

Equality between Monge and Kantorovich multimarginal problems with Coulomb cost

Maria Colombo ^{*} Simone Di Marino [†]

April 11, 2013

Abstract

A standard question arising in Optimal Transport Theory is whether the Monge problem and the Kantorovich relaxation have the same infimum; the positive answer means that we can pass to the relaxed problem without loss of information. In the classical case with two marginals, this happens when the cost is positive, continuous, possibly infinite and the first marginal has no atoms. We study a similar multimarginal symmetric problem, arising naturally in density functional theory, motivated by a recent paper by Buttazzo, De Pascale and Gori Giorgi. The cost is the potential interaction between n charged particles (hence it is symmetric, positive, continuous, and infinite whenever $x_i = x_j$) and the marginals are all equal with no atoms. We prove that also in this case there is equality between the infimum in the cyclical Monge problem (the natural Monge problem in this context) and in the classical Kantorovich problem. This result is new even for 2 marginals, because we consider only transport maps which are involutions. The result is generalized to every symmetric continuous cost function on a Polish space.

1 Introduction

In a recent paper by Buttazzo, De Pascale and Gori Giorgi [3] a mathematical model for the strong interaction limit in the density functional theory is studied (see also [6, 14]). In particular they observe that the minimal interaction of n electrons in \mathbb{R}^d whose density is μ can be represented as the minimum of a multimarginal Monge cyclical problem (the cyclicity is due to the symmetries of the problem).

The problem can be described as follows. We consider the cost function $c : (\mathbb{R}^d)^n \rightarrow \mathbb{R}$ given by

$$c(x_1, \dots, x_n) = \sum_{1 \leq i < j \leq n} \frac{1}{|x_i - x_j|} \quad \forall (x_1, \dots, x_n) \in (\mathbb{R}^d)^n \quad (1.1)$$

^{*}Scuola Normale Superiore, Pisa. email: maria.colombo@sns.it

[†]Scuola Normale Superiore, Pisa. email: simone.dimarino@sns.it

and the cyclical multimarginal problem (of Monge type)

$$(M) = \inf \left\{ \int_{\mathbb{R}^d} c(x, T(x), \dots, T^{(n-1)}(x)) d\mu(x) : T : \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ Borel, } T_{\#}\mu = \mu, T^{(n)} = Id \right\}, \quad (1.2)$$

where $\mu \in \mathcal{P}(\mathbb{R}^d)$ is a given probability measure on \mathbb{R}^d , $T^{(i)}$ stands for the i -th composition of T with itself, and $T_{\#}\mu$ represents the pushforward measure of the measure μ through the Borel map T .

A natural question arising both from the modeling and from a mathematical point of view is whether there exists a map T such that the minimum in (1.2) is achieved under suitable assumptions on μ . This question is rather difficult and up to now a positive answer is known only in dimension 1, where an explicit map T can be provided (see [5]). For some particular costs generated by vector fields the existence of an optimal maps has been provided by N. Ghoussoub, A. Moameni [8] and by N. Ghoussoub, B. Maurey [9]. In general, the problem of finding multimarginal optimal transport maps has been solved only under special assumptions on the local behavior of the cost (see [11, 13, 7]).

However, following the theory of optimal transportation (see [16, 13]) one can introduce the cost of a plan $\pi \in \mathcal{P}((\mathbb{R}^d)^n)$

$$\mathcal{C}(\pi) = \int_{(\mathbb{R}^d)^n} c(x_1, \dots, x_n) d\pi(x_1, \dots, x_n)$$

and a problem of Kantorovich type

$$(K) := \inf \left\{ \mathcal{C}(\pi) : \pi \in \mathcal{P}((\mathbb{R}^d)^n), (e_i)_{\#}\pi = \mu \forall i \in \{1, \dots, n\} \right\}, \quad (1.3)$$

where $e_i : (\mathbb{R}^d)^n \rightarrow \mathbb{R}^d$ are the projections on the i -th component. Roughly speaking, through the problem (K) we allow the splitting of mass. It can be easily seen that the problem (K) is linear with respect to plans and that the admissible plans form a bounded, weakly* compact subset of the set of measures on $(\mathbb{R}^d)^n$; hence the existence of a minimum in (K) follows easily from the lower semicontinuity of the cost (and therefore of \mathcal{C}).

Since to every transport map T corresponds a transport plan $(Id, T, \dots, T^{(n-1)})_{\#}\mu$, we have obviously that $(K) \leq (M)$. A natural question regarding the relation between (K) and (M), is whether the weaker problem (K) is consistent with (M), namely if $(K) = (M)$ in the case of a non-atomic probability μ . In the following we give a positive answer, proving that the cost of any transport plan can be well approximated through maps with the same marginals.

Theorem 1.1. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a non-atomic probability measure. Let c be the cost function (1.1). Then*

$$(K) = (M). \quad (1.4)$$

Remark 1.2. *Since the cost c is symmetric, the Kantorovich problem can be done considering only symmetric plans π . In other words,*

$$(K) = \inf \left\{ \mathcal{C}(\pi) : \pi \in \mathcal{P}((\mathbb{R}^d)^n), (e_i)_{\#}\pi = \mu \forall i \in \{1, \dots, n\}, \sigma_{\#}\pi = \pi \forall \sigma \in \mathcal{O}_n \right\}, \quad (1.5)$$

where \mathcal{O}_n is the set of maps which permute the coordinates of $(\mathbb{R}^d)^n$, formally defined below. Indeed, given a transport plan by the linearity of the transport problem the plan

$$\frac{1}{n!} \sum_{\sigma \in \mathcal{O}_n} \sigma_{\#} \pi$$

has the same cost and the same marginals as π and it is symmetric.

In a work of A. Pratelli [12], it is shown that in every Polish space there is equivalence between Monge problem and Kantorovich problem with two marginals, in the sharp case of a continuous cost, possibly unbounded and infinite, with the first marginal with no atoms. Theorem 1.1, which holds true as well in metric spaces, provides a version of this result to multimarginal problems with cyclic costs and maps. We remark that the generalization of [12] without requiring any cyclical structure on the maps would be straightforward, as pointed out below.

