

Improved Metric Distortion via Threshold Approvals

Elliot Anshelevich¹, Aris Filos-Ratsikas², Christopher Jerrett¹, Alexandros A. Voudouris³

¹Rensselaer Polytechnic Institute, USA

²University of Edinburgh, United Kingdom

³University of Essex, United Kingdom

eanshel@cs.rpi.edu, aris.filos-ratsikas@ed.ac.uk, jerrec@rpi.edu, alexandros.voudouris@essex.ac.uk

Abstract

We consider a social choice setting in which agents and alternatives are represented by points in a metric space, and the cost of an agent for an alternative is the distance between the corresponding points in the space. The goal is to choose a single alternative to (approximately) minimize the social cost (total cost of all agents) or the maximum cost of any agent, when only limited information about the preferences of the agents is given. Previous work has shown that the best possible distortion one can hope to achieve is 3 when access to the ordinal preferences of the agents is given, even when the distances between alternatives in the metric space are known. We improve upon this bound of 3 by designing deterministic mechanisms that exploit limited cardinal information. We show that it is possible to achieve distortion $1 + \sqrt{2}$ by using the ordinal preferences of the agents, the distances between alternatives, and a threshold approval set per agent that contains all alternatives for whom her cost is within an appropriately chosen factor of her cost for her most-preferred alternative. We show that this bound is the best possible for any deterministic mechanism in general metric spaces, and also provide improved bounds for the fundamental case of the line metric.

1 Introduction

Social choice theory is concerned with aggregating the heterogeneous preferences of individuals over a set of outcomes into a single decision (Brandt et al. 2016). Besides its many applications in traditional domains, such as political elections or voting for policy issues, social choice theory has also been at the epicenter of areas such as *multi-agent systems* (Ephrati and Rosenschein 1991), *recommendation systems* (Ghosh et al. 1999), *search engines* (Dwork et al. 2001), and *crowdsourcing* applications (Mao, Procaccia, and Chen 2012). Aggregation methods inspired from social choice theory have also been used in relation to machine learning, such as for *regression* and *estimation* tasks (Caragiannis, Procaccia, and Shah 2016; Chen et al. 2018; Chen, Shen, and Zheng 2020; Kahng, Kehne, and Procaccia 2020), as well as *virtual democracy* (Noothigattu et al. 2018; Kahng et al. 2019; Peters et al. 2020).

A central theme underpinning many of these applications is the interplay between the amount of available information

and the efficiency of the implemented decisions (Xia 2022; Zhao and Xia 2019; Mandal et al. 2019; Amanatidis et al. 2022a; Boutilier et al. 2015). Indeed, it is often the case that access to the preferences of the participants (or *agents*) is restricted due to lack of enough statistical data, or due to computational or cognitive limitations in the elicitation process. This leads to the natural question of whether we can design aggregation rules (or *mechanisms*) that achieve high efficiency even when presented with informational restrictions.

This is the main question studied in the literature of *distortion* (Procaccia and Rosenschein 2006; Anshelevich et al. 2021), which measures the deterioration of an aggregate social objective due to limited information about the preferences of the agents. In its original setup (e.g., see (Boutilier et al. 2015)), limited information is interpreted as having access only to the *ordinal preference rankings* over the possible outcomes, instead of a complete *cardinal utility* structure that fully captures the intensity of the preferences. Over the years, the notion of distortion has been refined to capture different kinds of informational limitations, including restricted ordinal information (Gross, Anshelevich, and Xia 2017; Kempe 2020), communication complexity (Mandal et al. 2019; Mandal, Shah, and Woodruff 2020), and query complexity (Amanatidis et al. 2021, 2022a).

The literature on *metric distortion* considers scenarios where the agents and the outcomes exist in a (possibly high-dimensional) metric space, and the costs (rather than utilities) are given by distances that satisfy the triangle inequality. The economic interpretation of these distances is typically in terms of the proximity for political or ideological issues along different axes (e.g., liberal to conservative, or libertarian to authoritarian). The metric distortion of social choice mechanisms, firstly studied in the works of Anshelevich and Postl (2017) and Anshelevich et al. (2018), is one of the most well-studied settings in the context of this literature, with many variants of the main setting being considered in recent years (see Section 3 of the survey of Anshelevich et al. (2021) for an overview). Since its inception in 2015 (the conference version of (Anshelevich et al. 2018)), the “holy grail” of this literature was a mechanism with a metric distortion of 3 for the *social cost objective* (i.e., the sum of costs of all agents), which would match the corresponding lower bound shown by Anshelevich et al. (2018). The up-

per bound was finally established by Gkatzelis, Halpern, and Shah (2020) via the PLURALITY MATCHING mechanism, and then later also by Kizilkaya and Kempe (2022) via the much simpler PLURALITY VETO mechanism. For the other most natural objective, that of minimizing the *maximum cost* of any agent, the best possible bound can also easily be seen to be 3.

Crucially, these tight bounds of 3 are achieved by mechanisms without access to any information of a cardinal nature, i.e., they work by using only the ordinal preference rankings of the agents. It can also be observed (e.g., see (Anshelevich and Zhu 2021)) that if one also assumes access to the distances between the outcomes in the metric space, then 3 is still the best bound one can hope for. At the same time, a growing recent literature advocates to take the natural next step, and study the distortion when a small amount of cardinal information about the preferences is also available, which, in many cases, is reasonable to elicit (Abramowitz, Anshelevich, and Zhu 2019; Amanatidis et al. 2021, 2022b,a; Benade et al. 2021; Bhaskar, Dani, and Ghosh 2018). This motivates the following question:

Is it possible to beat the barrier of 3 on the metric distortion if we also have access to limited information of a cardinal nature?

