Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence

# Approximating Optimal Social Choice under Metric Preferences

**Elliot Anshelevich** and **Onkar Bhardwaj** and **John Postl**
Rensselaer Polytechnic Institute
110 8th Street, Troy NY 12180
eanshel@cs.rpi.edu, bhardo@rpi.edu, postlj@rpi.edu

**Abstract**

We examine the quality of social choice mechanisms using a *utilitarian* view, in which all of the agents have costs for each of the possible alternatives. While these underlying costs determine what the optimal alternative is, they may be unknown to the social choice mechanism; instead the mechanism must decide on a good alternative based only on the ordinal preferences of the agents which are induced by the underlying costs. Due to its limited information, such a social choice mechanism cannot simply select the alternative that minimizes the total social cost (or minimizes some other objective function). Thus, we seek to bound the *distortion*: the worst-case ratio between the social cost of the alternative selected and the optimal alternative. Distortion measures how good a mechanism is at approximating the alternative with minimum social cost, while using only ordinal preference information. The underlying costs can be arbitrary, implicit, and unknown; our only assumption is that the agent costs form a *metric space*, which is a natural assumption in many settings. We quantify the distortion of many well-known social choice mechanisms. We show that for both total social cost and median agent cost, many positional scoring rules have large distortion, while on the other hand *Copeland* and similar mechanisms perform optimally or near-optimally, always obtaining a distortion of at most 5. We also give lower bounds on the distortion that could be obtained by *any* deterministic social choice mechanism, and extend our results on median agent cost to more general objective functions.

## 1 Introduction

*Social choice theory* deals with aggregating agent preferences over a set of alternatives into a collective decision via a *social choice mechanism*. The social choice mechanism takes as input the preferences of agents, which are usually total orderings over the set of alternatives, and typically outputs a single alternative as the winner. It is natural to now ask about the quality of different social choice mechanisms; to do this one needs to define what it means for a chosen alternative to be "good" or to accurately represent the consensus of the agent preferences. A popular way of achieving this is to define criteria or axioms for social choice mechanisms, which guarantee that the alternatives selected by these mechanisms satisfy desirable properties (see Related Work). Another common approach in fields like welfare economics and algorithmic mechanism design, and which we follow in this paper, is to use a *utilitarian* view (Boutilier et al. 2012). Instead of assuming that agents only have ordinal preferences over the alternatives, this approach assumes that every agent has (possibly latent or implicit) utility or cost values over the alternatives. These values are cardinal, and represent how happy the agent is with each alternative. The quality of an alternative can then be defined simply as the sum (or some other objective function) of the utility received by each agent for that alternative. Thus the best, or optimal, alternative is simply the one that maximizes the total social welfare (or minimizes total cost), as measured by the total utility received by the agents.

Utilitarian approach has recently received renewed attention in the study of social choice (Procaccia and Rosenschein 2006; Caragiannis and Procaccia 2011; Boutilier et al. 2012; Branzei et al. 2013). Indeed, as argued in Boutilier et al. (2012), although not all social choice problems are amenable to the utilitarian approach (especially the ones where it is unnatural to assume that agent utilities or costs can be compared) there are many real-life settings which fit the utilitarian view. For example, in recommender systems and many similar domains from mechanism design and e-commerce, the computational agents typically use real-valued rather than ordinal utilities (see Related Work and Boutilier et al. (2012)).

If the social choice mechanism knew exactly what utilities the agents receive from each alternative, then it could simply pick the alternative maximizing social welfare directly. An important point here, however, is that while we assume that some underlying utility structure exists, it is unreasonable to assume that we (or even the agents themselves) know exactly what it is. As discussed in Boutilier et al. (2012), it is often difficult for agents to determine their exact cardinal utilities, and most social choice mechanisms thus take only the ordinal preference orderings of the agents as input, even when latent utilities exist. Thus, a social choice function will not simply output the alternative that maximizes global utility, but instead may choose another alternative, since it only has access to ordinal preferences. As a result, one can think of a social choice function as an approximation algorithm which attempts to choose the best possible alternative (maximize social welfare or minimize social cost), but only has

| | Sum | Median |
|---|---|---|
| Plurality | $2m-1$ | $\infty$ |
| Borda | $2m-1$ | $\infty$ |
| $k$-approval | $2n-1$ | $\infty$ |
| Veto | $2n-1$ | $\infty$ |
| Copeland | 5 | 5 |
| Uncovered Set | 5 | 5 |
| Lower Bound | 3 | 5 |

Table 1: The worst-case distortion of various social choice mechanisms for both the sum and the median objective functions. All of the above bounds are provably tight, meaning that we provide example instances where the social choice function cannot achieve a better bound. The lower bounds of 3 and 5 are for any deterministic social choice functions.

access to limited information (ordinal preferences instead of cardinal utilities). To denote the approximation factor of a social choice function, Procaccia and Rosenschein (2006) introduced the term *distortion* which we will continue to use, although we will define it in terms of social cost instead of social welfare. Informally, the distortion of a social choice function is the worst-case ratio of the social cost of the alternative selected by the social choice function over the cost of the optimal alternative.

