Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)

Randomized Social Choice Functions Under Metric Preferences

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

Abstract

We determine the quality of randomized social choice mechanisms in a setting in which the agents have *metric preferences*: every agent has a cost for each alternative, and these costs form a metric. We assume that these costs are unknown to the mechanisms (and possibly even to the agents themselves), which means we cannot simply select the optimal alternative, i.e. the alternative that minimizes the total agent cost (or median agent cost). However, we do assume that the agents know their ordinal preferences that are induced by the metric space. We examine randomized social choice functions that require only this ordinal information and select an alternative that is good in expectation with respect to the costs from the metric. To quantify how good a randomized social choice function is, we bound the *distortion*, which is the worst-case ratio between expected cost of the alternative selected and the cost of the optimal alternative. We provide new distortion bounds for a variety of randomized mechanisms, for both general metrics and for important special cases. Our results show a sizable improvement in distortion over deterministic mechanisms.

1 Introduction

The goal of social choice theory is usually to aggregate the preferences of many agents with conflicting interests, and produce an outcome that is suitable to the whole rather than to any particular agent. This is accomplished via a *social choice mechanism* which maps the preferences of the agents, usually represented as total orders over the set of alternatives, to a single winning alternative. There is no agreed upon "best" social choice mechanism; it is not obvious how one can even make this determination. Because of this, much of social choice literature is concerned with defining normative or axiomatic criteria, so that a social choice mechanism is "good" if it satisfies many useful criteria.

Another method of determining the quality of a social choice function is the *utilitarian* approach, which is often used in welfare economics and algorithmic mechanism design. Here agents have an associated utility (or cost, as in this paper) with each alternative that is a measure of how desirable (or undesirable) an alternative is to an agent. We can define the quality of an alternative to be a function of these agent utilities, for example as the sum of all agent utilities for a particular alternative. Other objective functions such as the median or max utility of the agents for a fixed alternative can be used as well. The utilitarian approach has received a lot of attention recently in the social choice literature ([Caragiannis and Procaccia, 2011; Filos-Ratsikas *et al.*, 2014; Harsanyi, 1976; Feldman *et al.*, 2015], see especially [Boutilier *et al.*, 2012] for a thorough discussion of this approach, its strengths, and its weaknesses).

A frequent criticism of the utilitarian approach is that it is unreasonable to assume that the mechanism, or even the agents themselves, know what their utilities are. Indeed, it can be difficult for an agent to quantify the desirability of an alternative into a single number, but there are arguments in favor of cardinal utilities [Boutilier *et al.*, 2012; Harsanyi, 1976]. Even if the agents were capable of doing this for each alternative, it could be difficult for us to elicit these utilities in order to compute the optimal alternative. It is much more reasonable, and much more common, to assume that the agents know the preference rankings induced by their utilities over the alternatives, however. That is, it might be difficult for an agent to express exactly how she feels about alternatives X and Y, but she should know if she prefers X to Y. Because of this, recent work considers how well social choice mechanisms can perform when they only have access to *ordinal preferences* of the agents, i.e., their rankings over the alternatives, instead of the true underlying (possibly latent) utilities [Procaccia and Rosenschein, 2006; Boutilier *et al.*, 2012; Caragiannis and Procaccia, 2011; Anshelevich *et al.*, 2015; Feldman *et al.*, 2015]. The *distortion* of a social choice function is defined here as the worst-case ratio of the cost of the alternative selected by the social choice function and the cost of the truly optimal alternative.

Our goal in this work is to design social choice mechanisms that minimize the worst-case distortion for the sum and median objective functions when the agents have *metric preferences* [Anshelevich *et al.*, 2015]. That is, we assume that the costs of agents over alternatives form an arbitrary *metric space* and that their preferences are induced by this metric space. Assuming such metric or spacial preferences is common [Enelow and Hinich, 1984], has a natural interpretation of agents liking candidates/alternatives which are most similar to them, such as in facility location literature [Campos Rodríguez and Moreno Pérez, 2008; Escoffier *et al.*, 2011; Feldman *et al.*, 2015], and our setting is sufficiently general that it does not impose any restrictions on the set of allowable preference profiles. Anshelevich *et al.* [2015] provide distortion bounds for this setting using well-known deterministic mechanisms such as plurality, Copeland, and ranked pairs. We improve on these results by providing distortion guarantees for *randomized social choice functions*, which output a probability distribution over the set of alternatives rather than a single winning alternative. We show that our randomized mechanisms perform better than any deterministic mechanism, and provide optimal randomized mechanisms for various settings.

