Distortion in metric matching with ordinal preferences

NIMA ANARI, Stanford University, USA
MOSES CHARIKAR, Stanford University, USA
PRASANNA RAMAKRISHNAN, Stanford University, USA

Suppose that we have n agents and n items which lie in a shared metric space. We would like to match the agents to items such that the total distance from agents to their matched items is as small as possible. However, instead of having direct access to distances in the metric, we only have each agent's ranking of the items in order of distance. Given this limited information, what is the minimum possible worst-case approximation ratio (known as the *distortion*) that a matching mechanism can guarantee?

Previous work by Caragiannis et al. [2016] proved that the (deterministic) Serial Dictatorship mechanism has distortion at most $2^n - 1$. We improve this by providing a simple deterministic mechanism that has distortion $O(n^2)$. We also provide the first nontrivial lower bound on this problem, showing that any matching mechanism (deterministic or randomized) must have worst-case distortion $\Omega(\log n)$.

In addition to these new bounds, we show that a large class of truthful mechanisms derived from Deferred Acceptance all have worst-case distortion at least $2^n - 1$, and we find an intriguing connection between *thin matchings* (analogous to the well-known thin trees conjecture) and the distortion gap between deterministic and randomized mechanisms.

CCS Concepts: • **Theory of computation** → **Algorithmic game theory and mechanism design**; *Approximation algorithms analysis*.

ACM Reference Format:

1 INTRODUCTION

Social choice theory aims to understand the ways in which the preferences of individuals over several alternatives can be aggregated and used to make a collective decision for the group. Typically the goal is to design mechanisms that maximize some notion of the total welfare of the participants or equivalently minimize some cost. One cost framework that has received increased attention over the last decade is *metric distortion* (see Anshelevich et al. [2021] for a detailed survey of the area). In this model, the individuals and the alternatives lie in a shared metric space, which is unknown to the mechanism. The participants indicate their preferences over the alternatives by giving an ordinal ranking of the alternatives in order of distance. Based on this information, the mechanism makes a collective decision with the goal of minimizing the *distortion* – the worst-case approximation ratio between the cost of the mechanism's choice and the optimal choice with full knowledge of the distances in the metric.

The metric distortion framework has been most well understood for elections, where the individuals represent voters, the alternatives represent candidates, and the goal is to select a single candidate with a small total distance to the voters. A long line of work [Anshelevich et al., 2015,

Anshelevich and Postl, 2017, Charikar and Ramakrishnan, 2022, Gkatzelis et al., 2020, Kempe, 2020a,b, Kizilkaya and Kempe, 2022, Munagala and Wang, 2019, Pulyassary and Swamy, 2021] pinned down the optimal distortion for deterministic mechanisms at 3, and showed that the optimal distortion for randomized mechanisms is between 2.1126 and $3 - o(1)$.

In this paper, we turn our attention to understanding the metric distortion model in a less understood but similarly ubiquitous social choice problem: minimum-cost matching. Now the individuals are *n agents*, and the alternatives are *n items* that we would like to match to the agents. The agents and items lie in a shared metric space, unknown to the mechanism, which only receives each agent's ranking of the items in order of distance. The cost of a matching is the total of the distances between the matched pairs, and once again the *distortion* is the worst-case approximation ratio between the cost of the mechanism's chosen matching and the optimal matching.

Previous work [Caragiannis et al., 2016] first studied this problem with respect to two well-known mechanisms: Serial Dictatorship (SD) and Random Serial Dictatorship (RSD). In these mechanisms, the agents are matched sequentially (either in a fixed order or a random order) to their favorite unmatched item. These mechanisms are particularly attractive because they are *truthful* – no agent ever has the incentive to misreport their preferences. Caragiannis et al. [2016] showed that SD guarantees distortion $2^n - 1$, and provided a matching lower bound. They also showed that RSD guarantees distortion n, and proved a lower bound of $\Omega(n^{0.29})$. This work motivates the following question, raised by Anshelevich et al. [2021].

Question 1 (Anshelevich et al. [2021, Open problem 2]). What is the distortion of the best ordinal rule for the minimum-cost metric matching problem? What if we also require truthfulness?

The metric distortion framework has been attractive since the expressive power of metric spaces allows it to reflect a host of real-world circumstances. In a variety of matching problems with a notion of distance, algorithms for Question 1 may be useful. Some examples include assigning students to schools, cars to parking spots, patients to doctors, employees to office buildings, jobs to servers, and so on. Note that in many of these examples, the items can accommodate more than one agent, but this can be reduced to the perfect matching setting by having several copies of an item corresponding to its capacity. It is also worth noting that oftentimes, the metric space is not completely hidden – for example, school districts often know where students live. However, the demand that the algorithm works for any metric space can provide a stronger guarantee that allows for greater flexibility for the agents. For example, if a family finds it more convenient that their child's school is closer to a parent's place of work rather than their home address, a mechanism that only asks for preferences over schools would be able to accommodate this. In other settings, participants may prefer not to reveal their location for privacy reasons. Beyond practical considerations, the metric distortion model gives us a new lens for understanding and evaluating the efficacy of existing social choice mechanisms, and also provides fertile ground motivating the development of novel mechanisms that might find use beyond this framework.

1.1 Our contributions and technical overview

New distortion bounds. Our main contributions are new upper and lower bounds for Question 1. In Section 3, we describe a simple, deterministic matching mechanism that guarantees distortion $O(n^2)$. This is an exponential improvement over SD. The idea of the mechanism is to partition the agents into sets, and for each set designate a representative who is close to all of the other agents in the set. This means that the top choices of the representative are good choices for other agents in the set. If two representatives have a common item among their top choices, then they must actually be close to each other, and so the sets can be merged with one of the two representatives becoming the new representative for the merged set. Otherwise, all of the top several choices of

the representatives are disjoint, and we can match the members of a set to the top choices of their representative.

In Section 4 we prove the first nontrivial lower bounds for Question 1, showing that any matching mechanism, regardless of whether or not it is randomized, must have distortion $\Omega(\log n)$ in the worst case. The lower bound is based on a metric space where the agents are the leaves of a binary tree, and the items are internal nodes of the tree. The key is that in every subtree, there is one fewer item than agents, and so any matching must match one agent outside of the subtree. By considering a variety of different underlying metric spaces, it can be arranged so that in the optimal matching, only one agent incurs any cost for being matched outside of their subtree, but the mechanism pays for one agent at each of the $\log n$ levels of the tree.

The impact of truthfulness on distortion. Truthfulness is a crucial feature of matching mechanisms in applications where agents report their own preferences. Since truthfulness is an additional constraint on a mechanism, a natural question is whether the optimal distortion for truthful mechanisms is worse than that for non-truthful mechanisms. We begin to explore this question in Section 5 by showing that a large class of deterministic truthful mechanisms have distortion at least $2^n - 1$, which is the same as the lower bound proven by Caragiannis et al. [2016] for SD.

Our proof works by identifying a general condition (Definition 1) on mechanisms where the SD lower bound instance due to Caragiannis et al. [2016] still applies. Interestingly, this condition captures a wide class of truthful mechanisms that are derived by artificially creating preferences for the items over the agents and running Deferred Acceptance. Such mechanisms are commonly used, especially in the school choice literature [Abdulkadiroğlu and Sönmez, 2003, Allman et al., 2022, Shi, 2022].

This result can be viewed either as an indication that the optimal distortion for truthful mechanisms could be substantially worse than that for non-truthful mechanisms or as a partial characterization of the way in which a deterministic truthful mechanism must look different from SD and Deferred Acceptance in order for it to have low distortion.

Thin matchings and rounding randomized mechanisms. In Section 6, we seek to understand whether any randomized mechanism with low distortion can be converted into a low distortion deterministic mechanism in a black box fashion. A priori, one may not expect this to be possible. For instance, Serial Dictatorship has exponentially worse distortion than Random Serial Dictatorship. However, we show that it is possible to do such a conversion, with a relatively small blow up in the distortion.

The general idea is to find a matching among agent-item pairs that the randomized mechanism is likely to match. If such a matching had high distortion, then the randomized mechanism would also have high distortion. We use this to show that any randomized mechanism with distortion D can be rounded to a deterministic mechanism with distortion Dn^2. Moreover, we show that the blow-up factor in this rounding scheme is characterized by what we call the *thin matching problem*. This problem asks if given a fractional matching, one can always find a perfect matching such that across any cut, the weight of the perfect matching is at most a small factor larger than the weight of the fractional matching. Thin matchings are analogous to the notion of *thin trees* which played a key role in recent developments improving approximation algorithms for the Asymmetric Traveling Salesman Problem [An et al., 2010, Anari and Gharan, 2015, Asadpour et al., 2017, Gharan and Saberi, 2011]. The thin matching problem is a clean combinatorial question and might be of independent interest.

