



**HAL**  
open science

# Minimum-weight perfect matching for non-intrinsic distances on the line

Julie Delon, Julien Salomon, Andrei Sobolevski

► **To cite this version:**

Julie Delon, Julien Salomon, Andrei Sobolevski. Minimum-weight perfect matching for non-intrinsic distances on the line. 2011. hal-00564173v1

**HAL Id: hal-00564173**

**<https://hal.science/hal-00564173v1>**

Preprint submitted on 8 Feb 2011 (v1), last revised 26 Mar 2011 (v2)

**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.

# MINIMUM-WEIGHT PERFECT MATCHING FOR NON-INTRINSIC DISTANCES ON THE LINE

JULIE DELON, JULIEN SALOMON, AND ANDREI SOBOLEVSKI

ABSTRACT. We consider a minimum-weight perfect matching problem on the line and establish a “bottom-up” recursion relation for weights of partial minimum-weight matchings.

## 1. INTRODUCTION

For a given undirected graph whose edges have nonnegative weights, the *minimum weight perfect matching problem* consists in finding a perfect matching of vertices, i.e., a set of disjoint edges such that each vertex belongs to one edge and the sum of edge weights is minimal. In this paper we treat the particular case where all vertices belong to a line and edge weights are related to distances between the vertices.

A bipartite version of this problem reduces to transport optimization for cost functions of concave type. The latter problem received thorough treatment in the measure-theoretic setting in [7]; similar problems have also long been considered in the algorithmics literature for the particular case of the distance  $|x - y|$ , assuming that the measures are discrete [1, 6, 8]. An algorithm for a general weight of a concave type has been proposed recently in [3, 4].

In the present note we consider the minimum-weight perfect matching problem on a complete graph, which is not bipartite. However, for the class of weight functions generated by distances, the bipartite and non-bipartite problems turn out to be essentially equivalent due to a “no-crossing” property of the optimal matching. This observation, which we owe to S. Nechaev, allows to employ the theory developed for the bipartite case in [3, 4] to the non-bipartite matching.

The main contribution of the present note is a specific “bottom-up” recursion relation (6) for partial minimum-weight matchings. This recursion provides a new perspective on the construction of the two latter references. The note is organized as follows. Section 2 is a review of the basic construction of metrics on the real line. In Section 3 we recall the problem of minimum-weight perfect matching and cite a few useful results about the structure of its solution. Section 4 contains the key technical result, a kind of localization principle for minimum-weight matching. Finally in Section 5 we derive the recursive relation for partial minimum-weight matchings.

It is our pleasure to thank the organizers of the conference *Optimization and Stochastic Methods for Spatially Distributed Information* (St Petersburg, EIMI, May 2010), where an earlier version of this work was presented.

---

Supported by Agence Nationale de la Recherche under project ANR07-01-0235 OTARIE and the Russian Foundation for Basic Research under project 07-01-92217-CNRSL-a.

## 2. INTRINSIC AND NON-INTRINSIC DISTANCES ON THE LINE

Recall the usual axioms for a distance  $d(\cdot, \cdot)$  on the real line  $\mathbf{R}$ : for all  $x, y, z \in \mathbf{R}$ ,

- (D1)  $d(x, y) \geq 0$  with  $d(x, y) = 0$  iff  $x = y$ ;
- (D2)  $d(x, y) = d(y, x)$ ;
- (D3)  $d(x, y) + d(y, z) \geq d(x, z)$ .

These axioms are satisfied by the distance  $d(x, y) = |x - y|$  as well as by any distance of the form

$$(1) \quad d_g(x, y) = g(|x - y|),$$

where  $g(\cdot)$  is a nonnegative concave function defined for all  $x \geq 0$  such that  $g(x) = 0$  iff  $x = 0$ . Here concavity means that  $g(\lambda x + (1 - \lambda)y) \geq \lambda g(x) + (1 - \lambda)g(y)$  for all  $x, y \geq 0$  and all  $\lambda \in [0, 1]$ . Note that the distance  $d_g$  is *homogeneous* with respect to shifts:

$$(D4) \quad d_g(x + t, y + t) = d_g(x, y) \text{ for all } x, y, t \in \mathbf{R}.$$

Conversely, any distance satisfying the axioms (D1)–(D4) has the form (1) for a suitable nonnegative concave function  $g$ : indeed, take  $g(x) = d(0, x)$  for  $x \geq 0$  and check that concavity follows from (D3).

An example of this construction is the distance  $|x - y|^\alpha$  with  $0 \leq \alpha \leq 1$ , where

$$|x - y|^\alpha = \begin{cases} 0, & x = y, \\ 1, & x \neq y \end{cases}$$

is the “discrete distance”; here we are mostly interested in the case  $0 < \alpha < 1$ .

An important property of a distance  $d(\cdot, \cdot)$  is whether it is *intrinsic*. To recall the corresponding definition, take two distinct points  $x, y \in \mathbf{R}$  and connect them with parameterized curves taking values  $x(t)$  in  $\mathbf{R}$ , i.e., suppose that  $0 \leq t \leq 1$ ,  $x(0) = x$ , and  $x(1) = y$ . By the triangle inequality,

$$(2) \quad d(x, y) \leq \inf \sum_{0 \leq i < N} d(x(t_i), d(x(t_{i+1}))),$$

where the infimum is taken over all curves connecting  $x$  to  $y$  and all meshes  $0 = t_0 < t_1 < t_2 < \dots < t_N = 1$  with  $N \geq 1$ . The distance  $d$  is called intrinsic if (2) is an equality for all  $x, y$  (see, e.g., [2]).