We also examine the distortion of randomized mechanisms in important specialized settings. Many of our worst-case examples occur when many agents are indifferent between their top alternative and the optimal alternative. In many settings, however, agents are more *decisive* about their top choice, and prefer it much more than any other alternative. We introduce a formal notion of *decisiveness*, which is a measure of how strongly an agent feels about her top preference relative to her second choice. If an agent is very decisive, then she is very close to her top choice compared to her second choice in the metric space. In the extreme case, this means that the set of agents and alternatives is identical [Goel and Lee, 2012], as can occur for example when proposal writers rank all the other proposals being submitted, or when a committee must choose one of its members to lead it. We demonstrate that when agents are decisive, the distortion greatly improves, and quantify the relation between decisiveness and the performance of social choice mechanisms. Finally, we consider other natural special cases, such as when preferences are 1-Euclidean and when alternatives are vertices of a simplex. 1-Euclidean preferences are already recognized as a well-studied and well-motivated special case [Elkind and Faliszewski, 2014; Procaccia and Tennenholtz, 2009]. The setting in which alternatives form a simplex corresponds to the case in which alternatives share no similarities, i.e., when all alternatives are equally different from each other.

1.1 Our Contributions

In this paper, we bound the worst-case distortion of several randomized social choice functions in many different settings. Recall that the distortion is the worst-case ratio of the expected value of the alternative selected by the randomized mechanism and the optimal alternative. We use two different objective functions for the purpose of defining the quality of an alternative. The first is the sum objective, which defines the social cost of an alternative to be the sum of agent costs for that particular alternative. We also consider the median objective, which defines the quality of an alternative as the median agent's cost for that alternative.

A metric space is α-decisive if for every agent, the cost of her first choice is less than α times the cost of her second choice, for some $\alpha \in [0, 1]$. In other words, this provides a constraint on how indifferent an agent can be between her first and second choice. By definition, every metric space is 1-decisive. Considering α-decisive metric spaces allows us to immediately give results for important subcases, such as 0-decisive metric spaces in which every agent has distance 0 to her top alternative, i.e., every agent is also an alternative.

For the sum objective function, we begin by giving a lower bound of $1 + \alpha$ for *all* randomized mechanisms, which corresponds to a lower bound of 2 for general metric spaces. This is smaller than the lower bound of 3 for deterministic mechanisms from [Anshelevich *et al.*, 2015]. One of our first results is to show randomized dictatorship has worst-case distortion strictly better than 3, which is better than any possible deterministic mechanism. Furthermore, we show that a generalization of the "proportional to squares" mechanism is the optimal randomized mechanism when there are two alternatives, i.e., it has a distortion of $1 + \alpha$.

We also examine how well randomized mechanisms perform in important subcases. We consider the well-known case in which all agents and alternatives are points on a line with the Euclidean metric, known as 1-Euclidean preferences [Elkind and Faliszewski, 2014]. We give an algorithm, which heavily relies on proportional to squares, to achieve the optimal distortion bound of $1 + \alpha$ for any number of alternatives. We also consider a case first described by [Anshelevich *et al.*, 2015] known as the $(m - 1)$-simplex setting in which the alternatives are vertices of a simplex and the agents lie in the simplex. This corresponds to alternatives sharing no similarities. We are able to show that proportional to squares achieves worst-case distortion of $\frac{1}{2}\left(1 + \alpha + \sqrt{2}\sqrt{\alpha^2 + 1}\right)$, which is fairly close to the optimal bound of $1 + \alpha$. For details, see Section 3.2.

Our other major contribution is defining a new randomized mechanism for the median objective which achieves a distortion of 4 in arbitrary metric spaces (we call this mechanism *Uncovered Set Min-Cover*). This requires forming a very specific distribution over all alternatives in the uncovered set, and then showing that this distribution ensures that no alternative "covers" more than half of the total probability of all alternatives.

1.2 Related Work

Embedding the unknown cardinal preferences of agents into an ordinal space and measuring the *distortion* of social choice functions that operate on these ordinal preferences was first done by [Procaccia and Rosenschein, 2006]. Additional papers [Boutilier *et al.*, 2012; Caragiannis and Procaccia, 2011; Oren and Lucier, 2014; Anshelevich *et al.*, 2015; Feldman *et al.*, 2015] have since studied distortion and other related concepts of many different mechanisms with various assumptions about the utilities/costs of the agents. In this context, Anshelevich *et al.* [2015] introduced the notion of *metric preferences* which assumes the costs of the agents and alternatives form a metric.

Using mechanisms to select alternatives from a metric space when the true locations of agents is unknown is also reminiscent of facility location games [Campos Rodríguez and Moreno Pérez, 2008; Escoffier *et al.*, 2011]. However, we select only a single winning alternative in our setting, while in these papers, they select multiple facilities.

Assuming that the preferences of agents are induced by a metric is a type of spatial preference [Enelow and Hinich, 1984; Merrill and Grofman, 1999]. There are many other notions of spatial preferences that are prevalent in social choice, such as 1-Euclidean preferences [Elkind and Faliszewski, 2014; Procaccia and Tennenholtz, 2009], single-peaked preferences [Sui *et al.*, 2013], and single-crossing [Gans and Smart, 1996]. We consider 1-Euclidean preferences as an important special case of the metric preferences we study in this paper.

