Peeking Behind the Ordinal Curtain: Improving Distortion via Cardinal Queries

Georgios Amanatidis,[1] Georgios Birmpas,[2] Aris Filos-Ratsikas,[3] Alexandros A. Voudouris[2]
[1]Department of Computer, Control and Management Engineering, Sapienza University of Rome
amanatidis@diag.uniroma1.it
[2]Department of Computer Science, University of Oxford
{georgios.birmpas, alexandros.voudouris}@cs.ox.ac.uk
[3]Department of Computer Science, University of Liverpool
aris.filos-ratsikas@liverpool.ac.uk

Abstract

The notion of *distortion* was introduced by Procaccia and Rosenschein (2006) to quantify the inefficiency of using only ordinal information when trying to maximize the social welfare. Since then, this research area has flourished and bounds on the distortion have been obtained for a wide variety of fundamental scenarios. However, the vast majority of the existing literature is focused on the case where nothing is known beyond the ordinal preferences of the agents over the alternatives. In this paper, we take a more expressive approach, and consider mechanisms that are allowed to further ask *a few cardinal queries* in order to gain partial access to the underlying values that the agents have for the alternatives. With this extra power, we design new *deterministic* mechanisms that achieve significantly improved distortion bounds and outperform the best-known randomized ordinal mechanisms. We draw an almost complete picture of the number of queries required to achieve specific distortion bounds.

1 Introduction

Social choice theory (Brandt et al. 2016) is concerned with aggregating the preferences of individuals into a joint decision. In an election, for instance, the winner should represent well (in some precise sense) the viewpoints of the voters. Similarly, the expenditure of public funds is typically geared towards projects that increase the well-being of society. Most traditional models assume that the preferences of individuals are expressed through *ordinal preference rankings*, where each agent sorts all alternatives from the most to the least preferred. Underlying these ordinal preferences, it is often assumed that there exists a *cardinal* utility structure, which further specifies the intensity of the preferences. That is, there exist numerical values that indicate how much an agent prefers an outcome to another. Given this cardinal utility structure, usually expressed via *valuation functions*, one can define meaningful quantitative objectives, with the most prominent one being the *social welfare*, i.e., the total value of the agents for the chosen outcome.

The main rationale justifying the dominance of ordinal preferences in the classical economics literature is that the task of asking individuals to express their preferences in terms of numerical values is arguably quite demanding. In contrast, performing simple comparisons between the different options is more easily conceivable. To quantify how much the lack of cardinal information affects the maximization of social welfare, Procaccia and Rosenschein (2006) defined the notion of *distortion* as the worst-case ratio between the optimal social welfare (achievable by using all cardinal information) and the social welfare of the outcome, which is chosen having access only to ordinal information. Following their agenda, a plethora of subsequent works studied the distortion in different settings.

Somewhat surprisingly, these variants of the distortion framework studied in this rich line of work differentiate between two extremes: we either have complete cardinal information or only ordinal information. Driven by the original motivation for using ordinal preferences, it seems quite meaningful to ask whether improved distortion guarantees can be obtained if one has access to *limited* cardinal information, especially in settings for which the worst-case distortion bounds are already quite discouraging (Boutilier et al. 2015). We formulate this idea via *cardinal queries*, which elicit cardinal information from the agents. These queries can be as simple as asking the value of an agent for a possible outcome, or even asking an agent *whether an outcome is at least x times better than some other outcome*. Note that questions of the latter form are much less demanding than eliciting a complete cardinal utility structure, and thus are much more realistic as an elicitation device.

In this paper, we enhance the original distortion setting of Procaccia and Rosenschein (2006) and Boutilier et al. (2015) on single winner elections, by allowing the use of cardinal queries. In this setting, there are n agents that have cardinal values over m alternatives, and the goal is to elect a single alternative that (approximately) maximizes the social welfare, while having access *only* to ordinal information. Caragiannis et al. (2017) proved that no deterministic mechanism can achieve a distortion better that $\Omega(m^2)$ when agents have *unit-sum* normalized valuation functions (i.e., the sum of the values of each agent for all alternatives is 1). Under this assumption, Boutilier et al. (2015) proved that the distortion of any randomized mechanism is between

$\Omega(\sqrt{m})$ and $O(\sqrt{m} \cdot \log^* m)$. Here we show how—with only a limited number of cardinal queries, on top of the ordinal information—deterministic mechanisms can significantly outperform any mechanism that has access only to ordinal information, even randomized ones.