1.1 Our Contribution

We consider a setting in which a set N of agents and a set A of alternatives are positioned on a metric space, and the preferences of the agents are captured by the distances between their positions and that of the different alternatives. We are interested in the distortion of deterministic mechanisms in terms of the social cost and the maximum cost objectives, i.e., the ratio of the cost of the alternative chosen by the mechanism over the smallest possible cost over all alternatives, taken over all possible instances. Note that with full cardinal information, achieving an optimal distortion of 1 is trivial.

We answer the main question posed above in the affirmative, by assuming access to the ordinal preferences, the distances between the alternatives in the metric space, and some additional limited cardinal information about the costs of the agents. In particular, for each agent $i \in N$, we have access to an α -threshold approval set (α -TAS) A_i , which contains all the alternatives for which the agent has cost at most a factor α times the cost for her most-preferred alternative, with α being a value that can be chosen by the mechanism designer. For general metric spaces, we design mechanisms that achieve a distortion of $1 + \sqrt{2}$ in terms of the social cost and the maximum cost objectives, thus beating the aforementioned barrier of 3. We complement these results with lower bounds on the distortion of *any* deterministic mechanism that uses this amount of information.

For the fundamental special case of a line metric, we provide refined tight bounds on the distortion for both objectives, showing an even larger separation from the bound of 3 that holds in the absence of any cardinal information and even when the metric is a line. More generally, we fine-tune our analysis to the presence/absence of all three types of in-

formation (i.e., only α -TAS, α -TAS + ordinal preferences, α -TAS + known alternative distances, or α -TAS + ordinal preferences + known alternative distances) and show that, in most cases, the best possible distortion bounds can be obtained even when only using two types of information. Our results are summarized in Table 1. Due to space constraints, some proofs and statements are deferred to the full version.

1.2 Related Work and Discussion

The distortion literature is rather extensive and has been exposed in the recent survey of Anshelevich et al. (2021). Below we discuss the works that are mostly relevant to the setting that we study here.

The term “approval threshold” (in fact, “approval threshold vote”) was introduced in the conference version of the paper of Benade et al. (2021) (and was later also used by Bhaskar, Dani, and Ghosh (2018)) to refer to an elicitation device that returns a set of alternatives that an agent values higher than a given threshold. Amanatidis et al. (2021) introduced a setting in which limited cardinal information is elicited via a set of *value* or *comparison queries*, where an agent is asked to provide their value for a given alternative, or indicate whether she prefers one alternative over another by more than a given factor, respectively. An approval threshold in their setting can be constructed via a single value query (for the agent’s most-preferred alternative), and then a number of comparison queries which is logarithmic in the number of alternatives.

The α -TAS that we consider are closely related to the comparison query model of Amanatidis et al. (2021), in that they do not encode information about any absolute value for the cost of the agents, but rather only information that is *relative* to their cost for their most-preferred alternatives. In this sense, these sets are *relative*, and can be viewed as the outcomes of a single relative threshold query, or, equivalently, the outcome of logarithmically-many comparison queries. From a cognitive standpoint, this is an even more conceivable elicitation device since the agents are only asked to perform (cardinal) comparisons rather than to come up with absolute cost numbers. Such comparisons are motivated by and rooted in the ideas of the Von Neumann-Morgenstern utility theory (Von Neumann and Morgenstern 2007); see also (Amanatidis et al. 2021).

A common characteristic of all aforementioned works is that they do not consider the distortion in the metric setting. Abramowitz, Anshelevich, and Zhu (2019) studied the metric distortion of mechanisms that use limited cardinal information, namely information about whether the ratio between the costs of an agent for any two candidates is larger or smaller than a chosen threshold τ . Their elicitation method is different from ours, and requires information about the relative costs for all pairs of alternatives. In contrast, our α -TAS only require aggregate information about the set A_i of alternatives. The two settings coincide only in the case where there are just 2 alternatives overall, which allows us to obtain a lower bound of $2\sqrt{2} - 1 \approx 1.83$ for the social cost for the line metric from their work, for which we also show a matching upper bound.

		TAS $\cap$ ORD $\cap$ DIST	TAS $\cap$ ORD	TAS $\cap$ DIST	TAS
Social cost	General	$[2, 1 + \sqrt{2}]$	$[1 + \sqrt{2}, 3]$	$1 + \sqrt{2}$	$\Theta(n)$
	Line	$2\sqrt{2} - 1$	2	$2\sqrt{2} - 1$	
Max cost	General	$1 + \sqrt{2}$	$[1 + \sqrt{2}, 3]$		3
	Line	$1 + \sqrt{2}$			

Table 1: An overview of our distortion bounds for deterministic mechanisms that use different types of information. TAS is used to refer to the class of mechanism that have access to the ordinal preferences of the agents, DIST is used for mechanisms that have access to the distances between alternatives, and TAS is used for the mechanism with access to the α -threshold approval sets.

2 Preliminaries

We consider a single-winner social choice setting with a set N of n agents and a set A of m alternatives. Agents and alternatives are represented by points in a metric space. For any $x, y \in N \cup A$, let $d(x, y)$ denote the distance between the points representing x and y . The distances satisfy the standard conditions for metric spaces, namely that $d(x, y) = d(y, x)$, $d(x, y) = 0$ if $x = y$, and $d(x, y) \leq d(x, z) + d(z, y)$; the last condition is called the *triangle inequality*. We will use the tuple $I = (N, A, d)$ to denote an *instance* of our setting, and will use $\mathcal{I}$ to denote the set of all instances.

