delta}$  is a nested sequence of sets,  $(\nu_{\delta})_{\delta > 0}$  is a monotone sequence and taking the limit as  $\delta \rightarrow 0$  we have

$$(5.15) \quad \nu_{\text{exc}} = \sup_{\delta > 0} \nu_{\delta} + \rho_{\text{exc}} \llcorner \Sigma,$$

the second limit being  $\rho_{\text{exc}} \llcorner \Sigma$  because of Lemma 5.1 and since the only projection of a point in  $\Sigma$  is itself.

In the next Theorem 5.4 we show that the measures  $\nu_{\delta}$  have a uniformly bounded density w.r.t.  $\mathcal{H}^1$ . So when  $\rho_0$  is absolutely continuous w.r.t.  $\mathcal{H}^1$ , (5.15) shows that

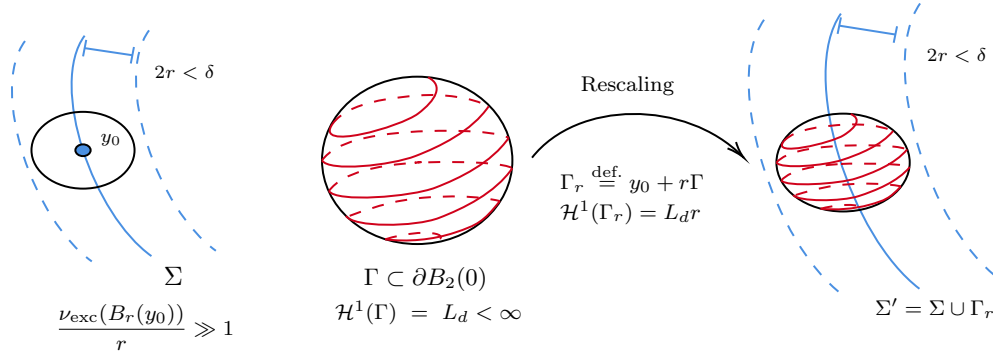


FIGURE 2. Scheme of the proof of Thm. 5.4. For the new competitor, created with the curve  $\Gamma$  from Lemma 5.3, we pay a little more in the transportation cost to generate  $\alpha^{-1}\mathcal{H}^1 \llcorner \Gamma_r$ , but pay much less by projecting the remaining mass onto it.

any optimal  $\nu \ll \mathcal{H}^1$ . The argument consists in crafting a competitor for the localized problem (5.13), built as a measure supported on a curve with controlled length, defined over small sphere, centered at an arbitrary point of the support of  $\nu_\delta$ . Letting the radius of this sphere go to zero, and comparing the energy of this competitor and the optimal measure, gives a uniform bound on the density. This strategy is illustrated in Figure 2.

**Lemma 5.3.** *Let  $B_2$  be the ball on  $\mathbb{R}^d$  centered at the origin. There exists a connected set  $\Gamma_d \subset \partial B_2$  with  $\mathcal{H}^1(\Gamma_d) < +\infty$  and such that*

$$\text{dist}(x, \Gamma_d) \leq |x - y| - \frac{1}{2}$$

for any  $x \notin B_2$  and for all  $y \in B_1$ .

*Proof.* We start by covering the sphere  $\partial B_2$  with finitely many balls  $(B_{1/2}(x_i))_{i=1}^{N_d}$ , each having radius  $1/2$ . The number of balls  $N_d$  being dependent on the dimension. In the sequel we define  $\Gamma_d$  with geodesics on  $\partial B_2$  connecting the centers  $(x_i)_{i=1}^{N_d}$ .

As we have finitely many points, we will also have finitely many curves and hence  $\mathcal{H}^1(\Gamma_d)$  must be a dimensional constant. We can even choose the connected set  $\Gamma_d$  with minimal length, which is a solution to Steiner's problem on the spheres and has a tree structure, so that we can bound  $\mathcal{H}^1(\Gamma_d) \leq (N_d - 1)D_d$ , where  $D_d$  is the diameter of  $\partial B_2$  in its Riemannian metric.

To prove the desired property, take  $x \notin B_2$  and  $y \in B_1$ . Let  $\{\hat{y}\} = [x, y] \cap \partial B_2$ . Then  $\hat{y} \in B_{1/2}(x_i)$  for some  $x_i$  while  $|x - \hat{y}| = |x - y| - |\hat{y} - y| \leq |x - y| - 1$ , and it follows:

$$\text{dist}(x, \Gamma_d) \leq |x - x_i| \leq |x - \hat{y}| + |\hat{y} - x_i| \leq |x - y| - \frac{1}{2},$$

which gives the desired construction.  $\square$



**Theorem 5.4.** *Given  $\rho_0 \in \mathcal{P}_p(\mathbb{R}^d)$ , let  $\nu$  be a solution to  $(\bar{P}_\Lambda)$ . Then it holds that the measures  $(\nu_\delta)_{\delta>0}$  are of the form*

$$\nu_\delta = \theta_\delta \mathcal{H}^1 \llcorner \Sigma, \text{ with } \|\theta_\delta\|_{L^\infty(\Sigma, \mathcal{H}^1)} \leq \frac{7}{2} \frac{C_d}{\mathcal{L}(\nu)},$$

for  $C_d = 2 + \mathcal{H}^1(\Gamma_d)$ ,  $\Gamma_d$  being the set from Lemma 5.3.

Therefore, if  $\rho_0 \ll \mathcal{H}^1$  or has a  $L^\infty$  density w.r.t.  $\mathcal{H}^1$ , so does  $\nu$ , in particular it is a rectifiable measure.

*Proof.* For  $y_0 \in \Sigma$ , let us define the one-dimensional upper density [3, Def. 2.55]

$$\theta_\delta(y_0) \stackrel{\text{def.}}{=} \limsup_{r \rightarrow 0} \frac{\nu_\delta(B_r)}{2r}.$$

We will show that  $\theta_\delta(y_0) \leq \frac{7}{2} \frac{C_d}{\mathcal{L}(\nu)}$ , so that thanks to [3, Thm. 2.56],  $\nu_\delta \ll \mathcal{H}^1 \llcorner \Sigma$ . Since  $\Sigma$  is 1-rectifiable, it follows that for  $\mathcal{H}^1$ -a.e.  $y_0 \in \Sigma$ ,  $\theta_\delta(y_0)$  is the Radon-Besicovitch derivative of  $\nu_\delta$  w.r.t.  $\mathcal{H}^1 \llcorner \Sigma$ , and the claim of the theorem follows.

From the optimality of  $\nu$ , the measure  $\nu_{\delta,r}$  solves problem (5.13). In order to build a competitor we consider the set  $\Gamma_d$  from Lemma 5.3, choose some point  $\bar{y} \in \Gamma_d$  and define

$$\Gamma_r \stackrel{\text{def.}}{=} [y_0, y_0 + r\bar{y}] \cup (y_0 + r\Gamma_d),$$

which is contained in  $\bar{B}_{2r}(y_0)$ . Notice that  $\Sigma \cup \Gamma_r$  is always a compact, connected and 1-rectifiable set and one has

$$\mathcal{H}^1(\Gamma_r) = C_d r,$$

where  $C_d = 1 + \mathcal{H}^1(\Gamma_d)$  is a constant depending only on the dimension.

In the sequel, setting  $\alpha = \mathcal{L}(\nu)$  we define the following parameter

$$m_r \stackrel{\text{def.}}{=} \frac{\mathcal{H}^1(\Gamma_r)}{\alpha \nu_\delta(B_r)}.$$

Suppose that  $C_d/\alpha < 2\theta_\delta(y_0)$ . Then,

$$1 > m_0 \stackrel{\text{def.}}{=} \frac{C_d}{2\alpha\theta_\delta(y_0)} = \liminf_{r \rightarrow 0} m_r.$$

Now, we consider a subsequence  $(r_k)_{k \in \mathbb{N}} \searrow 0$  such that  $\lim_{k \rightarrow \infty} m_{r_k} = \liminf_{r \rightarrow 0} m_r$ . In particular,  $m_{r_k} \in (0, 1)$  for  $r_k$  sufficiently small. For simplicity, in the sequel, we drop the subscript  $k$ , yet we consider only  $r \in \{r_k\}_{k \in \mathbb{N}}$ .

Let  $\gamma_{\Gamma_r}$  be an optimal transport plan between  $m_r \rho_{\delta,r}$  and  $\alpha^{-1} \mathcal{H}^1 \llcorner \Gamma_r$  for the Wasserstein- $p$  distance and define the new plan

$$\tilde{\gamma}_{\delta,r} \stackrel{\text{def.}}{=} \gamma_{\Gamma_r} + (1 - m_r) (\text{id}, \pi_{\Gamma_r})_\# \rho_{\delta,r}, \text{ and } \tilde{\nu}_{\delta,r} \stackrel{\text{def.}}{=} \pi_{1\#} \tilde{\gamma}_{\delta,r},$$

where  $\pi_{\Gamma_r}$  is a measurable selection of the projection operator onto  $\Gamma_r$ . This construction is illustrated in Figure 2. Therefore  $\tilde{\nu}_{\delta,r}$  is admissible for (5.13) and we have the following estimate

$$W_p^p(\rho_{\delta,r}, \tilde{\nu}_{\delta,r}) \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma_{\Gamma_r} + (1 - m_r) \int_{\mathbb{R}^d} \text{dist}(x, \Gamma_r)^p d\rho_{\delta,r}.$$

We will estimate each term of the previous inequality separately. For the first one, notice that as  $\text{supp } \gamma_{\Gamma_r} \subset \Pi_{\Sigma}^{-1}(B_r(y_0)) \times \overline{B}_{2r}(y_0)$ , it holds that

$$|x - y| \leq \text{dist}(x, \Sigma) + 3r, \quad \text{for } \gamma_{\Gamma_r}\text{-a.e. } (x, y).$$

For the second term, as the projection of  $x$  onto  $\Sigma$  is inside  $B_r(y_0)$ , it follows from Lemma 5.3 that

$$\text{dist}(x, \Gamma_r) \leq \text{dist}(x, \Sigma) - \frac{r}{2}, \quad \text{for } \text{dist}(x, \Sigma) > 2r.$$

Therefore, for a fixed  $\delta$  and taking  $2r < \delta$ , the Wasserstein distance is bounded by

$$W_p^p(\rho_{\delta,r}, \tilde{\nu}_{\delta,r}) \leq m_r \int_{\mathbb{R}^d} (\text{dist}(x, \Sigma) + 3r)^p d\rho_{\delta,r} + (1 - m_r) \int_{\mathbb{R}^d} (\text{dist}(x, \Sigma) - r/2)^p d\rho_{\delta,r}.$$

Notice that  $W_p^p(\rho_{\delta,r}, \nu_{\delta,r}) = \int_{\mathbb{R}^d} \text{dist}(x, \Sigma)^p d\rho_{\delta,r}$ , so in order to compare the Wasserstein distances we use the following inequalities

$$\begin{aligned} (\text{dist}(x, \Sigma) + 3r)^p &\leq \text{dist}(x, \Sigma)^p + 3rp(\text{dist}(x, \Sigma) + 3r)^{p-1} \\ \left(\text{dist}(x, \Sigma) - \frac{r}{2}\right)^p &\leq \text{dist}(x, \Sigma)^p - \frac{r}{2}p \left(\text{dist}(x, \Sigma) - \frac{r}{2}\right)^{p-1} \end{aligned}$$

which follow from the convexity of  $t \mapsto |t|^p$ . Then, given  $\varepsilon > 0$ , if  $r \leq \delta\varepsilon$  one deduces, for  $\text{dist}(x, \Sigma) \geq \delta$ , that:

$$\begin{aligned} (\text{dist}(x, \Sigma) + 3r)^p &\leq \text{dist}(x, \Sigma)^p + 3rp(1 + 3\varepsilon)^{p-1} \text{dist}(x, \Sigma)^{p-1} \\ \left(\text{dist}(x, \Sigma) - \frac{r}{2}\right)^p &\leq \text{dist}(x, \Sigma)^p - \frac{r}{2}p \left(1 - \frac{\varepsilon}{2}\right)^{p-1} \text{dist}(x, \Sigma)^{p-1}. \end{aligned}$$

Therefore it holds that

$$\begin{aligned} W_p^p(\rho_{\delta,r}, \tilde{\nu}_{\delta,r}) &\leq W_p^p(\rho_{\delta,r}, \nu_{\delta,r}) + pr\Delta_{r,\varepsilon} \int_{\mathbb{R}^d} \text{dist}(x, \Sigma)^{p-1} d\rho_{\delta,r} \\ \text{for } \Delta_{r,\varepsilon} &= 3m_r(1 + 3\varepsilon)^{p-1} - \frac{1 - m_r}{2} \left(1 - \frac{\varepsilon}{2}\right)^{p-1} \end{aligned}$$

