Improving Welfare in One-Sided Matchings using Simple Threshold Queries

Thomas Ma[1] , **Vijay Menon**[2] , **Kate Larson**[2*]

[1]Department of Computer Science, University of Toronto
[2]David R. Cheriton School of Computer Science, University of Waterloo

thomas.ma@mail.utoronto.ca, {vijay.menon, kate.larson}@uwaterloo.ca

Abstract

We study one-sided matching problems where each agent must be assigned at most one object. In this classic problem it is often assumed that agents specify only ordinal preferences over objects and the goal is to return a matching that satisfies some desirable property such as Pareto optimality or rank-maximality. However, agents may have cardinal utilities describing their preference intensities and ignoring this can result in welfare loss. We investigate how to elicit additional cardinal information from agents using simple threshold queries and use it in turn to design algorithms that return a matching satisfying some desirable matching property, while also achieving a good approximation to the optimal welfare among all matchings satisfying that property. Overall, our results show how we can improve welfare by even non-adaptively asking agents for just one bit of extra information per object.

1 Introduction

One-sided matching scenarios are ubiquitous in multiagent resource-allocation settings and have been well-studied, especially as the housing allocation or housing market problem both in economics [Shapley and Scarf, 1974; Hylland and Zeckhauser, 1979; Roth and Postlewaite, 1977; Abdulkadiroğlu and Sönmez, 1998; Abdulkadiroğlu and Sönmez, 1999; Sönmez and Ünver, 2010] and in computer science [Abraham *et al.*, 2004; Irving, 2004; Abraham *et al.*, 2006; Filos-Ratsikas *et al.*, 2014; Amanatidis *et al.*, 2021]. Other examples include assigning faculty members to school committees, workers to tasks, etc.

Much of the literature assumes that agents have an acceptable set of objects and that they submit an (ordinal) preference order over this set. Given this, the standard objective is to come up with an assignment of objects to agents (i.e., a matching) that satisfies some desirable property like Pareto optimality [Shapley and Scarf, 1974; Abraham *et al.*, 2004] or rank maximality [Irving, 2004; Irving *et al.*, 2006].

Although matchings that satisfy such desirable properties are better than arbitrary ones, one key drawback is that they may not take into account agents' preference intensities. To illustrate this, consider the following simple example where there are three agents $\{a_1, a_2, a_3\}$ and three objects $\{o_1, o_2, o_3\}$. Assume all agents agree that o_1 is preferred to o_2 which is preferred to o_3, but have different preference intensities (i.e., cardinal utilities) for the objects. In particular, a_1 and a_2 assign utility 0.9 to o_1 and 0.1 to o_2 while a_3 assigns utility 0.51 to a_1 and 0.49 to o_2. All agents assign zero utility for o_3. If only ordinal preferences are considered then any matching is, for example, Pareto optimal. However, any matching that assigns o_3 to a_3 leads to significant loss in overall social welfare compared to other Pareto optimal matchings and so is, in some sense, less desirable.

The observation that there might be a loss in welfare due to ignoring preference intensities (henceforth, cardinal utilities) is not new, and in particular, has been a much debated issue surrounding various *school-choice* mechanisms ([Abdulkadiroğlu *et al.*, 2011; Abdulkadiroğlu *et al.*, 2015]). This has also lead to proposals for new school choice mechanisms that ask agents to provide some extra information along with their ordinal preferences [Abdulkadiroğlu *et al.*, 2015]. Our work here is partially motivated by this line of work, but takes a more computational approach that is similar in style to the work that looks at *distortion*—which is essentially the cost of using only ordinal information—in various settings [Procaccia and Rosenschein, 2006; Boutilier *et al.*, 2015; Anshelevich and Sekar, 2016; Anshelevich and Zhu, 2017; Goel *et al.*, 2017; Abramowitz and Anshelevich, 2018].

Given a one-sided matching instance (the set of agents, objects, and agents' ordinal preferences), our goal is to find matchings that satisfy some particular property, say, X, while also accounting for agents' cardinal utilities. We accomplish this by designing algorithms that are guaranteed to return matchings that satisfy property X while also achieving a good approximation to the optimal welfare amongst all matchings that satisfy X. One way to achieve this is to ask agents to directly provide their cardinal utilities for objects. Such an approach, however, places a high burden on the agents themselves as they are required to articulate and communicate precise cardinal information. Instead we propose a middle-ground and use simple binary queries to elicit relevant information from agents. Our goal is to ask each agent a small

---

*Contact Author

number of such queries and return a matching that achieves a good approximation as described above.

In particular, we consider the following four well-studied matching properties: Pareto optimal matchings, rank-maximal matchings, max-cardinality rank-maximal matchings, and fair matchings, and two rich cardinal utility models (unit-sum and unit-range valuations). We first explore adaptive algorithms—algorithms that are able to change their queries depending on how agents answer previous queries—and show how for each of the properties mentioned above and for any $\epsilon > 0$, there is a deterministic algorithm that asks $O(c \log n)$ queries per agent, where $c = \left\lceil \frac{\log(n^2 \cdot 1/\epsilon)}{\log(1+\epsilon/2)} \right\rceil$, and returns a matching that achieves a $(1 + \epsilon)$-approximation to the optimal welfare among all matchings that satisfy the property of interest. We then focus on non-adaptive algorithms which, we argue, have many practical advantages over adaptive algorithms, and explore what is possible to achieve in the special case where the algorithm is allowed at most one query per (agent, object) pair. Table 1 summarizes our results.