Bounds for near-perfect matchings. The distortion of the deterministic mechanism in this work, as well as the best-known randomized mechanism [Caragiannis et al., 2016], are both polynomially

large: $O(n^2)$ and $O(n)$ respectively. In light of this, it is natural to ask if there is a way to relax the problem and achieve much better distortion. Indeed, one way to relax the problem called resource augmentation was studied in the prior work of Caragiannis et al. [2016]. Here we study another relaxation: instead of outputting a perfect matching, what if our algorithm is allowed to output a *near*-perfect matching, a matching between $1 - \varepsilon$ fraction of the agents and items?[1] In Section 7, we show that this relaxation allows us to lower the distortion significantly. For any constant $\varepsilon > 0$, there is a randomized mechanism, a simple truncation of the Random Serial Dictatorship, that achieves a distortion of $O(1)$. We also show that this mechanism can be modified to get a high-probability distortion guarantee instead of an in-expectation one. We show that there is a randomized mechanism that outputs a near-perfect matching M, that with high probability has distortion poly $\log(n)$ simultaneously on all metrics compatible with the input preferences. As a simple corollary, we also get deterministic, albeit exponential time, mechanisms that achieve distortion poly $\log(n)$.

For an alternate view of these results, note that if we have a near-perfect matching with cost c and a perfect matching with cost C, then we can extend the near-perfect matching to a perfect matching with cost $2c + C$. We simply overlay the two matchings, and match any previously unmatched agent to the unmatched item it is connected two via a path. This means that we can also view these results as giving a black box way to get a granular upgrade to any distortion result in the perfect matching setting. For instance, combining this with the mechanism in Section 3, we can obtain a deterministic mechanism where the total cost is within a $O(n^2)$ factor of the optimum cost, and where all but an ε fraction of the agents are matched at a cost within a poly $\log(n)$ factor of the optimum.

1.2 Related work

Since it was originally introduced by [Hylland and Zeckhauser, 1979], the problem of matching under incomplete information has been studied under a variety of contexts and models. [Filos-Ratsikas et al., 2014] first studied distortion in ordinal matching, but rather than considering costs induced by a metric space, they considered *unit-sum* and *unit-range* valuations (for each agent, their value of the items sums to 1 or their value for each item is in the interval $[0, 1]$ respectively). We note that the switch from costs to valuations means that the objective is maximum-matching rather than minimum-matching. In this setting, they showed that RSD guarantees distortion $O(\sqrt{n})$ and that this is the best possible for randomized mechanisms. For deterministic mechanisms with unit-sum valuations, Amanatidis et al. [2022b] showed that distortion $\Theta(n^2)$ is the best possible.

Another interesting variant considers unrestricted valuation functions but allows the mechanism to query a small number of the true valuations of each agent. Without these queries, one cannot achieve bounded distortion, but surprisingly with just a few queries per agent, the problem becomes tractable. Amanatidis et al. [2022b] showed that with $O(\lambda \log n)$ queries per agent, a mechanism can guarantee distortion $O(n^{1/(\lambda+1)})$, and Amanatidis et al. [2022a] showed that even with just two queries per agent, one can get distortion $O(\sqrt{n})$.

The metric assumption has also been applied to maximum matching with ordinal preferences rather than minimum matching. There, the best-known distortion bounds are small constants, even for truthful mechanisms. Anshelevich and Sekar [2016a,b] gave a mechanism that guarantees distortion 1.6 and a truthful mechanism that guarantees distortion 1.76 (both mechanisms are randomized).

---

[1]Note that the cost of the near-perfect matching is still compared to the optimal perfect matching, otherwise this would not be a relaxation.

Finally, Anshelevich and Zhu [2021] studied the min-cost ordinal matching problem in metric spaces where the locations of the agents are unknown, but the locations of the items are known. This is natural in some applications such as facility assignment. They showed that this setting is substantially easier, giving a deterministic mechanism with distortion 3 which is optimal. Their mechanism works by imagining that each agent is at the same location as their favorite item, and then outputting the optimal solution in this "projected" instance where all the distances are known.

2 PRELIMINARIES

Notation and terminology. Throughout the paper we will use $A = \{a_1, a_2, \ldots, a_n\}$ to denote the set of agents, and $B = \{b_1, b_2, \ldots, b_n\}$ to denote the set of items. An *ordinal matching problem instance* is given by n *preference lists*, one provided by each agent. Each preference list is a permutation of the items, given in increasing order of distance. Agent a_i *prefers* b_j over b_k if $d(a_i, b_j) \leq d(a_i, b_k)$, and a_i's *favorite* item in a set $S \subseteq B$ is $\arg\min_{b \in S} d(a_i, b)$. Similarly, a_i's favorite t items are the first t items on their preference list.

Metric spaces. Let $d : (A \cup B) \times (A \cup B) \rightarrow \mathbb{R}_{\geq 0}$ be the distance function of the metric space. By definition, d satisfies the following three properties.

(1) Identity of indiscernibles: $d(x, y) = 0$ if $x = y$,
(2) Symmetry: $d(x, y) = d(y, x)$,
(3) Triangle inequality: $d(x, y) \leq d(x, z) + d(z, y)$.

In certain definitions of metric spaces, it is required that different elements have positive distance, but we do not require this. This is justified by allowing different agents or items to "occupy" the same point in the metric space, in which case their distance is 0, but different points do indeed have positive distance. We note that just as easily, we can replace 0 distances with sufficiently small ε's without changing any results.

Matchings. We will use bijective functions $M : A \rightarrow B$ to represent matchings. Let

$$\text{cost}(M) := \sum_{i \in [n]} d(a_i, M(a_i))$$

denote the cost of the matching M. Let $M_{\text{OPT}} : A \rightarrow B$ denote optimal matching, and let $d_i = d(a_i, M_{\text{OPT}}(a_i))$ so that $\text{cost}(M_{\text{OPT}}) = \sum_i d_i$. With slight abuse of notation, we will allow M to also denote a set of matched pairs (often treated as edges in a weighted bipartite graph), so that $(a_i, b_j) \in M$ means that $M(a_i) = b_j$.

Given a problem instance, the *distortion* of a particular matching M is $\frac{\text{cost}(M)}{\text{cost}(M_{\text{OPT}})}$. The worst-case distortion of a deterministic matching mechanism is its maximum distortion over all problem instances (parametrized by n), and the worst-case distortion for a randomized mechanism is the maximum *expected* distortion ($\frac{\mathbb{E}[\text{cost}(M)]}{\text{cost}(M_{\text{OPT}})}$) over all problem instances.

When considering randomized mechanisms, we often find it more convenient to view the output as a *fractional* matching rather than a distribution over matchings. In particular, a randomized mechanism is defined by n^2 nonnegative real numbers $p_{i,j}$ for $i, j \in [n]$ (semantically, the probability that a_i is matched to b_j) such that for each j, $\sum_{i=1}^{n} p_{i,j} = 1$, and for each i, $\sum_{j=1}^{n} p_{i,j} = 1$. In this view, the cost of a matching is

$$\sum_{i=1}^{n} \sum_{j=1}^{n} p_{i,j} d(a_i, b_j) \tag{1}$$

3 DISTORTION UPPER BOUND

THEOREM 1. *There exists a deterministic matching mechanism that guarantees distortion $O(n^2)$.*

PROOF. We will maintain a partition of the agents into sets $S_1, S_2, \ldots$, and each set will have a representative $r_1, r_2, \ldots$ and so on. We will maintain the invariant that for every $a_i \in S_j$,

$$d(a_i, r_j) \leq -d_i + 2^{\lambda_j} \sum_{a_k \in S_j} d_k \qquad (2)$$

for some λ_j associated with each set S_j. Note that this implies that for each $a_i \in S_j$, $d(M_{\text{OPT}}(a_i), r_j) \leq d(M_{\text{OPT}}(a_i), a_i) + d(a_i, r_j) \leq 2^{\lambda_j} \sum_{a_k \in S_j} d_k$. It follows that there are at least $|S_j|$ distinct items b for which

$$d(b, r_j) \leq 2^{\lambda_j} \sum_{a_k \in S_j} d_k \qquad (3)$$

In particular, the above inequality holds when b is any of r_j's favorite $|S_j|$ items.

With this in mind, the matching mechanism is the following.

Algorithm 1. REPMATCH

Initially, there are n singleton sets $S_1, S_2, \ldots, S_n$, with $S_i = \{a_i\}$, $r_i = a_i$, and $\lambda_i = 0$.
While there exist two sets S_i and S_j such that some item is both one r_i's favorite $|S_i|$ items and one of r_j's favorite $|S_j|$ items:

- Replace S_i and S_j with $S_i \cup S_j$.
- The representative for $S_i \cup S_j$ is r_i if $\lambda_i \geq \lambda_j$ and r_j otherwise.
- The new λ for this set is $\max\{\lambda_i, \lambda_j\}$ if $\lambda_i \neq \lambda_j$, and $\lambda_i + 1$ otherwise.

Arbitrarily match the agents in each S_i to the top $|S_i|$ choices of r_i.