Our Contributions

We initiate the study of trade-offs between the number of cardinal queries per agent that a mechanism uses on top of the ordinal information, and the distortion that it can achieve. In particular, we show results of the following type:

> *The distortion $\mathcal{D}(\mathcal{M})$ of a mechanism $\mathcal{M}$ that makes at most λ queries per agent is $O(g(m, \lambda))$.*

Our results suggest that the distortion can be drastically reduced by exploiting only minimal cardinal information.

Query Model. We consider two different types of cardinal queries, namely *value queries* and *comparison queries*.

- A *value query* takes as input an agent i and an alternative j, and returns the agent's value for that alternative.
- A *comparison query* takes as input an agent i, two alternatives j, ℓ and a real number d and returns "yes" if the value of agent i for alternative j is at least d times her value for alternative ℓ, and "no" otherwise.

Note that value queries are in general stronger than comparison ones, as they reveal much more detailed information. On the other hand, comparison queries are quite attractive as an elicitation device, since the cognitive complexity of the question they pose is not much higher than that of forming preference rankings.

Results and Techniques. We devise a class of sophisticated mechanisms that achieve much improved trade-offs between the distortion and the number of queries. In particular, our class contains a mechanism that achieves *constant* distortion using at most $O(\log^2 m)$ value queries per agent, and a mechanism that achieves a distortion of $O(\sqrt{m})$ using $O(\log m)$ value queries. This matches the performance of the best possible randomized ordinal mechanism, and actually outperforms all known ones. Surprisingly, under the standard unit-sum normalization assumption,[1] we can approximate each agent's value for her favorite alternative using only $O(\log^2 m)$ comparison queries, and hence we can achieve constant distortion with only comparison queries at no extra cost, asymptotically. These results are in Section 3.

Next, in Section 4, we significantly improve on our $O(\sqrt{m})$ result for the fundamental case $n = \Theta(m)$. We present a mechanism that achieves this bound using *only* 2 value queries per agent. Finally, in Section 5, we present several lower bounds on the possible achievable trade-offs between the number of queries and distortion. An overview of our main results can be found in Table 1.

[1]We remark here that all of our other upper bounds hold *without any normalization assumption on the cardinal values*.

Related Work

The distortion has been studied extensively in numerous settings, such as normalized valuations (Caragiannis and Procaccia 2011; Boutilier et al. 2015; Filos-Ratsikas and Miltersen 2014; Filos-Ratsikas, Micha, and Voudouris 2019), metric preferences (Anshelevich et al. 2018; Anshelevich and Postl 2017; Goel, Krishnaswamy, and Munagala 2017; Fain et al. 2019; Pierczynski and Skowron 2019; Munagala and Wang 2019), committee elections (Caragiannis et al. 2017; Bhaskar, Dani, and Ghosh 2018), participatory budgeting (Goel et al. 2019; Benade et al. 2017), and matching and facility location (Filos-Ratsikas, Frederiksen, and Zhang 2014; Anshelevich and Sekar 2016; Feldman, Fiat, and Golomb 2016; Anshelevich and Zhu 2018).

Very recently, Mandal et al. (2019) study a model in which agents are asked to provide cardinal information, but there is a restriction on the number of bits to be communicated. Hence, they study trade-offs between the number of transmitted bits and distortion. This is markedly quite different from what we do here, as a query in their setting has access to the (approximate) values of an agent for many alternatives simultaneously, and is therefore much too expressive when translated to our setting. On the other hand, their setting does not assume "free" access to the ordinal preferences. We consider our work complementary to theirs, as they are mostly motivated by the computational limitations of elicitation (corresponding to a communication complexity approach), whereas we are motivated by the cognitive limitations of eliciting cardinal values (corresponding to a query complexity approach).