Related Research. Motivation for our work is derived from the school choice problem which addresses the loss in welfare due to not taking preference intensities into account [Abdulkadiroğlu *et al.*, 2011; Abdulkadiroğlu *et al.*, 2015]. Our concern here is similar, but we take a computational approach, reminiscent to the work on distortion [Procaccia and Rosenschein, 2006; Anshelevich and Zhu, 2017; Abramowitz and Anshelevich, 2018]. However, unlike this body of work which aims to calculate the worst-case loss in welfare due to only having ordinal preferences, we assume that, in addition to ordinal preferences, it is also possible to obtain information about agents' cardinal utilities. This in turn is similar to an approach that has been explored in the voting context [Abramowitz *et al.*, 2019; Amanatidis *et al.*, 2020], and one-sided matching [Amanatidis *et al.*, 2021]. While this latter work also looks at one-sided matching, the objective is different with its focus on distortion, as opposed to finding "good" matchings satisfying certain properties. We also observe that the query models used differ (with ours being significantly weaker).

Our work is also related to the study of communication complexity of voting protocols [Mandal *et al.*, 2020], to the work on participatory budgeting which compares different elicitation methods based on the distortion achieved ([Goel *et al.*, 2019; Benade *et al.*, 2020]), and is more broadly in line with the growing body of work that explicitly aims to make mechanisms or algorithms more robust, by either making use of coarse preference information [Chiesa *et al.*, 2012; Chiesa *et al.*, 2014; Menon and Larson, 2019], or by making sure that the algorithms designed produce solutions that work "well" (in the approximation sense) even under slightly modified inputs [Shiryaev *et al.*, 2013; Bredereck *et al.*, 2017; Menon and Larson, 2018; Chen *et al.*, 2019].

2 Model

For $k \in \mathbb{Z}^+$, let $[k]$ denote the set $\{1, \ldots, k\}$. We use $\mathcal{N}$, where $|\mathcal{N}| = n$, to denote the set of agents $\{a_1, \ldots, a_n\}$, and $\mathcal{H}$, where $|\mathcal{H}| = n$, to denote the set of objects $\{h_1, \ldots, h_n\}$.

We refer to a_i as agent i and h_j as object j. Every agent a_i has a weak order, P_i, over a subset of objects $A_i \subseteq \mathcal{H}$, where A_i, $|A_i| \geq 1$, is the set of objects a_i is willing to be matched to, the *acceptable set* of a_i. We use $\mathcal{P} = (P_1, \ldots, P_n)$ to refer to the weak orders of all the agents in $\mathcal{N}$ and refer to $\mathcal{P}$ as the *preference profile* of the agents. For an agent a_i, and for two objects $h_j, h_k \in A_i$, we use $h_j \succ_i h_k$ to denote that a_i strictly prefers h_j over h_k, and use $h_j \succeq_i h_k$ to indicate that h_j is either strictly preferred or considered to be equivalent to h_k. We refer to $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P} = (P_1, \ldots, P_n))$ as an *instance*, which encodes all the information about the agents, objects, and the agents' preferences, and use $\mathbb{I}$ to denote the set of all possible instances.

Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, $\mathbb{G}_{\mathcal{I}} = (\mathcal{N} \cup \mathcal{H}, \mathcal{E})$ is the induced bipartite graph with edges $(a_i, h_j) \in \mathcal{E}$ if $h_j \in A_i$. We refer to $e = (a_i, h_j) \in \mathcal{E}$ as a rank-k edge if $|\mathcal{U}_{ij}| = k - 1$, where $\mathcal{U}_{ij} = \{h_\ell \in A_i \mid h_\ell \succ_i h_j\}$. We also use $rank(a_i, h_j)$ to denote the k such that (a_i, h_j) is a rank-k edge and refer to an object h_j as a_i's rank-k object if $rank(a_i, h_j) = k$.

We additionally assume that each agent a_i has a cardinal utility function $v_i \colon \mathcal{H} \to [0, 1]$, which is consistent with the preference order P_i (meaning, $h_1 \succeq_i h_2 \Leftrightarrow v_i(h_1) \geq v_i(h_2)$); we assume that if $h \notin A_i$, then $v_i(h) = 0$. We use $v = (v_1, \ldots, v_n)$ to denote the valuation profile of agents and $\mathcal{V}_{\mathcal{I}}$ to denote the set of all possible valuation profiles that are consistent with the given preference profile in $\mathcal{I}$. In this work we consider two specific classes of valuation functions;

Unit-sum valuations: For each agent i, v_i is such that $\sum_{h \in \mathcal{H}} v_i(h) = 1$.

Unit-range valuations: Agents are said to have unit-range valuations if for each agent i, there exists $h_j, h_k \in A_i$ such that $h_j \succ_i h_k$, and $\max_{h \in A_i} v_i(h) = 1$ and $\min_{h \in A_i} v_i(h) = 0$. In words, the most preferred objects have value 1, the least preferred objects have value 0, and every other acceptable object has value between 0 and 1.

Note that information about the cardinal utilities is not part of an instance $\mathcal{I}$. Given $\mathcal{I}$, we are interested in matchings of agents to objects, namely bijections $\mu \colon \mathcal{N} \to \mathcal{H}$. For $c \in \mathcal{N} \cup \mathcal{H}$, we refer to $\mu(c)$ as c's *partner* in μ or as c's *allocation* in μ. Alternatively, a matching is also defined as a collection of edges μ in $\mathbb{G}_{\mathcal{I}}$ such that each vertex is part of at most one edge in μ. We use $\mathcal{M}_{\mathbb{G}_{\mathcal{I}}}$ to denote the set of all possible matchings in $\mathbb{G}_{\mathcal{I}}$.

