European Journal of Applied Mathematics (2025), **36**, pp. 68–81
doi:10.1017/S0956792524000202

PAPER

Sufficiency of *c*-cyclical monotonicity in a class of multi-marginal optimal transport problems

Luigi De Pascale and Anna Kausamo

Dipartimento di Matematica e Informatica, Università di Firenze, Firenze, Italy
Corresponding author: Luigi De Pascale; Email: luigi.depascale@unifi.it

Received: 16 March 2023; **Revised:** 20 March 2024; **Accepted:** 24 March 2024; **First published online** 14 May 2024

Keywords: Monge–Kantorovich problem; Optimality condition; c-cyclical monotonicity; ∞-cyclical monotonicity; Kantorovich potentials

2020 Mathematics Subject Classification: 49J45, 49N15 (Primary); 49K21, 49K27, 49Q22 (Secondary)

Abstract
c-cyclical monotonicity is the most important optimality condition for an optimal transport plan. While the proof of necessity is relatively easy, the proof of sufficiency is often more difficult or even elusive. We present here a new approach, and we show how known results are derived in this new framework and how this approach allows to prove sufficiency in situations previously not treatable.

1. Introduction

1.1 Optimal transport problems

Consider $N \geq 2$ Polish probability spaces $(X^1, \mu^1), \dots, (X^N, \mu^N)$ and $X = \prod_{i=1}^N X^i$. Let $c : X \to [0, \infty]$ and consider the *Multi-marginal optimal transport* problems

$$\min_{\gamma\in\Pi(\mu^1,\dots,\mu^N)} C[\gamma] := \min_{\gamma\in\Pi(\mu^1,\dots,\mu^N)} \int_X c d\gamma, \tag{P}$$

and

$$\min_{\gamma\in\Pi(\mu^1,\dots,\mu^N)} C_\infty[\gamma] := \min_{\gamma\in\Pi(\mu^1,\dots,\mu^N)} \gamma - \operatorname*{ess\,sup}_{(x^1,\dots,x^N)\in X} c, \tag{P$_\infty$}$$

in the set

$$\Pi(\mu^1, \dots, \mu^N) := \{\gamma \in \mathcal{P}(X) \mid \pi^i(\gamma) = \mu^i \text{ for all } i = 1, \dots, N\},$$

that is, in the set of *couplings* or *transport plans* between the N marginals $\mu^1, \dots, \mu^N$. We refer to the second problem as *the* sup *case*. The first of these problems is widely encountered in the literature of the last thirty years. The second, although also old, gained popularity only more recently thanks to the applications of optimal transportation in machine learning (see, for example [17]).

Some general results on the multi-marginal optimal transport problem are available in refs. [10, 28, 30–32], and results for special costs are available, for example in ref. [23] for the quadratic cost with some generalisations in ref. [27] and in ref. [12] for the determinant cost. More applications appeared in ref. [24]. Applications to economics of the multi-marginal optimal transportation problems include, for example, the problem of team-matching which is a generalisation of the classical marriage problem [11, 13]. Applications to physics are related to quantum chemistry and the strong interacting regime for particles which are described in refs. [33–35]. By now, there are several papers on the transport theory for the Coulomb cost and some more general repulsive costs, a selection is given by [8, 14–16, 20, 21].

This paper is concerned with an optimality condition for the problems above, introduced in the next subsection. In particular, we will study the sufficiency of such optimality condition. We will give a new, easier, and in our opinion easier-to-understand proof of some known results, and we will show that this new approach allows to extend sufficiency results to a wider setting.

1.2 c-cyclical monotonicity, ∞-c-cyclical monotonicity and the main theorem

In this context, the c-cyclical monotonicity takes the following form.

Definition 1.1. We say that a set $\Gamma \subset \prod_{i=1}^N X^i$ is c-cyclically monotone (CM), if for every k-tuple of points $(x^{1,j}, \ldots, x^{N,j})_{j=1}^k$ and every $(N-1)$-tuple of permutations $(\sigma^2, \ldots, \sigma^N)$ of the set $\{1, \ldots, k\}$ we have

$$\sum_{j=1}^k c(x^{1,j}, x^{2,j}, \ldots, x^{N,j}) \le \sum_{j=1}^k c(x^{1,j}, x^{2,\sigma^2(j)}, \ldots, x^{N,\sigma^N(j)}) \, .$$

We also say that $\gamma \in \Pi(\mu^1, \ldots, \mu^N)$ is c-cyclically monotone if it is concentrated on a c-cyclically monotone set.

Definition 1.2. We say that a set $\Gamma \subset \prod_{i=1}^N X^i$ is infinitely c-cyclically monotone (ICM), if for every k-tuple of points $(x^{1,j}, \ldots, x^{N,j})_{j=1}^k$ and every $(N-1)$-tuple of permutations $(\sigma^2, \ldots, \sigma^N)$ of the set $\{1, \ldots, k\}$ we have

$$\max\{c(x^{1,j}, x^{2,j}, \ldots, x^{N,j}) \mid j \in \{1, \ldots, k\}\} \le \max\{c(x^{1,j}, x^{2,\sigma^2(j)}, \ldots, x^{N,\sigma^N(j)}) \mid j \in \{1, \ldots, k\}\} \, .$$

We also say that a coupling $\gamma \in \Pi(\mu^1, \ldots, \mu^N)$ is infinitely cyclically monotone if it is concentrated on an ICM set.

We will use the expression c-cyclically monotone for both conditions above. The main theorem of this paper is the following

Theorem 1.3. *Let $\mu^i \in \mathcal{P}(X^i)$ with compact support for $i = 1, \ldots, N$, let $c : X \to \mathbb{R} \cup \{+\infty\}$ be continuous. If $\gamma \in \Pi(\mu^1, \ldots, \mu^N)$ is an ICM plan for c, then γ is optimal.*

Along the way, we will give a new proof of the following known result due, in a more general setting, to Griessler [25].

Theorem 1.4. *Let $\mu^i \in \mathcal{P}(X^i)$ with compact support for $i = 1, \ldots, N$, let $c : X \to \mathbb{R}$ be continuous. If $\gamma \in \Pi(\mu^1, \ldots, \mu^N)$ is a CM plan for c, then γ is optimal.*

We now discuss an important characterisation of c-cyclically monotone transport plans. With this aim we define, still according to Griessler [25],

Definition 1.5. Let γ be a positive and finite Borel measure on X. We say that γ is finitely optimal if all its finitely supported submeasures are optimal with respect to their marginals. Here by submeasure we mean any probability measure α satisfying $\operatorname{supp}(\alpha) \subset \operatorname{supp}(\gamma)$.