Input information. A (deterministic) *mechanism* M takes as input some information related to the distances between agents and alternatives, and outputs a single alternative as the *winner*. We will consider combinations of the following three types of information:

- Inf.1 The *ordinal preferences* $(\succ_i)_{i \in N}$ of the agents which are induced by the distances, where $\succ_i$ is a complete ordering over the elements of A , and $x \succ_i y$ implies that $d(i, x) \leq d(i, y)$ for any $i \in N$ and $x, y \in A$.
- Inf.2 The distances $d(x, y)$ between any pair of alternatives $x, y \in A$.
- Inf.3 A set of α -threshold approval sets (α -TAS) $(A_i)_{i \in N}$ such that $d(i, x) \leq \alpha \cdot d(i, o_i)$ for any $x \in A_i$, where o_i is the most-preferred alternative of agent i (i.e., $o_i \succ_i y$ for any $y \in A$). We will say that agent i *approves* alternative x in case $x \in A_i$.

Clearly, all of Inf.1 to Inf.3 can be inferred from an instance I . We will use the term $\text{Inf}(I)$ to denote the information that is available to a mechanism for instance I . For ease of reference, we classify mechanisms by the type(s) of information that they use at inputs. We associate Inf.1 with ORD, Inf.2 with DIST, and Inf.3 with TAS. We say that a mechanism

- is in $X \in \{\text{ORD}, \text{DIST}, \text{TAS}\}$ if it has access only to the information associated with X ,
- is in $X \cap Y$ for some $X, Y \in \{\text{ORD}, \text{DIST}, \text{TAS}\}$ if it has access only to the information associated with X and Y , and
- is in $\text{ORD} \cap \text{DIST} \cap \text{TAS}$ if it has access to all three types of information.

As we already explained in the Introduction, the bulk of the metric distortion literature considers mechanisms that are only in ORD, which results in a best-possible distortion

of 3 (Anshelevich et al. 2018; Gkatzelis, Halpern, and Shah 2020; Kizilkaya and Kempe 2022) (see below for the formal definition). For mechanisms in $\text{ORD} \cap \text{DIST}$, the best possible distortion is still 3 (Anshelevich and Zhu 2021). Our main results are for the case of having access to all types of information (i.e., mechanisms in $\text{ORD} \cap \text{DIST} \cap \text{TAS}$), but we also provide results for mechanisms in $\text{ORD} \cap \text{TAS}$ and $\text{DIST} \cap \text{TAS}$ independently, as well as TAS in isolation.

Objectives and Distortion. We will consider the two well-known minimization objectives, the *social cost* SC and the *maximum cost* MC, which for an alternative x are defined as

$$\text{SC}(x) = \sum_{i \in N} d(i, x)$$

and

$$\text{MC}(x) = \max_{i \in N} d(i, x).$$

Given an objective $F \in \{\text{SC}, \text{MC}\}$, the *distortion* of a mechanism M for F is defined as

$$\sup_{I \in \mathcal{I}} \frac{F(M(\text{Inf}(I)))}{\min_{j \in A} F(j|I)},$$

where $M(\text{Inf}(I))$ denotes the alternative chosen by the mechanism when given as input the information $\text{Inf}(I)$ for instance I . Clearly, the distortion is at least 1 for any mechanism, and our goal is to determine the best possible distortion when given combinations of Inf.1 to Inf.3.

3 Social Cost

In this section, we show bounds on the distortion for the social cost objective for mechanisms in $\text{ORD} \cap \text{DIST} \cap \text{TAS}$. Our main result is a mechanism, coined $(1 + \sqrt{2})$ -MINISUM-TAS-DISTANCE, which achieves a distortion of $1 + \sqrt{2} \approx 2.42$, thus breaking the barrier of 3, which is the best possible without access to the α -TAS. In fact, this mechanism does not require access to the ordinal preferences of the agents, i.e., it is in $\text{DIST} \cap \text{TAS}$. We complement this result with a lower bound of 2 on the distortion of any deterministic mechanism $M \in \text{ORD} \cap \text{DIST} \cap \text{TAS}$.

For mechanisms in $\text{ORD} \cap \text{TAS}$ and $\text{DIST} \cap \text{TAS}$, we provide stronger lower bounds of $1 + \sqrt{2}$; this establishes that $(1 + \sqrt{2})$ -MINISUM-TAS-DISTANCE is the best possible among mechanisms in $\text{DIST} \cap \text{TAS}$. We also prove that the distortion of any mechanism in TAS is bound to be bad,

Mechanism 1: α -MINISUM-TAS-DISTANCE**Input:** Distances between alternatives, α -TAS;**Output:** Winner w ;

$$w \in \arg \min_{x \in A} \left\{ \sum_{i \in N} \min_{j \in A_i} d(j, x) \right\};$$

namely $\Theta(n)$. For the special case of the line metric, we provide more refined *tight* bounds depending on the amount of information available in the input.

3.1 Results for General Metric Spaces

We start by showing an upper bound of $1 + \sqrt{2}$ for general metric spaces using a mechanism to which we refer as α -MINISUM-TAS-DISTANCE. This mechanism is quite intuitive and does not require any information about the ordinal preferences, i.e., it is in the class $\text{DIST} \cap \text{TAS}$. The mechanism chooses as the winner an alternative that minimizes the sum of distances from the α -TAS of the agents, where the distance between an alternative x and a set A_i is defined as the distance of x from its the closest alternative in A_i . See Mechanism 1 for a description using pseudocode.

