



HAL
open science

Local matching indicators for transport problems with concave costs

Julien Salomon, Julie Delon, Andrei Sobolevskii

► **To cite this version:**

Julien Salomon, Julie Delon, Andrei Sobolevskii. Local matching indicators for transport problems with concave costs. 2009. hal-00525994v1

HAL Id: hal-00525994

<https://hal.science/hal-00525994v1>

Preprint submitted on 13 Oct 2010 (v1), last revised 8 Dec 2011 (v3)

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

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

LOCAL MATCHING INDICATORS FOR TRANSPORT PROBLEMS WITH CONCAVE COSTS

JULIE DELON[†] JULIEN SALOMON[‡] AND ANDREĬ SOBOLEVSKI^{§¶}

Key words. Optimal transport, Assignment problems, Concave costs, Local matching indicators,

AMS subject classifications. 90C08, 68Q25, 90C25

Abstract. In this paper, we introduce a class of indicators that enable to compute efficiently optimal transport plans associated to arbitrary distributions of N demands and M supplies in \mathbb{R} in the case where the cost function is concave. The computational cost of these indicators is small and independent of N . A hierarchical use of them enables to obtain an efficient algorithm.

1. Introduction. The origins of optimal transportation go back to the late eighteenth century, when Monge [12] published his *Mémoire sur la théorie des déblais et des remblais* (1781). The problem, which was rediscovered and further studied by Kantorovich in the late 30's, can be described in the following way. Given two probability distributions μ and ν on X and c a measurable cost function on $X \times X$, find a joint probability measure π on $X \times X$ with marginals μ and ν and which minimizes the transportation cost

$$\int \int_{X \times X} c(x, y) d\pi(x, y). \quad (1.1)$$

Probability measures π with marginals μ and ν are called *transportation plans* or *transport plans*. A transport plan that minimizes the cost (1.1) is said to be *optimal*.

When the measures μ and ν are discrete (linear combinations of Dirac masses), the problem can be recast as a linear optimization one. For $N \geq 1$, consider two discrete distributions of mass, or *histograms*, given on \mathbb{R}^N : $\{(p_i, s_i)\}$, which represents “supplies” at locations p_i with weights s_i and $\{(q_j, d_j)\}$, which represents “demands” at locations q_j with weights d_j (notation from [1]) and assume that all values of d_i and s_j are positive reals with $D := \sum_i d_i$ and $S := \sum_j s_j$. The problem consists in minimizing the transport cost

$$\sum_{i,j} c(p_i, q_j) \gamma_{ij},$$

where γ_{ij} is the amount of mass going from p_i to q_j , subject to the conditions

$$\gamma_{ij} \geq 0, \quad \sum_j \gamma_{ij} \leq d_i, \quad \sum_i \gamma_{ij} \leq s_j, \quad \sum_{i,j} \gamma_{ij} = \min(S, D).$$

The matrix of values $\gamma = \{\gamma_{ij}\}$ is still called *transport plan*. When $S = D$, the problem is said to be *balanced* and is only a reformulation of (1.1) for discrete measures.

[†]LTCI CNRS, Télécom ParisTech, 46 rue Barrault F-75634 Paris cedex 13, France (julie.delon@enst.fr).

[‡]CEREMADE, UMR CNRS 7534, Université de Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, F-75775 Paris cedex 16, France (salomon@ceremade.dauphine.fr).

[§]Physics department, M.V. Lomonosov Moscow State University, Russia.

[¶]UMI 2615 CNRS “Laboratoire J.-V. Poncelet” (sobolevski@iitp.ru), Russia.

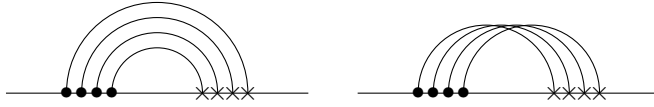


FIG. 1.1. On the left: optimal plans associated to a concave cost. On the right: optimal plans associated to a convex cost. Supplies are represented by points and demands by crosses.

When $S \neq D$, the problem is said to be *unbalanced*. The cases $S < D$ and $S > D$ can be treated in the same way. This paper deals with balanced problems and unbalanced problems of the form $S > D$.

In the *unitary case*, *i.e.* when all the masses s_i and d_i are equal to a single value v , it turns out that if γ is optimal, for all i, j , $\gamma_{i,j} \in \{0, v\}$ and for all j there exists only one i such that $\gamma_{i,j} = v$ (each demand receives all the mass from one supply). In the balanced case, the matrix γ is thus a permutation matrix. In the unbalanced case, the permutation matrix is truncated. As a consequence, this particular case boils down to an *assignment problem*. Such problems have been thoroughly studied by the combinatorial optimization community [5].

Optimal transportation problems appear in many fields, such as economy or physics for instance, see e.g. [9, 6, 4]. In economic examples optimal transport is often related to the field of logistic where supplies are furnished by producers at specific places p_i and in specific quantities d_i , while demands corresponds to consumers locations and needs. Depending on the application, various cost functions c can be used. For instance, concave functions of the distance appear as more realistic cost functions in many economic situations. Indeed, as underlined by McCann [11], a concave cost “*translates into an economy of scale for longer trips and may encourage cross-hauling.*”

During the last decades, many authors have taken interest in the study of existence, uniqueness and properties of optimal plans [10, 8, 2], with a specific interest for convex costs. Detailed descriptions of these results can be found in the books [14, 15]. One case of particular interest is the one-dimensional case, which, when c is a convex function of the distance on the line, has been completely understood [13]. Indeed, this problem has an explicit solution that does not depend on c (provided that it is convex) and consisting in a monotone rearrangement (see Chapter 2.2 of [14]). Recently, an efficient method has been introduced to tackle this issue on the circle [7], where the starting point for the monotone rearrangement is not known, and its choice and hence the optimal plan itself, unlike in the case of an interval, do depend on the cost function. Unfortunately, these results do not extend to non-convex costs, in particular to concave costs (see Figure 1.1 for an example). Although it is of broad interest for many applications, few papers have taken interest in this case [11] and computing solutions is far from obvious in general. Indeed, and contrary to the convex case on the line, optimal plans strongly depends on the choice of the function c . Consider the case of two unitary supplies at positions $p_1 = 0$ and $p_2 = 1.2$ and two unitary demands at positions $q_1 = 1$ and $q_2 = 2.2$ on the line, as drawn on Figure 1.2. If the cost function is $c(x, y) = |x - y|^{0.9}$, the left solution will be optimal, whereas the other one will be chosen for $c(x, y) = |x - y|^{0.5}$. For a convex cost, the left solution would always be chosen.

In practice, when no analytic solution is given (*i.e.* most of the time), finding optimal plans can be a tedious task. As underlined before, in a discrete setting, the problem can be written as a linear programming problem, and optimal plans can

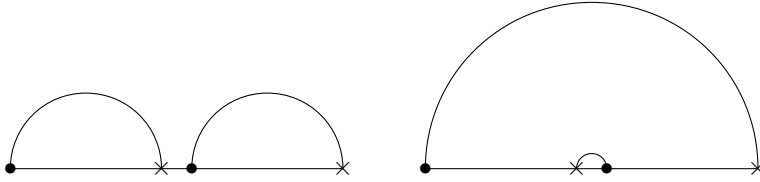


FIG. 1.2. On the left: solutions associated to the concave cost $c(x, y) = |x - y|^{0.9}$, and on the right to the cost $c(x, y) = |x - y|^{0.5}$. Supplies are represented by points and demands by crosses.

be estimated numerically, by using simplex or auction algorithms [3] for instance. Most of the time, these estimations lead to expansive computations. The assignment problem can be solved with dedicated approaches, such as the Hungarian method or more involved algorithms (see [5] for details).

The goal of this paper is to introduce a class of functions that reveals the local structure of optimal transport plans (either on the line, or on the circle), when the cost c is a concave function of the distance. The study will be first limited to the case of unitary masses, either in the balanced or unbalanced case. As a by-product, we build an algorithm that permits to obtain in both cases optimal transport plans in less than $O(N^2)$ operations, where N is the number of demands under consideration. However, let us insist that our aim is not to compete with recent linear assignment algorithms, which may be more interesting in practice, at least for balanced problems. Observe that our algorithm complements the method suggested by McCann [11], although the approach we follow here is closer to the purely combinatorial approach of [1]. The results of this last work are extended here to the general framework of concave cost functions.

The paper is organized as follows. In Section 2, we present the main result of the paper, which states that consecutive matching points in the optimal plan can be found thanks to local indicators, independently of other points on the line. Section 3 is devoted to different technical results, necessary to the proof of this result, which is itself presented in Section 4. Thanks to the low number of evaluations of the cost function required to apply the indicators, we derive in Section 5 an algorithm that finds an optimal transport plan in $O(N^2)$ operations in the worst case. Finally, Section 6 is devoted to extensions of these results to the circle and to the non-integer case.

2. Setting of the problem and main result.

2.1. The optimal transport problem. This paper deals with the problem of finding an optimal transport plan in the case where the problem contains possibly more supplies than demands and the transport cost is concave: the larger the distance to cover is, the less the transport costs per unit distance.