In this work, we are primarily interested in determining the quality of outcomes chosen by social choice mechanisms, as measured by their distortion. We prove bounds on the distortion of many well-known social choice functions for both the sum and median objective functions. Our results show that while the distortion is high for some mechanisms, the distortion of many important social choice functions is bounded by a small constant, assuming that the preferences of the agents are *spatial*. Specifically, we assume that the costs of agents for various alternatives form an arbitrary *metric space*. Such *metric costs* have a very natural interpretation – in the context of voting, as described in the classic Downsian proximity model (Merrill and Grofman 1999), we can think of the cost experienced by voter $i$ due to candidate $j$ being elected as the distance between $i$ and $j$'s beliefs in some high-dimensional space, as the number of issues they disagree on, etc. Such spatial preferences have been extensively studied (see Related Work), although usually the metric space is assumed to be simple, e.g., Euclidean with only one or two dimensions. In contrast, we make no assumptions about the metric space, other than the fact that it is a metric space. To see how general our metric assumption is, note that, unlike many common assumptions on spacial agent preferences, our metric assumption does not restrict the set of possible ordinal preferences in any way (see Proposition 1 and discussion before it).

## 1.1 Our Contributions

In this work, we bound the worst-case distortion of many well-known social choice functions. In other words, we show how closely social choice functions approximate the optimal alternative when they are given only the ordinal preference orderings, instead of the underlying metric costs which generate these preferences.[1] We consider two general objective functions to quantify the quality of alternatives, and give distortion results for both. The first is the sum objective function which defines the social cost of an alternative as the sum of all agent costs for that alternative. This function is very natural, and is the most common measure of social cost. Our other objective function defines the quality of an alternative as the median of agent costs for that alternative: this captures the objective that the best alternative is the one in which the cost of the median voter is minimized, instead of the average voter.

Most of our results are summarized in Table 1. First, we consider how well *any* social choice function could do when it only knows the ordinal preferences, but is supposed to approximate the social optimum. We show that no deterministic social choice mechanism can have worst-case distortion better than 3 (for the sum objective), or better than 5 (for the median objective). With these lower bounds established, we can nevertheless ask: do there exist social choice rules which meet this lower bound? Are there rules which obtain the minimum possible distortion?

We begin with the bad news: for common positional scoring rules such as plurality, Borda, $k$-approval, and veto, we prove that the worst case distortion can be high: either $2m-1$ or $2n-1$ where $m$ is the number of alternatives and $n$ is the number of agents/voters. There is good news as well, however. For the Copeland social choice rule, we prove that the distortion is always at most 5. This means that, although the Copeland social choice mechanism knows nothing about the metric costs other than the ordinal preferences induced by them, and cannot possibly find the true optimal alternative, it nevertheless *always* selects an alternative whose quality is only a factor of 5 away from optimal! Moreover, due to our lower bound, no deterministic mechanism can do better than Copeland for the median objective, and no deterministic mechanism can do much better than Copeland for the sum objective, because the distortion lower bound for any deterministic mechanism is 3.

While this bound of 5 holds for both the sum and median objectives, different techniques are required to prove it for the two cases. In fact, this bound holds not just for Copeland, but for similar voting rules as well, such as uncovered set (Moulin 1986). Since Copeland does not perform as well as the lower bound for the sum objective, we also analyze the distortion of the ranked pairs mechanism. We show that it performs even better than Copeland, but only when certain conditions on the agent preference profiles are satisfied (see Theorem 7).

In addition to the results in Table 1, we also analyze more general objective functions. Specifically, instead of the median objective, which sets the quality of an alternative $W$ to be the cost to the median voter, we consider more general *percentile* objectives, where the quality of an alternative $W$ is set to be the cost of the voter at the $x$'th percentile.

[1]This is assuming that agents submit their true ordinal preferences. We leave questions about non-truthful agents as future work.

We show how the distortion of various mechanisms changes with $x$, and establish that Copeland remains the mechanism with the best possible distortion for most values of $x$.

### 1.2 Related Work

The focus of much of the existing literature in social choice theory is the design and analysis of social choice functions with respect to various normative criteria (See for example (Faliszewski and Procaccia 2010; Bartholdi III, Tovey, and Trick 1989; Conitzer and Sandholm 2002; Lang and Xia 2009)). Results like Arrow's impossibility theorem and Gibbard-Satterthwaite theorem demonstrate nonexistence of social choice functions satisfying certain desirable criteria, and additional assumptions must be made in order to circumvent these results (e.g., (Moulin 1980; Gans and Smart 1996; Myerson 1996)).

