

MAXIMIZERS OF NONLOCAL INTERACTIONS OF WASSERSTEIN TYPE

ALMUT BURCHARD, DAVIDE CARAZZATO, AND IHSAN TOPALOGLU

ABSTRACT. We characterize the maximizers of a functional involving the minimization of the Wasserstein distance between equal volume sets. This functional appears as a repulsive interaction term in some models describing biological membranes. We combine a symmetrization-by-reflection technique with the uniqueness of optimal transport plans to prove that balls are the only maximizers. Further, in one dimension, we provide a sharp quantitative version of this maximality result.

1. INTRODUCTION

In this paper we study a variational problem involving the Wasserstein distance between equal volume sets. Specifically, for any $p > 1$ we consider the following energy defined on subsets of \mathbb{R}^N :

$$\mathcal{W}_p(E) := \inf \left\{ W_p(\mathcal{L}^N \llcorner E, \mathcal{L}^N \llcorner F) : |F| = |E|, |E \cap F| = 0 \right\}, \quad (1.1)$$

where $W_p(\mu_1, \mu_2)$ is the p -Wasserstein distance between two measures $\mu_1, \mu_2 \in \mathbf{M}_+(\mathbb{R}^N)$ with $\mu_1(\mathbb{R}^N) = \mu_2(\mathbb{R}^N) < +\infty$. Here \mathcal{L}^N denotes the Lebesgue measure in \mathbb{R}^N , and for any measurable set $E \subset \mathbb{R}^N$, we use the notation $|E| = \mathcal{L}^N(E)$.

The functional (1.1) appears in [BCL20], where Buttazzo, Carlier and Laborde investigate the Wasserstein distance between two mutually singular measures for any $p \geq 1$. In particular, given a measure μ they prove that the infimum is achieved among measures that are singular with respect to μ . They also show that, when the admissible class consists of densities bounded by 1, the optimal solution is given by the characteristic function of a set.

In [BCL20] the authors also introduce and analyze the perimeter regularization of (1.1). Namely, they consider the problem

$$\inf \left\{ P(E) + \lambda W_p(\mathcal{L}^n \llcorner E, \mathcal{L}^n \llcorner F) : E, F \subset \mathbb{R}^N, |E \cap F| = 0, |E| = |F| = 1 \right\}, \quad (1.2)$$

and show, for any $\lambda > 0$, the existence of minimizers when admissible sets E and F are required to be subsets of a bounded domain Ω . This problem (with $p = 1$) is introduced by Peletier and Röger as a simplified model for lipid bilayer membranes where the sets E and F represent the densities of the hydrophobic tails and hydrophilic heads of the two part lipid molecules, respectively [PR09, LPR14]. The perimeter term accounts for an interfacial energy arising from hydrophobic effects, while the Wasserstein term models the weak bonding between the head and tail particles.

Date: September 12, 2023.

When posed over the unbounded space, Buttazzo, Carlier and Laborde prove the existence of minimizers for the problem (1.2) in two dimensions. Xia and Zhou [XZ21] extend this result to higher dimensions but under the additional assumptions that λ is sufficiently small and that $p < n/(n - 2)$. Recently, Novack, Venkatraman and the third author [NTV23] prove that minimizers to (1.2) exist in any dimension *and* for all values of $\lambda > 0$ and $p \in [1, \infty)$. Simultaneously, Candau-Tilh and Goldman [CTG22] also obtain the existence of minimizers via an alternative argument and characterize global minimizers in the small λ regime. The analysis in [CTG22] and [NTV23] show that there is a direct competition between the perimeter and the Wasserstein terms in (1.2). This, also as pointed out by Rupert Frank to the third author, leads to the question whether the functional (1.1) is *maximized* when the set E is a ball. We investigate this question in this paper *when* $p > 1$.

It often happens that we need to relax a functional to exploit some compactness. We denote by A_m the class of admissible densities with mass m that we use to relax the problem, i.e.,

$$A_m := \left\{ \rho \in L^1(\mathbb{R}^N) : 0 \leq \rho \leq 1, \int \rho dx = m \right\}.$$

We will use the shorthand notation $A := A_1$ when we deal with probability densities. We define the relaxation of (1.1) to densities ρ with $0 \leq \rho \leq 1$ as follows:

$$\mathcal{W}_p(\rho) := \inf \left\{ W_p(\rho, \rho') : 0 \leq \rho', 0 \leq \rho + \rho' \leq 1, \int \rho' dx = \int \rho dx \right\}. \quad (1.3)$$

Our main result in this paper is the following theorem.

Main Theorem. *The only maximizer of (1.3) in the class A_m , up to translations, is the characteristic function of a ball B with $|B| = m$.*

By [DPMSV16, Proposition 5.2] in the case $p = 2$, and by the same result combined with [BCL20, Theorem 3.10] and [CTG22, Proposition 2.1] in the case $p \neq 2$, the expression (1.3) extends the definition on sets given in (1.1). By these results, we also have that for any $\rho \in A_m$ there is a unique density η_ρ realizing (1.3) when $p > 1$. Note that, for $p > 1$ [Vil03, Theorem 2.44] guarantees that there is only one optimal transport plan π_ρ between ρ and η_ρ , and it is induced by a map.

The class of transport plans, which we will call *admissible plans*, that play a role in the definition of $\mathcal{W}_p(\rho)$ is given by

$$AP_\rho := \left\{ \pi \in M_+(\mathbb{R}^N \times \mathbb{R}^N) : (p_1)_\# \pi = \rho \mathcal{L}^N, \rho \mathcal{L}^N + (p_2)_\# \pi \leq \mathcal{L}^N \right\},$$

where $M(\mathbb{R}^N)$ denotes the set of signed Borel measures in \mathbb{R}^N , and $M_+(\mathbb{R}^N) \subset M(\mathbb{R}^N)$ denotes the set of non-negative measures. Here p_1 and p_2 are the two usual projections from $\mathbb{R}^N \times \mathbb{R}^N$ in \mathbb{R}^N . Notice that, thanks to the properties of the push-forward, it is automatically true that the density of $(p_2)_\# \pi$ with respect to \mathcal{L}^N belongs to A_m whenever $\rho \in A_m$ and $\pi \in AP_\rho$.

Remark 1.1. We point out that the energy $\mathcal{W}_p(\rho)$ can be defined whenever we have a metric space with a reference measure (in our case, the euclidean space \mathbb{R}^N endowed with \mathcal{L}^N). If (X, d) is a Polish metric space, and $\gamma \in \mathbf{M}_+(X)$ is a Borel measure, then for any density $\rho: X \rightarrow [0, 1]$ we can define its Wasserstein energy as

$$\mathcal{W}_p(\rho) := \inf \left\{ W_p(\rho\gamma, \rho'\gamma) : 0 \leq \rho', \rho + \rho' \leq 1, \int \rho' d\gamma = \int \rho d\gamma \right\},$$

and the p -Wasserstein distance can be defined in any metric space. We continue to denote by \mathbf{AP}_ρ the set of admissible plans, i.e.

$$\mathbf{AP}_\rho = \{ \pi \in \mathbf{M}_+(X \times X) : (p_1)_\# \pi = \rho\gamma, \rho\gamma + (p_2)_\# \pi \leq \gamma \}.$$

We cannot expect to have many invariance properties in an abstract setting, but some analytic-flavoured features could be retrieved in wide generality. We will not use this abstract formulation in this paper, with the exception of Proposition 3.3 where we consider the space $X = \mathbb{R}^+$ with a weight. This appears because in Section 3 we reduce to radial densities, and it is convenient to look at them as 1-dimensional densities (a weight pops up because of the coarea formula).

Plan of the paper. In Section 2 we introduce some preliminary results that are useful for the problem. After recalling briefly some well-known theorems about the existence and uniqueness of the optimal transport map, we introduce some very simple properties of the functional \mathcal{W}_p that were essentially already present in the literature for slightly different problems. In particular, Lemma 2.6 is devoted to the saturation of the constraint in a certain region, and Corollary 2.7 provides a uniform control on the transport distance. These two results are quite robust, as they do not require any geometric property of the Euclidean space, but just its metric-measure structure. Lemma 2.9 and Lemma 2.10 are an original contribution. The first one shows the continuity of the functional \mathcal{W}_p with respect to the weak* convergence (when there is no loss of mass), and it is fundamental to prove the existence of maximizers for \mathcal{W}_p . The second one, instead, shows that some symmetries of a density ρ can be inherited by the optimal plan π_ρ that realizes $\mathcal{W}_p(\rho)$. In Section 3 we deal with the maximizers of \mathcal{W}_p , whose existence is proved in Proposition 3.2 applying the concentration compactness principle. This is a building block also for our successive characterization of the maximizers, since we combine a symmetrization technique and the uniqueness of the optimal transport plan to show that the maximizers have some symmetry. In fact, our plan to characterize them is the following:

- (i) prove that the segments maximize a 1-dimensional weighted version of \mathcal{W}_p , in Proposition 3.3;
- (ii) prove that, if ρ is a given maximizer, then the optimal transport plan realizing $\mathcal{W}_p(\rho)$ is radial. This is contained in Corollary 3.5, as a consequence of Lemma 3.4;
- (iii) combine the first two points to show that the maximizers have to be star-shaped sets, and then conclude that the ball is the only possible maximizer thanks to the

saturation of the constraint exposed in Lemma 2.6. This is contained in Theorem 3.6, and it is our main contribution.

Finally, in Section 4 we prove a quantitative version of this maximality result in one dimension, where we show that the deficit of maximality is controlled from below by the square of an asymmetry given as the L^1 distance between the ball and any density. Our inequality is asymptotically sharp, in the sense that the exponent of the asymmetry cannot be lowered.

A few days before submitting this paper, we became aware of the independent work by Candau-Tilh, Goldman and Merlet [CTGM] (posted on arXiv on September 6, 2023) studying the same maximization problem. Their result is more general, as it considers a broader class of cost functions in the transport problem. Our strategy, pursued in Section 3, instead, is more geometric, and we circumvent the need to introduce Kantorovich potentials to deal with the transport problem.

Notation. Throughout the paper, with an abuse of notation, we will denote the Wasserstein distance between two disjoint set, $W_p(\mathcal{L}^N \llcorner E, \mathcal{L}^N \llcorner F)$, by $W_p(E, F)$. By $B_r(x)$ we will denote the open ball of center x and radius r , and we will write B_r for $B_r(0)$. The cube of side length $2l$ centered at the origin will be denoted by $Q_l = [-l, l]^N \subset \mathbb{R}^N$; hence, $Q_l(x) = x + Q_l$. For $\rho \in \mathbf{A}$ by η_ρ we will denote any density in \mathbf{A} such that $\mathcal{W}_p(\rho) = W_p(\rho, \eta_\rho)$. Note that for $p > 1$ we have that η_ρ is unique (cfr. [BCL20, Remark 3.11]). Similarly, for $\rho \in \mathbf{A}$, π_ρ will denote the optimal plan $\mathcal{W}_p^p(\rho) = \int |x - y|^p d\pi_\rho(x, y)$, and T_ρ is the optimal transport map that induces π_ρ . If we have a density f , we will sometimes use the short-hand notation $T_\# f$ to denote the push forward of the measure $T_\#(f \mathcal{L}^N)$.

2. PRELIMINARY RESULTS

2.1. The optimal transport problem. We introduce in this section the optimal transport problem. The general theory is well developed, and goes far beyond the needs of this paper. We state some results, and we define the optimal transport problem, just in the setting that we need. The interested reader may find much more general statements, and much deeper developments, in the references that we cite, as well as in other books on the subject. Instead, one of the crucial restrictions that we impose is to work with cost $c(x) = |x|^p$ with $p > 1$ and (mostly) in the Euclidean space \mathbb{R}^N . This is necessary when we characterize the maximizers of \mathcal{W}_p since we use some uniqueness result valid for these special cost functions, while some parts of our strategy work also for $p = 1$ with a slightly different discussion. The next definitions describe rigorously our framework.

A quite general setting for the optimal transport problem is that of Polish metric spaces, that are defined as follows.

Definition 2.1 (Polish metric space). A metric space (X, d) is *Polish* if it is complete and separable.