Proposition 1.6. *If $\gamma \in \Pi(\mu^1, \ldots, \mu^N)$ is CM or ICM, then it is finitely optimal for the problem (P) or (P_∞), respectively.*

Lemma 1.7. *Let $\alpha = \sum_{i=1}^l m^i \delta_{(x^{1,i}, \ldots, x^{N,i})}$ and $\overline{\alpha} = \sum_{i=1}^{\bar{l}} \overline{m}^i \delta_{(\overline{x}^{1,i}, \ldots, \overline{x}^{N,i})}$ be two discrete measures with positive, integer coefficients and the same marginals. Let us denote by $\tilde{l} = m^1 + \ldots + m^l$ the number of rows of the following table*

https://doi.org/10.1017/S0956792524000202 Published online by Cambridge University Press

$$\begin{array}{ccc} x^{1,1} & \dots & x^{N,1} \\ \vdots & & \vdots \\ x^{1,1} & \dots & x^{N,1} \end{array} \left.\vphantom{\begin{array}{c} x \\ \vdots \\ x \end{array}}\right\} m^1\text{- times}$$
$$\begin{array}{ccc} \dots & \dots & \dots \end{array}$$
$$\begin{array}{ccc} x^{1,l} & \dots & x^{N,l} \\ \vdots & & \vdots \\ x^{1,l} & \dots & x^{N,l} \end{array} \left.\vphantom{\begin{array}{c} x \\ \vdots \\ x \end{array}}\right\} m^l\text{- times}$$

where the first m^1 rows are equal among themselves, the following m^2 rows are equal among themselves and so on. Let $\overline{A}$ be the analogous table associated with $\overline{\alpha}$. Then, $\overline{A}$ has $\tilde{l}$ rows, and there exist $(N-1)$ permutations of the set $\{1, \dots, \tilde{l}\}$ such that $\overline{A}$ is equal to

$$\begin{array}{ccc} x^{1,1} & \dots & x^{N,\sigma^N(1)} \\ \vdots & & \vdots \\ x^{1,1} & \dots & x^{N,\sigma^N(m_1)} \\ x^{1,2} & \dots & x^{N,\sigma^N(m_1+1)} \\ \vdots & & \vdots \\ x^{1,l} & \dots & x^{N,\sigma^N(m_1+\ldots+m_{l-1}+1)} \\ \vdots & & \vdots \\ x^{1,l} & \dots & x^{N,\sigma^N(\tilde{l})} \end{array}$$

Proof. For each $k \in \{1, \dots, N\}$, the k-th marginal of α is given by the sum of the Dirac masses centred on the points of the k-th column of the table A with multiplicity. Analogously, the k-th marginal of $\overline{\alpha}$ is given by the sum of the Dirac masses centred on the points of the i-th column of the table $\overline{A}$ with multiplicity. Since the marginals of α and $\overline{\alpha}$ are the same, each point $x^{k,i}$ appearing in the k-th marginal must appear in both matrices the same number of times, proving the existence of the bijections $\sigma^2, \dots, \sigma^N$ as required. This also implies that $\overline{A}$ has $\tilde{l}$ rows. □

Proof. (of Proposition 1.6) We fix a finitely supported submeasure $\alpha = \sum_{i=1}^{l} a_i \delta_{X^i}$ of γ. We need to show that α is an optimal coupling of its marginals. To do this, we fix another coupling, $\overline{\alpha} = \sum_{i=1}^{\tilde{l}} \overline{a}_i \delta_{\overline{X}^i}$, with the same marginals as α. We have to show that

$$\tilde{C}[\alpha] \leq \tilde{C}[\overline{\alpha}], \tag{1}$$

where $\tilde{C}$ is any of the two costs under consideration. Let us first assume that the discrete measures α and $\overline{\alpha}$ have rational coefficients. We consider the measures $M\alpha$ and $M\overline{\alpha}$, where M is the product of the denominators of the coefficients of α and $\overline{\alpha}$. They are discrete measures having positive, integer coefficients and the same marginals, so we can apply Lemma 1.7 to find permutations $\sigma^2, \dots, \sigma^N$ such that $M\alpha$ and $M\overline{\alpha}$ have representations A and $\overline{A}$, respectively. If $\tilde{C} = C$ we have using the c-cyclical monotonicity of α

$$MC[\alpha] = \sum_{i=1}^{\tilde{l}} c(x^{1,i}, \dots, x^{N,i}) \leq \sum_{i=1}^{\tilde{l}} c(x^{1,i}, x^{2,\sigma^2(i)}, \dots, x^{N,\sigma^N(i)}) = MC[\overline{\alpha}],$$

proving the optimality of α. If $\tilde{C} = C_\infty$, the conclusion is immediate:

$$C_\infty[\alpha] = \max_{1 \leq i \leq \tilde{k}} c(x^{1,i}, \dots, x^{N,i}) \leq \max_{1 \leq i \leq \tilde{k}} c(x^{1,i}, x^{2,\sigma^2(i)}, \dots, x^{N,\sigma^N(i)}) = C_\infty[\overline{\alpha}].$$

Now, assume that α and $\overline{\alpha}$ have real (not necessarily rational) coefficients,

$$\alpha := \sum_{i=1}^{l} a_i \delta_{X^i}, \ \overline{\alpha} = \sum_{i=1}^{\bar{l}} \overline{a}_i \delta_{\overline{X}^i}.$$

We show that for all $\varepsilon > 0$ there exist two discrete measures

$$\beta := \sum_{i=1}^{l} q_i \delta_{X^i} \text{ and } \overline{\beta} = \sum_{i=1}^{\bar{l}} \overline{q}_i \delta_{\overline{X}^i},$$

with the same marginals, $q_i, \overline{q}_i \in \mathbb{Q}$ and

$$|a_i - q_i| < \varepsilon, \ |\overline{a}_i - \overline{q}_i| < \varepsilon.$$

Being concentrated on $X^1, \ldots, X^l$ and $\overline{X}^1, \ldots \overline{X}^{\bar{l}}$ is equivalent to the fact that the vector $\underline{\mathbf{a}} := (a_1, \ldots, a_l, \overline{a}_1, \ldots, \overline{a}_{\bar{l}})$ is a solution of

$$\mathcal{A}\underline{\mathbf{a}} = 0,$$

where $\mathcal{A}$ is a matrix with coefficients $1, 0, -1$. Indeed, if we write, for example, the equality between the first two marginals we obtain

$$\sum_{i=1}^{l} a_i \delta_{x^{1,i}} = \sum_{i=1}^{\bar{l}} \overline{a}_i \delta_{\overline{x}^{1,i}}.$$

so some of the points $\overline{x}^{1,i}$ must coincide with, for example, $x^{1,1}$ and this gives, for two sets of indices

$$\sum_{i \in I} a_i = \sum_{j \in J} \overline{a}_j.$$

Since the matrix $\mathcal{A}$ has integer coefficients

$$\overline{Ker_{\mathbb{Q}} \mathcal{A}} = Ker_{\mathbb{R}} \mathcal{A},$$

and this allows to choose β and $\overline{\beta}$. Since $C[\alpha] \approx C[\beta]$, $C[\overline{\alpha}] \approx C[\overline{\beta}]$, $C_\infty[\alpha] = C_\infty[\beta]$ and $C_\infty[\overline{\alpha}] = C[\overline{\beta}]$, we conclude. □

2. Essential background and preliminary results

The c-cyclical monotonicity is the most important optimality condition for a transport plan. Giving a satisfactory historical background requires a paper on his own and we refer the reader to the survey [19]. Originally born in convex analysis as characterisation of sub-differential of convex functions, for $N = 2$ it first appeared as optimality condition in ref. [29] in an equivalent formulation of the Kantorovich's problem. In that context, which was partly motivated by some models appearing in financial mathematics, the authors started by characterising optimal random variables using c-cyclical monotonicity.