In this work, we instead adopt a *utilitarian* view of social choice as described in the Introduction. Social choice with utilitarian viewpoint has its advocates in welfare economics (Roemer 1998; Ng 1997) and has recently received attention from the AI community (Procaccia and Rosenschein 2006; Caragiannis and Procaccia 2011; Boutilier et al. 2012). The utilitarian approach has also been investigated in recommender systems (Ghosh et al. 1999), information extraction (Sigletos et al. 2005), etc. While assuming that agent utilities can be compared does not make sense for all settings (Harsanyi 1976), it is nevertheless reasonable in many applications of interest: see Boutilier et al. (2012) for much more discussion on this subject.

*Distortion* as a measure of performance of a social choice function in utilitarian settings was introduced first in Procaccia and Rosenschein (2006) and later used in Boutilier et al. (2012). In both these works, the worst-case distortion of social choice functions was shown to be unbounded or very high. In our work, however, we show that considering agent costs that form an *unknown metric* immediately greatly reduces the distortion of many mechanisms, from unbounded to only a small constant. Caragiannis and Procaccia (2011) use an analogous notion of distortion to analyze the worst-case distortion of embeddings into voting rules: these embeddings are functions that take as input an agent's utility function and determine which alternative the agent should select. Apart from the classic normative criteria, other papers have also used related interpretations of what makes a good social choice function, such as distance rationalizability (Elkind, Faliszewski, and Slinko 2009), rank approximation (Chakrabarty and Swamy 2014), and dynamic price of anarchy (Branzei et al. 2013).

In our paper, we assume that agents have spatial preferences resulting from metric agent costs. Spatial preferences and utility theory in the context of voting have a strong tradition in social choice and political science (Enelow and Hinich 1984; Merrill and Grofman 1999). Common assumptions include single-peaked preferences (Moulin 1980; Sui, Francois-Nienaber, and Boutilier 2013) and single-crossing preferences (Saporiti 2009; Myerson 1996); often preferences are assumed to be one-dimensional, while we consider metrics with arbitrary dimension.

Finally, the concept of distortion is related to many other notions of approximation, as it compares the optimal solution with the solution obtained given only limited information. This is similar, for example, to the competitive ratio in online algorithms, which is a measure of how an algorithm performs with limited information (not knowing the future), compared to how an all-knowing algorithm would perform (Borodin and El-Yaniv 1998; Oren and Lucier 2014).

## 2 Preliminaries

**Social Choice with Ordinal Preferences.** Let $N = \{1, 2, \dots, n\}$ be the set of agents, and let $\mathcal{A} = \{A_1, A_2, \dots, A_m\}$ be the set of alternatives. Let $\mathcal{S}$ be the set of all total orders on the set of alternatives $\mathcal{A}$. We will typically use $i, j$ to refer to agents and $W, X, Y, Z$ to refer to alternatives. Every agent $i \in N$ has a *preference ranking* $\sigma_i \in \mathcal{S}$; by $X \succ_i Y$ we will mean that $X$ is preferred over $Y$ in ranking $\sigma_i$. We call the vector $\sigma = (\sigma_1, \dots, \sigma_n) \in \mathcal{S}^n$ a *preference profile*. We say that an alternative $X$ *pairwise defeats* $Y$ if $|\{i \in N : X \succ_i Y\}| > \frac{n}{2}$.

Once we are given a preference profile, we want to aggregate the preferences of the agents and select a single alternative as the winner. A *social choice function* $f : \mathcal{S}^n \to \mathcal{A}$ is a mapping from a preference profile to an alternative. Some well-known social choice functions which we consider in this paper are as follows.

- **Positional scoring rules.** We are given a scoring vector $\vec{s} = (s_1, s_2, \dots, s_m)$ with $s_1 \geq s_2 \geq \cdots \geq s_m$. If an agent ranks an alternative in position $l$, then the alternative receives $s_l$ points. The total score $s(X, \sigma)$ of an alternative $X$ for a preference profile $\sigma$ is the total number of points that $X$ receives. The positional scoring rule is $f(\sigma) = \arg\max_{X \in \mathcal{A}} s(X, \sigma)$; that is, it selects the alternative with the highest total score. Many well-known voting rules can be thought of as positional scoring rules, for example:
  - **Plurality:** $\vec{s} = (1, 0, \dots, 0)$
  - **Veto:** $\vec{s} = (1, 1, \dots, 1, 0)$
  - **Borda:** $\vec{s} = (m-1, m-2, \dots, 1, 0)$
  - $k$**-approval** $(1 < k < m)$: $\vec{s} = (\underbrace{1, 1, \dots, 1}_{k}, 0, \dots, 0)$
- **Copeland:** The score of an alternative $X$ is $|\{Y \in \mathcal{A} : X \text{ pairwise defeats } Y\}|$. The alternative with the highest score, i.e., the alternative with the largest number of pairwise victories, is the winner.
- **Ranked pairs:** Construct a graph $G$ in the following manner. Let every alternative be a node in $G$. For every pair of alternatives $X, Y$, let $w(X, Y) = |\{i \in N : X \succ_i Y\}|$. Sort these $w(X, Y)$ values in non-increasing order and iterate over them. For each $w(X, Y)$ value, add the directed edge $(X, Y)$ to $G$ if it won't create a cycle, and do nothing otherwise. The winner is the source node of the resulting directed acyclic graph.