2.1 Pareto Optimal and Signature-Based Matchings

Although for a given instance there are several possible matchings, we are interested in matchings which also satisfy some additional desirable property. In particular, we consider the following well-studied properties: Pareto optimal matchings [Shapley and Scarf, 1974; Abraham *et al.*, 2004], rank-maximal matchings [Irving, 2004; Irving *et al.*, 2006], max-cardinality rank-maximal matchings [Mehlhorn and Michail, 2005; Abraham *et al.*, 2006], and fair matchings [Mehlhorn and Michail, 2005; Huang *et al.*, 2013]. The latter three

	Ordinal algorithms	Adaptive threshold query algorithms (for any $\epsilon > 0$, $O(c \log n)$ queries per agent, where $c = \left\lceil \frac{\log(n^2 \cdot 1/\epsilon)}{\log(1+\epsilon/2)} \right\rceil$)	Non-adaptive threshold query algorithms (at most 1 query per (agent, object) pair)
unit-sum valuations	UB: $O(n^2)$ [Theorem 1] LB: $\Omega(n^2)$ [Theorem 1]	$1 + \epsilon$ [Theorem 3]	UB: $O(n^{2/3})$ [Theorems 5 and 4] LB: $\Omega(\sqrt{n})$ [Theorem 8]
unit-range valuations	UB: $O(n)$ [Theorem 1] LB: $\Omega(n)$ [Theorem 1]	$1 + \epsilon$ [Theorem 3]	UB: $O(\sqrt{n})$ [Theorems 6 and 7] LB: $\Omega(\sqrt{n})$ [Theorem 8]

Table 1: Summary of our results. For X, where X is one of the properties in the set {Pareto optimal, rank-maximal, max-cardinality rank-maximal, fair}, an upper bound (UB) of α indicates that there is a deterministic algorithm that always produces a matching that satisfies X and achieves an α-approximation to the optimal welfare among matchings that satisfy X. A lower bound (LB) of β indicates that there is no deterministic algorithm that produces a matching that satisfies X and achieves a β-approximation to the optimal welfare among matchings that satisfy X.

are different ways to strengthen Pareto optimality and are together referred to as *signature-based* matchings.

Definition 1. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, a matching $\mu \in \mathcal{M}_{\mathbb{G}_\mathcal{I}}$ is **Pareto optimal (PO)** w.r.t. $\mathcal{I}$ if $\forall \mu' \in \mathcal{M}_{\mathbb{G}_\mathcal{I}}$*

$$(\exists a_i \in \mathcal{N}, \mu'(a_i) \succ_i \mu(a_i)) \Rightarrow (\exists a_j \in \mathcal{N}, \mu'(a_j) \prec_j \mu(a_j))$$

Definition 2. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, and a matching $\mu \in \mathcal{M}_{\mathbb{G}_\mathcal{I}}$, let s_i denote the number of agents that are matched to a rank-i edge in μ. Then, μ is*

- ***rank-maximal*** *if μ maximizes the number of agents who are matched to a rank-1 edge and, subject to that, it maximizes the number of agents who are matched to rank-2 edges, and so on. Formally, for each μ' in $\mathcal{M}_{\mathbb{G}_\mathcal{I}}$ define its signature to be the n-tuple $s_{\mu'} = (s_1, \ldots, s_n)$. Then μ is the matching with the lexicographically optimal signature.*

- ***max-cardinality rank-maximal*** *if μ is a maximum cardinality matching and, subject to that, is also rank-maximal. Formally, for each μ' in $\mathcal{M}_{\mathbb{G}_\mathcal{I}}$ define its signature to be the $(n+1)$-tuple $(\sum_{i=1}^{n} s_i, s_1, \ldots, s_n)$. Then μ is the matching with the lexicographically optimal signature.*

- ***fair*** *if μ is a maximum cardinality matching and, subject to that, minimizes the number of agents who are matched to a rank-n edge and, subject to that, minimizes the number of agents who are matched to a rank-$(n-1)$ edge, and so on. Formally, for each μ' in $\mathcal{M}_{\mathbb{G}_\mathcal{I}}$ define its signature to be the $(n+1)$-tuple $(\sum_{i=1}^{n} s_i, -s_n, -s_{n-1}, \ldots, -s_1)$. Then μ is the matching with the lexicographically optimal signature.*

Signature-based matchings can be reduced to an instance of the following problem, which we refer to as priority-**p** matchings, for a given $\mathbf{p} = (\mathbf{p_1}, \cdots, \mathbf{p_n})$ [Irving, 2004; Irving *et al.*, 2006; Mehlhorn and Michail, 2005; Huang *et al.*, 2013; Michail, 2007].

Definition 3. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P} = (P_1, \ldots, P_n))$ and a priority vector $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$, where*

$\forall i \in [n], p_i \in \mathbb{Z}_{\geq 0}$ *and $\exists j, k \in [n]$ such that $p_j \neq p_k$, a matching $\mu \in \mathcal{M}_{\mathbb{G}_\mathcal{I}}$ is said to be a priority-$\mathbf{p}$ matching if μ is a matching of maximum weight in $\mathcal{M}_{\mathbb{G}_\mathcal{I}}$, where a rank-$r$ edge in $\mathbb{G}_\mathcal{I}$ is assigned the weight p_r.*

In particular, given an instance $\mathcal{I}$, we can show that,[1]

- when $p_j = n^{2(n-j+1)}$ for all $j \in [n]$, a matching is a priority-$\mathbf{p}$ matching if and only if it is rank-maximal matching w.r.t. $\mathcal{I}$.