Consider two integers $M \geq N$ and $P = (p_i)_{i=1, \dots, M}$ and $Q = (q_i)_{i=1, \dots, N}$, two sets of points in \mathbb{R} that represent respectively unitary supply and unitary demand locations. The problem we deal with consists in minimizing the cost

$$\mathcal{C}(\gamma) = \sum_{i,j} c(p_i, q_j) \gamma_{ij}, \quad (2.1)$$

where $c(p_i, q_j) \in \mathbb{R}^+$ is the cost resulting from transport of a unit mass between p_i and q_j . The quantity γ_{ij} is the amount of mass going from p_i to q_j , subject for all

i, j to the conditions

$$\gamma_{ij} \geq 0, \quad \sum_j \gamma_{ij} \leq 1, \quad \sum_i \gamma_{ij} = 1. \quad (2.2)$$

The case $N = M$ will be called the *balanced case*, while the case $M > N$ will be called the *unbalanced case*. Observe that since the total number of supplies is larger than the total number of demands, some of the supplies may remain unmatched ($\forall j, \gamma_{ij} = 0$). It is well known that if γ minimizes the cost (2.1) under conditions (2.2), then without loss of generality one can assume that $\gamma_{ij} \in \{0, 1\} \forall i, j$, so that one can reformulate the problem as finding

$$\min_{\sigma \in \Sigma_M} C(\sigma),$$

where Σ_M is the set of all permutations of $\{1, \dots, M\}$ and where

$$C(\sigma) = \sum_{i/\sigma(i) \in \{1, \dots, N\}} c(p_i, q_{\sigma(i)}). \quad (2.3)$$

We focus on the case where the function c involves a concave function as stated in the next definition.

DEFINITION 2.1. *The cost function c in (2.3) is said to be concave if it is defined by $c(p, q) = g(|p - q|)$ with $p, q \in \mathbb{R}$, where $g: \mathbb{R}^+ \rightarrow \mathbb{R}$ is a concave non-decreasing function such that $g(0) := \lim_{x \rightarrow 0} g(x) \geq -\infty$.*

Some examples of such costs are given by $g(x) = \log(x)$ with $g(0) = -\infty$ and $g(x) = \sqrt{x}$ or $g(x) = |x|$ with $g(0) = 0$. If $g(0) > -\infty$, then we assume without loss of generality that $g(0) = 0$ (this changes the value of (2.1) by an amount $N g(0)$ independent of the transport plan). In what follows, we denote by σ^* the map associated to a given optimal transport plan between P and Q : $C(\sigma^*) \leq C(\sigma)$ for all $\sigma \in \Sigma_M$.

Observe that if two points p_i and q_j have the same position, then there exists an optimal transport plan σ between P and Q such that $\sigma(i) = j$. This results from a well known property of concave costs, for which all the mass shared by the measures P and Q stays in place [14]: suppose that a supply p and a demand q located at the same point are not matched together but with other demand and supply p' and q' located at distances x and y respectively. Either if $g(0) = 0$ or $g(0) = -\infty$ and as soon as g is concave, one has

$$\forall x, y > 0, \quad g(0) + g(x + y) \leq g(x) + g(y),$$

which implies that matching p and q is cheaper. In the following and before Section 6.2, we will always assume that $P \cap Q = \emptyset$.

2.2. Non-crossing rule and Chains. In this section, we present a way to subdivide the initial set of points $P \cup Q$ into a family of particular subsets called *Chains* that are preserved by optimal transport plans. As a result, we obtain a precomputation step that enables to break down an optimal transport problem into a set of smaller and independent problems.

One significant feature of concave costs is that optimal trajectories do not cross each other, as described by the following lemma.

LEMMA 2.2. *Consider two pairs of points $(p_i, q_{\sigma^*(i)})$ and $(p_{i'}, q_{\sigma^*(i')})$ such that $p_i < q_{\sigma^*(i)}$, $p_{i'} < q_{\sigma^*(i')}$ and*

$$c(p_i, q_{\sigma^*(i)}) + c(p_{i'}, q_{\sigma^*(i')}) < c(p_{i'}, q_{\sigma^*(i)}) + c(p_i, q_{\sigma^*(i')}). \quad (2.4)$$

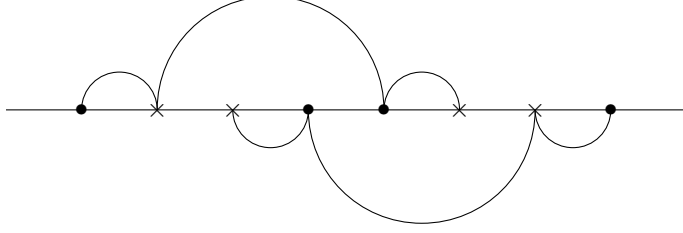


FIG. 2.1. Example of a problem containing two chains.

Then, the segments $[p_i, q_{\sigma^*(i)}], [p_{i'}, q_{\sigma^*(i')}]$ are either **disjoint** or **nested**, in the sense that following alternative holds:

1. $[p_i, q_{\sigma^*(i)}] \cap [p_{i'}, q_{\sigma^*(i')}] = \emptyset$,
2. $[p_i, q_{\sigma^*(i)}] \subset [p_{i'}, q_{\sigma^*(i')}]$ or $[p_{i'}, q_{\sigma^*(i')}] \subset [p_i, q_{\sigma^*(i)}]$.

The same conclusion holds if (2.4) is replaced by the analogous non-strict inequality but the function g is strictly monotone.

This result directly follows from the concavity of the cost function and is often referred to as the “non-crossing rule” [1, 11]. The ideas of the proof can be found in [1] or in the proof of Lemma 6.3.

In the unbalanced case, some points may not belong to a nested segment.

DEFINITION 2.3. A point $r \in P \cup Q$ is said to be isolated if

$$\forall i, \sigma^*(i) \in \{1, \dots, N\}, r \notin [\min(p_i, q_{\sigma^*(i)}), \max(p_i, q_{\sigma^*(i)})].$$

If g is strictly monotone, such points can be characterized as follows.

LEMMA 2.4. If g is strictly monotone, then unmatched supplies are isolated.

The proof is easy and left to the reader.

A first consequence of these rules is usually called the *local balance of supplies and demands*: in the unitary case, there are as many supplies as demands between any two matched points p_{i_0} and q_{j_0} . This consequence permits to conclude that the search for optimal transport plans can be restricted to chains, as defined in the following.

Given a supply point p_i , define its *left neighbor* q'_i as the nearest demand point on the left of p_i such that the numbers of supplies and demands between q'_i and p_i are equal; define the *right neighbor* q''_i of p_i in a similar way. Iterating this procedure, one obtains a subset that is preserved by σ^* .

DEFINITION 2.5. A chain is a maximal alternating sequence of supplies and demands of one of the forms

1. $(p_{i_1}, q_{j_1}, \dots, p_{i_k}, q_{j_k})$,
2. $(q_{j_1}, p_{i_1}, \dots, q_{j_k}, p_{i_k})$,
3. $(p_{i_1}, q_{j_1}, \dots, q_{j_{k-1}}, p_{i_k})$,

with $k \geq 1$ and such that each pair of consecutive points in the sequence is made of a point and its right neighbor.

Examples of chains are shown in Figure 2.1. Observe that because of Case (3), some chains can be composed of only one supply, and no demand. An extension of the proof of Lemma 3 of [1] shows that the family of chains forms a partition of $P \cup Q$. This partition is particularly adapted to our transport problem as one can easily prove that each chain is preserved by an optimal transport plan, hence a possible parallelization of the resolution. The construction of the set of chains only depends on relative positions of supplies and demands and does not involve any evaluation of the cost function. It can be achieved in $O(N + M)$ operations: a possible algorithm

consists in considering sequentially (say from the left to the right) the points of the set $P \cup Q$, and, at the same time, building all the chains iteratively either by adding the current point to an existing chains, or by initiating a new chain in such a way that the local balance of supplies and demands is fulfilled.

Note also that if a chain is composed of only one supply, this supply can not be matched in an optimal transport plan, and can consequently be dismissed of the original problem at this step.

Most of the rest of the paper deals with the research of optimal transport plans in the case of chains: except in Section 6.2, we focus in the sequel on the cases where P and Q satisfy $M = N$ (*balanced case*) and

$$p_1 < q_1 < \dots < p_i < q_i < p_{i+1} < q_{i+1} < \dots < p_N < q_N, \quad (2.5)$$

or $M = N + 1$ (*unbalanced case*) and

$$p_1 < q_1 < \dots < p_i < q_i < p_{i+1} < q_{i+1} < \dots < p_N < q_N < p_{N+1}. \quad (2.6)$$

In these cases the set $P \cup Q$ is called *balanced chain* and *unbalanced chain* respectively.

2.3. Main result. Thanks to the non-crossing rule, one knows that there exist at least two consecutive points (p_i, q_i) or (q_i, p_{i+1}) that are matched in any optimal transport plan. Starting from this remark, we take advantage of the structure of a chain to introduce a class of indicators that enable to detect a priori such pairs of points.

DEFINITION 2.6 (Local Matching Indicators of order k). *Given $k > 0$, consider $2k + 2$ consecutive points in a chain. If the first point is a supply p_i , define*

$$I_k^p(i) = c(p_i, q_{i+k}) + \sum_{\ell=0}^{k-1} c(p_{i+\ell+1}, q_{i+\ell}) - \sum_{\ell=0}^k c(p_{i+\ell}, q_{i+\ell}),$$

else denote the first point q_i and define

$$I_k^q(i) = c(p_{i+k+1}, q_i) + \sum_{\ell=1}^k c(p_{i+\ell}, q_{i+\ell}) - \sum_{\ell=0}^k c(p_{i+\ell+1}, q_{i+\ell}).$$

This definition is schematically depicted in Figure 2.2 in the case $k = 2$.

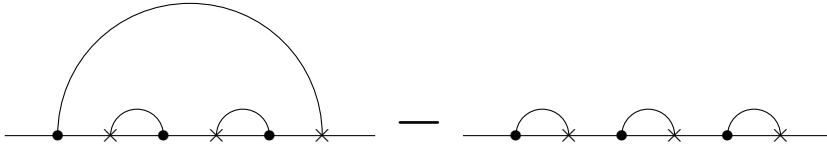


FIG. 2.2. Schematic representation of an indicator of order 2.

