Metric Distortion under Group-Fair Objectives

Georgios Amanatidis[1], Elliot Anshelevich[2],
Christopher Jerrett[2] and Alexandros A. Voudouris[1]

[1]University of Essex
[2]Rensselaer Polytechnic Institute

Abstract

We consider a voting problem in which a set of agents have metric preferences over a set of alternatives, and are also partitioned into disjoint groups. Given information about the preferences of the agents and their groups, our goal is to decide an alternative to approximately minimize an objective function that takes the groups of agents into account. We consider two natural group-fair objectives known as *Max-of-Avg* and *Avg-of-Max* which are different combinations of the max and the average cost in and out of the groups. We show tight bounds on the best possible *distortion* that can be achieved by various classes of mechanisms depending on the amount of information they have access to. In particular, we consider *group-oblivious full-information* mechanisms that do not know the groups but have access to the exact distances between agents and alternatives in the metric space, *group-oblivious ordinal-information* mechanisms that again do not know the groups but are given the ordinal preferences of the agents, and *group-aware* mechanisms that have full knowledge of the structure of the agent groups and also ordinal information about the metric space.

1 Introduction

One of the main subjects of study in (computational) social choice theory is to identify the capabilities and limitations of making appropriate collective decisions when given the preferences of individuals (or, *agents*) over alternative outcomes. This is done either by an axiomatic analysis of the potential decision-making mechanisms (which are also referred to as voting rules) [Brandt et al., 2016], or a qualitative analysis that aims to quantify the possible loss of efficiency when the agents have private cardinal utilities or costs for the alternatives but are only able to communicate partial information about their preferences, for example using ordinal information. This inefficiency is quantified by the notion of *distortion* which compares the quality of the computed outcome to that of the ideal outcome that could have been computed if full information about the underlying utilities of the agents was available. Since its introduction more than 15 year ago, distortion has been studied for many different social choice problems (such as voting applications, resource allocation, and facility location) and under different restrictions about the cardinal preferences of the agents (such as assuming unit-sum utilities or metric costs). For a more detailed overview see our discussion of the related work below and the survey of Anshelevich et al. [2021].

With few exceptions, the distortion literature has focused on voting settings in which the agents are assumed to be independent of each other. As such, the quality of the different outcomes is typically calculated using measures such as the *social welfare* (defined as the total or average utility of all agents) or the *egalitarian welfare* (defined as the minimum utility over all agents). However, there are social choice applications in which, while the agents can act autonomously, they are also part of larger groups and care about the overall welfare of the members of their groups, but not that much about other groups. As a toy example, consider the case of a university department, the academics of which are members of different research groups. For several matters, such as electing the head of the department, each academic participates individually in the voting process, but the outcome might affect different groups in different ways. Due to this, objectives such as the social and the egalitarian welfare do not fully capture the quality of an outcome according to the structure of the problem. Instead, we would like objectives that take into account the partition of the agents into groups to measure efficiency and also satisfy other desired properties such as fairness or some form of balance among different groups.

Two such natural objectives were first introduced by Anshelevich et al. [2022] who studied a metric district-based single-winner voting setting, where the agents have *costs* for the alternatives that are determined by their *distances* in a metric space, and the agents are furthermore partitioned into groups that are called *districts*. The first objective is *Max-of-Avg*, defined as the maximum over all groups of the average total distance of the agents within each group from the chosen alternative, and the second one is *Avg-of-Max*, defined as the average over all groups of the maximum distance among any agent within each group from the chosen alternative.[1] By their definition, to optimize them, we need to choose outcomes that strike a balance between the average or maximum cost of any group as a whole, thus achieving fairness among different groups, on top of absolute efficiency.

1.1 Our Contribution

We study a single-winner voting setting with n agents and m alternatives that lie in a metric space. Furthermore, the agents are partitioned into k disjoint groups. Given some information about the groups of agents, as well as the distances between agents and alternatives in the metric space, our goal is to choose an alternative as the winner that is (approximately) efficient with respect to the Max-of-Avg and Avg-of-Max objectives that were defined above. In particular, we show tight bounds on the distortion of different classes of *deterministic* mechanisms, depending on the type of information they have access to in order to decide the winner.

We start by considering *group-oblivious* mechanisms which have no knowledge of the groups of agents. In Section 3, we consider the class of *full-information* group-oblivious mechanisms which have complete information about the distances between agents and alternatives in the metric space. For such mechanisms, we show a tight bound of 3 for Max-of-Avg, a tight bound of 3 for Avg-of-Max on instances in which the groups are symmetric (i.e., all groups have the same size), and a tight bound of k for Avg-of-Max on general instances. In Section 4, we turn our attention to *ordinal-information* group-oblivious mechanisms which are given as input the ordinal preferences of the agents over the alternatives in the form of rankings from the smallest distance to the largest. We show a tight bound of 5 for Max-of-Avg, a tight bound of 5 for Avg-of-Max on instances with symmetric groups, and a tight bound of $2k + 1$ for Avg-of-Max on general instances. An overview of our results for group-oblivious mechanisms is given in Table 1.

In Section 5 we turn our attention to *group-aware* mechanisms which know the structure of the groups of agents. Having full information about the metric space on top of this knowledge about the groups makes the problem of optimizing the Max-of-Avg and the Avg-of-Max objectives trivial by simply calculating the cost of every alternative. Consequently, we consider group-aware mechanisms that have access to the ordinal preferences of the agents over the alternatives instead. For instances with two alternatives, we prove a tight bound of 3 on the distortion of such mechanisms for both objectives. For general instances, we show that the distortion is still 3 when we are allowed to exploit more information about the metric space for the upper bound. In particular, we assume access to the

[1] Observe that both of these objectives are essentially combinations of the *social cost* and the *egalitarian cost*, which are the analogues of the social welfare and egalitarian welfare when the agents have costs for the alternatives rather than utilities.

		Full-information	Ordinal-information
Max-of-Avg		3 (Theorems 3.1, 3.2)	5 (Theorems 4.1, 4.3)
Avg-of-Max	Symmetric	3 (Theorems 3.3, 3.4)	5 (Theorems 4.4, 4.5)
	Asymmetric	k (Theorems 3.5, 3.6)	$2k+1$ (Theorems 4.6, 4.7)

Table 1: An overview of our tight distortion bounds for the class of group-oblivious mechanisms.

distances between the alternatives. Resolving the distortion of group-aware mechanisms is probably the most challenging open question that our work leaves open; we discuss this in Section 6.

1.2 Related Work

Inspired by worst-case analysis, Procaccia and Rosenschein [2006] introduced the distortion as a means of quantifying the inefficiency of voting mechanisms that base their decisions on the ordinal preferences of the agents over the alternative outcomes. Without restricting the possible underlying cardinal utilities of the agents, the distortion can be shown to be unbounded for most natural mechanisms. This led to subsequent works to study voting settings where it is assumed that the agents have underlying normalized utilities [Boutilier et al., 2015, Caragiannis et al., 2017, Ebadian et al., 2022, 2023a], or costs determined by distances in an unknown metric space [Anshelevich et al., 2018, Gkatzelis et al., 2020, Kizilkaya and Kempe, 2022, Charikar and Ramakrishnan, 2022, Charikar et al., 2024, Caragiannis et al., 2022, Jaworski and Skowron, 2020], or combinations of the two [Gkatzelis et al., 2023]. The distortion has also been studied for other social choice problems, such as participatory budgeting [Benadè et al., 2021], matching [Filos-Ratsikas et al., 2014, Amanatidis et al., 2022], as well as clustering [Anshelevich and Sekar, 2016, Burkhardt et al., 2024] and other graph problems where only ordinal information is available [Abramowitz and Anshelevich, 2018]. We refer to the survey of Anshelevich et al. [2021] for a more detailed exposition of the distortion framework and the problems it has been applied to.