For the quadratic cost $c(x, y) = |x - y|^2$ in $\mathbb{R}^d$, the necessity of the condition is a basic result. See, for example, Prop. 2.24 of [36]. The classical structure of cyclical monotonicity of optimal plans was mentioned as a possible alternative tool in ref. [7] and explicitly exploited in ref. [9]. After that, in ref. [22], the authors extended the result to lower semi-continuous cost functions bounded from below. They showed that every finite optimal plan with respect to such a cost lies on a c-cyclically monotone set.

For more general settings there are, essentially, two arguments to prove that the support of the optimal plan must be c-cyclically monotone. The first one uses duality and appears in ref. [29], while the second one relies on modifying a transport plan that is not c-cyclically monotone and showing that its cost can be improved. The latter technique was introduced in ref. [1] and used, for example, in Proposition 2.24 of [36]. Both approaches can be extended to the multi-marginal case with few technical modifications.

To the best of our knowledge, for $N = 2$ the most general result was proved in ref. [4] who removed regularity assumptions on the cost proving that: if X, Y are Polish spaces equipped with Borel probability measures μ, ν and $c : X \times Y \to [0, \infty]$ is a Borel measurable cost function, then every optimal transport plan with finite total cost is c-cyclically monotone.

Concerning the sufficiency of the condition, we reported above Th. 1.4 which seems to be the most general available in the case $N > 2$. Much more is known for $N = 2$, and we will comment on this at the end of the paper.

2.1 Lower semi-continuity, compactness and existence of minimisers

Existence f the optimal transport problems above is usually obtained by the direct method of Calculus of Variations. Here, we shortly report the tools which we do not find elsewhere or that will be used substantially in our proofs. A useful convergence on the set of transport plans is the tight convergence.

Definition 2.1. Let X be a metric space and let $\gamma_n \in \mathcal{P}(X)$ we say that γ_n converges tightly to γ if for all $\phi \in C_b(X)$

$$\int \phi d\gamma_n \to \int \phi d\gamma.$$

The tight convergence will be denoted by $\overset{*}{\rightharpoonup}$.

Definition 2.2. Let Π be a set of Borel probability measures on a metric space X. We say that Π is tight (or uniformly tight) if for all $\varepsilon > 0$ there exists $K_\varepsilon \subset X$ compact such that

$$\gamma(K_\varepsilon) > 1 - \varepsilon \text{ or, equivalently, } \gamma(X \setminus K_\varepsilon) \leq \varepsilon$$

for all $\gamma \in \Pi$.

Theorem 2.3 (Prokhorov). *Let X be a complete and separable metric space (Polish space). Then, $\Pi \subset \mathcal{P}(X)$ is tight if and only if it is pre-compact with respect to the tight convergence.*

Remark 2.4.

1. The tight convergence is lower semi-continuous on open sets and upper semi-continuous on closed sets;
2. If X is complete and separable, then if Π is a singleton it is always tight.

The following compactness theorem will be used in this paper.

Theorem 2.5. *For $i = 1, \ldots, N$, let X^i be a Polish space. Let $X = X^1 \times \ldots \times X^N$. Let $\mathcal{M}^i \subset \mathcal{P}(X^i)$ be tight for all i. Then, the set*

$$\Pi = \{\gamma \in \mathcal{P}(X) \mid \pi^i_\sharp \gamma \in \mathcal{M}^i\}$$

is tight.

Proof. Let $\varepsilon > 0$. By the tightness of $\mathcal{M}^i$, we can fix a compact set $K^i \subset X^i$ such that for all $\mu^i \in \mathcal{M}^i$

$$\mu^i(X^i \setminus K^i) < \frac{\varepsilon}{N}.$$

Let $K = K^1 \times \ldots \times K^N$ and let $\gamma \in \Pi$. Since all the marginals $\pi^i_\sharp \gamma \in \mathcal{M}^i$, and since

$$X \setminus K \subset \left((X^1 \setminus K^1) \times \prod_{k=2}^{N} X^k\right) \cup \left(X^1 \times (X^2 \setminus K^2) \times \prod_{k=3}^{N} X^k\right) \cup \cdots \cup \left(\prod_{k=1}^{N-1} X^k \times (X^N \setminus K^N)\right),$$

one gets

$$\gamma(X \setminus K) \leq \varepsilon.$$

□

https://doi.org/10.1017/S0956792524000202 Published online by Cambridge University Press

Corollary 2.6. *By Prokhorov's theorem, a set $\Pi \subset \mathcal{P}(X)$ as in the theorem above is pre-compact for the tight convergence. This is, in particular, true if $\mathcal{M}^i = \{\mu^i\}$.*

If c is lower semi-continuous, then also the functionals C and C_∞ (see problems (P) and (P_∞) on the first page for the definition) are lower semi-continuous with respect to the tight convergence of measures. The lower semi-continuity of C is a standard result of optimal transport theory (see, e.g. [37] or [28] for the multi-marginal case). The next lemma proves the lower semi-continuity of C_∞.

Lemma 2.7. *If the function $c : X \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous, then also the functional C_∞ is lower semi-continuous.*

Proof. First, we note that, thanks to the lower semi-continuity of c, its γ-essential supremum can be written as

$$\gamma - \text{ess sup }\ c = \sup\{c(x^1, \ldots, x^N) \mid (x^1, \ldots, x^N) \in \text{supp}\ \gamma\}.$$

Fix $\gamma \in \Pi(\mu^1, \ldots, \mu^N)$ and let $(\gamma^n)_n$ be a sequence converging to γ. Now there exist a vector $v \in \text{supp}\ \gamma$ and a sequence $v^n = (x^{1,n}, \ldots, x^{N,n}) \in \prod_{i=1}^N X^i$ such that $v^n \in \text{supp}\ \gamma^n$ for all n and $v^n \to v$. Moreover,

$$\liminf_{n\to\infty} C_\infty[\gamma^n] \geq \liminf_{n\to\infty} c(v^n) \geq c(v)\,.$$

Since the above inequality holds for all $v \in \text{supp}\ \gamma$ and for all sequences converging to v, it also holds for the γ-essential supremum, and the claim follows. □

The use of compactness and semi-continuity theorems above gives the existence of optimal transport plans for both problems considered here.

2.2 Γ-convergence

A crucial tool that we will use in this paper is Γ-convergence. All the details can be found, for instance, in Braides's book [6] or in the classical book by Dal Maso [18]. In what follows, (X, d) is a metric space or a topological space equipped with a convergence.