- when $p_j = n^{2n} + n^{2(n-j)}$ for all $j \in [n]$, a matching is a priority-$\mathbf{p}$ matching if and only if it is a max-cardinality rank-maximal matching w.r.t. $\mathcal{I}$.

- when $p_j = 4n^{2n} - 2n^{j-1}$ for all $j \in [n]$, a matching is a priority-$\mathbf{p}$ matching if and only if it is a fair matching w.r.t. $\mathcal{I}$.

For ease of exposition, we sometimes use priority-$\mathbf{p}$, where $p_i = 0$ for all $i \in [n]$ to refer to Pareto optimal matchings. Note that this is purely for notational convenience since the algorithms we discuss in the context of Pareto optimal matchings are extensions to the ones for priority-$\mathbf{p}$ matchings. We also use $\mathbb{P}$ to denote the set of priority vectors of interest.

2.2 Worst-Case Welfare Lost

Given an instance $\mathcal{I}$, we are interested in deterministic algorithms which return a matching that satisfies one of the properties just defined. However, such matchings may not be unique. We argue that a principled way of selecting amongst all such matchings is to consider the cardinal utilities of the agents, returning a matching with small worst-case welfare loss. Formally, for an instance $\mathcal{I}$, consider the set of matchings $S \subseteq \mathcal{M}_{\mathbb{G}_\mathcal{I}}$ such that S is the set of all Pareto optimal/rank-maximal/max-cardinality rank-maximal/fair matchings in $\mathbb{G}_\mathcal{I}$. Next, for a matching $\mu \in S$, $v \in \mathcal{V}_\mathcal{I}$, and for an edge $e = (a_i, h_j) \in \mu$, let value$(e) =$

---

[1] The proof of this can be found in the full version of the paper (arXiv:2011.13977).

Algorithm 1 Welfare Optimal Priority-$\mathbf{p}$ Matching

1: **Input**: $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, priorities $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$, and $v = (v_1, \ldots, v_n)$, where $v_i \colon \mathcal{H} \to [0, 1]$
2: **Output**: welfare-optimal priority-$\mathbf{p}$ matching w.r.t. $\mathcal{I}$
3: $\mathbb{G}_\mathcal{I} = (\mathcal{N} \cup \mathcal{H}, \mathcal{E}) \leftarrow$ graph induced by $\mathcal{I}$
4: **for** $e = (a_i, h_j) \in \mathcal{E}$ **do**
5: $\quad r \leftarrow \text{rank}(a_i, h_j)$
6: $\quad w_e \leftarrow p_r + v_i(h_j)$
7: **end for**
8: $\mu \leftarrow$ max-weight matching in $\mathbb{G}_\mathcal{I}$ with weights $\{w_e\}_{e \in \mathcal{E}}$
9: **return** μ

$v_i(h_j)$ and $\text{SW}(\mu \mid v) = \sum_{e \in \mu} \text{value}(e)$, the social welfare of μ given the valuations v.[2] Given this, consider a deterministic algorithm $\mathcal{A}$ where, for all $\mathcal{I} \in \mathbb{I}$, $\mathcal{A}(\mathcal{I}) \in S$ and let $\mathcal{L}(\mathcal{A})$, which we refer to as the *worst-case welfare loss* of $\mathcal{A}$, be defined as $\mathcal{L}(\mathcal{A}) := \max_{\mathcal{I} \in \mathbb{I}} \mathcal{L}(\mathcal{A}, \mathcal{I})$, where

$$\mathcal{L}(\mathcal{A}, \mathcal{I}) := \sup_{v \in \mathcal{V}_\mathcal{I}} \frac{\max\limits_{\mu^* \in S} \text{SW}(\mu^* \mid v)}{\text{SW}(\mathcal{A}(\mathcal{I}) \mid v)}.$$

The objective is to design algorithms that return a matching with the desired property and minimize $\mathcal{L}(\mathcal{A})$. Throughout this paper, we say that, for an $\alpha \geq 1$, an algorithm $\mathcal{A}$ achieves an α-approximation to the optimal social welfare among Pareto-optimal/rank-maximal/max-cardinality rank-maximal/fair matchings if $\mathcal{L}(\mathcal{A}) \leq \alpha$.

We first observe that any purely ordinal algorithm, that is an algorithm that uses only the ordinal preferences of the agents, performs poorly with respect to worst-case welfare loss. Due to space limitations all proofs are in the full version (arXiv:2011.13977).

Theorem 1. *Let X denote one of the properties in the set {Pareto-optimal, rank-maximal, max-cardinality rank-maximal, and fair}. Let $\mathcal{A}$ be a deterministic ordinal algorithm that always produces a matching that satisfies property X. If there are n agents with unit-sum valuation functions, then $\mathcal{L}(\mathcal{A}) \in \Omega(n^2)$. If there are n agents with unit-range valuation functions then $\mathcal{L}(\mathcal{A}) \in \Omega(n)$. Moreover, these bounds are asymptotically tight.*

At the other extreme, an algorithm may have access to all utility information from the agents and, thus, is capable of returning the welfare-optimal matching subject to the underlying desired property. In particular, given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$ and valuation functions of the agents $v = (v_1, \ldots, v_n)$, where $v_i \colon \mathcal{H} \to [0, 1]$, the welfare-optimal priority-$\mathbf{p}$ problem is to find a matching of maximum welfare among the set of priority-$\mathbf{p}$ matchings. We observe that this reduces to an instance of the max-weight matching problem on $\mathbb{G}_\mathcal{I}$.