The key is in showing that the invariant (2) holds throughout the algorithm by observing how the λ's must change as sets merge. Clearly, the invariant holds at the start of the algorithm.

Let's suppose that we merge sets S_1 and S_2, and without loss of generality, $\lambda_1 \leq \lambda_2$. Since r_2 becomes the representative of the merged sets, we would like to bound the distance for any $a \in S_1$ to r_2. Suppose that b is an item that is both one of r_1's favorite $|S_1|$ items and one of r_2's favorite $|S_2|$ items. Then by (2) and (3),

$$d(a, r_2) \leq d(a, r_1) + d(r_1, b) + d(b, r_2) \leq \left(-d_i + 2^{\lambda_1} \sum_{a_k \in S_1} d_k \right) + 2^{\lambda_1} \sum_{a_k \in S_1} d_k + 2^{\lambda_2} \sum_{a_k \in S_2} d_k$$

$$\leq -d_i + 2^{\max\{\lambda_1 + 1, \lambda_2\}} \sum_{a_k \in S_1 \cup S_2} d_k$$

which means that our invariant is maintained, and the new λ for $S_1 \cup S_2$ is $\max\{\lambda_1 + 1, \lambda_2\}$, as specified by the algorithm.

This allows us to show another invariant – throughout the algorithm, $|S_j| \geq 2^{\lambda_j}$. This follows because when $\lambda_1 \leq \lambda_2$, $2^{\lambda_1} + 2^{\lambda_2} \geq 2^{\max\{\lambda_1 + 1, \lambda_2\}}$.

To conclude, the total cost of matching the elements of S_j to the top $|S_j|$ choices of r_j is at most the total distance from the elements of S_j to r_j plus the total distance from r_j to their top $|S_j|$ items. By (2) and (3), both of these terms are at most

$$|S_j| \cdot 2^{\lambda_j} \sum_{a_k \in S_j} d_k \leq |S_j|^2 \sum_{a_k \in S_j} d_k.$$

Therefore, the total cost of the final matching output by the algorithm is at most

$$\sum_j 2|S_j|^2 \sum_{a_k \in S_j} d_k \le 2 \max_j |S_j|^2 \text{cost}(M_{\text{OPT}}) \le 2n^2 \text{cost}(M_{\text{OPT}})$$

as desired.   □

4   DISTORTION LOWER BOUND

THEOREM 2. *Every (possibly randomized) matching mechanism has worst-case distortion* $\Omega(\log n)$.

PROOF. The goal will be to construct an ordinal matching instance with 2^k agents and items (i.e., specify the preference lists of the agents over the items), along with a collection of metric spaces that are consistent with the specified preferences, such that any matching mechanism must have distortion $\Omega(k)$ on one of the metric spaces in the collection.

Our matching instances are constructed as follows. To simplify the description of the preference lists, we will define a metric space such that the preference lists are easy to decipher by ordering distances in the metric space. The metric space is defined in terms of the distances between vertices on a weighted undirected graph. The graph is structured as a tree with 2^k leaves[2], where the root has one child, and every other internal vertex has two children. Note that this graph has 2^k internal vertices. The items occupy the internal vertices and the agents occupy the leaves. The weight of an edge (u, v) for which u is the parent of v is 2^{k-t}, where t is the distance from u to the root. See Figure 1 for a helpful diagram.
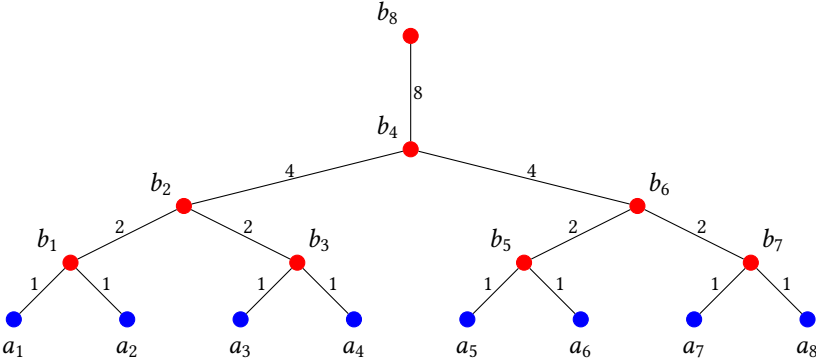

Fig. 1. The metric space that defines the preferences in our ordinal matching instance for $k = 3$. b_8 is the root of the tree. The red (internal) nodes are the items and the blue (leaf) nodes are the agents.

We will repeatedly refer to the structure of this tree as a tool for understanding the structure of the preference lists throughout the proof. The crucial property afforded to us by this construction is the following.

CLAIM 1. *Consider any vertex v (besides the root) whose parent is u. Then the agents in the subtree rooted at v prefer the items in that subtree over u and prefer u over all other items.*

To see the first part of the claim, note that for any agent a_i and item b_j in the subtree of v, we have that

$$d(a_i, u) = d(a_i, v) + d(v, u) > d(a_i, v) + d(v, b_j) \ge d(a_i, b_j)$$

---

[2]We define a leaf to be a vertex with no children and an internal vertex to be one with at least one child, so the root is not a leaf despite having degree 1.

using the fact that the distance from v to u is greater than the distance from v to any of the vertices in its subtree. The second part follows simply because for any agent a_i in the subtree of v, and any b_j not in this subtree, and not u, the path from a_i to b_j must go through u.

We now define a set of 2^k metric spaces such that any mechanism must have a large distortion with respect to one of these metric spaces. We will have one metric corresponding to each agent a_i. Once again, the metrics are defined by distances on a weighted undirected graph. We start with the graph defined above. Then, if P is the set of $k + 1$ items along the path from a_i to the root (including the root), we delete the edges between vertices in P and add edges from a_i to all the items in P. The newly added edges have a weight of 1, and the other edges have a weight of 0. See Figure 2.
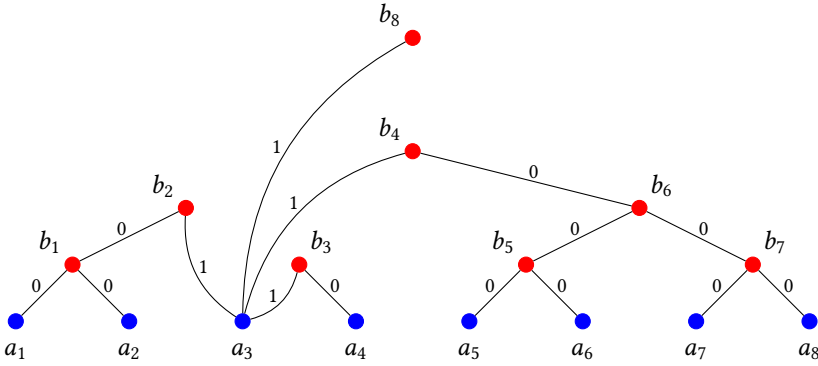

Fig. 2. The consistent metric corresponding to a_3.

We claim that these metrics are consistent with the preference lists derived before. Agent a_i has a distance of 1 to every item, so this is consistent with any preference list. For any other agent a_j, let u be the first common vertex on the paths from a_i and a_j to the root (in the original tree). Let v be the child of u such that a_j is in the subtree rooted at v. Then a_j has distance 0 to u and all the items in the subtree rooted at v, and distance 2 to all other items. By Claim 1 these distances are consistent with the preference list for a_j that we defined.

Finally, it remains to show that any mechanism must have a large distortion on one of these metrics. For each of these metrics, $\text{cost}(M_{\text{OPT}}) = 1$. This is because we can match a_i to the root (which is at distance 1), and then within every component where the distances are zero, there are an equal number of agents and items.

We will start by proving the result for deterministic mechanisms and then explain how to extend this to randomized mechanisms.

Suppose we have a matching M output by a deterministic mechanism. Consider the following process for choosing an agent. Eventually, the metric that causes a large distortion for the mechanism will be the metric corresponding to that agent.

- Initially the current vertex is the child of the root.
- Until we reach a leaf:
  - Let the left and right children of the current vertex be ℓ and r respectively.
  - In the subtree rooted at the current vertex, there is one more agent than item. Therefore, M matches some agent in this subtree to an item outside the subtree. Mark this agent *unlucky*.
  - If this agent is in the subtree rooted at ℓ, then update the current vertex to be r, and otherwise update it to be ℓ.
- Choose the agent at the leaf.

The process designates one unlucky agent at each step, so there are a total of k unlucky agents. Moreover, in the metric corresponding to the chosen agent, all of the unlucky agents are matched to an item that is distance 2 away. Hence, these contribute $2k$ to the cost of the mechanism, and the chosen agent also has a cost of 1, so the total cost is $2k + 1 = \Omega(k)$ as desired.