Note that in the first alternative of this definition, we have necessarily $1 \leq k \leq N - 1$, $1 \leq i \leq N - k$. In the second alternative, we have necessarily $1 \leq k \leq N - 2$ and $1 \leq i \leq N - k - 1$ in the balanced case and $1 \leq k \leq N - 1$ and $1 \leq i \leq N - k$ in the unbalanced case. The interest of these functions lies in the next result.

THEOREM 2.7 (Negative Local Matching Indicators of order k). *Let $k_0 \in \mathbb{N}$ with $1 \leq k_0 \leq N - 1$ and $i_0 \in \mathbb{N}$, such that $1 \leq i_0 \leq N - k_0$. In the unbalanced case, suppose in addition that g is strictly monotone.*

Assume that

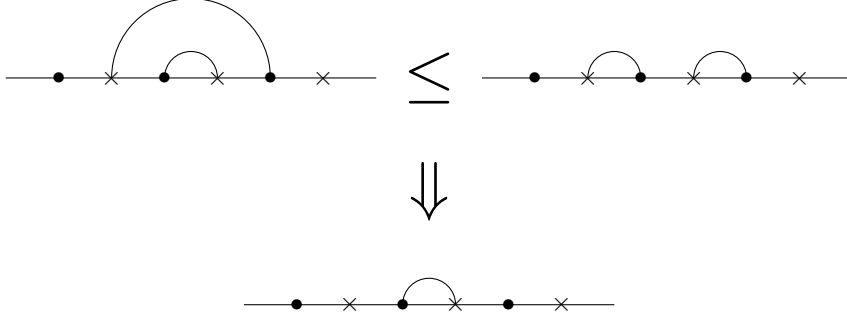


FIG. 2.3. Schematic representation of the result of Theorem 2.7 in the case $k = 1$.

1. $I_k^p(i) \geq 0$ for $k = 1, \dots, k_0 - 1$, $i_0 \leq i \leq i_0 + k_0 - k$,
2. $I_k^q(i') \geq 0$ for $k = 1, \dots, k_0 - 1$, $i_0 \leq i' \leq i_0 + k_0 - k - 1$, (resp. $1 \leq i' \leq i_0 + k_0 - k$ in the unbalanced case)
3. $I_{k_0}^p(i_0) < 0$.

Then any permutation σ associated to an optimal transport plan satisfies $\sigma(i) = i - 1$ for $i = i_0 + 1, \dots, i_0 + k_0$.

If the third condition is replaced by $I_{k_0}^q(i_0) < 0$ (with the same bounds on k_0 and i_0 in the unbalanced case, and with $1 \leq k_0 \leq N - 2$ and $1 \leq i_0 \leq N - k_0 - 1$ in the balanced case), then any permutation σ associated to an optimal transport plan satisfies $\sigma(i) = i$ for $i = i_0 + 1, \dots, i_0 + k_0$.

This result is represented in broad outline in Figure 2.3. For practical purposes, these indicators allow to find pairs of neighbors that are matched in an optimal transport plan.

3. Technical results. These section aims at introducing technical results that are required to prove Theorem 2.7. We keep the notations introduced therein. We start with a basic result that plays a significant role in the proof of Theorem 2.7. As for the non-crossing rule, the concavity of the cost function is an essential assumption of this lemma.

LEMMA 3.1. We keep the previous notations. For $x, y \in \mathbb{R}^+$, define

$$\varphi_{k,i}^p(x, y) = g(x + y + q_{i+k} - p_i) + \sum_{\ell=0}^{k-1} c(p_{i+\ell+1}, q_{i+\ell}) - g(x) - g(y) - \sum_{\ell=1}^{k-1} c(p_{i+\ell}, q_{i+\ell}),$$

for $k, i \in \mathbb{N}$, such that $1 \leq k \leq N - 1$ and $1 \leq i \leq N - k$, and

$$\varphi_{k,i}^q(x, y) = g(x + y + p_{i+k+1} - q_i) + \sum_{\ell=1}^k c(p_{i+\ell}, q_{i+\ell}) - g(x) - g(y) - \sum_{\ell=1}^{k-1} c(p_{i+\ell+1}, q_{i+\ell}),$$

for $k, i \in \mathbb{N}$, such that $1 \leq k \leq N - 2$ and $1 \leq i \leq N - k - 1$ in the balanced case and $1 \leq k \leq N - 1$ and $1 \leq i \leq N - k$ in the unbalanced case. Both functions $\varphi_{k,i}^p(x, y)$ and $\varphi_{k,i}^q(x, y)$ are decreasing with respect to each of their two variables.

This lemma is a direct consequence of the concavity of the function g . To deal with unbalanced chains, we need two additional lemmas, one of them requiring that g is strictly monotone. The first result is a simple extension of a lemma usually referred as "The rule of three" in the literature [11].

LEMMA 3.2 (“rule of three”). *Suppose that g is strictly monotone. Under the assumptions (1–3) of Theorem 2.7, the following inequalities are satisfied*

$$|q_j - p_{j+1}| < \min(|p_{i_0} - q_j|, |p_{j+1} - q_{i_0+k_0}|), \quad \forall j \in \{i_0, \dots, i_0 + k_0 - 1\}.$$

If $I_{k_0}^q(i'_0) < 0$ instead of $I_{k_0}^p(i_0) < 0$, the inequalities become

$$|p_j - q_j| < \min(|q_{i_0} - p_j|, |q_j - p_{i_0+k_0+1}|), \quad \forall j \in \{i_0 + 1, \dots, i_0 + k_0\}.$$

Proof: Assumption (3) of Theorem 2.7 implies that

$$c(q_j, p_{j+1}) + c(p_{i_0}, q_{i_0+k_0}) < \sum_{i=i_0}^{i_0+k_0} c(p_i, q_i) - \sum_{i=i_0}^{i_0+k_0-1} c(p_{i+1}, q_i) + c(q_j, p_{j+1}).$$

Now, because of Assumption (1), we have $I_{j-i_0}^p(i_0) \geq 0$ and $I_{i_0+k_0-j-1}^p(j+1) \geq 0$, which means that

$$\sum_{i=i_0}^j c(p_i, q_i) \leq c(p_{i_0}, q_j) + \sum_{i=i_0}^{j-1} c(p_{i+1}, q_i)$$

and

$$\sum_{i=j+1}^{i_0+k_0} c(p_i, q_i) \leq c(p_{j+1}, q_{i_0+k_0}) + \sum_{i=j+1}^{i_0+k_0-1} c(p_{i+1}, q_i).$$

Thus,

$$c(q_j, p_{j+1}) + c(p_{i_0}, q_{i_0+k_0}) < c(p_{i_0}, q_j) + c(p_{j+1}, q_{i_0+k_0}).$$

Since g is strictly increasing and since $|p_{i_0} - q_{i_0+k_0}| \geq \max(|p_{i_0} - q_j|, |p_{j+1} - q_{i_0+k_0}|)$, this implies that $|p_{j+1} - q_j| < \min(|p_{i_0} - q_j|, |p_{j+1} - q_{i_0+k_0}|)$. The result in the case $I_{k_0}^q(i'_0) < 0$ can be deduced by symmetry. \square

Note that in this proof, only the fact that the cost is an strictly increasing function of the distance is necessary. In particular the result also holds in the case where the cost function is increasing and convex.

LEMMA 3.3 (“partial sums”). *Under the assumptions of Theorem 2.7, for any ℓ in $\{i_0 + 1, \dots, i_0 + k_0\}$ and ℓ' in $\{i_0, \dots, i_0 + k_0 - 1\}$, the following inequalities are satisfied:*

$$\sum_{i=i_0}^{\ell-1} c(p_i, q_i) > \sum_{i=i_0}^{\ell-1} c(p_{i+1}, q_i), \quad (3.1)$$

and

$$\sum_{i=\ell'+1}^{i_0+k_0} c(p_i, q_i) > \sum_{i=\ell'}^{i_0+k_0-1} c(p_{i+1}, q_i). \quad (3.2)$$

Proof: In order to prove inequality (3.1), remark that since $I_{i_0}^p(k_0) < 0$

$$\begin{aligned} \sum_{i=i_0}^{\ell-1} c(p_i, q_i) &= \sum_{i=i_0}^{i_0+k_0} c(p_i, q_i) - \sum_{i=\ell}^{i_0+k_0} c(p_i, q_i) \\ &> c(p_{i_0}, q_{i_0+k_0}) + \sum_{i=i_0}^{i_0+k_0-1} c(p_{i+1}, q_i) - \sum_{i=\ell}^{i_0+k_0} c(p_i, q_i), \end{aligned}$$

for ℓ such that $i_0+1 \leq \ell \leq i_0+k_0$. Moreover, since $I_{i_0+k_0-\ell}^p(\ell) \geq 0$ and g is increasing,

$$\sum_{i=i_0}^{\ell-1} c(p_i, q_i) > c(p_{i_0}, q_{i_0+k_0}) + \sum_{i=i_0}^{i_0+k_0-1} c(p_{i+1}, q_i) - c(p_\ell, q_{i_0+k_0}) - \sum_{i=\ell}^{i_0+k_0-1} c(p_{i+1}, q_i),$$

which leads to the inequality (3.1). The proof of Equation (3.2) follows the same path. \square

On the contrary to the previous results, the next lemma will not be used in the proofs of this paper. We state it since it permits to detect isolated point, hence, it can be used to save computational time.

LEMMA 3.4 (“isolation rule”). *Suppose that g is strictly monotone. If p_i is an unmatched point in the unbalanced chain (2.6), then if $i > 1$*

$$c(p_i, q_{i-1}) \geq c(p_{i-1}, q_{i-1})$$

and if $i < N$

$$c(p_i, q_i) \geq c(p_{i+1}, q_i).$$

Proof: Assume for instance that $i > 1$ and $c(p_i, q_{i-1}) < c(p_{i-1}, q_{i-1})$. Thanks to Lemma 2.4, p_i is isolated, and consequently $\sigma^{-1}(i-1) \leq i-1$. Thus, $c(p_i, q_{i-1}) < c(p_{i-1}, q_{i-1}) \leq c(p_{\sigma^{-1}(i-1)}, q_{i-1})$. It is then cheaper to exclude $p_{\sigma^{-1}(i-1)}$ and match p_i with q_{i-1} , which contradicts the optimality of σ . \square

4. Proof of Theorem 2.7. We are now in the position to prove our main result. In a first part we focus on the balanced case, and then go to the unbalanced case, which requires more efforts.