Hence from the optimality of  $\nu_{\delta,r}$  we have  $\Delta_{r,\varepsilon} \geq 0$ , so that letting  $r \rightarrow 0$  and then  $\varepsilon \rightarrow 0$ , it must hold that  $3m_0 \geq (1 - m_0)/2$ , that is:

$$\theta_\delta(y_0) \leq \frac{7C_d}{2\alpha}.$$

As a result, the family  $(\nu_\delta)_{\delta>0}$  has a uniform  $L^\infty$  density bounds, and so does the limit measure  $\sup_{\delta>0} \nu_\delta = (\sup_{\delta>0} \theta_\delta) \mathcal{H}^1 \llcorner \Sigma$ . But as the exceeding measure can be decomposed as (5.15) we deduce that whenever the initial measure  $\rho_0 \ll \mathcal{H}^1$  or has a  $L^\infty$  density w.r.t.  $\mathcal{H}^1$ , so does the solution  $\nu$ .  $\square$

## 6. EXISTENCE OF SOLUTIONS TO $(P_\Lambda)$

This section is dedicated to the proof of Theorem 1.1, item (2). Knowing that the excess measure is absolutely continuous (Theorem 5.4), we use a blow up argument near a rectifiability point  $y_0$  of  $\Sigma$ . From Lemma 5.2, the blow-ups of  $\nu_{\text{exc}}$  minimize a family of functionals  $(F_r)_{r>0}$ , which in turn  $\Gamma$ -converge to some functional  $F$ . Since these blow-ups also converge (for  $\mathcal{H}^1$ -a.e.  $y_0$ ) to a uniform density on  $T_{y_0}\Sigma$ , this limit measure must also minimize the  $\Gamma$ -limit  $F$ . Yet if it is not zero, we can build

a better competitor (Lemma 6.3 below), giving a contradiction to the minimality of the uniform measure. We deduce that  $\nu_{\text{exc}}$  vanishes.

**6.1. Blow-up and  $\Gamma$ -convergence.** In the sequel, we assume that  $\rho_0 \ll \mathcal{H}^1$ , so that from Theorem 5.4 any minimizer  $\nu$ , as well as  $(\nu_\delta)_{\delta>0}$  (defined in (5.14)), are rectifiable measures and we can write

$$\nu_\delta = \theta_\delta \mathcal{H}^1 \llcorner \Sigma, \text{ for } \theta_\delta \in L^1(\mathcal{H}^1 \llcorner \Sigma).$$

Observe that  $\nu_\delta$ -a.e.  $y \in \Sigma$  is a rectifiability point, and we choose  $y_0 \in \Sigma$  such that:

$$(6.1) \quad T_{y_0} \Sigma \text{ exists} \quad \text{and} \quad y_0 \text{ is a Lebesgue point of } \theta_\delta.$$

We then use Lemma 5.2 with the choice  $\mathcal{S}_0 \times \mathcal{S}_1 = D_\delta \times B_r(y_0)$ , and we focus on the variational problem (5.9): we obtain the families of measures  $(\nu_{\delta,r})_{r>0}$  and  $(\sigma_{\delta,r})_{r>0}$  as  $\nu_{\delta,r} \stackrel{\text{def.}}{=} \nu_\delta \llcorner B_r(y_0)$  and  $\sigma_{\delta,r} \stackrel{\text{def.}}{=} \pi_{(1-r)\sharp} \gamma_{\delta,r}$ , where  $(\sigma_{\delta,t})_{t \in [0,1]}$  is a family of geodesic interpolations, as in Lemma 5.2, so that

$$(6.2) \quad \nu_{\delta,r} \in \arg \min \left\{ W_p^p(\sigma_{\delta,r}, \nu') : \begin{array}{l} \text{there is } \Gamma \text{ such that} \\ \nu' \in \mathcal{M}_+(\Sigma \cup \Gamma), \\ \nu' \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus \Sigma), \\ \Sigma \cup \Gamma \in \mathcal{A}, \nu'(\mathbb{R}^d) = \nu_{\delta,r}(\mathbb{R}^d) \end{array} \right\}.$$

From Lemma 5.1 the optimal transport plan between  $\nu_{\delta,r}$  and  $\sigma_{\delta,r}$  is supported on  $\text{graph}(\Pi_\Sigma)$ .

The sequence of measures  $\nu_{\delta,r}$  are essentially a localization of  $\nu_\delta$  around  $y_0$  so, by the blow-up Theorem 2.4 (see also [3, Theo. 2.83]), it holds that

$$(6.3) \quad r^{-1} \Phi_\sharp^{y_0,r} \nu_{\delta,r} \xrightarrow[r \rightarrow 0]{*} \theta_\delta(y_0) \mathcal{H}^1 \llcorner [-\tau, \tau], \text{ where } \mathbb{R}\tau = T_{y_0} \Sigma.$$

Up to a subsequence (not labelled) we also have:

$$(6.4) \quad r^{-1} \Phi_\sharp^{y_0,r} \sigma_{\delta,r} \xrightarrow[r \rightarrow 0]{*} \bar{\sigma}_\delta$$

for some measure  $\bar{\sigma}_\delta$ . By construction  $\sigma_{\delta,r}$  is supported on  $\{r\delta^{-1} \geq \text{dist}(\cdot, \Sigma) \geq r\delta\}$ , so that  $\text{supp } \bar{\sigma}_\delta \subset \{x : \delta^{-1} \geq \text{dist}(x, \mathbb{R}\tau) \geq \delta\}$ .

In view of (6.3) and (6.4), we introduce the blow-ups of the measures  $\nu_{\delta,r}$  and  $\sigma_{\delta,r}$ ,

$$(6.5) \quad \bar{\nu}_{\delta,r} \stackrel{\text{def.}}{=} \frac{1}{r} \Phi_\sharp^{y_0,r} \nu_{\delta,r}, \quad \bar{\sigma}_{\delta,r} \stackrel{\text{def.}}{=} \frac{1}{r} \Phi_\sharp^{y_0,r} \sigma_{\delta,r}, \text{ and the set } \Sigma_r \stackrel{\text{def.}}{=} \frac{\Sigma - y_0}{r} \cap \overline{B_1(0)}.$$

In addition, we define a family of functionals  $(F_r)_{r>0}$  as

$$(6.6) \quad F_r(\nu') \stackrel{\text{def.}}{=} \begin{cases} W_p^p(\bar{\sigma}_{\delta,r}, \nu'), & \begin{array}{l} \text{there is } \Gamma \subset \overline{B_1(0)} \text{ such that} \\ \nu' \in \mathcal{M}_+(\Sigma_r \cup \Gamma), \nu' \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus \Sigma_r), \\ \left(\frac{\Sigma - y_0}{r}\right) \cup \Gamma \text{ closed and connected,} \\ \nu'(\overline{B_1(0)}) = \frac{\nu_\delta(B_r(y_0))}{r}, \end{array} \\ +\infty, & \text{otherwise,} \end{cases}$$

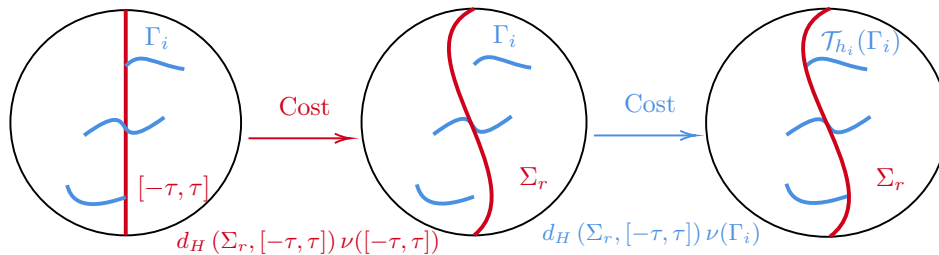


FIGURE 3. Transportation argument for the construction of a recovery sequence in the  $\Gamma$  convergence of  $(F_r)_{r>0}$ . Both operations have a transportation cost of the order  $d_H(\Sigma_r, [-\tau, \tau])$ , and hence converge to 0.

where  $\alpha = \mathcal{L}(\nu)$ . Observing that for any given measures  $\mu', \nu'$  we have

$$(6.7) \quad W_p^p \left( \frac{1}{r} \Phi_{\#}^{y_0, r} \mu', \frac{1}{r} \Phi_{\#}^{y_0, r} \nu' \right) = \frac{1}{r^{p+1}} W_p^p(\mu', \nu').$$

and recalling (6.2), we see that  $\bar{\nu}_{\delta, r} \in \arg \min F_r$  for any  $r > 0$ .

The natural candidate for the limit of this family is the following:

$$(6.8) \quad F(\nu') \stackrel{\text{def.}}{=} \begin{cases} W_p^p(\bar{\sigma}_{\delta}, \nu'), & \begin{aligned} & \text{there is } \Gamma \subset \overline{B_1(0)} \text{ such that} \\ & \nu' \in \mathcal{M}_+([-\tau, \tau] \cup \Gamma), \nu' \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus [-\tau, \tau]), \\ & \mathbb{R}\tau \cup \Gamma \text{ closed and connected,} \\ & \nu'(\overline{B_1(0)}) = 2\theta_{\delta}(y_0), \end{aligned} \\ +\infty, & \text{otherwise.} \end{cases}$$

We prove in Theorem 6.1 below that  $F_r$   $\Gamma$ -converges to  $F$  as  $r \rightarrow 0^+$ . We refer to [9, 6] and in particular to [6, Def. 1.24]) for the definition of the (lower and upper)  $\Gamma$ -limit. From the properties of the  $\Gamma$ -convergence, see [9, Cor. 7.20], it follows that  $\theta_{\delta}(y_0) \mathcal{H}^1 \llcorner [-\tau, \tau]$  must be a minimizer of  $F$  (as the limit of minimizers of  $F_r$ ). The estimate from below of the  $\Gamma$ -liminf is obtained with the tools developed so far, while estimating the  $\Gamma$ -limsup will require an appropriate construction illustrated in Figure 3.

**Theorem 6.1.** *The family of functionals  $(F_r)_{r>0}$   $\Gamma$ -converges to  $F$  as  $r \rightarrow 0^+$ , in the narrow topology.*

*Proof.*  $\Gamma$ -liminf: we consider an infinitesimal sequence  $(r_n)_{n \in \mathbb{N}}$  such that  $(\nu'_n)_{n \in \mathbb{N}}$  converges to  $\nu'$  in the narrow sense in  $\overline{B_1(0)}$ , and that  $\liminf_{n \rightarrow \infty} F_{r_n}(\nu'_n) < \infty$  for all  $n \in \mathbb{N}$ , otherwise there is nothing to prove.

First we look at the first marginals in the definition of  $F_{r_n}$ . From (6.4) we know that  $\bar{\sigma}_{\delta, r_n} \xrightarrow[n \rightarrow \infty]{*} \bar{\sigma}_{\delta}$ . By the lower semi-continuity of the Wasserstein distance w.r.t. the narrow convergence, if we prove that  $F(\nu') < \infty$ , that is, if the limit

satisfies the constraints in the definition of  $F$ , we will have that

$$F(\nu') \leq \liminf_{n \rightarrow \infty} F_{r_n}(\nu'_n).$$

As  $\alpha\nu'_n \geq \mathcal{H}^1 \llcorner (\Gamma_n \setminus \Sigma_{r_n})$  for some  $\Gamma_n \subset \overline{B_1(0)}$  such that  $\left(\frac{\Sigma - y_0}{r_n}\right) \cup \Gamma_n \in \mathcal{A}$ , Blaschke's Theorem [3, Thm. 6.1] and Lemma 2.6 imply that, up to a subsequence,  $\Gamma_n \xrightarrow[n \rightarrow \infty]{d_H} \Gamma$  for some closed set  $\Gamma \subset \overline{B_1(0)}$  and  $\frac{\Sigma - y_0}{r_n} \xrightarrow[n \rightarrow \infty]{K} \mathbb{R}\tau$ . Hence,

$$\Xi_n \stackrel{\text{def.}}{=} \left(\frac{\Sigma - y_0}{r_n}\right) \cup \Gamma_n \xrightarrow[n \rightarrow \infty]{K} \Xi \stackrel{\text{def.}}{=} \mathbb{R}\tau \cup \Gamma.$$

Let us check that  $\Xi$  is connected (which is not immediate since the Kuratowski limit of connected sets is not necessarily connected). Assume by contradiction that there are two disjoint open sets  $U, V \subset \mathbb{R}^d$  such that  $U \cap \Xi$  and  $V \cap \Xi$  form a partition of  $\Xi$ . Since  $\mathbb{R}\tau \subset \Xi$  is connected, it is contained in either  $U$  or  $V$  (say,  $U$ ). As a result,  $V \cap \Xi \subset \Gamma \subset \overline{B_1(0)}$  is bounded, and possibly replacing  $V$  with  $V \cap B_2(0)$ , we may assume that  $V$  is bounded too, so that  $\partial V$  is compact. Since  $\Xi \subset V \cap (\mathbb{R}^d \setminus \overline{V})$ , we note that  $\partial V \cap \Xi = \emptyset$ , and we deduce that  $\min_{x \in \partial V} \text{dist}(x, \Xi) > 0$ .