Randomized social choice was first studied in [Zeckhauser, 1969; Fishburn, 1972; Intriligator, 1973]. A similar setting was considered by [Fishburn and Gehrlein, 1977] in which agents are uncertain about their preferences and express their preferences using probability distributions. We consider several randomized mechanisms, such as randomized dictatorship [Chatterji *et al.*, 2014]. The use of randomized mechanisms is seen very frequently in literature concerning one-sided matchings. Random serial dictatorship and probabilistic serial are perhaps the most well-studied randomized mechanisms, and there is a significant amount of literature on them (e.g. [Bogomolnaia and Moulin, 2001; Aziz *et al.*, 2013; Aziz and Stursberg, 2014; Chakrabarty and Swamy, 2014; Christodoulou *et al.*, 2016; Filos-Ratsikas *et al.*, 2014]). In particular, the results of [Filos-Ratsikas *et al.*, 2014] are analogous to finding the distortion of one-sided matching mechanisms.

Finally, independently from us, Feldman *et al.* [2015] have also recently considered the distortion of randomized social choice functions. While they mostly focus on truthful mechanisms (i.e., the "strategic" setting), there is some intersection between our results. Specifically, Feldman *et al.* [2015] also give a bound of 3 (and a lower bound of 2) for arbitrary metric spaces in the sum objective, and also provide a mechanism with distortion 2 for the 1-Euclidean case. However, Feldman *et al.* [2015] do not consider either α-decisive voters or the median objective: showing better performance for decisive voters and designing better mechanisms for the median objective are two of our major contributions.

2 Preliminaries

Social Choice with Ordinal Preferences. Let $N = \{1, 2, \ldots, n\}$ be the set of agents, and let $M = \{A_1, A_2, \ldots, A_m\}$ be the set of alternatives. Let $\mathcal{S}$ be the set of all total orders on the set of alternatives M. 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 σ_i. We call the vector $\sigma = (\sigma_1, \ldots, \sigma_n) \in \mathcal{S}^n$ a *preference profile*. We say that an alternative X *pairwise defeats* Y if $|\{i \in N : X \succ_i Y\}| > \frac{n}{2}$. Furthermore, we use the following notation to describe sets of agents with particular preferences: $XY = \{i \in N : X \succ_i Y\}$ and $X^* = \{i \in N : X \succ_i Y \text{ for all } Y \neq X\}$.

Once we are given a preference profile, we want to aggregate the preferences of the agents and select a single alternative as the winner or find a probability distribution over the alternatives and pick a single winner according to that distribution. A *deterministic social choice function* $f : \mathcal{S}^n \to M$ is a mapping from the set of preference profiles to the set of alternatives. A *randomized social choice function* $f : \mathcal{S}^n \to \Delta(M)$ is a mapping from the set of preference profiles to the space of all probability distributions over the alternatives $\Delta(M)$. For example, the winning alternative may be selected according to *Randomized Dictatorship*, i.e., the probability of selecting an alternative Y is $p(Y) = \frac{|Y^*|}{n}$. We will also use heavily in this paper generalizations of the *Proportional to Squares* rule, i.e., $p(Y) = \frac{|Y^*|^2}{\sum_{Z \in M} |Z^*|^2}$.

Cardinal Metric Costs. In our work we take the utilitarian view, and study the case when the ordinal preferences σ are in fact a result of the underlying cardinal agent costs. Formally, we assume that there exists an arbitrary metric $d : (N \cup M)^2 \to \mathbb{R}_{\geq 0}$ on the set of agents and alternatives (or more generally a *pseudo-metric*, since we allow distinct agents and alternatives to be identical and have distance 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 σ is *consistent* with d if $\forall i \in N, \forall X, Y \in M$, 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 $\rho(d)$ denote the set of preference profiles consistent with d ($\rho(d)$ may include several preference profiles if the agent costs have ties). Similarly, we define $\rho^{-1}(\sigma)$ to be the set of metrics such that $\sigma \in \rho(d)$.

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, which is defined as $\mathrm{SC}(X, d) = \sum_{i \in N} d(i, X)$ for an alternative X. We also study the median objective function, $\mathrm{med}(X, d) = \mathrm{med}_{i \in N}(d(i, X))$. Since we have defined the cost of alternatives, we can now give the cost of an outcome of a deterministic social choice function f as $\mathrm{SC}(f(\sigma), d)$ or $\mathrm{med}(f(\sigma), d)$. For randomized functions, we define the cost of an outcome, which is a probability distribution over alternatives, as follows: $\mathrm{SC}(f(\sigma), d) = \mathbb{E}[\mathrm{SC}(X, d)] = \sum_{X \in M} p(X)\, \mathrm{SC}(X, d)$ and $\mathrm{med}(f(\sigma), d) = \mathbb{E}[\mathrm{med}(X, d)] = \sum_{X \in M} p(X)\, \mathrm{med}(X, d)$, where $p(X)$ is the probability of alternative X being selected, according to $f(\sigma)$. When the metric d is obvious from context, we will use $\mathrm{SC}(X)$ and $\mathrm{med}(X)$ as shorthand.