Definition 2.2 (Push forward). Let (X, d_X) and (Y, d_Y) be two Polish metric spaces. Given $f : X \rightarrow Y$ a Borel function, and given a measure $\mu \in \mathbf{M}(X)$, the *push forward* of μ induced by f is a new measure denoted by $f_{\#}\mu$. It is defined as follows: for every $A \subset Y$ Borel, we have that

$$(f_{\#}\mu)(A) = \mu(f^{-1}(A)).$$

Given (X, d) a Polish metric space, $p > 1$ a real exponent, and given $\mu_1, \mu_2 \in \mathbf{M}_+(X)$ with $\mu_1(X) = \mu_2(X) < +\infty$, we can consider the optimal transport problem with cost $c(x, y) = |x - y|^p$:

$$W_p^p(\mu_1, \mu_2) = \inf \left\{ \iint_{X \times X} |x - y|^p d\pi(x, y) : \pi \in \mathbf{M}_+(X \times X) : (p_1)_{\#}\pi = \mu_1, (p_2)_{\#}\pi = \mu_2 \right\}.$$

It is well known that for every couple of marginals μ_1 and μ_2 the infimum is attained (see [Vil03, Theorem 1.3] for a more general result). In some special cases, there are some structure theorems for the optimal transport plans, i.e. those measures π that realize the aforementioned infimum. The following is such a result that holds for strictly convex costs.

Theorem 2.3. [Vil03, Theorem 2.44] *Let $p > 1$ be given, and $\mu_1, \mu_2 \in \mathbf{M}_+(\mathbb{R}^N)$ be two measures with $\mu_1(\mathbb{R}^N) = \mu_2(\mathbb{R}^N) < +\infty$. Suppose that $\mu_1 \ll \mathcal{L}^N$ and that $W_p(\mu_1, \mu_2) < +\infty$. Then, there is a unique optimal transport plan π , and it is of the form*

$$\pi = (\text{Id}, T)_{\#}\mu_1,$$

where T denotes the unique optimal transport map.

In Section 3 it is crucial to characterize the maximizers in one dimension to later pass to higher dimension. Our task is simplified in one dimension because the transport problem has a very easy solution.

Theorem 2.4. [Vil03, Remarks 2.19] *Let $p > 1$ be given, and let $\mu_1, \mu_2 \in \mathbf{M}_+(\mathbb{R})$ be two measures with $\mu_1(\mathbb{R}) = \mu_2(\mathbb{R}) < +\infty$. If they are non-atomic, then the only optimal transport map realizing $W_p(\mu_1, \mu_2)$ is monotone.*

2.2. Properties of \mathcal{W}_p . The most basic fact is the following existence theorem.

Theorem 2.5. [DPMSV16, Section 5] *Let $p > 1$ be given. For any $m > 0$ and for any $\rho \in \mathbf{A}_m$, there exists a unique density, called $\eta_\rho \in \mathbf{A}_m$, realizing the infimum in (1.3).*

Combining this result with Theorem 2.3 we obtain the existence and uniqueness of the optimal transport plan π_ρ and the map inducing it, called T_ρ , which satisfy

$$\mathcal{W}_p^p(\rho) = W_p^p(\rho, \eta_\rho) = \int |x - y|^p d\pi_\rho(x, y) = \int |x - T_\rho(x)|^p \rho(x) dx.$$

We point out that the objects η_ρ , π_ρ and T_ρ all depend implicitly on p . We do not stress that dependence because we suppose $p > 1$ to be fixed in the whole paper.

One important result contains a geometric property of the optimal plan π_ρ . The proof of the following lemma is purely metric, and concerns mostly the structure of η_ρ , rather than

the optimal transport problem that is hidden in \mathcal{W}_p . Indeed, we do not exploit the c -cyclical monotonicity of the optimal plans. This result is a natural generalization of [DPMSV16, Lemma 5.1].

Lemma 2.6. *Let (X, d) be a Polish metric space, and let $\gamma \in \mathbf{M}_+(X)$ be a given measure. Let $\rho: X \rightarrow [0, 1]$ be a Borel density. If π is an optimal plan to compute $\mathcal{W}_p(\rho)$ and $(x, y) \in \text{spt}\pi$, then*

$$(p_2)_\# \pi = (1 - \rho)\gamma \quad \gamma - \text{a.e. in } B_{|y-x|}(x). \quad (2.1)$$

Moreover, we have that $(p_2)_\# \pi \geq \min\{1 - \rho, \rho\}\gamma$.

Proof. We start proving that $(p_2)_\# \pi$ saturates the constraint in the ball, and the second statement will follow easily. The idea is very simple: if π does not saturate the constraint in that ball, then we can lower the energy of ρ adding some mass close to x . We define $r = |y - x|$. Let us suppose by contradiction that there exist $\varepsilon, \delta > 0$ and a set $E \subset B_{r-4\delta}(x)$ with $\gamma(E)$ strictly positive and finite and such that

$$(1 - \rho)\gamma - (p_2)_\# \pi \geq \varepsilon\gamma \quad \text{in } E.$$

We take $\mu_1 = (p_1)_\#(\pi \llcorner B_\delta(x) \times B_\delta(y))$ and $\mu_2 = \varepsilon\gamma \llcorner E$, and we modify π in the following way: we take $0 < t < \min\{1, \mu_1(X)/\mu_2(X)\}$, and we take

$$\tilde{\pi} = \pi - t \frac{\mu_2(X)}{\mu_1(X)} \pi \llcorner (B_\delta(x) \times B_\delta(y)) + \frac{t}{\mu_1(X)} \mu_1 \times \mu_2.$$

One can check that $\tilde{\pi} \in \text{AP}_\rho$ thanks to our choice of t . Since π is an optimal plan to compute $\mathcal{W}_p(\rho)$, we have that

$$\begin{aligned} 0 &\leq \int |x - y|^p (d\tilde{\pi} - d\pi) = -t \frac{\mu_2(X)}{\mu_1(X)} \int_{B_\delta(x) \times B_\delta(y)} |x - y|^p d\pi + \frac{t}{\mu_1(X)} \int |x - y|^p d\mu_1 d\mu_2 \\ &\leq -t \frac{\mu_2(X)}{\mu_1(X)} (r - 2\delta)^p \mu_1(X) + \frac{t}{\mu_1(X)} (r - 4\delta + \delta)^p \mu_1(X) \mu_2(X) \\ &= t \mu_2(X) [(r - 3\delta)^p - (r - 2\delta)^p] < 0, \end{aligned}$$

and thus we reach a contradiction.

We now address the last inequality. Suppose by contradiction that the opposite inequality holds in a set $E \subset X$ with $\int_E \rho d\gamma > 0$. Then, thanks to what we have proved so far, we know that the set

$$\{x \in E: \text{spt}\pi \cap (\{x\} \times X) = (x, x)\} \quad (2.2)$$

has full γ -measure in E . In fact, if this was not the case, then we could find $E' \subset E$ with $\gamma(E') > 0$ and such that, for every $x \in E'$, there exists $y \in X \setminus \{x\}$ such that $(x, y) \in \text{spt}\pi$. Then, using (2.1) we find an open covering of E' where the contradiction hypothesis is not satisfied, against the definition of E . Condition (2.2) means that we are not moving mass in E , and thus

$$(p_2)_\#(\pi \llcorner (E \times X)) = (p_1)_\#(\pi \llcorner (E \times X)) = \rho \chi_E \gamma.$$

This is sufficient to conclude since $(p_2)_\# \pi \geq (p_2)_\#(\pi \llcorner (E \times X)) = \rho \chi_E \gamma$, that is incompatible with our contradiction hypothesis. \square

Corollary 2.7. *Let us consider the functional \mathcal{W}_p on the Euclidean space \mathbb{R}^N with the usual metric and the Lebesgue measure \mathcal{L}^N . There exists a constant $C_N < +\infty$ such that, for any $\rho \in \mathbf{A}_m$ and for any $(x, y) \in \text{spt}\pi_\rho$, we have that*

$$|x - y| \leq C_N m^{\frac{1}{N}}, \quad (2.3)$$

where π_ρ is any optimal transport plan associated to ρ and η_ρ . Therefore, we also have that $\mathcal{W}_p^p(\rho) \leq C m^{1+\frac{p}{N}}$.

Proof. This is a consequence of Lemma 2.6. In fact, if we fix $r > 0$ such that $|B_r| = 3m$, then for any $\rho' \in \mathbf{A}_m$ and for any $z \in \mathbb{R}^N$ we have that

$$\int_{B_r(z)} \rho + \rho' dx \leq 2m.$$

Therefore, the condition (2.1) is not satisfied for any couple of points $(x, y) \in \text{spt}\pi_\rho$ with $|x - y| > r$. This is precisely the required estimate, since we bound the transport distance with a quantity proportional to the radius of a ball with mass m . \square

Remark 2.8. We report here the scaling property of the energy \mathcal{W}_p , that is already stated in [NTV23, Lemma 2.5] for sets. Let ρ be a density satisfying the constraint $0 \leq \rho \leq 1$ and let $t > 0$ be a given constant. If we consider $\tilde{\rho}(x) = \rho(x/t)$, then we have that $\mathcal{W}_p^p(\tilde{\rho}) = t^{p+N} \mathcal{W}_p^p(\rho)$. In fact, it is sufficient to consider the density $\eta_\rho(\cdot/t)$, rescaling appropriately the transport map.

Lemma 2.9 (Continuity of \mathcal{W}_p). *Let $\rho \in \mathbf{A}_m$ be a given density and let $\{\rho_n\}_{n \in \mathbb{N}} \subset \mathbf{A}_m$ be a sequence such that $\rho_n \xrightarrow{*} \rho$. Then, the limit of $\mathcal{W}_p(\rho_n)$ exists and $\mathcal{W}_p(\rho) = \lim_n \mathcal{W}_p(\rho_n)$.*

Proof. We prove this proposition in two steps. In the first step we establish that for any $p \geq 1$ (1.3) is the lower semicontinuous envelope of the functional in (1.1) in the class \mathbf{A}_m with respect to the weak-* topology. As a consequence, \mathcal{W}_p is lower semicontinuous in \mathbf{A}_m . In the second step we obtain the upper semicontinuity of \mathcal{W}_p in \mathbf{A}_m .

Step 1. Thanks to Remark 2.8 we can consider only the case $m = 1$. Let $\{E_n\}_{n \in \mathbb{N}}$ be a sequence of sets with $|E_n| = 1$ such that $E_n \xrightarrow{*} \rho$ for some $\rho \in \mathbf{A}$, and let us call $\rho_n = \chi_{E_n}$. Since we preserve the total mass, we know that for any $\varepsilon > 0$ there exist $R > 0$ and $k \in \mathbb{N}$ such that $\int_{B_R} \rho_n dx > 1 - \varepsilon$ for every $n > k$. Using Corollary 2.7 we know that the transport distance is uniformly bounded by a constant C , and thus $\int_{B_{R+C}} \eta_{\rho_n} dx \geq 1 - \varepsilon$ for any $n > k$. Therefore, up to a subsequence, we have that also $\eta_{\rho_n} \xrightarrow{*} \rho'$ for some density ρ' with $\int \rho' dx = 1$. It is then easy to see that $\rho + \rho' \leq 1$ almost everywhere, and thus

$$\mathcal{W}_p(\rho) \leq W_p(\rho, \rho') \leq \liminf_n W_p(\rho_n, \eta_{\rho_n}) = \mathcal{W}_p(\rho_n),$$

where we used the well-known lower semicontinuity of the Wasserstein distance (it is sufficient to take the weak limit of the optimal transport plans). This proves that the functional in (1.3) is smaller than the lower semicontinuous envelope of \mathcal{W}_p with respect to the weak* topology.

Next, we will find a sequence that realizes the equality, proving that our definition of $\mathcal{W}_p(\rho)$ in **A** is the lower semicontinuous envelope of the functional defined in (1.1).