Now, the Kuratowski convergence of  $\Xi_n$  towards  $\Xi$  implies that, for all  $n$  large enough,  $\Xi_n$  intersects both  $V$  and  $U \subset \mathbb{R}^d \setminus \overline{V}$ , hence, by the connectedness of  $\Xi_n$ , there exists  $x_n \in \Xi_n \cap \partial V$ . But the Kuratowski convergence also implies that  $\text{dist}(\cdot, \Xi_n) \rightarrow \text{dist}(\cdot, \Xi)$  locally uniformly (hence uniformly on  $\partial V$ ), which contradicts that  $\min_{x \in \partial V} \text{dist}(x, \Xi) > 0$ . As a result,  $\Xi$  is connected.

The fact that  $\text{supp } \nu' \subset [-\tau, \tau] \cup \Gamma$  comes from the weak convergence of  $\nu'_n$  to  $\nu'$ . As this convergence takes place in a compact set it also holds that  $\nu'(\overline{B_1(0)}) = \lim_{n \rightarrow \infty} \nu'_n(\overline{B_1(0)}) = 2\theta_\delta(y_0)$  since  $\theta_\delta(y_0)$  is the density of  $\nu_\delta$  at  $y_0$ .

It only remains to verify the density constraints,  $\alpha\nu' \geq \mathcal{H}^1 \llcorner (\Gamma \setminus [-\tau, \tau])$ . We cannot apply Gołab's Theorem to  $\nu'_n$  since, although  $\alpha\nu'_n \geq \mathcal{H}^1 \llcorner (\Gamma_n \setminus \Sigma_{r_n})$ , we do not have an upper bound on the number of connected components of  $\Gamma_n \setminus \Sigma_{r_n}$ . What we do know is that the sequence  $\Xi_n = r_n^{-1}(\Sigma - y_0) \cup \Gamma_n$  satisfies the assumptions of Theorem 2.2, so we apply it to the measures  $\mathcal{H}^1 \llcorner \Xi_n$  instead, remembering that

$$\mathcal{H}^1 \llcorner \left(\frac{\Sigma - y_0}{r_n}\right) + \alpha\nu'_n \geq \mathcal{H}^1 \llcorner \left(\frac{\Sigma - y_0}{r_n} \cup \Gamma_n\right).$$

The left-hand side converges in the local weak- $\star$  sense to  $\mathcal{H}^1 \llcorner \mathbb{R}\tau + \alpha\nu'$ . The right-hand side (which is bounded by the left-hand side) converges in the same sense, up to a subsequence. We let  $\lambda$  denote a limit and Theorem 2.2 implies that  $\lambda \geq \mathcal{H}^1 \llcorner (\mathbb{R}\tau \cup \Gamma)$ , which gives  $\mathcal{H}^1 \llcorner \mathbb{R}\tau + \alpha\nu' \geq \mathcal{H}^1 \llcorner (\mathbb{R}\tau \cup \Gamma)$ , and thus

$$\alpha\nu' \geq \mathcal{H}^1 \llcorner (\Gamma \setminus [-\tau, \tau]).$$

$\Gamma$ -limsup: Let  $(r_n)_{n \in \mathbb{N}}$  be an infinitesimal sequence. By Lemma 2.6, we know that  $(\Sigma - y_0)/r_n$  converges in the Kuratowski sense towards  $\mathbb{R}\tau$ , and  $\Sigma_{r_n} \stackrel{\text{def.}}{=} (\Sigma - y_0)/r_n \cap \overline{B_1(0)}$  converges towards  $[-\tau, \tau]$  for the Hausdorff distance.

The strategy to prove the limsup is illustrated in Figure 3, and roughly explained as follows. We concatenate three steps. First we renormalize  $\nu'$  to satisfy the mass constraint in  $F_{r_n}$ . But this normalization may break the condition  $\alpha\nu'_n \geq$

$\mathcal{H}^1 \llcorner (\Gamma \setminus [-\tau, \tau])$ , so we slightly shrink the support to satisfy this constraint again. We also need the measure  $\nu'_n$  to be supported on some connected set  $\Sigma_{r_n} \cup \Gamma_n$ , hence we move the mass of  $\nu'$  from  $[-\tau, \tau]$  to  $\Sigma_{r_n}$  by projection, and we translate the mass of each connected component of the (shrunk)  $\Gamma \setminus [-\tau, \tau]$  so that it is connected to  $\Sigma_{r_n}$ . Eventually, by doing so, some parts of the support may get out of  $\overline{B_1(0)}$ , so we project the residual mass onto  $\overline{B_1(0)}$ .

To be more precise, we first address the case  $\theta_\delta(y_0) = 0$ . As  $F(\nu') < +\infty$  if and only if  $\nu' = 0$ , we need only prove the result for  $\nu' = 0$ . Let  $P_n$  be any measurable selection of the projection onto  $\Sigma_{r_n}$ , and define  $\nu'_n \stackrel{\text{def.}}{=} P_{n\#} \bar{\sigma}_{\delta, r_n}$ . With  $\Gamma = \emptyset$ , and since  $|x - P_n(x)| \leq \delta^{-1}$  for all  $x \in \text{supp } \bar{\sigma}_{\delta, r_n}$ , we observe that

$$F_{r_n}(\nu'_n) \leq W_p^p(\bar{\sigma}_{\delta, r_n}, \nu'_n) \leq \delta^{-p} \frac{\nu_\delta(B_{r_n})}{r_n} \xrightarrow{n \rightarrow +\infty} 0 = F(\nu').$$

Moreover, as  $\nu'_n \xrightarrow{n \rightarrow +\infty} \nu'$  in the narrow topology, we have built a recovery sequence for  $\nu'$ .

Now, we deal with the case  $\theta_\delta(y_0) > 0$ . Let  $\nu'$  such that  $F(\nu') < +\infty$ , and let  $\Gamma$  be a set as in (6.8). Observe that  $[-\tau, \tau] \cup \Gamma$  is connected, being the projection of  $\mathbb{R}\tau \cup \Gamma$  onto  $\overline{B_1(0)}$ , and since it has finite  $\mathcal{H}^1$  measure, it is arcwise connected, by [10, Prop. 30.1, Cor. 30.2]. As a result,  $\mathbb{R}\tau \cup \Gamma$  is arcwise connected too.

Let  $(C_i)_{i \in I}$  denote the arcwise connected components of  $\Gamma \setminus (\mathbb{R}\tau)$ . For each  $i \in I$ , as the set  $\mathbb{R}\tau \cup \Gamma$  is arcwise connected, one may check that there exists some  $z_i \in [-\tau, \tau]$  such that  $\{z_i\} \cup C_i$  is arcwise connected. As a result, the set  $C_i \subset \mathbb{R}^d \setminus (\mathbb{R}\tau)$  cannot consist of one single point, and  $\mathcal{H}^1(C_i) > 0$ . Therefore, the index set  $I$  is at most countable.

Let us construct a recovery sequence  $(\nu'_n)_{n \in \mathbb{N}}$ . By the Kuratowski (even Hausdorff) convergence of  $\Sigma_{r_n}$  towards  $[-\tau, \tau]$ , for each  $i \in I$ , there exists a sequence  $(z_{n,i})_{n \in \mathbb{N}}$  such that  $z_{n,i} \in \Sigma_{r_n}$  for each  $n \in \mathbb{N}$ , and  $z_{n,i} \rightarrow z_i$ . We then define

$$a_n \stackrel{\text{def.}}{=} \frac{\nu_\delta(B_{r_n})}{2r_n \theta_\delta(y_0)}, \quad \text{and} \quad s_n \stackrel{\text{def.}}{=} \max(1, a_n^{-1}),$$

noting that  $a_n \rightarrow 1$  and  $s_n \rightarrow 1$ , and we introduce the map  $T_n$ ,

$$T_n(y) \stackrel{\text{def.}}{=} \begin{cases} P_n(y/s_n), & \text{if } y \in [-\tau, \tau], \\ (y - z_i)/s_n + z_{n,i}, & \text{if } y \in C_i, \end{cases}$$

where, as before,  $P_n$  is some measurable selection of the projection onto  $\Sigma_{r_n}$ . The map  $T_n$  shrinks each connected component  $C_i$  and translates it to the corresponding  $z_{n,i} \in \Sigma_{r_n}$  so as to ensure connectedness (see below). Letting  $P_B$  denote the projection onto the unit ball  $\overline{B_1(0)}$ , we eventually define

$$\nu'_n \stackrel{\text{def.}}{=} (P_B \circ T_n)\#(a_n \nu').$$

Let us check that  $\nu'_n$  converges to  $\nu'$  in the narrow topology. We note that for  $y \in [-\tau, \tau]$ ,

$$|y/s_n - P_n(y/s_n)| = \text{dist}(y/s_n, \Sigma_{r_n}) \leq d_H([-\tau, \tau], \Sigma_{r_n}) \xrightarrow{n \rightarrow +\infty} 0,$$

so that  $T_n(y) \rightarrow y$ , and for  $y \in C_i$ ,

$$|y - T_n(y)| \leq |y|(1 - 1/s_n) + |z_i/s_n - z_{n,i}| \xrightarrow{n \rightarrow +\infty} 0.$$

As a result, for  $y \in [-\tau, \tau] \cup \Gamma$ ,  $T_n(y) \rightarrow y$ , and eventually  $P_B \circ T_n(y) \rightarrow y$ . By the dominated convergence theorem, we get that for any  $\phi \in C_b(\mathbb{R}^d)$ ,

$$\int \phi d\nu'_n = a_n \int_{[-\tau, \tau] \cup \Gamma} \phi(P_B(T_n(y))) d\nu'(y) \xrightarrow{n \rightarrow +\infty} \int_{[-\tau, \tau] \cup \Gamma} \phi(y) d\nu'(y)$$

so that  $\nu'_n \xrightarrow{n \rightarrow +\infty} \nu'$  in the narrow topology.

Let us now check the constraints in  $F_{r_n}$ . From the properties of image measures, we see that  $\text{supp } \nu'_n \subset \overline{B_1(0)}$ , and that  $\nu'_n(\overline{B_1(0)}) = \nu'_n(\mathbb{R}^d) = a_n \nu'(\mathbb{R}^d) = \nu_\delta(B_{r_n})/r_n$ , so that  $\nu'_n$  has the mass prescribed by  $F_{r_n}$ . Consider the set

$$(6.9) \quad \Gamma_n \stackrel{\text{def.}}{=} \bigcup_{i \in I} \Gamma_{n,i} \quad \text{where} \quad \Gamma_{n,i} \stackrel{\text{def.}}{=} \overline{(P_B \circ T_n)(C_i)}.$$

In addition, the mass of  $\nu'_n$  is concentrated in  $\Sigma_{r_n} \cup \Gamma_n$ , and we prove below that satisfies all the constraints in  $F_{r_n}$ .

First let us show that  $\frac{\Sigma - y_0}{r_n} \cup \Gamma_n$  is connected. For each  $i \in I$ , as the set  $\{z_i\} \cup C_i$  is arcwise connected, so is its image by the map  $y \mapsto (y - z_i)/s_n + z_{n,i}$ , which is equal to  $\{z_{n,i}\} \cup T_n(C_i)$ . As a result  $\{z_{n,i}\} \cup \overline{P_B \circ T_n(C_i)} = \{z_{n,i}\} \cup \Gamma_{n,i}$  is connected, as well as  $\frac{\Sigma - y_0}{r_n} \cup \Gamma_n$ .