As described in the Introduction, we can view social choice mechanisms in our setting as attempting to find the optimal alternative (one that minimizes cost), but only having access to the ordinal preference profile σ, instead of the full underlying costs d. 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.

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

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

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

Decisive Voters. Many of our worst-case examples occur when many agents are indifferent between their top alternative and the optimal alternative. In many settings, however, agents are more *decisive* about their top choice, and prefer it much more than any other alternative. Formally, we say that an agent i whose top choice is W and second choice is X is α-*decisive* if $d(i,W) \le \alpha \cdot d(i,X)$ such that $\alpha \in [0,1]$. We say that a metric space is α-*decisive* if for some fixed α, every agent is α-*decisive*. Every metric space is 1-decisive, while a metric space in which every agent has distance 0 to her top alternative is 0-decisive. In fact, 0-decisive metric spaces are interesting in their own right: they correspond to the case when each voter must exactly coincide with some alternative, and so capture the settings where the set of voters and alternatives is the same.

3 Distortion of the Sum of Agent Costs

In this section, we examine the sum objective and provide mechanisms with low distortion. We begin by addressing the question of how well *any* randomized social choice function can perform. Our first theorem shows that no randomized mechanism can find an alternative that is in expectation within a factor strictly smaller than $1+\alpha$ from the optimum alternative for α-decisive metric spaces. Thus no mechanism can have distortion better than 2 for general metric spaces. We observe that 2 is better than the worst-case distortion lower bound of 3 for deterministic mechanisms from [Anshelevich *et al.*, 2015].

Theorem 1 *The worst-case distortion of any randomized mechanism when the metric space is α-decisive is at least $1+\alpha$.*

Proof Sketch. We will consider a preference profile with $m = 2$ alternatives W, X and n agents (n is even) where $\frac{n}{2}$ agents prefer W over X and $\frac{n}{2}$ agents prefer X over W. We will consider an α-decisive metric space that induces the preference profile and where X is optimal. All agents i who prefer X have $d(i,X) = 0$ and $d(i,W) = 1$. The remaining agents have $d(i,W) = \frac{\alpha}{1+\alpha}, d(i,X) = \frac{1}{1+\alpha}$. Thus, $\mathrm{SC}(X) = \frac{n}{2} \cdot \frac{1}{1+\alpha}$ and $\mathrm{SC}(W) = \frac{n}{2}(\frac{\alpha}{1+\alpha}+1)$. For any randomized mechanism, the distortion is $p(X)+p(W)(1+2\alpha)$, where $p(X), p(W)$ are the probabilities of selecting X and W, respectively. Similarly, there exists a metric with distortion $p(W)+p(X)(1+2\alpha)$. Thus, the worst-case distortion is minimized when $p(X) = p(W) = \frac{1}{2}$. ■

The following lemma allows us to quickly upper bound the worst-case distortion of randomized mechanisms by simply plugging in the probability that an alternative is selected. It is used crucially in the proofs of Theorems 3 and 4. Any missing proofs can be found in the full version on arXiv [Anshelevich and Postl, 2015].

Lemma 2 *For any instance σ, social choice function f, and α-decisive metric space,*

$$\mathrm{dist}_{\sum}(f,\sigma) \le 1 + \frac{(1+\alpha)\sum p(Y)(n - \frac{2}{1+\alpha}|Y^*|)d(X,Y)}{\sum_{Y\in M}|Y^*|d(X,Y)},$$

where X is the optimal alternative and $p(Y)$ is the probability that alternative Y is selected by f given profile σ.

The following theorem is our main result of this section. It states that in the worst case, the distortion of randomized dictatorship is strictly better than 3 (in fact, it is at most $3-\frac{2}{n}$, which occurs when $\alpha = 1, |W^*| = 1$ in the theorem below). Thus, this simple randomized mechanism has better distortion than *any* deterministic mechanism, since no deterministic mechanism can have distortion strictly better than 3 in the worst case, as shown by [Anshelevich *et al.*, 2015]. This is surprising because randomized dictatorship operates only on the first preferences of every agent: the full preference ranking is not required.

Theorem 3 *If a metric space is α-decisive, then the distortion of randomized dictatorship is $\le 2+\alpha-\frac{2|W^*|}{n}$, where $W = \arg\min_{Y\in M:|Y^*|>0}|Y^*|$, and this bound is tight.*

α-Generalized Proportional to Squares. While randomized dictatorship performs well, it still does not achieve the best possible distortion of $1+\alpha$ for randomized mechanisms. We will now define an optimal mechanism for α-decisive metric spaces when there are $m = 2$ alternatives, which is a generalization of proportional to squares that is parameterized by α. For $\alpha = 1$, the mechanism is in fact ordinary proportional to squares. In this mechanism, an alternative Y is selected with probability

$$p(Y) = \frac{(1+\alpha)|Y^*|^2 - (1-\alpha)|X^*||Y^*|}{(1+\alpha)(|X^*|^2+|Y^*|^2) - 2(1-\alpha)|X^*||Y^*|},$$

where X is the second alternative.