Theorem 3.1. *In general metric spaces, the distortion of α -MINISUM-TAS-DISTANCE for the social cost objective is at most $\max \left\{ \alpha, 2 + \frac{1}{\alpha} \right\}$.*

Proof. Recall that o_i is the most-preferred alternative of agent i ; hence, $d(i, j) \leq \alpha \cdot d(i, o_i)$ for any $j \in A_i$. Let o be the optimal alternative (i.e., the alternative with the smallest social cost) and let w be the alternative chosen by the mechanism. For any agent i , denote by $x_i \in A_i$ the alternative that is closest to w , and by $y_i \in A_i$ the alternative that is closest to o . By the definition of w , we have that $\sum_{i \in N} d(x_i, w) \leq \sum_{i \in N} d(y_i, o)$. Using the triangle inequality, we now have:

$$\begin{aligned} \text{SC}(w) &= \sum_{i \in N} d(i, w) \\ &\leq \sum_{i \in N} d(i, x_i) + \sum_{i \in N} d(x_i, w) \\ &\leq \sum_{i \in N} d(i, x_i) + \sum_{i \in N} d(y_i, o). \end{aligned}$$

We make the following observations:

- For every agent i such that $o \in A_i$, $y_i = o$, and thus $d(y_i, o) = 0$, as well as $d(i, o_i) \leq d(i, o)$.
- For every agent i such that $o \notin A_i$, $d(y_i, o) \leq d(o_i, o)$ and $d(i, o_i) < \frac{1}{\alpha} \cdot d(i, o)$.

Combining the above with the fact that $d(i, x_i) \leq \alpha \cdot d(i, o_i)$ and using the triangle inequality, we obtain

$$\begin{aligned} \text{SC}(w) &\leq \sum_{i \in N} d(i, x_i) + \sum_{i \in N} d(y_i, o) \\ &= \sum_{i: o \notin A_i} \left(d(i, x_i) + d(y_i, o) \right) \end{aligned}$$

$$\begin{aligned} &+ \sum_{i: o \notin A_i} \left(d(i, x_i) + d(y_i, o) \right) \\ &\leq \sum_{i: o \in A_i} d(i, x_i) + \sum_{i: o \notin A_i} \left(d(i, x_i) + d(o_i, o) \right) \\ &\leq \alpha \cdot \sum_{i: o \in A_i} d(i, o_i) \\ &+ \sum_{i: o \notin A_i} \left(\alpha \cdot d(i, o_i) + d(i, o_i) + d(i, o) \right) \\ &\leq \alpha \cdot \sum_{i: o \in A_i} d(i, o) + \left(2 + \frac{1}{\alpha} \right) \cdot \sum_{i: o \notin A_i} d(i, o) \\ &\leq \max \left\{ \alpha, 2 + \frac{1}{\alpha} \right\} \cdot \text{SC}(o). \end{aligned}$$

Hence, the distortion is at most $\max \left\{ \alpha, 2 + \frac{1}{\alpha} \right\}$. $\square$

By optimizing over α , we obtain the following corollary.

Corollary 3.2. *In general metric spaces, the distortion of $(1 + \sqrt{2})$ -MINISUM-TAS-DISTANCE for the social cost objective is at most $1 + \sqrt{2}$.*

We complement the upper bound of $1 + \sqrt{2}$ by a nearly tight lower bound of 2 for mechanisms that use all three types of information. We state the main theorem here and defer its proof and the details to the full version.

Theorem 3.3. *For the social cost objective, the distortion of any mechanism $M \in \text{ORD} \cap \text{DIST} \cap \text{TAS}$ is at least 2.*

Refined Bounds for General Metric Spaces

While the bounds of Theorem 3.1 and Theorem 3.3 are not matching, we are able to show lower bounds of $1 + \sqrt{2}$ for mechanisms either in $\text{ORD} \cap \text{TAS}$ or $\text{DIST} \cap \text{TAS}$. That shows that α -MINISUM-TAS-DISTANCE is the best possible among mechanisms in the former class. We state the lower bounds below.

Theorem 3.4. *For the social cost objective, the distortion of any mechanism M in either $\text{ORD} \cap \text{TAS}$ or $\text{DIST} \cap \text{TAS}$ is at least $1 + \sqrt{2}$.*

Refined Bounds for the Real Line

For the real line metric, we are able to obtain tight bounds for any subset of the different types of information. We summarize those results in the following theorem.

Theorem 3.5. *On the line metric, for the social cost objective, the best possible distortion for mechanisms in $\text{ORD} \cap \text{DIST} \cap \text{TAS}$ is $2\sqrt{2} - 1 \approx 1.83$, and this is achieved by a mechanism in $\text{ORD} \cap \text{TAS}$. The best possible distortion for mechanisms in $\text{DIST} \cap \text{TAS}$ is 2, and this is achieved by $(1 + \sqrt{2})$ -MINISUM-TAS-DISTANCE.*

The lower bound of $2\sqrt{2} - 1$ in Theorem 3.5 follows from the work of Abramowitz, Anshelevich, and Zhu (2019), via a connection between their setting and ours, when there are only two alternatives. For the upper bound, we propose a mechanism that we refer to as α -ELIMINATION-WEIGHTED-MAJORITY. It achieves a distortion of $2\sqrt{2} - 1$

as an upper bound for $\alpha = \sqrt{2} - 1$, and is thus best possible. In fact, our mechanism does not need to have access to the distances between alternatives, only to the α -TAS and the ordinal preferences of the agents. We describe the mechanism and the intuition behind its performance below.

Remark 1. Since the metric space is a line, we may assume that we have access to an ordering of the agents and alternatives on the line, which is unique up to permutations of identical agents and reversal; this ordering can be determined via the ordinal preferences of the agents (Elkind and Faliszewski 2014).

Now, the mechanism works in two steps.