Now we will see how to extend this to randomized mechanisms which can now output a fractional matching rather than a perfect matching. We use a similar process to the one described above. The crucial fact will be that within the subtree of the current vertex (any vertex besides the root), the matching gives total weight at least 1 to pairs (a_i, b_j) where a_i is in the subtree but b_j is not. This is again because the subtree has one more agent than items.

Instead of saying that this comes from either the left subtree or the right subtree, we can say that at least $1/2$ of this weight must come from either the left or right subtree. If more of the weight comes from the left subtree, the next current vertex is the right child, and otherwise, it is the left child.

All of these pairs (a_i, b_j) have distance 2, so the total contribution to the cost at each step is at least 1. Since there are k steps, the total cost is $k + 1 = \Omega(k)$ (where the extra 1 again comes from the cost of the chosen agent, who has a distance of 1 to every item). □

5 TRUTHFULNESS

Consider the following class of deterministic matching mechanisms.

Definition 1. Say that a mechanism is *serializable* if there exist permutations π and σ of $[n]$ with the following property. If the mechanism is given an ordinal matching problem instance where for each $1 \le i < j \le n$, $a_{\pi(i)}$ prefers $b_{\sigma(i)}$ over $b_{\sigma(j)}$, then it matches $a_{\pi(i)}$ to $b_{\sigma(i)}$ for all i.

Intuitively, in these problem instances, the mechanism operates like a Serial Dictatorship where the agents choose in the order π, and σ defines the matching that is outputted. It is easy to see then that every deterministic Serial Dictatorship is serializable. We'll show that all serializable mechanisms fall victim to the same distortion lower bound that applies to Serial Dictatorship.

CLAIM 2. *If a mechanism is serializable, then its worst-case distortion is at least $2^n - 1$.*

PROOF. Our proof is essentially a redux of of the lower bound given by Caragiannis et al. [2016] for Serial Dictatorship.

Suppose that the underlying metric space is the real number line. For each i, the agent $a_{\pi(i)}$ sits at 2^{i-1}. For $i < n$, the item $b_{\sigma(i)}$ sits at 2^i, and $b_{\sigma(n)}$ sits at $-\varepsilon$.

The optimal matching would match $a_{\pi(1)}$ to $b_{\sigma(n)}$, and $a_{\pi(i)}$ to $b_{\sigma(i-1)}$ for $i > 1$, for a cost of $1 + \varepsilon$. However, Definition 1 tells us that for all i, $a_{\pi(i)}$ is matched to $b_{\sigma(i)}$. This match costs 2^{i-1} for $1 \le i < n$, and $2^{n-1} + \varepsilon$ for $i = n$. Thus, the total cost is $2^n - 1 + \varepsilon$. As ε tends to 0, the distortion tends to $2^n - 1$ (or equivalently, we can set $\varepsilon = 0$, and break ties in the preferences appropriately). □

Interestingly, the class of serializable mechanisms captures a wide variety of truthful mechanisms derived from Deferred Acceptance [Gale and Shapley, 1962].

Recall that Deferred Acceptance is a two-sided matching mechanism where participants on each side of the matching provide ranked preferences of the other side. Imagining that the items also provide a ranking over the agents, the algorithm works as follows. In rounds, each agent without a current match "proposes" to their favorite item that they have not yet proposed to. Then each item compares its proposals to its current match (if it has one), keeps its favorite, and rejects the rest. This continues until all of the agents are matched.

A natural way to convert this two-sided mechanism into a one-sided mechanism is to artificially introduce some preferences for the items. This can be done arbitrarily, or with some structure intended to encourage certain outcomes. For example, in the school choice literature [Abdulkadiroğlu

and Sönmez, 2003, Allman et al., 2022, Shi, 2022], this is used to prioritize students within walking distance from schools or students from historically disadvantaged neighborhoods. Moreover, since Deferred Acceptance is truthful from the side of the proposers, any one-sided mechanism derived in this way is also a truthful mechanism. This captures a wide variety of truthful mechanisms, by the sheer number of ways to fix each item's preferences.

However, no such mechanism can have good distortion, by the following.

THEOREM 3. *If a mechanism is implemented by Deferred Acceptance with fixed preferences for the items, it is serializable.*

PROOF. Suppose we run a Serial Dictatorship on the *items* (in the order $b_1, b_2, \ldots, b_n$), using their preferences to match them to agents. If b_i is matched to a_j, let $\pi(i) = j$. If we make σ the identity permutation, we claim that π and σ satisfy the property in Definition 1.

Suppose that per Definition 1, we are given an ordinal matching problem instance, where for each $1 \le i < j \le n$, $a_{\pi(i)}$ prefers b_i over b_j. We will show by induction on k that $a_{\pi(k)}$ gets matched to b_k.

For the base case, note that $a_{\pi(1)}$ and b_1 are each other's first choice, so $a_{\pi(1)}$ will propose to b_1 in the first round, and this match will be unbreakable.

Now suppose that we know that $a_{\pi(i)}$ will be matched to b_i for $i \le k$. Now consider $a_{\pi(k+1)}$. By the assumption in Definition 1, $a_{\pi(k+1)}$ prefers b_{k+1} over all of $b_{k+2}, \ldots, b_n$. This means that the only items that $a_{\pi(k+1)}$ may propose to before b_{k+1} are $b_1, b_2, \ldots, b_k$. But we know that these will be matched to $a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(k)}$ by the inductive hypothesis, and so any matches between $a_{\pi(k+1)}$ and b_i for $i \le k$ will eventually be broken.

Thus, $a_{\pi(k+1)}$ eventually proposes to b_{k+1}. Now, because of the Serial Dictatorship that was used to define π, we know that the only candidates that b_{k+1} might reject $a_{\pi(k+1)}$ for are $a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(k)}$ (those that were matched in the Serial Dictatorship before b_{k+1}'s turn). But each for each $i \le k$, $a_{\pi(i)}$ prefers b_i over b_{k+1}, so $a_{\pi(i)}$ would have to propose to b_i before b_{k+1}. Since we know that $a_{\pi(i)}$ is matched to b_i by the inductive hypothesis, $a_{\pi(i)}$ never proposes to b_{k+1}. This means that when $a_{\pi(k+1)}$ proposes to b_{k+1}, the match is made and never changes, which completes the induction.

Thus, our chosen π and σ satisfy the requirements of Definition 1, as desired.   □

6 ROUNDING RANDOMIZED MECHANISMS AND THIN MATCHINGS

Suppose that we have a randomized matching mechanism that guarantees some distortion D. Can we always round the fractional matching that it outputs to a perfect matching in a way that the cost of the perfect matching is not much larger than the cost of the fractional matching? In particular, we are concerned with the following question.

Question 2. What is the smallest $\alpha = \alpha(n)$ such that the following holds? If a fractional matching gives weight $p_{i,j}$ to each pair $(a_i, b_j) \in A \times B$, then there exists a perfect matching M such that for any metric d we have

$$\sum_{(a_i, b_j) \in M} d(a_i, b_j) \le \alpha \sum_{1 \le i, j \le n} p_{i,j} d(a_i, b_j).$$

If a randomized mechanism guarantees distortion D, then this means that on any input, the fractional matching it outputs satisfies $\sum_{1 \le i, j \le n} p_{i,j} d(a_i, b_j) \le D \cdot \text{cost}(M_{\text{OPT}})$. This means that if the above statement is true for some α, then any randomized mechanism that guarantees distortion D can be rounded to a deterministic mechanism with distortion αD.

Note that we can get $\alpha \le n^2$ quite easily. By Hall's theorem, the subgraph that consists of the edges (a_i, b_j) with $p_{i,j} \ge \frac{1}{n^2}$ must have a perfect matching[3], and any such matching M will have

$$\sum_{1 \le i, j \le n} p_{i,j} d(a_i, b_j) \ge \sum_{(a_i, b_j) \in M} p_{i,j} d(a_i, b_j) \ge \frac{1}{n^2} \sum_{(a_i, b_j) \in M} d(a_i, b_j).$$

In particular, since RSD guarantees distortion $O(n)$ [Caragiannis et al., 2016], by computing the matching probabilities of RSD and taking a matching from the edges that are matched with probability at least $\frac{1}{n^2}$, we can get a deterministic mechanism that gets distortion $O(n^3)$. One interesting feature of this mechanism is that it entirely offloads the problem of dealing with the preference lists to RSD.

It remains to see what the best factor of α we can guarantee is. We will show that up to an $O(\log n)$ factor, the optimal α is actually given by the following combinatorial question.

Question 3 (Thin matchings). What is the smallest $\beta = \beta(n)$ such that the following holds? Suppose we have a fractional matching of a balanced bipartite graph $G = (U \cup V, E)$ that gives weight $p_{u,v}$ to each edge (u, v). Then there exists a perfect matching M such that for any cut $(S, \overline{S})$, the weight of M across the cut is within a factor of β of the weight of the fractional matching across the cut. i.e.,

$$|M \cap (S \times \overline{S})| \le \beta \cdot \sum_{(u, v) \in S \times \overline{S}} p_{u,v}.$$

We say that a matching M satisfying the above inequality for all cuts is β-thin.