Remark 1.3. For a general n , it follows easily from the results in [12] that the cost of any plan can be approximated through $n - 1$ maps, in particular

$$(K) = \inf \left\{ \int_{(\mathbb{R}^d)^n} c(x, T_2(x), \dots, T_{n-1}(x)) d\mu(x) : T_i : \mathbb{R}^d \rightarrow \mathbb{R}^d, T_{i\#} \mu = \mu \forall i \in \{2, \dots, n\} \right\}. \quad (1.6)$$

Indeed one can consider the latest problem as a two-marginal problem between X and X^{n-1} . The main purpose of Theorem 1.1 is to show that we can approximate the cost of each transport plan with special maps of the form $(Id, T, T^{(2)}, \dots, T^{(n-1)})$, where $T^{(n)} = Id$.

Finally, we point out that the extension of Theorem 1.1 to metric spaces is almost straightforward. More precisely, we prove that the infimum of Monge multimarginal problem with cyclical maps is the same as the minimum of Kantorovich problem under the sharp assumption that the cost c is cyclical and continuous. Given $\tau : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ a permutation, we define $\sigma_\tau : (\mathbb{R}^d)^n \rightarrow (\mathbb{R}^d)^n$ as

$$\sigma_\tau(x_1, \dots, x_n) = (x_{\tau(1)}, \dots, x_{\tau(n)}) \quad \forall (x_1, \dots, x_n) \in (\mathbb{R}^d)^n. \quad (1.7)$$

We call \mathcal{O}_n the collection of these functions. We say that a possibly infinity-valued cost function c on $(\mathbb{R}^d)^n$ is symmetric if $c(\sigma) = c$ for every $\sigma \in \mathcal{O}_n$. We say that a cost function is cyclical if $c(\bar{\sigma}) = c$ for

$$\bar{\sigma}(x_1, \dots, x_n) = (x_2, \dots, x_n, x_1) \quad \forall (x_1, \dots, x_n) \in (\mathbb{R}^d)^n. \quad (1.8)$$

We notice that if $n = 2$ a cost is cyclical if and only if it is symmetric.

Theorem 1.4. Let (X, d) be a Polish space and $\mu \in \mathcal{P}(X)$ a non-atomic probability measure on X . Let $c : X^n \rightarrow [0, \infty]$ be a continuous cyclical cost. Then

$$\begin{aligned} & \min \left\{ \int_{X^n} c(x_1, \dots, x_n) d\pi(x_1, \dots, x_n) : \pi \in \mathcal{P}(X^n), (e_i)_{\#} \pi = \mu \forall i \in \{1, \dots, n\} \right\} \\ & = \inf \left\{ \int_X c(x, T(x), \dots, T^{(n-1)}(x)) d\mu(x) : T : X \rightarrow X \text{ Borel map, } T_{\#} \mu = \mu, T^{(n)} = Id \right\}. \end{aligned} \quad (1.9)$$

To make more clear the key ideas of the proof avoiding the technical details, we present in Section 2 a toy situation, namely the case of a bounded uniformly continuous symmetric cost on \mathbb{R}^d with two marginals. Also in this case the statement does not follow directly from more classical results, or from [12], because we require the almost-optimal maps to be involutions. Section 3 is devoted to the proof of Theorem 1.1. Finally, in Section 4 we generalize Theorem 1.1 to multimarginal problems with a cyclical continuous cost on a metric space.

Acknowledgements. We wish to thank Luigi Ambrosio for pointing out the problem. We also thank Luigi Ambrosio and Luigi De Pascale for a careful reading of the paper.

2 A uniformly continuous cost with 2 marginals: a toy case

In this section we prove the main result with a simplified cost function and with 2 marginals. We remark that, under standard assumptions on the cost implying that a unique optimal transport map exists between $\mu \ll \mathcal{L}^d$ and itself (namely the twist condition, see for example [16]), the optimal map is an involution, hence the minimum in the symmetric Monge problem is achieved. Indeed, since T and T^{-1} are both optimal transport maps between μ and itself; since the optimal transport map is unique under these assumptions, we obtain that $T = T^{-1}$ almost everywhere, which proves that T is an involution.

Theorem 2.1. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a non-atomic probability measure. Let $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$ be a uniformly continuous, symmetric cost. Then $(K) = (M)$.*

An analogous result without assuming any symmetry of the cost function and of the constructed map has been proved in [1].

Before proving the result, we present two lemmas. In the first one we show the existence of an invertible transport map between any couple of nonatomic measures with the same mass in \mathbb{R}^d . We remark that, if we assume that the two measures are absolutely continuous with respect to the Lebesgue measure, then the next lemma can be proved simply by considering an optimal transport map with respect to the quadratic cost $|x - y|^2$. Indeed this map is invertible almost everywhere and its inverse is given by the inverse transport between the two measures (see [16]).

Lemma 2.2. *Let μ be a non-atomic measure on \mathbb{R}^d and let A, B be Borel sets such that $\mu(A) = \mu(B) < \infty$. Then there exist two Borel maps $T : A \rightarrow B$ and $S : B \rightarrow A$ with the following properties:*

- $T_{\#}(1_A \mu) = 1_B \mu$ and $S_{\#}(1_B \mu) = 1_A \mu$;
- the maps T and S are μ -almost inverse in the sense that $T \circ S(x) = x$ for μ -almost every $x \in B$ and $S \circ T(x) = x$ for μ -almost every $x \in A$.

Proof. From [12, Theorem 1.4] applied with $X = Y = \mathbb{R}^d$ to the two measures $1_A\mu$ and $1_B\mu$, we obtain a transport map ϕ between two sets of full measure $\tilde{A} \subseteq \mathbb{R}^d$ and $\tilde{B} \subseteq \mathbb{R}^d$ that is also a homeomorphism; let now $\hat{A} = \phi^{-1}(\tilde{B})$ and $\hat{B} = \phi(\tilde{A})$. Let us note that since ϕ is a transport between $1_A\mu$ and $1_B\mu$ we have $\mu(\hat{A} \cap A) = \mu(\tilde{B}) = \mu(\tilde{A}) = \mu(A)$ and so $\hat{A} \cap A$ has full measure; similarly also $\hat{B} \cap B$ has full measure. Given $a_0 \in A$ and $b_0 \in B$, we define

$$T(x) = \begin{cases} \phi(x) & \text{if } x \in \hat{A} \cap A \\ b_0 & \text{otherwise} \end{cases} \quad S(x) = \begin{cases} \phi^{-1}(x) & \text{if } x \in \hat{B} \cap B \\ a_0 & \text{otherwise,} \end{cases}$$

and we claim that these functions satisfy the requested properties.