4.1. The balanced case. Consider the balanced case, i.e., the situation corresponding to (2.5). We focus on the case where $I_{k_0}^p(i_0) < 0$. The case $I_{k_0}^q(i'_0) < 0$ can be treated the same way.

The proof consists in proving that Assumptions (1-3) of Theorem 2.7 imply that neither demand nor supply points located between p_{i_0} and $q_{i_0+k_0}$ can be matched with points located outside this interval, i.e. that the set $\mathcal{S}_{i_0} = \{p_i, i_0+1 \leq i \leq i_0+k_0\} \cup \{q_i, i_0 \leq i \leq i_0+k_0-1\}$ is stable under an optimal transport plan. In this case, the result follows from Assumption (1-2).

Suppose that \mathcal{S}_{i_0} is not preserved by an optimal transport plan σ^* . Three cases can occur:

- a) There exists $i_1 \in \mathbb{N}$, such that $1 \leq i_1 \leq i_0$ and $i_0 \leq \sigma^*(i_1) \leq i_0 + k_0 - 1$ and there exists $i'_1 \in \mathbb{N}$, such that $\sigma^*(i_1) + 1 \leq i'_1 \leq i_0 + k_0$ and $i_0 + k_0 \leq \sigma^*(i'_1) \leq N$.
- b) There exists $i_2 \in \mathbb{N}$, with $i_0 + 1 \leq i_2 \leq i_0 + k_0$ such that $1 \leq \sigma^*(i_2) \leq i_0 - 1$.
- c) There exists $i_2 \in \mathbb{N}$, with $i_0 + k_0 < i_2 \leq N$ such that $i_0 \leq \sigma^*(i_2) < i_0 + k_0$.

We first prove that Case a) cannot occur.

In Case a), one can assume without loss of generality that $\sigma^*(i_1)$ is the largest index such that $1 \leq i_1 \leq i_0$, $i_0 \leq \sigma^*(i_1) \leq i_0 + k_0 - 1$ and that i'_1 is the smallest index such that $\sigma^*(i_1) + 1 \leq i'_1 \leq i_0 + k_0$, $i_0 + k_0 \leq \sigma^*(i'_1) \leq N$. Assume also that we are not in Cases b) or c). With such assumptions, the (possibly empty) subset $\{p_i, \sigma^*(i_1) + 1 \leq i \leq i'_1 - 1\} \cup \{q_i, \sigma^*(i_1) + 1 \leq i \leq i'_1 - 1\}$ is stable by σ^* . Because of Assumptions (1–2), no nesting (*i.e.* no pair of nested matchings) can occur in this subset, and $\sigma^*(i) = i$ for $i = \sigma^*(i_1) + 1, \dots, i'_1 - 1$.

On the other hand, since σ^* is supposed to be optimal, one has:

$$c(p_{i_1}, q_{\sigma^*(i_1)}) + c(p_{i'_1}, q_{\sigma^*(i'_1)}) + \sum_{i=\sigma^*(i_1)+1}^{i'_1-1} c(p_i, q_i) \leq c(p_{i_1}, q_{\sigma^*(i'_1)}) + \sum_{i=\sigma^*(i_1)}^{i'_1-1} c(p_{i+1}, q_i).$$

Thanks to Lemma 3.1, one deduces from this last inequality that:

$$c(p_{i_0}, q_{\sigma^*(i_1)}) + c(p_{i'_1}, q_{i_0+k_0}) + \sum_{i=\sigma^*(i_1)+1}^{i'_1-1} c(p_i, q_i) \leq c(p_{i_0}, q_{i_0+k_0}) + \sum_{i=\sigma^*(i_1)}^{i'_1-1} c(p_{i+1}, q_i),$$

and then:

$$\begin{aligned} c(p_{i_0}, q_{\sigma^*(i_1)}) + \sum_{i=i_0}^{\sigma^*(i_1)-1} c(p_{i+1}, q_i) + c(p_{i'_1}, q_{i_0+k_0}) + \sum_{i=i'_1}^{i_0+k_0-1} c(p_{i+1}, q_i) \\ + \sum_{i=\sigma^*(i_1)+1}^{i'_1-1} c(p_i, q_i) \leq c(p_{i_0}, q_{i_0+k_0}) + \sum_{i=i_0}^{i_0+k_0-1} c(p_{i+1}, q_i). \end{aligned} \quad (4.1)$$

According to Assumption (1), $I_{\sigma^*(i_1)-i_0}^p(i_0) \geq 0$ and $I_{i_0+k_0-i'_1}^p(i'_1) \geq 0$, so that:

$$\sum_{i=i_0}^{\sigma^*(i_1)} c(p_i, q_i) \leq c(p_{i_0}, q_{\sigma^*(i_1)}) + \sum_{i=i_0}^{\sigma^*(i_1)-1} c(p_{i+1}, q_i)$$

$$\sum_{i=i'_1}^{i_0+k_0} c(p_i, q_i) \leq c(p_{i'_1}, q_{i_0+k_0}) + \sum_{i=i'_1}^{i_0+k_0-1} c(p_{i+1}, q_i).$$

Combining these last inequalities with (4.1) one finds that:

$$\sum_{i=i_0}^{i_0+k_0} c(p_i, q_i) \leq c(p_{i_0}, q_{i_0+k_0}) + \sum_{i=i_0}^{i_0+k_0-1} c(p_{i+1}, q_i),$$

which contradicts Assumption (3).

Let us now prove that Cases b) and c) contradict the assumptions. As Cases b) and c) can be treated in the same way, we only consider Case b). Without loss of

generality, one can assume that i_2 is the smallest index such that $i_0 + 1 \leq i_2 \leq i_0 + k_0$ and $\sigma^*(i_2) \leq i_0 - 1$. Because there are necessarily as many demands as supplies between q_{i_0} and p_{i_2} , there exists one and only one index i'_2 such that $i_0 \leq \sigma^*(i'_2) \leq i_2 - 1$ and $1 \leq i'_2 \leq i_0$. Consequently, the (possibly empty) subsets $\{p_i, i_0 + 1 \leq i \leq \sigma^*(i'_2)\} \cup \{q_i, i_0 \leq i \leq \sigma^*(i'_2) - 1\}$ and $\{p_i, \sigma^*(i'_2) + 1 \leq i \leq i_2 - 1\} \cup \{q_i, \sigma^*(i'_2) + 1 \leq i \leq i_2 - 1\}$ are stable by an optimal transport plan. Because of Assumptions (1–2), no nesting can occur in these subsets, and $\sigma^*(i) = i - 1$ for $i = i_0 + 1, \dots, \sigma^*(i'_2)$ and $\sigma^*(i) = i$ for $i = \sigma^*(i'_2) + 1, \dots, i_2 - 1$.

On the other hand, since σ^* is supposed to be optimal, one has

$$\begin{aligned} c(p_{i_2}, q_{\sigma^*(i_2)}) + c(p_{i'_2}, q_{\sigma^*(i'_2)}) + \sum_{i=i_0+1}^{\sigma^*(i'_2)} c(p_i, q_{i-1}) + \sum_{i=\sigma^*(i'_2)+1}^{i_2-1} c(p_i, q_i) \\ \leq c(p_{i'_2}, q_{\sigma^*(i_2)}) + \sum_{i=i_0+1}^{i_2} c(p_i, q_{i-1}). \end{aligned}$$

Thanks to Lemma 3.1, one deduces from this last inequality that:

$$\begin{aligned} c(p_{i_2}, q_{\sigma^*(i_2)}) + c(p_{i_0}, q_{\sigma^*(i'_2)}) + \sum_{i=i_0+1}^{\sigma^*(i'_2)} c(p_i, q_{i-1}) + \sum_{i=\sigma^*(i'_2)+1}^{i_2-1} c(p_i, q_i) \\ \leq c(p_{i_0}, q_{\sigma^*(i_2)}) + \sum_{i=i_0+1}^{i_2} c(p_i, q_{i-1}). \quad (4.2) \end{aligned}$$

Because the cost is supposed to be increasing with respect to the distance, one finds that $c(p_{i_0}, q_{\sigma^*(i_2)}) \leq c(p_{i_2}, q_{\sigma^*(i_2)})$, so that (4.2) implies:

$$c(p_{i_0}, q_{\sigma^*(i'_2)}) + \sum_{i=i_0+1}^{\sigma^*(i'_2)} c(p_i, q_{i-1}) + \sum_{i=\sigma^*(i'_2)+1}^{i_2-1} c(p_i, q_i) \leq \sum_{i=i_0+1}^{i_2} c(p_i, q_{i-1}),$$

and then:

$$\begin{aligned} c(p_{i_0}, q_{\sigma^*(i'_2)}) + \sum_{i=i_0+1}^{\sigma^*(i'_2)} c(p_i, q_{i-1}) + \sum_{i=\sigma^*(i'_2)+1}^{i_2-1} c(p_i, q_i) + \sum_{i=i_2+1}^{i_0+k_0} c(p_i, q_{i-1}) \\ \leq \sum_{i=i_0+1}^{i_0+k_0} c(p_i, q_{i-1}). \quad (4.3) \end{aligned}$$

According to Assumption (1), $I_{\sigma^*(i'_2)-i_0}^p(i_0) \geq 0$, so that:

$$\sum_{i=i_0}^{\sigma^*(i'_2)} c(p_i, q_i) \leq c(p_{i_0}, q_{\sigma^*(i'_2)}) + \sum_{i=i_0}^{\sigma^*(i'_2)-1} c(p_{i+1}, q_i).$$

Combining these last inequalities with (4.3) one finds that:

$$\sum_{i=i_0}^{i_0+k_0} c(p_i, q_i) \leq c(p_{i_0}, q_{i_0+k_0}) + \sum_{i=i_0}^{i_0+k_0-1} c(p_{i+1}, q_i),$$

which contradicts Assumption (3).