Given $\rho \in \mathbf{A}$, for any $n \in \mathbb{N}$ we consider a partition of \mathbb{R}^N with a family of cubes $\mathcal{F}_n = \{Q_n^k\}_{k \in \mathbb{N}}$ with diameter $1/n$. Thanks to the compatibility condition $\rho + \eta_\rho \leq 1$, for any n we can find two sets E_n and F_n with $|E_n \cap F_n| = 0$ and such that

$$|E_n \cap Q_n^k| = \int_{Q_n^k} \rho \, dx, \quad |F_n \cap Q_n^k| = \int_{Q_n^k} \eta_\rho \, dx, \quad \forall Q_n^k \in \mathcal{F}_n.$$

It is immediate to see that $E_n \xrightarrow{*} \rho$ and $F_n \xrightarrow{*} \eta_\rho$ as $n \rightarrow +\infty$. Recalling $m = 1$, we also note that $W_p(E_n, \rho) \leq \text{diam}(Q_n^k)$ and $W_p(\eta_\rho, F_n) \leq \text{diam}(Q_n^k)$. To see this, it is sufficient to consider the (non-optimal) transport plan given by

$$\pi_n = \sum_{k \in \mathbb{N}} \frac{1}{|E_n \cap Q_n^k|} (\chi_{E_n \cap Q_n^k} \mathcal{L}^N) \times (\rho \chi_{Q_n^k} \mathcal{L}^N) \in \mathcal{P}(\mathbb{R}^N \times \mathbb{R}^N), \quad (2.4)$$

and notice that $|x - y| \leq \text{diam}(Q_n^k) = 1/n$ for any $(x, y) \in \text{spt} \pi_n$. The proof of the inequality for F_n and η_ρ is analogous, and thus we obtain that

$$W_p(E_n, F_n) \leq W_p(E_n, \rho) + W_p(\rho, \eta_\rho) + W_p(\eta_\rho, F_n) \leq \frac{2}{n} + W_p(\rho, \eta_\rho).$$

This, combined with the first part, shows that

$$\mathcal{W}_p(\rho) = \inf_{E_n \xrightarrow{*} \rho, |E_n|=m} \liminf_n W_p(E_n) \quad \forall \rho \in \mathbf{A}.$$

Step 2. We remind that, thanks to Theorem 2.3, there exists an optimal transport map for every transport problem that we consider in this paper. Up to taking a subsequence, we can suppose that $\lim_n W_p(\rho_n)$ exists, and we prove that $\mathcal{W}_p(\rho) = \lim_n W_p(\rho_n)$. Since we can extract one of such subsequence from any subsequence of $\{\rho_n\}_n$, this guarantees the existence of that limit for the whole sequence. We proceed by contradiction, and we suppose that there exists $\delta > 0$ such that $\mathcal{W}_p(\rho) < \lim_n W_p(\rho_n) - \delta$. The idea is to modify η_ρ and produce a competitor to compute $W_p(\rho_n)$, proving that we cannot have a strict inequality. To proceed with this plan we first truncate the densities to guarantee a convergence in Wasserstein distance. Up to taking another subsequence, we can suppose that $\eta_{\rho_n} \xrightarrow{*} \rho'$ for some $\rho' \in \mathbf{A}$ with $\rho + \rho' \leq 1$ (using the same argument as in Step 1). Since the sequences $\{\rho_n\}_n$ and $\{\eta_{\rho_n}\}_n$ do not lose mass, for any $\varepsilon < 1/2$ there exists $\bar{n}, k_1 \in \mathbb{N}$ such that

$$\int_{\mathbb{R}^N \setminus Q_{3k_1}} (\rho_n + \eta_{\rho_n}) \, dx < \varepsilon \quad \forall n > \bar{n}. \quad (2.5)$$

We will choose ε later on in order to make some approximations precise enough to obtain a contradiction out of the strict inequality.

Now take $k_2 = \lceil 3/\varepsilon \rceil$, so that $k_2 \varepsilon \in [3, 3 + \varepsilon]$, and we consider the cube $\bar{Q} = [-k_1 k_2 \varepsilon, k_1 k_2 \varepsilon]^N$. It is easy to see that we can partition \mathbb{R}^N with a family $\mathcal{F} = \{Q^k\}_{k \in \mathbb{N}}$ of cubes with side length equal to ε and such that $|Q^k \cap \bar{Q}| \in \{0, \varepsilon^N\}$ (i.e. \mathcal{F} contains two disjoint subfamilies that partition \bar{Q} and $\mathbb{R}^N \setminus \bar{Q}$). Moreover, it is also possible to find a partition of $\mathbb{R}^N \setminus \bar{Q}$

with a family $\tilde{\mathcal{F}} = \{\tilde{Q}^k\}_{k \in \mathbb{N}}$ of cubes with side length $k_2\varepsilon$. We will use the first partition to control the cost of an approximation of η_ρ inside \bar{Q} , where we move mass at short distance. The second one, instead, will be used to estimate the energy carried by the mass outside of that cube (thanks to (2.5), that mass is small). We call T the optimal transport map between ρ and η_ρ , and for any n we define the truncated densities $\tilde{\rho}_n = \rho_n \chi_{\bar{Q}}$. For any n we also take $L_n > 0$ such that $\int_{Q_{L_n}} \rho \, dx = \int \tilde{\rho}_n \, dx$, and we define the densities $\zeta_n := \rho \chi_{Q_{L_n}}$ and $\zeta'_n := (T_\rho)_\# \zeta_n$. Since $\rho_n \xrightarrow{*} \rho$, then $\tilde{\rho}_n \xrightarrow{*} \rho \chi_{\bar{Q}}$ and we can choose the sequence $\{L_n\}_n$ to be bounded. Moreover, we have that $\zeta_n \xrightarrow{*} \rho \chi_{\bar{Q}}$. Since the supports of the truncated densities are equibounded, then the p th-moment of ζ_n converges, as well as the p th-moment of $\tilde{\rho}_n$, and thus $W_p(\tilde{\rho}_n, \zeta_n) \rightarrow 0$ (see e.g. [Vil03, Theorem 7.12])

We take h_n^1 any non-negative density such that $\rho_n + h_n^1 \leq 1$ and for any $k \in \mathbb{N}$

$$\int_{Q^k} h_n^1 \, dx = \min \left\{ \int_{Q^k} \zeta'_n \, dx, \int_{Q^k} 1 - \rho_n \, dx \right\}.$$

Since $\zeta'_n = (T_\rho)_\# \zeta_n$, we can apply Corollary 2.7 and see that $\text{spt} h_n^1$ is contained in Q_{L_n+C} for any n , where C is a constant depending only on N . Since $\tilde{\rho}_n \xrightarrow{*} \rho \chi_{\bar{Q}}$ and $\zeta'_n \xrightarrow{*} (T_\rho)_\#(\rho \chi_{\bar{Q}})$, then we have that $\|h_n^1\|_1 - \|\zeta'_n\|_1 \rightarrow 0$ (notice that here only a finite number of cubes in \mathcal{F} play an active role). We choose any non-negative density h_n^2 with $\text{spt} h_n^2 \subset 3\bar{Q}$ and such that

$$\rho_n + h_n^1 + h_n^2 \leq 1 \quad \text{and} \quad \|h_n^1 + h_n^2\|_1 = \|\tilde{\rho}_n\|_1,$$

and our candidate to compute $\mathcal{W}_p(\tilde{\rho}_n)$ will be $\tilde{\rho}'_n := h_n^1 + h_n^2$. Observe that, by definition of h_n^1 and thanks to the properties of the push-forward of measures, we have that $\|h_n^1\|_1 \leq \|\zeta'_n\|_1 = \|\zeta_n\|_1 = \|\tilde{\rho}_n\|_1$. Thanks to the triangle inequality for the p -Wasserstein distance, we have that

$$W_p(\tilde{\rho}_n, \tilde{\rho}'_n) \leq W_p(\tilde{\rho}_n, \zeta_n) + W_p(\zeta_n, \zeta'_n) + W_p(\zeta'_n, \tilde{\rho}'_n).$$

The first term on the right hand side is going to 0 because, as we already noticed, the sets $\text{spt} \tilde{\rho}_n$ and $\text{spt} \zeta_n$ are uniformly bounded and these densities are converging to $\rho \chi_{\bar{Q}}$. Hence, up to taking \bar{n} large enough, we can suppose that $W_p(\tilde{\rho}_n, \zeta_n) < \varepsilon$. Likewise, the last term is controlled by ε , and we use a plan similar to (2.4) to show this.

We choose a density $\zeta''_n \leq \zeta'_n$ such that

$$\int_{Q^k} \zeta''_n \, dx = \int_{Q^k} h_n^1 \, dx \quad \forall k \in \mathbb{N},$$

and we consider the plan

$$\tilde{\pi}_n = \sum_{k \in \mathbb{N}} \frac{1}{\|h_n^1 \chi_{Q^k}\|_1} (\zeta''_n \chi_{Q^k} \mathcal{L}^N) \times (h_n^1 \chi_{Q^k} \mathcal{L}^N) + \frac{1}{\|h_n^2\|_1} ((\zeta'_n - \zeta''_n) \mathcal{L}^N) \times (h_n^2 \mathcal{L}^N),$$

where the sum is intended to run only on the indices for which $h_n^1 \chi_{Q^k}$ is not identically zero. Using $\tilde{\pi}_n$ as test plan to compute $W_p(\zeta'_n, \tilde{\rho}'_n)$ we obtain the following upper bound:

$$W_p^p(\zeta'_n, \tilde{\rho}'_n) \leq \int |x - y|^p \, d\tilde{\pi}_n(x, y) \leq C\varepsilon^p + \text{diam}(\text{spt} h_n^2 + \zeta'_n) (\|\zeta'_n\|_1 - \|h_n^1\|_1) \leq C\varepsilon^p,$$

where we used that the mass of h_n^1 remains inside the small cubes with side length ε , and the remaining mass is transported at finite distance in any case (the constant C depends only on N and p). The last inequality holds if we take \bar{n} , and thus n , large enough, and if we adjust the constant C . Adding up the various terms, we conclude that for any $n > \bar{n}$ there is an optimal transport plan π_n for $\tilde{\rho}_n$ and $\tilde{\rho}'_n$ such that

$$W_p(\tilde{\rho}_n, \tilde{\rho}'_n) = \left(\int |x - y|^p d\pi_n(x, y) \right)^{\frac{1}{p}} \leq W_p(\zeta_n, \zeta'_n) + C\varepsilon.$$

To conclude, we observe that the cubes in $\tilde{\mathcal{F}}$ are so large that we can find a non-negative density h_n^3 such that $\rho_n + \tilde{\rho}'_n + h_n^3 \leq 1$ and

$$\int_{\tilde{Q}^k} h_n^3 dx = \int_{\tilde{Q}^k} \rho_n dx \quad \forall k \in \mathbb{N}.$$

Therefore, we consider the plan γ_n associated to ρ_n and $\tilde{\rho}'_n + h_n^3$ defined as

$$\gamma_n = \pi_n + \sum_{k \in \mathbb{N}} \frac{1}{\|\rho_n \chi_{\tilde{Q}^k}\|_1} (\rho_n \chi_{\tilde{Q}^k} \mathcal{L}^N) \times (h_n^3 \chi_{\tilde{Q}^k} \mathcal{L}^N),$$

again summing only on the cubes with non-trivial measure. This gives the following estimate for $W_p^p(\rho_n, \tilde{\rho}'_n + h_n^3)$:

$$\begin{aligned} W_p^p(\rho_n, \tilde{\rho}'_n + h_n^3) &\leq \left(W_p(\zeta_n, \zeta'_n) + C\varepsilon \right)^p + C \|h_n^3\|_1 \\ &\leq \left(\mathcal{W}_p(\rho) + C\varepsilon \right)^p + C \|h_n^3\|_1 \\ &\leq \left(\mathcal{W}_p(\rho_n) - \delta + C\varepsilon \right)^p + C\varepsilon. \end{aligned}$$

Since $\delta > 0$ is fixed and since the constant C in that estimate depends only on N and p , we can find ε small enough so that $W_p^p(\rho_n, \tilde{\rho}'_n + h_n^3) < \mathcal{W}_p^p(\rho_n)$, and this is impossible since $\tilde{\rho}'_n + h_n^3$ is a competitor in the definition of $\mathcal{W}_p(\rho_n)$. \square

The next lemma describes particular symmetries of the problem (1.3) which are crucial in proving properties of maximizers of \mathcal{W}_p in the next section.

Lemma 2.10 (Symmetries of the transport problem). *Let $F: \mathbb{R}^N \rightarrow \mathbb{R}^N$ be an isometry and let $\rho \in \mathbf{A}$ be a given density such that $F_{\#}(\rho \mathcal{L}^N) = \rho \mathcal{L}^N$. Then the following hold:*

- (i) $F_{\#}(\eta_{\rho} \mathcal{L}^N) = \eta_{\rho} \mathcal{L}^N$ and $\tilde{F}_{\#} \pi_{\rho} = \pi_{\rho}$, where \tilde{F} is the map from $\mathbb{R}^N \times \mathbb{R}^N$ into itself defined as $\tilde{F}(x, y) = (F(x), F(y))$.
- (ii) If F is a reflection of the form $F(x) = x - 2\langle x, \nu \rangle \nu$ for some $\nu \in \mathbb{S}^{N-1}$, then we have that

$$\pi_{\rho}(\{(x, y) : \langle x, \nu \rangle \langle y, \nu \rangle < 0\}) = 0. \quad (2.6)$$

In other words, π_{ρ} does not transport mass from one side of the reflection hyperplane $\{x : \langle x, \nu \rangle = 0\}$ to the other.

