Tradeoffs Between Information and Ordinal Approximation for Bipartite Matching

Elliot Anshelevich and Wennan Zhu

Abstract

We study ordinal approximation algorithms for maximum-weight bipartite matchings. Such algorithms only know the ordinal preferences of the agents/nodes in the graph for their preferred matches, but must compete with fully omniscient algorithms which know the true numerical edge weights (utilities). Ordinal approximation is all about being able to produce good results with only limited information. Because of this, one important question is how much better the algorithms can be as the amount of information increases. To address this question for forming high-utility matchings between agents in $\mathcal{X}$ and $\mathcal{Y}$, we consider three ordinal information types: when we know the preference order of only nodes in $\mathcal{X}$ for nodes in $\mathcal{Y}$, when we know the preferences of both $\mathcal{X}$ and $\mathcal{Y}$, and when we know the total order of the edge weights in the entire graph, although not the weights themselves. We also consider settings where only the top preferences of the agents are known to us, instead of their full preference orderings. We design new ordinal approximation algorithms for each of these settings, and quantify how well such algorithms perform as the amount of information given to them increases.

1 Introduction

Many important settings involve agents with preferences for different outcomes. Such settings include, for example, social choice and matching problems. Although the quality of an outcome to an agent may be measured by a numerical utility, it is often not possible to obtain these exact utilities when forming a solution. This can occur because eliciting numerical information from the agents may be too difficult, the agents may not want to reveal this information, or even because the agents themselves do not know the exact numerical values. On the other hand, eliciting *ordinal* information (i.e., the preference ordering of each agent over the outcomes) is often much more reasonable. Because of this, there has been a lot of recent work on *ordinal approximation algorithms*: these are algorithms which only use ordinal preference information as their input, and yet return a solution provably close to the optimum one (e.g., [3–5, 9–12, 17]). In other words, these are algorithms which only use limited ordinal information, and yet can compete in the quality of solution produced with omniscient algorithms which know the true (possibly latent) numerical utility information.

Ordinal approximation is all about being able to produce good results with only limited information. Because of this, it is important to quantify how well algorithms can perform as more information is given. If the quality of solutions returned by ordinal algorithms greatly improves when they are provided more information, then it may be worthwhile to spend a lot of resources in order to acquire such more detailed information. If, on the other hand, the improvement is small, then such an acquisition of more detailed information would not be worth it. Thus the main question we consider in this paper is: *How does the quality of ordinal algorithms improve as the amount of information provided increases?*

This work was partially supported by NSF award CCF-1527497.

This work was published in Proc. of 10th International Symposium on Algorithmic Game Theory (SAGT 2017).

In this paper, we specifically consider this question in the context of computing a maximum-utility matching in a metric space. Matching problems, in which agents have preferences for which other agents they want to be matched with, are ubiquitous. The maximum-weight metric matching problem specifically provides solutions to important applications, such as forming diverse teams and matching in friendship networks (see [4, 5] for much more discussion of this). Formally, there exists a complete undirected bipartite graph for two sets of agents $\mathcal{X}$ and $\mathcal{Y}$ of size N, with an edge weight $w(x, y)$ representing how much utility $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ derive from their match; these edge weights satisfy the triangle inequality. The algorithms we consider, however, do not have access to such numerical edge weights: they are only given ordinal information about the agent preferences. The goal is to form a perfect matching between $\mathcal{X}$ and $\mathcal{Y}$, in order to approximate the maximum weight matching as much as possible using only the given ordinal information. We compare the weight of the matching returned by our algorithms with the true maximum-weight perfect matching in order to quantify the performance of our ordinal algorithms.

Types of Ordinal Information Ordinal approximation algorithms for maximum weight matching have been considered before in [4,5], although only for complete graphs; algorithms for bipartite graphs require somewhat different techniques. Our main contribution, however, lies in considering many types of ordinal information, forming different algorithms for each, and quantifying how much better types of ordinal information improve the quality of the matching formed. Specifically, we consider the following types of ordinal information.

- The most restrictive model we consider is *one-sided preferences*. That is, only preferences for agents in $\mathcal{X}$ over agents in $\mathcal{Y}$ are given to our algorithm. These preferences are assumed to be consistent with the (hidden) agent utilities, i.e., if x prefers y_1 to y_2, then it must be that $w(x, y_1) \geq w(x, y_2)$. Such one-sided preferences may occur, for example, when $\mathcal{X}$ represents people and $\mathcal{Y}$ represents houses. People have preferences over different houses, but houses do not have preferences over people. These types of preferences also apply to settings in which both sides have preferences, but we only have access to the preferences of $\mathcal{X}$, e.g., because the agents in $\mathcal{Y}$ are more secretive.

- The next level of ordinal information we consider is *two-sided preferences*, that is, both preferences for agents in $\mathcal{X}$ over $\mathcal{Y}$ and agents in $\mathcal{Y}$ over $\mathcal{X}$ are given. This setting could apply to the situation that two sets of people are collaborating, and they have preferences over each other, or of a matching between job applicants and possible employers. As we consider the model in a metric space, the distance (weight) between two people could represent the diversity of their skills, and a person prefers someone with most diverse skills from him/her in order to achieve the best results of collaboration.