Theorem 2. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, a vector of priorities $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$, where $\mathbf{p} \in \mathbb{P}$, and $v = (v_1, \ldots, v_n)$, where $v_i \colon \mathcal{H} \to [0, 1]$, Algorithm 1 returns a welfare-optimal priority-$\mathbf{p}$ matching w.r.t. $\mathcal{I}$.*

---

[2]For notational convenience, when v is clear from the context, we just write $\text{SW}(\mu)$ instead of $\text{SW}(\mu \mid v)$.

3 Binary Threshold Queries

In this section we look at the central question of this paper: *How can one improve social welfare in one-sided matching problems by asking only a small number of queries?* We believe that asking directly for cardinal utility information places a high cognitive burden on agents. Therefore, we aim for a middle-ground between solely ordinal and fully cardinal algorithms. We do this by analysing the power of using very simple queries, namely *binary threshold queries*.

Definition 4. *For an agent a_i, object h_j, and a real number $t_k \in [0, 1]$, a binary threshold query, $\mathcal{Q}(a_i, h_j, t_k)$, asks agent a_i to return 1 (alternatively, asks them to say "Yes") if $v_i(h_j) \geq t_k$, and 0 (alternatively, asks them to say "No") otherwise.*

Given an instance $\mathcal{I}$ and answers to a certain number of binary threshold queries, our goal is to design deterministic algorithms $\mathcal{A}$ that minimize the worst-case welfare loss $\mathcal{L}(\mathcal{A})$ and, for all $\mathcal{I} \in \mathbb{I}$, produces a matching in S (i.e., $\mathcal{A}(\mathcal{I}) \in S$), where S is the set of all Pareto optimal/rank-maximal/max-cardinality rank-maximal/fair matchings in $\mathbb{G}_\mathcal{I}$. Towards this end, we begin by considering adaptive algorithms—algorithms that are allowed to change its queries based on the agents' responses—and show how, when considering each of the four properties of interest, one can obtain a $(1 + \epsilon)$-approximation to the optimal welfare. Following this, we look at, what we believe is the more interesting and practical, case of non-adaptive algorithms. In particular, we restrict ourselves to algorithms that can ask at most one query per (agent, object) pair and show upper and lower bounds on the approximation achievable. Unless explicitly specified, all results hold for both unit-sum and unit-range valuations.

3.1 Adaptive Algorithm to Achieve $(1 + \epsilon)$-Approximation

Given an instance $\mathcal{I}$ and a property encoded as a priority-vector. Algorithm 2 returns a matching with the desired property. The high-level idea behind the algorithm is straightforward. For a specific choice of parameter c, it associates a partition of objects with every agent, where for $k \in [c]$, an object is in $\mathcal{E}_{ik}$ if agent a_i's value for the object is within some defined interval B_k. Using these partitions, carefully computed weights are assigned to the edges in the induced bipartite graph, $\mathbb{G}_\mathcal{I}$. A max-weight matching on the resulting weighted graph is then computed. Below we show for $c = \left\lceil \frac{\log(n^2 \cdot 1/\epsilon)}{\log(1 + \epsilon/2)} \right\rceil$, this results in an $(1 + \epsilon)$-approximation algorithm that uses $O(c \log n)$ queries per agent. In particular, this means that one can achieve a 2-approximation using $O(\log^2 n)$ queries per agent.

Theorem 3. *Given an $\epsilon > 0$, an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, and a priority vector $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$, Algorithm 2 adaptively asks $O(c \log n)$ queries per agent, where $c = \left\lceil \frac{\log(n^2 \cdot 1/\epsilon)}{\log(1 + \epsilon/2)} \right\rceil$, and returns a*

1. *Pareto optimal matching μ that achieves a $(1 + \epsilon)$-approximation to the optimal welfare among all Pareto optimal matchings when $p_i = 0$ for all $i \in [n]$.*

Algorithm 2

1: **Input:** $\epsilon > 0$, $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, and $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$
2: **Output:** a PO matching when $p_i = 0$ for all $i \in [n]$ and a priority-$\mathbf{p}$ matching when $\mathbf{p} \in \mathbb{P}$
3: $\mathbb{G}_{\mathcal{I}} = (\mathcal{N} \cup \mathcal{H}, \mathcal{E}) \leftarrow$ graph induced by $\mathcal{I}$
4: $c \leftarrow \left\lceil \frac{\log(n^2 \cdot 1/\epsilon)}{\log(1+\epsilon/2)} \right\rceil$
5: $t_i \leftarrow (\frac{2}{2+\epsilon})^i$, for $i \in [c]$
6: **for** $a_i \in \mathcal{N}$ **do**
7:    **for** $k \in [c]$ **do**
8:       $\mathcal{E}_{ik} \leftarrow \{(a_i, h_j) \in \mathcal{E} \mid \mathcal{Q}(a_i, h_j, t_k) = 1$ and, if $k \geq 2, \mathcal{Q}(a_i, h_j, t_{k-1}) = 0\}$
9:       **for** $e = (a_i, h_j) \in \mathcal{E}_{ik}$ **do**
10:          $r \leftarrow \text{rank}(a_i, h_j)$
11:          $\text{value}'(e) \leftarrow t_k$
12:          $w_e \leftarrow p_r + \text{value}'(e)$
13:       **end for**
14:    **end for**
15: **end for**
16: $\mu \leftarrow$ max-weight matching in $\mathbb{G}_{\mathcal{I}}$ with weights $\{w_e\}_{e \in \mathcal{E}}$
17: **if** $p_i = 0$ for all $i \in [n]$ **then**
18:    $\mu \leftarrow$ run top-trading cycles (TTC) algorithm with μ (from line 16) as the initial endowment
19: **end if**
20: **return** μ