Now, let us show that  $\frac{\Sigma - y_0}{r_n} \cup \Gamma_n$  is closed. If  $I$  is finite, then, by (6.9),  $r_n^{-1}(\Sigma - y_0) \cup \Gamma_n$  is closed as the finite union of closed sets. Otherwise,  $I$  is countable, and from [28, Lemma 2.6], we have

$$\mathcal{H}^1(\Gamma_{n,i}) = \mathcal{H}^1(P_B \circ T_n(C_i)) \leq \mathcal{H}^1(T_n(C_i)) = s_n^{-1} \mathcal{H}^1(C_i) \xrightarrow{i \rightarrow \infty} 0.$$

Let  $(x_k)_{k \in \mathbb{N}}$  be a sequence contained in  $r_n^{-1}(\Sigma - y_0) \cup \Gamma_n$ , such that  $x_k \rightarrow x$ . If there is an infinite amount of terms of this sequence in either  $\frac{\Sigma - y_0}{r_n}$  or any of the  $\Gamma_{n,i}$ , since these sets are closed, then  $x \in \frac{\Sigma - y_0}{r_n} \cup \Gamma_n$ . Otherwise, we can find a sub-sequence  $x_{k'} \in \Gamma_{n,i_{k'}}$ , so that

$$\text{dist} \left( x, \frac{\Sigma - y_0}{r_n} \right) = \lim_{k' \rightarrow \infty} \text{dist} \left( x_{k'}, \frac{\Sigma - y_0}{r_n} \right) \leq \lim_{k' \rightarrow \infty} \mathcal{H}^1(\Gamma_{n,i_{k'}}) = 0,$$

and we conclude that  $\frac{\Sigma - y_0}{r_n} \cup \Gamma_n$  is closed.

To show it satisfies the density constraints, take any non-negative  $\phi \in C_b(\mathbb{R}^d)$ ,

$$\begin{aligned}
\alpha \int \phi d\nu'_n &= \alpha a_n \int_{[-\tau, \tau] \cup \Gamma} \phi(P_B(T_n(y))) d\nu'(y) \\
&\geq \alpha a_n \sum_{i \in I} \int_{C_i} \phi(P_B(T_n(y))) d\nu'(y) \\
&\geq a_n \sum_{i \in I} \int_{C_i} \phi(P_B((y - z_i)/s_n + z_{n,i})) d\mathcal{H}^1(y) \\
&= a_n s_n \sum_{i \in I} \int_{\Gamma_{n,i}} \phi(P_B(y')) d\mathcal{H}^1(y') \\
&\geq \int_{\Gamma_n} \phi d\mathcal{H}^1.
\end{aligned}$$

It follows that  $\alpha \nu'_n \geq \mathcal{H}^1 \llcorner \Gamma_n$  and we conclude that  $F_{r_n}(\nu_n) < \infty$ , for all  $n \in \mathbb{N}$ .

By the continuity of the Wasserstein distance with respect to the narrow convergence (provided the measures are supported in some common compact set), we have that:

$$F_{r_n}(\nu'_n) \xrightarrow{n \rightarrow \infty} F(\nu').$$

The  $\Gamma$ -convergence follows.  $\square$

Now that we have characterized the limit problem, we show that the optimal transportation is given by projections as the blow-up family.

**Lemma 6.2.** *If  $\theta_\delta(y_0) > 0$ , the optimal transport plan between the measure  $\sigma_\delta$ , defined in (6.4), and  $\bar{\nu} = \theta_\delta(y_0) \mathcal{H}^1 \llcorner [-\tau, \tau]$ , defined in (6.3), is unique and given by the projection map  $\Pi_{[-\tau, \tau]}$ .*

*Proof.* Consider a family  $\bar{\gamma}_r$  of optimal transport plans from  $\bar{\sigma}_{\delta,r}$  to  $\bar{\nu}_{\delta,r}$ . Up to a subsequence it converges to some  $\bar{\gamma}$ , which, by the stability of optimal transport plans, also transports  $\sigma_\delta$  to  $\bar{\nu}$  optimally. Since  $\bar{\sigma}_{\delta,r}, \bar{\nu}_{\delta,r}$  are generated by the pushforward of  $\nu_{\text{exc}} \llcorner B_r(y_0)$  by  $\Phi^{y_0,r}$ , from Lemma 5.1 we know that

$$\text{supp } \bar{\gamma}_r \subset \text{graph}(\Pi_{\Sigma_r}).$$

Let us show that  $\text{supp } \bar{\gamma} \subset \text{graph}(\Pi_{[-\tau, \tau]})$ . Indeed if  $(x, p) \in \text{supp } \bar{\gamma}$ , there is an open ball  $B$  centered at  $(x, p)$  such that

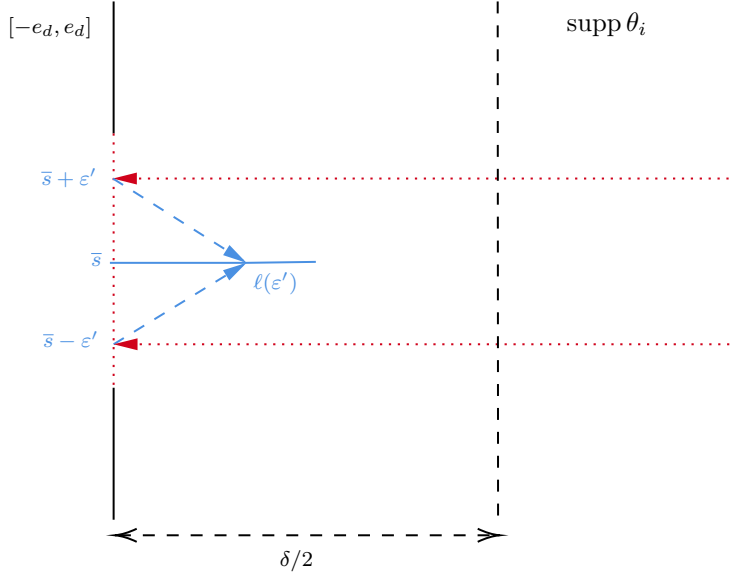
$$0 < \bar{\gamma}(B) \leq \liminf_{r \rightarrow 0} \bar{\gamma}_r(B).$$

In particular, we can find  $\text{supp } \bar{\gamma}_r \ni (x_r, p_r) \xrightarrow{r \rightarrow 0} (x, p)$ . So it holds that

$$|x - p| = \lim_{r \rightarrow 0} |x_r - p_r| = \lim_{r \rightarrow 0} \text{dist}(x_r, \Sigma_r) = \text{dist}(x, [-\tau, \tau]),$$

where the last equality comes from the uniform convergence of the distance functions, recalling from Lemma 2.6 that  $\Sigma_r \xrightarrow[r \rightarrow 0]{d_H} [-\tau, \tau]$ .



FIGURE 4. Construction of a competitor for the minimization of  $F$ .

Now we show that this property is true for any other optimal plan. Consider  $\gamma$  transporting  $\sigma_\delta$  to  $\bar{\nu}$  optimally, then by the optimality of  $\bar{\gamma}$  it holds that

$$\begin{aligned} \int_{\mathbb{R}^d} (\text{dist}(x, [-\tau, \tau]))^p d\sigma_\delta &= \int |x - y|^p d\bar{\gamma} = \int |x - y|^p d\gamma \\ &\geq \int (\text{dist}(x, [-\tau, \tau]))^p d\gamma = \int_{\mathbb{R}^d} \text{dist}(x, [-\tau, \tau])^p d\sigma_\delta. \end{aligned}$$

Since  $|x - y| - \text{dist}(x, [-\tau, \tau]) \geq 0$  for  $\gamma$ -a.e.  $(x, y)$  and the inequality above must be an equality, we must have  $\text{supp } \gamma \subset \text{graph}(\Pi_{[-\tau, \tau]})$  for any optimal  $\gamma$ . In particular, as  $\Pi_{[-\tau, \tau]}$  is uni-valued, it means that the optimal transport plan is unique and given by the projection map.  $\square$

**6.2. Competitor for the limit problem and existence for  $(P_\Lambda)$ .** Given  $y_0 \in \Sigma$  such that (6.1) holds, it follows from Theorem 6.1 that:

$$\bar{\nu}_\delta \stackrel{\text{def.}}{=} \theta_\delta(y_0) \mathcal{H}^1 \llcorner [-\tau, \tau] \in \arg \min F,$$

where  $F$  is defined in (6.8). In addition, Lemma 6.2 shows that if  $\theta_\delta(y_0) > 0$ , the optimal transportation of  $\bar{\sigma}_\delta$  to  $\bar{\nu}_\delta$  is given by the orthogonal projection. We show that in this case, we can lower the energy by projecting part of the mass to a (closer) horizontal line as in Figure 4.

This contradicts the existence of rectifiability points of  $\Sigma$  such that  $\theta_\delta(y_0) > 0$  so that  $\nu_\delta \equiv 0$ , and shows the following Lemma:

**Lemma 6.3.** *For any  $\delta > 0$ , the measures  $\nu_\delta$  defined in (5.14) vanish.*

*Proof.* Up to a rotation, we may assume that  $\tau = e_d$ , where  $(e_i)_{i=1}^d$  is a basis of  $\mathbb{R}^d$ . Since  $\bar{\sigma}_\delta$  is supported on  $\{x = (x', x_d) \in \mathbb{R}^d : |x'| > \delta, |x_d| \leq 1\}$ , we can cover

its support with finitely many sets  $(E_i)_{i=1}^N$  defined as:

$$E_i \stackrel{\text{def.}}{=} \{x = (x', x_d) \in \mathbb{R}^d : \langle \xi_i, x \rangle > \delta/2, |x_d| \leq 1\}$$

where  $\xi_i \in \mathbb{S}^{d-1} \cap [e_d]^\perp$  are ‘‘horizontal’’ unit vectors and  $N$  depends only on the dimension. We then define a disjoint family

$$F_1 = E_1, \quad F_{i+1} = E_{i+1} \setminus \bigcup_{j=1}^i F_j \text{ for } i \geq 1$$

and decompose our measures  $\bar{\sigma}_\delta$  and  $\bar{\nu}_\delta$  as

$$\bar{\sigma}_\delta = \sum_{i=1}^N \bar{\sigma}_{\delta,i}, \quad \bar{\nu}_\delta = \sum_{i=1}^N \bar{\nu}_{\delta,i} \text{ where } \bar{\sigma}_{\delta,i} \stackrel{\text{def.}}{=} \bar{\sigma}_\delta \llcorner F_i \text{ and } \bar{\nu}_{\delta,i} \stackrel{\text{def.}}{=} \text{proj}_{d\sharp} \bar{\sigma}_{\delta,i},$$

with  $\text{proj}_d : x \mapsto x_d e_d$  the projection onto the vertical axis. By Radon-Besicovitch’s differentiation theorem,  $\bar{\nu}_{\delta,i} = \theta_i \mathcal{H}^1 \llcorner [-e_d, e_d]$ , where  $\theta_i(s) = \theta_i(s e_d) \geq 0$  are such that

$$\sum_{i=1}^N \theta_i = \theta_\delta(y_0).$$

Consider  $\bar{s} \in (-1, 1)$  a common Lebesgue point of all  $\theta_i$ ,  $i = 1, \dots, N$ .

Let  $i$  be the index for which  $\theta_i(\bar{s})$  is maximal: then  $\theta_i(\bar{s}) \geq \theta_\delta(y_0)/N$ .