Finally, at the same time and independently of our work, Abramowitz, Anshelevich, and Zhu (2019) also introduce a setting in which the mechanism designer has access to some cardinal information on top of the ordinal preferences. This enables the design of improved mechanisms in terms of distortion. While the motivation of their paper is the same as ours, the two approaches are inherently different. Besides the fact that they study a metric distortion setting, the access to the cardinal information in their paper is not via queries, but is given explicitly as part of the input.

2 The model

In the standard *utilitarian* social choice setting, there is a set A of m alternatives and a set N of n agents. Our goal is to elect a *single* alternative based on the preferences of the agents, which are expressed through *valuation functions* $v_i : A \to \mathbb{R}_{\geq 0}$ that map alternatives to non-negative real numbers. For notational convenience, we use v_{ij} instead of $v_i(j)$ to denote the *cardinal value* of agent i for alternative j, and refer to the matrix $\mathbf{v} = (v_{ij})_{i \in N, j \in A}$ as a *valuation profile*. By $\mathbf{V}$ we denote the set of all possible valuation profiles. Clearly, the valuation function v_i also defines a *preference ranking* for agent i, i.e., a linear ordering $\succ_i$ of A such that $j \succ_i j'$ if $v_{ij} \geq v_{ij'}$. We assume that ties are broken according to a deterministic tie-breaking rule, e.g., according to a fixed global ordering of the alternatives.[2] We refer

[2]It would be equivalent to allow ties, get pre-linear orderings instead, and leave the tie-breaking to the mechanisms.

Number of queries	Upper Bounds	Lower Bounds
0 (ordinal, deterministic)	$O(m^2)$ (Caragiannis and Procaccia 2011)	$\Omega(m^2)$ (Caragiannis et al. 2017)
0 (ordinal, randomized)	$O(\sqrt{m}\log^* m)$ (Boutilier et al. 2015)	$\Omega(\sqrt{m})$ (Boutilier et al. 2015)
1 (value)	$O(m)$ [1-PRV, Theorem 1]	$\Omega(m)$ [Theorem 5] $\Omega(\sqrt{m})$ [Theorem 7]
$\lambda \geq 2$, constant (value)	$O(\sqrt{m})$ [$\sqrt{m}$-TRV, Theorem 4]*	$\Omega(m^{1/2(\lambda+1)})$ [Corollary 3]
$O\left(\frac{\log m}{\log\log m}\right)$ (value)	$O(\sqrt{m})$ [$\sqrt{m}$-TRV, Theorem 4]*	$\Omega(\log\log m)$ [Corollary 3]
$O(\log m)$ (value)	$O(\sqrt{m})$ [$O(1)$-ARV, Corollary 1]	$\Omega(1)$
$O(\log^2 m)$ (value)	$O(1)$ [$O(\log m)$-ARV, Corollary 1]	$\Omega(1)$
$O(\log^2 m)$ (comparison)	$O(1)$ [$O(\log m)$-ARV, Corollary 2]	$\Omega(1)$

Table 1: A table showing the most important results in the paper. All our results are for *deterministic* mechanisms. Results marked by $\star$ hold for $n = \Theta(m)$. Results for unit-sum valuation functions are highlighted; everything else is for unrestricted valuation functions.

to $\succ_{\mathbf{v}} = (\succ_1, \ldots, \succ_n)$ as an (ordinal) *preference profile*.

In this work, we consider the following two families of valuation functions:

- *Unrestricted valuation functions*, which may take any non-negative real values.
- *Unit-sum valuation functions*, which are such that $\sum_{j\in A} v_{ij} = 1$ for every agent $i \in N$.