Because of the definitions of \hat{A} and \hat{B} the image of T is contained in B and the image of S is contained in A . Moreover we have that $T \circ S(x) = x$ on $\hat{B} \cap B$ and $S \circ T(x) = x$ on $\hat{A} \cap A$, and so μ -almost everywhere on B and A . Since we modified ϕ in sets of null measure, T and S are still transport maps between $1_A\mu$ and $1_B\mu$. \square

In the following lemma we estimate the difference of the cost of two transport plans with the oscillation of the cost on their support.

Lemma 2.3. *Let $Q \subseteq (\mathbb{R}^d)^n$ be a measurable set and let π and γ be transport plans concentrated on Q and such that $\pi(Q) = \gamma(Q)$.*

Then we have that

$$\mathcal{E}(\pi) - \mathcal{E}(\gamma) \leq \pi(Q) \sup_{\mathbf{x}, \mathbf{y} \in Q} |c(\mathbf{x}) - c(\mathbf{y})|.$$

Proof. We have that

$$\begin{aligned} \mathcal{E}(\pi) - \mathcal{E}(\gamma) &= \int_{(\mathbb{R}^d)^n} c(\mathbf{x}) d\pi(\mathbf{x}) - \int_{(\mathbb{R}^d)^n} c(\mathbf{y}) d\gamma(\mathbf{y}) \\ &= \frac{1}{\gamma(Q)} \int_{(\mathbb{R}^d)^n} \int_{(\mathbb{R}^d)^n} (c(\mathbf{x}) - c(\mathbf{y})) d\pi(\mathbf{x}) d\gamma(\mathbf{y}) \leq \pi(Q) \sup_{\mathbf{x}, \mathbf{y} \in Q} |c(\mathbf{x}) - c(\mathbf{y})|. \end{aligned}$$

\square

The proof of Theorem 2.1 is based on a decomposition of $\mathbb{R}^d \times \mathbb{R}^d$ in square cells where the cost has a small oscillation. Before proving the Theorem we introduce our simple partitions. For every $k \in \mathbb{N}$ we partition \mathbb{R}^d in cubic cells of side length $1/2^k$

$$\mathcal{S}^k(\mathbb{R}^d) = \left\{ \left[\frac{i_1}{2^k}, \frac{i_1 + 1}{2^k} \right) \times \dots \times \left[\frac{i_d}{2^k}, \frac{i_d + 1}{2^k} \right) : i_1, \dots, i_d \in \mathbb{Z}^d \right\}. \quad (2.1)$$

We consider a partition of $\mathbb{R}^d \times \mathbb{R}^d$ made of squares which are products of cubic cells with side 2^{-k} , namely

$$\mathcal{Q}^k(\mathbb{R}^d \times \mathbb{R}^d) = \{C \times C' : C, C' \in \mathcal{S}^k(\mathbb{R}^d)\}.$$

Proof of Theorem 2.1. We prove that, given a plan π (namely, $\pi \in \mathcal{P}((\mathbb{R}^d)^2)$) such that $e_{1\sharp}\pi = \mu$ and $e_{2\sharp}\pi = \mu$) and an $\varepsilon > 0$, there exists a Borel map $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$T_{\sharp}\mu = \mu, \quad (2.2)$$

$$T(T(x)) = x \text{ for } \mu - a.e. x \in \mathbb{R}^d, \quad (2.3)$$

and

$$\mathcal{C}((Id, T)_{\sharp}\mu) \leq \mathcal{C}(\pi) + \varepsilon\|\pi\|. \quad (2.4)$$

We notice that we can assume π to be symmetric (namely $\pi(A \times B) = \pi(B \times A)$ for every $A, B \subseteq \mathbb{R}^d$ measurable sets) as observed in Remark 1.2.

Due to the uniform continuity of the cost, there exists a k sufficiently large such that in every square of $\mathcal{Q}^k(\mathbb{R}^d \times \mathbb{R}^d)$ the cost c oscillates less than ε .

Since \mathcal{S}^k is countable, we enumerate its elements $\mathcal{S}^k = \{C_i\}_{i \in \mathbb{N}}$. For every cell C_i we consider the squares in $\mathcal{Q}^k(\mathbb{R}^d \times \mathbb{R}^d)$ with projection C_i on the first \mathbb{R}^d , namely $\{C_i \times C_j : j \in \mathbb{N}\}$. Since $e_{1\sharp}\pi = \mu$, their masses $\{\pi(C_i \times C_j) : j \in \mathbb{N}\}$ satisfy

$$\sum_{j=1}^{\infty} \pi(C_i \times C_j) = \mu(C_i) \quad \forall i \in \mathbb{N}.$$

We partition C_i in disjoint sets $E_{i,j}$ such that $\mu(E_{i,j}) = \pi(C_i \times C_j)$: it is possible since μ has no atoms. We consider the set $E_{i,j} \times E_{j,i}$; due to the symmetry of the plan, we have that $\pi(C_i \times C_j) = \pi(C_j \times C_i)$, and so we have that $\mu(E_{i,j}) = \mu(E_{j,i})$. If $\mu(E_{i,j}) \neq 0$ and $i \leq j$ let $T_{i,j}$ and $T_{j,i}$ be the maps given by Lemma 2.2 applied to the sets $E_{i,j}$ and $E_{j,i}$ (we consider the identity map if $i = j$). For every $i, j \in \mathbb{N}$ the graph of the map $T_{i,j}$ lies in $E_{i,j} \times E_{j,i}$ and we have that $T_{j,i} \circ T_{i,j} = Id$ almost everywhere on $E_{i,j}$. Gluing all these maps together we define a map $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$, thanks to the fact that the sets $E_{i,j}$ are disjoint and their union is μ -a.e. \mathbb{R}^d

$$T(x) = \begin{cases} T_{i,j}(x) & \text{if } x \in E_{i,j} \\ 0 & \text{otherwise.} \end{cases}$$