Algorithm 3

1: **Input:** an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$ and priorities $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$
2: **Output:** a priority-$\mathbf{p}$ matching when $\mathbf{p} \in \mathbb{P}$
3: $\mathbb{G}_{\mathcal{I}} = (\mathcal{N} \cup \mathcal{H}, \mathcal{E}) \leftarrow$ graph induced by $\mathcal{I}$
4: $t_1 \leftarrow \frac{1}{n^{1/3}}$
5: $t_i \leftarrow \frac{1}{\min\{i, n^{1/3}\} \cdot n^{2/3}}$, for all $i \in \{2, \ldots, n\}$
6: **for** $e = (a_i, h_j) \in \mathcal{N} \times \mathcal{H}$ **do**
7:    $r \leftarrow \text{rank}(a_i, h_j)$
8:    **if** $\mathcal{Q}(a_i, h_j, t_r)$ **then**
9:       $w_e \leftarrow p_r + t_r$
10:    **else**
11:       $w_e \leftarrow p_r$
12:    **end if**
13: **end for**
14: $\mu \leftarrow$ max-weight matching in $\mathbb{G}_{\mathcal{I}}$, where weights are $(w_e)_{e \in \mathcal{E}}$
15: **return** μ

Algorithm 4

1: **Input:** an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P} = (P_1, \ldots, P_n))$
2: **Output:** a Pareto optimal matching
3: $\mu'_{MM} \leftarrow$ matching returned by Alg 3 on $\mathcal{I}$ and $\mathbf{p} = (\mathbf{0}, \ldots, \mathbf{0})$
4: $\mu_{MM} \leftarrow \mu'_{MM} \setminus \{e \in \mu_{MM} \mid w_e = 0\}$
5: **if** $|\mu_{MM}| == 0$ **then**
6:    $\mu'_{aux} \leftarrow$ matching in $\mathbb{G}_{\mathcal{I}}$ that maximizes the number of agents who are matched with an edge of rank at most $\lfloor \sqrt[3]{n}/2 \rfloor$
7: **else**
8:    $\mu'_{aux} \leftarrow$ matching in $\mathbb{G}_{\mathcal{I}}$ where as many agents as possible to a rank-1 edge
9: **end if**
10: $\mu_{aux} \leftarrow \mu'_{aux} \setminus \{(a, o) \mid (a, o) \in \mu'_{aux}$ and either a or o is matched in $\mu_{MM}\}$
11: $\mu_{\text{rest}} \leftarrow$ arbitrarily match the acceptable (agent, object) pairs that are not matched in $\mu_{MM} \cup \mu_{aux}$
12: $\mu \leftarrow$ run TTC with $\mu_{MM} \cup \mu_{aux} \cup \mu_{\text{rest}}$ as initial endowments and return the resulting matching.
13: **return** μ

---

2. *priority-$\mathbf{p}$ matching μ that achieves a $(1 + \epsilon)$-approximation to the optimal welfare among all priority-$\mathbf{p}$ matchings when $\mathbf{p} \in \mathbb{P}$.*

We observe that Theorem 3 immediately informs us about distortion. Any algorithm that returns a Pareto-optimal matching and achieves an α-approximation to the optimal welfare amongst all Pareto-optimal matchings must have distortion of α since welfare-optimal matchings are necessarily Pareto-optimal. Thus, we are able to automatically confirm and extend earlier distortion results [Amanatidis *et al.*, 2021], using a weaker query model. Furthermore, we again emphasize our interest in broader classes of matchings.

3.2 Non-Adaptive Algorithms: Asking One Query Per (agent, object) Pair

We now turn our attention to non-adaptive algorithms, in particular looking at algorithms that can only ask one query per (agent, object) pair and cannot change these queries depending on earlier responses. We believe that this is the more interesting and practical setting to consider for this problem, since such an algorithm does not have to wait for the agents to respond and also does not require an agent to answer multiple queries with respect to the same object—doing which would in turn entail that the agent is somewhat sure about their cardinal utilities.

We present two algorithms for when agents have unit-sum valuations, first in the context of priority-$\mathbf{p}$ matchings and second for Pareto optimal matchings. The latter is an extension of the former. In Algorithm 3, first the thresholds for the queries are carefully chosen. Then, in the induced bipartite graph, $\mathbb{G}_{\mathcal{I}}$, weights for the edges are determined by agents' responses to queries. A max-weight matching in this graph is returned and is guaranteed to be a priority-$\mathbf{p}$ matching. If a Pareto-optimal matching is required then a little bit of additional work is needed to handle the situation where Algorithm 3 returns a matching where not all agents and objects are matched. Algorithm 4 is the extension that handles this case, by carefully handling the initial unmatched agents through the construction of an auxiliary matching.

