Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

Optimal Metric Distortion for Matching on the Line

Aris Filos-Ratsikas[1], **Vasilis Gkatzelis**[2], **Mohamad Latifian**[1], **Emma Rewinski**[2] and **Alexandros A. Voudouris**[3]

[1]University of Edinburgh, UK
[2]Drexel University, USA
[3]University of Essex, UK

Abstract

We study the distortion of one-sided and two-sided matching problems on the line. In the one-sided case, n agents need to be matched to n items, and each agent's cost in a matching is their distance from the item they were matched to. We propose an algorithm that is provided only with ordinal information regarding the agents' preferences (each agent's ranking of the items from most- to least-preferred) and returns a matching aiming to minimize the social cost with respect to the agents' true (cardinal) costs. We prove that our algorithm simultaneously achieves the best-possible approximation of 3 (known as distortion) with respect to a variety of social cost measures which include the utilitarian and egalitarian social cost. In the two-sided case, where the agents need be matched to n other agents and both sides report their ordinal preferences over each other, we show that it is always possible to compute an optimal matching. In fact, we show that this optimal matching can be achieved using even less information, and we provide bounds regarding the sufficient number of queries.

1 Introduction

Matching problems are prevalent in everyday life and have played a central role in (computational) social choice for multiple decades, starting from the pioneering works of Gale and Shapley [1962], Shapley and Scarf [1974], and Hylland and Zeckhauser [1979]. These works consider settings in which a set of n *agents* need to be matched to a set of n items (one-sided matching), or to another set of n agents (two-sided matching), in some fair and efficient manner with respect to the agents' preferences. Some important applications of these problems include college admissions [Gale and Shapley, 1962], national residency matching [NRMP, ; Roth, 1984; Roth and Peranson, 1999]), school choice [Abdulkadiroğlu *et al.*, 2005b; Abdulkadiroğlu *et al.*, 2005a], and organ exchange [UNOS, ; Roth *et al.*, 2004].

To accurately capture the, often complicated, preferences of an agent over a discrete set of outcomes, the common approach is to use a Von Neumann–Morgenstern utility function which assigns a numerical value to each outcome, indicating how strongly it is liked or disliked [Von Neumann and Morgenstern, 1944]. Although the significance of these *cardinal* preferences of the agents is recognized in the classic literature on matching (e.g., [Hylland and Zeckhauser, 1979; Zhou, 1990; Bogomolnaia and Moulin, 2001]), most matching algorithms that have been proposed are *ordinal*: They ask each agent to report only their ranking over the outcomes from most- to least-preferred. This is motivated by the fact that it can be cognitively or computationally prohibitive for the agents to come up with these values, whereas simply ranking outcomes is an easier, often routine, task. However, given access only to this limited (ordinal) information regarding the agents' preferences, to what extent can one approximate social objectives which are functions of the cardinal values?

This is precisely the question studied in the vibrant literature of *distortion in social choice*, whose goal is to design ordinal algorithms with low *distortion*, i.e., strong worst case approximation guarantees with respect to cardinal social cost objectives (see [Anshelevich *et al.*, 2021] for a recent survey). One of the most well-studied settings in this literature since its inception [Procaccia and Rosenschein, 2006] is that of *metric distortion*, where the agents' cardinal preferences correspond to distances in some metric space [Merrill III *et al.*, 1999; Enelow and Hinich, 1984]. The metric distortion literature has been very successful in a plethora of different settings, leading to (near-)optimal distortion bounds. Examples include single-winner voting [Anshelevich *et al.*, 2018; Gkatzelis *et al.*, 2020; Kizilkaya and Kempe, 2022], multiwinner voting [Caragiannis *et al.*, 2022], and probabilistic social choice [Feldman *et al.*, 2016; Anshelevich and Postl, 2017; Charikar and Ramakrishnan, 2022; Charikar *et al.*, 2024]. However, a notable exception that remains poorly-understood, despite its fundamental nature and wealth of applications, is the metric distortion of matching problems.

In the *metric matching* setting, agents and items are embedded in a metric space, and an agent's cost in a matching is the distance between them and their match. The distortion of the metric (one-sided) matching problem was first studied by Caragiannis *et al.* [2024][1] who focused on the utilitarian social cost (the total cost) and, among other results, observed that the distortion of the well-known *serial dictatorship* algo-

[1]The conference version of this paper was published in 2016.

rithm is 2^{n-1}. For its randomized counterpart, the *random serial dictatorship*, they showed that its distortion is between $n^{0.29}$ and n. Their work also implies a lower bound of 3 on the distortion of any (possibly randomized) ordinal algorithm with respect to the utilitarian social cost. More recently, Anari *et al.* [2023] achieved improvements on both fronts: They designed a deterministic algorithm with distortion $O(n^2)$ and complemented it with a lower bound of $\Omega(\log n)$ on the distortion of any ordinal algorithm, even randomized ones. Despite these improvements, there is still a very large gap between the lower and the upper bound, and closing this gap is one of the main open problems in this literature.

In this paper, we resolve the metric distortion of the matching problem for the interesting case of the line metric, which has received a significant amount of attention, both because it models natural scenarios of interest, and because it gives rise to worst-case instances for many metric distortion settings.

In terms of motivating applications for the line metric, widely known as 1-*Euclidean*, one of the natural ones is the fact that it provides an abstraction of political or ideological beliefs along different axes, encompassed into a spectrum, e.g., between "left" and "right". This idea is inherent in the pioneering works of [Hotelling, 1929] and [Downs, 1957]. The line metric can also be used to capture preferences over policy issues, such as the extent to which a government should be involved in the economy, ranging from full involvement to zero intervention [Stokes, 1963]; see [Filos-Ratsikas *et al.*, 2024] for a more elaborate discussion.

It turns out in most of the fundamental metric distortion settings mentioned above (namely, single-winner and multi-winner voting), the best possible distortion on the line metric is either the same as or within a small constant from the best possible distortion on any metric. In the metric matching problem, however, this is not the case; the lower bound of $\Omega(\log n)$ shown by Anari *et al.* [2023] requires a highly involved tree metric. In particular, this lower bound does not rule out the existence of an algorithm with *constant* distortion on the line. Is a such a constant achievable?