We prove that (2.2), (2.3), and (2.4) are satisfied. Since the sets $E_{i,j}$ are disjoint and since by definition $T_{i,j\sharp}(\mu\chi_{E_{i,j}}) = \mu\chi_{E_{j,i}}$ for every couple $i, j \in \mathbb{N}$, we have that

$$T_{\sharp}(\mu) = T_{\sharp}\left(\mu \sum_{i,j=1}^{\infty} \chi_{E_{i,j}}\right) = \sum_{i,j=1}^{\infty} T_{i,j\sharp}(\mu\chi_{E_{i,j}}) = \sum_{i,j=1}^{\infty} \mu\chi_{E_{j,i}} = \mu,$$

which proves (2.2). To prove (2.3) it is sufficient to note that on $E_{i,j}$ the map T is $T_{i,j}$ and maps $E_{i,j}$ on $E_{j,i}$, thus is clear that $T \circ T = T_{j,i} \circ T_{i,j}$ on $E_{i,j}$. By construction we have that $T_{i,j} \circ T_{j,i} = Id$ μ -a.e. in $E_{j,i}$. Finally, we apply Lemma 2.3 to the measures $\pi 1_{C_i \times C_j}$ and $(Id, T_{i,j})_{\sharp}(\mu 1_{E_{i,j}})$, both concentrated on $C_i \times C_j$, where the cost oscillates less than ε ; we get

$$\mathcal{C}((Id, T_{i,j})_{\sharp}(\mu 1_{E_{i,j}})) \leq \mathcal{C}(\pi 1_{Q_{i,j}}) + \varepsilon\|\pi 1_{Q_{i,j}}\|.$$

Adding these inequalities over $i, j \in \mathbb{N}$ we obtain (2.4). \square

3 The multimarginal problem with a Coulomb cost

The proof of Theorem 2.1 was based on a decomposition of $\mathbb{R}^d \times \mathbb{R}^d$ in cubes of the same size where the cost has a small oscillation. Since the Coulomb cost (1.1) is not uniformly continuous, in the following lemma we do a partition of $(\mathbb{R}^d)^n$ in cubes adapted to the cost.

Before stating the lemma, we introduce a family of subsets of $(\mathbb{R}^d)^n$ where we will choose the partition. Given the family of cells in \mathbb{R}^d defined in (2.1) we consider the family of cubes in $(\mathbb{R}^d)^n$ such that each projection on \mathbb{R}^d is a cubic cell of side 2^{-k} for some k

$$\mathcal{Q}^k((\mathbb{R}^d)^n) = \{C_1 \times \dots \times C_n : C_1, \dots, C_n \in \mathcal{S}^k(\mathbb{R}^d)\}, \quad \forall k \in \mathbb{N}$$

and we set

$$\mathcal{Q}((\mathbb{R}^d)^n) = \bigcup_{k=1}^{\infty} \mathcal{Q}^k((\mathbb{R}^d)^n).$$

Lemma 3.1 (Partition of $\{c < \infty\}$). *Let $\varepsilon > 0$ and let c be as in (1.1). There exists a subset \mathcal{F} of $\mathcal{Q}((\mathbb{R}^d)^n)$ such that the following properties hold true.*

1. \mathcal{F} is a disjoint covering of $\{c < \infty\}$

$$\bigcup_{Q \in \mathcal{F}} Q = \{c < \infty\}. \quad (3.1)$$

2. The cost c oscillates at most of ε inside a cube, namely for every $Q \in \mathcal{F}$

$$|c(x) - c(y)| \leq \varepsilon \quad \forall x, y \in Q. \quad (3.2)$$

3. The family is cyclical, namely for every $Q \in \mathcal{F}$ we have

$$\bar{\sigma}(Q) \in \mathcal{F}, \quad (3.3)$$

where $\bar{\sigma}$ is defined as in (1.8).

Proof. For every $\mathbf{x} = (x_1, x_2, \dots, x_n) \in (\mathbb{R}^d)^n$ we consider all the cubes $Q \in \mathcal{Q}((\mathbb{R}^d)^n)$ to which \mathbf{x} belongs: there is exactly one cube in every $\mathcal{Q}^k((\mathbb{R}^d)^n)$ and those cubes form a chain by inclusion. We associate to every $\mathbf{x} \in \{c < \infty\}$ the biggest cube $Q_{\mathbf{x}}$ in $\mathcal{Q}((\mathbb{R}^d)^n)$ containing \mathbf{x} and where the oscillation of the cost is less than ε , namely where (3.2) holds true. Let $\mathcal{F} = \{Q_{\mathbf{x}} : c(\mathbf{x}) < \infty\}$. By construction it enjoys (3.2); moreover it is made of disjoint cubes. Indeed the elements of $\mathcal{Q}((\mathbb{R}^d)^n)$ (and hence of \mathcal{F}) are either disjoint or one contained in another; the second case cannot happen in \mathcal{F} because of the maximality condition on the oscillation of the cost. Furthermore, because of the property of bounded oscillations (3.2), for each $Q \in \mathcal{F}$ we have that $Q \subseteq \{c < \infty\}$ and so we get that

$$\{c < \infty\} \subseteq \bigcup_{\{c(\mathbf{x}) < \infty\}} Q_{\mathbf{x}} = \bigcup_{Q \in \mathcal{F}} Q \subseteq \{c < \infty\}$$

which proves (3.1). Given $\bar{\sigma}$ defined as in (1.8), since $\mathcal{Q}((\mathbb{R}^d)^n)$ is $\bar{\sigma}$ -invariant and $osc_Q(c) = osc_{\bar{\sigma}(Q)}(c)$ (because the cost is cyclical), we have that $Q_{\bar{\sigma}(\mathbf{x})} = \bar{\sigma}(Q_{\mathbf{x}})$ for every $\mathbf{x} \in \{c < \infty\}$. Hence \mathcal{F} is cyclical. \square

In the following lemma we construct a partition of \mathbb{R}^d adapted to a transport plan starting from the partition of $\{c < \infty\}$ constructed in Lemma 3.1.