While the bulk of the distortion literature has focused on settings where ordinal or even less than ordinal information is available about the preferences of the agents, there has been recent interest in settings where it is also possible to elicit some cardinal information. For example, the agents might be able to communicate a number of bits about their preferences [Mandal et al., 2019, 2020, Kempe, 2020], or answer value queries related to their utilities about the alternatives [Amanatidis et al., 2021, 2022, 2024, Ma et al., 2021, Caragiannis and Fehrs, 2023, Burkhardt et al., 2024], or provide more information in the form of intensities [Abramowitz et al., 2019, Kahng et al., 2023] or threshold approvals [Bhaskar et al., 2018, Benadè et al., 2021, Ebadian et al., 2023b, Anshelevich et al., 2024, Latifian and Voudouris, 2024]. In our work, we also consider more than ordinal information in the case of full-information group-oblivious mechanisms, where the main source of inefficiency comes from not knowing the structure of the groups of agents.

As already previously mentioned, the particular objective functions (Max-of-Avg and Avg-of-Max) that we consider in this paper have been studied in the context of distortion by Anshelevich et al. [2022] for single-winner distributed metric voting, and subsequently by Voudouris [2023] for the same setting, and by Filos-Ratsikas et al. [2024] for distributed facility location on the line. In those settings, similarly to our model here, the agents are partitioned into groups that are called districts, and a mechanism works in two steps: First, for each district, it decides a representative alternative or location based on given information about the preferences of the agents in the district, and then it decides a winner or a facility location based on information about the district representatives. Such distributed mechanisms can be thought of as members of the class of group-aware mechanisms in our setting when the groups are assumed to be known. However, they are very restricted as they essentially forget any detailed in-group information in the second step and instead rely only on the group representatives to make final decisions. The Max-of-Avg and Avg-of-Max objectives have also been considered in the context of mechanism design without money for altruistic facility location problems by Zhou et al. [2022, 2024].

2 Preliminaries

An instance I of our voting problem consists of a set N of $n \geq 2$ *agents* and a set A of $m \geq 2$ *alternatives*. Agents and alternatives are represented by points in a metric space. We denote by $d(x, y)$ the *distance* between any two points x and y in the metric space; the distance function satisfies the properties $d(x, x) = 0$, $d(x, y) = d(y, x)$, and the triangle inequality $d(x, y) \leq d(x, z) + d(z, y)$ for any $x, y, z \in N \cup A$. The agents are partitioned into $k \geq 2$ pairwise disjoint *groups* which may be known or unknown; Let $G := \{g_1, \dots, g_k\}$ be the set of groups, and denote by n_g the *size* of any group $g \in G$. If the groups are *symmetric*, to simplify our notation we write $n_g = \lambda = n/k$.

A *mechanism* M takes as input some information $\text{info}(I)$ related to the groups of agents and the distances between agents and alternatives in the metric space. Based on this information, it outputs one of the alternatives as the *winner*, denoted by $M(\text{info}(I))$. When the groups are assumed to be unknown (Sections 3 and 4), we consider two different classes of *group-oblivious* mechanisms depending on the type of information related to the metric space they have access to:

- *Full-information* group-oblivious mechanisms have complete knowledge of the metric space, that is, they have access to the distances between all agents and alternatives.
- *Ordinal-information* group-oblivious mechanisms have access to the ordinal preferences of the agents over the alternatives according to their distances; that is, if $d(i, x) < d(i, y)$ for an agent i and alternatives x and y, then i ranks x higher y.

When the groups are assumed to be known (Section 5), we consider *group-aware* mechanisms that have access to the ordinal preferences of the agents and—potentially—information related to the distances between alternatives (but not between agents, or between agents and alternatives).

We are interested in designing socially efficient mechanisms according to collective cost objective functions that take the groups of the agents into account. In particular, we focus on the following two objectives:

- The *Max-of-Avg* cost of an alternative x in a given instance I is the maximum over all groups of the average total distance of the agents within each group from x, that is,
$$\text{Max-of-Avg}(x \mid I) = \max_{g \in G} \left(\frac{1}{n_g} \sum_{i \in g} d(i, x) \right).$$
- The *Avg-of-Max* cost of an alternative x in a given instance I is the average over all groups of the maximum distance of any agent within each group from x, that is,
$$\text{Avg-of-Max}(x \mid I) = \frac{1}{k} \sum_{g \in G} \max_{i \in g} d(i, x).$$

Whenever the cost objective is clear from context, we will simplify our notation and write $\text{cost}(x \mid I)$ for the cost of alternative x in a given instance I. We will simplify our notation even more and write $\text{cost}(x)$ when the instance is also clear from context.

Since the mechanisms we consider only have partial information about the groups of agents or the metric space, they cannot always identify the alternatives that optimize cost objectives which depend on the structure of the groups, like Max-of-Avg and Avg-of-Max. The loss of efficiency of a mechanism M is captured by its *distortion*, which is the worst-case ratio (over all possible instances) of the cost of the alternative chosen by M over the minimum possible cost of *any* alternative, that is

$$\sup_I \frac{\text{cost}(M(\text{info}(I)) \mid I)}{\min_x \text{cost}(x \mid I)}.$$

Observe that the distortion of any mechanism is always at least 1; we define $0/0 = 1$ for consistency. We aim to reveal the best possible distortion that can be achieved by mechanisms in this group voting setting.

3 Full-Information Group-Oblivious Mechanisms

We start the presentation of our technical results with the class of full-information group-oblivious mechanisms; recall that such mechanisms have complete access to the distances between all agents and alternatives, which means that their inefficiency is solely due to being oblivious to how the agents are partitioned into groups. For the Max-of-Avg objective, we show a tight bound of 3 on the distortion of full-information mechanisms (Section 3.1). For the Avg-of-Max objective, we first show a tight bound of 3 for instances in which the groups are symmetric, and a tight bound of k for general instances with asymmetric groups (Section 3.2).

3.1 Max-of-Avg

We begin by showing a lower bound of 3 on the distortion of full-information group-oblivious mechanisms for the Max-of-Avg objective using an instance with symmetric groups.

Theorem 3.1. *For Max-of-Avg, the distortion of any full-information group-oblivious mechanism is at least $3 - \varepsilon$ for any $\varepsilon > 0$, even when there are only two alternatives and the groups are symmetric.*

Proof. Let $\varepsilon > 0$ be any constant and $\lambda \in \mathbb{N}$ be such that $\lambda > \frac{6}{\varepsilon} - 2$. Consider the following instance with $n = \lambda(\lambda + 1)$ agents and two alternatives with known locations on the line of real numbers:

- Alternative a is at 1 and alternative b is at 3;
- There are λ agents at 0, $\lambda(\lambda - 1)$ agents at 2, and λ agents at 4.

Due to the symmetric locations of the alternatives and the agents, any of the two alternatives can be chosen as the winner. We assume the winner is a, without loss of generality. The agents might be partitioned into the following $k = \lambda + 1$ symmetric groups of size λ each:

- The first group consists of all the λ agents at 4;
- Each of the remaining λ groups consists of one agent at 0 and $\lambda - 1$ agents at 2.

The total distance of the agents in the first group is 3λ from a and λ from b, whereas the total distance of the agents in each of the remaining groups is λ from a and $\lambda + 2$ from b. Hence, $\text{cost}(a) = 3$ and $\text{cost}(b) = 1 + \frac{2}{\lambda}$, leading to a distortion of at least $\frac{3\lambda}{\lambda+2} = 3 - \frac{6}{\lambda+2} > 3 - \varepsilon$. □

It is not hard to obtain a matching upper bound of 3 by using a mechanism that chooses the winner to be any alternative that minimizes the total distance of all agents. In Appendix A we present a refined analysis of this mechanism, by characterizing the worst-case distortion instances, and we obtain a distortion upper bound of $3 - \frac{2\mu}{n}$, where μ is the smallest group size and n is the number of agents.

Theorem 3.2. *For Max-of-Avg, the distortion of a mechanism that returns an alternative who minimizes the total distance from all agents is at most* 3.