Our main result provides a positive answer to this question by providing an ordinal algorithm that achieves a distortion of 3 with respect to the well-known *k-centrum cost* (for different values of $k \in \{1, \dots, n\}$), which is equal to the sum of the k largest agent costs [Tamir, 2001]; see also [Han *et al.*, 2023]. This objective naturally interpolates between the maximum (egalitarian) cost when $k = 1$ and the utilitarian social cost when $k = n$. In fact, we show that this distortion is the best possible for each of these cost functions and even for randomized algorithms, which fully resolves this question.

Informal Theorem 1. *There is a deterministic ordinal algorithm for the metric one-sided matching problem on the line that simultaneously achieves a distortion of 3 with respect to the k-centrum cost for all $k \in \{1, \dots, n\}$. This distortion is the best possible for any ordinal algorithm, including randomized ones.*

This result provides a clear, asymptotic separation between the distortion of one-sided matching on the line metric and its distortion in more general metric spaces.

Building on our main result above, we then consider the distortion of the metric two-sided matching problem. Despite the ubiquitous nature of this problem, it is perhaps surprising that, to the best of our knowledge, its metric distortion has not been considered prior to our work. For this setting, we prove a seemingly surprising result, namely that, using only ordinal information, we can construct a matching that *exactly* minimizes the k-centrum cost.

Informal Theorem 2. *There is an* optimal *(i.e., with distortion 1) deterministic ordinal algorithm for the metric two-sided matching problem on the line.*

From a technical standpoint, what enables us to improve from a distortion of 3 in the one-sided matching case to achieving optimality in the two-sided matching setting is the additional information that we have from the preferences of the agents on the other side. We explore this further and show that an optimal matching can be computed using even less information about the ordinal preferences of the agents, which is acquired by appropriate *queries*.

1.1 Further Related Work

Prior to the work of Anari *et al.* [2023], Anshelevich and Zhu [2021] considered the metric one-sided matching problem as part of a class of more general facility assignment problems, and showed a bound of 3 on the distortion of ordinal algorithms for the social cost and the maximum cost objectives, but under the restriction that the item locations in the metric space are known. Other works [Anshelevich and Zhu, 2019; Anshelevich and Sekar, 2016b; Anshelevich and Sekar, 2016a] considered related matching settings where the goal is to maximize the *metric utilities* of the agents rather than minimizing their costs.

In the non-metric setting, where the agents are assumed to have normalized utilities over the items, the best possible distortion for one-sided matching has been identified for both deterministic [Amanatidis *et al.*, 2022] and randomized algorithms [Filos-Ratsikas *et al.*, 2014]. The distortion of one-sided matching (and generalizations of it) has also been studied when more information can be elicited via queries [Amanatidis *et al.*, 2022; Amanatidis *et al.*, 2024; Ma *et al.*, 2021; Latifian and Voudouris, 2024; Ebadian and Shah, 2025]. The query models in those works are inherently different from the one we consider here, as they elicit *cardinal* information, on top of the ordinal preferences which are considered to be known. In contrast, our queries in the two-sided model elicit only *ordinal* information about the agents' preferences on each side.

2 Preliminaries

For any positive integer ℓ, let $[\ell] = \{1, \dots, \ell\}$. An instance of our problem consists of a set of n *agents* $A = \{a_1, \dots, a_n\}$ and a set of n *items* $G = \{g_1, \dots, g_n\}$. We assume that agents and items are located on distinct points on the same line metric[2], and let $d(a, g)$ denote the *distance* between any

[2]The assumption that agent and item locations are all distinct is without loss of generality: it suffices to assume that agent's ordinal

agent a and item g. Let $\succ_a$ be the *ordinal preference* of agent a over the set of items G, such that $g \succ_a g'$ implies $d(a,g) \leq d(a,g')$. Let $\succ := (\succ_a)_{a \in A}$ be the *ordinal profile* consisting of the ordinal preferences of all agents. Note that many different distance metrics may induce the same ordinal preferences for the agents. We write $d \rhd \succ$ to denote the event that the metric d is consistent to the ordinal profile $\succ$. We also denote by $\text{top}(a)$ the *favorite item* of agent a, that is, $\text{top}(a) \succ_a g$ for every $g \in G \setminus \{\text{top}(a)\}$. Let $\text{plu}(g)$ be the plurality score of an item $g \in G$, which is equal to the number of agents whose favorite item is g, i.e., $\text{plu}(g) = |\{a \in A : \text{top}(a) = g\}|$.

Our goal is to choose a matching M between the agents and the items. Given such a matching, we denote by $M(a)$ the item matched to agent a, and by $M(g)$ the agent matched to item g. We evaluate the quality of a matching using a variety of social cost measures, captured by the k-centrum cost. Given some $k \in [n]$, the *k-centrum cost* of a matching M is the sum of the k largest distances between the agents and their matched items, among all agents:

$$\text{SC}_k(M|d) = \sum_{q=1}^{k} \max_{a \in A}{}^{q} d(a, M(a)),$$

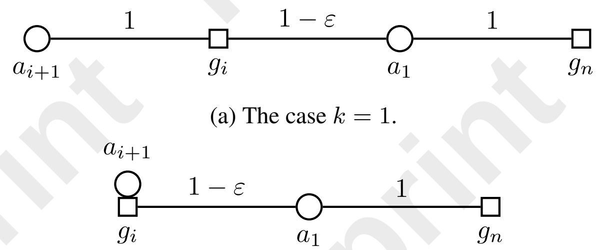
where $\max^q$ returns the q-th largest value. Note that this captures the two most well-studied social cost functions, the egalitarian and the utilitarian social cost, as special cases. Specifically, $\text{SC}_n(M|d) = \sum_a d(a, M(a))$ corresponds to the utilitarian social cost, which evaluates on the total cost over all agents, while $\text{SC}_1(M|d) = \max_a d(a, M(a))$ corresponds to the egalitarian social cost, which focuses on the agent that suffers the largest cost. In general, smaller values of k put more emphasis on fairness by focusing on the least happy agents. When the metric d is clear from context, we will simplify our notation and write $\text{SC}_k(M)$ instead of $\text{SC}_k(M|d)$.