Lemma 3.2. *Let $\varepsilon > 0$, c be as in (1.1), and \mathcal{F} be the partition introduced in Lemma 3.1. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a non-atomic probability measure and let $\pi \in \mathcal{P}((\mathbb{R}^d)^n)$ be a cyclical transport plan, namely a cyclical probability measure such that $e_{i\sharp}\pi = \mu$ for every $i = 1, \dots, n$.*

Then for every $Q \in \mathcal{F}$ there exists a Borel set $A_Q \subseteq \mathbb{R}^d$ such that the following properties hold true:

1. A_Q is contained in the first projection on Q and has measure $\pi(Q)$,

$$A_Q \subseteq e_1(Q), \quad \pi(Q) = \mu(A_Q) \quad \forall Q \in \mathcal{F}; \quad (3.4)$$

2. Sets associated with different cubes are essentially disjoint

$$\mu(A_Q \cap A_{Q'}) = 0 \quad \forall Q, Q' \in \mathcal{F}, Q \neq Q'. \quad (3.5)$$

To prove the lemma, we define a (total) order of construction on \mathcal{F} and a geometric order.

First order (\prec) Let $\mathcal{F}^k = \mathcal{F} \cap \mathcal{Q}^k((\mathbb{R}^d)^n)$ be the cubes in the family \mathcal{F} that have side equal to $1/2^k$, named in the following the k -th generation. Since every \mathcal{F}^k is at most countable, we enumerate its elements

$$\mathcal{F}^k = \{Q_{k,1}, Q_{k,2}, \dots\}.$$

We say that $Q \succ Q'$ (Q is older than Q') if Q belongs to a generation strictly smaller than Q' , or if it has a greater index in the same generation \mathcal{F}^k . In other words we have that $Q_{i,j} \succ Q_{i',j'}$ if $i < i'$ or if $i = i'$ and $j > j'$. We notice that \prec is a well-order if restricted to $\cup_{i=1}^k \mathcal{F}^i$ for some k , so that we can use induction on it.

Second order (\triangleleft) We say that that $Q' \triangleleft Q$ if $Q' \prec Q$ and $e_1(Q') \subseteq e_1(Q)$; this is not a total order. By construction we note that Q is older than Q' then either $e_1(Q') \cap e_1(Q) = \emptyset$ or $Q' \triangleleft Q$.

Heuristics for Lemma 3.2 We notice that once we fix A_Q for some Q then, if we consider all $Q' \prec Q$, there might be one such that $e_1(Q') \setminus A_Q$ is too little and so we haven't "space" to have a $A_{Q'} \subseteq e_1(Q')$ disjoint from A_Q and with the right measure.

Our strategy will be to choose A_Q for the smaller Q and then inductively for bigger and bigger cubes. In this way, we will have always "space" to do so. However, since we don't have a smallest Q we have to do an approximate construction. For every k we build inductively A_Q^k that satisfy our conditions (namely they are disjoint and with the right measure) for every $Q \in \cup_{j \leq k} \mathcal{F}^j$, from the smallest to the largest; we do a coherent construction, i.e. such that A_Q^k converges to some set A_Q in measure as $k \rightarrow \infty$. We conclude passing to the limit the properties of A_Q^k .

Before proving the result, we prove a simple measure-theoretic tool. In the following, given a Borel set $A \subseteq \mathbb{R}^d$ we denote by $\mathcal{B}(A)$ the family of Borel sets in \mathbb{R}^d which are contained in A .

Lemma 3.3. *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a non-atomic probability measure. Then for every $A, B \in \mathcal{B}(\mathbb{R}^d)$ such that $B \subseteq A$ there exists a function $\mathbb{B}(A, B, \cdot) : [0, \mu(A \setminus B)] \rightarrow \mathcal{B}(A)$ such that*

$$\mu(\mathbb{B}(A, B, t) \setminus B) = t \quad \forall t \in [0, \mu(A \setminus B)] \quad (3.6)$$

and

$$\mathbb{B}(A, B', t) \subseteq \mathbb{B}(A, B, t) \quad \forall B' \subseteq B. \quad (3.7)$$

Proof. We first deal with the case $B = \emptyset$. We can assume that $\mu(A) = 1$. Using the Lyapunov convexity theorem for non-atomic measures [10], we construct a Borel set $A_{1/2} \subseteq A$ such that $\mu(A_{1/2}) = 1/2$; using again this theorem, we inductively construct sets $A_{i/2^n}$ for every $i = 1, \dots, 2^n$, with $A_0 = \emptyset$ and $A_1 = A$, such that $A_{i/2^n} \subseteq A_{j/2^n}$ if $i \leq j$ and such that $\mu(A_{i/2^n}) = i/2^n$. In this way we define the function $\mathbb{B}(A, \emptyset, \cdot)$ on the dyadic numbers. Since it is 1-Lipschitz on a dense set, and $\mathcal{B}(A)$ is complete when endowed with the L^1 topology, it can be extended to a continuous function; this function satisfies (3.6) and

$$\mathbb{B}(A, \emptyset, t) \subseteq \mathbb{B}(A, \emptyset, t') \quad \forall 0 \leq t \leq t' \leq 1, \quad (3.8)$$

For every $B \in \mathcal{B}(\mathbb{R}^d)$ and $t \in [0, \mu(A \setminus B)]$ we define s to be a solution to

$$\mu(\mathbb{B}(A, \emptyset, s) \setminus B) = t \quad (3.9)$$

(note that the left-hand side is continuous and nondecreasing as a function of s , it is 0 at $s = 0$, it is $\mu(A \setminus B)$ at $s = 1$) and we set

$$\mathbb{B}(A, B, t) = \mathbb{B}(A, \emptyset, s).$$

Property (3.6) is true by definition. (3.7) follows from (3.9) and (3.8). \square

Proof of Lemma 3.2. Description of the k -th construction: We define $A_Q^k = \emptyset$ for every $Q \in \mathcal{F} \setminus \bigcup_{j \leq k} \mathcal{F}^k$. Then we construct A_Q^k by induction in the set $\bigcup_{j \leq k} \mathcal{F}^k$, which is well-ordered with respect to \prec . Given $Q \in \bigcup_{j \leq k} \mathcal{F}^k$, let $B_Q^k \subseteq \mathbb{R}^d$ the possibly empty set

$$B_Q^k = \bigcup_{Q' \prec Q}^{(k)} A_{Q'}^k,$$

(where in the following the notation $\bigcup^{(k)}$ means that we are considering only the cubes up to the k -th generation, namely in $\bigcup_{j \leq k} \mathcal{F}^k$). Since we have that $Q' \prec Q$ implies that $e_1(Q') \subseteq e_1(Q)$ and so $Q' \subseteq e_1^{-1}(e_1(Q))$ and since all the cubes in \mathcal{F} are disjoint, we have

$$\pi(Q) \leq \pi(e_1^{-1}(e_1(Q))) - \pi\left(\bigcup_{Q' \prec Q}^{(k)} Q'\right) = \mu(e_1(Q)) - \mu(B_Q^k) = \mu(e_1(Q) \setminus B_Q^k),$$

where we used that $(e_1)_\# \pi = \mu$. Now, thanks to Lemma 3.3 we can define

$$A_Q^k = \mathbb{B}(e_1(Q), B_Q^k, \pi(Q)) \setminus B_Q^k,$$

which is disjoint from B_Q^k .