- *Elimination step:* It identifies the most-preferred alternative x of the median agent, the alternative ℓ that is directly the left of x , and the alternative r that is directly to the right of x . It then eliminates one of ℓ and r by comparing the number $n(\ell, x)$ of agents that prefer ℓ to x and the number $n(r, x)$ of agents that prefer r to x as in the work of Anshelevich and Postl (2017). In particular, if $n(\ell, x) \leq n(r, x)$, then it eliminates ℓ , otherwise it eliminates r , and stores the non-eliminated alternative as y .
- *Weighted-majority step:* Afterwards, the mechanism runs a weighted majority between x and y by assigning a weight of 1 to each agent i such that $x, y \in A_i$, and a weight of $\frac{\alpha+1}{\alpha-1}$ to every other agent. This step will be shown to be equivalent to the algorithm used by Abramowitz, Anshelevich, and Zhu (2019) to show a bound of $2\sqrt{2} - 1 \approx 1.83$ in their setting when there are two alternatives.

4 Maximum Cost

In this section we consider the other most natural objective, i.e., to minimize the maximum cost of any agent, abbreviated as MC. For mechanisms in $\text{ORD} \cap \text{DIST} \cap \text{TAS}$ we show a tight bound of $1 + \sqrt{2}$. In the real line metric, we show that there in fact exist mechanisms in either $\text{ORD} \cap \text{TAS}$ and $\text{DIST} \cap \text{TAS}$ that achieve the $1 + \sqrt{2}$ distortion guarantee, without having access to the third type of information. We also show that the best distortion achievable by mechanisms in TAS is 3.

4.1 Results for General Metric Spaces

We start by proving an upper bound of $1 + \sqrt{2}$ for general metric spaces. This is achieved by a novel mechanism in $\text{ORD} \cap \text{DIST} \cap \text{TAS}$, which we refer to as α -MOST-COMPACT-SET, for $\alpha = 1 + \sqrt{2}$. In fact, the mechanism requires less information than a typical mechanism in $\text{ORD} \cap \text{DIST} \cap \text{TAS}$, as it only requires access to the most-preferred alternatives of the agents, rather than their complete ordinal preferences.

Before we define the mechanism, we will prove the following useful lemma relating the choices of any mechanism with the distortion for the maximum cost objective. To differentiate between alternatives that lie at minimum distance from an agent i and alternatives that appear first in the preference ranking of the agent (which might be different due to ties), we will refer to the former as *agent i 's min-distance*

Mechanism 2: α -MOST-COMPACT-SET

Input: Most-preferred alternatives of agents, distances between alternatives, α -TAS;

Output: Winner w ;

if $\exists x \in A_i, \forall i \in N$ **then**

$w \leftarrow x$;

else

for $i \in N$ **do**

$\rho_i = \max_{x \in A_i} d(x, o_i)$;

$w \in \arg \min_{i \in N} \rho_i$;

alternative (rather than agent i 's most-preferred alternative). In other words, a min-distance alternative for agent i is an alternative x for which $d(i, x) \leq d(i, x')$ for any $x' \in A$.

Lemma 4.1. *Consider a mechanism M that chooses the winner to be an alternative w that satisfies at least one of the following two conditions:*

1. $w \in A_j$ for all $j \in N$.
2. w is a min-distance alternative for some agent $i \in N$, and $o \notin A_i$, where o is an alternative that minimizes the maximum cost.

Then, for the maximum cost objective, the distortion of M is at most $\max\{\alpha, 2 + \frac{1}{\alpha}\}$.

Proof. For the first condition, if $w \in A_j$ for all $j \in N$, then

$$d(j, w) \leq \alpha \cdot d(j, o_j) \leq \alpha \cdot d(j, o),$$

where o_j is the most-preferred alternative of agent j . This implies that

$$\text{MC}(w) \leq \alpha \cdot \text{MC}(o).$$

For the second condition, let t be an agent that is at maximum distance from w . We have that

$$\begin{aligned} \text{MC}(w) &= d(t, w) \\ &\leq d(t, o) + d(o, i) + d(i, w) \\ &< d(t, o) + d(i, o) + \frac{1}{\alpha} \cdot d(i, o) \\ &\leq \left(2 + \frac{1}{\alpha}\right) \text{MC}(o), \end{aligned}$$

where the first inequality is due to the triangle inequality, and the second inequality follows from the fact that $o \notin A_i$. $\square$

Now, our mechanism, coined α -MOST-COMPACT-SET, works as follows: It first checks if there exists some alternative x that is approved by all agents, and, if this is indeed the case, it then chooses x as the winner w . Otherwise, for each agent i , it computes the maximum distance ρ_i of the most-preferred alternative o_i of i to any other alternative in A_i ; then it chooses the most-preferred alternative o_i of the agent i with minimum ρ_i . Essentially, in the otherwise case, the mechanism aims to choose the most-preferred alternative of an agent in order to minimize the radius to all other alternatives approved by the agent. See Mechanism 2 for a description using pseudocode.

Theorem 4.2. *In general metric spaces, for the max cost objective, the distortion of α -MOST-COMPACT-SET is at most $\max\{\alpha, 2 + \frac{1}{\alpha}\}$.*

Proof. If there is an alternative w such that $w \in A_j$ for all $j \in N$, the mechanism chooses it as the winner w . In that case, by Theorem 4.1, the distortion of α -MOST-COMPACT-SET is at most $\max\{\alpha, 2 + \frac{1}{\alpha}\}$. If such an alternative does not exist, then the mechanism chooses the most-preferred alternative of some agent i , in particular the agent with the smallest ρ_i , i.e., the smallest maximum distance from any of her approved alternatives to her most-preferred alternative o_i . If $o \notin A_i$, then by Theorem 4.1, the distortion of the mechanism is again at most $\max\{\alpha, 2 + \frac{1}{\alpha}\}$. Therefore we will consider the case when $o \in A_i$. By the definition of ρ_i , this implies that $d(o, o_i) \leq \rho_i$.