Up to a change of horizontal coordinates, we assume that  $\xi_i = e_1$ , and we introduce the notation:  $\mathbb{R}^d \ni x = (x_1, x'', x_d)$  for  $x'' \in \mathbb{R}^{d-2}$ . Let now:

$$C_\varepsilon \stackrel{\text{def.}}{=} F_i \cap \{x \in \mathbb{R}^d : |x_d - \bar{s}| < \varepsilon\} \subset \{x = (x_1, x'', x_d) : x_1 > \delta/2, |x_d - \bar{s}| < \varepsilon\}.$$

We obtain, from the fact that  $(\text{proj}_d)_\sharp \bar{\sigma}_{\delta,i} = \theta_i \mathcal{H}^1 \llcorner [-e_d, e_d]$ , that

$$\frac{\bar{\sigma}_{\delta,i}(C_\varepsilon)}{2\varepsilon} = \frac{1}{2\varepsilon} \int_{\bar{s}-\varepsilon}^{\bar{s}+\varepsilon} \theta_i(t) dt \xrightarrow{\varepsilon \rightarrow 0} \theta \stackrel{\text{def.}}{=} \theta_i(\bar{s}) \geq \frac{\theta_\delta(y_0)}{N}.$$

Now, assume by contradiction that  $\theta > 0$ . If  $\varepsilon$  is small enough, we have:

$$(6.10) \quad \theta \leq \frac{\bar{\sigma}_{\delta,i}(C_{\varepsilon'})}{\varepsilon'} \leq 3\theta.$$

for all  $\varepsilon' \leq \varepsilon$ . Now let us exploit the fact that, from Lemma 6.2, the optimal transport is given by projections to propose a new transport map, sending the mass in  $C_\varepsilon$  to a segment pointing towards  $e_1$ :

$$T(x) \stackrel{\text{def.}}{=} \begin{cases} \ell(|x_d - \bar{s}|)e_1 + \bar{s}e_d, & \text{if } x \in C_\varepsilon \\ \text{proj}_d(x), & \text{otherwise,} \end{cases}$$

where  $\ell : [0, \varepsilon] \rightarrow \mathbb{R}_+$  is defined via the conservation of mass relation, for  $0 \leq \varepsilon' \leq \varepsilon$ :

$$(6.11) \quad \ell(\varepsilon') = \alpha \bar{\sigma}_{\delta,i}(C_{\varepsilon'}).$$

In other words, the mass that was sent to the vertical segment  $[\bar{s} - \varepsilon', \bar{s} + \varepsilon']e_d$  is now sent to the horizontal segment  $\bar{s}e_d + [0, \ell(\varepsilon')]e_1$ , for each  $\varepsilon' \in [0, \varepsilon]$ . This construction is illustrated in Figure 4.

Thanks to (6.11), the map  $T$  sends  $\bar{\sigma}_{\delta,i} \llcorner C_\varepsilon$  to the measure  $\alpha^{-1} \mathcal{H}^1 \llcorner L$  where  $L \stackrel{\text{def.}}{=} \bar{s}e_d + [0, \ell(\varepsilon)]e_1$ , hence, the transported measure  $T_\sharp \bar{\sigma}_\delta$  satisfies the constraints in the definition (6.8) of the limiting functional  $F$  and one has  $F(T_\sharp \bar{\sigma}_\delta) < +\infty$ .

We shall now see that for each point  $x \in C_\varepsilon$  with  $x_d \neq \bar{s}$ , it holds that

$$(6.12) \quad |x - \text{proj}_d(x)|^p > |x - T(x)|^p.$$

To show (6.12), recalling the notation  $x = (x_1, x'', x_d)$ , it suffices that

$$\begin{aligned} |x - \text{proj}_d(x)|^2 &> |x - T(x)|^2 \\ \iff x_1^2 + |x''|^2 &> (x_1 - \ell(|x_d - \bar{s}|))^2 + |x''|^2 + (x_d - \bar{s})^2 \\ \iff 2x_1\ell(|x_d - \bar{s}|) &> \ell(|x_d - \bar{s}|)^2 + (x_d - \bar{s})^2. \end{aligned}$$

In addition to (6.10), we choose  $\varepsilon$  in such a way that for any  $x \in C_\varepsilon$  we have

$$\alpha\theta|x_d - \bar{s}| \leq \ell(|x_d - \bar{s}|) = \alpha\bar{\sigma}_{\delta,i}(C_{|x_d - \bar{s}|}) \leq 3\alpha\theta\varepsilon < \left(1 + \frac{1}{(\alpha\theta)^2}\right)^{-1} \delta$$

and hence

$$\ell(|x_d - \bar{s}|)^2 + (x_d - \bar{s})^2 \leq \left(1 + \frac{1}{(\alpha\theta)^2}\right) \ell(|x_d - \bar{s}|)^2 < \delta\ell(|x_d - \bar{s}|) \leq 2x_1\ell(|x_d - \bar{s}|),$$

for all  $x \in C_\varepsilon$ , with  $x_d \neq \bar{s}$ , so that (6.12) holds. Since  $\theta = \theta_i(\bar{s}) > 0$ , it follows that

$$F(T_{\sharp}\bar{\sigma}_\delta) = W_p^p(\bar{\sigma}_\delta, T_{\sharp}\bar{\sigma}_\delta) < W_p^p(\bar{\sigma}_\delta, \bar{\nu}_\delta) = F(\bar{\nu}_\delta).$$

This contradicts the fact that  $\theta_\delta(y_0)\mathcal{H}^1 \llcorner [-e_d, e_d]$  is a minimizer of  $F$ , showing that we must have  $\theta = \theta_i(\bar{s}) = 0$  and, in turn,  $\theta_\delta(y_0) = 0$ . As this holds for  $\mathcal{H}^1$ -a.e. point  $y_0 \in \Sigma$ , we deduce that  $\nu_\delta \equiv 0$ .  $\square$

The previous lemma, combined with the characterization of solutions, as in (5.15),

$$\nu = \alpha^{-1}\mathcal{H}^1 \llcorner \Sigma + \sup_{\delta>0} \nu_\delta + \rho_{\text{exc}} \llcorner \Sigma$$

proves the following result, showing in particular point (2) of Theorem 1.1.

**Theorem 6.4.** *Let  $\rho_0 \in \mathcal{P}_p(\mathbb{R}^d)$  and suppose that the parameter  $\Lambda < \Lambda_*$ . Then the solution to the relaxed problem  $(\bar{P}_\Lambda)$  is of the form*

$$\nu = \mathcal{L}(\nu)^{-1}\mathcal{H}^1 \llcorner \Sigma + \rho_{\text{exc}} \llcorner \Sigma,$$

where  $\rho_{\text{exc}}$  was defined in (5.4).

In addition, if  $\rho_0$  does not give mass to 1-rectifiable sets, any solution of the relaxed problem  $(\bar{P}_\Lambda)$  corresponds to a solution of the original shape optimization problem  $(P_\Lambda)$ .

## 7. AHLFORS REGULARITY

In this section we prove that whenever the initial measure  $\rho_0 \in L^{\frac{d}{d-1}}(\mathbb{R}^d)$ , the optimal solutions to the relaxed problem  $(\bar{P}_\Lambda)$  have an Ahlfors regular support.

**Definition 7.1.** We say that a set  $\Sigma \subset \mathbb{R}^d$  is *Ahlfors regular* whenever there exist  $r_0 > 0$  and  $c, C > 0$  such that for  $r \leq r_0$  it holds that

$$cr \leq \mathcal{H}^1(\Sigma \cap B_r(x)) \leq Cr, \text{ for all } x \in \Sigma.$$

We prove in this section the following result.

**Theorem 7.2.** *If  $\rho_0 \in L^{\frac{d}{d-1}}(\mathbb{R}^d)$ , let  $\nu$  be a solution of the relaxed problem  $(\bar{P}_\Lambda)$  and  $\Sigma$  its support. Then  $\Sigma$  is Ahlfors-regular: there exist  $\bar{r}_0 > 0$  and  $\bar{C} > 0$  such that, for all  $\bar{x} \in \Sigma$  and  $r \leq \bar{r}_0$ ,*

$$r \leq \mathcal{H}^1(\Sigma \cap B_r(\bar{x})) \leq \bar{C}r.$$

Moreover,  $\bar{r}_0$  depends only on  $d, p, \rho_0$  and  $\alpha \stackrel{\text{def.}}{=} \mathcal{L}(\nu)$ , while  $\bar{C}$  depends only on  $d$  and  $p$ .

The lower bound (with  $c = 1$  and  $r_0 = \text{diam } \Sigma$ ) follows directly from the connectedness of  $\Sigma$ . The upper bound will follow as a corollary of Lemma 7.3 below. Let us describe the strategy for proving this estimate. We point out that the construction in this section, although different, follows similar steps as the proof of Ahlfors' regularity in [27, Lem. 6.1, Thm. 6.4].

The idea is similar to proving the  $L^\infty$  bound on the excess measure: if in a small ball  $B_r(\bar{x})$  the measure  $\nu$  has too much mass, we build another "closer" 1D structure onto which the mass is transferred at a smaller cost. Yet there is an additional difficulty: when replacing  $\Sigma \cap B_r(\bar{x})$  with another set we must preserve the connectedness. The proof of Theorem 5.4, required to rearrange only the excess mass and this was not an issue. We now need to control the number of connected components of  $\Sigma \setminus B_r(\bar{x})$  and connect them back without adding too much length.

This number of connected components is controlled by the quantity  $\mathcal{H}^0(\Sigma \cap \partial B_r(\bar{x}))$ , which we can control on average by means of the *generalized area formula* [3, Theorem 2.91]: If  $f : \mathbb{R}^M \rightarrow \mathbb{R}^N$  is a Lipschitz function and  $E \subset \mathbb{R}^M$  is a  $k$ -rectifiable set then it holds that

$$(7.1) \quad \int_{\mathbb{R}^N} \mathcal{H}^0(E \cap f^{-1}(y)) d\mathcal{H}^k(y) = \int_E J_k d^E f_x d\mathcal{H}^k(x),$$

where  $d^E f_x$  is the restriction of  $\nabla f(x)$  (when  $f$  is smooth) to the approximate tangent space of  $E$ . Hence, choosing  $E = \Sigma \cap (B_{r_1}(\bar{x}) \setminus B_{r_2}(\bar{x}))$  and  $f : x \mapsto |x - \bar{x}|$ , we deduce from (7.1) that

$$(7.2) \quad \int_{r_2}^{r_1} \mathcal{H}^0(\Sigma \cap \partial B_s(\bar{x})) ds \leq \mathcal{H}^1(\Sigma \cap B_{r_1}(\bar{x})) - \mathcal{H}^1(\Sigma \cap B_{r_2}(\bar{x}))$$

Using this we first prove the following lemma:

**Lemma 7.3.** *Assume  $\rho_0 \in L^{\frac{d}{d-1}}(\mathbb{R}^d)$ . There exist  $\bar{C}(d, p) > 0$  and  $r_0$  depending on  $\rho_0, \alpha, d, p$ , such that for any  $C \geq \bar{C}$ , if  $r \leq r_0$  and  $x \in \Sigma$ , then either  $\mathcal{H}^1(\Sigma \cap B_r(x)) \leq Cr$  or  $\mathcal{H}^1(\Sigma \cap B_{2r}(x)) \geq 10Cr$ .*

*Proof.* Let  $r > 0$  and  $C \geq 1$ , and let  $\bar{x} \in \Sigma$  such that

$$(7.3) \quad \mathcal{H}^1(\Sigma \cap B_r(\bar{x})) > Cr \text{ and } \mathcal{H}^1(\Sigma \cap B_{2r}(\bar{x})) < 10Cr.$$

We show that if  $r \leq r_0$  and  $C \geq \bar{C}$ , which will both be chosen later, then we can construct a better competitor to the minimizer  $\nu$ .

The function  $f : s \mapsto \mathcal{H}^1(\Sigma \cap B_s(\bar{x}))$  is nondecreasing, hence in  $BV(\mathbb{R}_+)$  and satisfies, thanks to (7.2), that  $\mathcal{H}^0(\Sigma \cap \partial B_s(\bar{x})) ds \leq Df$  in the sense of measures (equivalently,  $\mathcal{H}^0(\Sigma \cap \partial B_s(\bar{x}))$  is less than, or equal to  $f'(s) ds$ , the absolutely continuous part of  $Df$ ).