Definition 2.8. Let $(F_n)_n$ be a sequence of functions $X \mapsto \bar{\mathbb{R}}$. We say that $(F_n)_n$ Γ-converges to F if for any $x \in X$ we have

- for any sequence $(x^n)_n$ of X converging to x

$$\liminf_n F_n(x^n) \geq F(x) \qquad (\Gamma\text{-liminf inequality});$$

- there exists a sequence $(x^n)_n$ converging to x and such that

$$\limsup_n F_n(x^n) \leq F(x) \qquad (\Gamma\text{-limsup inequality}).$$

This definition is actually equivalent to the following equalities for any $x \in X$:

$$F(x) = \inf\left\{\liminf_n F_n(x^n) : x^n \to x\right\} = \inf\left\{\limsup_n F_n(x^n) : x^n \to x\right\}$$

The function $x \mapsto \inf\left\{\liminf_n F_n(x^n) : x^n \to x\right\}$ is called Γ-liminf of the sequence $(F_n)_n$ and the other one its Γ-limsup. A useful result is the following (which for instance implies that a constant sequence of functions does not Γ-converge to itself in general).

Proposition 2.9. *The Γ-liminf and the Γ-limsup of a sequence of functions $(F_n)_n$ are both lower semi-continuous on X.*

The main interest of Γ-convergence resides in its consequences in terms of convergence of minima.

https://doi.org/10.1017/S0956792524000202 Published online by Cambridge University Press

Theorem 2.10. *Let $(F_n)_n$ be a sequence of functions $X \to \bar{\mathbb{R}}$ and assume that F_n Γ-converges to F. Assume moreover that there exists a compact and non-empty subset K of X such that*

$$\forall n \in \mathbb{N}, \ \inf_X F_n = \inf_K F_n$$

(we say that $(F_n)_n$ is equi-mildly coercive on X). Then, F admits a minimum on X and the sequence $(\inf_X F_n)_n$ converges to $\min F$. Moreover, if $(x_n)_n$ is a sequence of X such that

$$\lim_n F_n(x_n) = \lim_n (\inf_X F_n)$$

and if $(x_{\phi(n)})_n$ is a subsequence of $(x_n)_n$ having a limit x, then $F(x) = \inf_X F$.

3. Discretisation of transport plans (Dyadic-type decomposition in Polish spaces)

Let γ be a Borel probability measure on $X = (X^1, d_1) \times \cdots \times (X^N, d_N)$ with marginals $\mu^1, \ldots, \mu^N$. The space X will be equipped with the sup metric

$$d(w, z) = \max_{1 \le i \le N} d_i(w^i, z^i).$$

Let $\varepsilon_n = \frac{1}{n}$. Since $\{\mu^i\}_{i=1}^N$ are Borel probability measures, they are inner regular. Hence for all n, there exist compact sets $K^{1,n} \subset \operatorname{supp} \mu^1, K^{2,n} \subset \operatorname{supp} \mu^2, \ldots, K^{N,n} \subset \operatorname{supp} \mu^N$ such that

$$\mu^k(X^k \setminus K^{k,n}) < \frac{\varepsilon_n}{N}, \tag{2}$$

for all $k = 1, \ldots, N$. We may assume that, for all k and n, $K^{k,n} \subset K^{k,n+1}$.

We denote $K^n := \prod_{k=1}^N K^{k,n}$. Since

$$X \setminus K^n \subset \left((X^1 \setminus K^{1,n}) \times \prod_{k=2}^N X^k \right) \cup \left(X^1 \times (X^2 \setminus K^{2,n}) \times \prod_{k=3}^N X^k \right) \cup \cdots \cup \left(\prod_{k=1}^{N-1} X^k \times (X^N \setminus K^{N,n}) \right),$$

one gets

$$\gamma(X \setminus K^n) \le \varepsilon_n.$$

The cost c is uniformly continuous on each K^n, and for all n, we can fix $\delta_n \in (0, \varepsilon_n)$ such that the sequence (δ_n) is decreasing in n and

$$|c(u) - c(z)| < \varepsilon_n \quad \text{for all } u, z \in K^n \text{ for which } d(u, z) < \delta_n.$$

Next we fix, for all n, finite Borel partitions for the sets $K^{1,n}, \ldots, K^{N,n}$. We denote these by $\{\tilde{B}_i^{k,n}\}_{i=1}^{\tilde{m}^{k,n}}$, $k = 1, \ldots, N$, and we choose them in such a way that for all $n \in \mathbb{N}$ and $k \in \{1, \ldots, N\}$

$$diam(\tilde{B}_i^{k,n}) < \frac{1}{2}\delta_n,$$

for all $i \in \{1, \ldots, \tilde{m}^{k,n}\}$.

We form a new, possibly finer, partition $\{B_i^{k,n}\}_{i=1}^{m^{k,n}}$ for each $K^{k,n}$ by intersecting (if the intersection if non-empty) each element $\tilde{B}_i^{k,n}$ successively first with the set $K^{k,1}$, then with $K^{k,2}$, and so on up until intersecting with the set $K^{k,n-1}$. So that for $j \in \{1, \ldots n\}$ either $B_i^{k,n} \cap K^{k,j}$ is empty or it is the entire $B_i^{k,n}$. The products

$$\mathcal{Q}^n = \{B_{i_1}^{1,n} \times B_{i_2}^{2,n} \times \cdots \times B_{i_N}^{N,n}, \ i_k \in \{1, \ldots, m^{k,n}\} \text{ for all } k = 1, \ldots, N\}$$

form a partition of the set K^n with

$$diam(B_i^{k,n}) < \frac{1}{2}\delta_n,$$

for all $i \in \{1, \ldots, m^{k,n}\}$.

We denote

$$I^n = \{(i_1, \dots, i_N) \mid \gamma(B^{1,n}_{i_1} \times B^{2,n}_{i_2} \times \cdots \times B^{N,n}_{i_N}) > 0\},$$

and for all $\mathbf{i} := (i_1, \dots, i_N) \in I^n$, we use the notation $Q^n_{\mathbf{i}} := B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N}$. We fix points $z^n_{\mathbf{i}} = z^n_{i_1,\dots,i_N} \in \prod_{k=1}^N B^{k,N}_{i_k} \cap \operatorname{supp}\gamma$ (i.e. $z^n_{\mathbf{i}} \in Q^n_{\mathbf{i}} \cap \operatorname{supp}\gamma$). We define

$$\tilde{\alpha}^n = \sum_{(i_1,\dots,i_N)\in I^n} \gamma(B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N})\delta_{z^n_{i_1,\dots,i_N}} \quad \text{and} \quad \alpha^n = \frac{1}{\gamma(K^n)}\tilde{\alpha}^n; \tag{3}$$

since $\tilde{\alpha}^n(X) = \gamma(K^n)$, the measures α^n are probability measures.

To each multi-index $\mathbf{i} = (i_1, \dots, i_N)$ and thus to each point $z^n_{\mathbf{i}}$ correspond N points

$$x^{1,n}_{\mathbf{i}} \in B^{1,n}_{i_1}, \dots, x^{N,n}_{\mathbf{i}} \in B^{N,n}_{i_N},$$

which are 'coordinates' in the spaces X_i of $z^n_{\mathbf{i}}$. The marginals of α^n are supported by the Dirac measures given by these points. We denote these marginals by $\mu^{1,n}, \dots, \mu^{N,n}$. More precisely, they can be described as

$$\mu^{k,n} = \frac{1}{\gamma(K^n)} \sum_{i=1}^{m_{k,n}} \sum_{\substack{\mathbf{i}\in I^n \\ i=\mathbf{i}_k}} \gamma(Q_{\mathbf{i}})\delta_{x^{k,n}_{\mathbf{i}}}. \tag{4}$$

Proposition 3.1. $\alpha^n \rightharpoonup \gamma$.