Existence of the limit: We show by induction on the order \prec that

$$A_Q^k \subseteq A_Q^{k+1} \cup B_Q^{k+1}, \quad B_Q^k \subseteq B_Q^{k+1}. \quad (3.10)$$

Indeed by inductive assumption the same relation olds true for every $Q' \prec Q$ and hence

$$B_Q^k = \bigcup_{Q' \prec Q}^{(k)} A_{Q'}^k \subseteq \bigcup_{Q' \prec Q}^{(k)} (A_{Q'}^{k+1} \cup B_{Q'}^{k+1}) = B_Q^{k+1}. \quad (3.11)$$

Thanks to (3.7) and (3.11) we therefore have

$$A_Q^k = \mathbb{B}(e_1(Q), B_Q^k, \pi(Q)) \setminus B_Q^k \subseteq \mathbb{B}(e_1(Q), B_Q^{k+1}, \pi(Q)) \subseteq A_Q^{k+1} \cup B_Q^{k+1},$$

which proves (3.10).

Since $\mu(A_Q^k) = \mu(A_Q^{k+1})$ by (3.6), we obtain that

$$\mu(A_Q^k \Delta A_Q^{k+1}) = 2\mu(A_Q^k \setminus A_Q^{k+1}).$$

From (3.10) and since $A_Q^k \cap B_Q^k = \emptyset$ we obtain that $A_Q^k \setminus A_Q^{k+1} \subseteq B_Q^{k+1} \setminus B_Q^k$ and therefore

$$\mu(A_Q^k \setminus A_Q^{k+1}) \leq \mu(B_Q^{k+1} \setminus B_Q^k) \leq \sum_{Q' \prec Q}^{(k+1)} \pi(Q') - \sum_{Q' \prec Q}^{(k)} \pi(Q') = \sum_{\substack{Q' \prec Q \\ Q' \in \mathcal{F}^{k+1}}} \pi(Q').$$

Hence for every $Q \in \mathcal{F}$ the sequence A_Q^k converges in measure (i.e. the characteristic functions converge in L^1) to some set A_Q that has the same mass; furthermore since A_Q^k and $A_{Q'}$ are disjoint we deduce that A_Q and $A_{Q'}$ are essentially disjoint, i.e. their intersection is μ -negligible. \square

3.1 Proof of Theorem 1.1

In the following we prove that, given a transport plan π , there exists a map whose cost differs less then ε from the cost of π . The construction of the map is modeled on the following property of graphs of functions. Given $A_{Q,1}, A_{Q,2}, \dots, A_{Q,n}$ disjoint subsets of $(\mathbb{R}^d)^n$ and $(T_2(x), \dots, T_n(x)) : A_{Q,1} \rightarrow A_{Q,2} \times \dots \times A_{Q,n}$ so that T_i is bijective between $A_{Q,1}$ and $A_{Q,i}$ for every $i = 2, \dots, n$, we set G the graph of this map, namely

$$G = \{(x, T_2(x), \dots, T_n(x)) : x \in A_{Q,1}\}.$$

Then we have that, given $\bar{\sigma}$ defined as in (1.8)

$$\begin{aligned} \bar{\sigma}^i(G) &= \{(T_i(x), \dots, T_n(x), x, T_2(x), \dots, T_{i-1}(x)) : x \in A_{Q,1}\} \\ &= \{(x, T_{i+1} \circ T_i^{-1}(x), \dots, T_n \circ T_i^{-1}(x), T_i^{-1}(x), \dots, T_i \circ T_i^{-1}(x)) : x \in A_{Q,i}\}. \end{aligned} \quad (3.12)$$

Proof of Theorem 1.1. Let $\varepsilon > 0$ and let π be an admissible plan for the Kantorovich problem (1.3); thanks to Remark 1.2, we can assume π to be symmetric (namely $\sigma_{\#}\pi = \pi$ for every $\sigma \in \mathcal{O}_n$) and in particular cyclic (namely $\bar{\sigma}_{\#}\pi = \pi$ for $\bar{\sigma}$ defined in (1.8)). We prove that there exists a map $T_2 : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $T_{2\#}\mu = \mu$, $T_2^{(n)}(x) = x$ for every $x \in \mathbb{R}^d$, and

$$\mathcal{C} \left((Id, T_2, T_2^{(2)}, \dots, T_2^{(n-1)})_{\#}\mu \right) \leq \mathcal{C}(\pi) + \varepsilon \|\pi\|. \quad (3.13)$$

We consider a partition \mathcal{F} as in Lemma 3.1. For every $Q \in \mathcal{F}$ we consider $A_{Q,1}$ constructed in Lemma 3.2. Then we consider $\bar{\sigma}$ defined as in (1.8) and we define for every $i = 2, \dots, n$

$$A_{Q,i} = A_{\bar{\sigma}^i(Q),1}. \quad (3.14)$$

Now, clearly \mathcal{S}^k is countable for every k and so we can fix a total order \leq on it. For every $i = 1, \dots, n$ let $\mathcal{F}_i \subseteq \mathcal{F}$ be the set

$$\mathcal{F}_i = \{Q \in \mathcal{F} : e_i(Q) \leq e_j(Q) \quad \forall j \neq i\}.$$