We note that

$$\begin{aligned} \inf_{s \in (3r/2, 2r)} \left( \frac{s\mathcal{H}^0(\Sigma \cap \partial B_s(\bar{x}))}{\mathcal{H}^1(\Sigma \cap B_s(\bar{x}))} \right) &\leq \frac{2}{r} \int_{3r/2}^{2r} \frac{s\mathcal{H}^0(\Sigma \cap \partial B_s(\bar{x}))}{\mathcal{H}^1(\Sigma \cap B_s(\bar{x}))} ds \\ &\leq 4 \int_{3r/2}^{2r} \frac{1}{f(s)} f'(s) ds \\ &\leq 4 \ln \left( \frac{f(2r)}{f(3r/2)} \right), \end{aligned}$$

where we have used the classical chain rule at almost every point and [3, Cor. 3.29]. Since  $f(2r)/f(3r/2) < (10Cr)/(Cr) = 10$ , we deduce that there exists  $\bar{s} \in (3r/2, 2r)$  such that

$$(7.4) \quad \bar{s}\mathcal{H}^0(\Sigma \cap \partial B_{\bar{s}}(\bar{x})) \leq \mathcal{H}^1(\Sigma \cap B_{\bar{s}}(\bar{x})) \quad \text{where } \bar{\delta} \stackrel{\text{def.}}{=} \frac{1}{4 \ln 10}.$$

Now, we let

$$(7.5) \quad M = 2 \left( 1 + 10 \cdot \left( \frac{40}{17} \right)^{p-1} \right)$$

(this choice will be made clear at the end of this proof) and we consider

$$(7.6) \quad \delta \stackrel{\text{def.}}{=} \frac{\bar{\delta}}{10M} < \bar{\delta} < \frac{1}{2}.$$

We define a set  $\Gamma$  as follows: we choose a finite covering of  $\partial B_1(0)$  with balls  $B(x_i, \delta/2)$  centered at points  $(x_i)_{i=1}^N$  (the minimal number  $N$  depends only on  $d$  and  $p$ , through  $\delta$ ). Then, we find a minimal tree connecting the points  $(x_i)_{i=1}^N$  through geodesics on the sphere. We add to this minimal tree the segments  $[x_i, (1+\delta)x_i]$ ,  $i = 1, \dots, N$ . We call  $\Gamma$  the resulting (connected) set, whose total length  $L \stackrel{\text{def.}}{=} \mathcal{H}^1(\Gamma)$  is of order at most  $2N\delta$  and depends only on  $d$  and  $p$ . Notice that each point of  $\partial B_1$  is at distance at most  $\delta$ , along the geodesic curve on the sphere, to a point of  $\Gamma$ , and that thanks to the ‘‘spikes’’  $[x_i, (1+\delta)x_i]$ , any point with, say,  $|x| \geq 10$  is closer to a point of  $\Gamma$  than from any point in  $B_1(0)$ .

Now, we define

$$\Gamma_{\bar{s}} \stackrel{\text{def.}}{=} (\bar{x} + \bar{s}\Gamma) \cup \bigcup_{x \in \Sigma \cap \partial B_{\bar{s}}} S_x,$$

where  $S_x$  denotes a geodesic connecting  $x$  to  $\bar{x} + \bar{s}\Gamma$ , of length at most  $\mathcal{H}^1(S_x) \leq \bar{s}\delta$ . Since  $\bar{s} < 2r$  and  $\delta < 1/2$ , it follows that  $\Gamma_{\bar{s}} \subset B_{3r}(\bar{x})$ . We define the competitor set as

$$\Sigma' \stackrel{\text{def.}}{=} \Sigma \setminus B_{\bar{s}}(\bar{x}) \cup \Gamma_{\bar{s}}.$$

The addition of the geodesics  $S_x$  ensures that  $\Sigma'$  remains connected, and using (7.4), we estimate the length of  $\Gamma_{\bar{s}}$  as

$$(7.7) \quad \begin{aligned} \mathcal{H}^1(\Gamma_{\bar{s}}) &\leq L\bar{s} + \bar{s}\mathcal{H}^0(\Sigma \cap \partial B_{\bar{s}}(\bar{x})) \leq 2Lr + \frac{1}{10M}\mathcal{H}^1(\Sigma \cap B_{\bar{s}}(\bar{x})) \\ &< (2L + \frac{C}{M})r, \end{aligned}$$

where we have used (7.3) in the last estimate. Now we define a new competitor  $\nu'$  whose support is  $\Sigma'$ . If  $\gamma$  denotes an optimal transport plan from  $\rho_0$  to  $\nu$ , given

$s > 0$  let

$$\rho_s \stackrel{\text{def.}}{=} \pi_{0\sharp} (\gamma \llcorner (\mathbb{R}^d \times B_s))$$

denote the portion of the measure  $\rho_0$  which is transported to the ball  $B_s$ . In particular, the above length estimates imply that

$$(7.8) \quad Lr \leq \mathcal{H}^1(\Gamma_{\bar{s}}) < (2L + \frac{C}{M})r \leq (2\frac{L}{C} + \frac{1}{M})\alpha\nu(B_r) \leq \alpha\rho_r(\mathbb{R}^d) \leq \alpha\rho_{\bar{s}}(\mathbb{R}^d),$$

where  $\alpha \stackrel{\text{def.}}{=} \mathcal{L}(\nu)$ , and using that  $M \geq 2$  (see (7.5)) and assuming  $\bar{C} \geq 4L$  (which we recall depends only on  $d$  and  $p$ ). But, if  $r$  is small enough (not depending on  $\bar{x}$ , by uniform equi-integrability of  $\rho_0^{d/(d-1)}$ ) Hölder's inequality implies that

$$(7.9) \quad \alpha\rho_{\bar{s}}(B_{10r}(\bar{x})) \leq \alpha\|\rho_0\|_{L^{\frac{d}{d-1}}(B_{10r}(\bar{x}))} |B_{10r}(\bar{x})|^{\frac{1}{d}} < Lr.$$

We fix  $r_0 > 0$ , which depends only on the dimension (through  $L$ ), the integrability of  $\rho_0$ , and  $\alpha$ , such that the above inequality holds for  $r \leq r_0$ .

Equations (7.8)-(7.9) show that for  $r$  small enough, part of the mass transported to  $\nu \llcorner B_{\bar{s}}$  must come from outside of the ball  $B_{10r}$ . In particular, since  $t \mapsto \rho_{\bar{s}}(B_t(\bar{x}))$  is continuous, there is  $R > 10r$  such that

$$(7.10) \quad \rho_{\bar{s}}(B_R(\bar{x})) = \alpha^{-1}\mathcal{H}^1(\Gamma_{\bar{s}}).$$

To form the new competitor we proceed as follows: the mass sent to  $\Sigma \setminus B_{\bar{s}}$  remains untouched, the mass  $\rho_{\bar{s}} \llcorner B_R$  previously used to form  $\nu \llcorner B_{\bar{s}}$  is transported to  $\alpha^{-1}\mathcal{H}^1 \llcorner \Gamma_{\bar{s}}$  and the remaining mass is projected onto  $\Gamma_{\bar{s}}$ .

So, letting  $\tilde{\gamma}$  an optimal transport plan between  $\rho_{\bar{s}} \llcorner B_R$  and  $\alpha^{-1}\mathcal{H}^1 \llcorner \Gamma_{\bar{s}}$ , we define the plan

$$\gamma' = \gamma \llcorner \mathbb{R}^d \times B_{\bar{s}}(\bar{x})^c + \tilde{\gamma} \llcorner B_R \times \mathbb{R}^d + (\text{id}, \pi_{\Gamma_{\bar{s}}})_{\sharp} (\rho_{\bar{s}} \llcorner B_R^c),$$

and the new competitor  $\nu'$  as its second marginal. By construction,  $\alpha\nu' \geq \mathcal{H}^1 \llcorner \Sigma'$  so that  $\mathcal{L}(\nu') \leq \mathcal{L}(\nu)$ . We now estimate the gain in terms of transportation cost.

- For  $(x, y) \in B_R \times B_{\bar{s}}$  and for any  $y' \in \Gamma_{\bar{s}} \subset B_{3r}$ , as  $\bar{s} \leq 2r$  and  $10r < R$ , the convexity of  $t \mapsto t^p$  yields

$$\begin{aligned} |x - y'|^p &\leq (|x - y| + 5r)^p \leq |x - y|^p + 5rp(|x - y| + 5r)^{p-1} \\ &\leq |x - y|^p + 5rp(2R)^{p-1}. \end{aligned}$$

Hence integrating w.r.t. the transport plans we get

$$\int_{B_R \times \Gamma_{\bar{s}}} |x - y'|^p d\tilde{\gamma} \leq \int_{B_R \times B_{\bar{s}}} |x - y|^p d\gamma + 5rp(2R)^{p-1} \rho_{\bar{s}}(B_R),$$

(this can be checked by disintegration w.r.t. their common first marginal, which is the measure  $\rho_{\bar{s}} \llcorner B_R$ ).

- Similarly, for  $x \in B_R^c$  and  $y \in B_{\bar{s}} \setminus B_r$  the addition of the spikes ensures that

$$|x - \pi_{\Gamma_{\bar{s}}}(x)| \leq |x - y|.$$

However if  $x \in B_R^c$  and  $y \in B_r$  it holds that

$$|x - \pi_{\Gamma_{\bar{s}}}(x)| \leq |x - y| - \frac{r}{2} \text{ and } |x - y| \geq R - r,$$

so that once again using the convexity of  $t \mapsto t^p$  we have

$$\begin{aligned} |x - \pi_{\Gamma_{\bar{s}}}(x)|^p &\leq \left(|x - y| - \frac{r}{2}\right)^p \leq |x - y|^p - p\frac{r}{2} \left(|x - y| - \frac{r}{2}\right)^{p-1} \\ &\leq |x - y|^p - p\frac{r}{2} \left(\frac{17}{20}R\right)^{p-1}. \end{aligned}$$

So, decomposing the integration for the points going to  $B_r$  and to  $B_{\bar{s}} \setminus B_r$ , this time the transportation cost can be bound by:

$$\begin{aligned} \int_{B_{\bar{R}}^c} |x - \pi_{\Gamma_{\bar{s}}}(x)|^p d\rho_{\bar{s}} &= \int_{B_{\bar{R}}^c} |x - \pi_{\Gamma_{\bar{s}}}(x)|^p d(\rho_{\bar{s}} - \rho_r) + \int_{B_{\bar{R}}^c} |x - \pi_{\Gamma_{\bar{s}}}(x)|^p d\rho_r \\ &\leq \int_{B_{\bar{R}}^c \times B_{\bar{r}}} |x - y|^p d\gamma - p\frac{r}{2} \left(\frac{17}{20}R\right)^{p-1} \rho_r(B_{\bar{R}}^c). \end{aligned}$$

We get:

$$W_p^p(\rho_0, \nu') \leq W_p^p(\rho_0, \nu) + 5rp(2R)^{p-1} \rho_{\bar{s}}(B_R) - p\frac{r}{2} \left(\frac{17}{20}R\right)^{p-1} \rho_r(B_{\bar{R}}^c).$$

As  $\mathcal{L}(\nu') \leq \mathcal{L}(\nu)$ , the optimality of  $\nu$  gives that  $W_p^p(\rho_0, \nu) \leq W_p^p(\rho_0, \nu')$ , which, along with the previous estimates, implies

$$0 \leq 5 \cdot 2^{p-1} \rho_{\bar{s}}(B_R) - \frac{1}{2} \left(\frac{17}{20}\right)^{p-1} \rho_r(B_{\bar{R}}^c) \Leftrightarrow \rho_r(B_{\bar{R}}^c) \leq 10 \cdot \left(\frac{40}{17}\right)^{p-1} \rho_{\bar{s}}(B_R).$$

On the other hand, since

$$\rho_r(B_R(\bar{x})^c) = \nu(B_r(\bar{x})) - \rho_r(B_R(\bar{x})) \geq \alpha^{-1}Cr - \rho_r(B_R(\bar{x})) \geq \alpha^{-1}Cr - \rho_{\bar{s}}(B_R(\bar{x})),$$

and recalling (7.7) and (7.10), we deduce:

$$C \leq \left(1 + 10 \cdot \left(\frac{40}{17}\right)^{p-1}\right) \left(2L + \frac{C}{M}\right)$$

We conclude that with the choice (7.5) of  $M$ , one has  $C \leq 2ML$ , which depends only on  $p$  and  $d$  and a contradiction follows if we choose  $\bar{C} = 1 + 2ML$ .  $\square$

*Proof of Theorem 7.2.* Consider  $\bar{C}$ ,  $r_0$  from Lemma 7.3. Fix  $x \in \Sigma$  and assume there is  $r \in (0, r_0)$  such that  $\mathcal{H}^1(\Sigma \cap B_r(x)) > \bar{C}r$ . Then the thesis of the lemma applies and it must hold that  $\mathcal{H}^1(\Sigma \cap B_{2r}(x)) \geq 10\bar{C}r$ . By induction, we find that for  $k \geq 1$ , one of the following holds:

- either  $2^k r > r_0$ ;
- or we apply the lemma again (with  $C' = 5^k \bar{C}$  and  $r' = 2^k r$ ), using that  $\mathcal{H}^1(\Sigma \cap B_{2^k r}(x)) > 5^k \bar{C}(2^k r)$ , and we get

$$\mathcal{H}^1(\Sigma \cap B_{2^{k+1}r}(x)) > 5^{k+1} \bar{C}(2^{k+1}r).$$

Let  $k \geq 1$  be the first integer such that  $2^k r > r_0$ , so that  $2^{k-1}r \leq r_0$  and

$$5^k \bar{C}(2^k r) < \mathcal{H}^1(\Sigma \cap B_{2^k r}(x)).$$

Hence,  $r_0 < 2^k r \leq 5^{-k} \bar{C}^{-1} \mathcal{H}^1(\Sigma)$  and it holds that  $k \leq k_0 \stackrel{\text{def.}}{=} \log_5(\mathcal{H}^1(\Sigma)/\bar{C}r_0)$ , and

$$r > r_0 2^{-k} \geq \bar{r}_0 \stackrel{\text{def.}}{=} r_0 \cdot 2^{-k_0}.$$

This shows that, if  $r \leq r_0$  is such that  $\mathcal{H}^1(\Sigma \cap B_r(x)) > \bar{C}r$ , then  $r > \bar{r}_0$ . As a result, for every  $r \leq \bar{r}_0$  and every  $x \in \Sigma$ , we have  $\mathcal{H}^1(\Sigma \cap B_r(x)) \leq \bar{C}r$ .  $\square$

*Remark 7.4.* It is interesting to observe here that the regularity constant  $\bar{C}$  depends only on  $d$  and  $p$ , while the scale  $\bar{r}_0$  at which the Ahlfors-regularity holds gets smaller as  $\rho_0$  gets more singular or when  $\alpha$  (or  $\mathcal{H}^1(\Sigma)$ ) increases (which is when  $\Lambda$  decreases).