A matching *algorithm* ALG takes as input an ordinal profile $\succ$ and computes a one-to-one matching $\text{ALG}(\succ)$ of the items to the agents. We want to design matching algorithms with as small distortion as possible. The *distortion* of a matching algorithm ALG with respect to the k-centrum cost SC_k for some $k \in [n]$ is the worst-case ratio over all possible ordinal profiles and consistent metrics between the objective value of the matching returned by the algorithm and the minimum possible objective value over all possible matchings:

$$\sup_{\succ} \sup_{d: d \rhd \succ} \frac{\text{SC}_k(\text{ALG}(\succ)|d)}{\min_M \text{SC}_k(M|d)}.$$

M^*: A well-structured optimal matching

For each instance there may exist multiple matchings that minimize the k-centrum cost. Among these optimal matchings, we are particularly interested in one that greedily matches agents to items according to their true ordering on the line; the optimality of this matching is established in the following theorem. Throughout the paper, we denote this specific matching by M^* and refer to it simply as the optimal matching. Also, for each agent a we refer to $M^*(a)$ as a's optimal item. Due to space constraints, the proof of the following theorem (as well as of others in the rest of the paper) are omitted.



(a) The case $k = 1$.

(b) The case $k \geq 2$.

Figure 1: The metrics used in the proof of Theorem 3.1 to give a lower bound of 3 on the distortion of (randomized) algorithms for (a) $k = 1$ and (b) $k \geq 2$. Circles correspond to agents, rectangles correspond to items, and $i \in [n-1]$. The weight of each edge is the distance between its endpoints.

Theorem 2.1. *Given the true ordering of all agents and items on the line, greedily matching the leftmost agent to the leftmost item leads to an optimal matching M^* with respect to SC_k for all $k \in [n]$.*

A structural property

Given an ordinal profile $\succ$, we can observe some structure regarding the locations of the agents and the items on the line.

Lemma 2.2. *If item g_x is to the left of item g_y, then all agents whose favorite item is g_x are to the left of all agents whose favorite item is g_y.*

3 One-Sided Case: Optimal Distortion of 3

In this section we first prove that no algorithm can achieve a distortion better than 3 with respect to the k-centrum cost for any $k \in [n]$, even if it is randomized. Then, we present an optimal (deterministic) algorithm that simultaneously guarantees a distortion upper bound of 3 with respect to all k-centrum costs.

3.1 Distortion Lower Bound for any Algorithm

Here we prove a lower bound on the distortion of matching algorithms, even randomized ones.

Theorem 3.1. *For any $k \in [n]$, no deterministic algorithm can achieve a distortion better than 3 with respect to the k-centrum cost SC_k, and no randomized algorithm can achieve a distortion better than $3 - 2/n$.*

Proof. Consider an instance with n agents $\{a_1, \dots, a_n\}$ and n items $\{g_1, \dots, g_n\}$. The ordinal profile is such that all agents have the same ranking over the items: $g_1 \succ \dots \succ g_n$. For any $i \in [n]$, let p_i be the probability that agent a_i is matched to item g_n according to an arbitrary randomized matching algorithm. Without loss of generality, suppose that $p_1 \leq 1/n$; note that if the algorithm is deterministic, then $p_1 = 0$.

preferences consistently tie-break over items with the same location.

Consider the metric spaces that are illustrated in Figure 1a for $k = 1$ and in Figure 1b for $k \geq 2$, where $\varepsilon > 0$ is an infinitesimal. For any $k \in [n]$, the optimal matching consists of the pairs (a_1, g_n) and (a_{i+1}, g_i) for $i \in [n-1]$, leading to a k-centrum cost of $d(a_1, g_n) = 1$. However, since the algorithm matches a_1 to g_n with probability p_1, it matches agents in $\{a_1, \dots, a_n\}$ to g_n with the remaining probability $1 - p_1$. Whenever an agent in $\{a_1, \dots, a_n\}$ is matched to g_n, the k-centrum cost of this matching is $3 - \varepsilon$, thus leading to an expected k-centrum cost of

$$p_1 \cdot 1 + (1 - p_1) \cdot (3 - \varepsilon) = 3 - \varepsilon - (2 - \varepsilon) p_1.$$

Consequently, the distortion is at least $3 - \varepsilon$ for deterministic algorithms (since $p_1 = 0$) and $3 - 2/n - \varepsilon$ for randomized algorithms, for any $\varepsilon > 0$. □

3.2 Distortion-optimal Matching Algorithm

We now provide a deterministic algorithm that simultaneously achieves the optimal distortion of 3 with respect to SC_k for any $k \in [n]$. Our algorithm consists of the following three key steps: The first one is to appropriately partition the items into two sets G_{in} and G_{out} using the ordinal preferences of the agents. The second step determines the true ordering (up to reversal) of the items in G_{in}, and it also uses this ordering to define an ordering over the set of agents. The third step matches the items in G_{in} to agents using the aforementioned orderings, and then arbitrarily matches the remaining items to the remaining agents. The rest of this section provides more details regarding the implementation of these steps. A description of the algorithm using pseudocode is given as Algorithm 1.

Step 1: Partitioning the items into G_{in} and G_{out}
Let $G_+ := \{g \in G : \text{plu}(g) \geq 1\}$ be the subset of items in G that have a positive plurality score. By considering each agent's ranking of the items in G_+, at most two items in G_+ can be ranked last by the agents (these would be the two extreme items in G_+, i.e., the leftmost and the rightmost one). We use g_ℓ and g_r to denote these items, and without loss of generality[3], we assume that the former is the leftmost item in G_+ and the latter is the rightmost one. If every agent ranks the same item of G_+ last, that means $|G_+| = 1$ and $g_\ell = g_r$.