- The most informative model which we consider in this paper is that of *total-order*. That is, the order of all the edges in the bipartite graph is given to us, instead of only local preferences for each agent. In this model, global ordinal information is available, compared to the preferences of each agent in the previous two models. Studying this setting quantifies how much efficiency is lost due to the fact that we only know ordinal information, as opposed to the fact that we only know *local* information given to us by each agent.

Comparing the results for the above three information types allows us to answer questions like: "Is it worth trying to obtain two-sided preference information or total order information when only given one-sided preferences?" However, above we always assumed that for an agent x, we are given their entire preferences for all the agents in $\mathcal{Y}$. Often, however, an agent would not give their preference ordering for all the agents they could match with, and instead would only give an ordered list of their top preferences. Because of this, in addition to the three models described above, we also consider the case of *partial* ordinal preferences, in which only the top α fraction of a preference list is given by each agent of $\mathcal{X}$. Thus for $\alpha = 0$ no information at all is given to us, and for $\alpha = 1$ the full preference ordering of an agent is given. Considering partial preferences tells us when, if there is a cost to buying information, we might choose to buy only part of the ordinal preferences. We establish tradeoffs between the percentage of available preferences and the possible approximation ratio for all three models of information above, and thus quantify when a specific amount of ordinal information is enough to form a high-quality matching.

Our Contributions We show that as we obtain more ordinal information about the agent preferences, we are able to form better approximations to the maximum-utility matching, even without knowing the true numerical edge weights. Our main results are shown in Figure 1 and Table 1.

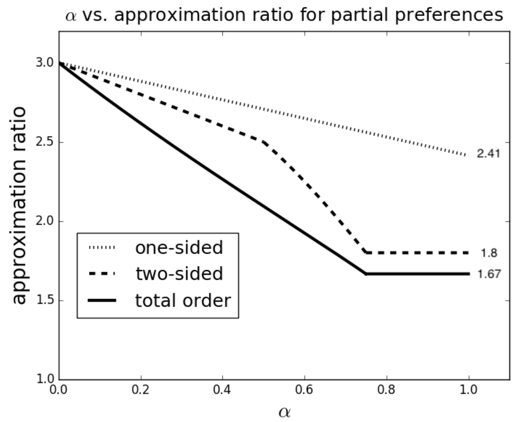

Figure 1: α vs. approximation ratio for partial information. As we obtain more information about the agent preferences (α increases), we are able to form better approximation to the maximum-weight matching. The tradeoff for one-sided preferences is linear, while it is more complex for two-sided and total order.

approximation ratio	$\alpha = 0$	$0 < \alpha < 1$	$\alpha = 1$
one-sided	3	$(3 - (2 - \sqrt{2})\alpha)$	2.41
two-sided	3	$((3 - 2\alpha)(3 - \alpha))/(2\alpha^2 - 3\alpha + 3)$	1.8
total order	3	$(2 + \sqrt{1 - \alpha})/(2 - \sqrt{1 - \alpha})$	1.67

Table 1: α vs. approximation ratio for partial information.

Using only one-sided preference information, with only the order of top αN preferences given for agents in $\mathcal{X}$, we are able to form a $(3 - (2 - \sqrt{2})\alpha)$-approximation. We do this by combining random serial dictatorship with purely random matchings. When $\alpha = 1$, the algorithm yields a $(\sqrt{2} + 1)$-approximation. This is the first non-trivial analysis for the performance of RSD on maximum bipartite matching in a metric space, and this analysis is one of our main contributions.

Given two-sided information, with the order of top αN preferences for agents in both $\mathcal{X}$ and $\mathcal{Y}$, we can do significantly better. When $\alpha \geq \frac{1}{2}$, adopting an existing framework in [4], by mixing greedy and random algorithms, and adjusting it for bipartite graphs, we get a $\frac{(3-2\alpha)(3-\alpha)}{2\alpha^2-3\alpha+3}$-approximation. When $\alpha \leq \frac{1}{2}$, the framework would still work, but would not produce a good approximation. We instead design a different algorithm to get better results. Inspired by RSD, we take advantage of the information of preferences from both sets of agents, adjust RSD to obtain "undominated" edges in each step, and finally combine it with random matchings to get a $(3-\alpha)$-approximation. When $\alpha \geq \frac{3}{4}$, the algorithm yields a 1.8-approximation.

For the total-ordering model, the order of top αN^2 heaviest edges in the bipartite graph is given. We use the framework in [4] again to obtain a $\frac{2+\sqrt{1-\alpha}}{2-\sqrt{1-\alpha}}$-approximation. Here we must re-design the framework to deal with the cases that $\alpha \leq \frac{3}{4}N$, which is not a straight-forward adjustment. When $\alpha \geq \frac{3}{4}N$ the algorithm yields a $\frac{5}{3}$-approximation.