PROPOSITION 3. *Let α and β be the best answers to Question 2 and Question 3 respectively. Then*

$$\beta \le \alpha \le O(\log n) \cdot \beta.$$

PROOF. To prove that $\beta \le \alpha$, we need to show that if there exists a matching for Question 2 that gets a factor of α, then there exists a matching for Question 3 that gets a factor of α as well. Indeed, we can use the same matching, and just apply the guarantee for Question 2 to each of the cut metrics. These metrics assign distance 0 to edges within a cut, and 1 to those across the cut, i.e.,

$$d_S(u, v) := \mathbf{1}[(u, v) \in S \times \overline{S}].$$

Next, we prove that $\alpha \le O(\log n) \cdot \beta$. We use a well-known consequence of Bourgain's embedding [Bourgain, 1985], which says that any metric can be embedded as a conic combination of cut metrics with distortion $O(\log n)$. In particular, for any metric d, there exist non-negative weights γ_S such that

$$\sum_S \gamma_S d_S(u, v) \le d(u, v) \le O(\log n) \cdot \sum_S \gamma_S d_S(u, v).$$

---

[3]For any $S \subseteq A$, if $|N(S)| < |S|$, then $\sum_{a_i \in S} \sum_{b_j \notin N(S)} p_{i,j} \ge 1$. But this means for some $(a_i, b_j) \in S \times (B \setminus N(S))$ has $p_{i,j} \ge \frac{1}{n^2}$ which is a contradiction.

This means that if we can get a factor of β for Question 3, then we have for any d,

$$\sum_{(u,v)\in M} d(u,v) \leq O(\log n) \sum_{(u,v)\in M} \sum_S \gamma_S d_S(u,v)$$

$$= O(\log n) \sum_S \gamma_S |M \cap (S \times \overline{S})|$$

$$\leq O(\log n) \cdot \beta \sum_S \gamma_S \sum_{(u,v)\in S\times\overline{S}} p_{u,v}$$

$$= O(\log n) \cdot \beta \sum_S \gamma_S \sum_{u,v} p_{u,v} d_S(u,v)$$

$$= O(\log n) \cdot \beta \sum_{u,v} p_{u,v} \sum_S \gamma_S d_S(u,v)$$

$$\leq O(\log n) \cdot \beta \sum_{u,v} p_{u,v} d(u,v),$$

which means exactly that we can get a factor of $O(\log n) \cdot \beta$ for Question 2.   □

In summary, we see if we can prove strong bounds on Question 3, then we can round any randomized mechanism to a deterministic one without losing much in terms of distortion.

6.1   Relationship with thin trees

The definition of *thinness* in Question 3 closely mirrors the well-known notion of *thin trees* which concerns finding spanning trees, rather than perfect matchings, with few edges across each cut. A definition in the style of Question 3 is given below.

Definition 2 (Thin trees). Suppose we have a weighted graph G with edge weights $p_{u,v}$. We say that a spanning tree T is β-thin if for any cut $(S, \overline{S})$ of the graph,

$$|T \cap (S \times \overline{S})| \leq \beta \cdot \sum_{(u,v)\in S\times\overline{S}} p_{u,v}.$$

With the above definition, the thin tree conjecture [Goddyn, 2004] states that if the edge weights $p_{u,v}$ form a point in the spanning tree polytope,[4] then there exists an $O(1)$-thin tree T. We remark that in the literature this conjecture is sometimes stated with the alternative assumption that the graph G is 1-edge-connected, i.e., the weights $p_{u,v}$ sum to at least 1 across each cut. Up to constant factors in thinness, these assumptions are equivalent due to a classic result of Nash-Williams [1961].

Thin trees originated in the study of graph embeddings [Goddyn, 2004], where they were proposed as a path towards Jaeger's flow conjecture [Jaeger, 1984, 1988]. They received significant attention much later on when Asadpour et al. [2017] showed that an efficient algorithm for finding β-thin trees implies an $O(\beta)$ approximation algorithm for the Asymmetric Traveling Salesman Problem (ATSP). They also gave such an algorithm for $\beta = O(\log n/\log\log n)$. Anari and Gharan [2015] gave a proof that $O(\text{poly}\log\log n)$-thin trees always exist, though their proof is non-constructive. The thin trees conjecture postulates that $O(1)$-thin trees always exist (see for example Anari and Gharan [2015, Conjecture 1.3] and the discussion that follows). See An et al. [2010], Gharan and Saberi [2011] for some other applications of thin trees in the ATSP literature.

Just as thin trees have found surprising relevance in very different areas, we find it similarly intriguing that thin matchings are connected to the problem of ordinal metric matching, which on

---

[4]That is, the edge weights are a convex combination of indicators of spanning trees, analogous to a fractional matching, but for spanning trees.

the surface seems unrelated. Note that in Question 3, the notions of preference lists and metric spaces are completely abstracted away, even though they are central to the original problem .

In light of the strong known bounds for thin trees, and the thin trees conjecture, we make the following analogous conjecture.

CONJECTURE 1. *The optimal β for Question 3 is $O(1)$.*

Since Proposition 3 already has a logarithmic factor, for our rounding application an even weaker version of the above conjecture that just shows a $O(\text{poly} \log n)$ bound would suffice – it implies that the optimal randomized and deterministic mechanisms for ordinal metric matching have the same distortion up to poly log factors.

A natural approach to proving strong bounds for Question 3 would be to adapt the tools used to make progress on the thin tree conjecture. In particular, the technique used by Asadpour et al. [2017] was to write the weights $p_{u,v}$ as the mean of a distribution μ over spanning trees – this is possible by the assumption that $p_{u,v}$ is in the spanning tree polytope. Amongst all possible μ with this mean, Asadpour et al. [2017] pick the one with the maximum entropy. This ensures that for a random sample $T \sim \mu$, each cut value $|T \cap (S \times \overline{S})|$ is highly concentrated around its mean, and by union bounding tail deviation bounds over all the cuts, $O(\log n / \log \log n)$-thinness follows.

An analogous attempt would be to similarly write a fractional matching as the mean of a distribution over perfect matchings that exhibits concentration properties. However, the challenge with matchings is that distributions over them, even max-entropy ones, exhibit strong correlations that prevent useful concentration inequalities. We demonstrate this with an example depicted in Figure 3. Consider a graph G which is a cycle of length $4k$ with vertices numbered $1, 2, \ldots, 4k$ in the order they appear on the cycle. Let $p_{u,v} = 1/2$ on every edge of the cycle – a fractional matching. There are exactly two perfect matchings in G, M_{odd} formed by the odd edges, and M_{even} formed by the even edges. There is a unique distribution μ over perfect matchings with mean p: μ has to put $1/2$ mass on each one of M_{odd} and M_{even}. Now, consider the cut $(S, \overline{S})$ where $S = \{1, 2, 5, 6, \ldots, 4i+1, 4i+2, \ldots\}$. This cut has all the edges $(2i, 2i+1)$ and none of the edges $(2i+1, 2i+2)$. For a random sample $M \sim \mu$, the value of the cut $|M \cap (S \times \overline{S})|$ is uniformly distributed over the set $\{0, 2k\}$, which shows that no meaningful concentration around the mean holds here.

While this example rules out strong concentration around the mean, both M_{odd} and M_{even} still produce cut values within a factor of 2 of the mean, which tempts one to guess that a perhaps weaker multiplicative tail bound can be used to prove Conjecture 1. However, we can modify this example to make it even worse. Let $q \in [0, 1]$, and let $p_{u,v}$ be the convex combination of M_{odd} and M_{even} with coefficients q and $1 - q$ respectively. Then with probability q, the sampled perfect matching will have a cut value $1/q$ times as large as the fractional matching. We can boost this failure probability q by parallel repetition. Consider a disjoint union of $n/4k$ cycles of length $4k$ each, where on each cycle the edge weights alternate between q and $1 - q$. A max-entropy distribution would independently choose the matching in each cycle. Thus with probability at least $1 - (1-q)^{n/4k}$ we will get a cut where the sampled perfect matching has $1/q$ times as many edges as the fractional matching. We can set $k = \Theta(1)$ and $q = \Theta(1/n)$ to get a blow-up by a factor of $\Theta(n)$ in some cut with overwhelming probability.