Since $g_\ell, g_r \in G_+$, both of these items have at least one agent who ranks them at the top. Given two agents a_i and a_j such that $\text{top}(a_i) = g_\ell$ and $\text{top}(a_j) = g_r$, we define the following subset of items:

$$G_{\text{in}}(a_i, a_j) := \{g \in G : g \succ_{a_i} g_r \text{ or } g \succ_{a_j} g_\ell\}.$$

We define $(a_\ell, a_r) \in \arg\max_{a_i, a_j} |G_{\text{in}}(a_i, a_j)|$ to be two agents that maximize the size of $G_{\text{in}}(a_i, a_j)$, we henceforth partition the set of items into the following two sets:

$$G_{\text{in}} := G_{\text{in}}(a_\ell, a_r) \quad \text{and} \quad G_{\text{out}} := G \setminus G_{\text{in}}. \tag{1}$$

In words, G_{in} consists of the items that either agent a_ℓ prefers over g_r or agent a_r prefers over g_ℓ, and G_{out} consists of all the remaining items. We remark that defining $(a_\ell, a_r) \in \arg\max_{a_i, a_j} |G_{\text{in}}(a_i, a_j)|$ is vital to achieving a distortion of 3. If a_ℓ and a_r were to be chosen arbitrarily among the agents whose favorite items are g_ℓ and g_r, a distortion of 3 would not be achieved with respect to SC_k for any $k \in [n]$.

[3]Note that both the behavior of our algorithm and the analysis is invariant to the reversal of agent and item locations; so, we only fix g_ℓ to be the leftmost item to make the exposition more intuitive.

Algorithm 1: ORDERMATCH

Input: Sets A and G and ordinal profile $\succ$
Output: A matching M between A and G
Identify the extreme items g_ℓ and g_r using $\succ$.
Partition G into G_{in} and G_{out} according to (1).
Compute ordering π_g over G_{in} and π_a over A.
Match the ith item in π_g to the ith agent in π_a.
Arbitrarily match items in G_{out} to unmatched agents.

Note that if $g_\ell = g_r$, i.e., there is only one item in G_+, then $G_{\text{in}} = \varnothing$ since no other item is preferred over that item.

Lemma 3.2. *If $G_{in} \neq \varnothing$, then $G_+ \subseteq G_{in}$.*

Step 2(a): Ordering π_g of the items in G_{in}
Although we cannot fully uncover the true ordering of all items in G given the ordinal preferences of the agents, we can extract the true ordering π_g of the items in $G_{\text{in}} \subseteq G$ using the algorithm of [Elkind and Faliszewski, 2014] which do so for a superset of G_{in}.

Lemma 3.3. *The true ordering π_g of the items in G_{in} can be computed.*

Step 2(b): Ordering π_a of the agents in A
Given the ordering π_g of the items in G_{in}, we also define an ordering of all the agents in A based on what their top item is. Note that every agent's top item is in G_+ which, by Lemma 3.2 is a subset of G_{in}, so by Lemma 2.2, π_g implies a partial order over A. First, we let $\tilde{\pi}_a$ be the partial ordering of the agents in A such that $\tilde{\pi}_a(a_i) = \tilde{\pi}_a(a_j)$ if $\text{top}(a_i) = \text{top}(a_j)$ and $\tilde{\pi}_a(a_i) > \tilde{\pi}_a(a_j)$ if $\pi_g(\text{top}(a_i)) > \pi_g(\text{top}(a_j))$. Then, we define a total order π_a by breaking the ties of the partial order in some arbitrary way.

Step 3: Matching agents to items using π_a and π_g
Our algorithm then determines how to match the items to agents starting with the items in G_{in}. It considers them based on the ordering π_g and assigns them to the agents using the ordering of π_a. Once all the items in G_{in} have been assigned, our algorithm arbitrarily matches the items in G_{out} to the agents that remain unmatched.

3.3 Analysis of ORDERMATCH

We now prove the main result of this section.

Theorem 3.4. ORDERMATCH *achieves a distortion of* 3 *with respect to the k-centrum cost SC_k for any $k \in [n]$.*

While our algorithm only has access to the ordinal preferences of the agents through $\succ$, our analysis in this section uses the exact locations of the agents and items along with the optimal matching M^* (e.g., to further partition the agents using this information, as well as to define a graph induced by each matching). For example, let $A_{\text{in}} = \{a \in A : M^*(a) \in$

$G_{\text{in}}\}$ be the set of agents whose optimal items (according to the optimal matching M^*) are in G_{in} and $A_{\text{out}} = A \setminus A_{\text{in}}$ be the set of all the remaining agents, i.e., the agents whose optimal items are in G_{out}.

Permutation graph P_M induced by matching M

Given an instance with optimal matching M^*, we define the *permutation-graph* of any matching M as $P_M = (A, E)$, where each vertex corresponds to an agent in A, and each directed edge (a_i, a_j) exists in E if and only if $M(a_i) = M^*(a_j)$ (i.e., there exists an edge from agent a_i to the agent a_j if and only if M assigns a_j's optimal item to a_i; note that it may be the case that $a_i = a_j$). In other words, this graph captures the permutation of items from M^* to M. Note that for any given instance (i.e., for a fixed M^*) every permutation graph P_M corresponds to a unique matching M. Also, note that every vertex of a permutation-graph has in-degree and out-degree equal to 1 and P_M is a collection of disjoint cycles.

To categorize some of the edges of this graph, we further refine the set of agents A_{in} as follows: Let $A_x \subseteq A_{\text{in}}$ be the subset of agents in A_{in} whose favorite item is the xth-leftmost item in G_{in}. Due to Lemma 2.2, every agent in A_x is to the left of every agent in A_y for any $x < y$. We refer to an edge $(a_i, a_j) \in E$ as *forward* if $a_i \in A_x$ and $a_j \in A_y$ for some $x < y$; *backward* if $a_i \in A_x$ and $a_j \in A_y$ for some $x > y$; *internal* if $a_i \in A_x$ and $a_j \in A_x$; *inward* if $a_i \in A_{\text{out}}$ and $a_j \in A_{\text{in}}$. See the example in Figure 2.