Finally, in Section 6 we analyze the case when edge weights cannot be too different: the highest weight edge is at most β times the lowest weight edge in one-sided model. When the edge weights have this relationship, we can extend our analysis to give a $(\sqrt{\beta - \frac{3}{4}} + \frac{1}{2})$-approximation, even without assuming that edge weights form a metric.

Discussion and Related Work Previous work on forming good matchings can largely be classified into the following classes. First, there is a large body of work assuming that numerical weights or utilities don't exist, only ordinal preferences. Such work studies many possible objectives, such as forming stable matchings (see e.g., [15, 16]), or maximizing objectives determined only by the ordinal preferences (e.g., [2, 8]). Second, there is work assuming that numerical utilities or weights exist, and are *known* to the matching designer. Unlike the above two settings, we consider the case when numerical weights *exist*, but are latent or *unknown*, and yet the goal is to approximate the true social welfare, i.e., maximum weight of a perfect matching. Note that although some previous work assumes that all numerical utilities are known, they often still use algorithms which only require ordinal information, and thus fit into our framework; we discuss some of these results below.

Similar to our one-sided model, house allocation [1] is a popular model of assigning n agents to n items. [6] studied the ordinal welfare factor and the linear welfare factor of RSD and other ordinal algorithms. [14] studied both maximum matching and maximum vertex weight matching using an extended RSD algorithm. These either used objectives depending only on ordinal preferences, such as the size of the matching formed, or used node weights (as opposed to edge weights). [11] and [9] assumed the presence of numerical agent utilities and studied the properties of RSD. Crucially, this work assumed normalized agent utilities, such as unit-sum or unit-range. This allowed [9, 11] to prove approximation ratios of $\Theta(\sqrt{n})$ for RSD. Instead of assuming that agent utilities are normalized, we consider agents in a metric space; this different correlation between agent utilities allows us to prove much stronger results, including a constant approximation ratio for RSD. Kalyanasundaram et al. studied serial dictatorship for maximum weight matching in a metric space [13], and gave a 3-approximation for SD in this, while we are able to get a tighter bound of 2.41-approximation using RSD.[1]

Besides maximizing social welfare, minimizing the social cost of a matching is also popular. [7] studied the approximation ratio of RSD and augmentation of serial dictatorship (SD) for minimum weight matching in a metric space. Their setting is very similar to ours, except that we consider the maximization problem, which has different applications [4, 5], and allows for a much better approximation factor (constant instead of linear in n) using different techniques.

Another area studying ordinal approximation algorithms is social choice, where the goal is to decide a single winner in order to maximize the total social welfare. This is especially related to our work when the hidden utilities of voters are in a metric space (see e.g., [3, 10, 12, 17]),

The work most related to ours is [4, 5]. As mentioned above, we use an existing framework [4] for the two-sided and the total-order model. While the goal is the same: to approximate the maximum weight matching using ordinal information, this paper is differ-

[1] Note that many of the papers mentioned here specifically attempt to form *truthful* algorithms. While RSD is certainly truthful, in this paper we attempt to quantify what can be done using ordinal information in the presence of latent numerical utilities, and leave questions of truthfulness to future work.

ent from [4] in several aspects. [4] only considered approximating the true maximum weight matching for non-bipartite complete graphs. We instead focus on bipartite graphs, and especially on considering different levels of ordinal information by analyzing three models with increasing amount of information, and also consider partial preferences. Although we use similar techniques for parts of two-sided and total-order model analysis, they need significant adjustments to deal with bipartite graphs and partial preferences; moreover, the method used for analyzing the one-sided model is quite different from [4].

2 Model and Notation

For all the problems studied in this paper, we are given as input two sets of agents $\mathcal{X}$ and $\mathcal{Y}$ with $|\mathcal{X}| = |\mathcal{Y}| = N$. $G = (\mathcal{X}, \mathcal{Y}, E)$ is an undirected complete bipartite graph with weights on the edges. We assume that the agent preferences are derived from a set of underlying hidden edge weights $w(x, y) \in \mathbb{R}_{\geq 0}$ for each edge (x, y), $x \in \mathcal{X}, y \in \mathcal{Y}$. $w(x, y)$ represents the utility of the match between x and y, so if x prefers y_1 to y_2, then it must be that $w(x, y_1) \geq w(x, y_2)$. $w(G)$ of any bipartite graph G is the total edge weight of the graph, and $w(M)$ of any perfect matching M is the total weight of edges in the matching. Let $OPT(G)$ denote the complete bipartite matching that gives the maximum total edge weights, i.e., $OPT(G) = \text{argmax}_M w(M)$. The approximation ratio of a perfect matching M is the worst-case ratio between $w(OPT)$ and $w(M)$. If the matchings returned by an approximation algorithm f for any instance of a setting have a approximation ratio of at most ω, we say f gives a ω-approximation to this problem setting. The agents lie in a metric space, by which we will only mean that, $\forall x_1, x_2 \in \mathcal{X}, \forall y_1, y_2 \in \mathcal{Y}, w(x_1, y_1) \leq w(x_1, y_2) + w(x_2, y_1) + w(x_2, y_2)$. We assume this property in all sections except for Section 6.