Theorem 4 *If a metric space is α-decisive and $m = 2$, then the distortion of α-generalized proportional to squares is $1+\alpha$, and this is tight.*

3.1 1-Euclidean Preferences

We now consider a well-known and well-studied special case of 1-Euclidean preferences [Elkind and Faliszewski, 2014; Procaccia and Tennenholtz, 2009] in which all agents and alternatives are on the real number line and the metric is defined to be the Euclidean distance. First, we observe that in this setting, a Condorcet winner (an alternative which pairwise defeats all others) always exists, so the distortion is at most 3, and this is tight for deterministic mechanisms. In designing an optimal randomized mechanism, we heavily use properties of this metric space from [Elkind and Faliszewski, 2014].

Namely, using only the preference profile, we can determine the ordering of the agents on the line (which is unique up to reversal and permutations of identical voters) and the unique ordering of the alternatives that are between the top preference of the first agent and the top preference of the last agent.

Algorithm 1 Optimal randomized mechanism for the α-decisive, 1-Euclidean space

Input: A preference profile σ
Output: A probability distribution over the alternatives
$>_N \leftarrow$ ordering of agents ([Elkind and Faliszewski, 2014])
$>_M \leftarrow$ ordering of alternatives ([Elkind and Faliszewski, 2014])
$i' \leftarrow$ median agent of $>_N$
$X \leftarrow$ top preference of i'
$Y \leftarrow$ alternative directly before X in $>_M$
$Z \leftarrow$ alternative directly after X in $>_M$
if $|YX| < |ZX|$ **then**
$p(Z) \leftarrow \frac{(1+\alpha)|ZX|^2-(1-\alpha)|XZ||ZX|}{(1+\alpha)(|XZ|^2+|ZX|^2)-2(1-\alpha)|XZ||ZX|}$
$p(X) \leftarrow \frac{(1+\alpha)|XZ|^2-(1-\alpha)|XZ||ZX|}{(1+\alpha)(|XZ|^2+|ZX|^2)-2(1-\alpha)|XZ||ZX|}$
else if $|YX| > |ZX|$ **then**
$p(Y) \leftarrow \frac{(1+\alpha)|YX|^2-(1-\alpha)|XY||YX|}{(1+\alpha)(|XY|^2+|YX|^2)-2(1-\alpha)|XY||YX|}$
$p(X) \leftarrow \frac{(1+\alpha)|XY|^2-(1-\alpha)|XY||YX|}{(1+\alpha)(|XY|^2+|YX|^2)-2(1-\alpha)|XY||YX|}$
else
$p(X) \leftarrow 1$
end if

We show that this mechanism has worst-case distortion at most $1+\alpha$ through a series of steps in which we reduce the set of possible optimal alternatives from m to 2. First, we claim that the optimal alternative must be one of the two alternatives on either side of the median agent from our agent ordering. One of these alternatives must be the top preference X of the median agent. However, since we do not know if the median agent's top preference is to the left or right of it, we must consider three alternatives: her top preference X and the two alternatives to the left and right of X, which we call Y, Z, respectively. This reduces our set of possible optimal alternatives from m to 3. The following lemma further reduces the possible optimal alternatives to 2.

Lemma 5 *If $|YX| \leq |ZX|$, then Y cannot be better than X, and if $|ZX| \leq |YX|$, Z cannot be better than X.*

Finally, we can use the α-generalized proportional to squares mechanism on the restricted set of alternatives X and one of Y, Z to achieve a distortion of $1+\alpha$, which is tight since our lower bound example from Theorem 1 occurs in the 1-Euclidean setting.

Theorem 6 *In the 1-Euclidean setting, Algorithm 1 has distortion at most $1+\alpha$, and thus has the best possible worst-case distortion.*

3.2 Distortion in the $(m-1)$-Simplex

In this section we consider a specialized, yet natural, setting inspired by [Anshelevich *et al.*, 2015], known as the $(m-1)$-*simplex setting*. In this setting, we assume that the m alternatives are all distance 1 from each other and for every agent i, $\forall Y \in M$, $d(i,Y) \leq 1$. This includes the case when m alternatives are the vertices of the $(m-1)$-simplex and all of the agents lie inside this simplex. Although this is a very constrained setting, it is a reasonable assumption in the case when all of the alternatives are uncorrelated, i.e., when all the alternatives are equally different from one another. In this setting, the distortion of randomized dictatorship does not improve because the worst case occurs on a line.