Proof. Let $\varepsilon > 0$ and $\varphi \in C_b(X)$. We have to find $n_0 \in \mathbb{N}$ such that

$$\left| \int_X \varphi d\gamma - \int_X \varphi d\alpha^n \right| < \varepsilon, \quad \text{for all } n \geq n_0. \tag{5}$$

Let $M > 0$ be such that

$$|\varphi(z)| \leq M \quad \text{for all } z \in X.$$

We fix $\bar{n} \in \mathbb{N}$ such that

$$\gamma(X \setminus K^n) < \min\left\{\frac{1}{2}, \frac{\varepsilon}{5M}\right\}, \quad \text{for all } n \geq \bar{n}$$

Since $\varphi \in C_b(X)$, it is uniformly continuous on the set $K^{\bar{n}}$, there exists $\delta > 0$ such that

$$|\varphi(z) - \varphi(v)| < \frac{\varepsilon}{5} \quad \text{for all } z, v \in K^{\bar{n}} \text{ such that } d(z, v) < \delta.$$

Moreover, the decomposition $\mathcal{Q}^n$ has been constructed so that there exists $n_0 \geq \bar{n}$ such that for all $k \in \{1, \dots, N\}$ and $n \geq n_0$

$$diam(B^{k,n}_i) < \delta \quad \text{for all } i \in \{1, \dots, m^{k,n}\}.$$

We start from

$$\left| \int_X \varphi \, \mathrm{d}\gamma - \int_X \varphi \, \mathrm{d}\alpha^n \right| \leq \left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\alpha^n \right| + \left| \int_{X\setminus K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{X\setminus K^{\bar{n}}} \varphi \, \mathrm{d}\alpha^n \right| \tag{6}$$

https://doi.org/10.1017/S0956792524000202 Published online by Cambridge University Press

and we evaluate separately the two terms on the RHS. For all $n \geq n_0$, the first term can be estimated as follows: (we recall that, by construction, $K^{\bar{n}} \subset K^n$)

$$
\begin{aligned}
\left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\alpha^n \right| &= \left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \frac{1}{\gamma(K^n)} \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\tilde{\alpha}^n \right| \\
&\overset{a)}{\leq} \left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\tilde{\alpha}^n \right| + \frac{\gamma(X \setminus K^n)}{1 - \gamma(X \setminus K^n)} \int_{K^{\bar{n}}} |\varphi| \, \mathrm{d}\tilde{\alpha}^n \\
&< \left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\tilde{\alpha}^n \right| + M \cdot 2 \cdot \frac{\varepsilon}{5M} \\
&< \left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\tilde{\alpha}^n \right| + \frac{2\varepsilon}{5}.
\end{aligned} \tag{7}
$$

Above in a), we have written

$$
\frac{1}{\gamma(K^n)} = \frac{1}{1 - \gamma(X \setminus K^n)} = 1 + \frac{\gamma(X \setminus K^n)}{1 - \gamma(X \setminus K^n)}
$$

and then estimated the numerator from above by $\frac{\varepsilon}{5M}$ and the term $\gamma(X \setminus K^n)$ of the denominator from below by $\frac{1}{2}$. By construction, since $n \geq \bar{n}$, there exist a subset $\bar{I}^n \subset I^n$ such that

$$
K^{\bar{n}} = \bigcup_{(i_1,\ldots,i_N) \in \bar{I}^n} B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N}.
$$

So we write

$$
\int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\tilde{\alpha}^n = \sum_{\mathbf{i} \in \bar{I}^n} \left(\int_{(B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N})} \varphi \, \mathrm{d}\gamma - \int_{(B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N})} \varphi \, \mathrm{d}\tilde{\alpha}^n \right).
$$

We simplify the notations for the next few lines and, for all $\mathbf{i} \in \bar{I}^n$, we denote by $Q := B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N}$ and by $u_0 = z_{i_1,\ldots,i_N} \in Q$ the point in which $\tilde{\alpha}^n$ is concentrated. Then for each 'cube' Q

$$
\begin{aligned}
\left| \int_Q \varphi(u) \, \mathrm{d}\gamma - \int_Q \varphi(u) \, \mathrm{d}\tilde{\alpha}^n \right| &= \left| \int_Q \varphi(u) \, \mathrm{d}\gamma - \varphi(u_0)\gamma(Q) \right| \\
&\leq \int_Q |\varphi(u) - \varphi(u_0)| \, \mathrm{d}\gamma \leq \gamma(Q) \cdot \frac{\varepsilon}{5},
\end{aligned}
$$

and in the last passage, we have used the uniform continuity of φ on $K^{\bar{n}}$. Summing the estimate above over all cubes $Q = B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N}$, $\mathbf{i} \in \bar{I}^n$, gives

$$
\left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\tilde{\alpha}^n \right| < \gamma(K^{\bar{n}}) \cdot \frac{\varepsilon}{5} \leq \frac{\varepsilon}{5}.
$$

Combining this estimate with (7) gives us the estimate

$$
\left| \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\gamma - \int_{K^{\bar{n}}} \varphi \, \mathrm{d}\alpha^n \right| < \frac{1}{5}\varepsilon + \frac{2\varepsilon}{5} = \frac{3}{5}\varepsilon. \tag{8}
$$

Finally, the 'tail' term in (6). Using the set $\bar{I}^n$ defined above one gets

$$
\begin{aligned}
\alpha^n(X \setminus K^{\bar{n}}) &= 1 - \frac{1}{\gamma(K^n)} \sum_{(i_1,\ldots,i_N) \in \bar{I}^n} \gamma(B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N}) \\
&= 1 - \frac{\gamma(K^{\bar{n}})}{\gamma(K^n)} \leq 1 - \gamma(K^{\bar{n}}) < \frac{\varepsilon}{5M}.
\end{aligned}
$$

https://doi.org/10.1017/S0956792524000202 Published online by Cambridge University Press

Using this we get

$$\left| \int_{X\setminus K^{\bar{n}}} \varphi \, d\gamma - \int_{X\setminus K^{\bar{n}}} \varphi \, d\alpha^n \right| \le \int_{X\setminus K^{\bar{n}}} |\varphi| \, d\gamma + \int_{X\setminus K^{\bar{n}}} |\varphi| \, d\alpha^n$$
$$< M\frac{\varepsilon}{5M} + M\frac{\varepsilon}{5M} = \frac{2}{5}\varepsilon. \tag{9}$$

Together estimates (8) and (9) prove the claim (5). □

Remark 3.2. If supp μ^k is compact for $k = 1, \dots, N$, then the dependence on n of K^n is not needed anymore since one can take $K^n \equiv K :=$ supp $\mu^1 \times \dots \times$ supp μ^N. This also simplifies the analytic expressions of α^n and their marginal measures.

In line with the previous Remark, we prove the following:

Proposition 3.3. *If supp μ^k is compact for $k = 1, \dots, N$ then for all k, n and all i*

$$\mu^{k,n}(B_i^{k,n}) = \mu^k(B_i^{k,n}),$$

where, we recall, $\mu^{k,n}$ is defined in (4) above and is the k-th marginal of the discretisation α^n of γ defined in (3).