The *social welfare* of alternative $j \in A$ with respect to $\mathbf{v}$ is the total value of the agents for j: $\text{SW}(j|\mathbf{v}) = \sum_{i\in N} v_{ij}$. Our goal is to output one of the alternatives who maximize the social welfare, i.e., an alternative in $\arg\max_{j\in A} \text{SW}(j|\mathbf{v})$. This is clearly a trivial task if one has full access to the valuation profile. However, we assume *limited access* to these cardinal values. In particular, we assume that we only have access to the preference profile $\succ_{\mathbf{v}}$ and can also learn cardinal information by asking queries. We consider two types of queries: *value queries* that reveal the value of an agent for a given alternative, and *comparison queries* that reveal whether the value of an agent for an alternative is a multiplicative factor larger than her value for some other alternative.

Definition 1. Given a preference profile, a query about the underlying cardinal values is called

- A *value query*, if it takes as input an agent i and an alternative j and returns the agent's value v_{ij} for that alternative. This is implemented via the function $\mathcal{V} : N \times A \to \mathbb{R}_{\geq 0}$. We say that *agent i is queried at position k*, if alternative j is ranked k-th in $\succ_i$ and we make the query $\mathcal{V}(i, j)$.
- A *comparison query*, if it takes as input an agent i, two alternatives j, ℓ and a real number d, and returns YES if $v_{ij} \geq d \cdot v_{i\ell}$, and NO otherwise. This is implemented via the function $\mathcal{C} : N \times A \times A \times \mathbb{R}_{\geq 0} \to \{\text{YES}, \text{NO}\}$.

Note that the information obtained by a comparison query can be obtained by at most two value queries. On the other hand, however, without any cardinal information or any normalization assumption, it is impossible to even approximate the information obtained by a value query using only comparison queries. In this sense, value queries are considerably stronger than comparison queries.

Definition 2. A *mechanism* $\mathcal{M} = (\mathcal{Q}, f)$ with access to a (value or comparison) oracle takes as input a preference profile $\succ_{\mathbf{v}}$ and returns an alternative. In particular, it consists of the following two parts:

- An algorithm $\mathcal{Q}$ that takes as input the preference profile $\succ_{\mathbf{v}}$, adaptively makes queries to the oracle, and returns the set of answers to these queries.
- A mapping f that takes as input the preference profile $\succ_{\mathbf{v}}$ and the set $\mathcal{Q}(\succ_{\mathbf{v}})$ of answers to the queries above, and outputs a single alternative $j \in A$. Such a mapping is called a *social choice function*.

By the description of $\mathcal{Q}$ above, it is clear that the mechanism is free to choose the positions at which each agent will be queried, and those can depend not only on $\succ_{\mathbf{v}}$, but on the answers to the queries already asked as well. The performance of a mechanism is measured by its *distortion*.

Definition 3. The *distortion* of a mechanism $\mathcal{M}$ is

$$\mathcal{D}(\mathcal{M}) = \sup_{\mathbf{v}\in\mathbf{V}} \frac{\max_{j\in A}\text{SW}(j|\mathbf{v})}{\text{SW}(\mathcal{M}(\succ_{\mathbf{v}})|\mathbf{v})}.$$

Due to space constraints, some details have been omitted and can be found in the full version (Amanatidis et al. 2019).

3 Achieving Constant Distortion

Before we dive into the main result of this section, let us first discuss probably the most obvious idea, to query each agent at the first $\lambda \geq 1$ positions; we refer to such queries as *prefix*. Among such mechanisms, we show that *λ-Prefix Range Voting* (λ-PRV), which elects the alternative who maximizes the total value of the agents according to the answers to the queries, is asymptotically the best possible.

Theorem 1. *The distortion of* λ-PRV *is* $O(m/\lambda)$*, even for unrestricted valuation functions, and is best possible among mechanisms that make prefix queries.*

Observe that λ-PRV is a good first step, but it needs to make a large number of queries per agent in order to provide good improvements on the distortion. In particular, it achieves distortion $O(\sqrt{m})$ for $\lambda = \Theta(\sqrt{m})$ and constant distortion for $\lambda = \Theta(m)$. It is natural to ask whether it is possible to design mechanisms that achieve similar distortion bounds, but require fewer queries per agent. We manage to answer this question positively.