## 8. CONCLUSION

In this paper we have proposed a new variational problem, which serves as a method for approximating a probability measure with a measure uniformly distributed over a one-dimensional continuum. In order to prove existence, we have passed through a relaxed problem and the definition of a new functional on the space of probability measures, the length functional, that generalizes the notion of length of the support of a measure. As a tool for our analysis we have also generalized Gołab's Theorem to the case of a sequence of possibly unbounded sets converging in the Kuratowski sense. We then have shown that solutions of the relaxed problems are, in fact, solutions to the original one whenever the original measure does not give mass to 1-rectifiable sets of  $\mathbb{R}^d$ . We also have proved an elementary regularity properties of the optimal sets, in the form of an Ahlfors regularity estimate.

There are still many open questions left, such as:

- Does the support of minimizers have loops or are they trees?
- What is the regularity of the optimal  $\Sigma$ ? Can we adapt the theory in [25] and conclude they are locally  $C^{1,\alpha}$  curves?
- If  $\nu_\Lambda$  is a solution to  $(\bar{P}_\Lambda)$ , what is the rate of convergence of  $\nu_\Lambda \xrightarrow[\Lambda \rightarrow 0]{*} \rho_0$ ?
- The blow-up analysis in Section 6 is very similar to the arguments in [31] for the blow-up of average distance minimizers. However, the argument is applied to the excess measure and not to the entire solution. Can we use similar tools to study the blow-ups of the optimal networks in our problem as well?
- What are the Euler-Lagrange equations of  $(\bar{P}_\Lambda)$ ?
- Could we find (efficient) numerical algorithms to solve this problem?

Some progress has been made on a few of these questions: for instance in [21], it is proven in a simplified setting (when  $\rho_0$  is a finite sum of Dirac masses) that the solution is supported on a tree; in [22], a phase-field approach is suggested to approximate Problem  $(P_\Lambda)$ , which could lead to (still complicated) numerical methods and simulations.

## APPENDIX A. LOCALIZED VARIATIONAL PROBLEM

In this section, we prove Lemma 5.2, which states that the optimality of  $\nu$  implies that the exceeding measure  $\nu_{\text{exc}}$ , or a slight modification of it, must satisfy a localized optimization problem. Before proceeding we review the notation introduced in the statement of the Lemma. Given an optimal transportation plan  $\gamma$  between  $\rho_0$  and the minimizer  $\nu$ , we recall the definition of  $\gamma_{\text{exc}}$  in (5.3) and we fix a general



Borel set  $\mathcal{S} = \mathcal{S}_0 \times \mathcal{S}_1$  to define

$$\gamma_{\mathcal{S}} \stackrel{\text{def.}}{=} \gamma_{\text{exc}} \llcorner \mathcal{S}_0 \times \mathcal{S}_1$$

along with its marginals

$$\rho_{\mathcal{S}} \stackrel{\text{def.}}{=} \pi_{0\#} \gamma_{\mathcal{S}} \quad \nu_{\mathcal{S}} \stackrel{\text{def.}}{=} \pi_{1\#} \gamma_{\mathcal{S}},$$

**Proof of Lemma 5.2:** First, we fix some arbitrary  $\Gamma$  such that  $\Sigma \cup \Gamma \in \mathcal{A}$ . We consider measures  $\nu' \in \mathcal{M}_+(\Sigma \cup \Gamma)$  such that  $\nu'(\mathbb{R}^d) = \nu_{\mathcal{S}}(\mathbb{R}^d)$  and  $\nu' \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus \Sigma)$ , and we build competitors to  $\nu$  of the form  $\nu - \nu_{\mathcal{S}} + \nu'$ . Such measures are supported over  $\Sigma \cup \Gamma \in \mathcal{A}$  and

$$\begin{aligned} \nu - \nu_{\mathcal{S}} + \nu' &= \nu_{\mathcal{H}^1} + (\nu_{\text{exc}} - \nu_{\mathcal{S}}) + \nu' \\ &\geq \alpha^{-1} \mathcal{H}^1 \llcorner \Sigma + \alpha^{-1} \mathcal{H}^1 \llcorner (\Gamma \setminus \Sigma) \geq \alpha^{-1} \mathcal{H}^1 \llcorner (\Sigma \cup \Gamma), \end{aligned}$$

so that  $\mathcal{L}(\nu - \nu_{\mathcal{S}} + \nu') \leq \alpha = \mathcal{L}(\nu)$ . By optimality of  $\nu$ , we deduce that

$$W_p^p(\rho_0, \nu) \leq W_p^p(\rho_0, \nu - \nu_{\mathcal{S}} + \nu').$$

Given any transport plan  $\gamma'$  from  $\rho_{\mathcal{S}}$  to  $\nu'$ ,  $\gamma - \gamma_{\mathcal{S}} + \gamma'$  is a transport plan from  $\rho_0$  to  $\nu - \nu_{\mathcal{S}} + \nu'$  and it follows, from the optimality of  $\nu$  and  $\gamma$ :

$$\int |x-y|^p d(\gamma - \gamma_{\mathcal{S}}) + \int |x-y|^p d\gamma_{\mathcal{S}} = \int |x-y|^p d\gamma \leq \int |x-y|^p d(\gamma - \gamma_{\mathcal{S}}) + \int |x-y|^p d\gamma',$$

so that:

$$(A.1) \quad \int |x-y|^p d\gamma_{\mathcal{S}} \leq \int |x-y|^p d\gamma'.$$

Observe that in case  $\nu' = \nu_{\mathcal{S}}$  (and  $\Gamma = \emptyset$ ), we find that  $\gamma_{\mathcal{S}}$  is an optimal plan. In particular the left-hand side of this equation is  $W_p^p(\rho_{\mathcal{S}}, \nu_{\mathcal{S}})$ .

Since the same argument applies to  $\gamma - \gamma_{\mathcal{S}}$ , we observe that:

$$(A.2) \quad W_p^p(\rho_0, \nu) = W_p^p(\rho_0 - \rho_{\mathcal{S}}, \nu - \nu_{\mathcal{S}}) + W_p^p(\rho_{\mathcal{S}}, \nu_{\mathcal{S}}).$$

Considering an optimal transport plan  $\gamma'$  in (A.1), we get in addition that  $W_p^p(\rho_{\mathcal{S}}, \nu_{\mathcal{S}}) \leq W_p^p(\rho_{\mathcal{S}}, \nu')$  for all the admissible variations  $\nu'$  of the excess measure.

As  $\gamma_{\mathcal{S}}$  is an optimal transportation plan between  $\rho_{\mathcal{S}}$  and  $\nu_{\mathcal{S}}$ , from [30, Theorem 5.27] one can define a constant speed geodesic between such measures as

$$\sigma_{\mathcal{S},t} \stackrel{\text{def.}}{=} \pi_{(1-t)\#} \gamma_{\mathcal{S}}, \quad \text{where } \pi_t(x, y) \stackrel{\text{def.}}{=} (1-t)x + ty.$$

Hence for any variation  $\nu'$ , admissible in the sense of the previous problem, and for any  $t \in [0, 1]$ , it holds that

$$\begin{aligned} W_p(\rho_{\mathcal{S}}, \sigma_{\mathcal{S},t}) + W_p(\sigma_{\mathcal{S},t}, \nu_{\mathcal{S}}) &= W_p(\rho_{\mathcal{S}}, \nu_{\mathcal{S}}) \leq W_p(\rho_{\mathcal{S}}, \nu') \\ &\leq W_p(\rho_{\mathcal{S}}, \sigma_{\mathcal{S},t}) + W_p(\sigma_{\mathcal{S},t}, \nu'). \end{aligned}$$

Where the equality comes from general properties of constant speed geodesics in metric spaces, while the inequalities come from the minimality of  $\nu_{\mathcal{S}}$  and the triangle inequality, respectively. We conclude that in fact, the measure  $\nu_{\mathcal{S}}$  also minimizes the Wasserstein distance to any measure  $\sigma_{\mathcal{S},t}$  along the geodesic.  $\square$

## APPENDIX B. KURATOWSKI CONVERGENCE AND GOLAB'S THEOREM

In this appendix we give a proof of Lemma 2.1. We then give a simple proof of the local version of Gołab's, Theorem 2.2. We use the notation  $B_R = \{x : |x| < R\}$  and  $\bar{B}_R = \{x : |x| \leq R\}$ .

*Proof of Lemma 2.1.* Notice that, up to a translation, it suffices to prove the result for  $x_0 = 0$ . We can also assume that  $C \neq \emptyset$ , otherwise for any  $R > 0$ ,  $C_n \cap B_R = \emptyset$  for  $n$  large enough and the result holds. Defining  $R_0 = \inf\{R > 0 : C \cap \bar{B}_R \neq \emptyset\}$ , we have that if  $R < R_0$ , one has  $C_n \cap \bar{B}_R = \emptyset$  for  $n$  large enough and the Hausdorff limit is empty, as expected.

Now we take  $R \geq R_0$  and consider a subsequence  $(C_{n_k})_{k \in \mathbb{N}}$  and a closed set  $C^R$  such that

$$C_{n_k} \cap \bar{B}_R \xrightarrow[n \rightarrow \infty]{d_H} C^R.$$

Since  $C_{n_k} \cap \bar{B}_R \subset C_{n_k}$ , it holds that  $C^R \subset C$ . On the other hand, given  $x \in C \cap B_R$ , if there exists  $x_n \in C_n \cap \bar{B}_R$  with  $x_n \rightarrow x$ , then  $x \in C^R$ . Therefore

$$\overline{C \cap \bar{B}_R} \subset C^R \subset C \cap \bar{B}_R$$

and to finish the proof it suffices to show that there is a countable set  $I \subset [R_0, +\infty)$  such that if  $R \notin I$ ,  $R > R_0$ , then  $C \cap \bar{B}_R = \overline{C \cap \bar{B}_R}$ .

Let  $\xi \in \partial B_1$  and consider the function  $R \mapsto \text{dist}(R\xi, C \cap \bar{B}_R)$ . If  $R > R' \geq R_0$  it holds that

$$\text{dist}(R\xi, C \cap \bar{B}_R) \leq \text{dist}(R'\xi, C \cap \bar{B}_{R'}) + R - R'.$$

Indeed, let  $x_{R'}$  be the point minimizing the distance from  $R'\xi$  to  $C \cap \bar{B}_{R'}$ , then

$$\begin{aligned} \text{dist}(R\xi, C \cap \bar{B}_R) &\leq d(R\xi, x_{R'}) \leq d(R\xi, R'\xi) + d(R'\xi, x_{R'}) \\ &= \text{dist}(R'\xi, C \cap \bar{B}_{R'}) + R - R'. \end{aligned}$$

Hence the function  $\varphi_\xi : R \mapsto \text{dist}(R\xi, C \cap \bar{B}_R) - R$ , is nonincreasing in  $[R_0, +\infty)$  and in particular it has at most a countable number of discontinuity points. In addition, given  $\xi, \xi' \in \partial B_1$ , it holds that

$$\begin{aligned} |\varphi_\xi(R) - \varphi_{\xi'}(R)| &= \left| \inf_{x \in \bar{B}_R} d(x, R\xi) - \inf_{x \in \bar{B}_R} d(x, R\xi') \right| \\ &\leq \sup_{x \in \bar{B}_R} |d(x, R\xi) - d(x, R\xi')| \leq R|\xi - \xi'|. \end{aligned}$$

Therefore if  $R$  is a point of discontinuity for  $\varphi_\xi$ , then for all  $\xi'$  in a neighborhood of  $\xi$ ,  $R$  is a point of discontinuity for  $\varphi_{\xi'}$ .