For any fixed instance, each graph P_M corresponds to a distinct matching M, so we can also express the k-centrum cost of a graph P_M and its corresponding matching M as:

$$\text{SC}_k(P_M) = \text{SC}_k(M) = \sum_{q=1}^{k} \max_{(a_i,a_j)\in E}{}^{q}\, d(a_i, M^*(a_j))$$

We will now prove that the permutation graph P_M of our algorithm has no backward edges and can be reduced to a permutation graph $P_{M'}$ with no forward or backward edges with a weakly larger k-centrum cost for all $k \in [n]$. We will then prove the desired distortion of 3 for $P_{M'}$.

Lemma 3.5. *If M is the matching computed by* ORDERMATCH*, then the induced graph P_M does not include backward edges.*

Proof. If G_{in} is empty (i.e., $g_\ell = g_r$), then P_M cannot include any backward edges. If G_{in} is non-empty, then Lemma 3.3 shows that π_g correctly orders the items in G_{in} from left to right. In the ordering π_a, agents are ordered earlier than other agents if their favorite item is to the left of the favorite item of those other agents according to π_g. For any $x < y$, Lemma 2.2 shows that all agents in A_x are to the left of all agents in A_y. Thus, the optimal item of every agent in A_x is to the left of the optimal item of every agent in A_y due to Theorem 2.1. Consequently, since all agents in A_x are matched to items before the agents in A_y, no agent in A_y can be matched to the optimal item of an agent in A_x, and thus there are no backward edges in P_M. ☐

Algorithm 2: REMOVEFORWARDEDGES

Input: Graph $P_M = (A, E)$ with no backward edges.
Output: Graph $P_{M'}$.
while there exists a forward edge in E **do**
  Let $(a_3, a_4) \in E$ be a forward edge such that top(a_3) is weakly to the right of top(a_i) for any other forward edge $(a_i, a_j) \in E$. Find a forward or inward edge $(a_1, a_2) \in E$ such that top$(a_2) =$ top(a_3) .
  $E \leftarrow E \setminus \{(a_1, a_2), (a_3, a_4)\}$.
  $E \leftarrow E \cup \{(a_1, a_4), (a_3, a_2)\}$.
return $P_{M'} \leftarrow (A, E)$.

Reduction to graphs without forward edges

If P_M is the permutation-graph induced by the matching M computed by ORDERMATCH, Lemma 3.5 shows that this graph will not contain any backward edges. We now prove that, even though this graph can, in general, contain forward edges, we can without loss of generality assume that it does not. Specifically, we provide a reduction (REMOVEFORWARDEDGES) that takes as input any permutation-graph P_M without backward edges and transforms it into a graph $P_{M'}$ that contains neither forward nor backward edges, while ensuring that $\text{SC}_k(P_M) \le \text{SC}_k(P_{M'})$ for any $k \in [n]$. This reduction iteratively removes inward or forward edges from the graph and replaces them with other inward or internal edges, until no forward edges remain.

Theorem 3.6. *Let M be the matching computed by* ORDERMATCH *when given an ordinal profile $\succ$. If its permutation-graph P_M contains a forward edge, then there exists another matching M' whose permutation-graph $P_{M'}$ does not contain any forward or backward edges, and has weakly larger k-centrum cost, i.e., $SC_k(M') \ge SC_k(M)$ for any $k \in [n]$.*

Bounding the cost of all remaining edges

Having addressed forward and backward edges, we now provide an upper bound for the cost of all remaining edges.

Lemma 3.7. *Consider any permutation-graph $P_M = (A, E)$ without forward or backward edges. For any $(a_i, a_j) \in E$,*

$$d(a_i, M^*(a_j)) \le d(a_i, M^*(a_i)) + 2 \cdot d(a_j, M^*(a_j)).$$

Proof Sketch. Since (a_i, a_j) is neither a forward nor a backward edge it could be an internal or inward edge, or it could also be the case that $a_i \in A_{\text{in}}$ and $a_j \in A_{\text{out}}$, or both $a_i, a_j \in A_{\text{out}}$. In the appendix we prove the claimed inequality for each case separately. Here, we only prove the case of internal edges.

If (a_i, a_j) is an internal edge, both agents share the same favorite item, say g^*. Using the triangle inequality and the fact that $d(a, g^*) \le d(a, g)$ for any $a \in \{a_i, a_j\}$ and $g \in G$, we have:

$$\begin{aligned} d(a_i, M^*(a_j)) &\le d(a_i, g^*) + d(a_j, g^*) + d(a_j, M^*(a_j)) \\ &\le d(a_i, M^*(a_i)) + 2 \cdot d(a_j, M^*(a_j)). \end{aligned}$$ ☐

We can now prove the main result of this section.

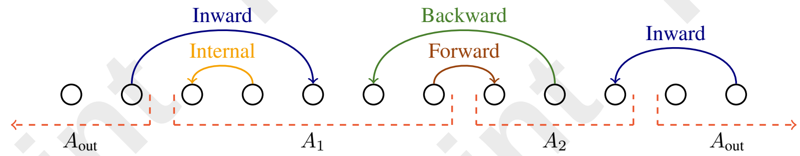


Figure 2: A snapshot of a graph P_M induced by some matching M, with its nodes appearing in the order of the corresponding agents' true locations on the line. The dashed lines below the nodes exhibit how these agents are partitioned into A_1, A_2, and A_{out}. A subset of the graph's edges appear above the nodes, labeled by their type (forward, backward, internal, and inward).