7 NEAR-PERFECT MATCHINGS

In this section, we study a relaxation of the problem, where our mechanism is allowed to output a non-perfect matching M, matching a subset of agents to a subset of items. We still measure the performance of the mechanism in terms of distortion $\left(\frac{\text{cost}(M)}{\text{cost}(M_{\text{OPT}})}\right)$ or expected distortion $\left(\frac{\mathbb{E}[\text{cost}(M)]}{\text{cost}(M_{\text{OPT}})}\right)$, where M_{OPT} is the optimal *perfect* matching.
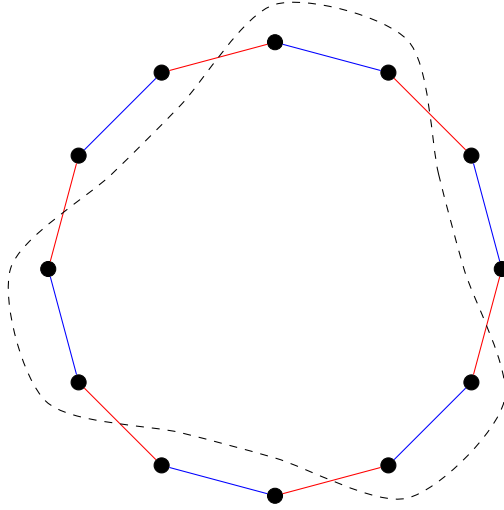
Fig. 3. The uniform distribution over the blue and red perfect matchings shows no meaningful concentration for the number of edges of a sample that cross the indicated cut.

Our first result shows that a simple truncation of the Random Serial Dictatorship (RSD) rule achieves $O(1)$ distortion while matching a $1 - \varepsilon$ fraction of the agents to items. Random Serial Dictatorship was studied by Caragiannis et al. [2016] for the non-relaxed problem and was shown to achieve expected distortion $\leq n$. Our analysis of the truncated version follows closely the ideas in their analysis.

Algorithm 2. *TRUNCATEDRSD*

The goal is to match m out of n agents to items.
We start with an empty matching. While fewer than m agents are matched:

- Pick an agent a_i uniformly at random from all unmatched agents.
- Match a_i to their most preferred unmatched item b_j.

THEOREM 4. *If M is the matching output by TruncatedRSD, then*

$$\mathbb{E}[\text{cost}(M)] \leq \frac{m}{n + 1 - m} \cdot \text{cost}(M_{\text{OPT}}).$$

Note that by setting $m = (1 - \varepsilon)n$, we get expected distortion $\leq (1 - \varepsilon)/\varepsilon = O_\varepsilon(1)$, yielding the following simple corollary.

COROLLARY 4. *For every constant $\varepsilon > 0$, there is a randomized mechanism that outputs a matching M of size $(1 - \varepsilon)n$ with expected distortion $O(1)$.*

PROOF OF THEOREM 4. Let M_i be the partial matching after the first i agents have been matched. We will show by induction that there exists a (random) matching N_i between the $n - i$ unmatched agents and the $n - i$ unmatched items ($M_i \cup N_i$ is a perfect matching) such that $\mathbb{E}[\text{cost}(N_i)] \leq \frac{n+1}{n+1-i} \cdot \text{cost}(M_{\text{OPT}})$ and $\mathbb{E}[\text{cost}(M_i)] \leq \frac{i}{n+1-i} \cdot \text{cost}(M_{\text{OPT}})$. At the end, by setting $i = m$, we get that $\text{cost}(M_m) \leq \frac{m}{n+1-m} \text{cost}(M_{\text{OPT}})$ as desired.

For the base case $i = 0$, we can take N_0 to be M_{OPT} and note that M_0 is the empty matching. Clearly $\text{cost}(N_0) = \frac{n+1}{n+1} \text{cost}(M_{\text{OPT}})$ and $\text{cost}(M_0) = 0$.

Now assume that we have proved the claim for i. We will prove it for $i + 1$. In iteration $i + 1$, the agent we pick, let us call it a, is uniformly randomly picked from all unmatched ones, i.e., those agents matched in N_i. Now a gets matched to their top available item; the cost of this edge is at most the cost of edge incident to a in N_i. Therefore the average cost of this edge, averaged over all unmatched a, is upper-bounded by $\text{cost}(N_i)/(n - i)$. Therefore

$$\mathbb{E}[\text{cost}(M_{i+1})] \le \mathbb{E}[\text{cost}(M_i)] + \frac{\mathbb{E}[\text{cost}(N_i)]}{n - i} \le \left(\frac{i}{n + 1 - i} + \frac{n + 1}{(n + 1 - i)(n - i)} \right) \text{cost}(M_{\text{OPT}})$$

$$= \frac{i + 1}{n - i} \cdot \text{cost}(M_{\text{OPT}}).$$

It remains to construct N_{i+1} and bound its cost. Suppose that agent a gets matched to the item b by the mechanism. Since both a and b are unmatched at the beginning of the iteration, they must both be endpoints of edges in N_i. Let the other endpoints of these edges be b' and a' respectively. Now we simply let N_{i+1} be the matching obtained from N_i by deleting the edges (a, b') and (a', b), and instead adding the edge (a', b'); if $a = a'$ and $b = b'$, we simply delete the edge. It is easy to see that N_{i+1} is a matching with one fewer edge that matches all agents and items still available after this iteration. It remains to bound its cost. Here we use the triangle inequality: $d(a', b') \le d(a, b') + d(a, b) + d(a', b)$ to get that $\text{cost}(N_{i+1}) \le \text{cost}(N_i) + d(a, b)$; the inequality is still valid in the special case where $a = a'$ and $b = b'$. We already established that $\mathbb{E}[d(a, b)] \le \mathbb{E}[\text{cost}(N_i)]/(n - i)$. Therefore

$$\mathbb{E}[\text{cost}(N_{i+1})] \le \left(1 + \frac{1}{n - i} \right) \text{cost}(N_i) \le \frac{n + 1}{n - i} \cdot \text{cost}(M_{\text{OPT}}),$$

as desired.   □

Next, we show that by losing some logarithmic factors, we can get a mechanism whose output near-perfect matching M satisfies a much stronger guarantee: with high probability, it has *distortion* (not expected distortion) at most $O(\log^2(n))$ simultaneously w.r.t. all metrics d compatible with the input preference rankings. Note that this is not a trivial implication of Theorem 4. A randomized mechanism is allowed to hedge its bets, i.e., choose a mixed strategy against an adversary who chooses the worst metric d. A priori, the gap between pure strategies and mixed strategies can be very large.

THEOREM 5. *For any constant $\varepsilon > 0$, there exists a computationally efficient randomized mechanism that outputs a matching M of size $(1 - \varepsilon)n$ such that with probability $\ge 1 - 1/\text{poly}(n)$, M has distortion at most $O(\log^2(n))$ w.r.t. the optimal perfect matching in all metrics d compatible with the input preferences.*

PROOF. If we run TruncatedRSD with $m = (1 - \varepsilon/10)n$, at the end we get a random matching of size m with expected distortion $O(1/\varepsilon)$. Let x be the fractional (non-perfect) matching which is the expectation of the output of TruncatedRSD. In other words, x is a vector indexed by pairs of agents and items, such that $x_{a,b} = \mathbb{P}[a \text{ matched to } b]$. We know that

$$\text{cost}(x) = \sum_{(a,b) \in A \times B} d(a, b)x_{a,b} = \mathbb{E}[\text{cost(output of TruncatedRSD)}] \le O(1/\varepsilon) \cdot \text{cost}(M_{\text{OPT}}).$$

Our goal is to round x to an integral matching, without blowing up the cost by much. We employ a strategy based on thin matchings, similar to the proof of Proposition 3. Before we explain the rounding, we have to deal with a technical issue. We do not know how to efficiently compute the exact vector x; it is the expectation of the output of TruncatedRSD, but naively we have to enumerate $\simeq n!$ permutations describing the order in which the agents are picked, in order to compute x. Instead, we run TruncatedRSD h independent times, and we let x' be the average of the

indicators of the h output matchings. Instead of working with x, we work with x'. As $h \to \infty$, almost surely $x' \to x$. We will eventually set h to be only polynomially large, making the mechanism computationally efficient.

As in the proof of Proposition 3, we use Bourgain's embedding [Bourgain, 1985] to write, up to a distortion of $O(\log(n))$, a hypothetical metric d as a conic combination of cut metrics d_S. It is enough to find a (non-perfect) matching M that guarantees for every subset S of our metric space:

$$\frac{M(S, \overline{S})}{x(S, \overline{S})} := \frac{|M \cap (S \times \overline{S})|}{\sum_{(u,v) \in S \times \overline{S}} x_{u,v}} \leq \beta.$$

An argument very similar to the proof of Proposition 3 would then show that the distortion of M w.r.t. any metric compatible with the input rankings is $\leq O(\beta \log n/\varepsilon)$.