For the setting of one-sided preferences, $\forall x \in \mathcal{X}$, we are given a strict preference ordering P_x over the agents in $\mathcal{Y}$. When dealing with partial preferences, only top αN agents in P_x are given to us in order. We assume αN is an integer, $\alpha \in [0, 1]$. Of course, when $\alpha = 0$, nothing can be done except to form a completely random matching. For two-sided partial preferences, we are given both the top α fraction of preferences P_x of agents x in $\mathcal{X}$ over those in $\mathcal{Y}$, and vice versa. For the total order setting, we are given the order of the highest-weight αN^2 edges in the complete bipartite graph $G = (\mathcal{X}, \mathcal{Y}, E)$.

3 One-sided Ordinal Preferences

For one-sided preferences, our problem becomes essentially a house allocation problem to maximize social welfare, see e.g., [9, 11, 14]. Before we proceed, it is useful to establish a baseline for what approximation factor is reasonable. Simply picking a matching uniformly at random immediately results in a 3-approximation (see Theorem 3.3), and there are examples showing that this bound is tight. Other well-known algorithms, such as Top Trading Cycle, also cannot produce better than a 3-approximation to the maximum weight matching for our setting. Serial Dictatorship, which uses only one-sided ordinal information, is also known to give a 3-approximation to the maximum weight matching for our problem [13]. Serial Dictatorship simply takes an arbitrary agent from $x \in \mathcal{X}$, assigns it x's favorite unallocated agent from $\mathcal{Y}$, and repeats. Note that [13] used a greedy algorithm for the online maximum weight matching problem, and the algorithm is actually SD because the arbitrary arriving order in online problems describes how we pick agents in an arbitrary order. Unfortunately, it is not difficult to show that this bound of 3 is tight. Our first major result in this paper is to prove that *Random* Serial Dictatorship always gives a $(\sqrt{2} + 1)$-approximation in expectation, no matter what the true numerical weights are, thus giving a significant improvement to all the algorithms mentioned above.

Algorithm 1: Random Serial Dictatorship for Perfect Matching of one-sided ordering.

Initialize $M = \emptyset$, $G = (\mathcal{X}, \mathcal{Y}, E)$;
while $E \neq \emptyset$ **do**
  Pick an agent x uniformly at random from $\mathcal{X}$;
  Let y denote x's most preferred agent in $\mathcal{Y}$;
  Take $e = (x, y)$ from E and add it to M ;
  Remove x, y, and all edges containing x or y from the graph G ;
end
Final Output: Return M.

Theorem 3.1 *Suppose $G = (\mathcal{X}, \mathcal{Y}, E)$ is a complete bipartite graph on the set of nodes $\mathcal{X}, \mathcal{Y}$ with $|\mathcal{X}| = |\mathcal{Y}| = N$. Then, the expected weight of the perfect matching M returned by Algorithm 1 is $E[w(M)] \geq \frac{1}{\sqrt{2}+1} w(OPT(G))$.*

Proof Sketch. We give a proof sketch here; full proofs for all our results can be found in the full version of this paper at `http://www.cs.rpi.edu/~eanshel/`. Let $Min(G)$ denote a *minimum* weight perfect matching on G, and $RSD(G)$ denote the expected weight returned by Algorithm 1 on graph G. For any $x \in \mathcal{X}$, we use $\lambda(x)$ to denote the edge between x and its most preferred agent in $\mathcal{Y}$. Define $R(x)$ as the remaining graph after removing x, x's most preferred agent, and all the edges containing x or x's most preferred agent from G. We now state the main technical lemma which allows us to prove the result. This lemma gives a bound on the maximum weight matching in terms of the quantities defined above.

Lemma 3.2 *For any given graph $G = (\mathcal{X}, \mathcal{Y}, E)$, one of the following two cases must be true:*
Case 1: $w(OPT(G)) \leq \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} w(OPT(R(x))) + \frac{\sqrt{2}+1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} w(\lambda(x))$
Case 2: $w(OPT(G)) \leq (\sqrt{2} + 1) w(Min(G))$

We will prove this lemma below, but first we discuss how the rest of the proof proceeds. When Case 1 above holds, we know that at any step of the algorithm, the change in the weight of the optimum solution in the remaining graph is not that different from the weight of the edge selected by our algorithm. This allows us to compare the weight of OPT with the weight of the matching returned by our algorithm. In fact, this is the technique used in a previous paper [5] to analyze RSD for complete graphs (i.e., non-bipartite graphs), and show that RSD gives a 2-approximation for perfect matching on complete graphs. It is important to note here that this *does not* work for bipartite graphs. In bipartite matching, there are examples in which using only this method will not give an approximation ratio better than 3. We get around this problem by adding Case 2 to our lemma, and then using this to prove the theorem.