Since there is no alternative x such that $x \in A_i$ for all $i \in N$, there must exist at least one agent j such that $o \notin A_j$. Also, let t be the agent that has the maximum distance from w . We have that

$$\begin{aligned} \text{MC}(w) &= d(t, o_i) \\ &\leq d(t, o) + d(o, o_i) \\ &\leq \text{MC}(o) + \rho_i \\ &\leq \text{MC}(o) + \rho_j. \end{aligned}$$

where the last inequality follows from that fact that i was chosen to be an agent that minimizes $\rho_k = \max_{x \in A_k} d(x, o_k)$ by the mechanism. Let y be an alternative in A_j with maximum distance from o_j . Then,

$$\begin{aligned} \rho_j &= d(y, o_j) \\ &\leq d(j, o_j) + d(j, y) \\ &\leq \left(1 + \frac{1}{\alpha}\right) \text{MC}(o), \end{aligned}$$

where the first inequality is due to the triangle inequality, and the second one follows since $d(j, o_j) \leq d(j, o) \leq \text{MC}(o)$, and since $o \notin A_j$ implies that $d(j, o_j) < \frac{1}{\alpha} d(j, o) \leq \frac{1}{\alpha} \text{MC}(o)$. Putting everything together, we obtain

$$\text{MC}(w) \leq \left(2 + \frac{1}{\alpha}\right) \text{MC}(o).$$

Overall, the distortion of the mechanism is at most $\max\{\alpha, 2 + \frac{1}{\alpha}\}$, as desired. $\square$

By optimizing over α , we obtain the following corollary.

Corollary 4.3. *In general metric spaces, for the max cost objective, the distortion of $(1 + \sqrt{2})$ -MOST-COMPACT-SET is at most $1 + \sqrt{2}$.*

Next, we prove that the bound of $1 + \sqrt{2}$ is the best possible for the maximum cost objective, even when all three types of information are available, and even when the metric space is as simple as a line.

Theorem 4.4. *For the max cost, the distortion of any mechanism $M \in \text{ORD} \cap \text{DIST} \cap \text{TAS}$ is at least $1 + \sqrt{2}$, even on a line metric.*

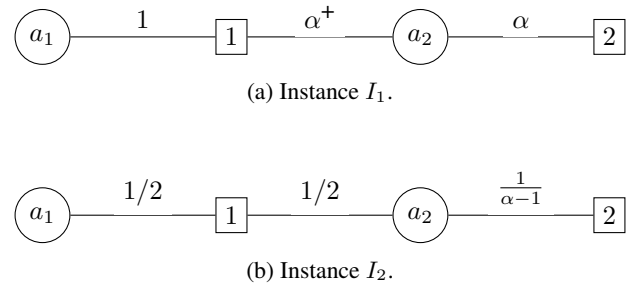


Figure 1: The line metrics used in the proof of Theorem 4.4. Agents are represented by rectangles and alternatives are represented by circles. A value above an edge represents the distance between the edge's endpoints.

Proof. Consider any mechanism M that has access to all three types of information. To show the lower bound, we will consider the following two cases depending on the value of α that M uses.

Case 1: $\alpha \leq 1 + \sqrt{2}$. Consider an instance I_1 with two agents $\{1, 2\}$ and two alternatives $\{a_1, a_2\}$. We have

- $A_1 = \{a_1\}$;
- $A_2 = \{a_2\}$;
- $a_1 \succ_1 a_2$;
- $a_2 \succ_2 a_1$.

The distance between the two alternatives is equal to $1 + \alpha + \epsilon$, for some infinitesimal $\epsilon > 0$. Given this information, M may choose any of the two alternatives as the winner, say a_1 . We define the following line metric (see Fig. 1a):

- Alternative a_1 is at 0;
- Agent 1 is at 1;
- Alternative a_2 is at $1 + \alpha + \epsilon$
- Agent 2 is at $1 + 2\alpha + \epsilon$.

It is not hard to verify that these positions of the agents and the alternatives are consistent to the information given to the mechanism for any $\alpha \leq 1 + \sqrt{2}$: Agent 2 approves only alternative a_2 as $d(2, a_1) = 1 + 2\alpha + \epsilon > \alpha^2 = \alpha \cdot d(2, a_2)$; all other information is clearly consistent. We have that

$$\text{MC}(a_1) = 1 + 2\alpha + \epsilon$$

and

$$\text{MC}(a_2) = \alpha + \epsilon,$$

thus obtaining a lower bound of $2 + 1/\alpha \geq 1 + \sqrt{2}$ when ϵ tends to 0.

Case 2: $\alpha > 1 + \sqrt{2}$. Consider an instance I_2 with two agents $\{1, 2\}$ and two alternatives $\{a_1, a_2\}$. We have

- $A_1 = A_2 = \{a_1, a_2\}$;
- $a_1 \succ_1 a_2$;
- $a_2 \succ_2 a_1$.

The distance between the two alternatives is equal to 1. Given this information, M may choose any of the two alternatives as the winner, say a_1 . We define the following line metric (see Fig. 1b):

- Alternative a_1 is at 0;
- Agent 1 is at $1/2$;
- Alternative a_2 is at 1;
- Agent $\min\{3/2, \frac{\alpha}{\alpha-1}\}$.