Proof of Theorem 3.4. By Lemma 3.5, the permutation-graph of the matching returned by ORDERMATCH has no backward edges, and by Lemma 3.6, the k-centrum cost of this matching can be upper-bounded by the k-centrum cost of another matching whose permutation-graph has no forward or backward edges. We now show that for any matching M whose permutation-graph $P_M = (A, E)$ contains no forward or backward edges, we have $\text{SC}_k(M) = \text{SC}_k(P_M) \leq 3 \cdot \text{SC}_k(M^*)$ for all $k \in [n]$. Specifically, if we let $E_k \subseteq E$ be the k edges (a_i, a_j) of E with the largest $d(a_i, M^*(a_j))$ cost, then

$$\begin{aligned}\text{SC}_k(M) &= \sum_{(a_i,a_j)\in E_k} d(a_i, M^*(a_j)) \\ &\leq \sum_{(a_i,a_j)\in E_k} \Big(d(a_i, M^*(a_i)) + 2 \cdot d(a_j, M^*(a_j)) \Big) \\ &= \sum_{a_i:(a_i,a_j)\in E_k} d(a_i, M^*(a_i)) + \sum_{a_j:(a_i,a_j)\in E_k} 2 \cdot d(a_j, M^*(a_j)) \\ &\leq 3 \cdot \sum_{q=1}^{k} \max_{a\in A}{}^q d(a, M^*(a)) = 3 \cdot \text{SC}_k(M^*),\end{aligned}$$

where the first inequality uses Lemma 3.7 for every edge in E_k and the subsequent equality breaks down the contribution of each edge $(a_i, a_j) \in E_k$ to the contribution of a_i and the contribution of a_j. The final inequality uses the fact that: *(i)* all nodes have in- and out-degree 1 in any permutation-graph, so each agent a participates in at most two edges in E_k, (a, a_j) and (a_i, a), and *(ii)* the total cost of the k agents $a_i : (a_i, a_j) \in E_k$ (and that of the k agents $a_j : (a_i, a_j) \in E_k$) is at most the total cost of the k maximum-cost agents. □

4 Two-Sided Case: Reaching the Optimal Matching

In this section we consider a two-sided matching, in which instead of n agents that have preferences of n items, we have a set of n *takers* denoted by A and a set of n *givers* denoted by B. We have access to the ranking of each taker over the givers, and of each giver over the takers. We will refer to each of these sets as a side of the matching problem, and will use the term *agent* to refer to a taker or a giver.

We will show that using the ordinal rankings of the takers and the givers, we can compute an optimal matching. In particular, to do so, it suffices to show that we can recover the true ordering of the takers and the givers on the line. Then, we can greedily compute an optimal matching using Theorem 2.1 by considering one side as the items.

Theorem 4.1. *For the two-sided matching problem on a line metric, we can compute an optimal matching with respect to the k-centrum cost SC_k for any $k \in [n]$.*

Proof. We first show that we can recover the true relative ordering of the takers as well as the true relative ordering of the givers. We do so for the takers; the argument for the givers is symmetric. We consider two cases: Either all takers agree on their least-preferred giver, or there are two givers that appear at the bottom of the rankings of the takers.

In the first case, let b be the giver that appears at the bottom of every taker's ranking. We claim that all takers are on the same side of b. If not, then there is a taker i that is to the left of b and a taker j that is to the right of b. This implies that i weakly prefers b to top(j), which contradicts the fact that b is the least-preferred giver of all takers. Now, since all takers are on the same side of b, $\succ_b$ is the true ranking of the takers.

In the second case, let b_r and b_ℓ be the two extreme givers, and for the ease of exposition assume that b_r is to the right of b_ℓ. First, we claim that there is no giver to the left of b_ℓ, and similarly to the right of b_r. For the sake of contradiction, assume that there is a giver b to the left of b_ℓ, and let i be a taker with $b_r \succ_i b_\ell$. This means that this taker is to the right of b_ℓ, and hence $b_r \succ_i b_\ell \succ_i b$, which contradicts the fact that only b_r and b_ℓ appear at the bottom of the takers' rankings.

Let A^{b_ℓ} be the set of takers that have b_ℓ as their top choice. Since b_ℓ is the leftmost giver, all the takers to the left of it are in A^{b_ℓ}. Hence, all the takers in $A \setminus A^{b_\ell}$ are to the right of b_ℓ and to the right of all the takers in A^{b_ℓ}. Consequently, the ordering $\succ_{b_\ell}$ over $A \setminus A^{b_\ell}$ is the true ordering of these takers. Similarly, $\succ_{b_r}$ recovers the true ordering of A^{b_ℓ}. By putting these two together, we have the true ordering of the whole set of takers.

Now that we have the true relative ordering of the takers and the givers, we can greedily match them to get the optimal matching M^*. The only issue is that these orderings are correct up to reversal and can be oriented in opposite directions. In other words, one ordering can be from left to right and the other from right to left. So, we have to make them consistent. We consider two cases depending on if at least one of the two sides have two different agents at the bottom of their lists (case 1), or both sides have the same (case 2).

Case 1. Assume that the takers have two different givers as their least preferred. Let a be an extreme taker (given the ordering of the takers, which we now know), and assume that a is the rightmost one. Then, a's least preferred giver must be the leftmost one. We use this information to give the same direction to the relative orderings of the takers and the givers.

Case 2. Here, each side has the same least-preferred agent in the other side. Let a be an extreme taker (according to the ordering of the takers), and assume that a is to the rightmost one. Then, a is to the same side of all the givers, and thus top(a) has to be one of the extreme takers. Hence, we order the givers so that top(a) is the rightmost one.

Now that the orderings are in the same direction we can greedily match them from left to right, thus computing the optimal matching M^*. □

4.1 Query Model

We now consider a setting where we are given some prior information about the preferences of the agents, and we can obtain further information by making specific types of *queries*. In particular, we consider two types of prior information:

- *One-sided setting*: We are given the complete ordinal preferences of the takers in A over the givers in B.
- *Zero-knowledge setting*: No initial information about the preferences of the agents is given.

Similarly, we consider two types of queries that we can make to gain more information:

- *Full-preference queries*: We ask an agent to reveal its ordinal preference over all the agents on the other side.
- *Rank queries*: We ask an agent to reveal its t-th preferred agent on the other side.

For each combination of prior information and type of queries as defined above, we are interested in pinpointing the number of queries that are sufficient to compute an optimal matching. Given that an optimal matching can be computed given for the two-sided matching problem as we showed above, we are essentially aiming to find out how many queries we need to make in each case to reveal the true ordering of the agents of each side on the line.