The distance  $|x - y|$  and its scalar multiples are the only homogeneous intrinsic distances in  $\mathbf{R}$ . For distances  $|x - y|^\alpha$  with  $0 < \alpha < 1$ , connecting curves have infinite length unless  $x = y$ , and therefore these distances are not intrinsic.

The intuition behind this notion is that whereas the geometry of  $\mathbf{R}$  equipped with an intrinsic distance is fully determined by the distance, the same line  $\mathbf{R}$  with a non-intrinsic distance should be viewed as embedded in an auxiliary space of a larger dimension, and geometry on  $\mathbf{R}$  is induced by that in the embedding space. According to the Assouad embedding theorem (see, e.g., [5]), for the distance  $|x - y|^\alpha$  on  $\mathbf{R}$  the embedding space has dimension of the order  $1/\alpha$  for small  $\alpha$ .

## 3. MINIMUM-WEIGHT PERFECT MATCHINGS

Consider an even number of points  $x_1 < x_2 < \dots < x_{2n}$  on the real line  $\mathbf{R}$  equipped with a distance  $d$  and look for a *minimum-weight perfect matching* in a complete graph  $K_{2n}$  on these points, where the weight of an edge connecting  $x_i$

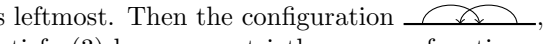
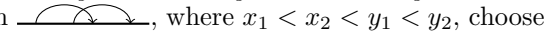
to  $x_j$  is  $d(x_i, x_j)$ . It is convenient to represent  $\mathbf{R}$  with a horizontal interval and use arcs in the upper halfplane to show the edges of the graph  $K_{2n}$ .

The following lemma, adapted from McCann [7, Lemma 2.1], allows to describe the structure of minimum-weight matchings.

**Lemma 1.** *Suppose the distance has the form  $d_g$  (1) with a strictly concave function  $g$  (i.e.,  $g(\lambda x + (1 - \lambda)y) > \lambda g(x) + (1 - \lambda)g(y)$  for all  $x, y \geq 0$ ,  $x \neq y$ , and  $0 < \lambda < 1$ ). Then the inequality*

$$(3) \quad d_g(x_1, y_1) + d_g(x_2, y_2) \leq d_g(x_1, y_2) + d_g(x_2, y_1)$$

*implies that the intervals connecting  $x_1$  to  $y_1$  and  $x_2$  to  $y_2$  are either disjoint or one of them is contained in the other.*

*Proof.* Let arcs representing the matching be directed from  $x$ 's to  $y$ 's and assume, without loss of generality, that  $x_1$  is leftmost. Then the configuration , where  $x_1 < x_2 < y_1 < x_2$ , cannot satisfy (3) because a strictly concave function  $g$  must be strictly growing and therefore  $d_g(x_1, y_2) < d_g(x_1, y_1)$  and  $d_g(x_2, y_1) < d_g(x_2, y_2)$ . To rule out configuration , where  $x_1 < x_2 < y_1 < y_2$ , choose  $0 < \lambda < 1$  such that

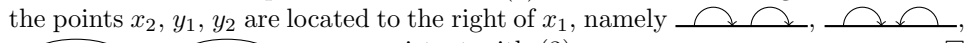
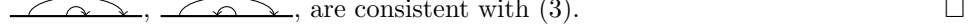
$$y_1 - x_1 = (1 - \lambda)(y_2 - x_1) + \lambda(y_1 - x_2).$$

Then, since  $(y_1 - x_1) + (y_2 - x_2) = (y_2 - x_1) + (y_1 - x_2)$ , we obtain

$$y_2 - x_2 = \lambda(y_2 - x_1) + (1 - \lambda)(y_1 - x_2)$$

and can further use the strict concavity of  $g$  to get

$$\begin{aligned} d_g(x_1, y_1) &> (1 - \lambda)d_g(x_1, y_2) + \lambda d_g(x_2, y_1), \\ d_g(x_2, y_2) &> \lambda d_g(x_1, y_2) + (1 - \lambda)d_g(x_2, y_1). \end{aligned}$$

The sum of these inequalities contradicts (3). All the other configurations where the points  $x_2, y_1, y_2$  are located to the right of  $x_1$ , namely , , are consistent with (3).  $\square$

It is possible to show that conversely, for an arbitrary bivariate function  $d(x, y)$  the property of Lemma 1 together with homogeneity (D4) imply that  $d$  has the form (1) for a suitable concave function  $g$  [7, Lemma B4].

We call a matching *nested* if, for any two arcs  $(x_i, x_j)$  and  $(x_{i'}, x_{j'})$  that are present in the matching, the corresponding intervals in  $\mathbf{R}$  are either disjoint or one of them is contained in the other.