Proof. Let w be an alternative that minimizes the total distance from all agents, and let o be an optimal alternative (that minimizes the Max-of-Avg cost according to the unknown groups of the agents). By the definition of w, there must exist some group γ such that $\sum_{i\in\gamma} d(i,w) \leq \sum_{i\in\gamma} d(i,o)$; otherwise, the total distance of o from all agents would be strictly less than that of w, thus contradicting the choice of w. By the definition of the objective function, we also have that $\text{cost}(o) \geq \frac{1}{n_g}\sum_{i\in g} d(i,o)$ for every group g. Denoting by g_w the group that determines the cost of w and using the triangle inequality, we have

$$\text{cost}(w) = \frac{1}{n_{g_w}}\sum_{i\in g_w} d(i,w) \leq \frac{1}{n_{g_w}}\sum_{i\in g_w}(d(i,o)+d(w,o)) \leq \text{cost}(o) + d(w,o).$$

Using the triangle inequality and the property of group γ mentioned above, we further have that

$$d(w,o) = \frac{1}{n_\gamma}\sum_{i\in\gamma} d(w,o) \leq \frac{1}{n_\gamma}\sum_{i\in\gamma}(d(i,w)+d(i,o)) \leq 2\cdot\frac{1}{n_\gamma}\sum_{i\in\gamma} d(i,o) \leq 2\cdot\text{cost}(o).$$

Combining the two, we obtain $\text{cost}(w) \leq 3\cdot\text{cost}(o)$, i.e., the desired upper bound of 3. □

3.2 Avg-of-Max

For the Avg-of-Max objective, we first focus on instances where the groups are symmetric (that is, every group consists of the same number $\lambda = n/k$ of agents) and show a tight bound of 3.

Theorem 3.3. *For Avg-of-Max, the distortion of any full-information group-oblivious mechanism is at least $3-\varepsilon$ for any $\varepsilon > 0$, even when there are two alternatives and the groups are symmetric.*

Proof. Let $\varepsilon > 0$ be any constant and $\lambda \in \mathbb{N}$ be such that $\lambda > \frac{8}{\varepsilon} - 3$. We consider the same instance construction as in the proof of Theorem 3.1 on the line of real numbers. Recall that:

- Alternative a is at 1 and alternative b is at 3;
- There are λ agents at 0, $\lambda(\lambda-1)$ agents at 2, and λ agents at 4.

We assumed that the winner is a, which is without loss of generality due to symmetry. The agents are partitioned into the $k = \lambda + 1$ symmetric groups:

- The first group consists of all the λ agents at 4;
- Each of the remaining λ groups consists of one agent at 0 and $\lambda - 1$ agents at 2.

Therefore, $\text{cost}(a) = (3+\lambda)/(\lambda+1)$ and $\text{cost}(b) = (1+3\lambda)/(\lambda+1)$, leading to a distortion of at least $\frac{3\lambda+1}{\lambda+3} = 3 - \frac{8}{\lambda+3} > 3 - \varepsilon$. □

The tight upper bound follows again by choosing any alternative that minimizes the total distance from all agents; hence, this very simple mechanism is best possible in terms of both the Max-of-Avg objective for general instances and the Avg-of-Max objective for instances with symmetric groups.

Theorem 3.4. *For Avg-of-Max and symmetric groups, the distortion of a mechanism that returns an alternative who minimizes the total distance from all agents is at most* 3.

Proof. Let w be an alternative that minimizes the total distance from all agents, and denote by o an optimal alternative (that minimizes the Avg-of-Max cost according to the k unknown groups of agents). Let $S_1, \dots, S_\lambda$ be any partition of the agents into $\lambda = n/k$ disjoint sets of size k such that each set consists of one agent per group; note that there are multiple such partitions of the agents since the groups are symmetric. By the definition of w, there must exist some $\ell \in [\lambda]$ such that $\sum_{i \in S_\ell} d(i, w) \leq \sum_{i \in S_\ell} d(i, o)$ since, otherwise, the total distance of o from the agents would be strictly less than that of w, thus contradicting the choice of w. Let i_g be a most-distant agent in group g from w, i.e., $i_g \in \arg\max_{i \in g} d(i, w)$. By matching each i_g to a unique agent $f(i_g) \in S_\ell$ (i.e., $f : \{i_{g_1}, \dots, i_{g_k}\} \to S_\ell$ is a bijection), we can rewrite the property of set S_ℓ as

$$\sum_{g \in G} d(f(i_g), w) \leq \sum_{g \in G} d(f(i_g), o).$$

In addition, by the definition of the objective function, we have that

$$\text{cost}(o) \geq \frac{1}{k} \sum_{g \in G} d(i_g, o)$$

and

$$\text{cost}(o) \geq \frac{1}{k} \sum_{i \in S_\ell} d(i, o) = \frac{1}{k} \sum_{g \in G} d(f(i_g), o).$$

Hence, by applying the triangle inequality twice, we obtain

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{k} \sum_{g \in G} d(i_g, w) \\
&\leq \frac{1}{k} \sum_{g \in G} \big(d(i_g, o) + d(f(i_g), o) + d(f(i_g), w) \big) \\
&\leq 3 \cdot \text{cost}(o),
\end{aligned}$$

which shows the desired upper bound of 3. ☐

We now turn our attention to the general case where the groups might be asymmetric and show a tight bound of k.

Theorem 3.5. *For Avg-of-Max, the distortion of any full-information group-oblivious mechanism is at least k, even when there are two alternatives.*

Proof. Consider the following instance with $n = 2k$ agents and two alternatives on the line of real numbers:

- Alternative a is at 0 and alternative b is at 1;
- There are k agents at 0 and k agents at 1.

Due to symmetry, given only this information, any of the two alternatives can be chosen as the winner. Without loss of generality, we assume the winner is a. In that case, however, the agents might be split into k groups as follows:

- The first group consists of all agents at 0 and one agent at 1;
- Each of the remaining $k-1$ groups consists of a single agent at 1.

Hence, cost$(a) = 1$ and cost$(b) = 1/k$, leading to a distortion of k. □

For the upper bound, we first remark that choosing any alternative that minimizes the total distance from all agents (as we did in the case of Avg-of-Max, or Max-of-Avg with symmetric groups) leads to a distortion of at least $2k+1$. Nevertheless, we can achieve a matching bound of k by choosing any alternative that minimizes the maximum distance from the agents.

Theorem 3.6. *For Avg-of-Max, the distortion of a mechanism that returns an alternative who minimizes the maximum distance from any agent is at most k.*

Proof. Let w be the chosen alternative and o an optimal alternative. Let i_w and i_o be the most distant agents from w and o, respectively. Then, by the definition of w, $d(i_w, w) \leq d(i_o, o)$. By the definition of i_w, $d(i, w) \leq d(i_w, w)$ for every agent i. Hence,

$$\text{cost}(w) = \frac{1}{k} \sum_{g \in G} \max_{i \in g} d(i, w) \leq \frac{1}{k} \sum_{g \in G} d(i_w, w) = d(i_w, w).$$

On the other hand,

$$\text{cost}(o) = \frac{1}{k} \sum_{g \in G} \max_{i \in g} d(i, o) \geq \frac{1}{k} d(i_o, o) \geq \frac{1}{k} d(i_w, w).$$

Consequently, the distortion is at most k. □

4 Ordinal-Information Group-Oblivious mechanisms

We now consider mechanisms that are given access to ordinal information about the preferences of the agents over the alternatives, but are still oblivious to how the agents are partitioned into groups. Recall that every agent i reports a ranking of the alternatives such that, if $d(i, x) < d(i, y)$ for alternatives x and y, then i ranks x higher y. For the Max-of-Avg objective, we show a tight bound of 5 on the distortion of ordinal-information group-oblivious mechanisms. For the Avg-of-Max objective, we show that the distortion is exactly 5 when the groups are symmetric, and is exactly $2k+1$ when the groups are asymmetric.

4.1 Max-of-Avg

We start by showing a lower bound of 5 on the distortion of any mechanism.

Theorem 4.1. *For Max-of-Avg, the distortion of any ordinal-information group-oblivious mechanism is at least $5 - \varepsilon$ for any $\varepsilon > 0$, even when there are only two alternatives and the groups are symmetric.*

Proof. Let $\varepsilon > 0$ be any constant and $\lambda \geq 3$ be an odd integer such that $\lambda > \frac{4}{\varepsilon} - 1$. Consider an instance with $n = \lambda^2 + \lambda = \lambda(\lambda + 1)$ agents and two alternatives $\{a, b\}$; clearly $\lambda^2 + \lambda$ is an even number. Half of the agents prefer a and the other half prefer b. With this information, any of the two alternatives can be chosen as the winner, so assume that the winner is a. The agents are partitioned into $k = \lambda + 1$ symmetric groups of λ agents each. Consider the scenario where the metric space is the line of real numbers and the grouping of the agents is as follows:

- a is at 0 and b is at 2;
- In the first group, all λ agents prefer b and are all positioned at $2 + \frac{\lambda+1}{2\lambda}$.
- In each of the remaining λ groups, there are $\frac{\lambda+1}{2}$ agents that prefer a and are positioned at 1, and $\frac{\lambda-1}{2}$ agents that prefer b and are positioned at 2.