In the one-sided setting, n full-preference queries are clearly sufficient to compute the optimal matching as we would then have the ordinal preference of all agents, leading to the two-sided matching problem. We further show that this amount of such queries is necessary.

Theorem 4.2. *In the one-sided setting, at least $n-1$ full-preference queries are required to compute an optimal matching.*

Due to Theorem 4.2, since $n-1$ full-preference queries are required even when the whole ordering of one side is given, by applying the argument to each side independently, we obtain a lower bound of $2n-2$ on the number of required full-preference queries for the zero-knowledge setting, which is nearly tight as $2n$ full-preference queries reveal the orderings of both sides of agents.

Corollary 4.3. *In the zero-knowledge setting, we need at least $2n-2$ full-preference queries to find the optimal matching.*

Since we can simulate a full-preference query using at most $n-1$ rank queries, we can clearly compute an optimal matching by using at most $n(n-1)$ rank queries in the one-sided setting. In the following theorem, we show that we can do this using only $3n-4$ rank queries.

Theorem 4.4. *In the one-sided setting, we can compute an optimal matching by making $3n-4$ rank queries.*

For the zero-knowledge setting, we show that only a constant factor more rank queries are required to compute the optimal matching.

Theorem 4.5. *In the zero-knowledge setting, we can find the optimal matching with $5n-4$ rank queries.*

5 Conclusion and Future Directions

We showed a tight bound of 3 on the distortion of ordinal algorithms for the metric one-sided matching problem on the line with respect to the k-centrum cost, the sum of the k-largest costs among all n agents, for any $k \in [n]$. We further showed that, for the two-sided matching problem, the ordinal preferences of the agents are sufficient to reveal their true ordering on the line which can then be used to compute an optimal matching; in fact, even less information is sufficient to achieve optimality.

Although we have resolved the distortion of matching on the line, there are several natural directions for future research. For example, it would be interesting to investigate the one-sided matching problem on the line using the learning-augmented framework [Lykouris and Vassilvitskii, 2021]. According to this framework, which was recently also used for the metric distortion of single-winner voting [Berger *et al.*, 2024], the designer is provided with some (potentially inaccurate) *prediction* regarding the agents' cardinal preferences. The goal is to achieve an improved distortion whenever the prediction is accurate (known as consistency) while maintaining a good distortion even if the prediction is arbitrarily inaccurate (known as robustness). In our setting, the main question would be whether a consistency better than 3 can be combined with a robustness of 3.

For more general metrics, the distortion of the one-sided matching problem still remains wide open, and bridging the gap between the known lower bound of $\Omega(\log n)$ and the upper bound of $O(n^2)$ for deterministic algorithms or the upper bound of $O(n)$ for randomized algorithms is quite possibly one of the most challenging open questions in the distortion literature at the moment. To obtain positive results, as we did in this paper for the line metric, one could next consider other specific metric spaces, such as 2-dimensional Euclidean spaces or trees (on which the lower bound holds). Additionally, the two-sided matching problem has not been considered before our work under the metric distortion framework, and it is thus a natural problem to study for general metrics, especially as it seems to be somewhat easier than the one-sided variant (at least for the line, as we have shown).

Acknowledgments

Aris Filos-Ratsikas and Mohamad Latifian were supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/Y003624/1. Vasilis Gkatzelis and Emma Rewinski were supported by NSF CAREER award CCF-2047907 and NSF grant CCF-2210502.

References

[Abdulkadiroğlu *et al.*, 2005a] Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. The new york city high school match. *American Economic Review*, 95(2):364–367, 2005.

[Abdulkadiroğlu *et al.*, 2005b] Atila Abdulkadiroğlu, Parag A Pathak, Alvin E Roth, and Tayfun Sönmez. The boston public school match. *American Economic Review*, 95(2):368–371, 2005.

[Amanatidis *et al.*, 2022] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. A few queries go a long way: Information-distortion tradeoffs in matching. *Journal of Artificial Intelligence Research*, 74, 2022.

[Amanatidis *et al.*, 2024] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. Don't roll the dice, ask twice: The two-query distortion of matching problems and beyond. *SIAM Journal on Discrete Mathematics*, 38(1):1007–1029, 2024.

[Anari *et al.*, 2023] Nima Anari, Moses Charikar, and Prasanna Ramakrishnan. Distortion in metric matching with ordinal preferences. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC)*, pages 90–110, 2023.

[Anshelevich and Postl, 2017] Elliot Anshelevich and John Postl. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research*, 58:797–827, 2017.

[Anshelevich and Sekar, 2016a] Elliot Anshelevich and Shreyas Sekar. Blind, greedy, and random: Algorithms for matching and clustering using only ordinal information. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, pages 390–396, 2016.

[Anshelevich and Sekar, 2016b] Elliot Anshelevich and Shreyas Sekar. Truthful mechanisms for matching and clustering in an ordinal world. In *Proceedings of the 12th International Conference Web and Internet Economics (WINE)*, pages 265–278, 2016.

[Anshelevich and Zhu, 2019] Elliot Anshelevich and Wennan Zhu. Tradeoffs between information and ordinal approximation for bipartite matching. *Theory of Computing Systems*, 63(7):1499–1530, 2019.

[Anshelevich and Zhu, 2021] Elliot Anshelevich and Wennan Zhu. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. *ACM Transactions on Economics and Computation*, 9(2):9:1–9:24, 2021.

[Anshelevich *et al.*, 2018] Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264:27–51, 2018.

[Anshelevich *et al.*, 2021] Elliot Anshelevich, Aris Filos-Ratsikas, Nisarg Shah, and Alexandros A. Voudouris. Distortion in social choice problems: The first 15 years and beyond. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 4294–4301, 2021.

[Berger *et al.*, 2024] Ben Berger, Michal Feldman, Vasilis Gkatzelis, and Xizhi Tan. Learning-augmented metric distortion via (p, q)-veto core. In *Proceedings of the 25th ACM Conference on Economics and Computation, EC 2024, New Haven, CT, USA, July 8-11, 2024*, page 984. ACM, 2024.