**Cardinal Metric Costs.** In our work we take the utilitarian view, and study the case when the ordinal preferences $\sigma$ are in fact a result of the underlying cardinal agent costs.

Formally, we assume that there exists an arbitrary metric $d : (N \cup \mathcal{A})^2 \to \mathbb{R}_{\geq 0}$ on the set of agents and alternatives (or more generally a *semi-metric*, since we allow agents to be identical and have $d(i, j) = 0$). Here $d(i, X)$ is the cost incurred by agent $i$ of alternative $X$ being selected as the winner; these costs can be arbitrary but are assumed to obey the triangle inequality. The metric costs $d$ naturally give rise to a preference profile. Formally, we say that $\sigma$ is *consistent* with $d$ if $\forall i \in N, \forall X, Y \in \mathcal{A}$, if $d(i, X) < d(i, Y)$, then $X \succ_i Y$. In other words, if the cost of $X$ is less than the cost of $Y$ for an agent, then the agent should prefer $X$ over $Y$. Let $p(d)$ denote the set of preference profiles consistent with $d$ ($p(d)$ may include several preference profiles if the agent costs have ties). Similarly, we define $p^{-1}(\sigma)$ to be the set of metrics such that $\sigma \in p(d)$.

When making additional assumptions on how the preference rankings of the agents are generated, the set of possible preference profiles may become restricted. For example, if we restrict agents to one-dimensional single-peaked preferences, or to single-crossing preferences, then preference profiles with the Condorcet paradox can no longer be realized (Black 1948; Gans and Smart 1996; Saporiti 2009). However, having arbitrary metric costs in our model does not restrict the set of possible profiles $\sigma$ in any way: metrics are general enough that any preference profile in $\mathcal{S}^n$ can be induced.

**Proposition 1** *For every preference profile $\sigma$, there exists a metric $d$ such that $\sigma$ is consistent with $d$.*

Any missing proofs can be found in the full version on http://cs.rpi.edu/∼eanshel.

**Social Cost and Distortion.** We measure the quality of each alternative using the costs incurred by all the agents when this alternative is chosen. We use two different notions of social cost. First, we study the sum objective function, defined as $\mathrm{SC}_{\sum}(X, d) = \sum_{i \in N} d(i, X)$; this is the most common notion of social cost. We also study the median objective function, $\mathrm{SC}_{\mathrm{med}}(X, d) = \mathrm{med}_{i \in N}(d(i, X))$. As described in the Introduction, we can view social choice mechanisms in our setting as attempting to find the optimal alternative (one that minimizes social cost), but only having access to the ordinal preference profile $\sigma$, instead of the full underlying costs $d$. The following proposition establishes that this is impossible to do: the only way one can determine the optimal alternative while only having access to $\sigma$ is if there is a single alternative that is the top preference for *all* agents. In fact, we cannot even eliminate any alternative from consideration of being optimal, except in trivial cases.

**Proposition 2** *For any preference profile $\sigma$ and alternative $X$, there exists a metric $d \in p^{-1}(\sigma)$ such that $X$ is optimal with respect to the social cost function $\mathrm{SC}_{\sum}(X, d)$, except when there exists an alternative $Y$ such that for all $i \in N$, $Y \succ_i X$.*

Since it is impossible to compute the optimal alternative using only ordinal preferences, we would like to determine how well the aforementioned social choice functions select alternatives based on their social costs, despite only being given the preference profiles. In particular, we would like to quantify how the social choice functions perform in the worst-case. To do this, we use the notion of *distortion* from (Procaccia and Rosenschein 2006; Boutilier et al. 2012), defined as follows.

$$\mathrm{dist}_{\sum}(f, \sigma) = \sup_{d \in p^{-1}(\sigma)} \frac{\mathrm{SC}_{\sum}(f(\sigma), d)}{\min_{X \in \mathcal{A}} \mathrm{SC}_{\sum}(X, d)}$$

$$\mathrm{dist}_{\mathrm{med}}(f, \sigma) = \sup_{d \in p^{-1}(\sigma)} \frac{\mathrm{SC}_{\mathrm{med}}(f(\sigma), d)}{\min_{X \in \mathcal{A}} \mathrm{SC}_{\mathrm{med}}(X, d)}.$$

In other words, the distortion of a social choice mechanism $f$ on a profile $\sigma$ is the worst-case ratio between the social cost of $f(\sigma)$, and the social cost of the true optimum alternative. The worst-case is taken over all metrics $d$ which may have induced $\sigma$, since the social choice function does not and cannot know which of these metrics is the true one.