The total distance of the agents in the first group is $\lambda\left(2 + \frac{\lambda+1}{2\lambda}\right) = \frac{5\lambda+1}{2}$ from a and $\lambda \cdot \frac{\lambda+1}{2\lambda} = \frac{\lambda+1}{2}$ from b. In each of the remaining λ groups, the total distance of the agents therein is $\frac{\lambda+1}{2} + 2 \cdot \frac{\lambda-1}{2} = \frac{3\lambda-1}{2}$ from a and $\frac{\lambda+1}{2}$ from b. Consequently, $\text{cost}(a) = \frac{5\lambda+1}{2\lambda}$ (realized by the first group) and $\text{cost}(b) = \frac{\lambda+1}{2\lambda}$ (realized by any of the groups), leading to a distortion of at least $5 - \frac{4}{\lambda+1} > 5 - \varepsilon$. ☐

We now show that there are ordinal-information group-oblivious mechanisms which do achieve this best possible bound of 5. The *domination graph* of an alternative x is a bipartite graph $G_x = (N, N, E_x)$ with the set of agents on both sides and set of (directed) edges such that $(i, j) \in E_x$ if and only if i prefers x over the most-preferred alternative $\text{top}(j)$ of j, that is, $d(i, x) \leq d(i, \text{top}(j))$. We focus on alternatives whose domination graphs attain perfect matchings. There are several voting rules that compute alternatives with this property, such as PluralityMatching [Gkatzelis et al., 2020] and PluralityVeto [Kizilkaya and Kempe, 2022]. The distortion of these rules in terms of the social cost (the total distance of the agents) is known to be exactly 3. We show the following property of such alternatives, which will be useful in some of our upper bounds.

Lemma 4.2. *Given an instance, let x be some alternative whose domination graph attains a perfect matching, and y any other alternative. Then,*

$$d(x, y) \leq \frac{4}{n} \cdot \sum_{g \in G} \sum_{i \in g} d(i, y).$$

Proof. Let $\boldsymbol{\mu} = (\mu(i))_i$ be the perfect matching in the domination graph G_x of x; that is, agent i is matched to agent $\mu(i)$. By the triangle inequality, the property of the domination graph that $d(i, x) \leq d(i, \text{top}(\mu(i)))$, the fact that M is a perfect matching, and the fact that $d(i, \text{top}(i)) \leq d(i, y)$ for any i, we have

$$\begin{aligned}
n \cdot d(x, y) = \sum_{i \in N} d(x, y) &\leq \sum_{i \in N} d(i, x) + \sum_{i \in N} d(i, y) \\
&\leq \sum_{i \in N} d(i, \text{top}(\mu(i))) + \sum_{i \in N} d(i, y) \\
&\leq \sum_{i \in N} \big(d(i, y) + d(\mu(i), y) + d(\mu(i), \text{top}(\mu(i))) \big) + \sum_{i \in N} d(i, y) \\
&\leq \sum_{i \in N} d(i, y) + 2 \cdot \sum_{i \in N} d(\mu(i), y) + \sum_{i \in N} d(i, y) \\
&= 4 \cdot \sum_{g \in G} \sum_{i \in g} d(i, y).
\end{aligned}$$

The statement now follows by dividing each side of the inequality by n. ☐

We are now ready to show the upper bound of 5 for Max-of-Avg.

Theorem 4.3. *For Max-of-Avg, the distortion of a mechanism that returns an alternative whose domination graph has a perfect matching is at most* 5.

Proof. Let w be the chosen alternative (whose domination graph has a perfect matching), and o an optimal alternative. Let g_w be the group that determines the maximum cost of w. By the definition of Max-of-Avg, we have $n_g \cdot \text{cost}(o) \geq \sum_{i \in g} d(i, o)$ for any group g. Since $n = \sum_g n_g$, by adding all these inequalities together, we have

$$n \cdot \text{cost}(o) \geq \sum_{g \in G} \sum_{i \in g} d(i, o). \tag{1}$$

By the triangle inequality, we have

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, w) \\
&\leq \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, o) + \frac{1}{n_{g_w}} \sum_{i \in g_w} d(w, o) \\
&\leq \text{cost}(o) + d(w, o).
\end{aligned} \tag{2}$$

By Lemma 4.2 with $x = w$ and $y = o$, and using (1), we have

$$d(w, o) \leq \frac{4}{n} \cdot \sum_{g \in G} \sum_{i \in g} d(i, o) \leq 4 \cdot \text{cost}(o),$$

which, combined with (2), leads to

$$\text{cost}(w) \leq 5 \cdot \text{cost}(o),$$

which directly implies the desired upper bound. □

4.2 Avg-of-Max

For the Avg-of-Max cost, we first consider the case of symmetric groups, in which $n_g = \lambda$ for every g, and show a tight bound of 5 on the distortion of ordinal-information group-oblivious mechanisms.

Theorem 4.4. *For Avg-of-max, the distortion of any ordinal-information group-oblivious mechanism is at least* 5*, even when there are only two alternatives and the groups are symmetric.*

Proof. Let $\varepsilon > 0$ be any constant and $\lambda \geq 2$ be an integer such that $\lambda > \frac{4}{\varepsilon}$. Consider an instance with $n = 2\lambda(\lambda - 1)$ agents and two alternatives $\{a, b\}$; clearly, n is even. Half of the agents prefer a while the remaining half prefer b. With this information, any of the two alternatives can be chosen as the winner, so assume that the winner is a. The agents might be split into $k = 2(\lambda - 1)$ groups of λ agents each as follows:

- There are λ groups, each consisting of $\lambda - 1$ agents that prefer a and one agent that prefers b;
- There are $\lambda - 2$ groups, each consisting of λ agents that prefer b.

Further, consider the metric space being the line of real numbers and the positioning of the alternatives and the agents being as follows:

- a is at 0 and b is at 2;
- All agents that prefer a are at $1-\varepsilon/10$;
- The λ agents that prefer b and are part of the first λ groups (in which there are agents that prefer a) are at 3;
- The remaining $\lambda(\lambda-2)$ agents that prefer b are at 2.

We have that

$$k \cdot \text{cost}(a) = \lambda \cdot 3 + (\lambda-2)\cdot 2 = 5\lambda - 2$$

and

$$k \cdot \text{cost}(b) = \lambda \cdot (1+\varepsilon/10) + (\lambda-2)\cdot 0 = \lambda(1+\varepsilon/10),$$

leading to a distortion of at least $\frac{5}{1+\varepsilon/10} - \frac{2}{\lambda(1+\varepsilon/10)} > 5 - \frac{\varepsilon}{2} - \frac{2}{\lambda} > 5-\varepsilon$, where the first inequality is just a matter of simple calculations. ☐

For the upper bound, we consider again mechanisms that output alternatives whose domination graphs have perfect matchings, and show an upper bound of 5 with a proof similar to the one used for the Max-of-Avg objective.

Theorem 4.5. *For Avg-of-Max and symmetric groups, the distortion of a mechanism that returns an alternative whose domination graph has a perfect matching is at most* 5.