For any $k \in [m]$, we define a mechanism which we call *k-Acceptable Range Voting* (k-ARV). Let $\lambda_1, \ldots, \lambda_k$ be k thresholds such that $\lambda_\ell = m^{\frac{\ell}{k+1}}$ for $\ell \in [k]$. For every agent $i \in N$, we first query her value v_i^* for her favorite alternative $j_i(1)$. Then, using binary search we compute the maximal *λ_ℓ-acceptable set* $S_{i,\ell} = \{j \in A : v_{ij} \geq v_i^*/\lambda_\ell\}$ for every $\ell \in [k]$. The λ_ℓ-acceptable set consists of the alternatives that the agent finds at most λ_ℓ times worse than her favorite. We define $S_{i,0} = \{j_i(1)\}$ to contain only the favorite alternative of the agent. We continue by constructing a new approximate valuation profile $\tilde{\mathbf{v}}$, where the values of every agent i are

- $\tilde{v}_i^* = v_i^*$;
- $\tilde{v}_{ij} = v_i^*/\lambda_\ell$ for every $j \in S_{i,\ell} \setminus S_{i,\ell-1}$ with $\ell \in [k]$;
- $\tilde{v}_{ij} = 0$ for every $j \in A \setminus S_{i,k}$.

We finally elect the alternative $z \in A$ that maximizes the social welfare according to the approximate valuation profile.

Theorem 2. *The mechanism* k-ARV *makes* $O(k \log m)$ *value queries per agent, and has distortion* $O(\sqrt[k+1]{m})$.

Proof. Consider any instance with valuation profile $\mathbf{v}$. Since mechanism k-ARV executes a binary search in order to compute the λ_ℓ-acceptable sets for each $\ell \in [k]$, it requires a total of $O(k \log m)$ value queries per agent. The rest of the proof is dedicated to bounding the distortion of k-ARV. First, we define some useful notation:

- z is the alternative elected by k-ARV;
- y is a welfare-maximizing alternative for the valuation profile $\hat{\mathbf{v}}$, which is such that the value of agent $i \in N$ for alternative $j \in A$ is
$$\hat{v}_{ij} = \begin{cases} 0, & \text{if } j \in A \setminus S_{i,k} \\ v_{ij}, & \text{otherwise.} \end{cases}$$
That is, $y \in \arg\max_{j \in A} \sum_{i \in N} \hat{v}_{ij}$.
- x is the welfare-maximizing alternative for the true valuation profile $\mathbf{v}$. That is, $x \in \arg\max_{j \in A} \sum_{i \in N} v_{ij}$.

Also, let $N_j(\mathbf{v}) = \{i \in N : v_{ij} > 0\}$ be the set of agents with strictly positive value for alternative $j \in A$. We use the following easy fact about welfare-maximizing alternatives.

Fact 1. *If* $j \in \arg\max_{j \in A} \sum_{i \in N} v_{ij}$*, then* $j \in \arg\max_{j \in A} \sum_{i \in N_j(\mathbf{v})} v_{ij}$.

To prove the statement, we will bound the social welfare of x in terms of the social welfare of z for the true valuation profile $\mathbf{v}$. In particular, we will show that

$$\text{SW}(x|\mathbf{v}) \leq \left(\lambda_1 + \frac{m}{\lambda_k}\right) \text{SW}(z|\mathbf{v}) \,. \tag{1}$$

Then, the approximation ratio of k-ARV will be

$$\frac{\text{SW}(x|\mathbf{v})}{\text{SW}(z|\mathbf{v})} \leq \lambda_1 + \frac{m}{\lambda_k} = 2 \cdot m^{\frac{1}{k+1}} = O(\sqrt[k+1]{m}) \,.$$

We partition the social welfare of x into the following two quantities: the contribution of the agents that place x in the λ_k-acceptable set $S_{i,k}$, and the contribution of the remaining agents that have small value for x. By definition, we have that $i \in N_x(\hat{\mathbf{v}})$ for any agent i such that $x \in S_{i,k}$, and thus

$$\text{SW}(x|\mathbf{v}) = \sum_{i \in N_x(\hat{\mathbf{v}})} v_{ix} + \sum_{i \notin N_x(\hat{\mathbf{v}})} v_{ix}$$