## 3 Distortion of Total Agent Cost

In this section, we study the sum objective function, which measures the quality of an alternative to be the total agent cost when this alternative is chosen. We prove tight upper bounds for distortion of several well-known social choice functions. Our main result in this section is that the Copeland voting mechanism (as well as several others) exhibit a distortion of at most 5; this guarantee is independent of the number of agents or alternatives, and the underlying metric space is allowed to be completely arbitrary (and unknown).

Before proceeding with showing upper bounds on possible distortion, we ask the question: how well can *any* social choice function perform? The following simple theorem tells us that we cannot possibly hope to approximate the optimal alternative within a factor better than 3.

**Theorem 3** *No (deterministic) social choice function has worst-case distortion less than 3 for the sum objective.*

**Proof.** Suppose there are only two alternatives $X$ and $W$. Half of the agents prefer $X$ over $W$, and the other half prefer $W$ over $X$. Suppose without loss of generality that the given social choice function picks $W$ as the winner. The underlying metric can be as follows. All $n/2$ agents who prefer $X$ are located exactly at $X$, i.e., $d(i, X) = 0$ and $d(i, W) = 2$. All $n/2$ agents who prefer $W$ are approximately halfway between $X$ and $W$, i.e., $d(i, X) = 1 + \epsilon$ and $d(i, W) = 1 - \epsilon$ for some small $\epsilon > 0$. Then $\mathrm{SC}_{\sum}(X, d) = \sum_{i \in N} d(i, X) = (1 + \epsilon) \cdot n/2$ and $\mathrm{SC}_{\sum}(W, d) = \sum_{i \in N} d(i, W) = 2 \cdot n/2 + (1 - \epsilon) \cdot n/2$. Thus, the distortion approaches 3 as $\epsilon \to 0$. ■

In fact, it is easy to show that for only two alternatives, any social choice function that picks the winner preferred by the majority of agents has a distortion of 3, i.e., all such social choice functions achieve the optimal distortion bound for two alternatives. This is a corollary of Theorem 4. Unfortunately, as the number of agents and candidates becomes large, the distortion of many social choice mechanisms increases linearly.

**Theorem 4** *For plurality and Borda social choice functions, the distortion is at most* $2m-1$*; for* $k$*-approval and veto it is at most* $2n-1$*. Furthermore, these bounds are tight, i.e., they are achieved exactly in some instances.*

Before proving this theorem, we observe the helpful lemma below (see the full version for complete proof). For the remainder of this section, we will use the following notation: $WX = \{i \in N : W \succ_i X\}$ and $WXY = \{i \in N : W \succ_i X \succ_i Y\}$.

**Lemma 5** *For any instance* $\sigma$ *and social choice function* $f$*,* $\text{dist}_{\sum}(f, \sigma) \leq 1 + \frac{2(n-|WX|)}{|WX|}$*, where* $W = f(\sigma)$ *is the winning alternative and* $X$ *is the optimal alternative.*

**Proof Sketch of Theorem 4.** Let $W$ denote the winning alternative, and let $X$ denote an optimal alternative. Since the bound from Lemma 5 decreases with $|WX|$, bounding the smallest possible $|WX|$ for each scoring function will give us an upper bound on the worst-case distortion. For plurality and Borda, it is not difficult to show that $|WX| \geq \frac{n}{m}$. For $k$-approval and veto, it is enough that $|WX| \geq 1$. Figure 1 shows tight examples of these bounds for plurality (left-hand side) and Borda (right-hand side), achieving the distortion bound of $2m-1$ for $m=4$ alternatives (these examples can easily be generalized to any $m \geq 3$). □

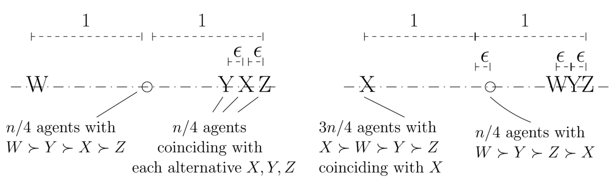


Figure 1: Examples showing tightness of distortion bound for plurality (left) and Borda (right) with $m=4$. Here $W$ is picked as winner and $X$ is the optimal alternative. As $\epsilon \to 0$, distortion approaches $2m-1$. Adding an extra agent coinciding with the center point makes $W$ the unique winner.