Proof. Consider any instance with k symmetric groups, each consisting of $\lambda = n/k$ agents. Let w be an alternative whose domination graph has a perfect matching, and o an optimal alternative. For every group g, let i_g and i_g^* be most-distant agents from w and o, respectively. Clearly,

$$\text{cost}(o) = \frac{1}{k} \cdot \sum_{g \in G} d(i_g^*, o) \geq \frac{1}{k} \cdot \sum_{g \in G} d(i_g, o).$$

By the triangle inequality, we have

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{k} \cdot \sum_{g \in G} d(i_g, w) \\
&\leq \frac{1}{k} \cdot \sum_{g \in G} d(i_g, o) + \frac{1}{k} \cdot \sum_{g \in G} d(w, o) \\
&\leq \text{cost}(o) + d(w, o) \qquad (3)
\end{aligned}$$

By Lemma 4.2 with $x = w$ and $y = o$, and since $k = n/\lambda$, we have

$$\begin{aligned}
d(w, o) &\leq \frac{4}{n} \cdot \sum_{g \in G} \sum_{i \in g} d(i, o) \\
&\leq \frac{4}{n} \cdot \sum_{g \in G} \lambda \cdot \max_{i \in g} d(i, o) \\
&= \frac{4}{k} \sum_{g \in G} d(i_g^*, o)
\end{aligned}$$

$$= 4 \cdot \text{cost}(o)$$

Using this, (3) becomes

$$\text{cost}(w) \leq 5 \cdot \text{cost}(o),$$

giving us the desired bound of 5 on the distortion. $\square$

For general instances with asymmetric groups, we show a tight bound of $2k+1$.

Theorem 4.6. *For Avg-of-Max, the distortion of any ordinal-information group-oblivious mechanism is at least* $2k+1$*, even when there are only two alternatives.*

Proof. Consider the following instance with $n = 2k$ agents and two alternatives located on the line of real numbers:

- Alternative a is at 0 and alternative b is at 2;
- There k agents that prefer alternative a and k agents that prefer alternative b.

Since there is no way of distinguish between the two alternative given the preferences of the agents, we may assume that the winner is a, without loss of generality. The agents might be partitioned into the following k groups:

- The first group consists of $k+1$ agents that includes those that prefer a who are located at 1 and one agent that prefers b who is located at 3;
- Each of the remaining $k-1$ groups consist of just one agent that prefers b who is located at 2.

Hence, $k \cdot \text{cost}(a) = 3 + (k-1) \cdot 2 = 2k+1$ and $k \cdot \text{cost}(b) = 1$, leading to a distortion of $2k+1$. $\square$

The matching upper bound follows easily by choosing any alternative who is ranked first by some agent.

Theorem 4.7. *For Avg-of-Max, the distortion of a mechanism that returns an alternative who is the most-preferred of some agent is at most* $2k+1$*.*

Proof. For any group g, let i_g and i_g^* be agents that are most-distant from the winner w and the optimal alternative o, respectively. Let S be the set of groups in which there is at least one agent with w as her most-preferred alternative, and observe that $|S| \geq 1$, and thus $|G \setminus S| \leq k-1$. We make the following observations:

- For any group $g \in S$, let j_g be an agent who ranks w first. By the triangle inequality, for any $g \in S$, we have that

  $$d(i_g, w) \leq d(i_g, o) + d(j_g, o) + d(j_g, w) \leq d(i_g, o) + 2d(j_g, o) \leq 3 \cdot d(i_g^*, o).$$

  In addition, since there is agent j_g that prefers w over o, then

  $$\frac{1}{2} \cdot d(w, o) \leq \frac{1}{2}\big(d(j_g, w) + d(j_g, o)\big) \leq d(j_g, o) \leq d(i_g^*, o).$$

- For any group $g \notin S$, by the triangle inequality, we have that

$$d(i_g, w) \leq d(i_g, o) + d(w, o) \leq d(i_g^*, o) + d(w, o).$$

Also $d(i_g^*, o) \geq 0$.

Using these, we can now bound the distortion as follows:

$$\begin{aligned}
\frac{\text{cost}(w)}{\text{cost}(o)} &= \frac{\sum_{g \in S} d(i_g, w) + \sum_{g \notin S} d(i_g, w)}{\sum_{g \in G} d(i_g^*, o)} \\
&\leq \frac{3 \sum_{g \in S} d(i_g^*, o) + \sum_{g \notin S} \big(d(i_g^*, o) + d(w, o)\big)}{\sum_{g \in G} d(i_g^*, o)} \\
&\leq 3 + \frac{\sum_{g \notin S} d(w, o)}{\sum_{g \in S} d(i_g^*, o)} \\
&\leq 3 + \frac{|G \setminus S| \cdot d(w, o)}{|S| \cdot \frac{1}{2} \cdot d(w, o)} \\
&\leq 3 + 2(k-1) = 2k + 1,
\end{aligned}$$

as desired. □

5 Group-Aware Mechanisms

In the previous two sections, we focused on mechanisms that are oblivious to the partition of the agents into groups. It is thus natural for one to wonder whether improved distortion bounds can be achieved by mechanisms that are *aware* of the groups. Clearly, we can optimize exactly both objectives if we are also given full information about the locations of the agents and the alternatives in the metric space, so this question makes sense when we only have access to partial information about the metric space, such as ordinal information. In this section, we consider such group-aware mechanisms and show tight bounds on the distortion in two cases: (1) there are only two alternatives; (2) there are $m \geq 2$ alternatives and the distances between them are known.

5.1 The Case of Two Alternatives

Here, we consider the case of two alternative a and b. For both objectives (Max-of-Avg and Avg-of-Max), we show a tight bound of 3 on the distortion of ordinal-information mechanisms. We start with the lower bounds, which are implied by the classic voting setting without groups.

Theorem 5.1. *For both Max-of-Avg and Avg-of-Max, the distortion of any ordinal-information group-aware mechanism is at least* 3*, even when there are only two alternatives and the groups are symmetric.*

Proof. The lower bounds for both objectives follow by considering instances in which the agents are partitioned into singleton groups. Then, the Max-of-Avg objective reduces to the egalitarian cost (the maximum distance over all agents), while the Avg-of-Max objectives reduces to the average social cost (the average total distance of the agents). When there are no groups (or, equivalently, there are singleton groups), the best possible distortion in terms of the egalitarian or the average social cost is 3, even where there are only two alternatives [Anshelevich et al., 2018, Gkatzelis et al., 2020, Kizilkaya and Kempe, 2022]. □

Next, we present the tight upper bounds. For the Max-of-Avg objective, we consider the Group-Proportional-Majority mechanism which chooses the winner w to be an alternative that has the largest proportional majority within any group. In particular, for any alternative $x \in \{a, b\}$, let $n_g(x)$ be the number of agents in group g that prefer x. Then,

$$w \in \arg\max_{x \in \{a,b\}} \max_{g \in G} \frac{n_g(x)}{n_g}.$$

Theorem 5.2. *For Max-of-Avg and two alternatives, the distortion of* Group-Proportional-Majority *is at most* 3.

Proof. For any group g, let $S_g(x)$ be the set of agents in g that prefer x; thus, $n_g(x) = |S_g(x)|$. By the definition of the mechanism, there is a group γ such that $\frac{n_\gamma(w)}{n_\gamma} \geq \frac{n_g(o)}{n_g}$ for every group g. Clearly, for any agent $i \in S_\gamma(w)$, $d(i, w) \leq d(i, o)$, and thus, by the triangle inequality, $d(i, o) \geq d(w, o)/2$. Using this, for any group g, we can bound the optimal cost as follows:

$$\text{cost}(o) \geq \frac{1}{n_\gamma} \sum_{i \in \gamma} d(i, o) \geq \frac{1}{n_\gamma} \sum_{i \in S_\gamma(w)} d(i, o) \geq \frac{n_\gamma(w)}{n_\gamma} \cdot \frac{d(w, o)}{2} \geq \frac{n_g(o)}{n_g} \cdot \frac{d(w, o)}{2}$$

or, equivalently,

$$\frac{n_g(o)}{n_g} \cdot d(w, o) \leq 2 \cdot \text{cost}(o). \tag{4}$$

Now, let g_w be the group that determines the cost of w. Using the fact that $d(i, w) \leq d(i, o)$ for every agent $i \in S_{g_w}(w)$ and the triangle inequality, we have

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, w) \\
&= \frac{1}{n_{g_w}} \sum_{i \in S_{g_w}(w)} d(i, w) + \frac{1}{n_{g_w}} \sum_{i \in S_{g_w}(o)} d(i, w) \\
&\leq \frac{1}{n_{g_w}} \sum_{i \in S_{g_w}(w)} d(i, o) + \frac{1}{n_{g_w}} \sum_{i \in S_{g_w}(o)} (d(i, o) + d(w, o)) \\
&\leq \text{cost}(o) + \frac{n_{g_w}(o)}{n_{g_w}} \cdot d(w, o).
\end{aligned}$$

Using (4) for $g = g_w$, we finally obtain $\text{cost}(w) \leq 3 \cdot \text{cost}(o)$, as desired. □

For Avg-of-Max, we consider the Group-Score mechanism which, for any alternative $x \in \{a, b\}$, assigns 2 points to x for any group in which all agents prefer x, and 1 point for any group in which some agents prefer x while the remaining agents prefer the other alternative. The winner w is the alternative with maximum score, breaking possible ties arbitrarily.