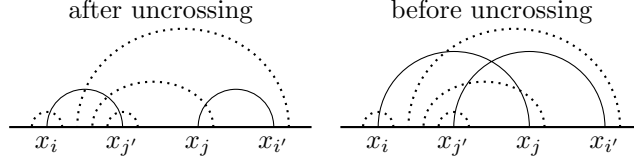
**Theorem 2** ([1, 7]). *A minimum-weight matching is nested.*

*Proof.* This statement essentially follows from Lemma 1, because whenever two arcs  $(x_i, x_j)$  and  $(x_{i'}, x_{j'})$  are crossed, the sum of distances corresponding to the “uncrossed” arcs  $(x_i, x_{j'})$ ,  $(x_{i'}, x_j)$  must be smaller according to (3).

However uncrossing may introduce new crossings with other arcs, and it remains to be checked that it does decrease the total number of crossings. This is done using an argument adapted from [1, Lemma 1].

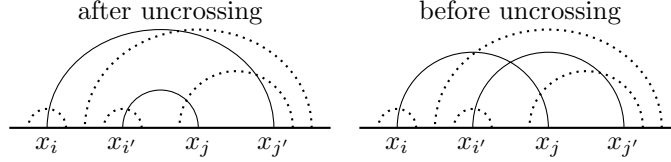
Let an arc  $(x_k, x_\ell)$  cross either of the arcs  $(x_i, x_{j'})$  or  $(x_{i'}, x_j)$  obtained after uncrossing the arcs  $(x_i, x_j)$  and  $(x_{i'}, x_{j'})$ ; we would like to prove that before uncrossing the arc  $(x_k, x_\ell)$  had the same number of crossings with  $(x_i, x_j)$  and  $(x_{i'}, x_{j'})$ . It suffices to consider the following two cases.

CASE  $x_i < x_{j'} < x_j < x_{i'}$ . Assume that  $(x_k, x_\ell)$  crosses  $(x_i, x_{j'})$ . Then  $(x_k, x_\ell)$ , shown as a dotted arc, may be situated in one of the following ways:



We see that the number of crossings between  $(x_k, x_\ell)$  and the other two arcs is the same before and after uncrossing. The case where  $(x_k, x_\ell)$  crosses  $(x_{i'}, x_j)$  is symmetrical.

CASE  $x_i < x_{i'} < x_j < x_{j'}$ . The following situations of the dotted arc  $(x_k, x_\ell)$  are possible:



Again the number of crossings between  $(x_k, x_\ell)$  and the other two arcs is the same before and after uncrossing.

It follows that each uncrossing removes exactly one crossing from the matching, and therefore any possible sequence of uncrossings leads in a finite number of steps to a nested matching with a strictly smaller weight. In other words, it suffices to look for the minimum-weight matching only among the nested ones.  $\square$

This result implies that the minimum-weight perfect matching problem is essentially bipartite, and it is indeed the bipartite setting that is considered in [1, 3, 4, 6–8]

**Corollary 3.** *In a minimum-weight perfect matching, points with even numbers are matched to points with odd numbers.*

*Proof.* Suppose on the contrary that  $x_i$  is matched to  $x_j$  with  $i$  and  $j$  both even or both odd. The interval connecting  $x_i$  to  $x_j$  contains an odd number  $|i - j| - 1$  of points, one of which has to be matched to a point outside and thus cause a crossing with  $(x_i, x_j)$ . Since a minimum-weight matching is nested, the result follows.  $\square$

Observe finally the following simple form of the Bellman optimality principle.

**Lemma 4.** *Any subset of arcs in a minimum-weight matching is itself a minimum-weight perfect matching on the set of endpoints of the arcs that belong to this subset.*

*Proof.* Indeed, if one could rematch these points achieving a smaller total weight, then the full original matching itself would not be minimum-weight: rematching the corresponding subset of arcs (and possibly uncrossing any crossed arcs that might result from the rematch) would give a matching with a strictly smaller weight.  $\square$

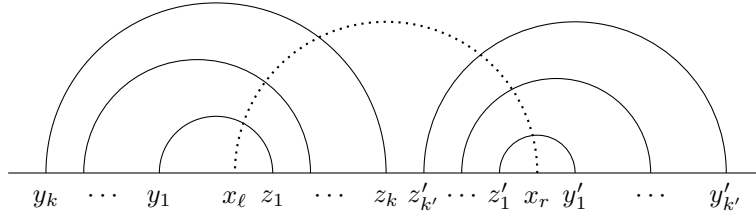


FIGURE 1. Notation used in the proof of Theorem 5. Note that in general  $x_\ell \leq z_1$  and  $z'_1 \leq x_r$ , but in the figure these pairs of points are shown to be distinct.

#### 4. PRESERVATION OF HIDDEN ARCS

Call an arc  $(x_i, x_j)$  in a nested matching *exposed* if it is not contained in any other arc, i.e., if there is no arc  $(x_{i'}, x_{j'})$  with  $x_i, x_j$  contained between  $x_{i'}$  and  $x_{j'}$ . We call all other arcs in a nested matching non-exposed or *hidden*. Intuitively, exposed arcs are those visible “from above” and hidden arcs are those covered with exposed ones.