We have then shown that neither demand nor supply points located between p_{i_0} and $q_{i_0+k_0+1}$ can be matched with points located outside this interval. The set \mathcal{S}_{i_0} is then stable by an optimal transport plan. According to Assumption (1–2), no nesting can occur in \mathcal{S}_{i_0} . The result follows. \square

4.2. The unbalanced case. In this section, we show that Theorem 2.7 still holds in the unbalanced case.

Observe first that none of the points p_j , $i_0 + 1 \leq j \leq i_0 + k_0$ can remain unmatched in an optimal transport plan. Indeed, assume on the contrary that there exists ℓ in $\{i_0 + 1, \dots, i_0 + k_0\}$ such that p_ℓ is unmatched in the optimal transport plan σ^* . According to Lemma 3.3

$$\sum_{i=i_0}^{\ell-1} c(p_i, q_i) > \sum_{i=i_0}^{\ell-1} c(p_{i+1}, q_i).$$

Therefore we cannot have $\sigma^*(i) = i$ for $i = i_0, \dots, \ell - 1$: otherwise it would be possible to rematch all the points q_i in this interval to their right neighbors and reduce the cost. Hence, as the point p_ℓ is unmatched and isolated, there exists m in $\{i_0, \dots, \ell - 1\}$ such that $(\sigma^*)^{-1}(m) < i_0$. Choose m to be the greatest value of the index with this property and observe that we have $\sigma^*(i) = i$ for all i in the (possibly empty) interval $m + 1 \leq i \leq \ell - 1$. Now, since g is an increasing function,

$$c(p_{(\sigma^*)^{-1}(m)}, q_m) + \sum_{i=m+1}^{\ell-1} c(p_i, q_i) > c(p_{i_0}, q_m) + \sum_{i=i_0}^{\ell-1} c(p_i, q_i) - \sum_{i=i_0}^m c(p_i, q_i)$$

Using again Eq. (3.1) of Lemma 3.3, one deduces from this last inequality that

$$c(p_{(\sigma^*)^{-1}(m)}, q_m) + \sum_{i=m+1}^{\ell-1} c(p_i, q_i) > c(p_{i_0}, q_m) + \sum_{i=i_0}^{\ell-1} c(p_{i+1}, q_i) - \sum_{i=i_0}^m c(p_i, q_i).$$

It follows from this and from $I_{m-i_0}^p(i_0) \geq 0$ that

$$\begin{aligned} c(p_{(\sigma^*)^{-1}(m)}, q_m) + \sum_{i=m+1}^{\ell-1} c(p_i, q_i) &> c(p_{i_0}, q_m) + \sum_{i=i_0}^{m-1} c(p_{i+1}, q_i) \\ &+ \sum_{i=m}^{\ell-1} c(p_{i+1}, q_i) - \sum_{i=i_0}^m c(p_i, q_i) \\ &\geq \sum_{i=m}^{\ell-1} c(p_{i+1}, q_i). \end{aligned}$$

In other words, it is cheaper to match each q_i , $m \leq i \leq \ell - 1$, with its right neighbor p_{i+1} and to exclude $p_{(\sigma^*)^{-1}(m)}$ than to match each q_i with its neighbor p_i and to exclude p_ℓ . In all cases, the point p_ℓ cannot remain unmatched.

If the point p_{i_0} is matched in the transport plan σ^* , then we can conclude by the already proved first part of Theorem 2.7 that $\sigma^*(i) = i - 1$ for $i = i_0 + 1, \dots, i_0 + k_0$ (according to Lemma 2.4 unmatched points are isolated, the existence of an unmatched p_j outside of $[p_{i_0}, q_{i_0+k_0}]$ has no consequence on this result).

Now, assume that p_{i_0} remains unmatched and that there exists m in $i = i_0, \dots, i_0 + k_0 - 1$ such that $(\sigma^*)^{-1}(m) \neq m + 1$. Since p_{i_0} is isolated, and since matched points are either neighbors or separated by more than $2k_0$ points, $(\sigma^*)^{-1}(m) > i_0 + k_0$. One can assume without loss of generality that m is the largest index in $\{i_0, \dots, i_0 + k_0 - 1\}$ satisfying $(\sigma^*)^{-1}(m) > i_0 + k_0$. Now, because of the rule of three, we know that $|p_{i_0+k_0} - q_{i_0+k_0-1}| < |p_{i_0+k_0} - q_{i_0+k_0}|$. Thus $(\sigma^*)^{-1}(i_0 + k_0 - 1) \leq i_0 + k_0$, otherwise there would exist q_j , with $j \geq i_0 + k_0$ such that $|p_{i_0+k_0} - q_j| < |p_{i_0+k_0} - q_{i_0+k_0-1}|$ (again, because of the rule of three), which contradicts the previous inequality. It follows that $m < i_0 + k_0 - 1$. Two cases can occur: either $(\sigma^*)(i) = i$ for all i in $\{m+1, \dots, i_0 + k_0\}$, or there exists one (and only one) supply p_k in $\{m+1, i_0 + k_0\}$ such that $\sigma^*(k) > i_0 + k_0$. This cannot happen for two different supplies in $\{m+1, i_0 + k_0\}$, otherwise there would be another demand q_ℓ between these supplies such that $(\sigma^*)^{-1}(\ell) > i_0 + k_0$.

In the first case, thanks to equation (3.2)

$$\begin{aligned} c(q_m, p_{(\sigma^*)^{-1}(m)}) + \sum_{i=m+1}^{i_0+k_0} c(p_i, q_i) &> c(q_m, p_{(\sigma^*)^{-1}(m)}) + \sum_{i=m}^{i_0+k_0-1} c(p_{i+1}, q_i) \\ &> c(q_{i_0+k_0}, p_{(\sigma^*)^{-1}(m)}) + \sum_{i=m}^{i_0+k_0-1} c(p_{i+1}, q_i), \end{aligned}$$

which contradicts the optimality of σ^* .

In the second case, since $I_{k_0}^p(i_0) < 0$,

$$\begin{aligned} c(q_m, p_{(\sigma^*)^{-1}(m)}) + \sum_{i=m+1}^{k-1} c(p_i, q_i) + c(p_k, q_{\sigma^*(k)}) &> c(q_m, p_{(\sigma^*)^{-1}(m)}) + c(p_k, q_{\sigma^*(k)}) \\ &+ c(p_{i_0}, q_{i_0+k_0}) + \sum_{i=i_0}^{i_0+k_0-1} c(p_{i+1}, q_i) - \sum_{i=i_0}^m c(p_i, q_i) - \sum_{i=k}^{i_0+k_0} c(p_i, q_i). \end{aligned}$$

Now, since $I_{m-i_0}^p(i_0) \geq 0$ and $I_{i_0+k_0-k}^p(k) \geq 0$, this inequality yields

$$\begin{aligned} c(q_m, p_{(\sigma^*)^{-1}(m)}) + \sum_{i=m+1}^{k-1} c(p_i, q_i) + c(p_k, q_{\sigma^*(k)}) &> c(q_m, p_{(\sigma^*)^{-1}(m)}) \\ &+ c(p_k, q_{\sigma^*(k)}) - c(p_k, q_{i_0+k_0}) + c(p_{i_0}, q_{i_0+k_0}) - c(p_{i_0}, q_m) + \sum_{i=m}^{k-1} c(p_{i+1}, q_i). \end{aligned}$$

The two differences that appear in the right-hand side are positive so that

$$\begin{aligned} c(q_m, p_{(\sigma^*)^{-1}(m)}) + \sum_{i=m+1}^{k-1} c(p_i, q_i) + c(p_k, q_{\sigma^*(k)}) &\geq c(q_{\sigma^*(k)}, p_{(\sigma^*)^{-1}(m)}) \\ &+ \sum_{i=m}^{k-1} c(p_{i+1}, q_i), \end{aligned}$$

which also contradicts the optimality of σ^* .

By symmetry, the theorem remains valid in the case where $I_{k_0}^q(i'_0) < 0$ instead of $I_{k_0}^p(i_0) < 0$. \square

5. Algorithm. In this section, we derive from Theorem 2.7 a simple algorithm to compute the optimal transport plan in the case of chains and give details about its implementation and complexity. For the sake of simplicity, we only consider the balanced case. The unbalanced case can be treated in the same way.

5.1. Computation of optimal transport plans for chains. The recursive use of the local matching indicators defined in Definition 2.6 is on the basis of the next algorithm.

Algorithm: Set $\mathcal{P} = \{p_1, \dots, p_N, q_1, \dots, q_N\}$, $\ell^p = (1, \dots, N)$, $\ell^q = (1, \dots, N)$, and $k = 1$.