Proof Sketch of Lemma 3.2. For any fixed $x \in \mathcal{X}$, denote x's most preferred agent in $\mathcal{Y}$ as y (so $\lambda(x) = (x, y)$). In $OPT(G)$, suppose x is matched to $b \in \mathcal{Y}$, and y is matched to $a \in \mathcal{X}$. In $Min(G)$, suppose b is matched to $m \in \mathcal{X}$. $\forall x \in \mathcal{X}$, there exist y, a, b, m as described above. Denote edge (x, y) by $\lambda(x)$, (x, b) by $P(x)$, (a, y) by $\bar{P}(x)$, and (a, b) by $D(x)$.

We'll prove Lemma 3.2 by showing that if **Case 2** is not true, then **Case 1** must be true. Suppose **Case 2** is not true, i.e., $w(OPT(G)) > (\sqrt{2} + 1) w(Min(G))$. Suppose that random serial dictatorship picks $x \in \mathcal{X}$. Then $OPT(R(x))$ is at least as good as the matching obtained by removing $P(x)$ and $\bar{P}(x)$, and adding $D(x)$ to $OPT(G)$ (the rest stay the same):

$$w(OPT(R(x))) \geq w(OPT(G)) - w(P(x)) - w(\bar{P}(x)) + w(D(x))$$

Summing this up over all nodes x, we obtain:

$$\frac{1}{|\mathcal{X}'|}\sum_{x\in\mathcal{X}'} w(OPT(R(x))) \geq (1-\frac{1}{|\mathcal{X}'|})w(OPT(G)) - \frac{1}{|\mathcal{X}'|}\sum_{x\in\mathcal{X}'}(w(\bar{P}(x)) - w(D(x))) \quad (1)$$

By the triangle inequality, we know that $w(a,y) \leq w(a,b) + w(m,b) + w(m,y)$ Because $\lambda(m)$ is the edge to m's most preferred agent, $w(m,y) \leq w(\lambda(m))$, and thus $w(\bar{P}(x)) \leq w(D(x)) + w(m,b) + w(\lambda(m)))$.

Summing this up for all $x \in \mathcal{X}$, note that each x is matched to a unique b in $OPT(G)$, and each b is matched to a unique m in $Min(G)$, so each agent in $\mathcal{Y}$ appears as b exactly once and each agent in $\mathcal{X}$ appears as m exactly once.

$$\sum_{x\in\mathcal{X}}(w(\bar{P}(x)) - w(D(x))) \leq w(Min(G)) + \sum_{x\in\mathcal{X}} w(\lambda(x))) \quad (2)$$

Combining Inequality 1 and Inequality 2,

$$\frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}} w(OPT(R(x))) \geq (1-\frac{1}{|\mathcal{X}|})w(OPT(G)) - \frac{1}{|\mathcal{X}|}[w(Min(G)) + \sum_{x\in\mathcal{X}} w(\lambda(x))] \quad (3)$$

$w(P(x)) \leq w(\lambda(x))$ since $\lambda(x)$ is the most preferred edge of x, so it is obvious that $w(OPT(G)) \leq \sum_{x\in\mathcal{X}} w(\lambda(x))$. Combining this with our assumption about $Min(G)$, we obtain the desired result. For detailed proof, see the full version.

Partial One-sided Ordinal Preferences

In this section, we consider the case when we are given even less information than in the previous one, i.e., only partial preferences. We begin by establishing the following easy result for the completely random algorithm.

Theorem 3.3 *The uniformly random perfect matching is a 3-approximation to the maximum-weight matching.*

Algorithm 2: Algorithm for Perfect Matching given partial one-sided ordering.

Run Algorithm 1, stop when $|M| = \alpha N$, then form random matches until all agents are matched. Return M.

Theorem 3.4 *Suppose $G = (\mathcal{X}, \mathcal{Y}, E)$ is a complete bipartite graph on the set of nodes $\mathcal{X}, \mathcal{Y}$ with $|\mathcal{X}| = |\mathcal{Y}| = N$. There is a strict preference ordering P_x over the agents in $\mathcal{Y}$ for each agent $x \in \mathcal{X}$. We are only given top αN agents in P_x in order. Then, the expected weight of the perfect matching M returned by Algorithm 2 is $E[w(M)] \geq \frac{1}{3-(2-\sqrt{2})\alpha} w(OPT(G))$, as shown in Figure 1.*

Proof Sketch. We establish a linear tradeoff as α increases. Note that this would not work for combining any two arbitrary algorithms. The key insight which makes this proof work is that, at every step, the expected weight of RSD is higher than in the following step, and that RSD always produces an edge weight which is better than random in expectation.

4 Two-sided Ordinal Preferences

For two-sided preferences, we give separate algorithms for the cases when $\alpha \geq \frac{1}{2}$ and when $\alpha \leq \frac{1}{2}$, as these require somewhat different techniques.