Theorem 4 implies that the distortion for plurality and Borda is unbounded in the number of candidates, and for $k$-approval and veto it is unbounded in the number of voters. Informally, this is because the optimal alternative can be preferred over the eventual winner by a relatively large fraction of the agents, and yet still lose. We now consider several social choice functions that escape this predicament, resulting in significantly better performance.

**Theorem 6** *For the Copeland social choice function, distortion is always* $\leq 5$*, and this bound is tight.*

**Proof Sketch.** We will give a general outline of the proof; see the full version for further details.

Let $W$ denote the winner under Copeland, and let $X$ denote the optimal alternative. We know that whenever $W$ pairwise defeats $X$, the distortion can be at most 3; hence we need to only consider the case when $X$ pairwise defeats $W$. In Copeland, this implies that there must exist an alternative $Y$ such that $W$ pairwise defeats $Y$ and $Y$ pairwise defeats $X$ (Moulin 1986). We will quantify precisely the level to which the existence of such an alternative $Y$ prevents the social cost of $X$ from being too small.

If $|WX|$ is large then the distortion cannot be too high, as seen from Lemma 5. Suppose instead that a large fraction of agents prefer $X$ over $W$, i.e., $|WX|$ is small. This together with the fact that $W$ pairwise defeats $Y$ implies that a non-trivial fraction of the agents are in $XWY$.

First, consider the easier case in which $d(X,Y) \geq d(X,W)$. Since $Y$ pairwise defeats $X$, at least half of the agents have a significant distance to $X$, i.e., $d(i,X) \geq d(X,Y)/2 \geq d(X,W)/2$. This implies that the cost of $X$ cannot be too small compared to the cost of $W$.

Now, consider the more difficult case when $d(X,Y) < d(X,W)$. Notice that the agents in $XWY$ cannot have $d(i,X) = 0$, because otherwise they would prefer $Y$ over $W$. In fact, for such agents $i$ we show that $d(i,X) \geq \frac{1}{2} \cdot (d(X,W) - d(X,Y))$. This lower bound, along with $d(i,X)$ lower bounds for agents who prefer $W$ or $Y$ over $X$, suffice to show that the social cost of $X$ must be large. While obtaining a loose upper bound using these insights is easy, combining this together to form a bound of 5 requires somewhat careful analysis: see the full version for further details. ■

**Remark:** *In fact, the result for distortion being at most* 5 *holds whenever for any other alternative* $Z$*, the winner* $W$ *either pairwise defeats* $Z$ *or there exists an alternative* $Y$ *whom* $W$ *pairwise defeats and* $Y$ *pairwise defeats* $Z$*. This precisely corresponds to the notion of* $W$ *being a member of the uncovered set (Moulin 1986). Thus the distortion is at most* 5 *for several notions of tournament winners other than Copeland such as the winner being selected from minimal covering set, bipartisan set, banks set, tournament equilibrium set, etc., as all these sets are a subset of the uncovered set (Laffond, Laslier, and Le Breton 1995).*

Recall that no social choice function can have distortion less than 3. Thus, Copeland is nearly optimal with a distortion of at most 5. We can show that the ranked pairs mechanism achieves the best possible distortion bound, but only in the special case when the majority graph (directed graph in which a link $(X,Y)$ denotes that $X$ pairwise defeats $Y$) has small circumference (i.e., maximum cycle size).

**Theorem 7** *The distortion of ranked pairs is* $\leq 3$*, as long as the majority graph has circumference* $\leq 4$*.*

## 4 Distortion of Median Agent Cost

In this section we study the distortion of social choice functions as measured by the median agent cost. We define the median social cost of an alternative $\text{SC}_{\text{med}}(Y,d) = \text{med}_{i \in N}\, d(i,Y)$ to be the median of the list of distances of all the agents to the alternative $Y$. If $n$ is even, we define it to be the $(\frac{n}{2}+1)^{th}$ smallest value of the distances. As a shorthand, we will refer to this as $\text{med}(Y)$ when the cost metric $d$ is fixed. The distortion of a social choice function is now $\text{dist}_{\text{med}}$ as defined in Section 2. We begin by establishing lower bounds on the distortion achieved by any deterministic social choice function; this bound is higher than in the sum case.