While $\mathcal{P} \neq \emptyset$ and $k \leq N - 1$ do

1. Compute $I_k^p(i)$ and $I_k^q(i')$ for $i = 1, \dots, N - k$ and $i' = 1, \dots, N - k - 1$.
2. Define

$$\mathcal{I}_k^p = \{i_0, 1 \leq i_0 \leq N - k, I_k^p(i_0) < 0\},$$

$$\mathcal{I}_k^q = \{i_0, 1 \leq i_0 \leq N - k - 1, I_k^q(i_0) < 0\},$$

and do

- (a) If $\mathcal{I}_k^p = \emptyset$ and $\mathcal{I}_k^q = \emptyset$, set $k = k + 1$.
- (b) Else do
 - for all i_0 in \mathcal{I}_k^p and for $i = i_0 + 1, \dots, i_0 + k$, do
 - define $\sigma^*(\ell_i^p) = \ell_{i-1}^q$,
 - remove $\{p_{\ell_i^p}, q_{\ell_{i-1}^q}\}$ from \mathcal{P} ,
 - remove ℓ_i^p and ℓ_i^q from ℓ^p and ℓ^q respectively.
 - for all i'_0 in \mathcal{I}_k^q and for $i = i'_0 + 1, \dots, i'_0 + k$, do
 - define $\sigma^*(\ell_i^q) = \ell_i^p$,
 - remove $\{p_{\ell_i^q}, q_{\ell_i^p}\}$ from \mathcal{P} ,
 - remove ℓ_i^q and ℓ_i^p from ℓ^p and ℓ^q respectively.
 - set $N = \frac{1}{2} \text{Card}(\mathcal{P})$, and rename the points in \mathcal{P} such that

$$\mathcal{P} = \{p_1, \dots, p_N, q_1, \dots, q_N\},$$

$$p_1 < q_1 < \dots < p_i < q_i < p_{i+1} < q_{i+1} < \dots < p_N < q_N.$$

- set $k = 1$.

If $k = N - 1$, for $i = 1, \dots, N$ set $\sigma^*(\ell_i^p) = \ell_i^q$.

An alternative algorithm consists in testing the sign of each $I_k^p(i)$ and $I_k^q(i')$ as soon as they have been computed and remove the corresponding pairs of points if a negative value is found. What follows also holds for this variant.

5.2. About the implementation and the complexity. The cost of the algorithm can be estimated through the number of additions and evaluations of the cost function that are required to terminate the algorithm. These operations are only carried out in Step 1, when computing the indicators. This section aims at giving details about efficient ways to implement this step and about the complexity of the resulting procedure.

5.2.1. Implementation through a table of indicators. We first define a table that collect the values of indicators and then describe a way to update it. The aim of this structure is to avoid redundant computations. We present it in the balanced case (see (2.5)).

Consider a table of $N - 1$ lines, where the lines correspond to the values of the indicators of order k . The line k has $2N - 1 - 2k$ entries corresponding to the $N - k$ values of the indicators I_k^p and the $N - k - 1$ values of the indicators I_k^q . At the beginning of the algorithm, the table is empty and Step 1 consists in filling the line k of the table. Let us explain how to modify the table in case a negative indicator has been found.

Following the assumptions of Theorem 2.7, consider the case where all the indicators that have been computed currently are positive except the last one, denoted by $I_{k_0}^p(i_0) < 0$. According to Step 2b, k_0 pairs of supply and demand have to be matched and removed from the current list of points \mathcal{P} . Doing this, the table loses k_0 lines. In each line $k < N - k_0 - 1$ of the new table, only $\min(i_0 - k_0 + k, 2N - 1 - k) - \max(i_0 - k_0 - k, k + 1)$ values are not valid any more since the corresponding indicator involves points that have been modified. Other values are not affected by the withdrawal.

5.2.2. Bounds for the complexity . In the vein of the previous section, we assume up to now that all the numerical values computed during the algorithm are saved. In this framework and as in any assignment problem, the number of evaluations of the cost function cannot exceed $\frac{N(N+1)}{2}$. The most favorable case consists in finding a negative indicator at each step of the loop. In this case, all points are removed through indicators of order 1. This case requires $O(N)$ additions and evaluations of the cost function.

On the opposite, the worst case corresponds to the case where all the indicators are positive. In such a situation, no pairs are removed until the table is full. All possible transport costs $c(p_i, q_j)$ are computed. Consequently, this case requires $\frac{N(N+1)}{2}$ evaluations of the cost function. The number of additions is also bound by $O(N^2)$ as stated in the next theorem.

THEOREM 5.1. *Denote by $C^+(N)$ the number of additions required to compute an optimal transport plan between N supplies and N demands with the algorithm of Section 5.1. One has:*

$$C^+(N) \leq 3N^2 - 6N.$$

The proof of this result is given in Appendix.

5.2.3. Empirical complexity. In order to estimate the complexity of our algorithm, we have applied it to an increasing number N of pairs of points. For a fixed value of N , 100 samples of points have been chosen randomly in $[0, 1]$, and the mean of the number of additions and evaluations of g has been computed. The results are shown in Figures 5.1-5.2. They show that the less concave the cost function is, the more accurate the bound $O(N^2)$ is.

6. Extensions. In this section, we explain a way to adapt our result to tackle optimal transport problems on the circle and in the case of integer masses.

6.1. The case of the circle. In many applications, see [7] and the references therein, a rapid computation of optimal transport plans on the circle is required. This section is devoted to the adaptation of our main result to this case.

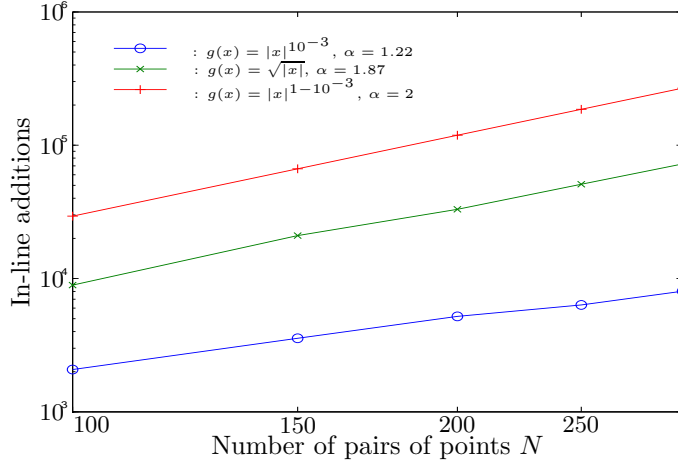


FIG. 5.1. Number of in-line additions with respect to the number of pairs, for various cost functions. The number α is the slope of the log-log graphs.

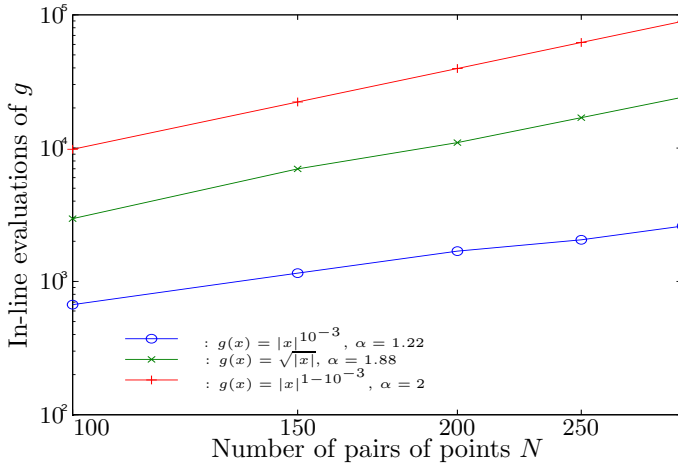


FIG. 5.2. Number of in-line evaluations of g with respect to the number of pairs, for various cost functions. The number α is the slope of the log-log graphs.

6.1.1. Notations and result. Define a unit circle \mathcal{T} as the segment $[0, 1]$ with identified endpoints. The positive direction on the segment is identified with the clockwise direction on the circle. Consider two integers $N \leq M$ and $P = (p_i)_{i=1, \dots, M}$ and $Q = (q_i)_{i=1, \dots, N}$, two sets of points in \mathcal{T} that represent respectively unitary supply and unitary demand locations. As before, we consider the problem of finding

$$\min_{\sigma \in \Sigma_M} C(\sigma),$$

where Σ_M is the set of all permutations of $\{1, \dots, M\}$ and where

$$C(\sigma) = \sum_{i/\sigma(i) \in \{1, \dots, N\}} c(p_i, q_{\sigma(i)}). \quad (6.1)$$

In this equation, the cost c is defined as follows.

DEFINITION 6.1. *The cost function in (6.1) is defined on $[0, 1]$ by*

$$c(p, q) = \min\{g(|p - q|), g(1 - |p - q|)\}$$

with $0 \leq p, q < 1$, where $g(\cdot)$ is a concave non-decreasing real-valued function of a real positive variable such that $g(0) := \lim_{x \rightarrow 0} g(x) \geq -\infty$.

Following Aggarwal *et al.* [1], we remark that any two distinct points p, q in \mathcal{T} split the circle into two arcs, one going clockwise from p to q and denoted $cw(p, q)$ and the other going counterclockwise from p to q and denoted $ccw(p, q)$. We call the shortest of the two arcs the *path between p and q* and denote it by $x(p, q)$: if $0 < p - q < \frac{1}{2}$ or if $\frac{1}{2} < q - p$, then $x(p, q) = ccw(p, q)$ and otherwise $x(p, q) = cw(p, q)$, including the case $|p - q| = \frac{1}{2}$.

Assume that the endpoints are *not* included into arcs.

DEFINITION 6.2. *We call two pairs (p, q) and (p', q') crossed when p', q' belong to different arcs in the splitting defined by p, q and nested otherwise [1]. Two nested pairs (p, q) and (p', q') are called weakly crossed if either $p = q'$ or $p' = q$ and properly nested if this is not the case.*

For the sake of sufficiency, we give below an equivalent of the non-crossing rule and its proof in the case of the circle.

LEMMA 6.3. *Let $p \neq q, p' \neq q'$ and suppose that (p, q) and (p', q') are crossed (possibly weakly). Then*

$$c(p, q) + c(p', q') \geq c(p, q') + c(p', q).$$

This holds with a strict inequality if the function g is strictly monotone.

This is equivalent to the following statement.