Theorem 5.3. *For Avg-of-Max and two alternatives, the distortion of* Group-Score *is at most* 3.

Proof. Let w be the alternative chosen by the mechanism, and o an optimal alternative; clearly, if $w = o$, the distortion is 1, so we assume that $w \neq o$. We partition the groups into three sets:

- S_w contains the groups that are *in favor of* w, in which all agents prefer w over o;
- S_o contains the groups that are *in favor of* o, in which all agents prefer o over w;

- S_m contains the groups that are *mixed*, in which some agents prefer w over o and some agents prefer o over w.

For any group g, let i_g be a most-distant agent from w and i_g^* a most-distant agent from o; hence, $\text{cost}(w) = \frac{1}{k}\sum_g d(i_g, w)$ and $\text{cost}(o) = \frac{1}{k}\sum_g d(i_g^*, o)$. We make the following observations:

- For any $g \in S_w$, both i_g and i_g^* prefer w over o. Hence, $d(i_g, w) \leq d(i_g, o) \leq d(i_g^*, o)$ and, using the triangle inequality, $d(i_g^*, o) \geq \frac{1}{2} \cdot d(w, o)$.
- For any $g \in S_o$, by the triangle inequality, $d(i_g, w) \leq d(i_g, o) + d(w, o) \leq d(i_g^*, o) + d(w, o)$. Also, recall that $d(i_g^*, o) \geq 0$.
- For any $g \in S_m$, like above, $d(i_g, w) \leq d(i_g^*, o) + d(w, o)$. Also, since there is at least one agent that prefers w over o, it must be the case that $d(i_g^*, o) \geq \frac{1}{2} \cdot d(w, o)$.

Using first the upper bounds on the distances from w, and then the lower bounds on the distances from o, we can write the distortion as follows:

$$\begin{aligned}
\frac{\text{cost}(w)}{\text{cost}(o)} &= \frac{\sum_g d(i_g, w)}{\sum_g d(i_g^*, o)} \\
&\leq \frac{\sum_g d(i_g^*, o) + (|S_o| + |S_m|) \cdot d(w, o)}{\sum_g d(i_g^*, o)} \\
&= 1 + \frac{(|S_o| + |S_m|) \cdot d(w, o)}{\sum_g d(i_g^*, o)} \\
&\leq 1 + \frac{(|S_o| + |S_m|) \cdot d(w, o)}{(|S_w| + |S_m|) \cdot \frac{1}{2} \cdot d(w, o)} \\
&= 1 + 2 \cdot \frac{|S_o| + |S_m|}{|S_w| + |S_m|}.
\end{aligned}$$

By the definition of the mechanism, w is chosen as the winner because $2|S_w| + |S_m| \geq 2|S_o| + |S_m|$ or, equivalently, $|S_w| \geq |S_o|$. Using this, the distortion is at most

$$1 + 2 \cdot \frac{|S_o| + |S_m|}{|S_w| + |S_m|} \leq 1 + 2 \cdot \frac{|S_o| + |S_m|}{|S_o| + |S_m|} = 3,$$

as claimed. ☐

5.2 Known Distances between Alternatives

We finally consider the general case of $m \geq 2$ but when slightly more information than just ordinal preferences is available. In particular, besides knowing the ordinal preferences of the agents over the alternatives, we assume that the distances between the alternatives in the metric space are also known. This is a natural assumption in various important applications (such as in facility location problems) and it has thus been examined in previous work on the distortion for different voting settings [Anshelevich and Zhu, 2021, Anshelevich et al., 2024]. Before we continue, we remark that the lower bound of 3, and even the lower bounds in the previous sections, still hold for this setting where the distances between the alternatives are known since they have been proven using instances with just two alternatives.

To show a tight bound of 3 for the two objectives, we consider mechanisms that virtually map each agent i to its most-preferred alternative $\text{top}(i)$, and then choose the winner to be an alternative that minimizes the objective under consideration for these most-preferred alternatives. In particular, the winner for the Max-of-Avg objective is

$$w \in \arg\max_{x \in A} \max_{g \in G} \left(\frac{1}{n_g} \sum_{i \in g} d(\text{top}(i), x) \right),$$

while the winner for the Avg-of-Max objective is

$$w \in \arg\max_{x \in A} \left(\frac{1}{k} \sum_{g \in G} \max_{i \in g} d(\text{top}(i), x) \right).$$

We will refer to these two mechanisms as VIRTUAL-MINIMAX-OF-AVG and VIRTUAL-MINIAVG-OF-MAX, respectively.

Theorem 5.4. *When the alternative locations are known, the distortion of* VIRTUAL-MINIMAX-OF-AVG *is at most* 3 *for Max-of-Avg, and the distortion of* VIRTUAL-MINIAVG-OF-MAX *is at most* 3 *for Avg-of-Max.*

Proof. We first show the bound for the Max-of-Avg objective. Let w be the alternative chosen by the VIRTUAL-MINIMAX-OF-AVG mechanism, and denote by o an optimal alternative. By definition, $\text{cost}(o) \geq \max_{g \in G} \left(\frac{1}{n_g} \sum_{i \in g} d(i, o) \right)$. Let g_w be the group that determines the cost of w. By the triangle inequality, the fact that $d(i, \text{top}(i)) \leq d(i, o)$ for any agent i, the definition of w (which minimizes the Max-of-Avg cost of the most-preferred alternatives of all agents), and the fact that the maximum of a set of additive functions is subadditive, we obtain

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, w) \\
&\leq \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, \text{top}(i)) + \frac{1}{n_{g_w}} \sum_{i \in g_w} d(\text{top}(i), w) \\
&\leq \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, o) + \max_{g \in G} \left(\frac{1}{n_g} \sum_{i \in g} d(\text{top}(i), w) \right) \\
&\leq \text{cost}(o) + \max_{g \in G} \left(\frac{1}{n_g} \sum_{i \in g} d(\text{top}(i), o) \right) \\
&\leq \text{cost}(o) + \max_{g \in G} \left(\frac{1}{n_g} \sum_{i \in g} d(i, \text{top}(i)) \right) + \max_{g \in G} \left(\frac{1}{n_g} \sum_{i \in g} d(i, o) \right) \\
&\leq 3 \cdot \text{cost}(o).
\end{aligned}$$

The proof for the Max-of-Avg objective is quite similar. Now let w be the alternative chosen by the VIRTUAL-MINIAVG-OF-MAX mechanism. For the optimal alternative o, by definition, we have $\text{cost}(o) \geq \frac{1}{k} \sum_{g \in G} \max_{i \in g} d(i, o)$. Let i_g be the most-distant agent from w in group g. Again, using the triangle inequality, the fact that $d(i, \text{top}(i)) \leq d(i, o)$ for any agent i, the definition of w (which now minimizes the Avg-of-Max cost of the most-preferred alternative of all agents), and the fact that max is a subadditive function, we obtain

$$\begin{aligned}
\text{cost}(w) &= \frac{1}{k} \sum_{g \in G} d(i_g, w) \\
&\leq \frac{1}{k} \sum_{g \in G} d(i_g, \text{top}(i_g)) + \frac{1}{k} \sum_{g \in G} d(\text{top}(i_g), w)
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{k}\sum_{g\in G} d(i_g, o) + \frac{1}{k}\sum_{g\in G}\max_{i\in g} d(\text{top}(i), w)\\
&\leq \frac{1}{k}\sum_{g\in G}\max_{i\in g} d(i, o) + \frac{1}{k}\sum_{g\in G}\max_{i\in g} d(\text{top}(i), o)\\
&\leq \text{cost}(o) + \frac{1}{k}\sum_{g\in G}\max_{i\in g}\big(d(i,\text{top}(i)) + d(i,o)\big)\\
&\leq \text{cost}(o) + \frac{1}{k}\sum_{g\in G}\max_{i\in g}\big(2\cdot d(i,o)\big)\\
&\leq 3\cdot\text{cost}(o),
\end{aligned}
$$

as claimed. □

6 Conclusion and Open problems

In this paper, we considered a metric voting setting in which the agents are partitioned into groups. When the groups are unknown, we showed tight bounds on the distortion of oblivious full-information and oblivious ordinal-information mechanisms in terms of two objectives that take the groups into account, the Max-of-Avg and the Avg-of-Max objectives. On the other hand, when the groups are known, we managed to show tight bounds on the distortion of group-aware ordinal mechanisms when there are just two alternatives or when we also have access to the locations of the alternatives in the metric space.