It can be easily seen that

$$\mathcal{F}_i = \{\bar{\sigma}^i(Q) : Q \in \mathcal{F}_1\}.$$

Since every $Q \in \mathcal{F}$ is contained in $\{c < \infty\}$ and thanks to the structure of this set (in particular we remark that $c(x_1, \dots, x_n) = \infty$ whenever $x_i = x_j$ for some $i \neq j$) we have that

$$\bigcup_{i=1}^n \mathcal{F}_i = \mathcal{F}. \quad (3.15)$$

Moreover thanks to the properties of our cost function $\bar{\sigma}^i(Q) \cap Q = \emptyset$ for every $i = 1, \dots, n-1$; using also (3.5) we have that

$$\mu(A_{Q,i} \cap A_{Q',j}) = 0 \quad \forall Q, Q' \in \mathcal{F}_1, i, j \in \{1, \dots, n-1\}, \text{ such that } (Q, i) \neq (Q', j). \quad (3.16)$$

Thanks to (3.16), (3.14), (3.15), and (3.4) we have that

$$\begin{aligned} \mu \left(\bigcup_{Q \in \mathcal{F}_1} \bigcup_{i=1}^n A_{Q,i} \right) &= \sum_{Q \in \mathcal{F}_1} \sum_{i=1}^n \mu(A_{Q,i}) = \sum_{i=1}^n \sum_{Q \in \mathcal{F}_1} \mu(A_{\bar{\sigma}^i(Q),1}) \\ &= \sum_{i=1}^n \sum_{Q \in \mathcal{F}_i} \mu(A_{Q,1}) = \sum_{Q \in \mathcal{F}} \mu(A_{Q,1}) = \sum_{Q \in \mathcal{F}} \pi(Q). \end{aligned} \quad (3.17)$$

We consider $Q \in \mathcal{F}_1$ and we want to define maps T_1, \dots, T_n on $A_{Q,1} \cup \dots \cup A_{Q,n}$. First we define them on $A_{Q,1}$ and then $A_{Q,2} \cup \dots \cup A_{Q,n}$ reminding (3.12). For every $x \in A_{Q,1}$, we consider $(T_1(x), \dots, T_n(x)) : A_{Q,1} \rightarrow A_{Q,1} \times \dots \times A_{Q,n}$ so that

$$\begin{cases} T_1(x) = x \\ T_{i\#}(1_{A_{Q,1}}\mu) = 1_{A_{Q,i}}\mu, \quad T_i \text{ is } \mu\text{-a.e. invertible on } A_{Q,1} \text{ for every } i = 2, \dots, n; \end{cases} \quad (3.18)$$

these maps exist thanks to Lemma 2.2. With an abuse of notation we set T_i^{-1} the almost everywhere inverse of T_i given by Lemma 2.2.

For every $i = 1, \dots, n$ we define $(T_1(x), T_2(x), \dots, T_n(x)) : A_{Q,i} \rightarrow A_{Q,i} \times \dots \times A_{Q,n} \times A_{Q,1} \times \dots \times A_{Q,i-1}$ so that $T_1(x) = x$ and $T_j(x) = T_{i+j-1} \circ T_i^{-1}(x)$, where the indices have to be intended modulo n (it is easy to check that these functions are well defined on $A_{Q,i}$).

We repeat the construction for every $Q \in \mathcal{F}_1$. Thanks to (3.16) and (3.17) the maps T_1, \dots, T_n are uniquely defined at μ -almost every point. It can be easily checked by induction that $T_j(x) = T_2^{(j)}(x)$ and that $x = T_2^{(n)}(x)$ for almost every $x \in \mathbb{R}^d$. Indeed for $i = 1$ the statement is trivial. We assume the statement for $j - 1$ and we prove it for j . For every $x \in A_{Q,i}$ (we remark that all the indices have to be intended modulo n) we have that $T_2(x) \in A_{Q,i+1}$; hence

$$T_2^{(j)}(x) = T_2^{(j-1)}(T_2(x)) = T_j(T_2(x)) = T_{i+j-1} \circ T_{i+1}^{-1} \circ T_{i+1} \circ T_i^{-1}(x) = T_j(x).$$

We prove that T_2 is a transport map between μ and μ , i.e. $T_{2\#}\mu = \mu$. By definition,

$$T_{2\#}(1_{A_{Q,i}}\mu) = T_{i+1\#} \left[(T_i^{-1})_{\#}(1_{A_{Q,i}}\mu) \right] = 1_{A_{Q,i+1}}\mu.$$

Summing up over $i = 1, \dots, n$ and $Q \in \mathcal{F}_1$, since the sets $A_{Q,i}$ are essentially disjoint and cover $(\mathbb{R}^d)^n$ in measure, we obtain that (the indices are to be intended modulo n)

$$T_{2\#}\mu = T_{2\#} \left(\sum_{Q \in \mathcal{F}_1} \sum_{i=1}^n 1_{A_{Q,i}}\mu \right) = \sum_{Q \in \mathcal{F}_1} \sum_{i=1}^n 1_{A_{Q,i+1}}\mu = \mu.$$

We are left to prove 3.13. We apply Lemma 2.3 to $(T_1 \times T_2 \times \dots \times T_n)_{\#}(1_{A_{Q,i}}\mu)$ and $1_{\bar{\sigma}^i(Q)}\pi$. The first plan is concentrated on $A_{Q,i} \times \dots \times A_{Q,n} \times A_{Q,1} \times \dots \times A_{Q,i-1}$, which is contained in $\bar{\sigma}^i(Q)$ by (3.4), and the second is concentrated on $\bar{\sigma}^i(Q)$. Thanks to (3.2) we have that

$$\begin{aligned} & \mathcal{C} \left((Id, T_2, T_2^{(2)}, \dots, T_2^{(n-1)})_{\#}\mu \right) - \mathcal{C}(\pi) = \mathcal{C}(T_1 \times T_2 \times \dots \times T_n)_{\#}\mu - \mathcal{C}(\pi) \\ &= \sum_{Q \in \mathcal{F}_1} \sum_{i=1}^n \int_{A_{Q,i}} c(x, T_2(x), \dots, T_n(x)) dx - \sum_{Q \in \mathcal{F}_1} \sum_{i=1}^n \int_{\bar{\sigma}^i(Q)} c(x_1, x_2, \dots, x_n) d\pi(x_1, \dots, x_n) \\ &\leq \sum_{Q \in \mathcal{F}_1} \sum_{i=1}^n \pi(\bar{\sigma}^i(Q))\varepsilon = \sum_{Q \in \mathcal{F}} \pi(Q)\varepsilon = \varepsilon, \end{aligned} \tag{3.19}$$

which proves (3.13). □

4 A generalization to the multimarginal problem with unbounded continuous costs on metric spaces

To perform a partition of X^n and of X as in Lemmas 3.1 and 3.2 we need a nested structure on X , playing the same role as the decomposition of \mathbb{R}^d introduced before. Following the same ideas in [2] and [4] we perform a dyadic decomposition.