Proof. We recall that the optimal plan π_ρ is unique (see Theorem 2.3). Also, notice that $F_\#(\rho\mathcal{L}^N)$ and $F_\#(\eta_\rho\mathcal{L}^N)$ are absolutely continuous with respect to the Lebesgue measure, and we have that $F_\#(\rho\mathcal{L}^N) = (\rho \circ F)\mathcal{L}^N$ and $F_\#(\eta_\rho\mathcal{L}^N) = (\eta_\rho \circ F)\mathcal{L}^N$. Therefore, it is trivial to see that $F_\#(\rho\mathcal{L}^N) \in \mathbf{A}$, $F_\#(\eta_\rho\mathcal{L}^N) \in \mathbf{A}$ and $F_\#((\rho + \eta_\rho)\mathcal{L}^N) \leq \mathcal{L}^N$.

It is easy to see that $\tilde{\pi}_\rho = (\tilde{F})_\#\pi_\rho$ is a transport plan associated to $F_\#(\rho\mathcal{L}^N)$ and $F_\#(\eta_\rho\mathcal{L}^N)$: by the properties of the push forward, we have that $(p_1 \circ \tilde{F})_\#\pi_\rho = (p_1)_\#(\tilde{F}_\#\pi_\rho)$, and $p_1 \circ \tilde{F} = F \circ p_1$, therefore $(p_1)_\#\tilde{\pi}_\rho = F_\#(\rho\mathcal{L}^N)$. An analogous property holds for the second projection p_2 , and thus $\tilde{\pi}_\rho$ has the correct marginals. Then, we consider the plan $(\pi_\rho + \tilde{\pi}_\rho)/2$, whose marginals are $\rho\mathcal{L}^N$ and $\frac{1}{2}(\eta_\rho + \eta_\rho \circ F)\mathcal{L}^N$, and we observe that

$$\begin{aligned} \mathcal{W}_p^p(\rho) &\leq \frac{1}{2} \int |x - y|^p d\pi_\rho(x, y) + \frac{1}{2} \int |x - y|^p d\tilde{F}_\#\pi_\rho(x, y) \\ &= \frac{1}{2} \int |x - y|^p d\pi_\rho(x, y) + \frac{1}{2} \int |F(x) - F(y)|^p d\pi_\rho(x, y) = \mathcal{W}_p^p(\rho, \eta_\rho), \end{aligned}$$

where we used that F is an isometry to obtain the last identity. This implies that $\eta_\rho \circ F$ is also an optimal density to compute $\mathcal{W}_p(\rho)$. Since there exists a unique density which realizes $\mathcal{W}_p(\rho)$, then $\eta_\rho\mathcal{L}^N = F_\#(\eta_\rho\mathcal{L}^N)$ and $\tilde{F}_\#\pi_\rho = \pi_\rho$.

In order to prove (ii), suppose that $F(x) = x - 2\langle x, \nu \rangle \nu$ for some $\nu \in \mathbb{S}^{N-1}$. From the previous point we know that π_ρ satisfies $\tilde{F}_\#\pi_\rho = \pi_\rho$. We want to prove that, whenever (2.6) does not hold, we can find a better plan, contradicting the definition of π_ρ . In fact, we consider the plan

$$\tilde{\pi}_\rho = \pi_\rho \llcorner (H_1 \times H_1) + \pi_\rho \llcorner (H_2 \times H_2) + (\text{Id}, F)_\#(\pi_\rho \llcorner (H_1 \times H_2)) + (\text{Id}, F)_\#(\pi_\rho \llcorner (H_2 \times H_1)),$$

where $H_1 = \{x : \langle x, \nu \rangle > 0\}$ and $H_2 = F(H_1) = \{x : \langle x, \nu \rangle < 0\}$. We observe that, since $(p_1)_\#\pi_\rho$ and $(p_2)_\#\pi_\rho$ are absolutely continuous with respect to Lebesgue measure, then π_ρ does not give mass to $\partial(H_i \times H_j)$ for any $i, j \in \{1, 2\}$. Therefore, $\tilde{\pi}_\rho$ is a probability measure, and the well-known properties of the push-forward operation guarantee that $(p_1)_\#\tilde{\pi}_\rho = \rho\mathcal{L}^N$. Since $\pi_\rho = \tilde{F}_\#\pi_\rho$ and $\tilde{F}(H_1 \times H_2) = H_2 \times H_1$, then $\pi_\rho \llcorner (H_1 \times H_2) = \tilde{F}_\#(\pi_\rho \llcorner (H_2 \times H_1))$. With this observation we arrive to

$$\begin{aligned} ((p_2)_\#\tilde{\pi}_\rho) \llcorner H_1 &= (p_2)_\#(\pi_\rho \llcorner (H_1 \times H_1)) + (p_2)_\#((\text{Id}, F)_\#(\pi_\rho \llcorner (H_1 \times H_2))) \\ &= (p_2)_\#(\pi_\rho \llcorner (H_1 \times H_1)) + (p_2)_\# \left((\text{Id}, F)_\# \tilde{F}_\#(\pi_\rho \llcorner (H_2 \times H_1)) \right) \\ &= (p_2)_\#(\pi_\rho \llcorner (H_1 \times H_1)) + (p_2)_\#((F, \text{Id})_\#(\pi_\rho \llcorner (H_2 \times H_1))) \\ &= (p_2)_\#(\pi_\rho \llcorner (H_1 \times H_1)) + (p_2)_\#(\pi_\rho \llcorner (H_2 \times H_1)) \\ &= (p_2)_\#(\pi_\rho \llcorner (\mathbb{R}^N \times H_1)) = ((p_2)_\#\pi_\rho) \llcorner H_1, \end{aligned}$$

where we used that $(\text{Id}, F) \circ \tilde{F} = (F, \text{Id})$ and the fact that F is an isometry to pass from the second to the third line. Arguing in the same way, one can also see that $((p_2)_\#\tilde{\pi}_\rho) \llcorner H_2 = ((p_2)_\#\pi_\rho) \llcorner H_2$. This is sufficient to say that $\rho\mathcal{L}^N + (p_2)_\#(\tilde{\pi}_\rho\mathcal{L}^N) \leq \mathcal{L}^N$, and thus $\tilde{\pi}_\rho \in \mathbf{AP}_\rho$. Now we can compare the costs associated to $\tilde{\pi}_\rho$ and π_ρ . Discarding the common terms, we

get that

$$\int |x - y|^p d(\tilde{\pi}_\rho - \pi_\rho) = \int_{(H_1 \times H_2) \cup (H_2 \times H_1)} (|x - F(y)|^p - |x - y|^p) d\pi_\rho(x, y), \quad (2.7)$$

and a simple geometric argument shows that the function inside the integral is strictly negative. Therefore, if the domain appearing in the right hand side of (2.7) has positive π_ρ measure, then $\tilde{\pi}_\rho$ is a strictly better competitor to compute $\mathcal{W}_p(\rho)$, in contradiction with the definition of π_ρ . To conclude, we observe that we have just proved that $\pi_\rho((H_1 \times H_2) \cup (H_2 \times H_1)) = 0$, and this is equivalent to (2.6). \square

3. MAXIMIZER OF \mathcal{W}_p

3.1. Existence of maximizers. In this section we first prove the existence of maximizers of the energies (1.3) in \mathbf{A} by applying the concentration compactness principle to a maximizing sequence of densities, where we consider them as measures. Even though we consider a maximization problem, our strategy works since \mathcal{W}_p is continuous with respect to the weak* convergence, as shown in Lemma 2.9. Here we state concentration compactness lemma for measures for the convenience of the reader.

Lemma 3.1 (Concentration compactness, [Str08]). *Let $\mu_n \in \mathcal{P}(\mathbb{R}^N)$ be a given sequence of probability measures. Then there exists a subsequence (not relabelled) such that one of the following holds:*

- (i) (Compactness) *There exists a sequence of points $x_n \in \mathbb{R}^N$ such that, for every $\varepsilon > 0$, there exists $L > 0$ large enough such that $\mu_n(Q_L(x_n)) > 1 - \varepsilon$.*
- (ii) (Vanishing) *For every $\varepsilon > 0$ and every $L > 0$ there exists $\bar{n} \in \mathbb{N}$ such that*

$$\mu_n(Q_L(x)) < \varepsilon \quad \forall x \in \mathbb{R}^N, \forall n > \bar{n}.$$

- (iii) (Dichotomy) *There exist $\lambda \in (0, 1)$ and a sequence of points $x_n \in \mathbb{R}^N$ with the following property: for any $\varepsilon > 0$, there exists $L > 0$ such that, for any $L' > L$ there exist two non-negative measures μ_n^1 and μ_n^2 that satisfy, for every n large enough, the following conditions*

$$\begin{aligned} \mu_n^1 + \mu_n^2 &\leq \mu_n, \\ \text{spt} \mu_n^1 &\subset Q_L(x_n), \quad \text{spt} \mu_n^2 \subset \mathbb{R}^N \setminus Q_{L'}(x_n), \\ |\mu_n^1(\mathbb{R}^N) - \lambda| + |\mu_n^2(\mathbb{R}^N) - (1 - \lambda)| &< \varepsilon. \end{aligned}$$

Theorem 3.2. *Let $p > 1$ be fixed. Then there exists a maximizer of \mathcal{W}_p in \mathbf{A} .*

Proof. Let us consider a maximizing sequence $\rho_n \in \mathbf{A}$ with $\mathcal{W}_p(\rho_n) \rightarrow \sup_{\rho \in \mathbf{A}} \mathcal{W}_p(\rho)$. Notice that, thanks to Corollary 2.7, we have that $\sup_{\rho \in \mathbf{A}} \mathcal{W}_p(\rho) \leq C < +\infty$ for some constant $C = C(p, N)$. We are going to apply the concentration compactness lemma to $\mu_n = \rho_n \mathcal{L}^N$, and show that the vanishing and dichotomy phenomena do not happen. Then exploiting the

invariance of the energy under translations and Lemma 2.9 we establish the existence of a maximizer.

We first exclude the vanishing case. Up to translations, we can suppose that the points x_n appearing in Lemma 3.1 all coincide with the origin. Suppose by contradiction that, for any $\varepsilon > 0$ and any $L > 0$ we can find $\bar{n} \in \mathbb{N}$ such that $\mu_n(Q_L(x)) < (\varepsilon/3)^N$ for every $x \in \mathbb{R}^N$. Then, we fix a partition $\mathcal{F} = \{Q^k\}_{k \in \mathbb{N}}$ of \mathbb{R}^N made of cubes with side length ε . Since by hypothesis $\mu_n(Q^k) < |Q^k|/3$ for every $n > \bar{n}$ and every $k \in \mathbb{N}$, then for every $n > \bar{n}$ there exists $\rho'_n \in \mathbf{A}$ such that $\rho_n + \rho'_n \leq 1$ and

$$\int_{Q^k} \rho_n dx = \int_{Q^k} \rho'_n dx \quad \forall k \in \mathbb{N}.$$

Using a transport plan similar to π_n defined in (2.4), it is immediate to see that

$$\mathcal{W}_p^p(\rho_n) \leq \mathcal{W}_p^p(\rho_n, \rho'_n) \leq \text{diam}(Q^k)^p = C_{N,p} \varepsilon^p.$$

If we take ε sufficiently small, we clearly have that ρ_n is not a maximizing sequence for \mathcal{W}_p , arriving to a contradiction.

Now we treat the dichotomy case. Suppose for a contradiction that there exists $\lambda \in (0, 1)$ such that, for any $\varepsilon > 0$ there exist $\bar{n} \in \mathbb{N}$, $L > 0$ and two sequences of non-negative densities ρ_n^1, ρ_n^2 that satisfy

$$\begin{aligned} \rho_n^1 + \rho_n^2 &\leq \rho_n \\ \text{spt} \rho_n^1 &\subset Q_L \quad \text{spt} \rho_n^2 \subset \mathbb{R}^N \setminus Q_{L+3C_N}, \\ \left| \int \rho_n^1 dx - \lambda \right| + \left| \int \rho_n^2 dx - (1 - \lambda) \right| &< \varepsilon, \end{aligned} \tag{3.1}$$

where C_N is the constant appearing in (2.3).