Suppose  $X = \{x_i\}_{1 \leq i \leq 2n}$  with  $x_1 < x_2 < \dots < x_{2n}$  and  $X' = \{x'_{i'}\}_{1 \leq i' \leq 2n'}$  with  $x'_1 < x'_2 < \dots < x'_{2n'}$  be two sets such that  $x_{2n} < x'_1$ , i.e.,  $X'$  lies to the right of  $X$ . We will refer to minimum-weight perfect matchings on  $X$  and  $X'$  as *partial matchings* and to the minimum-weight perfect matching on  $X \cup X'$  as *joint matching*. The following result is closely related to properties of “local matching indicators” introduced and studied in [3, 4].

**Theorem 5.** *Whenever an arc  $(x_i, x_j)$  [respectively  $(x'_{i'}, x'_{j'})$ ] is hidden in the partial matching on  $X$  [respectively on  $X'$ ], it belongs to the joint optimal matching and is hidden there too.*

Observe that exposed arcs in partial matchings are generally not preserved in the joint matching: they may disappear altogether or become hidden.

*Proof.* By contradiction, assume that some of hidden arcs in the partial matching on  $X$  do not belong to the joint matching. Then there will be at least one exposed arc  $(x_\ell, x_r)$  in the partial matching on  $X$  such that some points  $x_i$  with  $x_\ell < x_i < x_r$  are connected in the joint matching to points outside  $(x_\ell, x_r)$ .

Indeed, if points inside every exposed arc  $(x_\ell, x_r)$  would be matched in the joint matching only among themselves, then by Lemma 4 their matching would be exactly the same as in the partial matching on  $X$ , and therefore all hidden arcs between  $x_\ell$  and  $x_r$  would be preserved in the joint matching.

Suppose  $(x_\ell, x_r)$  is the leftmost arc of the above kind. Denote all the points in the segment  $[x_\ell, x_r]$  that are connected in the joint matching to points on the left of  $x_\ell$  by  $z_1 < z_2 < \dots < z_k$ ; denote the opposite endpoints of the corresponding arcs by  $y_1 > y_2 > \dots > y_k$ , where the inequalities follow from the fact that the joint matching is nested. Likewise denote those points from  $[x_\ell, x_r]$  that are connected in the joint matching to points on the right of  $x_r$  by  $z'_1 > z'_2 > \dots > z'_{k'}$  and their counterparts in the joint matching by  $y'_1 < y'_2 < \dots < y'_{k'}$  (fig. 1).

Although  $k$  or  $k'$  may be zero, the number  $k + k'$  must be positive and even. Indeed, by Corollary 3 the segment  $[x_\ell, x_r]$  contains an even number of points and all of them must be matched in a perfect matching; removing from the joint

matching all arcs whose ends both lie in  $[x_\ell, x_r]$ , we are left with an even number of points that are matched outside this segment.

Let us now restrict our attention to the segment  $[x_\ell, x_r]$  and consider a matching that consists of the following arcs: those arcs of the joint matching whose both ends belong to  $[x_\ell, x_r]$ ; the arcs  $(z_1, z_2), \dots, (z_{2\kappa-1}, z_{2\kappa})$ , where<sup>1</sup>  $\kappa = \lfloor k/2 \rfloor$ ; the arcs  $(z'_2, z'_1), \dots, (z'_{2\kappa'}, z'_{2\kappa'-1})$ , where  $\kappa' = \lfloor k'/2 \rfloor$ ; and  $(z_k, z'_{k'})$  if both  $k$  and  $k'$  are odd.

Denote by  $W'$  the weight of this matching. By Lemma 4, it cannot be smaller than the weight  $W'_0$  of the restriction of the partial matching on  $X$  to  $[x_\ell, x_r]$ . For the total weight  $W$  of the joint matching on  $X \cup X'$  we thus have

$$(4) \quad W \geq W - W' + W'_0.$$

The right-hand side of (4) is represented in fig. 2 (a) in the case when both  $k$  and  $k'$  are odd. It is a sum of positive terms corresponding to the arcs of the joint matching outside  $[x_\ell, x_r]$  (not shown), the arcs of the partial matching on  $X$  inside  $[x_\ell, x_r]$  (not shown, with exception of  $(x_\ell, x_r)$  represented with a solid arc in the upper halfplane), the arcs of the joint matching having one end inside  $[x_\ell, x_r]$  and the other end outside this segment (solid arcs in the upper halfplane), and negative terms that come from subtraction of  $W'$  and correspond to the arcs connecting the  $z$  points (solid arcs in the lower halfplane).

We now show that by a suitable sequence of uncrossings the right-hand side of (4) can be further reduced to a matching whose weight is strictly less than  $W$ .

STEP 1. Note that the arcs  $(z_1, y_1)$  and  $(x_\ell, x_r)$ , shown in fig. 2 (a) with thick lines, are crossing. Therefore

$$d(y_1, z_1) + d(x_\ell, x_r) > d(y_1, x_\ell) + d(z_1, x_r).$$

Uncrossing these arcs gives the matching represented in fig. 2 (b) and strictly reduces the right-hand side of (4):

$$W > W - W' + W'_0 - d(y_1, z_1) - d(x_\ell, x_r) + d(y_1, x_\ell) + d(z_1, x_r).$$

Now the arcs  $(y_2, z_2)$  and  $(z_1, x_r)$  are crossing, so

$$d(y_2, z_2) + d(z_1, x_r) - d(z_1, z_2) > d(y_2, x_r)$$

and therefore

$$W > W - W' + W'_0 - d(y_1, z_1) - d(y_2, z_2) - d(x_\ell, x_r) + d(y_1, x_\ell) + d(z_1, z_2) + d(y_2, x_r).$$

The right-hand side of this inequality is represented in fig. 2 (c).