COROLLARY 6.4 (Non-crossing rule). *Let $p \neq q, p' \neq q'$ and suppose that*

$$c(p, q) + c(p', q') < c(p, q') + c(p', q). \quad (6.2)$$

Then (p, q) and (p', q') are properly nested. The same conclusion holds if (6.2) is replaced by analogous non-strict inequality but the function g is strictly monotone.

Proof (of Lemma 6.3): We apply a minor modification of the proof of [11, Lemma 2.1 (i)]. There are in total sixteen cases to consider, depending on whether $x(p, q) = cw(p, q)$ or $ccw(p, q)$, whether $x(p', q') = cw(p', q')$ or $ccw(p', q')$, and whether $x(p, q)$ contains p' or q' in the case of proper crossing or whether $p' = q$ or $p = q'$ in the case of weak crossing. However all these cases fall within one of the two following crossing patterns.

PATTERN I. Let $x(p, q) = cw(p, q), x(p', q') = ccw(p', q')$ and either $q' \in x(p, q)$ (which implies $q \in x(p', q')$ because of the assumed crossing) or $p = q'$ in the case of weak crossing. Using homogeneity of the cost function (6.1) choose representatives of the four points in $[0, 1]$ such that $p \leq q' < q \leq p'$, where $0 \leq q' - p < q - p \leq \frac{1}{2}$ and $0 \leq p' - q < p' - q' \leq \frac{1}{2}$. Monotonicity of the function g implies

$$\begin{aligned} c(p, q) &= g(q - p) \geq g(q' - p) = c(p, q'), \\ c(p', q') &= g(p' - q') \geq g(p' - q) = c(p', q). \end{aligned}$$

It therefore follows that

$$c(p, q) + c(p', q') \geq c(p, q') + c(p', q). \quad (6.3)$$

The cases in which $p' \in x(p, q)$ or $p' = q$ and/or directions of the arcs are opposite can be treated similarly by interchanging ps and qs and/or primed and non-primed quantities: both of these transformations preserve the cost as defined in (6.1). Thus the present crossing pattern represents eight of sixteen possible cases.

PATTERN II. Assume that $x(p, q)$ and $x(p', q')$ both go clockwise and that $p' \in x(p, q)$ or $p' = q$. The four representatives in $[0, 1]$ may now be chosen such that $p < p' \leq q < q'$, where $0 < q - p \leq \frac{1}{2}$, $0 < q' - p' \leq \frac{1}{2}$, and $q' - p \leq 1$. Take $0 < t < 1$ such that

$$q - p = (1 - t)(q' - p) + t(q - p').$$

Subtracting both sides of this equality from the larger quantity $q + q' - p - p'$ gives

$$q' - p' = t(q' - p) + (1 - t)(q - p').$$

Concavity of the function g now implies that

$$\begin{aligned} g(q - p) &\geq (1 - t)g(q' - p) + t g(q - p'), \\ g(q' - p') &\geq t g(q' - p) + (1 - t)g(q - p'), \end{aligned}$$

from which

$$g(q - p) + g(q' - p') \geq g(q' - p) + g(q - p'). \quad (6.4)$$

Observe that since $0 < q - p \leq \frac{1}{2}$ and $0 < q' - p' \leq \frac{1}{2}$, one has $c(p, q) = g(q - p)$ and $c(p', q') = g(q' - p')$, whereas $g(q' - p) \geq c(p, q')$ and $g(p' - q) \geq c(p', q)$ by definition. Therefore (6.4) implies (6.3). The remaining cases are covered by interchanging ps and qs and/or primed and non-primed quantities.

Finally let g be strictly monotone. In this case, all the above arguments hold with strict inequalities and the contradiction is achieved even when Inequality (6.2) is non-strict. \square

Thanks to this adaptation of the non-crossing rule, one can also define chains in the case of the circle, and restrict oneself to the two cases where P and Q satisfy $M = N$ (*balanced case*) and

$$0 < p_1 < q_1 < \dots < p_i < q_i < p_{i+1} < q_{i+1} < \dots < p_N < q_N \leq 1, \quad (6.5)$$

or $M = N + 1$ (*unbalanced case*) and

$$0 < p_1 < q_1 < \dots < p_i < q_i < p_{i+1} < q_{i+1} < \dots < p_N < q_N < p_{N+1} \leq 1.$$

We still call the set $P \cup Q$ *balanced chain* in the first case and *unbalanced chain* in the second case.

6.1.2. The balanced case. In what follows for $r \in \mathbb{Z}$ and $1 \leq i \leq N$, the point p_i is also denoted by p_{i+rN} . The same notation is used for the points of Q .

In this setting, the indicators are defined as follows.

DEFINITION 6.5 (Local Matching Indicators of order k). *Given $0 < k \leq N - 1$, consider $2k + 2$ consecutive points in the chain (6.5) such that the path between the*

first and the last point of the chain is of length equal or smaller than $\frac{1}{2}$. If the first point is a supply p_i with $1 \leq i \leq N$, define

$$I_k^p(i) = c(p_i, q_{i+k}) + \sum_{\ell=0}^{k-1} c(p_{i+\ell+1}, q_{i+\ell}) - \sum_{\ell=0}^k c(p_{i+\ell}, q_{i+\ell}),$$

else denote the first point q_i with $1 \leq i \leq N$ and define

$$I_k^q(i) = c(p_{i+k+1}, q_i) + \sum_{\ell=1}^k c(p_{i+\ell}, q_{i+\ell}) - \sum_{\ell=0}^k c(p_{i+\ell+1}, q_{i+\ell}).$$

The condition $k \leq N-1$ guarantees that all the points used to define the indicators are distinct.

We can now give a version of Theorem 2.7 adapted to the circle case.

THEOREM 6.6 (Negative Local Matching Indicators of order k). *Let $k_0 \in \mathbb{N}$ with $1 \leq k_0 \leq N-1$ and $i_0 \in \mathbb{N}$ with $1 \leq i_0 \leq N$ (resp. $i'_0 \in \mathbb{N}$ with $1 \leq i'_0 \leq N$).*

Assume that

1. $I_k^p(i) \geq 0$ for $k = 1, \dots, k_0 - 1$, for all $i \in \mathbb{N}$, $1 \leq i \leq N$ such that $I_k^p(i)$ is defined,
2. $I_k^q(i') \geq 0$ for $k = 1, \dots, k_0 - 1$, for all $i' \in \mathbb{N}$, $1 \leq i' \leq N$, such that $I_k^q(i')$ is defined,
3. $I_{k_0}^p(i_0) < 0$ (resp. $I_{k_0}^q(i'_0) < 0$).

Then any permutation σ associated to an optimal transport plan satisfies $\sigma(i) = i - 1[N]$ for $i = i_0 + 1[N], \dots, i_0 + k_0[N]$ (resp. $\sigma(i) = i[N]$ for $i = i_0 + 1[N], \dots, i_0 + k_0[N]$).

In this Theorem, $i[N]$ denotes the rest of the integer division of i by N .

Proof: The proof of this result is similar to the one of Theorem 2.7. We consider the case where $I_{k_0}^p(i_0) < 0$. The case $I_{k_0}^q(i'_0) < 0$ can be treated the same way.

The proof consists again in proving that Assumptions (1–3) imply that neither demand nor supply points located in $x(p_{i_0}, q_{i_0+k_0+1})$ can be matched with points located outside this interval, i.e. that the set $\mathcal{S}_{i_0} = \{p_i, p_i \in x(p_{i_0}, q_{i_0+k_0+1})\} \cup \{q_i, q_i \in x(p_{i_0}, q_{i_0+k_0+1})\}$ is stable by an optimal transport plan. In this case, the result follows from Assumption (1–2). Note that because of the condition on the length of the path of the sequence considered in an indicator as defined in Definition 6.5, all indicators involving only points located in $x(p_i, q_{i+k})$ are defined as soon as $I_k^p(i)$ is defined.

Suppose that \mathcal{S}_{i_0} is not preserved by an optimal transport plan σ^* . Three cases can occur:

- a) There exists $i_1 \in \mathbb{N}$, such that $q_{\sigma^*(i_1)} \in x(p_{i_0}, q_{i_0+k_0})$ and $p_{i_0} \in x(p_{i_1}, q_{\sigma^*(i_1)})$ and there exists $i'_1 \in \mathbb{N}$, such that $p_{i'_1} \in x(q_{\sigma^*(i_1)}, p_{i_0+k_0})$ and $q_{i_0+k_0} \in x(p_{i'_1}, q_{\sigma^*(i'_1)})$.
- b) There exists $i_2 \in \mathbb{N}$, such that $p_{i_2} \in x(p_{i_0}, q_{i_0+k_0})$ and $p_{i_0} \in x(q_{\sigma^*(i_2)}, p_{i_2})$.
- c) There exists $i_2 \in \mathbb{N}$, such that $q_{\sigma^*(i_2)} \in x(p_{i_0}, q_{i_0+k_0})$ and $q_{i_0+k_0} \in x(q_{\sigma^*(i_2)}, p_{i_2})$.

Following the lines of the proof given in Section 4.1, one notes that all the lengths that are considered are smaller than $\frac{1}{2}$. Consequently $c(p, q) = g(|p - q|)$ for every pair (p, q) that is considered and Lemma 3.1 can be used together with the non-decreasing property of the cost. The result follows. \square

As it was the case for the real line, one can derive from this theorem an algorithm to compute an optimal plan. Actually, the only adaptation concerns the length of the

indicators, that must not exceed $\frac{1}{2}$ and the end of the algorithm: In the case of the circle, if all indicators are positive at one point, then only two possibilities can occur, each remaining supply being matched either with its left neighbor, or with its right one. Since a great part of the cost should be already computed at this step, this case does not damage the efficiency of the algorithm.