$\boldsymbol{\alpha \geq \frac{1}{2}}$ While for the case when $\alpha < \frac{1}{2}$ new techniques are necessary to obtain a good approximation, the approach for the case when $\alpha \geq \frac{1}{2}$ is essentially the same as the one used in [4]. We adopt this approach to deal with bipartite graphs and with partial preferences, giving us a 1.8-approximation for $\alpha = 1$. To do this, we re-state the definition of Undominated Edges from [4], and a standard greedy algorithm for forming a matching of size k.

Definition 4.1 *(Undominated Edges) Given a set E of edges, $(x, y) \in E$ is said to be an undominated edge if for all (x, a) and (y, b) in E, $w(x, y) \geq w(x, a)$ and $w(x, y) \geq w(y, b)$.*

Note that an undominated edge must always exist: either there are two nodes x and y such that they are each other's top preferences (and so (x, y) is undominated), or there is a cycle $x_1, x_2, \ldots$ in which x_{i+1} is the top preference of x_i, in which case all edges in the cycle must be the same weight, and thus all edges in the cycle are undominated. This also gives us an algorithm for determining if an edge (x, y) is undominated: either x and y prefer each other over all other agents, or it is part of such a cycle of top preferences.

Algorithm 3: Greedy Algorithm for Max k-Matching of two-sided ordering.

Given bipartite graph $G = (\mathcal{X}, \mathcal{Y}, E)$, and k, initialize a matching $M = \emptyset$. Pick an arbitrary undominated edge $e = (x, y)$ from E and add it to M. Remove x, y, and all edges containing x or y from E. Repeat until $|M| = k$. Return M.

Algorithm 4: Algorithm for two-sided matching with partial ordinal information ($\frac{1}{2} \leq \alpha \leq \frac{3}{4}$).

Input : $\mathcal{X}, \mathcal{Y}$, top αN of $P(\mathcal{X})$, top αN of $P(\mathcal{Y})$
Output: Perfect Bipartite Matching M
Initialize E to be complete bipartite graph on $\mathcal{X}, \mathcal{Y}$, and $M_1 = M_2 = \emptyset$;
Let M_0 be the output returned by Algorithm 3 for E, $k = \alpha N$;
Let $\mathcal{X}_T$ be the set of nodes in $\mathcal{X}$ matched in M_0, $\mathcal{Y}_T$ be the set of nodes in $\mathcal{Y}$ matched in M_0, and T be the complete bipartite graph on $\mathcal{X}_T, \mathcal{Y}_T$;
Let $\mathcal{X}_B = \mathcal{X} \backslash \mathcal{X}_T$, $\mathcal{Y}_B = \mathcal{Y} \backslash \mathcal{Y}_T$, and B be the complete bipartite graph on $\mathcal{X}_B, \mathcal{Y}_B$;
First Algorithm;
$M_1 = M_0 \cup$ (Uniformly random perfect matching on B);
Second Algorithm;
Choose $(2\alpha - 1)N$ edges from M_0 uniformly at random and add them to M_2 ;
Let X_A be the set of nodes in $\mathcal{X}_T$ and not in M_2, Y_A be the set of nodes in $\mathcal{Y}_T$ and not in M_2;
Let E_{AB} be the edges of the complete bipartite graph $(X_A, \mathcal{Y}_B)$ and E'_{AB} be the edges of the complete bipartite graph $(\mathcal{X}_B, Y_A)$;
Run random bipartite matching on the set of edges in E_{AB} and E'_{AB} separately to obtain perfect bipartite matchings and add the edges returned by the algorithm to M_2;
Final Output: Return M_1 with probability $\frac{3-2\alpha}{3-\alpha}$ and M_2 with probability $\frac{\alpha}{3-\alpha}$.

Note that for $\alpha > \frac{3}{4}$ this algorithm does not seem to provide better guarantees than for $\alpha = \frac{3}{4}$. Because of this, for $\alpha > \frac{3}{4}$, we simply run the same algorithm for $\alpha = \frac{3}{4}$.

Theorem 4.1 *Algorithm 4 returns a* $\frac{(3-2\alpha)(3-\alpha)}{2\alpha^2-3\alpha+3}$*-approximation to the maximum-weight perfect matching given two-sided ordering when* $\frac{1}{2} \leq \alpha \leq \frac{3}{4}$.

$\boldsymbol{\alpha \leq \frac{1}{2}}$ Unlike the case for $\alpha \geq \frac{1}{2}$, this case requires different techniques than in [4]. While the techniques above would still work, they will not give us a bound as good as the one we form below. The idea in this section is to do something similar to our one-sided algorithm for partial preferences: run the greedy algorithm for a while, and then switch to random. Unfortunately, if we simply run the greedy Algorithm 3 without any changes and then switch to random, this will not form a good approximation. The reason why this is true is that an undominated edge which is picked by the greedy algorithm may be much worse than the average weight of an edge, and so the approximation factor of the random algorithm will dominate, giving only a 3-approximation. Even taking an undominated edge uniformly at random has this problem. We can fix this, however, by picking each undominated edge with an appropriate probability, as described below. Such an algorithm results in matchings which are guaranteed to be better than either RSD or Random, thus allowing us to prove the result.