**Theorem 8** *No (deterministic) social choice function has worst-case distortion less than* $5$ *for the median objective.*

**Proof.** Suppose there are only three alternatives $W$, $X$, $Y$. Let there be $n/3$ agents corresponding to each of the preference rankings $W \succ Y \succ X$, $Y \succ X \succ W$ and $X \succ W \succ Y$. Without loss of generality, suppose that the given social choice function picks $W$ as the winner. Consider an underlying metric as shown in Figure 2. (The distances not shown in the figure can be chosen to be consistent with the metric and the preference profile). In this instance, we have $\text{med}(W) = 5 + \epsilon$ and $\text{med}(X) = 1 + \epsilon$. Thus, the distortion approaches 5 as $\epsilon \to 0$. ■

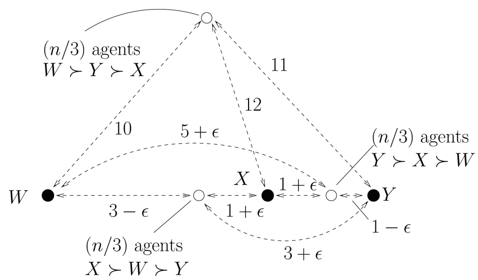


Figure 2: With median objective, $W$ being picked as the winning alternative leads to worst-case distortion arbitrarily close to 5 as $\epsilon \to 0$.

As for the sum objective function, the distortion of the common positional scoring rules remains high for the median objective. In fact, it becomes unbounded for any $m > 2$ number of alternatives.

**Theorem 9** *Plurality, Borda, $k$-approval, and veto have unbounded distortion for any number of alternatives* $m > 2$.

**Proof Sketch.** The same examples which are bad for the sum objective provide an unbounded distortion for the median objective. ■

Now we show that the Copeland social choice function achieves the optimal distortion bound: due to the lower bound in Theorem 8 no deterministic rule can have better median distortion than Copeland. Note that this result holds also for several other notions of tournament winners mentioned in the concluding remark in Section 3.

**Theorem 10** *For the Copeland social choice function, median distortion is always* $\leq 5$, *and this bound is tight.*

**Proof Sketch.** At first we proceed in the same way as in the proof of Theorem 6; let $W, X, Y$ be defined in the same way. The case when $\text{med}(W) > \frac{5}{4} \cdot d(X, W)$ is easy, since $d(i, X) \geq d(i, W) - d(X, W)$, and so the ratio between $\text{med}(W)$ and $\text{med}(X)$ cannot be too high. When instead $\text{med}(W) \leq \frac{5}{4} \cdot d(X, W)$, then we can show that

$$\text{med}(X) \geq \max\left(\frac{d(X,Y)}{2}, \frac{d(X,W) - d(X,Y)}{2}\right).$$

The first term in the above inequality is due to half of the agents preferring $Y$ over $X$, and the second to half of the agents preferring $W$ over $Y$. Finding the worst possible value for $\frac{\text{med}(W)}{\text{med}(X)}$ gives us the desired bound of 5. ■

### 4.1 Generalizing Median: Percentile Distortion

Instead of considering the happiness of the median voter or agent, it also makes sense to consider the happiness of the 25'th or 75'th percentile. We can generalize the median objective function $\text{med}(Y)$ above by using percentiles as follows. Let $\alpha$-PC$(Y)$ be the value from the set $\{d(i, Y) : i \in N\}$ below which lie an $\alpha$ fraction of the values. Thus $\alpha$-PC$(Y)$ with $\alpha = 1/2$ is the same as $\text{med}(Y)$. The distortion with $\alpha$-PC is defined analogously to Section 2.

For various ranges of $\alpha$, we now give lower bounds on the distortion that any social choice function must have in Theorem 11, and then give social choice functions that always achieve these bounds in Theorems 12 and 13.

**Theorem 11** *For any deterministic social choice function:*

*(a) For* $\alpha \in [\frac{2}{3}, 1)$, *worst-case $\alpha$-PC distortion is at least* $3$.

*(b) For* $\alpha \in [\frac{1}{2}, \frac{2}{3})$, *worst-case $\alpha$-PC distortion is at least* $5$.

*(c) For* $\alpha \in [0, \frac{1}{2})$, *worst-case distortion is unbounded.*

**Theorem 12** *For the plurality social choice function, distortion is always* $\leq 3$ *for $\alpha$-PC objective with* $\alpha \geq \frac{m-1}{m}$.

**Theorem 13** *For the Copeland social choice function, distortion is always* $\leq 5$ *for $\alpha$-PC objective with* $\frac{1}{2} \leq \alpha < 1$, *and this bound is tight.*

The proofs of Theorems 11 and 12 appear in the full version. For Theorem 13, the proof of Theorem 10 works verbatim with median replaced by $\alpha$-PC, and we show tightness in the full version. Together with the lower bound from Theorem 11, this shows that for $\alpha \geq \frac{m-1}{m}$, no deterministic rule can have better worst-case distortion than plurality, whereas Copeland achieves the optimal worst-case distortion for $\frac{1}{2} \leq \alpha < \frac{2}{3}$.