Theorem 4. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$ and a vector of priorities $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$, where $\mathbf{p} \in \mathbb{P}$, Algorithm 3 asks one non-adaptive query per (agent, object) and returns a priority-$\mathbf{p}$ matching that achieves an $O(n^{2/3})$-approximation to the optimal welfare among all priority-$\mathbf{p}$ matchings for the case when agents have unit-sum valuations.*

Theorem 5. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, Algorithm 4 asks one non-adaptive query per (agent, object) pair and returns a Pareto optimal matching that achieves an $O(n^{2/3})$-approximation to the optimal welfare among all Pareto optimal matchings for the case when agents have unit-sum valuations.*

Finally, we also consider the case when agents have unit-range valuations and show that it is possible to obtain an $O(\sqrt{n})$-approximation to the optimal social welfare among

Pareto optimal and priority-$\mathbf{p}$ matchings. The algorithms and analyses for this case share similarities with Algorithms 3 and 4, and can be found in the full version (arXiv:2011. 13977).

Theorem 6. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$ and a vector of priorities $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})$, where $\mathbf{p} \in \mathbb{P}$, there exists an algorithm that asks one non-adaptive query per (agent, object) and returns a priority-$\mathbf{p}$ matching that achieves an $O(\sqrt{n})$-approximation to the optimal welfare among all priority-$\mathbf{p}$ matchings for the case when agents have unit-range valuations.*

Theorem 7. *Given an instance $\mathcal{I} = (\mathcal{N}, \mathcal{H}, \mathcal{P})$, there exists an algorithm that asks one non-adaptive query per (agent, object) and returns a Pareto optimal matching that achieves an $O(\sqrt{n})$-approximation to the optimal welfare among all Pareto optimal matchings for the case when agents have unit-range valuations.*

Lower Bounds

We finally turn our attention to lower bounds for the case when an algorithm can ask at most one query per (agent, object). We show that, for the unit-sum and unit-range valuation cases, any deterministic algorithm $\mathcal{A}$ that asks at most one query per (agent, object) pair and produces a Pareto-optimal/rank-maximal/max-cardinality rank-maximal/fair matching has a worst-case welfare loss of $\Omega(\sqrt{n})$, i.e., $\mathcal{L}(\mathcal{A}) \in \Omega(\sqrt{n})$.

Theorem 8. *Let X denote one of the properties in the set {Pareto-optimal, rank-maximal, max-cardinality rank-maximal, and fair}. Let $\mathcal{A}$ be a non-adaptive deterministic algorithm that always produces a matching that satisfies property X and asks at most one query per (agent, object) pair. If there are n agents with unit-sum valuation functions or unit-range valuations, then $\mathcal{L}(\mathcal{A}) \in \Omega(\sqrt{n})$.*

4 Discussion

We investigated the benefit of eliciting a small amount of information about agents' cardinal utilities in the context of one-sided matching. We designed algorithms that used simple threshold queries and returned a matching satisfying some desirable matching property, while also achieving a good approximation to the optimal welfare among all matchings satisfying that property. Our results show how we can improve welfare by even non-adaptively asking agents for just one bit of extra information per object. Given a one-sided matching instance, there are often multiple matchings of interest and we view the methodology we presented here as providing a principled way of tie-breaking.

While adaptive algorithms provide better approximation guarantees, we believe the benefits of non-adaptive approaches outweigh the negatives. The argument for not insisting agents reveal full and exact cardinal utility information in the first place is that this places too high a cognitive demand on the agents. Our non-adaptive approach, which asks for only one bit of information for each (agent, object) pair, could be implemented using a simple menu – reducing the time and effort an individual agent must interact with the matching process.

There are a number of future research directions that this work can take. For example, we may want multiple agents to be assigned the same object, like when assigning students to courses or schools. Only minimal modifications are needed to address this case. In particular, every time we construct a graph in any of the algorithms, all that needs to be done is to create k_j copies for the node that corresponds to object h_j. Other open algorithmic problems include addressing the gap between the upper and lower bounds for the non-adaptive algorithms, expanding the set of properties of interest to include, for example, popular matchings [Abraham *et al.*, 2007], or asking similar questions in the context of two-sided matching problems. Finally, we are interested in better understanding the implications of deploying such an approach in practice, including evaluating how interface-design might best support queries, as well as better understanding what matching properties are deemed to be most important by users and designers of systems.

References

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

[Abdulkadiroğlu and Sönmez, 1999] Atila Abdulkadiroğlu and Tayfun Sönmez. House allocation with existing tenants. *Journal of Economic Theory*, 88(2):233–260, 1999.

[Abdulkadiroğlu *et al.*, 2011] Atila Abdulkadiroğlu, Yeon-Koo Che, and Yosuke Yasuda. Resolving conflicting preferences in school choice: The "Boston mechanism" reconsidered. *American Economic Review*, 101(1):399–410, 2011.

[Abdulkadiroğlu *et al.*, 2015] Atila Abdulkadiroğlu, Yeon-Koo Che, and Yosuke Yasuda. Expanding "choice" in school choice. *American Economic Journal: Microeconomics*, 7(1):1–42, 2015.

[Abraham *et al.*, 2004] David J. Abraham, Katarína Cechlárová, David F. Manlove, and Kurt Mehlhorn. Pareto optimality in house allocation problems. In *Proceedings of ISAAC-2004*, pages 3–15, 2004.

[Abraham *et al.*, 2006] David J. Abraham, Ning Chen, Vijay Kumar, and Vahab S Mirrokni. Assignment problems in rental markets. In *Proceedings of WINE-2006*, pages 198–213, 2006.