Proof. Again we prove the formula for the first marginal.

$$\alpha^n\left(B_i^{1,n} \times \prod_{k=2}^N X^k\right) = \sum_{\substack{\underline{\mathrm{i}} \in I^n \\ i = i_1}} \gamma(Q_{\underline{\mathrm{i}}}^n)\delta_{z_{\underline{\mathrm{i}}}^n}\left(B_i^{1,n} \times \prod_{k=2}^N X^k\right)$$
$$= \sum_{\substack{\underline{\mathrm{i}} \in I^n \\ i = i_1}} \gamma(Q_{\underline{\mathrm{i}}}^n) = \gamma\left(B_i^{1,n} \times \prod_{k=2}^N X^k\right).$$

□

4. Variational approximations and conclusions

In this section, we prove the discrete approximations of the functionals that will be used in the optimality proofs. Given a transport plan γ, we have introduced, in the previous section, the dyadic approximation $\{\alpha^n\}_{n\in\mathbb{N}}$ of γ.

4.1 The sup *case*

We define the functionals $\mathcal{F}_n, \mathcal{F} : \mathcal{P}(X) \to \mathbb{R} \cup \{+\infty\}$ by

$$\mathcal{F}_n(\beta) = \begin{cases} C_\infty[\beta] & \text{if } \beta \in \Pi(\mu^{1,n}, \dots, \mu^{N,n}), \\ +\infty & \text{otherwise;} \end{cases}$$

and

$$\mathcal{F}(\beta) = \begin{cases} C_\infty[\beta] & \text{if } \beta \in \Pi(\mu^{1}, \dots, \mu^{N}), \\ +\infty & \text{otherwise.} \end{cases}$$

For the rest of this subsection, we assume that c is continuous and that μ^i has compact support for $i = 1, \dots, N$. We prove the following

Proposition 4.1. *The functionals $\mathcal{F}_n$ are equi-coercive and*

$$\mathcal{F}_n \xrightarrow{\Gamma} \mathcal{F}. \tag{10}$$

Proof. Let $\beta \in \mathcal{P}(X)$. We recall that we need to prove the following:

$$\forall (\beta^n)_n \overset{*}{\rightharpoonup} \beta \text{ in } \mathcal{P}(X),\ \liminf_{n\to\infty} \mathcal{F}_n(\beta^n) \geq \mathcal{F}(\beta). \tag{I}$$

$$\exists (\beta^n)_n \overset{*}{\rightharpoonup} \beta \text{ in } \mathcal{P}(X) \text{ s.t. } \limsup_{n\to\infty} \mathcal{F}_n(\beta^n) \leq \mathcal{F}(\beta). \tag{II}$$

If $\mathcal{F}[\beta] < +\infty$, the Γ-lim inf inequality (Condition (I)) follows from the lower semi-continuity of the functional C_∞. If $\mathcal{F}[\beta] = +\infty$, then either $\beta \notin \Pi(\mu^1, \ldots, \mu^N)$ or $C_\infty(\beta) = +\infty$. In the first case, since $\beta^n \overset{*}{\rightharpoonup} \beta$ and $\mu^{i,n} \overset{*}{\rightharpoonup} \mu^i$ for $i = 1, \ldots, N$, there exists $n_0 \in \mathbb{N}$ such that $\beta^n \notin \Pi(\mu^{1,n}, \ldots, \mu^{N,n})$ for all $n \geq n_0$. Hence, $\mathcal{F}_n[\beta^n] = +\infty$ for all $n \geq n_0$. If $C_\infty(\beta) = +\infty$, then let $M > 0$ and let $\mathbf{x} \in \operatorname{spt} \beta$ and $r > 0$ be such that $B(\mathbf{x}, r) \subset \{c > M - \varepsilon\}$. Since the evaluation on open sets is lower semi-continuous with respect to the tight convergence, we have that, for n big enough, $\beta_n(B(\mathbf{x}, r)) > 0$ so that $C_\infty(\beta_n) > M - \varepsilon$ and since M is arbitrary we conclude.

For the Γ-lim sup inequality (Condition (II)), if $\mathcal{F}[\beta] = +\infty$, then any sequence with the right marginals and tightly converging to β will do. Therefore, we may assume that the measure β satisfies $\beta \in \Pi(\mu^1, \ldots, \mu^N)$ and $C_\infty[\beta] < +\infty$. To build the approximants, we use the Borel partitions $\{B_i^{k,n}\}_{i=1}^{m^{k,n}}$ and discrete measures introduced in Section 3. For all n, given a multi-index $\underline{\mathbf{i}} = (i_1, \ldots, i_N)$ we use, again, the 'cube'

$$Q^n_{\underline{\mathbf{i}}} := B^{1,n}_{i_1} \times \cdots \times B^{N,n}_{i_N}$$

and set

$$J^n := \{\underline{\mathbf{i}} \mid \beta(Q^n_{\underline{\mathbf{i}}}) > 0\}.$$

We then define the measures

$$\beta^n = \sum_{\underline{\mathbf{i}} \in J^n} \beta(Q^n_{\underline{\mathbf{i}}}) \frac{\mu^{1,n}_{|B^{1,n}_{i_1}}}{\mu^1(B^{1,n}_{i_1})} \otimes \ldots \otimes \frac{\mu^{N,n}_{|B^{N,n}_{i_N}}}{\mu^N(B^{N,n}_{i_N})}.$$

We show that β^n has marginals $\mu^{1,n}, \ldots, \mu^{N,n}$. For all Borel sets $A \subset X_1$, we have

$$\begin{aligned}
\beta^n\left(A \times \prod_{k=2}^N X_k\right) &= \sum_{\underline{\mathbf{j}} \in J^n} \beta(Q^n_{\underline{\mathbf{j}}}) \frac{\mu^{1,n}_{|B^{1,n}_{j_1}}(A)}{\mu^{1,n}(B^{1,n}_{j_1})} \\
&= \sum_{j_1 \in \pi^1(J^n)} \frac{\mu^{1,n}_{|B^{1,n}_{j_1}}(A)}{\mu^1(B^{1,n}_{j_1})} \sum_{\{(j_2,\ldots,j_N) \mid \underline{\mathbf{j}} \in J^n\}} \beta(Q^n_{\underline{\mathbf{j}}}) \\
&= \sum_{j_1 \in \pi^1(J^n)} \frac{\mu^{1,n}_{|B^{1,n}_{j_1}}(A)}{\mu^1(B^{1,n}_{j_1})} \mu^1(B^{1,n}_{j_1}) \\
&= \sum_{j_1 \in \pi^1(J^n)} \mu^{1,n}_{|B^{1,n}_{j_1}}(A) = \mu^{1,n}(A).
\end{aligned} \tag{11}$$

where the third inequality is due to Proposition 3.3. The computation is analogous for the other marginals.