Repeating this step  $\kappa = \lfloor k/2 \rfloor$  times gives the inequality

$$\begin{aligned} W > W - W' + W'_0 - d(x_\ell, x_r) - \sum_{1 \leq i \leq 2\kappa} d(y_i, z_i) \\ + \sum_{1 \leq i \leq \kappa} d(z_{2i-1}, z_{2i}) + \sum_{1 \leq i \leq \kappa} d(y_{2i-1}, y_{2i-2}) + d(y_{2\kappa}, x_r), \end{aligned}$$

where in the rightmost sum  $y_0$  is defined to be  $x_\ell$ . Note that at this stage all arcs coming to points  $z_1, z_2, \dots$  from outside  $[x_\ell, x_r]$  are eliminated from the matching, except possibly  $(y_k, z_k)$  if  $k$  is odd.

---

<sup>1</sup> $\lfloor \xi \rfloor$  is the largest integer  $n$  such that  $n \leq \xi$ .

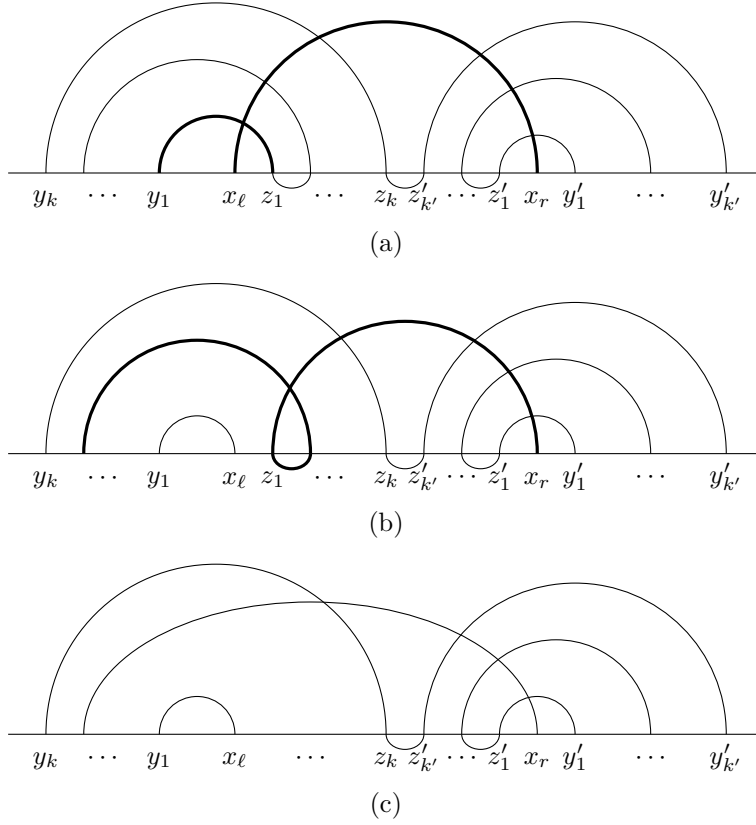


FIGURE 2. Step 1 of the proof (see explanation in the text).

STEP 2. It is now clear by symmetry that a similar reduction step can be performed on arcs going from  $z'_1, z'_2, \dots$  to the right. Repeating this  $\kappa' = \lfloor k'/2 \rfloor$  times gives the inequality

$$\begin{aligned}
 W > W - W' + W'_0 - d(x_\ell, x_r) - \sum_{1 \leq i \leq 2\kappa} d(y_i, z_i) - \sum_{1 \leq i' \leq 2\kappa'} d(z'_{i'}, y'_{i'}) \\
 &+ \sum_{1 \leq i \leq \kappa} d(z_{2i-1}, z_{2i}) + \sum_{1 \leq i \leq \kappa} d(y_{2i-1}, y_{2i-2}) \\
 &+ \sum_{1 \leq i' \leq \kappa'} d(z'_{2i'}, z'_{2i'-1}) + \sum_{1 \leq i' \leq \kappa'} d(y'_{2i'-2}, y'_{2i'-1}) + d(y_{2\kappa}, y'_{2\kappa'}),
 \end{aligned}$$

where  $y'_0 = x_r$ .

STEP 3. If  $k$  and  $k'$  are odd, we perform two more uncrossings shown in fig. 3.