Since the distance between $\text{spt} \rho_n^1$ and $\text{spt} \rho_n^2$ is larger than $3C_N$, then applying Corollary 2.7 we obtain that $\mathcal{W}_p^p(\rho_n^1 + \rho_n^2) = \mathcal{W}_p^p(\rho_n^1) + \mathcal{W}_p^p(\rho_n^2)$. Combining the first and the third conditions in (3.1), we get that $\|\rho_n - \rho_n^1 - \rho_n^2\|_1 < \varepsilon$, and we define $m_n^1 = \|\rho_n^1\|_1$ and $m_n^2 = \|\rho_n^2\|_1$. Using this fact, and that $\rho_n^1 + \rho_n^2 + \eta_{\rho_n^1 + \rho_n^2} \leq 1$, we deduce that

$$\int (\eta_{\rho_n^1 + \rho_n^2} - (1 - \rho_n))_+ dx \leq \varepsilon. \tag{3.2}$$

We denote by T_n the optimal transport map to compute $\mathcal{W}_p(\rho_n^1 + \rho_n^2)$, and we define

$$\zeta_n = \min\{\eta_{\rho_n^1 + \rho_n^2}, 1 - \rho_n\}, \quad \tilde{\rho}_n = (T_n^{-1})\# \zeta_n,$$

so that $\tilde{\rho}_n$ is an approximation of $\rho_n^1 + \rho_n^2$, and it is smaller than that sum. We let $\mathcal{F} = \{Q^k\}_{k \in \mathbb{N}}$ be a partition of \mathbb{R}^N made of cubes with side length equal to 3, and we can find, as we did before, a density ζ'_n such that $\rho_n + \zeta_n + \zeta'_n \leq 1$ and

$$\int_{Q^k} \zeta'_n dx = \int_{Q^k} \rho_n - \tilde{\rho}_n dx \quad \forall k \in \mathbb{N}.$$

Therefore, we estimate the energy of ρ_n with the plan

$$\tilde{\pi}_n = (\text{Id}, T_n)_\# \tilde{\rho}_n + \sum_{k \in \mathbb{N}} \frac{1}{\|\zeta'_n \chi_{Q^k}\|_1} ((\rho_n - \tilde{\rho}_n) \chi_{Q^k} \mathcal{L}^N) \times (\zeta'_n \chi_{Q^k} \mathcal{L}^N).$$

In fact, combining (3.2) and the fact that $\|\rho_n - \rho_n^1 - \rho_n^2\|_1 \leq \varepsilon$, we have that $\|\rho_n - \tilde{\rho}_n\|_1 \leq 2\varepsilon$, and thus

$$\begin{aligned} \mathcal{W}_p^p(\rho_n) &\leq \int |x - y|^p d\tilde{\pi}_n \leq \int |T_n(x) - x|^p \tilde{\rho}_n(x) dx + 2(\text{diam } Q^k)^p \varepsilon \\ &\leq \mathcal{W}_p^p(\rho_n^1 + \rho_n^2) + C_{N,p} \varepsilon \\ &= \mathcal{W}_p^p(\rho_n^1) + \mathcal{W}_p^p(\rho_n^2) + C_{N,p} \varepsilon \\ &\leq \sup \{ \mathcal{W}_p^p(\rho) : \rho \in \mathbf{A}_{m_n^1} \} + \sup \{ \mathcal{W}_p^p(\rho) : \rho \in \mathbf{A}_{m_n^2} \} + C_{N,p} \varepsilon. \end{aligned} \tag{3.3}$$

Using the rescaling exploited in Remark 2.8 we see that

$$\sup \{ \mathcal{W}_p^p(\rho) : \rho \in \mathbf{A}_m \} = m^{1+\frac{p}{N}} \sup \{ \mathcal{W}_p^p(\rho) : \rho \in \mathbf{A} \};$$

hence, (3.3) implies that

$$\mathcal{W}_p^p(\rho_n) \leq C_{N,p} \varepsilon + \left((m_n^1)^{1+\frac{p}{N}} + (m_n^2)^{1+\frac{p}{N}} \right) \sup \{ \mathcal{W}_p^p(\rho) : \rho \in \mathbf{A} \}.$$

If ε is small enough, this is incompatible with the fact that $\lim_n \mathcal{W}_p(\rho_n) = \sup_{\rho \in \mathbf{A}} \mathcal{W}_p(\rho)$. In fact, the function $t \mapsto t^{1+\frac{p}{N}}$ is strictly convex, and if $\varepsilon < \frac{1}{2} \min\{\lambda, 1 - \lambda\}$, then m_n^1 and m_n^2 are far away from 0. \square

3.2. The only maximizer is the ball. In the second part of this section we will characterize the maximizers of \mathcal{W}_p over \mathbf{A} . In fact, we prove that the only maximizer of \mathcal{W}_p is the characteristic function of a ball (with the correct volume). The intuition behind this result is that, if we have a set, and we create some holes in it (adding some mass somewhere else), we are lowering the energy since the additional mass can be transported at shorter distance. We obtain the main result in several steps: First we study the 1-dimensional case, possibly with a weight, where the structure of the transport plan is known explicitly. Then, using a symmetrization argument we show that the optimal plan associated to a maximizer has some geometric properties, and, in fact, it is radial. Next, using the 1-dimensional case, we prove that a maximizer has to be a star-shaped set, and via an optimality argument we deduce that a star-shaped maximizer must actually be a ball.

Proposition 3.3. *Let $m > 0$ be a given parameter. Let $w : (0, +\infty) \rightarrow (0, +\infty)$ be a non-decreasing weight and let $I = [0, \ell]$ be the unique segment such that $\int_I w dx = m$. For any density $\rho : \mathbb{R}^+ \rightarrow [0, 1]$ with $\int_{\mathbb{R}^+} \rho w dx = m$, we have that*

$$\mathcal{W}_p(\chi_I) \geq \mathcal{W}_p(\rho), \tag{3.4}$$

where \mathcal{W}_p is defined in the metric-measure setting with base space \mathbb{R}^+ endowed with the usual distance and reference measure equal to $w \mathcal{L}^1$. Moreover, the equality holds if and only if $\rho = \chi_I$ almost everywhere.

Proof. We note that, also in this weighted case, the transport distance is bounded (using again Lemma 2.6), and thus for any density the infimum in the definition of \mathcal{W}_p is achieved thanks to Theorem 2.3 and Theorem 2.5. Therefore, there exists η_ρ such that $\mathcal{W}_p(\rho) = W_p(\rho\gamma, \eta_\rho\gamma)$, where we use the notation $\gamma = w\mathcal{L}^1$. Moreover, since we have an increasing cost, we actually have that $\mathcal{W}_p(\chi_I) = W_p(\chi_I\gamma, \chi_{I'}\gamma)$, where $I' = [\ell, \ell']$ for some $\ell' > \ell$, and the transport plan is induced by a monotone map T (see Theorem 2.4).

Now we introduce an auxiliary problem that produces a non-optimal candidate to estimate $\mathcal{W}_p(\rho)$. The advantage of this modified problem is that it enforces a “geometric” constraint that clarifies some arguments. The auxiliary functional, which considers only plans which move mass “forward”, is given by

$$\mathcal{AW}_p^p(\rho) := \inf \left\{ \int |x - y|^p d\pi(x, y) : \pi \in \mathbf{AP}_\rho, \pi(\{(x, y) : y < x\}) = 0 \right\}.$$

We observe that the infimum is actually a minimum since the additional constraint is closed under weak* convergence. Moreover, applying the standard results for the one dimensional transport problem, we know that the optimal plan is induced by a non-decreasing map. Since we have already observed that $\mathcal{W}_p(\chi_I) = W_p(\chi_I\gamma, \chi_{I'}\gamma)$, then the monotonicity of the optimal map ensures that $\mathcal{W}_p(\chi_I) = \mathcal{AW}_p(\chi_I)$. For a general density ρ , instead, we have just the inequality $\mathcal{AW}_p(\rho) \geq \mathcal{W}_p(\rho)$ due to the introduction of the additional constraint. With these observations, we reduce to proving the following (stronger) inequality:

$$\mathcal{W}_p(\chi_I) \geq \mathcal{AW}_p(\rho),$$

and (3.4) simply follows.

From now on we denote by \tilde{T}_ρ the transport map appearing when we compute $\mathcal{AW}_p(\rho)$. We define the following “volume” functions with domain \mathbb{R}^+ :

$$V(x) := \int_0^x w(t) dt, \quad V_\rho(x) := \int_0^x \rho(t)w(t) dt.$$

We also denote by $d(v)$ (resp. $d_\rho(v)$) the transport distance of the point $V^{-1}(v)$ (resp. $V_\rho^{-1}(v)$) when we compute $\mathcal{W}_p(\chi_I)$ (resp. $\mathcal{W}_p(\rho)$), i.e.

$$d(v) := |T(V^{-1}(v)) - V^{-1}(v)|, \quad d_\rho(v) := |\tilde{T}_\rho(V_\rho^{-1}(v)) - V_\rho^{-1}(v)|.$$

Using the explicit expression of the optimal transport map in 1D (see for example [Vil03, Remarks 2.19 (iv)]), we have that

$$\gamma([V^{-1}(v), V^{-1}(v) + d(v)]) = m \quad \forall v \in [0, m].$$

One can easily adapt the proof of Lemma 2.6 to the auxiliary functional and see that, if x is a Lebesgue point for \tilde{T}_ρ and $r = |\tilde{T}_\rho(x) - x|$, then $(\tilde{T}_\rho)_\#(\rho\gamma) = (1 - \rho)\gamma$ in $[x, x + r]$. Moreover, since \tilde{T}_ρ is non-decreasing, we also have that

$$(\tilde{T}_\rho)_\#(\rho\gamma \llcorner [0, x]) = (1 - \rho)\gamma \quad \text{in } [x, x + r]. \quad (3.5)$$

We claim that $d_\rho \leq d$. In fact, suppose for a contradiction that there exists $v \in (0, m)$ such that $d_\rho(v) > d(v)$. Since $\rho \leq 1$, then $V^{-1} \leq V_\rho^{-1}$, and thus

$$\begin{aligned} \int_0^{V_\rho^{-1}(v)+d_\rho(v)} \rho(t)w(t) dt &\geq \int_{V_\rho^{-1}(v)}^{V_\rho^{-1}(v)+d_\rho(v)} w(t) dt \geq \int_{V^{-1}(v)}^{V^{-1}(v)+d_\rho(v)} w(t) dt \\ &> \int_{V^{-1}(v)}^{V^{-1}(v)+d(v)} w(t) dt = \gamma([V^{-1}(v), V^{-1}(v) + d(v)]) = m, \end{aligned}$$

where we used (3.5) applied to $x = V_\rho^{-1}(v)$ and $r = d_\rho(v)$ to get the first inequality, and the monotonicity of w to obtain the second one. This chain of inequalities of course leads to a contradiction since $m = \int \rho d\gamma$. Therefore $d_\rho \leq d$.

Since w and ρw are locally bounded in $[0, +\infty)$, then both V and V_ρ are locally Lipschitz, and we can apply the fundamental theorem of calculus: using $v = V_\rho(x)$ as variable in the computation of $\mathcal{AW}_p(\rho)$ we obtain that

$$\mathcal{AW}_p^p(\rho) = \int_{\mathbb{R}^+} |\tilde{T}_\rho(x) - x|^p \rho(x)w(x) dx = \int_0^m d_\rho(v)^p dv \leq \int_0^m d(v)^p dv = \mathcal{W}_p^p(\chi_I),$$

where the inequality follows from comparison between d and d_ρ , and this is the desired inequality. Finally, one can notice that the only way to obtain an equality in the previous chain of inequalities is that $\rho = \chi_{I''}$ for some segment I'' and w is constant in $\text{spt}(\rho + T_\# \rho)$. However, if $I'' \neq I$, then one can construct a better transport plan moving some mass to the left (this plan should belong to AP_ρ , but it is not admissible for the auxiliary problem). Therefore, the equality in (3.4) holds only for $\rho = \chi_I$. \square

Lemma 3.4. *Let $p > 1$ be given, and let $\rho \in \mathbf{A}$ be a maximizer of \mathcal{W}_p . If $\nu \in \mathbb{S}^{N-1}$ is such that*

$$\int_{\{x: \langle x, \nu \rangle > 0\}} \rho dx = \int_{\{x: \langle x, \nu \rangle < 0\}} \rho dx = \frac{1}{2}, \quad (3.6)$$

then the optimal plan π_ρ satisfies

$$\pi_\rho(\{(x, y) : \langle x, \nu \rangle \cdot \langle y, \nu \rangle < 0\}) = 0. \quad (3.7)$$

Proof. The idea is to consider an auxiliary functional, as in the proof of Proposition 3.3, and show that it coincides with \mathcal{W}_p when evaluated at ρ (due to the maximality of this density). This ensures that π_ρ has some additional structure due to the uniqueness of the optimal plan. We define the auxiliary functional

$$\mathcal{AW}_p^p(\rho, \nu) := \inf \left\{ \int |x - y|^p d\pi(x, y) : \pi \in \text{AP}_\rho, \pi(\{(x, y) : \langle x, \nu \rangle \cdot \langle y, \nu \rangle < 0\}) = 0 \right\}.$$

Loosely speaking, this auxiliary functional uses only plans that do not transport mass across the hyperplane $\{x : \langle x, \nu \rangle = 0\}$. As before, we are introducing an additional constraint that is closed under weak* convergence, and thus there exists an optimal plan in the definition of $\mathcal{AW}_p(\rho, \nu)$. Clearly, since we are introducing a constraint in the minimization process, we have that $\mathcal{W}_p(\rho) \leq \mathcal{AW}_p(\rho, \nu)$.