We construct this M by employing a randomized rounding mechanism due to Chekuri et al. [2011]. They showed that for any point y in the matching polytope and a desired parameter $\gamma \in (0, 1/2]$, one can efficiently produce a random matching M such that $\mathbb{E}[\mathbf{1}_M] = (1 - \gamma)y$ and for any vector $v \in [0, 1]^{A \times B}$, the value $\langle v, \mathbf{1}_M \rangle$ is tightly concentrated around its mean; in particular if $\mu \geq \langle v, y \rangle$ and $t \geq 2$, then

$$\mathbb{P}[\langle v, \mathbf{1}_M \rangle \geq t\mu] \leq \exp(-\mu\gamma(2t - 3)/20).$$

We would like to apply the above concentration bound with y set to be x' and v set to be the indicator of edges in a cut $(S, \overline{S})$; this means $\langle v, \mathbf{1}_M \rangle = |M \cap (S \times \overline{S})|$, the exact quantity we would like to control. However, we need to union bound over all such cuts and there are exponentially many. A common strategy here, also used in the context of the thin tree problem [Asadpour et al., 2017], is to use the well-known fact due to Karger [Karger, 1993] that the number of near-minimum cuts in any weighted graph is only polynomially large. In particular, if we define c to be the minimum cut

$$c := \min_{S \neq \emptyset, A \cup B} y(S, \overline{S}),$$

then the number of sets S where $y(S, \overline{S}) \leq \ell \cdot c$ is at most $(2n)^{2\ell}$ for any integer $\ell \geq 1$.

There is a problem with this approach. Unlike in the case of the thin tree problem, there is no a priori lower bound on the minimum cut c of x' or even x, which are fractional matchings, and not fractional spanning trees. Indeed, the minimum cut in x and/or x' can be 0, and the number of cuts of size 0 can be exponentially large; think of a fractional matching consisting of $\Omega(n)$ disjoint connected components. We overcome this issue by manually removing all small cuts.

Starting with x' we produce a new fractional matching x'' as follows: every time there is a cut $(S, \overline{S})$ with value in $(0, \varepsilon/10)$, we simply remove all the edges in that cut, i.e., set their value to 0. We repeat until no such cut can be found and we call the result x''. Note that every removal of a cut increases the number of connected components (counted w.r.t. the nonzero edges) by one, so the number of iterations can be at most $2n$. In particular, $\sum_e x''_e$ is only smaller than $\sum_e x'_e$ by at most $2n \cdot \varepsilon/10 = \varepsilon n/5$. Thus, x'' is a fractional matching of size at least $(1 - \varepsilon/10)n - \varepsilon n/5 = (1 - 3\varepsilon/10)n$.

We apply the rounding mechanism of Chekuri et al. [2011] with $y = x''$ and $\gamma = \varepsilon/10$ to get our random matching M. Note that the expected size of M is at least $(1 - \gamma)(1 - 3\varepsilon/10)n \geq (1 - \varepsilon/2)n$. Since the size of M can never be larger than n, by Markov's inequality, with probability at least $1/2$, the size of M will be $\geq (1 - \varepsilon)n$. Note that we can boost this probability by repeatedly applying the rounding mechanism of Chekuri et al. [2011] a number of times; we will show that the probability that M is not thin is negligibly small, so it remains negligibly small even after conditioning on $|M| \geq (1 - \varepsilon)n$. So repeating the process, at most a logarithmic number of times, until the first time we see a matching of size $\geq (1 - \varepsilon)n$, gives us the desired result. It just remains to show the probability of M not being thin is negligibly small.

The crucial observation is that it suffices to only consider cuts which nontrivially separate *exactly one* connected component of x''. We call such cuts *fundamental cuts*. More precisely, if $C_1, \ldots, C_k$ are the connected components of x'', then a fundamental cut is a pair (S, T) of disjoint sets such that $S \cup T = C_i$ for some C_i. By Karger's arguments, the number of fundamental cuts (S, T) with $x''(S, T) \leq \ell\varepsilon/10$ is bounded by $\sum_i |C_i|^{2\ell} \leq (2n)^{2\ell}$. We claim that if for each fundamental cut (S, T) we have, for a logarithmically large β to be set later, that

$$M(S, T) \leq \beta x(S, T),$$

then the same inequality holds over *all* global cuts $(U, \overline{U})$, that is

$$M(U, \overline{U}) \leq \beta x(U, \overline{U}).$$

Consider such a global cut; for each component C_i, the cut $(U, \overline{U})$ induces a fundamental cut (S_i, T_i).[5] Note that the edges of the matching M are a subset of the support of x'', because $\mathbb{E}[\mathbf{1}_M] = (1 - \gamma)x''$. This means that

$$M(U, \overline{U}) = \sum_i M(S_i, T_i) \leq \sum_i \beta x(S_i, T_i) \leq \beta x(U, \overline{U}),$$

where the last inequality follows from the fact that any edge with endpoints in S_i and T_i must cross $(U, \overline{U})$.

So it remains to show that the inequality $M(S, T) \leq \beta x(S, T)$ holds for all fundamental cuts with high probability. We derive these by chaining $M(S, T) \leq (\beta/2)x''(S, T)$ with $x''(S, T) \leq x'(S, T) \leq 2x(S, T)$.

The inequality $x'(S, T) \leq 2x(S, T)$ holds with high probability if the number of independent trials h of TruncatedRSD that we run to obtain x' is large enough. In particular, the value of the output matching of TruncatedRSD in the fundamental cut (S, T) is a random variable supported on $[0, n]$ with mean $x(S, T)$. So by the standard Hoeffding's inequality, we have

$$\mathbb{P}[x'(S, T) - x(S, T) \geq t] \leq \exp(-\Omega(ht^2/n^2)).$$

For any fundamental cut, we have $x'(S, T) \geq \varepsilon/10$. So if we set $t = \varepsilon/20$, we get

$$\mathbb{P}[x'(S, T) > 2x(S, T)] \leq \mathbb{P}[x'(S, T) - x(S, T) > \varepsilon/20] \leq \exp(-\Omega(h\varepsilon^2/n^2)).$$

There are at most $2^{O(n)}$ fundamental cuts, so we have

$$\mathbb{P}[\exists \text{fundamental cut } (S, T) : x'(S, T) > 2x(S, T)] \leq \exp(O(n) - \Omega(h\varepsilon^2/n^2)).$$

It is enough to set $h = \text{poly}(n/\varepsilon)$ for the above to be negligibly small: $\exp(-\text{poly}(n))$.

It remains to show that with high probability for all fundamental cuts (S, T) we have $M(S, T) \leq (\beta/2)x''(S, T)$. Here we use the tail bound of [Chekuri et al., 2011] together with the bound on the number of fundamental cuts of various sizes. Consider an integer $\ell \geq 1$. The number of fundamental cuts (S, T) with $x''(S, T) \in [\ell\varepsilon/10, (\ell + 1)\varepsilon/10)$ is at most $n^{O(\ell)}$. For each such cut, we get that

$$\mathbb{P}[M(S, T) \geq \beta x''(S, T)/2] \leq \mathbb{P}[M(S, T) \geq \beta\ell\varepsilon/20] \leq \exp(-\Omega(\beta\ell\varepsilon^2)).$$

So union bounding over all such cuts, and then summing over all ℓ, we get

$$\mathbb{P}[\exists \text{fundamental cut } (S, T) : M(S, T) \geq \beta x''(S, T)/2] \leq \sum_{\ell=1}^{\infty} n^{O(\ell)} \exp(-\Omega(\beta\ell\varepsilon^2)).$$

---

[5]The cut (S_i, T_i) could be trivial, i.e., one of S_i or T_i could be $\emptyset$, but the arguments still hold.

For any constant $C > 0$, there is some constant C' such that if we set $\beta = C' \log n / \varepsilon^2$, we can bound the above as

$$\sum_{\ell=1}^{\infty} \exp(O(\ell \log n) - \Omega(C' \ell \log n)) \leq \sum_{\ell=1}^{\infty} \exp(-C\ell \log n) = O(1/n^C). \qquad \square$$

Note that Theorem 5 also implies the existence of a single near-perfect matching whose distortion is at most $O(\log^2(n))$, a nontrivial fact by itself. As a simple corollary, we get a deterministic, albeit not computationally efficient, mechanism with a similar distortion.

COROLLARY 5. *For any constant $\varepsilon > 0$, there exists a deterministic exponential time mechanism that outputs a matching of size $(1 - \varepsilon)n$ guaranteed to have distortion at most $O(\log^2 n)$.*

PROOF. We can simply compute the expected value x of the TruncatedRSD mechanism, then enumerate all near-perfect matchings M, and for each one check whether $M(S, \overline{S}) \leq O(\log^2 n)x(S, \overline{S})$ by enumerating all cuts S. We output the first matching M which satisfies this thinness condition. By Theorem 5, such a near-perfect matching is guaranteed to exist. $\qquad \square$

8 DISCUSSION

We discuss some qualitative interpretations of our work, as well as various future directions. For ease of exposition, we split Question 1 into two open problems separating the question of truthfulness.

OPEN PROBLEM 1. *What is the optimal metric distortion for deterministic and randomized mechanisms for min-cost ordinal matching?*

On this problem, we narrow the interval of the optimal distortion for deterministic matchings to $[\Omega(\log n), O(n^2)]$ and for randomized matchings to $[\Omega(\log n), O(n)]$ (using the upper bound proven by [Caragiannis et al., 2016] for RSD). Note that our super-constant lower bound contrasts with several other related metric distortion problems (matching with known item locations, maximum matching, elections) where constant upper bounds are known. Our work in Section 6 also suggests that it could be quite possible that the optimal distortions for deterministic and randomized mechanisms are the same up to log factors.