Algorithm 5: Algorithm for two-sided matching with partial ordinal information ($0 \leq \alpha \leq \frac{1}{2}$).

```
Input  : 𝒳, 𝒴, top αN of P(𝒳) and P(𝒴)
Initialize M = ∅, G = (𝒳, 𝒴, E) ;
while E ≠ ∅ do
    Pick an agent x uniformly at random from 𝒳 ;
    Let y denote x's most preferred agent in 𝒴 ;
    x1 ← x, y1 ← y, c ← y1;
    while (x1, y1) is not an undominated edge do
        if c = y1 then
            x1 ← y1's most preferred agent in 𝒳 ;
            c ← x1;
        else
            y1 ← x1's most preferred agent in 𝒴 ;
            c ← y1;
        end
    end
    Take (x1, y1) from E and add it to M ;
    Remove x1, y1, and all edges containing x1 or y1 from the graph G ;
    if |M| = αN then
        break;
    end
end
Run the uniform random matching algorithm for the remaining graph G, add the
edges returned by the algorithm to M. Final Output: Return M.
```

This algorithm guarantees that an undominated edge is chosen for any x in any bipartite graph G. Now, before we reach an undominated edge, the weights of edges are non-decreasing in the order they are checked. Thus whenever a node x is picked, the algorithm adds an undominated edge (x_1, y_1) to the matching which is guaranteed to have higher weight than all edges leaving x.

Theorem 4.2 *Algorithm 5 returns a* $(3-\alpha)$*-approximation to the maximum-weight perfect matching given two-sided ordering when* $0 \leq \alpha \leq \frac{1}{2}$.

Proof Sketch. We use a similar method and the same notation as in Section 3 to prove this theorem. Essentially, because we are always picking undominated edges, we can form a linear interpolation between a factor of 2 and a factor of 3 for random matching, instead of between factors $\sqrt{2}+1$ and 3 as for one-sided preferences. The reason why we are able to form such an interpolation is entirely because of the probabilities with which we choose the undominated edges; if we simply chose arbitrary undominated edges or choose them uniformly at random, then there are examples where the random edge weights will dominate and result in a poor approximation, since undominated edges are only guaranteed to be within a factor of 3 of the average edge weight.

5 Total Ordering of Edge Weights

For the setting in which we are given the top αN^2 edges of G, we prove that for $\alpha = \frac{3}{4}$, we can obtain an approximation of $\frac{5}{3}$ in expectation. For larger α, however, more information does not seem to help, and so we simply use the algorithm for $\alpha = \frac{3}{4}$ for any $\alpha > \frac{3}{4}$.

Algorithm 6: Algorithm for Max k-Matching with total ordering of edge weights.

Given bipartite graph $G = (\mathcal{X}, \mathcal{Y}, E)$, and k, initialize a matching $M = \emptyset$. Pick the heaviest edge $e = (x, y)$ from E and add it to M. Remove x, y, and all edges containing x or y from E. Repeat until $|M| = k$. Return M.

The algorithm for bipartite matching with partial ordinal information is similar to that with partial two-sided ordinal information, except that we only need to consider the case that $k \leq \frac{1}{2}N$, $\alpha \leq \frac{3}{4}$.

Algorithm 7: Algorithm for matching given partial total ordering.

Input : $\mathcal{X}, \mathcal{Y}$, order of the top αN^2 edges in the graph.
Output: Perfect Bipartite Matching M
Initialize E to be complete bipartite graph on $\mathcal{X}, \mathcal{Y}$, and $M_1 = M_2 = \emptyset$;
Let M_0 be the output returned by Algorithm 6 for E, $k = (1 - \sqrt{1-\alpha})N$. Let $\alpha_1 = 1 - \sqrt{1-\alpha}$, then $k = \alpha_1 N$;
Let $\mathcal{X}_T$ be the set of nodes in $\mathcal{X}$ matched in M_0, $\mathcal{Y}_T$ be the set of nodes in $\mathcal{Y}$ matched in M_0, and T be the complete bipartite graph on $\mathcal{X}_T, \mathcal{Y}_T$;
Let $\mathcal{X}_B$ be the set of nodes in $\mathcal{X}$ not matched in M_0, $\mathcal{Y}_B$ be the set of nodes in $\mathcal{Y}$ not matched in M_0, and B is the complete bipartite graph on $\mathcal{X}_B, \mathcal{Y}_B$;
First Algorithm;
$M_1 = M_0 \cup$ (Uniformly random perfect matching on B);
Second Algorithm;
Choose $(1 - 2\alpha_1)N$ nodes both from $\mathcal{X}_B$ and $\mathcal{Y}_B$ uniformly at random, get the perfect matching output by the uniform random algorithm on these nodes and add the results to M_2 ;
Let X_A be the set of nodes in $\mathcal{X}_B$ and not in M_2, Y_A be the set of nodes in $\mathcal{Y}_B$ and not in M_2;
Let E_{AT} be the edges of the complete bipartite graph $(X_A, \mathcal{Y}_T)$ and E'_{AT} be the edges of the complete bipartite graph $(\mathcal{X}_T, Y_A)$;
Run random bipartite matching on the set of edges in E_{AT} and E'_{AT} separately to obtain perfect bipartite matchings and add the edges returned by the algorithm to M_2;
Final Output: Return M_1 with probability $\frac{2}{2+\sqrt{1-\alpha}}$ and M_2 with probability $\frac{\sqrt{1-\alpha}}{2+\sqrt{1-\alpha}}$.