Let $F(x) = x - 2\langle x, \nu \rangle \nu$ be the reflection map, and define the two symmetrizations of ρ with respect to ν :

$$\rho_1 = \rho \llcorner H_1 + F_{\#}(\rho \llcorner H_1), \quad \rho_2 = \rho \llcorner H_2 + F_{\#}(\rho \llcorner H_2),$$

where $H_1 = \{x: \langle x, \nu \rangle > 0\}$ and $H_2 = F(H_1) = \{x: \langle x, \nu \rangle < 0\}$. We denote by $\bar{\pi}_1$ and $\bar{\pi}_2$ the two optimal plans realizing $\mathcal{AW}_p(\rho_1, \nu)$ and $\mathcal{AW}_p(\rho_2, \nu)$, respectively. We claim that

$$\bar{\pi} = \bar{\pi}_1 \llcorner (H_1 \times H_1) + \bar{\pi}_2 \llcorner (H_2 \times H_2)$$

realizes $\mathcal{AW}_p(\rho, \nu)$. In fact, $\bar{\pi}$ is admissible to compute $\mathcal{AW}_p(\rho, \nu)$, and if we find a better candidate π to compute $\mathcal{AW}_p(\rho, \nu)$, then we can also construct the following plans that are good candidates to compute $\mathcal{AW}_p(\rho_1, \nu)$ and $\mathcal{AW}_p(\rho_2, \nu)$ respectively:

$$\pi_1 = \pi \llcorner (H_1 \times H_1) + \tilde{F}_{\#}(\pi \llcorner (H_1 \times H_1)), \quad \pi_2 = \pi \llcorner (H_2 \times H_2) + \tilde{F}_{\#}(\pi \llcorner (H_2 \times H_2)),$$

where $\tilde{F}(x, y) = (F(x), F(y))$. Then we observe that

$$\begin{aligned} \mathcal{AW}_p^p(\rho_1, \nu) &= \int |x - y|^p d\bar{\pi}_1 = 2 \int_{H_1 \times H_1} |x - y|^p d\bar{\pi}_1, \\ \mathcal{AW}_p^p(\rho_2, \nu) &= \int |x - y|^p d\bar{\pi}_2 = 2 \int_{H_2 \times H_2} |x - y|^p d\bar{\pi}_2, \\ \int |x - y|^p d\bar{\pi} &= \frac{1}{2} (\mathcal{AW}_p^p(\rho_1, \nu) + \mathcal{AW}_p^p(\rho_2, \nu)), \\ \int |x - y|^p d\pi &= \frac{1}{2} \left(\int |x - y|^p d\pi_1 + \int |x - y|^p d\pi_2 \right). \end{aligned}$$

If $\mathcal{AW}_p^p(\rho, \nu) < \int |x - y|^p d\bar{\pi}$, then at least one between π_1 and π_2 is a better competitor for $\mathcal{AW}_p(\rho_1, \nu)$ or $\mathcal{AW}_p(\rho_2, \nu)$, contradicting the definition of $\bar{\pi}_1$ and $\bar{\pi}_2$. Therefore, the following conditions hold:

$$\mathcal{W}_p^p(\rho) \leq \mathcal{AW}_p^p(\rho, \nu) = \frac{1}{2} (\mathcal{AW}_p^p(\rho_1, \nu) + \mathcal{AW}_p^p(\rho_2, \nu)) = \frac{1}{2} (\mathcal{W}_p^p(\rho_1) + \mathcal{W}_p^p(\rho_2)), \quad (3.8)$$

where we used the second part of Lemma 2.10 to obtain the last equality. Since ρ is a maximizer, then (3.8) guarantees that ρ_1 and ρ_2 are also maximizers. This, however, implies that $\mathcal{W}_p(\rho) = \mathcal{AW}_p(\rho, \nu)$. In other words, $\bar{\pi}$ realizes $\mathcal{W}_p(\rho)$ and satisfies (3.7). Therefore, necessarily, we have that $\pi_\rho = \bar{\pi}$, concluding the proof. \square

Corollary 3.5. *Let $p > 1$ be given, and let $\rho \in \mathbf{A}$ be a maximizer of \mathcal{W}_p . Then there exists $x_0 \in \mathbb{R}^N$ such that π_ρ has the following property:*

$$\pi_\rho(\{(x, y): \langle y - x_0, x - x_0 \rangle \neq |y - x_0||x - x_0|\}) = 0. \quad (3.9)$$

That is, π_ρ is radial with center x_0 .

Proof. By sliding each hyperplane $\{x: \langle x, e_i \rangle = 0\}$ until it splits the mass of ρ in half, and by taking the intersection of the N hyperplanes, we find a point $x_0 \in \mathbb{R}^N$ such that

$$\int_{\{x: \langle x - x_0, e_i \rangle > 0\}} \rho dx = \int_{\{x: \langle x - x_0, e_i \rangle < 0\}} \rho dx = \frac{1}{2} \quad \forall i \in \{1, \dots, N\}.$$

Up to translations, we suppose that $x_0 = 0$. By (3.8) we know that suitable symmetrizations of ρ with respect to the coordinate axes are again maximizers. Iterating this procedure, we obtain a maximizer $\tilde{\rho}$ taking successive reflections of the sector

$$\rho \llcorner \{x: \langle x, e_i \rangle > 0 \ \forall i = 1, \dots, N\}, \quad (3.10)$$

and the result is a density symmetric with respect to each coordinate direction. The symmetries of $\tilde{\rho}$ guarantee that

$$\int_{\{x: \langle x, \nu \rangle > 0\}} \tilde{\rho} dx = \int_{\{x: \langle x, \nu \rangle < 0\}} \tilde{\rho} dx = \frac{1}{2} \quad \forall \nu \in \mathbb{S}^{N-1}.$$

Hence, applying Lemma 3.4 to $\tilde{\rho}$ we obtain that $\pi_{\tilde{\rho}}$ satisfies the splitting condition (3.7) for any vector ν . Thus, the condition (3.9) holds for $\pi_{\tilde{\rho}}$. We finally conclude by uniqueness of the optimal plan, as we did in the last part of Lemma 3.4: we can use the same strategy starting from a different sector in (3.10), defining a different symmetric density $\tilde{\rho}$. The same conclusion holds for the new optimal plan associated to that density, namely $\pi_{\tilde{\rho}}$. By uniqueness of the optimal plan, we know that π_{ρ} can be obtained gluing together the plans of each sector, and thus also π_{ρ} satisfies (3.9). \square

Now we can state and prove our main result.

Theorem 3.6. *Let $p > 1$ be given. Then the only maximizer of \mathcal{W}_p in the class \mathbf{A} , up to translations, is the characteristic function of B with $|B| = 1$.*

Proof. We prove this result in two steps: in the first one we prove that any maximizer must be the characteristic function of a star-shaped set, while in the second one we exploit the inner-ball condition exposed in Lemma 2.6 to see that the length of the rays must be constant. Without loss of generality, we can suppose $N \geq 2$ since the 1-dimensional case has already been treated in Proposition 3.3.

Step 1. First we will apply Corollary 3.5 and decompose the transport along rays. Then, exploit the one dimensional result obtained in Proposition 3.3 to prove that, along the rays, we see only segments, and this is equivalent to saying that the maximizer is a star-shaped set.

Let ρ be any maximizer of \mathcal{W}_p in \mathbf{A} . We apply Corollary 3.5 to ρ , and we can suppose, without loss of generality, that the point x_0 coincides with the origin. Therefore, the optimal plan π_{ρ} is induced by a radial map T_{ρ} . Since in this proof we do not need to stress the dependence of η_{ρ} , π_{ρ} and T_{ρ} on the density ρ , we simplify the notation, and we denote those objects by η , π and T , respectively. We decompose every function in radial coordinates, and let $w(r) = r^{N-1}$ denote the coarea factor when we integrate in polar coordinates. For any $\omega \in \mathbb{S}^{N-1}$ we define the functions

$$\rho^{\omega}(r) = \rho(r\omega), \quad \eta^{\omega}(r) = \eta(r\omega), \quad T^{\omega}(r) = |T(r\omega)|$$

for every $r \in [0, +\infty)$. We consider them as functions defined (almost everywhere) on the metric-measure space (X, d, γ) , where $X = \mathbb{R}^+$, $\gamma = w\mathcal{L}^1$ and d is the usual distance.

We claim that, since $T(r\omega) = T^\omega(r)\omega$ and $T_\# \rho = \eta$, we have

$$(T^\omega)_\#(\rho^\omega \gamma) = \eta^\omega \gamma \quad \text{for a.e. } \omega \in \mathbb{S}^{N-1}. \quad (3.11)$$

For any $s > 0$ and any $E \subset \mathbb{S}^{N-1}$ we define the set $F = \{r\omega : 0 \leq r \leq s, \omega \in E\}$ and we have that

$$\begin{aligned} \int_E d\mathcal{H}_\omega^{N-1} \int_0^s \eta^\omega d\gamma &= \int_E d\mathcal{H}_\omega^{N-1} \int_0^s \eta(r\omega) r^{N-1} dr = \int_F \eta(x) dx \\ &= \int_F (T_\# \rho)(x) dx = \int_{T^{-1}(F)} \rho(x) dx \\ &= \int_E d\mathcal{H}_\omega^{N-1} \int_{(T^\omega)^{-1}([0,s])} \rho(r\omega) r^{N-1} dr \\ &= \int_E d\mathcal{H}_\omega^{N-1} \int_{(T^\omega)^{-1}([0,s])} \rho^\omega(r) d\gamma \\ &= \int_E d\mathcal{H}_\omega^{N-1} \int_0^s (T^\omega)_\# \rho^\omega d\gamma. \end{aligned}$$

Here we used that T is radial to pass from the second to the third line, in combination with the integration in polar coordinates. Since E and s are arbitrary, this proves (3.11).

We obtain the result of this first step by applying Proposition 3.3 separately for any $\omega \in \mathbb{S}^{N-1}$. In fact, we can integrate in polar coordinates the transport cost and obtain that

$$\begin{aligned} \int |T(x) - x|^p \rho(x) dx &= \int_{\mathbb{S}^{N-1}} \int_0^{+\infty} |T(r\omega) - r\omega|^p \rho(r\omega) r^{N-1} dr d\omega \\ &= \int_{\mathbb{S}^{N-1}} \left(\int_0^{+\infty} |T^\omega(r) - r|^p \rho^\omega(r) \omega(r) dr \right) d\omega. \end{aligned}$$

The inner integral in the last expression coincides with the transport cost of T^ω between $\rho^\omega \gamma$ and $\eta^\omega \gamma$, and since T is the optimal transport map between ρ and η , then also T^ω must be optimal between $\rho^\omega \gamma$ and $\eta^\omega \gamma$ for every $\omega \in \mathbb{S}^{N-1}$. This is properly justified by showing that gluing the optimizers ω -by- ω we obtain a measurable density. We sketch the proof of this fact in Appendix A. Therefore, if we denote by $m(\omega) = \int_{\mathbb{R}^+} \rho^\omega d\gamma$, then

$$\begin{aligned} \mathcal{W}_p^p(\rho) &= \int_{\mathbb{S}^{N-1}} \mathcal{W}_p^p(\rho^\omega) d\omega \\ &\leq \int_{\mathbb{S}^{N-1}} \sup \left\{ \mathcal{W}_p^p(\theta) : \theta : X \rightarrow [0, 1], \int_X \theta d\gamma = m(\omega) \right\} d\omega, \quad (3.12) \end{aligned}$$

where we use the metric-measure definition of \mathcal{W}_p in those integrals (see Remark 1.1). By Proposition 3.3, for every $\omega \in \mathbb{S}^{N-1}$, the supremum inside the last integral coincides with $\mathcal{W}_p^p(\chi_{I^\omega})$, where $I^\omega \subset X$ is the unique segment of the form $[0, \ell^\omega]$ with $\gamma(I^\omega) = m(\omega)$. Moreover, the inequality is strict whenever ρ^ω is not equivalent to χ_{I^ω} . Since the map $\omega \mapsto m(\omega)$ is measurable, we can glue the segments I^ω together and obtain another candidate to compute \mathcal{W}_p . The density ρ is a maximizer; hence, for almost every $\omega \in \mathbb{S}^{N-1}$ the density ρ^ω must be equivalent to χ_{I^ω} , concluding the proof of the first step.