Theorem 7 *If the $(m-1)$-simplex setting is α-decisive, then plurality has distortion $\leq 1+2\alpha$.*

Plurality, although a deterministic mechanism, does very well in this setting, because for $\alpha = 0$ the alternative with the most votes is clearly optimal. In general, as $\alpha \to 0$, the agents are forced closer to the vertices of the simplex, and plurality better approximates finding the optimal alternative. However, when α is not small, plurality fares poorly compared to the proportional to squares mechanism. Indeed, for $\alpha \leq \frac{1}{7}$, plurality is better than proportional to squares, but otherwise the opposite is true.

Theorem 8 *If the $(m-1)$-simplex setting is α-decisive, the proportional to squares mechanism has distortion $\leq \frac{1}{2}\left(1+\alpha+\sqrt{2}\sqrt{\alpha^2+1}\right)$.*

Unlike all of the previous distortion bounds we have provided, this is the first that is not linearly increasing in α. It increases slower than the distortion of plurality, which is $1+2\alpha$. For smaller values of α, such as $\alpha = 0$, which is where plurality has the largest advantage over proportional to squares, the distortion of proportional to squares is still at most $\frac{1+\sqrt{2}}{2} \approx 1.2071$, which is reasonably small. For $1 \geq \alpha \geq .5$, the values of $1+\alpha$ and $\frac{1}{2}\left(1+\alpha+\sqrt{2}\sqrt{\alpha^2+1}\right)$ are relatively close. This implies that proportional to squares is nearly optimal for sufficiently large values of α.

4 Median Agent Cost

In this section, we will examine the median objective function. Anshelevich *et al.* [2015] show that no deterministic mechanism can achieve a worst-case distortion of better than 5, and that the Copeland mechanism achieves this bound. We begin this section by showing that randomized mechanisms have a general worst-case distortion lower bound of 3 rather than 5 like deterministic mechanisms.

Theorem 9 *For $m \geq 2$, the worst-case median distortion is at least 3 for all randomized mechanisms.*

Unfortunately, it is not difficult to show that both randomized dictatorship and proportional to squares have unbounded distortion, even for $m = 2$. As a result, we focus on designing a randomized mechanism which will always achieve a distortion of at most 4 for the median objective. We claim that to design a randomized mechanism for the median objective, it makes sense to consider the uncovered set, which is the set of alternatives X that pairwise defeats every other alternative Y either directly (i.e., X pairwise defeats Y) or indirectly through another alternative Z (i.e., X does not pairwise defeat Y, but X pairwise defeats Z, which in turn pairwise defeats Y). We rely on the following two lemmas concerning the quality of alternatives in the uncovered set.

Lemma 10 ([Anshelevich *et al.*, 2015]) *If an alternative W pairwise defeats (or pairwise ties) the alternative X, then* $\mathrm{med}(W) \leq 3 \cdot \mathrm{med}(X)$.

Lemma 11 ([Anshelevich *et al.*, 2015]) *If an alternative W is in the uncovered set, then* $\mathrm{med}(W) \leq 5 \cdot \mathrm{med}(X)$*, where X is any alternative.*

Thus, we want to mix over the entire uncovered set and ensure that some alternatives that pairwise defeat the optimal alternative (i.e., alternatives only a factor of 3 away) are chosen with high probability to decrease the distortion. However, since we do not know the optimal alternative, we must have this property hold for every alternative. This is made precise in the following theorem. Let $G = (M, E)$ be the majority graph, i.e., a graph in which the alternatives are vertices and the edges denote pairwise victories: an edge $(Y, Z) \in E$ if Y is preferred to Z by a strict majority of the voters. Let S be the uncovered set, and p be some probability distribution over S. Finally, define $\pi(Y)$ for any alternative Y to be the total probability distribution "covered" by Y, i.e., $\pi(Y) = \sum_{(Y,Z)\in E} p(Z)$. Then, we have that

Theorem 12 *If a mechanism selects alternatives only from the uncovered set S with probability distribution p, and if for all alternatives X we have that $\pi(X) \leq \frac{1}{2}$, then the expected median distortion of this mechanism is at most 4.*

Proof Sketch. Let X be the optimal alternative. By Lemmas 10 and 11, we know expected distortion is at most $\pi(X) \cdot 5 + (1 - \pi(X)) \cdot 3 = 3 + \pi(X) \cdot 2 \leq 4$. ■

Thus, we want a mechanism that manages to ensure that for every alternative X, the alternatives that can be more than a factor of 3 away from X (i.e., the ones it pairwise defeats) are selected with probability at most $\frac{1}{2}$. The mechanism we describe, Uncovered Set Min-Cover, uses a linear program to accomplish this. We define the subset of the edges on the uncovered set as $E(S) = \{E = (Y, Z) : Y \in S, Z \in S\}$. We also give the LP (and its dual which is not used by the algorithm, but is necessary for our proofs), which is used as a subroutine by our mechanism.