6.1.3. The unbalanced case. In the unbalanced case, the precomputation of chains can reveal one unmatched point, i.e. a chain containing only one point, say p_{i_0} . In such a case, no path between two matched points in an optimal transport plan contains p_{i_0} . Consequently, one can simply recast the problem on the real line by cutting the circle in p_{i_0} and apply the unbalanced version of the algorithm of Section 5.1.

Otherwise, we can show that Theorem 6.6 can be generalized to the unbalanced case when assuming that g is strictly monotone. Indeed, the proof of Section 4.2 is still valid provided that the lengths of the chains that are considered are equal or smaller than $\frac{1}{2}$.

6.2. The integer case. We present now a simple strategy to tackle the case where the masses are not unitary but integer. In this section, we denote by $C_{P,Q}$ the cost defined in (2.3) where P and Q still stand for the locations of the possibly non-unitary masses. In the same way, we still denote by σ^* the optimal transport plan.

Given $\varepsilon > 0$, the method we follow consists in scattering each non-unitary integer mass in unitary masses located in different places in an interval of size ε around the original position. We denote by P^ε and Q^ε the corresponding location of unitary demands and supply respectively, and by σ^ε the optimal transport plan associated to this new configuration. Note that σ^ε can be computed by the algorithms presented in this paper.

LEMMA 6.7. *For any $\eta > 0$, there exists ε such that:*

$$C_{P,Q}(\sigma^*) + \eta \geq C_{P,Q}(\sigma^\varepsilon) \geq C_{P,Q}(\sigma^*).$$

Proof: Note first that given a transport plan σ , the mapping $\varepsilon \mapsto C_{P^\varepsilon, Q^\varepsilon}(\sigma)$ is a continuous function. Consequently, there exists ε_0 such that: $\forall \varepsilon > 0, \varepsilon_0 > \varepsilon$,

$$C_{P,Q}(\sigma^*) + \frac{1}{2}\eta \geq C_{P^\varepsilon, Q^\varepsilon}(\sigma^*). \quad (6.6)$$

Moreover, because the set of permutations of Σ_M is finite, one can also assume that

$$\forall \sigma \in \Sigma_M, C_{P^\varepsilon, Q^\varepsilon}(\sigma) \geq C_{P,Q}(\sigma) - \frac{1}{2}\eta.$$

hence

$$C_{P^\varepsilon, Q^\varepsilon}(\sigma^\varepsilon) \geq C_{P,Q}(\sigma^\varepsilon) - \frac{1}{2}\eta. \quad (6.7)$$

Because $C_{P^\varepsilon, Q^\varepsilon}(\sigma^*) \geq C_{P^\varepsilon, Q^\varepsilon}(\sigma^\varepsilon)$, Inequalities 6.6 and 6.7 give rise to

$$C_{P,Q}(\sigma^*) + \eta \geq C_{P,Q}(\sigma^\varepsilon).$$

The second inequality follows from the optimality of σ^* . □

In the case there exists a unique optimal transport plan, it follows from this lemma that for ε small enough $\sigma^\varepsilon = \sigma^*$.

Acknowledgments. This work was started during the visit of JD and JS at the Observatoire de Nice made possible by ANR through grant ANR-07-BLAN-0235 OTARIE (<http://www.mccme.ru/~ansobol/otarie/>); AS thanks the Ministry of National Education of France for supporting his visit to the Observatoire de la Côte d’Azur, where part of this text was written, and to Russian Fund for Basic Research for the partial support via grant RFBR 07–01–92217-CNRS-a.

Appendix : proof of Theorem 5.1. Before proving Theorem 5.1, let us state some intermediate results. In what follows, we denote by $c_k^+(N)$ the number of additions required to achieve Step 1 of the algorithm for an arbitrary value of k .

LEMMA 6.8. *Keeping the previous notations, we have:*

$$c_k^+(N) \leq 3(2(N - k) - 1). \quad (6.8)$$

Proof: The proof of (6.8) in the case $k = 1$ is left as exercise for the reader. Suppose that $k > 1$. Consider for example $I_k^p(i)$ and recall that:

$$I_k^p(i) = c(p_i, q_{i+k}) + \sum_{\ell=0}^{k-1} c(p_{i+\ell+1}, q_{i+\ell}) - \sum_{\ell=0}^k c(p_{i+\ell}, q_{i+\ell}). \quad (6.9)$$

The first term of this formula does not require any addition and most of the other terms have already been computed during the previous steps. Indeed, the first sum has been computed to evaluate $I_{k-1}^q(i)$ and the second one has been computed to evaluate $I_{k-1}^p(i)$. It remains to add $c(p_{i+k}, q_{i+k-1})$ to it to compute the last sum of (6.9). Since at given order k at most $2(N - k) - 1$ indicators have to be computed, the result follows. \square

We now consider the number of operations required between the beginning of the algorithm and the first occurrence of Step 2.2b.

LEMMA 6.9. *The operations required by the algorithm between its beginning and the first occurrence of Step 2.2b can be achieved with $\ell_{k_0}^+(N) := 3k_0(2N - k_0 - 2)$ additions, where k_0 denote the current value of k when Step 2.2b occurs.*

Proof: Between the beginning of the algorithm and the first occurrence of Step 2.2b, only positive indicators have been computed, except for the current value of $k = k_0$. This means that Step 1 has been carried out for $k = 1, \dots, k_0$ since the beginning. The corresponding number of additions is bounded by $\sum_{k=1}^{k_0} c_k^+(N)$. Thanks to Lemma 6.8, the result follows. \square

Recall now that after Step 2.2b having been achieved, the parameter k is set to 1. The previous arguments consequently applies to evaluate the number of additions between two occurrences of Step 2.2b, i.e. between two withdrawals. In this way, one finds that this number is bounded by $\ell_{k'_0}^+(N')$, where N' and k'_0 are the current values of N and k at the last occurrence of Step 2.2b. Note that $\ell_{k'_0}^+(N')$ is a coarse upper bound because we are not considering the first occurrence of this step and a part of the indicators has already been computed as explained in Section 5.2.1.

We are now in the position to prove Theorem 5.1.

Proof (of Theorem 5.1): Let k_0, k_1, \dots, k_s be the successive orders at which the Step 2.2b of the algorithm is visited. Observe that some of these numbers can be

equal. Assume also that only one negative indicator was found at each of these orders, which is the worst case for complexity. As a consequence, $\sum_{i=0}^s k_i = N$, and the number of additions required for the whole algorithm is lower than

$$C^+ \leq \sum_{i=0}^s \ell_{k_i}^+ (N - \sum_{j=0}^{i-1} k_j),$$

where ℓ_k^+ is defined in Lemma 6.9. Using Lemma 6.9, we compute

$$\begin{aligned} C^+ &\leq \sum_{i=0}^s 3k_i (2(N - \sum_{j=0}^{i-1} k_j) - k_i - 2) \\ &= \sum_{i=0}^{s-1} 3k_i (2(N - \sum_{j=0}^{i-1} k_j) - k_i - 2) + 3k_s (2(N - \sum_{j=0}^{s-1} k_j) - k_s - 2) \\ &= \sum_{i=0}^{s-1} 3k_i (2(N - \sum_{j=0}^{i-1} k_j) - k_i - 2) + 3(N - \sum_{j=0}^{s-1} k_j)(N - \sum_{j=0}^{s-1} k_j - 2) \\ &= 3N^2 - 6N - 6 \sum_{i=0}^{s-1} \sum_{j=0}^{i-1} k_i k_j - 3 \sum_{i=0}^{s-1} k_i^2 + 3(\sum_{j=0}^{s-1} k_j)^2 \\ &= 3N^2 - 6N. \end{aligned}$$

□

REFERENCES

- [1] A. Aggarwal, A. Bar-Noy, S. Khuller, D. Kravets, and B. Schieber. Efficient minimum cost matching using quadrangle inequality. In *Annual Symposium on Foundations of Computer Science*, volume 33, pages 583–583. IEEE Computer Society Press, 1992.
- [2] L. Ambrosio, L. Caffarelli, Y. Brenier, G. Buttazzo, and C. Villani. *Optimal Transportation and Applications*, volume 1813 of *Lecture Notes in Mathematics*. Springer, Berlin / Heidelberg, mathematics and statistics edition, 2003.
- [3] D. Bertsekas. Auction algorithms for network flow problems: A tutorial introduction. *Computational Optimization and Applications*, 1(1):7–66, 1992.
- [4] Y. Brenier. The least action principle and the related concept of generalized flows for incompressible perfect fluids. *Journal of the American Mathematical Society*, 2(2):225–255, 1989.
- [5] R. Burkard, M. Dell’Amico, and S. Martello. *Assignment Problems*. SIAM, 2008.
- [6] M. J. P. Cullen and R. J. Purser. Properties of the lagrangian semigeostrophic equations. *Journal of the Atmospheric Sciences*, 46(17):2684–2697, 1989.
- [7] J. Delon, J. Salomon, and A. Sobolevski. Fast transport optimization for monge costs on the circle. *SIAM Journal on Applied Mathematics*, 70(7):2239–2258, 2010.
- [8] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.
- [9] A. Lachapelle, J. Salomon, and G. Turinici. Computation of mean field equilibria in economics. *M3AS*, 20(4):567–588, 2010.
- [10] R. J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.
- [11] R. J. McCann. Exact solutions to the transportation problem on the line. *Proceedings: Mathematical, Physical and Engineering Sciences*, pages 1341–1380, 1999.
- [12] G. Monge. *Mémoire sur la théorie des déblais et des remblais*. Histoire de l’Académie Royale des Sciences, 1781.

- [13] S. Rachev. The Monge–Kantorovich mass transference problem and its stochastic applications. *Theory of Probability and its Applications*, 29:647, 1985.
- [14] C. Villani. *Topics in optimal transportation*. American Mathematical Society, 2003.
- [15] C. Villani. *Optimal transport: old and new*. Springer Verlag, 2008.