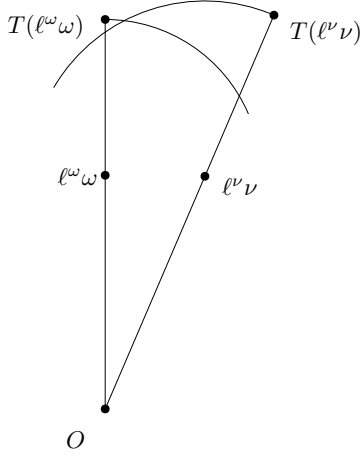


FIGURE 1. In this figure we depict two points $\ell^\omega \omega$ and $\ell^\nu \nu$ that belong to the support of ρ , and their images through the map T , which coincide with $T^\omega(\ell^\omega)\omega$ and $T^\nu(\ell^\nu)\nu$ respectively. The inner ball condition implies that the two image points have to lie outside of the circles depicted, whose radii coincide with the transport distances $T^\omega(\ell^\omega) - \ell^\omega$ and $T^\nu(\ell^\nu) - \ell^\nu$ respectively.

Step 2. For any $\omega \in \mathbb{S}^{N-1}$ we know that $T(\ell^\omega \omega) = (T^\omega(\ell^\omega))\omega$, and Lemma 2.6 guarantees that $\eta(x) = 1 - \rho(x)$ for every $x \in \mathbb{R}^N$ such that $|x - \ell^\omega \omega| \leq T^\omega(\ell^\omega) - \ell^\omega$. Let $\nu \in \mathbb{S}^{N-1}$ be another unit vector. Note that $T^\omega(\ell^\omega) = 2^{1/N} \ell^\omega$ (see e.g. [CTG22] where the transport map in the case of a ball is given explicitly). Thanks to the inner ball condition, we obtain that $T^\nu(\ell^\nu)$ is larger than t for any $t > 0$ such that $|t\nu - \ell^\omega \omega| \leq T^\omega(\ell^\omega) - \ell^\omega$.

In order to simplify the notation we define $c = 2^{1/N}$, $r = \ell^\omega$ and $s = \ell^\nu$. Taking the square of both sides of the inner-ball inequality (see Figure 1 for a geometric intuition of the inner ball condition in this situation), we get that $s \geq t$ for every $t > 0$ satisfying

$$c^2 t^2 - 2c \langle \nu, \omega \rangle r t + c(2 - c)r^2 = 0.$$

Solving the above equation in t one gets that

$$s \geq \frac{\langle \nu, \omega \rangle + \sqrt{\langle \nu, \omega \rangle^2 - c(2 - c)}}{c} r.$$

By the definition of c , the expression under the square root is non-negative whenever $\langle \nu, \omega \rangle$ is close enough to 1 since $c > 1$ and $1 - c(2 - c) = (c - 1)^2 > 0$. Swapping the roles of ν and ω we also arrive to the analogous inequality

$$r \geq \frac{\langle \nu, \omega \rangle + \sqrt{\langle \nu, \omega \rangle^2 - c(2 - c)}}{c} s.$$

Combining these two inequalities we can control the difference between s and r in terms of the distance between ν and ω :

$$\begin{aligned} s - r &\geq \frac{r}{c} \left(\langle \nu, \omega \rangle - c + \sqrt{\langle \nu, \omega \rangle^2 - c(2-c)} \right) = \frac{2r(1 - \langle \nu, \omega \rangle)}{\langle \nu, \omega \rangle - c - \sqrt{\langle \nu, \omega \rangle^2 - c(2-c)}}, \\ s - r &\leq \frac{s}{c} \left(c - \langle \nu, \omega \rangle - \sqrt{\langle \nu, \omega \rangle^2 - c(2-c)} \right) = \frac{2s(1 - \langle \nu, \omega \rangle)}{c - \langle \nu, \omega \rangle + \sqrt{\langle \nu, \omega \rangle^2 - c(2-c)}}. \end{aligned}$$

By Corollary 2.7 we have that $|T(x) - x| \leq C_N$ for a dimensional constant C_N ; hence, r and s are also uniformly bounded. Since $2(1 - \langle \nu, \omega \rangle) = |\nu - \omega|^2$, we can combine the previous estimates and obtain that

$$|\ell^\nu - \ell^\omega| \leq \tilde{C}_N |\nu - \omega|^2$$

for any ν and ω sufficiently close. This implies that the map $\omega \mapsto \ell^\omega$ is 2-Hölder continuous on the sphere, hence it is constant. This is equivalent to showing that the only maximizer is the ball, and thus the proof is concluded. \square

4. QUANTITATIVE INEQUALITY IN ONE DIMENSION

In this section we prove a quantitative inequality for \mathcal{W}_p in one dimension, so we manage to strengthen the result obtained in Section 3 adding a term that measures the displacement of a density ρ respect to the characteristic function of a ball. In order to measure that distance, we consider a version of the Frankel asymmetry that, loosely speaking, is the L^1 distance between a density and a ball. This choice is by no means new: for example, the asymmetry was used in the quantitative isoperimetric inequality (cfr. [FMP08, FMP10]) and in the quantitative Brunn-Minkowski inequality (cfr. [BJ17]). See also [FP20, FL21] for a quantitative inequality involving a functional of Riesz type.

Definition 4.1. We define the following quantity, that we will just call asymmetry in the sequel:

$$A(\rho) := \inf \left\{ \left\| \rho - \chi_{B_r(x)} \right\|_1 : x \in \mathbb{R}^N, |B_r(x)| = 1 \right\} \quad \forall \rho \in \mathbf{A}.$$

With this notion, our quantitative inequality reads as the following.

Theorem 4.2. *For $N = 1$ and $p > 1$ fixed, there exists a constant $C_p > 0$ such that*

$$\mathcal{W}_p^p(B) - \mathcal{W}_p^p(\rho) \geq C_p A(\rho)^2 \quad \text{for all } \rho \in \mathbf{A}.$$

Remark 4.3. We point out that the exponent 2 in our quantitative inequality is sharp, in the sense that the inequality would be false with a smaller exponent for densities with small asymmetry. This can be seen by taking $\rho = \chi_{[-1/2-\varepsilon, -1/2]} + \chi_{[-1/2+\varepsilon, 1/2-\varepsilon]} + \chi_{[1/2, 1/2+\varepsilon]}$ for ε small.

Proof. By definition of asymmetry, $A(\rho) \leq 2$ for every $\rho \in \mathbf{A}$, and without loss of generality we can suppose that $A(\rho) > 0$. Up to translations, we can suppose that

$$\int_{-\infty}^0 \rho dx = \int_0^{+\infty} \rho dx = \frac{1}{2}.$$

As we showed in Proposition 3.3, it is possible to find a transport plan $\bar{\pi} \in \mathbf{AP}_\rho$ such that $|x - y| \leq 1/2$ for any $(x, y) \in \text{spt}\bar{\pi}$. Loosely speaking, that transport plan moves mass “away from the origin”. Now we want to get a quantitative inequality modifying $\bar{\pi}$ and finding another plan $\pi \in \mathbf{AP}_\rho$ for which the transport distance is again controlled by $1/2$, and moreover

$$\pi(\{(x, y) : |x - y| \leq d\}) \geq \frac{A(\rho)}{100}, \quad \text{where } d := \frac{1}{2} - \frac{A(\rho)}{100}. \quad (4.1)$$

With this competitor, if $E = \{(x, y) \in \mathbb{R}^N \times \mathbb{R}^N : |x - y| \leq d\}$ is the set considered in the previous inequality, we have that

$$\begin{aligned} \mathcal{W}_p^p(\rho) &\leq \int |x - y|^p d\pi(x, y) \leq d^p \pi(E) + \frac{1}{2^p} (1 - \pi(E)) \\ &= \frac{1}{2^p} + \frac{\pi(E)}{2^p} \left[\left(1 - \frac{A(\rho)}{50}\right)^p - 1 \right] \leq \frac{1}{2^p} + \frac{\pi(E)}{2^p} (-C_p A(\rho)) \\ &= \mathcal{W}_p^p(B) - C_p A(\rho)^2, \end{aligned}$$

where C_p is a constant depending only on p . Therefore, we need to find such a plan π to complete the proof. We denote by \bar{T} the map that induces $\bar{\pi}$. Let us look at the set $\{x \geq 0\}$, and we define x_R as the smallest point that is moved at distance d , i.e. $x_R := \inf\{x > 0 : \bar{T}_\#(\rho \llcorner [0, x_R]) = 1 - \rho \text{ in } [x_R, x_R + d]\}$. Exploiting the same arguments of Proposition 3.3, it can be shown that $\bar{T}_\#(\rho \llcorner [0, x_R]) = 1 - \rho$ in $[x_R, x_R + d]$, and thus $\int_0^{x_R} \rho dx + \int_{x_R}^{x_R+d} \rho dx \geq d$. Now we explore the different cases that may appear.

Case 1. If we have that $\int_0^{x_R} \rho dx \geq \frac{A(\rho)}{100}$, then the plan $\bar{\pi}$ already satisfies (4.1) and there is nothing to do.

Case 2. Let us suppose that both of the following conditions hold

$$\int_0^{x_R} \rho dx < \frac{A(\rho)}{100}, \quad \int_0^{x_R} (1 - \rho) dx > \frac{A(\rho)}{100}.$$

In this case, we take a point $x_R^1 > x_R$ such that $\int_0^{x_R^1} \rho dx = \frac{A(\rho)}{100}$, and we try to move mass in the opposite direction in the segment $[0, x_R^1]$. This is necessary in order to take into account densities similar to the characteristic function of the union of two intervals: in that case, the optimal map actually moves mass toward the origin (see Figure 2).

To do this, we consider a transport plan tailored to ρ and depending on x_R^1 , and it is obtained again through a minimization process:

$$\min \left\{ \int |x - y|^p d\pi(x, y) : \pi \in \mathbf{AP}_\rho, \text{spt}\pi \subset D \right\}, \quad (4.2)$$

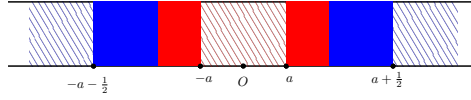


FIGURE 2. Denoting by $\rho = \chi_{[-a-1/2, -a] \cup [a, a+1/2]}$ for some $a > 0$, we represent in this picture the matching induced by a plan realizing every $\mathcal{W}_p(\rho)$. In fact, if $x \in \text{spt}\rho$ has a given color, it will be mapped to a point with the corresponding dashed color.

where $D \subset \mathbb{R} \times \mathbb{R}$ is the following domain:

$$D := \{(x, y) : x \notin (0, x_R^1), x \cdot (y - x) \geq 0\} \cup ([0, x_R^1] \times [0, x_R^1]).$$

Observe that, since $\int_0^{x_R^1} \rho dx = \frac{A(\rho)}{100} < \int_0^{x_R^1} (1 - \rho) dx$, then it is possible to find a minimizer π of (4.2). Applying again the structure theorem for optimal plans in one dimension, we find a map T that induces an optimal plan. This transport problem is actually decoupled, considering independently $\rho \llcorner [0, x_R^1]$ and $\rho - (\rho \llcorner [0, x_R^1])$. Hence, it is possible to adapt [DPMSV16, Lemma 5.1] separately to both pieces and see that $|T(x) - x| \leq d$ for every $x \in [0, x_R^1]$. In fact, if this is not the case, then $T_{\#}\rho = 1 - \rho$ in a segment $I \subset [0, x_R^1]$ longer than d . This is impossible since

$$d \leq |I| = \int_I (\rho + (1 - \rho)) dx = \int_I (\rho + T_{\#}\rho) dx \leq 2 \int_0^{x_R^1} \rho dx = 2 \cdot \frac{A(\rho)}{100} \leq \frac{1}{25},$$

and $d = \frac{1}{2} - \frac{A(\rho)}{100} > \frac{1}{3}$. Having this uniform bound on the transport length in $[0, x_R^1]$, then we see that π satisfies (4.1) because $\int_0^{x_R^1} \rho dx = \frac{A(\rho)}{100}$.