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

[Abramowitz and Anshelevich, 2018] Ben Abramowitz and Elliot Anshelevich. Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences. In *Proceedings of AAAI-2018*, pages 894–901, 2018.

[Abramowitz *et al.*, 2019] Ben Abramowitz, Elliot Anshelevich, and Wennan Zhu. Awareness of voter passion greatly

improves the distortion of metric social choice. In *Proceedings of WINE-2019*, pages 3–16, 2019.

[Amanatidis *et al.*, 2020] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. Peeking behind the ordinal curtain: Improving distortion via cardinal queries. In *Proceedings of AAAI-2020*, pages 1782–1789, 2020.

[Amanatidis *et al.*, 2021] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. A few queries go a long way: Information-distortion tradeoffs in matching. In *Proceedings of AAAI-2021*, 2021.

[Anshelevich and Sekar, 2016] Elliot Anshelevich and Shreyas Sekar. Blind, greedy, and random: Algorithms for matching and clustering using only ordinal information. In *Proceedings of AAAI-2016*, pages 383–389, 2016.

[Anshelevich and Zhu, 2017] Elliot Anshelevich and Wennan Zhu. Tradeoffs between information and ordinal approximation for bipartite matching. In *Proceedings of SAGT-2017*, pages 267–279, 2017.

[Benade *et al.*, 2020] Gerdus Benade, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. Preference elicitation for participatory budgeting. *Management Science, Articles in Advance*, pages 1–15, 2020.

[Boutilier *et al.*, 2015] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213, 2015.

[Bredereck *et al.*, 2017] Robert Bredereck, Piotr Faliszewski, Andrzej Kaczmarczyk, Rolf Niedermeier, Piotr Skowron, and Nimrod Talmon. Robustness among multiwinner voting rules. In *Proceedings of SAGT-2017*, pages 80–92, 2017.

[Chen *et al.*, 2019] Jiehua Chen, Piotr Skowron, and Manuel Sorge. Matchings under preferences: Strength of stability and trade-offs. In *Proceedings of ACM EC-2019*, pages 41–59, 2019.

[Chiesa *et al.*, 2012] Alessandro Chiesa, Silvio Micali, and Zeyuan Allen Zhu. Mechanism design with approximate valuations. In *Proceedings of ITCS-2012*, pages 34–38, 2012.

[Chiesa *et al.*, 2014] Alessandro Chiesa, Silvio Micali, and Zeyuan Allen Zhu. Knightian self uncertainty in the VCG mechanism for unrestricted combinatorial auctions. In *Proceedings of ACM EC-2014*, pages 619–620, 2014.

[Filos-Ratsikas *et al.*, 2014] Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, and Jie Zhang. Social welfare in one-sided matchings: Random priority and beyond. In *Proceedings of SAGT-2014*, pages 1–12, 2014.

[Goel *et al.*, 2017] Ashish Goel, Anilesh K. Krishnaswamy, and Kamesh Munagala. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings ACM EC-2017*, pages 287–304, 2017.

[Goel *et al.*, 2019] Ashish Goel, Anilesh K. Krishnaswamy, Sukolsak Sakshuwong, and Tanja Aitamurto. Knapsack

voting for participatory budgeting. *ACM Transactions on Economics and Computation*, 7(2):1–27, 2019.

[Huang *et al.*, 2013] Chien-Chung Huang, Telikepalli Kavitha, Kurt Mehlhorn, and Dimitrios Michail. Fair matchings and related problems. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, 2013.

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

[Irving *et al.*, 2006] Robert W. Irving, Telikepalli Kavitha, Kurt Mehlhorn, Dimitrios Michail, and Katarzyna E. Paluch. Rank-maximal matchings. *ACM Transactions on Algorithms*, 2(4):602–610, October 2006.

[Irving, 2004] Robert W. Irving. Greedy matchings. TR-2004-177, University of Glasgow, Computing Science Department Research Report, 2004.

[Mandal *et al.*, 2020] Debmalya Mandal, Nisarg Shah, and David P. Woodruff. Optimal communication-distortion tradeoff in voting. In *Proceedings of ACM-EC 2020*, pages 795–813, 2020.

[Mehlhorn and Michail, 2005] Kurt Mehlhorn and Dimitrios Michail. Network problems with non-polynomial weights and applications. Manuscript, 2005.

[Menon and Larson, 2018] Vijay Menon and Kate Larson. Robust and approximately stable marriages under partial information. In *Proceedings of WINE-2018*, pages 341–355, 2018.

[Menon and Larson, 2019] Vijay Menon and Kate Larson. Mechanism design for locating a facility under partial information. In *Proceedings of SAGT-2019*, pages 49–62, 2019.

[Michail, 2007] Dimitrios Michail. Reducing rank-maximal to maximum weight matching. *Theoretical Computer Science*, 389(1-2):125–132, 2007.

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

[Roth and Postlewaite, 1977] Alvin E. Roth and Andrew Postlewaite. Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics*, 4(2):131–137, 1977.

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

[Shiryaev *et al.*, 2013] Dmitry Shiryaev, Lan Yu, and Edith Elkind. On elections with robust winners. In *Proceedings of AAMAS-2013*, pages 415–422, 2013.

[Sönmez and Ünver, 2010] Tayfun Sönmez and M. Utku Ünver. House allocation with existing tenants: A characterization. *Games and Economic Behavior*, 69(2):425–445, 2010.