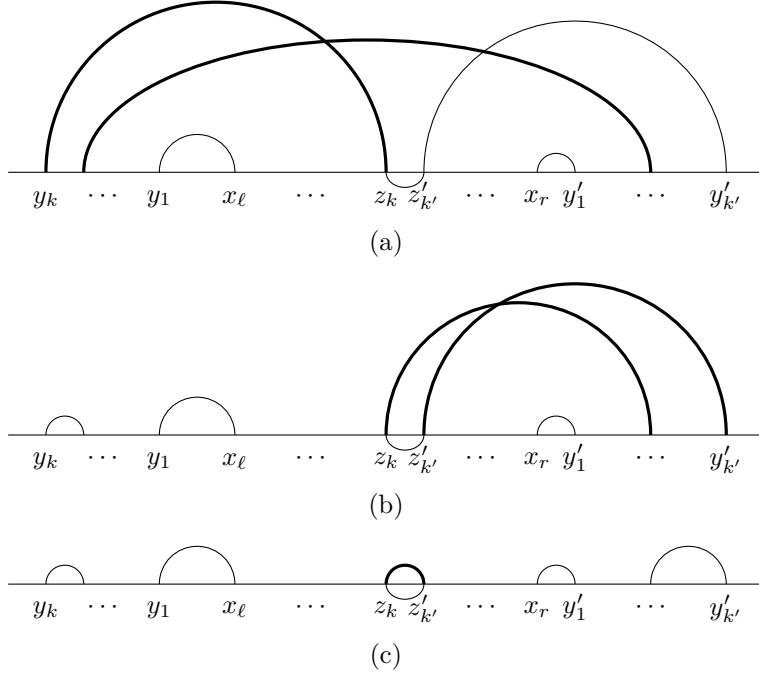


FIGURE 3. Step 3 of the proof. Note that in stage (c) the arc  $(z_k, z'_k)$  gives two contributions with positive and negative signs, which cancel out each other.

The final estimate for  $W$  has the form

$$\begin{aligned}
(5) \quad W &> W - W' + W'_0 - d(x_\ell, x_r) - \sum_{1 \leq i \leq k} d(y_i, z_i) - \sum_{1 \leq i' \leq k'} d(z'_{i'}, y'_{i'}) \\
&+ \sum_{1 \leq i \leq \kappa} d(z_{2i-1}, z_{2i}) + \sum_{1 \leq i' \leq \kappa'} d(z'_{2i'}, z'_{2i'-1}) + d(z_k, z'_k) \cdot [k, k' \text{ are odd}] \\
&+ \sum_{1 \leq i \leq \kappa} d(y_{2i-1}, y_{2i-2}) + \sum_{1 \leq i' \leq \kappa'} d(y'_{2i'-2}, y'_{2i'-1}) + d(y_k, y'_k) \cdot [k, k' \text{ are even}],
\end{aligned}$$

where notation such as  $[k, k' \text{ are odd}]$  means 1 if  $k, k'$  are odd and 0 otherwise.

The right-hand side of (5) contains four groups of terms: first,

$$W - \sum_{1 \leq i \leq k} d(y_i, z_i) - \sum_{1 \leq i' \leq k'} d(z'_{i'}, y'_{i'}),$$

corresponding to the joint matching without the arcs connecting points inside  $[x_\ell, x_r]$  to points outside this segment; second,

$$W' - \sum_{1 \leq i \leq \kappa} d(z_{2i-1}, z_{2i}) - \sum_{1 \leq i' \leq \kappa'} d(z'_{2i'}, z'_{2i'-1}) - d(z_k, z'_k) \cdot [k, k' \text{ are odd}],$$

which comes with a negative sign and corresponds to the arcs of the joint matching with both ends inside  $[x_\ell, x_r]$ , and cancels them from the total; third,

$$W'_0 - d(x_\ell, x_r),$$

with positive sign, which corresponds to the hidden arcs of the partial matching on  $X$  inside the exposed arc  $(x_\ell, x_r)$ , not including the latter; and finally the terms in the last line of (5), corresponding to the arcs matching  $x_\ell, x_r$ , and points  $y_1, \dots, y_k, y'_1, \dots, y'_{k'}$ , i.e., those points outside  $[x_\ell, x_r]$  that were connected in the joint matching to points inside this segment.

Gathering together contributions of these four groups of terms, we observe that all negative terms cancel out and what is left corresponds to a perfect matching with a weight strictly smaller than  $W$ , in which all arcs hidden by  $(x_\ell, x_r)$  in the partial matching on  $X$  are restored. There may still be some crossings caused by terms of the fourth group and *not* involving the hidden arcs in  $[x_\ell, x_r]$ ; uncrossing them if necessary gives a nested perfect matching whose weight is strictly less than that of the joint matching. This contradicts the assumption that the latter is the minimum-weight matching on  $X \cup X'$ . Therefore all hidden arcs in the partial matching on  $X$  (and, by symmetry, those in the partial matching on  $X'$ ) belong to the joint matching.  $\square$

5. RECURSION FOR MINIMUM WEIGHTS

In this section we show how to apply Theorem 5 to compute the minimum-weight perfect matching algorithmically. Let  $x_1 < x_2 < \dots < x_{2n}$  be a set of points on the real line  $\mathbf{R}$  equipped with a homogeneous distance  $d$  of the form (1). For indices  $i, j$  of opposite parity and such that  $i < j$ , let  $W_{i,j}$  be the weight of the minimum-weight perfect matching on the  $j - i + 1$  points  $x_i < x_{i+1} < \dots < x_j$ . It is convenient to organize weights  $W_{i,j}$  with  $i < j$  into a pyramidal table:

$$\begin{array}{cccccccc}
 & & & & & & & W_{1,2n} \\
 & & & & & & & \dots \\
 & & & & & & & \dots \\
 & & & & & & & W_{1,6} & W_{2,7} & W_{3,8} & W_{4,9} & \dots \\
 & & & & & & & W_{1,4} & W_{2,5} & W_{3,6} & W_{4,7} & W_{5,8} & \dots & W_{2n-3,2n} \\
 W_{1,2} & W_{2,3} & W_{3,4} & W_{4,5} & W_{5,6} & W_{6,7} & \dots & W_{2n-2,2n-1} & W_{2n-1,2n}
 \end{array}$$