Let  $(\xi_n)_{n \in \mathbb{N}}$  be a dense sequence in  $\partial B_1$ . For each  $n$  we can find a countable subset  $I_n \subset [R_0, +\infty)$ , such that  $\varphi_{\xi_n}$  is continuous at any  $R \in (R_0, +\infty) \setminus I_n$ . Finally, we define the countable set  $I$  as  $I = \bigcup_{n \in \mathbb{N}} I_n$ .

If  $R \notin I$ , then either  $R < R_0$  and  $C \cap \bar{B}_R = \overline{C \cap \bar{B}_R} = \emptyset$ , or  $R > R_0$ . In that case, for any  $\xi \in \partial B_1$ ,  $\varphi_\xi$  is continuous. Otherwise, there would be some  $\xi_n$ ,

close enough to  $\xi$ , such that  $\varphi_{\xi_n}$  is discontinuous, a contradiction. In particular, whenever  $x = R\xi \in C$  the continuity of  $\varphi_\xi$  implies that

$$\lim_{R' \uparrow R} \text{dist}(R'\xi, C \cap \overline{B_{R'}}) = 0.$$

Hence take  $R_n \uparrow R$ , set  $\varepsilon_n \stackrel{\text{def}}{=} \text{dist}(R_n\xi, C \cap \overline{B_{R_n}})$  and let  $x_n \in C \cap \overline{B_{R_n}}$  be a vector attaining this distance. As  $x_n \in C \cap B_R$  and  $|x - x_n| \leq \varepsilon_n + R - R_n$ ,  $x_n$  converges to  $x$ , and  $x \in \overline{C \cap B_R}$ . It follows that  $(C \cap \overline{B_R}) \setminus \overline{C \cap B_R} = \emptyset$ , completing the proof.  $\square$

*Proof of Theorem 2.2.* We will show that  $\mu(\Sigma \cap B_r(y_0)) \geq \mathcal{H}^1(\Sigma \cap B_r(y_0))$  for  $\mathcal{H}^1$ -a.e.  $y_0 \in \Sigma$  and for  $r > 0$  small enough. This implies that  $\Theta_1(\mu, y_0) \geq 1$ , and the result follows by integrating. Assume that  $\Sigma$  is not a singleton, otherwise there is nothing to prove. Since a compact and connected set with finite length is path-wise connected, see [10, Prop. 30.1 and Cor. 30.2] and [1, Thm. 4.4], for any  $y_0 \in \Sigma$ , for  $r > 0$  small enough  $\Sigma \cap B_r^c(y_0) \neq \emptyset$  and there is a path connecting  $y_0$  to the boundary  $\partial B_r(y_0)$  of length at least  $r$ . From the Kuratowski convergence, for  $n$  large enough, each set  $\Sigma_n$  has a point inside and another outside the ball  $\overline{B_r(y_0)}$ .

We start by fixing some  $0 < \delta < r$  and looking at the smaller ball  $B_{r-\delta}(y_0)$ . Consider the following class

$$\mathcal{A}_n \stackrel{\text{def}}{=} \left\{ \gamma \text{ connected component of } \Sigma_n \cap \overline{B_r(y_0)} \text{ which intersects } B_{r-\delta}(y_0) \right\}.$$

Each  $\gamma \in \mathcal{A}_n$  must be such that  $\mathcal{H}^1(\gamma) \geq \delta$ . Indeed, as for each  $n \in \mathbb{N}$  there is a point in  $\Sigma_n \cap B_r(y_0)^c$  and another in  $\gamma \cap \partial B_{r-\delta}(y_0)$ , the connectivity implies  $\gamma$  is contained in an arc joining these two points, but then it must have length at least  $\delta$ , as it is the smallest distance between the two balls. So define

$$\tilde{\Sigma}_n \stackrel{\text{def}}{=} \bigcup_{\gamma \in \mathcal{A}_n} \gamma,$$

which is a bounded sequence of closed sets, but not necessarily connected. However this sequence has a uniformly bounded number of connected components since

$$\delta \#\mathcal{A}_n \leq \sum_{\gamma \in \mathcal{A}_n} \mathcal{H}^1(\gamma) \leq \mathcal{H}^1(\Sigma_n \cap B_R(x_0)), \text{ hence } \#\mathcal{A}_n \leq \sup_{n \in \mathbb{N}} \frac{\mathcal{H}^1(\Sigma_n \cap B_R(y_0))}{\delta} < +\infty,$$

for  $R > 0$  large enough.

As  $\tilde{\Sigma}_n$  is a bounded sequence, by Blaschke's Theorem we can assume up to an extraction that  $\tilde{\Sigma}_n \xrightarrow[n \rightarrow \infty]{d_H} \tilde{\Sigma}$ . In fact, for a.e.  $0 < \delta < r$ , using Lemma 2.1, it holds that

$$(B.1) \quad \tilde{\Sigma} \cap \overline{B_{r-\delta}(y_0)} = \overline{\Sigma \cap B_{r-\delta}(y_0)},$$

since by the construction,  $\tilde{\Sigma}_n \cap \overline{B_{r-\delta}(y_0)} = \Sigma_n \cap \overline{B_{r-\delta}(y_0)}$  and choosing  $\delta$  such that  $\Sigma_n \cap \overline{B_{r-\delta}(y_0)} \xrightarrow[n \rightarrow \infty]{K} \Sigma \cap \overline{B_{r-\delta}(y_0)}$ .

This way, we can apply the global version of Gołab's Theorem with a uniformly bounded number of connected components to the sequence  $\tilde{\Sigma}_n \cap B_{r-\delta}(y_0)$  so that

we write

$$\begin{aligned}
\mu\left(\overline{B_r(y_0)}\right) &\geq \limsup_{n \rightarrow \infty} \mathcal{H}^1(\Sigma_n \cap B_r(y_0)) \geq \limsup_{n \rightarrow \infty} \mathcal{H}^1(\tilde{\Sigma}_n) \\
&\geq \liminf_{n \rightarrow \infty} \mathcal{H}^1(\tilde{\Sigma}_n \cap B_{r-\delta}) \\
&\geq \mathcal{H}^1(\tilde{\Sigma} \cap B_{r-\delta}(y_0)) = \mathcal{H}^1(\overline{\Sigma \cap B_{r-\delta}(y_0)}) \\
&\geq \mathcal{H}^1(\Sigma \cap B_{r-\delta}(y_0)),
\end{aligned}$$

where the first inequality is due to the local weak- $\star$  convergence of the measures and the forth is given by Gołab's Theorem. But as this estimate is true for any  $\delta > 0$ , it must hold that  $\mu\left(\overline{B_r(y_0)}\right) \geq \mathcal{H}^1(\Sigma \cap B_r(y_0))$  for any  $y_0 \in \Sigma$  and  $r > 0$ . To extend this to open balls as well we use the following estimates

$$\mu(B_r) = \lim_{n \rightarrow \infty} \mu\left(\overline{B_{r-1/n}}\right) \geq \lim_{n \rightarrow \infty} \mathcal{H}^1(\Sigma \cap B_{r-1/n}) = \mathcal{H}^1(\Sigma \cap B_r).$$

□

#### REFERENCES

- [1] Giovanni Alberti and Martino Ottolini. On the structure of continua with finite length and Gołab's semicontinuity theorem. *Nonlinear Analysis: Theory, Methods & Applications*, 153:35–55, 2017.
- [2] Luigi Ambrosio, Elia Brué, and Daniele Semola. Lectures on optimal transport, 2021.
- [3] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Courier Corporation, 2000.
- [4] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.
- [5] Luigi Ambrosio and Paolo Tilli. *Topics on analysis in metric spaces*, volume 25. Oxford University Press on Demand, 2004.
- [6] Andrea Braides. *Gamma-convergence for beginners*. Number 22 in Oxford lecture series in mathematics and its applications. Oxford University Press, New York, 2002.
- [7] Giuseppe Buttazzo and Eugene Stepanov. Optimal transportation networks as free dirichlet regions for the monge-kantorovich problem. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 2(4):631–678, 2003.
- [8] Nicolas Chauffert, Philippe Ciuciu, Jonas Kahn, and Pierre Weiss. A projection method on measures sets. *Constructive Approximation*, 45(1):83–111, 2017.
- [9] Gianni Dal Maso. *An Introduction to  $\Gamma$ -convergence*. Number v. 8 in Progress in nonlinear differential equations and their applications. Birkhäuser, Boston, MA, 1993.
- [10] Guy David. *Singular sets of minimizers for the Mumford-Shah functional*, volume 233. Springer Science & Business Media, 2006.
- [11] Camillo De Lellis. Lecture notes on rectifiable sets, densities, and tangent measures. *Preprint*, 23, 2006.
- [12] Sylvain Delattre and Aurélie Fischer. On principal curves with a length constraint. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 56, pages 2108–2140. Institut Henri Poincaré, 2020.
- [13] Martin Ehler, Manuel Gräf, Sebastian Neumayer, and Gabriele Steidl. Curve based approximation of measures on manifolds by discrepancy minimization. *Foundations of Computational Mathematics*, 21(6):1595–1642, 2021.
- [14] Herbert Federer. *Geometric measure theory*. Springer, 2014.
- [15] S. Gołab. Sur quelques points de la théorie de la longueur. *Annales de la Société Polonaise de Mathématiques*, VII:227–241, 1928. Rocznik polskiego tow.matematycznego.

- [16] Trevor Hastie and Werner Stuetzle. Principal curves. *Journal of the American Statistical Association*, 84(406):502–516, 1989.
- [17] Balázs Kégl, Adam Krzyzak, Tamás Linder, and Kenneth Zeger. Learning and design of principal curves. *IEEE transactions on pattern analysis and machine intelligence*, 22(3):281–297, 2000.
- [18] Léo Lebrat, Frédéric de Gournay, Jonas Kahn, and Pierre Weiss. Optimal transport approximation of 2-dimensional measures. *SIAM Journal on Imaging Sciences*, 12(2):762–787, 2019.
- [19] Antoine Lemenant. A presentation of the average distance minimizing problem. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 2011.
- [20] Xin Yang Lu and Dejan Slepčev. Average-distance problem for parameterized curves. *ESAIM: Control, Optimisation and Calculus of Variations*, 22(2):404–416, 2016.
- [21] João Miguel Machado. *Optimization in spaces of measures: Optimal Transport, Geometric Structures and Game Theory*. PhD thesis, École Doctorale SDOSE - PSL, 2024.
- [22] João Miguel Machado. Phase-field approximation for 1-dimensional shape optimization problems. Preprint HAL-04620380, June 2024.
- [23] Francesco Maggi. *Sets of finite perimeter and geometric variational problems: an introduction to Geometric Measure Theory*. Number 135. Cambridge University Press, 2012.
- [24] Jean-Michel Morel and Sergio Solimini. *Variational methods in image segmentation: with seven image processing experiments*, volume 14. Springer Science & Business Media, 2012.
- [25] Frank Morgan.  $(m, \varepsilon, \delta)$ -minimal curve regularity. *Proceedings of the American Mathematical Society*, pages 677–686, 1994.
- [26] Sebastian Neumayer and Gabriele Steidl. From optimal transport to discrepancy. *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging: Mathematical Imaging and Vision*, pages 1–36, 2021.
- [27] Emanuele Paolini and Eugene Stepanov. Qualitative properties of maximum distance minimizers and average distance minimizers in  $\mathbb{R}^n$ . *Journal of Mathematical Sciences*, 122(3):3290–3309, 2004.
- [28] Emanuele Paolini and Eugene Stepanov. Existence and regularity results for the steiner problem. *Calculus of Variations and Partial Differential Equations*, 46(3):837–860, 2013.
- [29] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.
- [30] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, 55(58-63):94, 2015.
- [31] Filippo Santambrogio and Paolo Tilli. Blow-up of optimal sets in the irrigation problem. *The Journal of Geometric Analysis*, 15:343–362, 2005.
- [32] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

*Email address:* `chambolle@ceremade.dauphine.fr`

*Email address:* `vincent.duval@inria.fr`

*Email address:* `joao-miguel.machado@ceremade.dauphine.fr`

CEREMADE, CNRS, UNIVERSITÉ PARIS-DAUPHINE, UNIVERSITÉ PSL, 75775 PARIS, FRANCE

MOKAPLAN, INRIA PARIS, 46 RUE BARRAULT, 75013 PARIS, FRANCE