[Bogomolnaia and Moulin, 2001] Anna Bogomolnaia and Hervé Moulin. A new solution to the random assignment problem. *Journal of Economic theory*, 100(2):295–328, 2001.

[Caragiannis *et al.*, 2022] Ioannis Caragiannis, Nisarg Shah, and Alexandros A. Voudouris. The metric distortion of multiwinner voting. *Artificial Intelligence*, 313:103802, 2022.

[Caragiannis *et al.*, 2024] Ioannis Caragiannis, Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, Kristoffer Arnsfelt Hansen, and Zihan Tan. Truthful facility assignment with resource augmentation: an exact analysis of serial dictatorship. *Mathematical Programming*, 203(1):901–930, 2024.

[Charikar and Ramakrishnan, 2022] Moses Charikar and Prasanna Ramakrishnan. Metric distortion bounds for randomized social choice. In *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2986–3004, 2022.

[Charikar *et al.*, 2024] Moses Charikar, Prasanna Ramakrishnan, Kangning Wang, and Hongxun Wu. Breaking the metric voting distortion barrier. In *Proceedings of the 35th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1621–1640, 2024.

[Downs, 1957] Anthony Downs. *An economic theory of democracy*, volume 28. Harper and Row, 1957.

[Ebadian and Shah, 2025] Soroush Ebadian and Nisarg Shah. Every bit helps: Achieving the optimal distortion with a few queries. In *Proceedings of the 39th Annual AAAI Conference on Artificial Intelligence (AAAI)*, 2025.

[Elkind and Faliszewski, 2014] Edith Elkind and Piotr Faliszewski. Recognizing 1-euclidean preferences: An alternative approach. In Ron Lavi, editor, *Proceedings of the 7th International Symposium Algorithmic Game Theory (SAGT)*, volume 8768, pages 146–157, 2014.

[Enelow and Hinich, 1984] James M. Enelow and Melvin J. Hinich. *The spatial theory of voting: An introduction*. CUP Archive, 1984.

[Feldman *et al.*, 2016] Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)*, pages 269–286, 2016.

[Filos-Ratsikas *et al.*, 2014] Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, and Jie Zhang. Social welfare in one-sided matchings: Random priority and beyond. In *Proceedings of the 7th International Symposium Algorithmic Game Theory (SAGT)*, pages 1–12, 2014.

[Filos-Ratsikas *et al.*, 2024] Aris Filos-Ratsikas, Panagiotis Kanellopoulos, Alexandros A. Voudouris, and Rongsen Zhang. The distortion of distributed facility location. *Artificial Intelligence*, 328:104066, 2024.

[Gale and Shapley, 1962] David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[Gkatzelis *et al.*, 2020] Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. Resolving the optimal metric distortion conjecture. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1427–1438, 2020.

[Han *et al.*, 2023] Yue Han, Christopher Jerrett, and Elliot Anshelevich. Optimizing multiple simultaneous objectives for voting and facility location. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5665–5672, 2023.

[Hotelling, 1929] Harold Hotelling. Stability in competition. *The Economic Journal*, 39(153):41–57, 1929.

[Hylland and Zeckhauser, 1979] Aanund Hylland and Richard Zeckhauser. The efficient allocation of individuals to positions. *Journal of Political economy*, 87(2):293–314, 1979.

[Kizilkaya and Kempe, 2022] Fatih Erdem Kizilkaya and David Kempe. Plurality veto: A simple voting rule achieving optimal metric distortion. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 349–355, 2022.

[Latifian and Voudouris, 2024] Mohamad Latifian and Alexandros A. Voudouris. The distortion of threshold approval matching. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2851–2859, 2024.

[Lykouris and Vassilvitskii, 2021] Thodoris Lykouris and Sergei Vassilvitskii. Competitive caching with machine learned advice. *Journal of the ACM*, 68(4):1–25, 2021.

[Ma *et al.*, 2021] Thomas Ma, Vijay Menon, and Kate Larson. Improving welfare in one-sided matchings using simple threshold queries. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 321–327, 2021.

[Merrill III *et al.*, 1999] Samuel Merrill III, Samuel Merrill, and Bernard Grofman. *A unified theory of voting: Directional and proximity spatial models*. Cambridge University Press, 1999.

[NRMP,] NRMP. The National Resident Matching Program. https://www.nrmp.org/. Accessed: 2025-01-23.

[Procaccia and Rosenschein, 2006] Ariel D. Procaccia and Jeffrey S. Rosenschein. The distortion of cardinal preferences in voting. In *Proceedings of the 10th International Workshop on Cooperative Information Agents (CIA)*, pages 317–331, 2006.

[Roth and Peranson, 1999] Alvin E Roth and Elliott Peranson. The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American economic review*, 89(4):748–780, 1999.

[Roth *et al.*, 2004] Alvin E Roth, Tayfun Sönmez, and M Utku Ünver. Kidney exchange. *The Quarterly journal of economics*, 119(2):457–488, 2004.

[Roth, 1984] Alvin E Roth. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of political Economy*, 92(6):991–1016, 1984.

[Shapley and Scarf, 1974] Lloyd Shapley and Herbert Scarf. On cores and indivisibility. *Journal of mathematical economics*, 1(1):23–37, 1974.

[Stokes, 1963] Donald E Stokes. Spatial models of party competition. *American political science review*, 57(2):368–377, 1963.

[Tamir, 2001] Arie Tamir. The k-centrum multi-facility location problem. *Discrete Applied Mathematics*, 109(3):293–307, 2001.

[UNOS,] UNOS. United Network for Organ Sharing. https://unos.org//. Accessed: 2025-01-23.

[Von Neumann and Morgenstern, 1944] J Von Neumann and O Morgenstern. Theory of games and economic behavior. 1944.

[Zhou, 1990] Lin Zhou. On a conjecture by gale about one-sided matching problems. *Journal of Economic Theory*, 52(1):123–135, 1990.