Case 3. Finally, let us suppose that the following inequalities hold at the same time:

$$\int_0^{x_R} \rho dx < \frac{A(\rho)}{100}, \quad \int_0^{x_R} (1 - \rho) dx \leq \frac{A(\rho)}{100}.$$

At this point, we can explore each of the previous cases on the left side of the real line, producing the analogous $x_L = \sup \{x < 0 : x - \bar{T}(x) > d\}$. Since in the first two cases we managed to construct the desired π , we can suppose without loss of generality that we are in Case 3 also on the left side. In other words, the following holds

$$\max \left\{ \int_0^{x_R} \rho dx, \int_0^{x_R} (1 - \rho) dx, \int_{x_L}^0 \rho dx, \int_{x_L}^0 (1 - \rho) dx \right\} \leq \frac{A(\rho)}{100}.$$

Combining these information we obtain an estimate on $|x_R - x_L|$:

$$x_R - x_L = \int_{x_L}^0 (\rho + (1 - \rho)) dx + \int_0^{x_R} (\rho + (1 - \rho)) dx \leq \frac{A(\rho)}{25},$$

and we will see that this is not possible because we can get an inequality for the asymmetry of ρ . We repeat here the argument of Proposition 3.3: adapting [DPMSV16, Lemma 5.1] we

obtain that $\bar{T}_\#(\rho \lfloor_{[x_L, 0]}) = 1 - \rho$ in $[x_L - d, x_L]$ and $\bar{T}_\#(\rho \lfloor_{[0, x_R]}) = 1 - \rho$ in $[x_R, x_R + d]$, and thus

$$\begin{aligned} \int_{x_L-d}^{x_R+d} \rho dx &= \int_{x_L-d}^{x_L} \rho dx + \int_{x_L}^0 \rho dx + \int_0^{x_R} \rho dx + \int_{x_R}^{x_R+d} \rho dx \\ &\geq \int_{x_L-d}^{x_L} \rho dx + \int_{x_L-d}^{x_L} \bar{T}_\# \rho dx + \int_{x_R}^{x_R+d} \bar{T}_\# \rho dx + \int_{x_R}^{x_R+d} \rho dx = 2d. \end{aligned}$$

This means that $\int_{x_L-d}^{x_R+d} \rho \geq 1 - \frac{A(\rho)}{50}$. If $x_R + d - (x_L - d) \leq 1$, then by definition of asymmetry

$$A(\rho) \leq 2 \int_{-\infty}^{x_L-d} \rho dx + 2 \int_{x_R+d}^{+\infty} \rho dx \leq \frac{A(\rho)}{25},$$

that is impossible. Hence, we know that $x_L - d + 1 < x_R + d$. Since we proved that $x_R - x_L \leq \frac{A(\rho)}{25}$, we obtain an inequality always valid in our case: $x_R + d - (x_L - d) = 1 - \frac{A(\rho)}{50} + x_R - x_L \leq 1 + \frac{A(\rho)}{50}$. Therefore, we have that

$$\begin{aligned} A(\rho) &\leq 2 \int_{x_L-d}^{x_L-d+1} (1 - \rho) dx \leq 2 \int_{x_L-d}^{x_R+d} (1 - \rho) dx \leq 2(x_R - x_L + 2d) - 2 \left(1 - \frac{A(\rho)}{50}\right) \\ &\leq 2 + \frac{A(\rho)}{25} - 2 + \frac{A(\rho)}{25} = \frac{2}{25} A(\rho), \end{aligned}$$

and thus we reach a contradiction, concluding the last remaining case. \square

APPENDIX A. SKETCH OF THE MEASURABILITY OF THE CONSTRUCTION IN THEOREM 3.6

In Theorem 3.6 we needed to check that the density

$$(r, \omega) \mapsto \bar{\zeta}^\omega(r)$$

is measurable, where $\bar{\zeta}^\omega$ satisfies $\mathcal{W}_p(\rho^\omega) = W_p(\rho^\omega, \bar{\zeta}^\omega)$. This is necessary to have the representation in (3.12). To do that, we approximate ρ in L^1 with densities $\rho_k \in \mathbf{A}$ that are piecewise constant along the sphere. In other words, for every k there exists a partition of the sphere $\mathbb{S}^{N-1} = \bigcup_j E_j^k$ with sets such that $\text{diam}(E_j^k) + |E_j^k| \leq 1/k$, and such that for every j

$$\rho_k(r\omega) = \rho_k(r\omega') \quad \forall \omega, \omega' \in E_j^k.$$

We construct the following densities: for every k and every $\omega \in \mathbb{S}^{N-1}$ we take ζ such that $\mathcal{W}_p(\rho_k^\omega) = W_p(\rho_k^\omega, \zeta)$ (in the metric-measure sense), and we define

$$\zeta_k(r, \omega) = \zeta(r).$$

In other words, ζ_k^ω is the optimal density to compute $\mathcal{W}_p(\rho_k^\omega)$. This density is measurable since it is piecewise constant along the sphere. Since $\rho_k \rightarrow \rho$ in L^1 , then $\rho_k^\omega \rightarrow \rho^\omega$ in L^1 for a.e. $\omega \in \mathbb{S}^{N-1}$. For this reason, we say that $\zeta_k^\omega \rightarrow \bar{\zeta}^\omega$ in weak* sense for a.e. ω .

To see this, notice that ζ_k^ω converges to some density ϕ^ω because the sequence ρ_k^ω is bounded in L^∞ , and the transport distance is bounded when the mass of ρ_k^ω is finite, that happens for a.e. ω . By lower semicontinuity of the transport distance we have that

$$\mathcal{W}_p(\rho^\omega) = W_p(\rho^\omega, \bar{\zeta}^\omega) \leq W_p(\rho^\omega, \phi^\omega) \leq \liminf_k W_p(\rho_k^\omega, \zeta_k^\omega) = \mathcal{W}_p(\rho^\omega),$$

where we used that $\rho^\omega + \phi^\omega \leq 1$ in the first inequality, and the continuity of \mathcal{W}_p with respect to the weak* convergence in the last equality. Since the optimal density to compute $\mathcal{W}_p(\rho^\omega)$ is unique, then $\bar{\zeta}^\omega = \phi^\omega = \lim_k \zeta_k^\omega$.

We finally conclude because $\zeta_k \rightarrow \zeta_\infty$ for some ζ_∞ in weak* sense, and ζ_∞ is therefore measurable. Moreover, a little argument shows that, whenever $f_k : X \times Y \rightarrow \mathbb{R}$ converges in weak* sense to f (X and Y being reasonable spaces, in our case $X = \mathbb{R}^+$ and $Y = \mathbb{S}^{N-1}$), then for almost every $y \in Y$ we have that

$$f_k \llcorner (X \times \{y\}) \xrightarrow{*} f \llcorner (X \times \{y\}).$$

Hence, for almost every $\omega \in \mathbb{S}^{N-1}$ we have that

$$\zeta_k^\omega \xrightarrow{*} \zeta_\infty^\omega,$$

and our previous argument shows also that

$$\zeta_k^\omega \xrightarrow{*} \bar{\zeta}^\omega \quad \text{for a.e. } \omega \in \mathbb{S}^{N-1}.$$

Combining these facts, we get that $\bar{\zeta} = \zeta_\infty$ almost everywhere, and thus $\bar{\zeta}$ is measurable, as we wanted.

Acknowledgments. D.C. is member of the Istituto Nazionale di Alta Matematica (INdAM), Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA), and is partially supported by the INdAM–GNAMPA 2023 Project *Problemi variazionali per funzionali e operatori non-locali*, codice CUP_E53C22001930001. D.C. wishes to thank Virginia Commonwealth University for the generous hospitality provided during his visit, which marked the beginning of this project. I.T.'s research was partially supported by a Simons Collaboration grant 851065 and an NSF grant DMS 2306962.

REFERENCES

- [BCL20] G. Buttazzo, G. Carlier, and M. Laborde, “On the Wasserstein distance between mutually singular measures,” *Adv. Calc. Var.*, vol. 13, no. 2, pp. 141–154, 2020. [Online]. Available: <https://doi.org/10.1515/acv-2017-0036>
- [BJ17] M. Barchiesi and V. Julin, “Robustness of the Gaussian concentration inequality and the Brunn-Minkowski inequality,” *Calc. Var. Partial Differential Equations*, vol. 56, no. 3, pp. Paper No. 80, 12, 2017. [Online]. Available: <https://doi.org/10.1007/s00526-017-1169-x>
- [CTG22] J. Candau-Tilh and M. Goldman, “Existence and stability results for an isoperimetric problem with a non-local interaction of Wasserstein type,” *ESAIM Control Optim. Calc. Var.*, vol. 28, pp. Paper No. 37, 20, 2022. [Online]. Available: <https://doi.org/10.1051/cocv/2022040>

- [CTGM] J. Candau-Tilh, M. Goldman, and B. Merlet, “An exterior optimal transport problem,” *arXiv preprint arXiv:2309.02806*, posted on September 6, 2023. [Online]. Available: <https://arxiv.org/abs/2309.02806>
- [DPMSV16] G. De Philippis, A. R. Mészáros, F. Santambrogio, and B. Velichkov, “BV estimates in optimal transportation and applications,” *Arch. Ration. Mech. Anal.*, vol. 219, no. 2, pp. 829–860, 2016. [Online]. Available: <https://doi.org/10.1007/s00205-015-0909-3>
- [FL21] R. L. Frank and E. H. Lieb, “Proof of spherical flocking based on quantitative rearrangement inequalities,” *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, vol. 22, no. 3, pp. 1241–1263, 2021.
- [FMP08] N. Fusco, F. Maggi, and A. Pratelli, “The sharp quantitative isoperimetric inequality,” *Ann. of Math. (2)*, vol. 168, no. 3, pp. 941–980, 2008. [Online]. Available: <https://doi.org/10.4007/annals.2008.168.941>
- [FMP10] A. Figalli, F. Maggi, and A. Pratelli, “A mass transportation approach to quantitative isoperimetric inequalities,” *Invent. Math.*, vol. 182, no. 1, pp. 167–211, 2010. [Online]. Available: <https://doi.org/10.1007/s00222-010-0261-z>
- [FP20] N. Fusco and A. Pratelli, “Sharp stability for the Riesz potential,” *ESAIM Control Optim. Calc. Var.*, vol. 26, pp. Paper No. 113, 24, 2020. [Online]. Available: <https://doi.org/10.1051/cocv/2020024>
- [LPR14] L. Lussardi, M. A. Peletier, and M. Röger, “Variational analysis of a mesoscale model for bilayer membranes,” *J. Fixed Point Theory Appl.*, vol. 15, no. 1, pp. 217–240, 2014. [Online]. Available: <https://doi.org/10.1007/s11784-014-0180-5>
- [NTV23] M. Novack, I. Topaloglu, and R. Venkatraman, “Least Wasserstein distance between disjoint shapes with perimeter regularization,” *J. Funct. Anal.*, vol. 284, no. 1, pp. Paper No. 109732, 26, 2023. [Online]. Available: <https://doi.org/10.1016/j.jfa.2022.109732>
- [PR09] M. A. Peletier and M. Röger, “Partial localization, lipid bilayers, and the elastica functional,” *Arch. Ration. Mech. Anal.*, vol. 193, no. 3, pp. 475–537, 2009. [Online]. Available: <https://doi.org/10.1007/s00205-008-0150-4>
- [Str08] M. Struwe, *Variational methods*, 4th ed., ser. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 2008, vol. 34, applications to nonlinear partial differential equations and Hamiltonian systems.
- [Vil03] C. Villani, *Topics in optimal transportation*, ser. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2003, vol. 58. [Online]. Available: <https://doi.org/10.1090/gsm/058>
- [XZ21] Q. Xia and B. Zhou, “The existence of minimizers for an isoperimetric problem with Wasserstein penalty term in unbounded domains,” *Advances in Calculus of Variations*, p. 000010151520200083, 2021. [Online]. Available: <https://doi.org/10.1515/acv-2020-0083>

(Almut Burchard) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, ON, CANADA
Email address: `almut@math.toronto.edu`

(Davide Carazzato) SCUOLA NORMALE SUPERIORE, PISA, ITALY
Email address: `davide.carazzato@sns.it`

(Ihsan Topaloglu) DEPARTMENT OF MATHEMATICS AND APPLIED MATHEMATICS, VIRGINIA COMMONWEALTH UNIVERSITY, RICHMOND VA, UNITED STATES
Email address: `iatopaloglu@vcu.edu`