It is not hard to verify that these positions of the agents and the alternatives are consistent to the information given to the mechanism for any $\alpha > 1 + \sqrt{2}$: Agent 1 is at distance $1/2$ from both a_1 and a_2 . When $\frac{\alpha}{\alpha-1} \leq 3/2 \Leftrightarrow \alpha \leq 3$, agent 2 is at distance $\frac{\alpha}{\alpha-1} - 1 = \frac{1}{\alpha-1}$ from a_2 and distance $\frac{\alpha}{\alpha-1}$ from a_1 ; clearly, the ratio of these distances is exactly α . When $\frac{\alpha}{\alpha-1} > 3/2 \Leftrightarrow \alpha > 3$, agent 2 is at distance $1/2$ from a_2 and distance $3/2$ from a_1 ; the ratio of these distances is exactly 3, and thus at most α . Since $\frac{1}{\alpha-1} < 1/2 < \frac{\alpha}{\alpha-1}$ for any $\alpha > 1$, we have that

$$MC(a_1) = \min\{3/2, \frac{\alpha}{\alpha-1}\}$$

and

$$MC(a_2) = 1/2,$$

thus obtaining a lower bound of 3 when $\alpha > 3$ and

$$\frac{2\alpha}{\alpha-1} \geq \alpha \geq 1 + \sqrt{2}$$

when $\alpha \in (1 + \sqrt{2}, 3]$. $\square$

We conclude the exposition of our results for general metric spaces by showing that the very simple mechanism ANY-APPROVED $\in$ TAS that chooses any alternative in the α -TAS of some agent achieves a distortion bound of $2 + \alpha$ in terms of the maximum cost objective. By setting $\alpha = 1$, we obtain a distortion bound of 3, which is the best possible for all mechanisms in TAS, and also matches the bound of 3 that is best possible for any mechanism in $ORD \cap DIST$. We defer the corresponding statements and proofs to the full version.

Refined Bounds for the Real Line

As in the case of the social cost objective, we also obtain tight bounds on the line for the maximum cost objective. The lower bound of Theorem 4.4 for mechanisms in $ORD \cap DIST \cap TAS$ already applies even to the line metric, so the question is whether there is a mechanism in this class that achieves a corresponding upper bound. We show something stronger than this, namely that mechanisms in either $ORD \cap TAS$ or $DIST \cap TAS$ indeed achieve this bound.

Theorem 4.5. *When the metric space is a line, for the maximum cost, there are mechanisms in $ORD \cap TAS$ and in $DIST \cap TAS$ with distortion $1 + \sqrt{2}$.*

For $ORD \cap TAS$, the mechanism that achieves the bound is coined MAX-TAS-LEFTMOST, and simply selects the leftmost alternative (which is possible to find given the ordinal preferences, see Remark 1) from those alternatives that appear in the largest number of α -TAS. For $DIST \cap ORD$, the natural translation of α -MINISUM-TAS-DISTANCE, which minimizes the maximum distance from the approval sets (coined α -MINIMAX-TAS-DISTANCE) achieves this upper bound. Interestingly, this bound is achieved even by

α -MINISUM-TAS-DISTANCE itself, although it is not designed with the maximum cost objective in mind. We defer the details to the full version.

5 Future Directions

The main problem that our work leaves open is that of identifying the best possible distortion bound in terms of the social cost for general metric spaces and mechanisms in $ORD \cap DIST \cap TAS$, for which we showed an upper bound of $1 + \sqrt{2}$ and a lower bound of 2. Another interesting open question is to show whether $1 + \sqrt{2}$ can be achieved with a mechanism in $ORD \cap TAS$ for general metric spaces. For the maximum cost objective, we were able to obtain tight bounds for the general class $ORD \cap DIST \cap TAS$, so the open question is whether these bounds can be achieved by mechanisms in $ORD \cap TAS$ and $DIST \cap TAS$, as we proved to be the case on the real line. For $DIST \cap TAS$, the $(1 + \sqrt{2})$ -MINIMAX-TAS-DISTANCE mechanism, which we showed to achieve the tight bound of $1 + \sqrt{2}$ on the line, seems like a very natural candidate. Unfortunately, however, we can construct counterexamples where the distortion of this mechanism is lower-bounded by 3 in general metric spaces. For $ORD \cap TAS$, MAX-TAS-LEFTMOST is obviously not even well-defined in general metric spaces, as the notion of a leftmost alternative is meaningless. That being said, we can design mechanisms in $ORD \cap TAS$ that are based on very similar principles and are not line-specific, which still achieve a distortion of $1 + \sqrt{2}$ on the line. Unfortunately, once considered in general metric spaces, these mechanisms also fall short, as similar examples to the aforementioned one show a lower bound of 3 on their distortion.

More generally, our paper advocates the study of the interplay between information and efficiency in the metric social choice setting, building on a growing literature that has recently gained significant momentum. As a next step, one could consider eliciting additional cardinal information, for example, by using multiple approval thresholds $\alpha_1, \dots, \alpha_k$, in conjunction with the other types of information that we use here, and quantify the effect on the distortion. More generally, it would make sense to consider different types of information structures and combinations between them; note, for example, that our α -MOST-COMPACT-SET mechanism for the maximum cost uses the distances between alternatives, the α -TAS, and information about the top-ranked alternative of each agent. Finally, a very meaningful avenue is to consider how randomization might aid in achieving even more improved distortion bounds, in the presence of the α -TAS.

References

- Abramowitz, B.; Anshelevich, E.; and Zhu, W. 2019. Awareness of voter passion greatly improves the distortion of metric social choice. In *Proceedings of the 15th International Conference on Web and Internet Economics (WINE)*, 3–16.
- Amanatidis, G.; Birmpas, G.; Filos-Ratsikas, A.; and Voudouris, A. A. 2021. Peeking behind the ordinal curtain: Improving distortion via cardinal queries. *Artificial Intelligence*, 296: 103488.