Lemma 4.1. *Let $B \subset X$ be a Borel subset of the Polish space (X, d) . Given $\delta > 0$ there exists a partition \mathcal{A}_B^δ of B made of disjoint sets of diameter at most δ , with*

$$\bigcup_{C \in \mathcal{A}_B^\delta} C = B$$

and for every $C \in \mathcal{A}_B^\delta$

$$d(x, y) \leq \delta \quad \forall x, y \in C.$$

Proof. Let us fix a countable dense set $D = \{x_i\}_{i \in \mathbb{N}} \subseteq X$. We define, inductively, $B_0 = B$ and for every $n \geq 1$

$$C_n = \{x \in B_{n-1} : d(x_n, x) \leq \delta/2\},$$

$$B_n = B_{n-1} \setminus C_n.$$

If for some k we have $B_k = \emptyset$ we end the construction.

We claim that the set $\{C_i\}_{i \in \mathbb{N}}$ satisfies the assumptions: it is clear that, by construction, for all natural numbers n the set C_n is contained in B and furthermore for every $x, y \in C_n$ we have

$$d(x, y) \leq d(x, x_n) + d(y, x_n) \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

It's also clear that the C_i are disjoint. We show that their union is the whole B . In fact, taking any point $x \in B$ we know that, by density, there exists a point x_i in D such that $d(x_i, x) \leq \delta/2$; now if $x \in B_{i-1}$ then x will belong to C_i otherwise if $x \notin B_{i-1}$ it means it belongs to some other C_j , with $j < i$. \square

In analogy with the previous situation, we define the basic cells of our decomposition. Let $0 < \eta < 1$ be a fixed parameter (in Section 3 we took $\eta = 1/2$). Then we define $\mathcal{S}^1 = \mathcal{A}_X^\eta$, and then for every $k \geq 2$ we inductively divide the space in cells whose diameter is less than η^k

$$\mathcal{S}^k(X) = \bigcup_{C \in \mathcal{S}^{k-1}} \mathcal{A}_C^{\eta^k}.$$

Finally we define

$$\mathcal{Q}^k(X^n) = \{C_1 \times \cdots \times C_n : C_1, \dots, C_n \in \mathcal{S}^k\},$$

$$\mathcal{Q}(X^n) = \bigcup_{k=1}^{\infty} \mathcal{Q}^k(X^n).$$

So $\mathcal{Q}(X^n)$ will be our family of cubes and $\mathcal{Q}^k(X^n)$ will be the k -th generation.

Therefore Lemmas 3.1 and 3.2 can be repeated verbatim in this context.

The proof of Theorem 1.4 follows the same lines as the one of Theorem 1.1; the only thing which should be taken into account is that in this case we cannot say any more that each cube has all different faces (which was obtained by the fact that $c(x_1, \dots, x_n) = \infty$ whenever $x_i = x_j$ for some $i \neq j$), but when two faces coincide we can define the map between these faces as the identity map. More precisely, in the definition of the maps T_1, \dots, T_n on $A_{Q,1}$ given in (3.18), we put $T_i(x) = T_j(x)$ if $A_{Q,i} = A_{Q,j}$ and $T_i(x) = x$ if $A_{Q,1} = A_{Q,i}$.

References

- [1] L. AMBROSIO, *Lecture notes on optimal transport problems in “Mathematical aspects of Evolving Interfaces”*, Lecture notes in mathematics, 1812 (2003), 1–52.
- [2] L. AMBROSIO, M. COLOMBO, S. DI MARINO, *Sobolev spaces in metric measure spaces: reflexivity and lower semicontinuity of slope*, preprint.
- [3] G. BUTTAZZO, L. DE PASCALE, P. GORI-GIORGI, *Optimal transport formulation of electronic density-functional theory*, Phys. Rev. A, 85 (2012), 062502.
- [4] M. CHRIST, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math., 60/61(1990), 601–628.
- [5] M. COLOMBO, S. DI MARINO, L. DE PASCALE, *Existence of multimarginal optimal transport maps in the 1-dimensional case with a repulsive cost*, paper in preparation.
- [6] C. COTAR, G. FRIESECKE, C. KLÜPPELBERG, *Density functional theory and optimal transportation with Coulomb cost*, preprint, <http://arxiv.org/abs/1104.0603>.
- [7] W. GANGBO, A. SWIECH, *Optimal maps for the multidimensional Monge-Kantorovich problem*, Commun. Pure Appl. Math., 51 (1998), 23–45.
- [8] N. GHOUSSOUB, A. MOAMENI, *Symmetric Monge-Kantorovich problems and polar decompositions of vector fields*, preprint.
- [9] N. GHOUSSOUB, B. MAUREY, *Remarks on multidimensional symmetric Monge-Kantorovich problems*, Discr. Cont. Dyn. Sys. A, special issue on Optimal Transport and Applications (2012).
- [10] A. A. LYAPUNOV, *Sur les fonctions-vecteurs completement additives*, Izv. Akad. Nauk SSSR Ser. Mat., 4(1940), 465-478 (in Russian).
- [11] B. PASS, *Uniqueness and Monge solutions in the multimarginal optimal transportation problem*, SIAM J. Math. Anal., 43 (2011), 2758–2775.
- [12] A. PRATELLI, *On the equality between Monge’s infimum and Kantorovich’s minimum in optimal mass transportation*, Ann. Inst. H. Poincaré’ Probab. Statist., 43 (2007), 1–13.
- [13] S. RACHEV, L. RÜSCHENDORF, *Mass Transportation Problems*, Springer-Verlag, New York (1998).
- [14] M. SEIDL, P. GORI-GIORGI, A. SAVIN, *Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities*. Phys. Rev. A, 75 (2007), 042511.
- [15] W. SIERPINSKI, *Sur les fonctions d’ensemble additives et continues*, Fund. Math., 3 (1922), 240–246.
- [16] C. VILLANI, *Optimal transport. Old and new*, vol. 338 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.