A priori, it is difficult to tell which of the two bounds is closer to the right answer. The lower bound could potentially be improved by considering hard metrics with a greater range of distances (ours only use 0, 1, and 2), but one challenge is that larger distances constrain the preference lists more severely. This makes it difficult to have a variety of very different possible underlying metric spaces that all have large distances in different places. On the other hand, the upper bound could be improved by formalizing this intuition for why lower bounds are difficult into a more clever charging scheme.

OPEN PROBLEM 2. *What is the optimal metric distortion for deterministic and randomized* truthful *mechanisms for min-cost ordinal matching?*

We showed that a large class of deterministic mechanisms, which capture a number of known truthful mechanisms, are all as bad as SD in terms of distortion. Could it be that SD is the best deterministic truthful mechanism for this problem?

It is not hard to see that for a mechanism to be truthful, it must give each agent a set of items that only depends on the other agent's preferences, and then assign the agent their favorite item in the set (the set is exactly all of the items that the agent could be matched to, over all of their possible preference lists). The challenge is that it is difficult to conjure such mechanisms that aren't serializable since the most natural way to prevent an agent from influencing their own set is for

decisions to be made in some sequential manner. One can find with a computer search that truthful non-serializable mechanisms do exist, but they are unnatural.

As a counterpoint, part of the issue may just be a lack of imagination, where our attempts to design truthful mechanisms tend to skew toward serialization. Resolving this question may require the development of exciting new kinds of truthful matching mechanisms.

ACKNOWLEDGEMENTS

REFERENCES

Atila Abdulkadiroğlu and Tayfun Sönmez. 2003. School choice: A mechanism design approach. *American economic review* 93, 3 (2003), 729–747.

Maxwell Allman, Itai Ashlagi, Irene Lo, Juliette Love, Katherine Mentzer, Lulabel Ruiz-Setz, and Henry O'Connell. 2022. Designing School Choice for Diversity in the San Francisco Unified School District. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 290–291.

Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A Voudouris. 2022a. Don't Roll the Dice, Ask Twice: The Two-Query Distortion of Matching Problems and Beyon. *arXiv preprint arXiv:2203.01872* (2022).

Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A Voudouris. 2022b. A few queries go a long way: information-distortion tradeoffs in matching. *Journal of Artificial Intelligence Research* 74 (2022), 227–261.

Hyung-Chan An, Robert D Kleinberg, and David B Shmoys. 2010. Approximation algorithms for the bottleneck asymmetric traveling salesman problem. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Springer, 1–11.

Nima Anari and Shayan Oveis Gharan. 2015. Effective-resistance-reducing flows, spectrally thin trees, and asymmetric TSP. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*. IEEE, 20–39.

Elliot Anshelevich, Onkar Bhardwaj, and John Postl. 2015. Approximating Optimal Social Choice under Metric Preferences. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 29.

Elliot Anshelevich, Aris Filos-Ratsikas, Nisarg Shah, and Alexandros A Voudouris. 2021. Distortion in Social Choice Problems: The First 15 Years and Beyond. *arXiv preprint arXiv:2103.00911* (2021).

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

Elliot Anshelevich and Shreyas Sekar. 2016a. Blind, greedy, and random: Algorithms for matching and clustering using only ordinal information. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Elliot Anshelevich and Shreyas Sekar. 2016b. Truthful mechanisms for matching and clustering in an ordinal world. In *International Conference on Web and Internet Economics*. Springer, 265–278.

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

Arash Asadpour, Michel X Goemans, Aleksander Madry, Shayan Oveis Gharan, and Amin Saberi. 2017. An $O(\log n/\log \log n)$-approximation algorithm for the asymmetric traveling salesman problem. *Operations Research* 65, 4 (2017), 1043–1061.

Jean Bourgain. 1985. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel Journal of Mathematics* 52, 1 (1985), 46–52.

Ioannis Caragiannis, Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, Kristoffer Arnsfelt Hansen, and Zihan Tan. 2016. Truthful facility assignment with resource augmentation: An exact analysis of serial dictatorship. In *International Conference on Web and Internet Economics*. Springer, 236–250.

Moses Charikar and Prasanna Ramakrishnan. 2022. Metric Distortion Bounds for Randomized Social Choice. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 2986–3004.

Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. 2011. Multi-budgeted matchings and matroid intersection via dependent rounding. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*. SIAM, 1080–1097.

Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, and Jie Zhang. 2014. Social welfare in one-sided matchings: Random priority and beyond. In *International Symposium on Algorithmic Game Theory*. Springer, 1–12.

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

Shayan Oveis Gharan and Amin Saberi. 2011. The asymmetric traveling salesman problem on graphs with bounded genus. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*. SIAM, 967–975.

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

Luis A. Goddyn. 2004. Some Problems I Like. (2004). https://www.sfu.ca/~goddyn/Problems/problems.html

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

François Jaeger. 1984. On circular flows in graphs. In *Finite and infinite sets*. Elsevier, 391–402.

François Jaeger. 1988. Nowhere-zero flow problems. *Selected topics in graph theory* 3 (1988), 71–95.

David R Karger. 1993. Global Min-cuts in RNC, and Other Ramifications of a Simple Min-Cut Algorithm.. In *Soda*, Vol. 93. 21–30.

David Kempe. 2020a. An analysis framework for metric voting based on LP duality. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 2079–2086.

David Kempe. 2020b. Communication, distortion, and randomness in metric voting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 2087–2094.

Fatih Erdem Kizilkaya and David Kempe. 2022. Plurality veto: A simple voting rule achieving optimal metric distortion. *arXiv preprint arXiv:2206.07098* (2022).

Kamesh Munagala and Kangning Wang. 2019. Improved metric distortion for deterministic social choice rules. In *Proceedings of the 2019 ACM Conference on Economics and Computation*. 245–262.

C St JA Nash-Williams. 1961. Edge-disjoint spanning trees of finite graphs. *Journal of the London Mathematical Society* 1, 1 (1961), 445–450.

Haripriya Pulyassary and Chaitanya Swamy. 2021. On the Randomized Metric Distortion Conjecture. *arXiv preprint arXiv:2111.08698* (2021).

Peng Shi. 2022. Optimal priority-based allocation mechanisms. *Management Science* 68, 1 (2022), 171–188.

A   LOWER BOUND AGAINST THE BOSTON MECHANISM

In this section we briefly discuss the distortion of the Boston mechanism [Abdulkadiroğlu and Sönmez, 2003], a commonly used mechanism in the school choice literature which seems very natural for this problem but unfortunately has exponential distortion.

The mechanism proceeds in rounds as follows. In the first round, each agent proposes to their favorite item. Then each item that has some proposers chooses one to irrevocably match to and rejects all others. In the second round, the agents that didn't get matched then propose to their second favorite items, and the mechanism continues in the same way until all agents are matched. Note that the way that the items choose a proposer can be done in a variety of ways: arbitrarily, randomly, according to some fixed or randomized preference ordering, or in the setting with two-sided preferences, according to the preference list of the item. For ease we will say that choices are made according to a fixed priority order, but this can easily be extended to the other settings.

This mechanism seems like a very natural alternative to Serial Dictatorship – by allowing matchings to happen in parallel, it might cut off the domino effect that causes Serial Dictatorship to have exponential distortion. However, the same effect can still happen, if to a lesser extent. Consider the following instance. The metric space is the line, with the agents are at the points $(1, 2, 4, ..., 2^{k-1})$ and the items lie at the points $(-\varepsilon, 2, 4, ..., 2^{k-1})$. For $t \geq 1$ there are t agents and t items at the point 2^t. There is one agent at 1, and one item at $-\varepsilon$. The co-located agents have identical preferences over the items. Agents have priority in order of their distance to 0.

Then the mechanism will operate as follows. In the first round, for every $t \geq 2$, the agents at 2^t all propose to the same item at 2^t, and that item chooses one of these agents. On the other hand, both the agents at 1 and 2 propose to the item at 2, and the item will choose the agent at 1. In the second round, for every $t \geq 3$ the agents at 2^t, the unmatched agents at 2^t all propose to the same item at 2^t, and that item chooses one of these agents. On the other hand, both the unmatched agents at 2 and 4 propose to the item at 4, and the item will choose the agent at 2. Thus, essentially

the same domino effect takes place as with serial dictatorship, only the number of agents and items at each location must scale linearly to prevent the effect from getting cut off.

It's not hard to see that the overall cost of the mechanism will be $2^k + 1$. As the number of agents is $n = k(k-1)/2$, this shows that the mechanism has distortion $2^{\Omega(\sqrt{n})}$. A similar bound can be proven for the randomized variants of the mechanism just by increasing the number of agents and items at each location. This makes the costly matchings likely enough to cause large distortion in expectation.