- Amanatidis, G.; Birmpas, G.; Filos-Ratsikas, A.; and Voudouris, A. A. 2022a. Don't Roll the Dice, Ask Twice: The Two-Query Distortion of Matching Problems and Beyond. In *Proceedings of the 36th Annual Conference on Neural Information Processing Systems (NeurIPS)*.
- Amanatidis, G.; Birmpas, G.; Filos-Ratsikas, A.; and Voudouris, A. A. 2022b. A few queries go a long way: Information-distortion tradeoffs in matching. *Journal of Artificial Intelligence Research*, 74: 227–261.
- Anshelevich, E.; Bhardwaj, O.; Elkind, E.; Postl, J.; and Skowron, P. 2018. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264: 27–51.
- Anshelevich, E.; Filos-Ratsikas, A.; Shah, N.; and Voudouris, A. A. 2021. Distortion in Social Choice Problems: The First 15 Years and Beyond. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*, 4294–4301.
- Anshelevich, E.; and Postl, J. 2017. Randomized Social Choice Functions Under Metric Preferences. *Journal of Artificial Intelligence Research*, 58: 797–827.
- Anshelevich, E.; and Zhu, W. 2021. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. *ACM Transactions on Economics and Computation*, 9(2): 1–24.
- Benade, G.; Nath, S.; Procaccia, A. D.; and Shah, N. 2021. Preference elicitation for participatory budgeting. *Management Science*, 67(5): 2813–2827.
- Bhaskar, U.; Dani, V.; and Ghosh, A. 2018. Truthful and near-optimal mechanisms for welfare maximization in multi-winner elections. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.
- Boutilier, C.; Caragiannis, I.; Haber, S.; Lu, T.; Procaccia, A. D.; and Sheffet, O. 2015. Optimal social choice functions. *Artificial Intelligence*, 227(C): 190–213.
- Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D. 2016. *Handbook of computational social choice*. Cambridge University Press.
- Caragiannis, I.; Procaccia, A.; and Shah, N. 2016. Truthful univariate estimators. In *International Conference on Machine Learning*, 127–135.
- Chen, Y.; Podimata, C.; Procaccia, A. D.; and Shah, N. 2018. Strategyproof linear regression in high dimensions. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, 9–26.
- Chen, Y.; Shen, Y.; and Zheng, S. 2020. Truthful data acquisition via peer prediction. *Advances in Neural Information Processing Systems*, 33: 18194–18204.
- Dwork, C.; Kumar, R.; Naor, M.; and Sivakumar, D. 2001. Rank aggregation methods for the web. In *Proceedings of the 10th international conference on World Wide Web*, 613–622.
- Elkind, E.; and Faliszewski, P. 2014. Recognizing 1-Euclidean Preferences: An Alternative Approach. In *Proceedings of the 7th International Symposium on Algorithmic Game Theory*, 146–157.
- Ephrati, E.; and Rosenschein, J. S. 1991. The Clarke Tax as a Consensus Mechanism Among Automated Agents. In *AAAI*, volume 91, 173–178.
- Ghosh, S.; Mundhe, M.; Hernandez, K.; and Sen, S. 1999. Voting for movies: the anatomy of a recommender system. In *Proceedings of the third annual conference on Autonomous Agents*, 434–435.
- Gkatzelis, V.; Halpern, D.; and Shah, N. 2020. Resolving the optimal metric distortion conjecture. In *Proceedings of the 61st IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, 1427–1438.
- Gross, S.; Anshelevich, E.; and Xia, L. 2017. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.
- Kahng, A.; Kehne, G.; and Procaccia, A. 2020. Strategyproof Mean Estimation from Multiple-Choice Questions. In *International Conference on Machine Learning (ICML)*, 5042–5052.
- Kahng, A.; Lee, M. K.; Noothigattu, R.; Procaccia, A.; and Psomas, C.-A. 2019. Statistical foundations of virtual democracy. In *International Conference on Machine Learning (ICML)*, 3173–3182. PMLR.
- Kempe, D. 2020. Communication, distortion, and randomness in metric voting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 2087–2094.
- Kizilkaya, F. E.; and Kempe, D. 2022. Plurality Veto: A Simple Voting Rule Achieving Optimal Metric Distortion. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, 349–355.
- Mandal, D.; Procaccia, A. D.; Shah, N.; and Woodruff, D. P. 2019. Efficient and thrifty voting by any means necessary. In *Proceedings of the 32nd Annual Conference on Neural Information Processing Systems (NeurIPS)*, 7178–7189.
- Mandal, D.; Shah, N.; and Woodruff, D. P. 2020. Optimal communication-distortion tradeoff in voting. In *Proceedings of the 21st ACM Conference on Economics and Computation*, 795–813.
- Mao, A.; Procaccia, A. D.; and Chen, Y. 2012. Social choice for human computation. In *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*.
- Noothigattu, R.; Gaikwad, S.; Awad, E.; Dsouza, S.; Rahwan, I.; Ravikumar, P.; and Procaccia, A. 2018. A voting-based system for ethical decision making. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 32.
- Peters, D.; Procaccia, A. D.; Psomas, A.; and Zhou, Z. 2020. Explainable voting. *Advances in Neural Information Processing Systems*, 33: 1525–1534.
- Procaccia, A. D.; and Rosenschein, J. S. 2006. The distortion of cardinal preferences in voting. In *Cooperative Information Agents X: 10th International Workshop, CIA 2006 Edinburgh, UK, September 11-13, 2006 Proceedings 10*, 317–331. Springer.
- Von Neumann, J.; and Morgenstern, O. 2007. Theory of games and economic behavior. In *Theory of games and economic behavior*. Princeton university press.

Xia, L. 2022. Group decision making under uncertain preferences: powered by AI, empowered by AI. *Annals of the New York Academy of Sciences*, 1511(1): 22–39.

Zhao, Z.; and Xia, L. 2019. Learning mixtures of plackett-luce models from structured partial orders. *Advances in Neural Information Processing Systems*, 32.