**Theorem 6.** For all indices  $i, j$  of opposite parity with  $1 \leq i < j \leq 2n$ , weights  $W_{i,j}$  satisfy the following second-order recursion

$$(6) \quad W_{i,j} = \min [d(x_i, x_j) + W_{i+1,j-1}, W_{i,j-2} + W_{i+2,j} - W_{i+2,j-2}]$$

with “initial conditions”

$$(7) \quad W_{i,i-1} = 0, \quad W_{i+2,i-1} = -d(x_i, x_{i+1}).$$

*Proof.* By an abuse of notation, we will refer to the minimum-weight perfect matching on points  $x_r < x_{r+1} < \dots < x_s$  as the “matching  $W_{r,s}$ .”

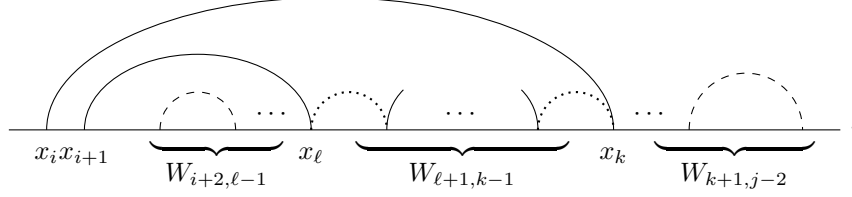
Consider first the matching that consists of the arc  $(x_i, x_j)$  and all arcs of the matching  $W_{i+1,j-1}$ , and observe that by Lemma 4 its weight  $d(x_i, x_j) + W_{i+1,j-1}$  is minimal among all matchings that contain  $(x_i, x_j)$ .

We now examine the meaning of the expression  $W_{i,j-2} + W_{i+2,j} - W_{i+2,j-2}$ . Denote the point connected in the matching  $W_{i,j-2}$  to  $x_i$  by  $x_k$  and the point connected to  $x_{i+1}$  by  $x_\ell$ . By Corollary 3, the pairs of indices  $i, k$  and  $i + 1, \ell$  both have opposite parity. Assume first that

$$(8) \quad x_{i+1} < x_\ell < x_k \leq x_{j-2}.$$

Applying Theorem 5 to the sets  $X = \{x_i, x_{i+1}\}$  and  $X' = \{x_{i+2}, \dots, x_{j-2}\}$  and taking into account parity of  $k$  and  $\ell$ , we see that  $x_k$  and  $x_\ell$  (as well as their

neighbors  $x_{k+1}$  and  $x_{\ell-1}$  if they are contained in  $[x_{i+2}, x_{j-2}]$ ) belong to exposed arcs of the matching  $W_{i+2, j-2}$ . Thus the matching  $W_{i, j-2}$  has the following structure:



where dashed (resp., dotted) arcs correspond to those exposed arcs of the matching  $W_{i+2, j-2}$  that belong (resp., do not belong) to  $W_{i, j-2}$ .

Since points  $x_{\ell-1}$  and  $x_{k+1}$  belong to exposed arcs in the matching  $W_{i+2, j-2}$ , by Lemma 4 we see that the (possibly empty) parts of this matching that correspond to points  $x_{i+2} < \dots < x_{\ell-1}$  and  $x_{k+1} < \dots < x_{j-2}$  coincide with the (possibly empty) matchings  $W_{i+2, \ell-1}$  and  $W_{k+1, j-2}$ . For the same reason the (possibly empty) part of the matching  $W_{i, j-2}$  supported on  $x_{\ell+1} < \dots < x_{k-1}$  coincides with  $W_{\ell+1, k-1}$ . Therefore

$$(9) \quad W_{i, j-2} = d(x_i, x_k) + d(x_{i+1}, x_{\ell}) + W_{i+2, \ell-1} + W_{\ell+1, k-1} + W_{k+1, j-2}.$$

Taking into account (7), observe that in the case  $k = i + 1$  and  $\ell = i$ , which was left out in (8), this expression still gives the correct formula  $W_{i, j-2} = d(x_i, x_{i+1}) + W_{i+2, j-2}$ .

Now assume that in the matching  $W_{i+1, j}$  the point  $x_j$  is connected to  $x_{\ell'}$  and the point  $x_{j-1}$  to  $x_{k'}$ . A similar argument gives

$$(10) \quad W_{i+2, j} = W_{i+2, \ell'-1} + W_{\ell'+1, k'-1} + W_{k'+1, j-2} + d(x_{\ell'}, x_j) + d(x_{k'}, x_{j-1});$$

in particular, if  $\ell' = j - 1$  and  $k' = j$ , then  $W_{i+2, j} = W_{i+2, j-2} + d(x_{j-1}, x_j)$ .