(Linear Program)

$$\begin{aligned} \text{minimize} \quad & p_{\max} \\ \text{subject to} \quad & p_Y \geq 0,\ Y \in S \\ & \sum_{\substack{(Y,Z)\\ \in E(S)}} p_Z \leq p_{\max}, Y \in S \\ & \sum_{Z \in S} p_Z = 1. \end{aligned}$$

(Dual)

$$\begin{aligned} \text{maximize} \quad & b_{\min} \\ \text{subject to} \quad & b_Y \geq 0,\ Y \in S \\ & \sum_{\substack{(Z,Y)\\ \in E(S)}} b_Z \geq b_{\min}, Y \in S \\ & \sum_{Z \in S} b_Z = 1. \end{aligned}$$

Theorem 13 *The expected median distortion of Uncovered Set Min-Cover is* ≤ 4.

This theorem is immediate from Theorem 12 if we can show that for the distribution formed by Uncovered Set Min-Cover, we have that $\pi(X) \leq \frac{1}{2}$ for all X. We prove this fact using the following two lemmas, the proofs of which can be found in the full version of this paper.

Algorithm 2 Uncovered Set Min-Cover

Input: A preference profile σ
Output: A probability distribution $\mathbf{p}$ over the alternatives of the uncovered set
  $G = (M, E) \leftarrow$ majority graph of σ
  $S \leftarrow$ uncovered set of G
  $\mathbf{p} \leftarrow$ solution to LP (see above)

Lemma 14 *Let $G = (M, E)$ be the majority graph in which ties are broken arbitrarily. For the dual of LP, there does not exist $\mathbf{b}$ such that for all $Y \in S$, $\sum_{Z\in S:(Z,Y)\in E} b_Z > \frac{1}{2}$, i.e., it must be that $b_{\min} \leq \frac{1}{2}$.*

Due to LP-Duality, the above lemma immediately implies that $p_{\max} \leq \frac{1}{2}$, and thus $\pi(X) \leq \frac{1}{2}$ for all $X \in S$. This does not complete the proof of Theorem 13, however, since it is possible that the optimal alternative X is outside of the uncovered set S.

Lemma 15 *Suppose we have a probability distribution p over alternatives in the uncovered set S, and for all $Y \in S$, we have that $\pi(Y) = \sum_{(Y,Z)\in E} p(Z) \leq \frac{1}{2}$. Then, this also must hold for alternatives outside of S, i.e., for all $X \notin S$, we also have that $\pi(X) \leq \frac{1}{2}$.*

The proof of the above lemma crucially relies on properties of the uncovered set S. This completes the proof of Theorem 13: by Lemma 14 we have that $\pi(X) \leq \frac{1}{2}$ for all $X \in S$, by Lemma 15 we have that this is true even for $X \notin S$, and by Theorem 12 we obtain the desired distortion bound.

Finally, we consider special metric spaces. For 1-Euclidean, we are trivially able to obtain an optimal mechanism: selecting the Condorcet winner, since it is guaranteed to exist for the 1-Euclidean setting. By Lemma 10, we know that the Condorcet winner is guaranteed to be within a factor of 3 of the optimal alternative.

In the $(m-1)$-simplex setting, almost any alternative is of high median quality. This is due to the fact that the alternatives are very spread out. Unless an alternative has at least $\frac{n}{2}$ agents very close to it, its median cost is guaranteed to be at least $\frac{1}{2}$ in general metric spaces. The following result shows than any mechanism which selects an alternative preferred by at at least $\frac{n}{2}$ voters as their top choice (if one exists) will have low median distortion. Thus, for example, plurality is a good mechanism for this setting.

Theorem 16 *If the $(m-1)$-simplex setting is α-decisive, any mechanism that satisfies the majority criterion has median distortion at most $1 + \alpha$.*

5 Conclusion

Many open questions still remain. While we were able to show that proportional to squares is an optimal mechanism for $m = 2$ alternatives in the sum setting, our best known mechanism for arbitrary m is randomized dictatorship, which has a distortion arbitrarily close to 3 in the worst case. We suspect there may exist a generalization of proportional to squares that is able to achieve a distortion of 2, but it is likely significantly more complex and may require the full preference profile instead of agents' top preferences.

Acknowledgements

We owe great thanks to Edith Elkind for many interesting and informative discussions on the topic of social choice. This work was supported in part by NSF awards CCF-1101495, CNS-1218374, and CCF-1527497.

References

[Anshelevich and Postl, 2015] Elliot Anshelevich and John Postl. Randomized social choice functions under metric preferences. *CoRR*, abs/1512.07590, 2015.

[Anshelevich *et al.*, 2015] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. Approximating optimal social choice under metric preferences. In *Proc. 29th AAAI Conf. Artificial Intell.*, pages 777–783. AAAI Press, 2015.

[Aziz and Stursberg, 2014] Haris Aziz and Paul Stursberg. A generalization of probabilistic serial to randomized social choice. In *Proc. 28th AAAI Conf. Artificial Intell.*, pages 559–565. AAAI Press, 2014.