Theorem 5.1 *Algorithm 7 returns a* $\frac{2+\sqrt{1-\alpha}}{2-\sqrt{1-\alpha}}$*-approximation to the maximum-weight matching in expectation for* $\alpha \leq \frac{3}{4}$*, as shown in Figure 1.*

6 One-sided Preferences with Restricted Edge Weights

In previous sections, we made the assumption that the agents lie in a metric space, and thus the edge weights, although unknown to us, must follow the triangle inequality. In this section we once again consider the most restrictive type of agent preferences — that of one-sided preferences — but now instead of assuming that agents lie in a metric space, we instead consider settings where edges weights cannot be infinitely different from each other. This applies to settings where the agents are at least somewhat indifferent and the items are somewhat similar; the least-preferred agent and the most-preferred items differ only by a constant factor to any agent. Indeed, when for example purchasing a house in a reasonable market (i.e., once houses that almost no one would buy have been removed from consideration), it is unlikely that any agent would like house x so much more than house y that they would be willing to pay hundreds of times more for x than for y.

More formally, for each agent $i \in \mathcal{X}$, we are given a strict preference ordering P_i over the agents in $\mathcal{Y}$. In this section we assume that the highest weight edge e_{max} is at most β times of the lowest weight edge e_{min}. We normalize the lowest weight edge e_{min} in the graph to $w(e_{min}) = 1$; then for any edge $e \in E$, $w(e) \leq \beta$. We use similar analysis as in Section 3, except that instead of getting bounds by using the triangle inequality, the relationships among edge weights are bounded by our assumption of the highest and lowest weight edge ratio. As stated above, we no longer assume the agents lie in a metric space in this section.

Theorem 6.1 *Suppose* $G = (\mathcal{X}, \mathcal{Y}, E)$ *is a complete bipartite graph on the set of nodes* $\mathcal{X}, \mathcal{Y}$ *with* $|\mathcal{X}| = |\mathcal{Y}| = N$*.* $w(e_{min}) = 1$*,* $\forall e \in E$*,* $w(e) \leq \beta$*. The expected weight of the perfect matching returned by Algorithm 1 is* $w(M) \geq \frac{1}{\sqrt{\beta - \frac{3}{4}} + \frac{1}{2}} w(OPT)$ *(see plot and proof in the full version).*

References

[1] Atila Abdulkadiroğlu and Tayfun Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.

[2] David J. Abraham, Robert W. Irving, Telikepalli Kavitha, and Kurt Mehlhorn. Popular matchings. *SIAM J. Comput.*, 37(4):1030–1045, 2007.

[3] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. Approximating optimal social choice under metric preferences. In *AAAI 2015.*

[4] Elliot Anshelevich and Shreyas Sekar. Blind, greedy, and random: Algorithms for matching and clustering using only ordinal information. In *AAAI 2016.*

[5] Elliot Anshelevich and Shreyas Sekar. Truthful mechanisms for matching and clustering in an ordinal world. In *WINE 2016.*

[6] Anand Bhalgat, Deeparnab Chakrabarty, and Sanjeev Khanna. Social welfare in one-sided matching markets without money. In *APPROX 2011.*

[7] 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. In *WINE 2016*.

[8] Deeparnab Chakrabarty and Chaitanya Swamy. Welfare maximization and truthfulness in mechanism design with ordinal preferences. In *ITCS 2014*.

[9] George Christodoulou, Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, Paul W Goldberg, Jie Zhang, and Jinshan Zhang. Social welfare in one-sided matching mechanisms. In *AAMAS 2016*.

[10] Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *EC 2016*.

[11] Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, and Jie Zhang. Social welfare in one-sided matchings: Random priority and beyond. In *SAGT 2014*.

[12] Ashish Goel, Anilesh Kollagunta Krishnaswamy, and Kamesh Munagala. Metric distortion of social choice rules: Lower bounds and fairness properties. In *EC 2017*.

[13] Bala Kalyanasundaram and Kirk Pruhs. On-line weighted matching. In *SODA*, volume 91, pages 234–240, 1991.

[14] Piotr Krysta, David Manlove, Baharak Rastegari, and Jinshan Zhang. Size versus truthfulness in the house allocation problem. In *EC 2014*.

[15] Baharak Rastegari, Anne Condon, Nicole Immorlica, and Kevin Leyton-Brown. Two-sided matching with partial information. In *EC 2013*.

[16] Alvin E Roth and Marilda Sotomayor. Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541, 1992.

[17] Piotr Skowron and Edith Elkind. Social choice under metric preferences: Scoring rules and stv. In *AAAI 2017*.