The sequence (β^n) converges tightly to β which can be seen in a manner analogous to the convergence of the sequence (α^n) to γ. It remains to prove that the sequence satisfies the Γ-lim sup inequality. We fix $\varepsilon > 0$. It suffices to show that

$$\limsup_{n\to\infty} C_\infty[\beta^n] \leq C_\infty[\beta] + \varepsilon.$$

https://doi.org/10.1017/S0956792524000202 Published online by Cambridge University Press

Since for all n the support of β^n is a finite set, we can fix $u^n \in \operatorname{supp} \beta^n$ such that $C_\infty[\beta^n] = c(u^n)$. Moreover, for all n there exists $z^n \in \operatorname{supp} \beta$ such that $d(u^n, z^n) \le \frac{1}{2}\delta_n$. Now for all n large enough to satisfy $\varepsilon_n < \varepsilon$, we have

$$C_\infty[\beta^n] = c(u^n) \le c(z^n) + \varepsilon_n \le C_\infty[\beta] + \varepsilon_n < C_\infty[\beta] + \varepsilon$$

and we are done.

By Corollary 2.6, $\Pi(\mu^1, \dots, \mu^N) \cup_n \Pi(\mu^{1,n}, \dots, \mu^{N,n})$ is compact and therefore the equi-coercivity follows. □

4.2 The integral case

We define the functionals $\mathcal{G}_n, \mathcal{G} : \mathcal{P}(X) \to \mathbb{R} \cup \{+\infty\}$ by

$$\mathcal{G}_n(\beta) = \begin{cases} C[\beta] & \text{if } \beta \in \Pi(\mu^{1,n}, \dots, \mu^{N,n}), \\ +\infty & \text{otherwise;} \end{cases}$$

and

$$\mathcal{G}(\beta) = \begin{cases} C[\beta] & \text{if } \beta \in \Pi(\mu^{1}, \dots, \mu^{N}), \\ +\infty & \text{otherwise.} \end{cases}$$

Also for the integral case, we assume that the measures $\mu^1, \dots, \mu^N$ have compact supports and that the cost function $c : X \to \mathbb{R}$ is continuous. We prove the following:

Proposition 4.2. *The functionals $\mathcal{G}_n$ are equi-coercive and*

$$\mathcal{G}_n \xrightarrow{\Gamma} \mathcal{G}. \tag{12}$$

Proof. The proof is analogous to that of Proposition 4.1. The only substantial difference is in the proof of the Γ-lim sup inequality in the case that the measure β belongs to the set $\Pi(\mu^1, \dots, \mu^N)$. We have to find a sequence (β^n), weakly* converging to β and satisfying Condition (II). Let (β^n) be the discretisation defined in the proof of Proposition 4.1. Since the supports of the measures $\mu^1, \dots, \mu^N$ are compact, also the set $K := \operatorname{spt} \mu^1 \times \cdots \times \operatorname{spt} \mu^N$ is compact. Note that for all $n \in \mathbb{N}$, we have $\operatorname{spt} \beta^n \subset K$. We set $T = \max_{z \in K} c(z)$. Now the function $c_T := \min\{c, T\}$ is continuous and bounded on X and by the weak*-convergence

$$\mathcal{G}(\beta^n) = \int_X c\, d\beta^n = \int_X c_T\, d\beta^n \to \int_X c_T\, d\beta = \int_X c\, d\beta = \mathcal{G}[\beta],$$

from which the Γ-lim sup inequality follows. □

4.3 Proof of the main theorems and a counterexample

Proof. (of Theorem 1.3) By Proposition 3.1 and Remark 3.2, we can find a sequence $(\alpha^n)_n$ with finite supports such that $\operatorname{spt} \alpha^n \subset \operatorname{spt} \gamma$ and $\alpha^n \overset{*}{\rightharpoonup} \gamma$. We define the functionals $\mathcal{F}$ and $\mathcal{F}_n$ of Subsection 4.1 using the marginals of γ and α^n. The plan γ is ICM; therefore by Proposition 1.6, it is finitely optimal. This means that each plan α^n is optimal between its marginals and thus a minimiser of the functional $\mathcal{F}_n$.

The Γ-convergence and equi-coercivity established in Proposition 4.1 imply, by Theorem 2.10, that the minimisers of the functionals $\mathcal{F}_n$ converge, up to subsequences, to a minimiser of $\mathcal{F}$. Therefore, since $\alpha^n \overset{*}{\rightharpoonup} \gamma$, the plan γ is optimal for the problem (P_∞). □

https://doi.org/10.1017/S0956792524000202 Published online by Cambridge University Press

Proof. (of Theorem 1.4) The proof is the same as that of Theorem 1.3. The Γ-convergence is now given by Proposition 4.2. □

In ref. [2], Ambrosio and Pratelli give, for the problem (P), an example of lower semi-continuous cost function $c : X \times X \to [0, \infty]$ (c assumes also the value $+\infty$), for which there exists a $c-$cyclically monotone transport plan which is not optimal. After that it has been shown in refs. [4] and [5] that, for $N = 2$, it is enough that c is Borel measurable and that the set $\{c = +\infty\}$ as a special structure. Actually, the measure theoretical tools introduced in ref. [5] could be applied in an even more general settings. We refer the reader to those papers for further details.

The next example, that is a slightly modified version of the example of [2], shows that also in the case of the problem (P_∞) the continuity of the cost may be required, even when the cost assumes only finite values.

Example 4.3. *Let us consider the two-marginal* L^∞*-optimal transportation problem with marginals* $\mu = \nu = \mathcal{L}|_{[0,1]}$ *and the cost function*

$$c(x, y) = \begin{cases} 1 & \text{if } x = y \\ 2 & \text{otherwise} \end{cases}.$$

We fix an irrational number α. We set $T_1 = Id_{[0,1]}$ and $T_2 : [0, 1] \to [0, 1]$, $T_2(x) = x + \alpha \pmod 1$. Now T_1 is an optimal transportation map for the problem (P_∞) with $C_\infty[T_1] = 1$. Since $C_\infty[T_2] = 2$, T_2 cannot be optimal. However, it is ICM.

In fact if we assume that T_2 is not ICM, we should find a minimal $K \in \mathbb{N}$ and a K-tuple of couples $\{x_i, y_i\}_{i=1}^K$, all belonging to the support of the plan given by T_2, such that

$$\max_{1 \le i \le K} c(x_i, y_i) > \max_{1 \le i \le K} c(x_{i+1}, y_i),$$

with the convention $x_{K+1} = x_1$. By the definition of the map T_2, we have $y_i = x_i + \alpha \pmod 1$ for all i. Given the form of c, the only form in which this inequality can hold is $2 > 1$. The right-hand side now tells us that $y_i = x_i + \alpha \pmod 1$ for all i, that is, $x_{i+1} = x_i + \alpha \pmod 1$ for all i. Summing up now gives us (keeping in mind that $x_{K+1} = x_1$) that $x_1 = x_1 + K\alpha \pmod 1$, contradicting the irrationality of α.

Acknowledgements. Both authors acknowledge the support of GNAMPA-INDAM and *'Fondi di ricerca di ateneo, ex 60 %'* of the University of Firenze.

Funding statement. The research of the first author is part of the project *Metodologie innovative per l'analisi di dati a struttura complessa* financed by the *Fondazione Cassa di Risparmio di Firenze*. The second author acknowledges the support of the PRIN (Progetto di ricerca di rilevante interesse nazionale) 2022J4FYNJ, *Variational methods for stationary and evolution problems with singularities and interfaces*.