[Aziz *et al.*, 2013] Haris Aziz, Felix Brandt, and Markus Brill. On the tradeoff between economic efficiency and strategyproofness in randomized social choice. In *Proc. 2013 Int. Conf. Autonomous Agents and Multi-agent Syst.*, pages 455–462, 2013.

[Bogomolnaia and Moulin, 2001] Anna Bogomolnaia and Hervé Moulin. A new solution to the random assignment problem. *J. Econ. Theory*, 100(2):295 – 328, 2001.

[Boutilier *et al.*, 2012] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 197–214. ACM, 2012.

[Campos Rodríguez and Moreno Pérez, 2008] Clara M. Campos Rodríguez and José A. Moreno Pérez. Multiple voting location problems. *Eur. J. Oper. Res.*, 191(2):437–453, December 2008.

[Caragiannis and Procaccia, 2011] Ioannis Caragiannis and Ariel D. Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intell.*, 175(9):1655–1671, June 2011.

[Chakrabarty and Swamy, 2014] Deeparnab Chakrabarty and Chaitanya Swamy. Welfare maximization and truthfulness in mechanism design with ordinal preferences. In *Proc. 5th Conf. Innovations in Theoretical Comput. Sci.*, pages 105–120. ACM, 2014.

[Chatterji *et al.*, 2014] Shurojit Chatterji, Arunava Sen, and Huaxia Zeng. Random dictatorship domains. *Games and Econ. Behavior*, 86:212–236, July 2014.

[Christodoulou *et al.*, 2016] George Christodoulou, Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, Paul W. Goldberg, Jie Zhang, and Jinshan Zhang. Social welfare in one-sided matching mechanisms. *CoRR*, abs/1502.03849, February 2016.

[Elkind and Faliszewski, 2014] Edith Elkind and Piotr Faliszewski. Recognizing 1-euclidean preferences: An alternative approach. In *Proc. 7th Int. Symp. Algorithmic Game Theory*, pages 146–157. Springer, 2014.

[Enelow and Hinich, 1984] James M. Enelow and Melvin J. Hinich. *The Spatial Theory of Voting: An Introduction*. Cambridge University Press, New York, NY, 1984.

[Escoffier *et al.*, 2011] Bruno Escoffier, Laurent Gourves, Nguyen Kim Thang, Fanny Pascual, and Olivier Spanjaard. Strategy-proof mechanisms for facility location games with many facilities. In *Proc. 2nd Int. Conf. Algorithmic Decision Theory*, pages 67–81. Springer, 2011.

[Feldman *et al.*, 2015] Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. *CoRR*, abs/1512.05868, December 2015.

[Filos-Ratsikas *et al.*, 2014] Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, and Jie Zhang. Social welfare in one-sided matchings: Random priority and beyond. In *Proc. 7th Int. Symp. Algorithmic Game Theory*, pages 1–12. Springer, 2014.

[Fishburn and Gehrlein, 1977] Peter C. Fishburn and William V. Gehrlein. Towards a theory of elections with probabilistic preferences. *Econometrica*, 45(8):1907–1924, 1977.

[Fishburn, 1972] Peter C. Fishburn. Lotteries and social choices. *J. Econ. Theory*, 5(2):189 – 207, 1972.

[Gans and Smart, 1996] Joshua S. Gans and Michael Smart. Majority voting with single-crossing preferences. *J. Public Econ.*, 59(2):219–237, February 1996.

[Goel and Lee, 2012] Ashish Goel and David Lee. Triadic consensus. In *Proc. 8th Workshop Internet and Network Econ.*, pages 434–447. Springer, 2012.

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

[Intriligator, 1973] Michael D. Intriligator. A probabilistic model of social choice. *Rev. Econ. Stud.*, 40(4):553–560, October 1973.

[Merrill and Grofman, 1999] Samuel Merrill and Bernard Grofman. *A unified theory of voting: Directional and proximity spatial models*. Cambridge University Press, 1999.

[Oren and Lucier, 2014] Joel Oren and Brendan Lucier. Online (budgeted) social choice. In *Proc. 28th AAAI Conf. Artificial Intell.*, pages 1456–1462. AAAI Press, 2014.

[Procaccia and Rosenschein, 2006] Ariel D. Procaccia and Jeffrey S. Rosenschein. The distortion of cardinal preferences in voting. In *Proc. 10th Int. Workshop Cooperative Inform. Agents X*, pages 317–331. Springer, 2006.

[Procaccia and Tennenholtz, 2009] Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. In *Proc. 10th ACM Conf. Electron. Commerce*, pages 177–186. ACM, 2009.

[Sui *et al.*, 2013] Xin Sui, Alex Francois-Nienaber, and Craig Boutilier. Multi-dimensional single-peaked consistency and its approximations. In *Proc. 23rd Int. Joint Conf. Artificial Intell.*, pages 375–382. AAAI Press, 2013.

[Zeckhauser, 1969] Richard Zeckhauser. Majority rule with lotteries on alternatives. *Quart. J. Econ.*, 83(4):696–703, 1969.