## 5 Conclusion and Future Directions

We analyzed the distortion of many common social choice mechanisms in the setting where the agent costs form a metric space. We showed that despite the process of winner determination having absolutely no extra information about the underlying metric space except the induced ordinal agent preferences, mechanisms like Copeland achieve a small constant-factor approximation to the optimal candidate (and in fact, for median objective function they achieve the best approximation to an optimal candidate that a deterministic mechanism can ever hope to achieve).

Nevertheless, some important open questions remain. Foremost among them is the question of a social choice rule which beats Copeland, and maybe achieves the best possible distortion of 3. While we showed some weaker results for the ranked pairs mechanism, we believe there is a good chance that it performs even better than we anticipate, and actually guarantees a distortion of 3 for all instances, not just the ones with small graph circumference. Exploring the space of randomized mechanisms could also be very fruitful. Randomized mechanisms still cannot get arbitrarily close to the optimal alternative (we can prove a lower bound of 2 on distortion instead of 3), but a small amount of randomization added to Copeland and ranked pairs has a chance to greatly improve their distortion properties.

## 6 Acknowledgments

This work began as a project in a class taught by Lirong Xia, and we would like to thank him greatly for many interesting discussions on the topic of social choice. This work was partially supported by NSF awards CCF-1101495 and CNS-1218374.

## References

Bartholdi III, J.; Tovey, C. A.; and Trick, M. A. 1989. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and welfare* 6(2):157–165.

Black, D. 1948. On the rationale of group decision-making. *The Journal of Political Economy* 23–34.

Borodin, A., and El-Yaniv, R. 1998. *Online computation and competitive analysis*. Cambridge University Press.

Boutilier, C.; Caragiannis, I.; Haber, S.; Lu, T.; Procaccia, A. D.; and Sheffet, O. 2012. Optimal social choice functions: A utilitarian view. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 197–214. ACM.

Branzei, S.; Caragiannis, I.; Morgenstern, J.; and Procaccia, A. 2013. How bad is selfish voting? In *AAAI*.

Caragiannis, I., and Procaccia, A. D. 2011. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence* 175(9):1655–1671.

Chakrabarty, D., and Swamy, C. 2014. Welfare maximization and truthfulness in mechanism design with ordinal preferences. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, 105–120.

Conitzer, V., and Sandholm, T. 2002. Vote elicitation: Complexity and strategy-proofness. In *AAAI/IAAI*, 392–397.

Elkind, E.; Faliszewski, P.; and Slinko, A. 2009. On distance rationalizability of some voting rules. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, 108–117.

Enelow, J. M., and Hinich, M. J. 1984. *The spatial theory of voting: An introduction*. CUP Archive.

Faliszewski, P., and Procaccia, A. D. 2010. Ai's war on manipulation: Are we winning? *AI Magazine* 31(4):53–64.

Gans, J. S., and Smart, M. 1996. Majority voting with single-crossing preferences. *Journal of Public Economics* 59(2):219–237.

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. ACM.

Harsanyi, J. C. 1976. *Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility*. Springer.

Laffond, G.; Laslier, J. F.; and Le Breton, M. 1995. Condorcet choice correspondences: A set-theoretical comparison. *Mathematical Social Sciences* 30(1):23–35.

Lang, J., and Xia, L. 2009. Sequential composition of voting rules in multi-issue domains. *Mathematical social sciences* 57(3):304–324.

Merrill, S., and Grofman, B. 1999. *A unified theory of voting: Directional and proximity spatial models*. Cambridge University Press.

Moulin, H. 1980. On strategy-proofness and single peakedness. *Public Choice* 35(4):437–455.

Moulin, H. 1986. Choosing from a tournament. *Social Choice and Welfare* 3(4):271–291.

Myerson, R. B. 1996. *Fundamentals of social choice theory*.

Ng, Y.-K. 1997. A case for happiness, cardinalism, and interpersonal comparability. *The Economic Journal* 107(445):1848–1858.

Oren, J., and Lucier, B. 2014. Online (budgeted) social choice. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*.

Procaccia, A. D., and Rosenschein, J. S. 2006. The distortion of cardinal preferences in voting. In *Cooperative Information Agents X*. Springer. 317–331.

Roemer, J. E. 1998. *Theories of distributive justice*. Harvard University Press.

Saporiti, A. 2009. Strategy-proofness and single-crossing. *Theoretical Economics* 4(2):127–163.

Sigletos, G.; Paliouras, G.; Spyropoulos, C. D.; and Hatzopoulos, M. 2005. Combining information extraction systems using voting and stacked generalization. *The Journal of Machine Learning Research* 6:1751–1782.

Sui, X.; Francois-Nienaber, A.; and Boutilier, C. 2013. Multi-dimensional single-peaked consistency and its approximations. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, 375–382. AAAI Press.