Suppose that  $x_k < x_{\ell'}$ . Using again Lemma 4 and taking into account that  $x_k, x_{k+1}, x_{\ell'-1}$ , and  $x_{\ell'}$  all belong to exposed arcs in  $W_{i+2, j-2}$ , we can write

$$(11) \quad W_{k+1, j-2} = W_{k+1, \ell'-1} + W_{\ell', j-2}, \quad W_{i+2, \ell'-1} = W_{i+2, k} + W_{k+1, \ell'-1}$$

and

$$(12) \quad W_{i+2, j-2} = W_{i+2, k} + W_{k+1, \ell'-1} + W_{\ell', j-2}.$$

Substituting (11) into (9) and (10) and taking into account (12), we obtain

$$\begin{aligned} W_{i, j-2} + W_{i+2, j} - W_{i+2, j-2} &= d(x_i, x_k) + d(x_{i+1}, x_{\ell}) + W_{i+2, \ell-1} + W_{\ell+1, k-1} \\ &\quad + W_{k+1, \ell'-1} + d(x_{\ell'}, x_j) + W_{\ell'+1, k'-1} + d(x_{k'}, x_{j-1}) + W_{k'+1, j-2}. \end{aligned}$$

The right-hand side of this expression corresponds to a matching that coincides with  $W_{i, j-2}$  on  $[x_i, x_k]$ , with  $W_{i+2, j-2}$  on  $[x_{k+1}, x_{\ell'-1}]$ , and with  $W_{i+1, j}$  on  $[x_{\ell'}, x_j]$ . By Lemma 4 this matching cannot be improved on any of these three segments and is therefore optimal among all matchings in which  $x_i$  and  $x_j$  belong to different exposed arcs.

It follows that under the assumption that  $x_k < x_{\ell'}$  the expression in the right-hand side of (6) gives the minimum weight of all matchings on  $x_i < x_{i+1} < \dots < x_j$ . Moreover, the only possible candidates for the optimal matching are those constructed above: one that corresponds to  $d(x_i, x_j) + W_{i+1, j-1}$  and one given by the right-hand side of the latter formula.

It remains to consider the case  $x_k \geq x_{\ell'}$ . Since  $x_k \neq x_{\ell'}$  for parity reasons, it follows that  $x_k > x_{\ell'}$ ; now a construction similar to the above yields a matching which corresponds to  $W_{i,j-2} + W_{i+2,j} - W_{i+2,j-2}$  and in which the arcs  $(x_i, x_k)$  and  $(x_{\ell'}, x_j)$  are crossed. Uncrossing them leads to a matching with strictly smaller weight, which contains the arc  $(x_i, x_j)$  and therefore cannot be better than  $d(x_i, x_j) + W_{i+1,j-1}$ . This means that (6) holds in this case too with  $W_{i,j} = d(x_i, x_j) + W_{i+1,j-1}$ .  $\square$

## REFERENCES

- [1] A. Aggarwal, A. Bar-Noy, S. Khuller, D. Kravets, and B. Schieber, *Efficient minimum cost matching using quadrangle inequality*, J. Algorithms **19** (July 1995), no. 1, 116–143, DOI doi:10.1006/jagm.1995.1030.
- [2] D. Burago, Y. Burago, and S. Ivanov, *A Course in Metric Geometry*, Graduate Studies in Mathematics, vol. 33, AMS, Providence, Rhode Island, 2001.
- [3] J. Delon, J. Salomon, and A. Sobolevski, *Local matching indicators for concave transport costs*, C. R. Acad. Sci. Paris, Série I Math. **348** (Aug. 2010), no. 15–16, 901–905, DOI doi:10.1016/j.crma.2010.07.010.
- [4] ———, *Local matching indicators for transport problems with concave costs*, 2011, in preparation.
- [5] J. Heinonen, *Lectures on Analysis in Metric Spaces*, Universitext, Springer-Verlag, 2001.
- [6] R. M. Karp and S. Y. R. Li, *Two special cases of the assignment problem*, Discrete Mathematics **13** (1975), 129–142.
- [7] R. J. McCann, *Exact solutions to the transportation problem on the line*, Proc. R. Soc. A: Math., Phys. and Eng. Sci. **455** (1999), 1341–1380.
- [8] M. Werman, S. Peleg, R. Meller, and T. Y. Kong, *Bipartite graph matching for points on a line or a circle*, J. Algorithms **7** (1986), no. 2, 277–284.

LTCI CNRS, TELECOM PARISTECH, 46 RUE BARRAULT, 75634 PARIS, FRANCE

CEREMADE, UNIVERSITÉ PARIS-DAUPHINE, PLACE DU MARECHAL LATTRE DE TASSIGNY, 75775 PARIS, FRANCE

INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS (KHARKEVICH INSTITUTE), 19 B. KARETNY PER., MOSCOW 127994, RUSSIA

LABORATOIRE J.-V. PONCELET (UMI 2615 CNRS), 11 B. VLASSIEVSKI PER., MOSCOW 119002, RUSSIA

*E-mail address:* sobolevski@iitp.ru