There are multiple avenues for further research in the group voting model we considered here. The most important problem that our work leaves open is to resolve the distortion of group-aware ordinal mechanisms for more than two alternatives. While this a very challenging task in general, we remark that achieving constant distortion can be done by using the two-step distributed mechanisms of Anshelevich et al. [2022] which are, by definition, group-aware. However, those mechanisms do not fully exploit the structure of the groups, and we therefore expect that better distortion bounds can be achieved by unlocking the full potential of group-aware mechanisms. Other interesting directions would be to consider randomized mechanisms and other objective functions that take the groups into account, beyond Max-of-Avg and Avg-of-Max.

References

Ben Abramowitz and Elliot Anshelevich. Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 894–901, 2018.

Ben Abramowitz, Elliot Anshelevich, and Wennan Zhu. Awareness of voter passion greatly improves the distortion of metric social choice. In *Proceedings of the 15th International Conference Web and Internet Economics (WINE)*, pages 3–16, 2019.

Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. Peeking behind the ordinal curtain: Improving distortion via cardinal queries. *Artificial Intelligence*, 296: 103488, 2021.

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

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

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

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

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

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

Elliot Anshelevich, Aris Filos-Ratsikas, and Alexandros A. Voudouris. The distortion of distributed metric social choice. *Artificial Intelligence*, 308:103713, 2022.

Elliot Anshelevich, Aris Filos-Ratsikas, Christopher Jerrett, and Alexandros A. Voudouris. Improved metric distortion via threshold approvals. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, pages 9460–9468, 2024.

Gerdus Benadè, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. Preference elicitation for participatory budgeting. *Management Science*, 67(5):2813–2827, 2021.

Umang Bhaskar, Varsha Dani, and Abheek Ghosh. Truthful and near-optimal mechanisms for welfare maximization in multi-winner elections. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence, (AAAI)*, pages 925–932, 2018.

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.

Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors. *Handbook of Computational Social Choice.* Cambridge University Press, 2016.

Jakob Burkhardt, Ioannis Caragiannis, Karl Fehrs, Matteo Russo, Chris Schwiegelshohn, and Sudarshan Shyam. Low-distortion clustering with ordinal and limited cardinal information. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, pages 9555–9563, 2024.

Ioannis Caragiannis and Karl Fehrs. Beyond the worst case: Distortion in impartial culture electorate. *CoRR*, abs/2307.07350, 2023.

Ioannis Caragiannis, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152, 2017.

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

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

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

Soroush Ebadian, Anson Kahng, Dominik Peters, and Nisarg Shah. Optimized distortion and proportional fairness in voting. In *Proceedings of the 23rd ACM Conference on Economics and Computation (EC)*, pages 523–600, 2022.

Soroush Ebadian, Aris Filos-Ratsikas, Mohamad Latifian, and Nisarg Shah. Explainable and efficient randomized voting rules. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2023a.

Soroush Ebadian, Mohamad Latifian, and Nisarg Shah. The distortion of approval voting with runoff. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1752–1760, 2023b.

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

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

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

Vasilis Gkatzelis, Mohamad Latifian, and Nisarg Shah. Best of both distortion worlds. In *Proceedings of the 24th ACM Conference on Economics and Computation, (EC)*, pages 738–758, 2023.

Michal Jaworski and Piotr Skowron. Evaluating committees for representative democracies: the distortion and beyond. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 196–202, 2020.

Anson Kahng, Mohamad Latifian, and Nisarg Shah. Voting with preference intensities. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5697–5704, 2023.

David Kempe. Communication, distortion, and randomness in metric voting. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*, pages 2087–2094, 2020.

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

Mohamad Latifian and Alexandros A. Voudouris. The distortion of threshold approval matching. *CoRR*, abs/2401.09858, 2024.

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

Debmalya Mandal, Ariel D. Procaccia, Nisarg Shah, and David P. Woodruff. Efficient and thrifty voting by any means necessary. In *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS)*, pages 7178–7189, 2019.

Debmalya Mandal, Nisarg Shah, and David P Woodruff. Optimal communication-distortion tradeoff in voting. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*, pages 795–813, 2020.

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

Alexandros A. Voudouris. Tight distortion bounds for distributed metric voting on a line. *Operations Research Letters*, 51(3):266–269, 2023.

Houyu Zhou, Minming Li, and Hau Chan. Strategyproof mechanisms for group-fair facility location problems. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 613–619, 2022.

Houyu Zhou, Hau Chan, and Minming Li. Altruism in facility location problems. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, pages 9993–10001, 2024.

A Refined Analysis of the Full-Information Group-Oblivious Mechanism for Max-of-Avg

In Theorem 3.2, we showed that choosing any alternative that minimizes the total distance from all agents achieved the best possible distortion of 3 in terms of the Max-of-Avg objective when taking the worst-case over all possible instances. In this appendix, we present a more detailed analysis of this mechanism and show that the distortion bound is $3 - 2\mu/n$, where μ is the smallest group size and n is the number of agents. While this is still 3 in the worst-case, it implies some improved upper bounds for cases in which the number of groups k is small or the smallest group size μ is rather large compared to n. In particular, for instances where there are k symmetric groups, the bound becomes $3 - 2/k$.

Theorem A.1. *For Max-of-Avg, the distortion of any alternative that minimizes the total distance from all agents is at most $3 - 2\mu/n$, where μ is the smallest group size and n is the number of agents.*

Proof. Let w be the alternative that minimizes the total distance from all agents. Suppose that $d(w, o) = 2$ without loss of generality. Let g_w be the group that determines the cost of w (that is, g_w is the group for which the average distance of the agents therein from w is maximized). We define z_w and z_o to be the average distance of all agents that do not belong to g_w from w and o, respectively, that is

$$z_w := \frac{1}{n - n_{g_w}} \sum_{g \neq g_w} \sum_{i \in g} d(i, w)$$

and

$$z_o := \frac{1}{n - n_{g_w}} \sum_{g \neq g_w} \sum_{i \in g} d(i, o).$$

Also, let $y = \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, o)$ and $\gamma = \text{cost}(w) = \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, w)$. Using this notation, we have

$$\sum_i d(i, w) = n_{g_w} \cdot \gamma + (n - n_{g_w}) \cdot z_w$$

and

$$\sum_i d(i, o) = n_{g_w} \cdot y + (n - n_{g_w}) \cdot z_o.$$

By applying the triangle inequality, and using our assumption that $d(w, o) = 2$, we obtain

$$\gamma \leq \frac{1}{n_{g_w}} \sum_{i \in g_w} (d(i, o) + d(o, w)) = 2 + y.$$

We also have that $y \leq \gamma$; otherwise, since $\gamma = \text{cost}(w)$ and $y \leq \text{cost}(o)$, the distortion would be 1. This implies that $\sum_{i \in g_w} d(i, w) \geq \sum_{i \in g_w} d(i, o)$. Also, by the definition of w (which minimizes the total distance from all agents), $\sum_i d(i, w) \leq \sum_i d(i, o)$. By these, we can conclude that $z_w \leq z_o$; otherwise the total distance from o would be strictly smaller than w, and w would not be the winner. So, there must exist $C, x \geq 0$ such that $z_w = C - x$ and $z_o = C + x$. In fact, using the triangle inequality,

$$2 = d(w, o) = \frac{1}{n - n_{g_w}} \sum_{g \neq g_w} \sum_{i \in g} d(w, o) \leq z_w + z_o = 2C$$

and thus $C \geq 1$.