Competing interests. There are no competing interest about this research.

References

[1] Abdellaoui, T. & Henich, H. (1994) Sur la distance de deux lois dans le cas vectoriel. Comptes rendus de l'Académie des sciences. *C. R. l'Acad. Sci. Sér 1, Math.* **319**(4), 397–400.

[2] Ambrosio, L. (2003) *Lecture notes on optimal transportation. Mathematical aspects of evolving interfaces (Funchal, 2000), Lecture notes in mathematics*, Springer, New York.

[3] Beiglböck, M., Léonard, C. & Schachermayer, W. (2012) A general duality theorem for the monge–kantorovich transport problem. *Stud. Math.* **209**(2), 2–167.

[4] Beiglböck, M., Goldstern, M., Maresch, G. & Schachermayer, W. (2009) Optimal and better transport plans. *J. Funct. Anal.* **256**(6), 1907–1927.

[5] Bianchini, S. & Caravenna, L. (2010) On optimality of c-cyclically monotone transference plans. *C. R. Math.* **348**(11-12), 613–618.

[6] Braides, A. (2002) *Gamma-convergence for beginners, vol. 22*, Oxford University Press, Oxford.

[7] Brenier, Y. (1991) Polar factorization and monotone rearrangement of vector-valued functions. *Commun. Pure Appl. Math.* **44**(4), 375–417.

[8] Buttazzo, G., De Pascale, L. & Gori-Giorgi, P. (2012) Optimal-transport formulation of electronic density-functional theory. *Phys. Rev. A* **85**(6), 062502.
[9] Caffarelli, L. (1992) The regularity of mappings with a convex potential. *J. Am. Math. Soc.* **5**(1), 99–104.
[10] Carlier, G. (2003) On a class of multidimensional optimal transportation problems. *J. Convex Anal.* **10**(2), 517–530.
[11] Carlier, G., Ekeland, I. (2010) Matching for teams. *Econom. Theory* **42**(2), 397–418.
[12] Carlier, G. & Nazaret, B. (2008) Optimal transportation for the determinant. *ESAIM Control Optim. Calc. Var.* **14**(04), 678–698.
[13] Chiapporri, P.-A., McCann, R. & Nesheim, L. (2010) Hedonic price equilibria, stable matching and optimal transport; equivalence, topology and uniqueness. *Econom. Theory* **42**(2), 317–354.
[14] Colombo, M., De Pascale, L. & Di Marino, S. (2015) Multimarginal optimal transport maps for 1-dimensional repulsive costs. *Can. J. Math.* **67**(2), 350–368.
[15] Colombo, M. & Di Marino, S. (2013) Equality between Monge and Kantorovich multimarginal problems with coulomb cost. *Ann. Matem. Pura Appl.*, 1–14.
[16] Cotar, C., Friesecke, G. & Klüppelberg, C. (2013) Density functional theory and optimal transportation with coulomb cost. *Commun. Pure Appl. Math.* **66**(4), 548–599.
[17] Peyré, G. & Cuturi, M. (2019) Computational optimal transport: With applications to data science. *Found. Trends® Mach. Learn.* **11**, 355–607.
[18] Dal Maso, G. (1993) *An Introduction to* Γ*-Convergence*, Birkhäuser, Boston, MA.
[19] De Pascale, L., Kausamo, A. & Wyczesany, K. 60 years of cyclic monotonicity: A survey. *J. Convex Anal.*, ArXiv: 2308.07682v2.
[20] Di Marino, S., Gerolin, A., & Nenna, L. (2017) Optimal transportation theory for repulsive costs. In: Bergounioux, M., Oudet, E., Rumpf, M., Carlier, G., Champion, T. & Santambrogio, F. (editors), *Topological Optimization and Optimal Transport: In the Applied Sciences*, De Gruyter, Berlin, Boston, pp. 204–256.
[21] Friesecke, G., Mendl, C. B., Pass, B., Cotar, C. & Klüppelberg, C. (2013) N-density representability and the optimal transport limit of the Hohenberg-Kohn functional. *J. Chem. Phys.* **139**(16), 164109.
[22] Gangbo, W. & McCann, R. J. (1996) The geometry of optimal transportation. *Acta Math.* **177**(2), 113–161.
[23] Gangbo, W. & Swiech, A. (1998) Optimal maps for the multidimensional Monge-Kantorovich problem. *Commun. Pure Appl. Math.* **51**(1), 23–45.
[24] Ghoussoub, N. & Moameni, A. (October 2012) A self-dual polar factorization for vector fields. *Commun. Pure Appl. Math.* **66**(6), 905–933. DOI: 10.1002/cpa.21430.
[25] Griessler, C. (2018) *c -cyclical monotonicity as a sufficient criterion for optimality in the multi-marginal Monge-Kantorovich problem*. *Proc. Am. Math. Soc.* **145**, 4735–4740.
[26] Jylhä, H. (2015) The L^∞ optimal transport: Infinite cyclical monotonicity and the existence of optimal transport maps. *Calc. Var. Partial Differ. Eq.* **52**, 303–326.
[27] Heinich, H. (2002) Problème de Monge pour n probabilités. *C. R. Math.* **334**(9), 793–795.
[28] Kellerer, H. G. (1984) Duality theorems for marginal problems. *Prob. Theory Relat. Fields* **67**(4), 399–432.
[29] Knott, M. & Smith, C. (1992) On Hoeffding-Fréchet bounds and cyclic monotone relations. *J. Multiv. Anal.* **40**(2), 328–334.
[30] Pass, B. (2011) Uniqueness and Monge solutions in the multi-marginal optimal transportation problem. *SIAM J. Math. Anal.* **43**(6), 2758–2775.
[31] Pass, B. (2012) On the local structure of optimal measures in the multi-marginal optimal transportation problem. *Calc. Var. Partial Differ. Eq.* **43**(3-4), 529–536.
[32] Rachev, S. T. & Rüschendorf, L. (1998) Mass transportation problems. In: *Vol. I. Probability and its Applications (New York). Theory*, Springer-Verlag, New York.
[33] Seidl, M. (1999) Strong-interaction limit of density-functional theory. *Phys. Rev. A* **60**(6), 4387–4395.
[34] Seidl, M., Gori-Giorgi, P. & Savin, A. (2007) Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities. *Phys. Rev. A* **75**(4), 042511.
[35] Seidl, M., Perdew, J. P. & Levy, M. (1999) Strictly correlated electrons in density-functional theory. *Phys. Rev. A* **59**(1), 51–54.
[36] Villani, C. (2003) *Topics in Optimal Transportation, vol. 58*, American Mathematical Society, Providence, RI.
[37] Villani, C. (2009) *Optimal transport. Old and new. vol. 338 of Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, Berlin.

Cite this article: De Pascale L. and Kausamo A. (2025). Sufficiency of ***c***-cyclical monotonicity in a class of multi-marginal optimal transport problems. *European Journal of Applied Mathematics*, **36**, 68–81. https://doi.org/10.1017/S0956792524000202