We now introduce a lemma that characterizes the worst-case instances we need to focus on.

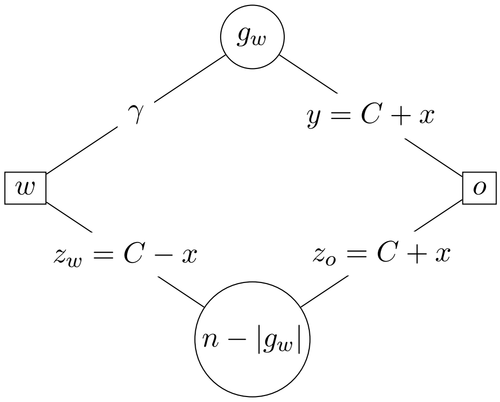

Figure 1: An illustration of the worst-case instance given by Lemma A.2.

Lemma A.2. *In a worst-case instance (in terms of distortion), $y = z_o = C + x$.*

Proof. The proof proceeds in two steps: We first transform the original instance I into a new instance I' in which there are only a few distinct points in the metric space where agents and alternatives are located. Then, we transform I' into another instance I'' with the desired property $y = C + x$. While doing these transformations we will show that the cost of o does not increase, that is,

$$\text{cost}(o \mid I) \geq \text{cost}(o \mid I') \geq \text{cost}(o \mid I''),$$

while the cost of w does not decrease, that is,

$$\text{cost}(w \mid I) \leq \text{cost}(w \mid I') \leq \text{cost}(w \mid I'').$$

We now introduce the first transformation. Consider the following new instance I' with two alternatives w and o:

- There is a group consisting of n_{g_w} agents all of whom are located at the same point with distance γ from w and y from o.
- There are also $k - 1$ groups with sizes equal to the sizes of the remaining groups in the original instance I. The $n - n_{g_w}$ agents in all those groups are located at the same point with distance $z_w = C - x$ from w and $z_o = C + x$ from o.

Observe that w still minimizes the total distance of all agents in I'; this follows since the total distance of the agents from w and o remain the same as in I. For the same reason, $\text{cost}(w \mid I') = \text{cost}(w \mid I)$.

We now argue that $\text{cost}(o \mid I') \leq \text{cost}(o \mid I)$. By the definition of y, the average distance of the agents in g_w from o is the same as in I. Consider any group $g_o \neq g_w$ that maximizes $\frac{1}{n_{g_o}} \sum_{i \in g_o} d(i, o)$. Since g_o maximizes the average distance out of all groups that are different than g_w, we have that

$$\frac{1}{n_{g_o}} \sum_{i \in g_o} d(i, o) \geq \frac{1}{n - n_{g_w}} \sum_{g \neq g_w} \sum_{i \in g} d(i, o) = z_o = C + x.$$

Observe now that

$$\text{cost}(o \mid I) = \max\left(\frac{1}{n_{g_o}} \sum_{i \in g_o} d(i, o), \frac{1}{n_{g_w}} \sum_{i \in g_w} d(i, o)\right) \geq \max\{C + x, y\} = \text{cost}(o \mid I'),$$

thus proving our claim.

Next, we transform I' into a new instance I'' with the desired property $C + x = y$ such that the distortion does not decrease. We consider the following two cases.

Case 1: $y \geq C + x$ in I'. Let $C' = y - x \geq C$ and consider the following instance I'':

- There is a group consisting of n_{g_w} agents all of whom are located at the same point with distance γ from w and y from o.
- There are also $k - 1$ groups with sizes equal to the sizes of the remaining groups in the original instance I. The $n - n_{g_w}$ agents in all those groups are located at the same point with distance $C' - x$ from w and $C' + x$ from o.

In this new instance w still minimizes the total distance from all agents; indeed, the location of any agent $i \in g_w$ is the same in both instances I' and I'', while any agent $i \notin g_w$ has been moved closer to w and o by exactly the same distance $C' - C = y - (x + C) \geq 0$ between the two instances. Observe that $\text{cost}(o \mid I'') = \text{cost}(o)$ since the average distance of any group from o is y in I'' and $\text{cost}(o \mid I') = \max\{C + x, y\} = y$ by our assumption for this case. In addition, the distances to w increase for some agents, and thus $\text{cost}(w \mid I'') \geq \text{cost}(w \mid I')$, which further means that the distortion does not decrease as we go from I' to I'' for which the desired property $y = C' + x$ holds.

Case 2: $y < C + x$ in I'. Let $y' = C + x$, $\gamma' = \gamma + (y' - y)$ and consider the following instance I'':

- There is a group consisting of n_{g_w} agents all of whom are located at the same point with distance γ' from w and y' from o.
- There are also $k - 1$ groups with sizes equal to the sizes of the remaining groups in the original instance I. The $n - n_{g_w}$ agents in all those groups are located at the same point with distance $C - x$ from w and $C + x$ from o.

As in the previous case, w still minimizes the total distance from all agents; indeed, the any agent $i \in g_w$ has been moved closer to w and o by the same distance $y' - y$ between the two instances, while any agent $i \notin g_w$ is at the same location in both instances. We also have that $\text{cost}(o \mid I'') = \text{cost}(o \mid I') = C + x$, and $\text{cost}(w \mid I'') \geq \text{cost}(w \mid I')$ since the distance to w has increased for some agents from $y < C + x$ to $y' = C + x$. Hence, the distortion again does not decrease, and the new instance again satisfies the desired property that $y' = C + x$. $\square$

By the above lemma, there exists a worst-case instance (where the ratio $\frac{\text{cost}(w)}{\text{cost}(o)}$ is maximized) with $C + x = y$. We consider the following two cases.

Case 1: $x \geq \frac{n_{g_w}}{n - n_{g_w}}$. Then, since $\text{cost}(w) = \gamma \leq 2 + y = 2 + C + x$ and $\text{cost}(o) = \max\{y, C + x\} = C + x$, the distortion is at most

$$\frac{\text{cost}(w)}{\text{cost}(o)} = \frac{2 + C + x}{C + x}.$$

This expression is a non-increasing function in terms of C and x. Hence, since $C \geq 1$ and $x \geq \frac{n_{g_w}}{n - n_{g_w}}$, we obtain an upper bound of

$$\frac{3 + \frac{n_{g_w}}{n - n_{g_w}}}{1 + \frac{n_{g_w}}{n - n_{g_w}}} = 3 - \frac{2n_{g_w}}{n} \leq 3 - \frac{2\mu}{n}.$$

Case 2: $x \leq \frac{n_{g_w}}{n-n_{g_w}}$. Since $z_w = C - x$, have that

$$\sum_i d(i,w) = n_{g_w} \cdot \gamma + (n - n_{g_w}) \cdot z_w = n_{g_w} \cdot \gamma + (n - n_{g_w}) \cdot (C - x).$$

Also, since $z_o = C + x = y$, we have that

$$\sum_i d(i,o) = n_{g_w} \cdot y + (n - n_{g_w}) \cdot z_o = n \cdot (C + x).$$

Using these and the definition of w, which is the alternative that minimizes the total distance from all agents, we further have that

$$\sum_i d(i,w) \leq \sum_i d(i,o) \Leftrightarrow \gamma \leq \frac{2nx}{n_{g_w}} + C - x.$$

Hence, the distortion in the worse case instance is at most,

$$\frac{\text{cost}(w)}{\text{cost}(o)} = \frac{\gamma}{\text{cost}(o)} \leq \frac{\frac{2nx}{n_{g_w}} + C - x}{C + x}$$

This expression is a non-decreasing function of x and a non-increasing function of C. Since $x \leq \frac{n_{g_w}}{n-n_{g_w}}$ and $C \geq 1$, we obtain an upper bound of

$$\frac{\frac{2n}{n-n_{g_w}} + 1 - \frac{n_{g_w}}{n-n_{g_w}}}{1 + \frac{n_{g_w}}{n-n_{g_w}}} = 3 - \frac{2n_{g_w}}{n} \leq 3 - \frac{2\mu}{n}.$$

The proof is now complete. □