We first consider the term $\sum_{i \in N_x(\hat{\mathbf{v}})} v_{ix}$, and have that

$$\begin{aligned} \sum_{i \in N_x(\hat{\mathbf{v}})} v_{ix} &\leq \sum_{i \in N_y(\hat{\mathbf{v}})} v_{iy} \leq \lambda_1 \cdot \sum_{i \in N_y(\hat{\mathbf{v}})} \tilde{v}_{iy} \\ &\leq \lambda_1 \cdot \sum_{i \in N_z(\tilde{\mathbf{v}})} \tilde{v}_{iz} \leq \lambda_1 \cdot \sum_{i \in N_z(\tilde{\mathbf{v}})} v_{iz} \\ &\leq \lambda_1 \cdot \text{SW}(z|\mathbf{v}) \,, \end{aligned} \tag{2}$$

where the first inequality follows by the definition of y, the simple fact that $N_x(\hat{\mathbf{v}}) = N_y(\hat{\mathbf{v}})$, and Fact 1; for the second inequality it suffices to notice that for any $i \in N_y(\hat{\mathbf{v}})$ there exists an $\ell \in [k]$ such that $y \in S_{i,\ell} \setminus S_{i,\ell-1}$, and thus $v_{ij} \leq \frac{v_i^*}{\lambda_{\ell-1}} = \lambda_1 \cdot \frac{v_i^*}{\lambda_\ell} = \lambda_1 \cdot \tilde{v}_{ij}$; the third inequality follows by the definition of z, the simple fact that $N_y(\hat{\mathbf{v}}) = N_z(\tilde{\mathbf{v}})$, and Fact 1; the fourth inequality follows by $v_{ij} \geq \tilde{v}_{ij}$, for any $i \in N$, $j \in A$; the last inequality is obvious.

Next, we consider the term $\sum_{i \notin N_x(\hat{\mathbf{v}})} v_{ix}$. By the definition of $N_x(\hat{\mathbf{v}})$, for every $i \notin N_x(\hat{\mathbf{v}})$ it holds that $x \notin S_{i,k}$, and hence $v_{ix} < v_i^*/\lambda_k$. Using this, we obtain

$$\begin{aligned} \sum_{i \notin N_x(\hat{\mathbf{v}})} v_{ix} &< \sum_{i \notin N_x(\hat{\mathbf{v}})} \frac{v_i^*}{\lambda_k} = \frac{1}{\lambda_k} \sum_{i \notin N_x(\hat{\mathbf{v}})} v_i^* \\ &\leq \frac{1}{\lambda_k} \sum_{j \in A \setminus \{x\}} \sum_{i \in T_1(j)} v_{ij} \,, \end{aligned} \tag{3}$$

where $T_1(j)$ is the set of agents whose favorite alternative is j, and for whom $v_i^* = \tilde{v}_i^* = v_{ij} = \tilde{v}_{ij}$. Since z it the alternative that maximizes the quantity $\sum_{i \in N} \tilde{v}_{ij}$, for every $j \neq z$ we have that

$$\sum_{i \in N} \tilde{v}_{iz} \geq \sum_{i \in N} \tilde{v}_{ij} = \sum_{i \in T_1(j)} v_{ij} + \sum_{i \in N \setminus T_1(j)} \tilde{v}_{ij} \geq \sum_{i \in T_1(j)} v_{ij} \,.$$

Combining the above inequality together with the fact that $v_{iz} \geq \tilde{v}_{iz}$ for every agent $i \in N$, we have that

$$\sum_{i \in N} v_{iz} \geq \sum_{i \in T_1(j)} v_{ij} \,.$$

Using this last inequality, (3) becomes

$$\begin{aligned} \sum_{i \notin N_x(\hat{\mathbf{v}})} v_{ix} &\leq \frac{1}{\lambda_k} \sum_{j \in A \setminus \{x\}} \sum_{i \in T_1(j)} v_{ij} \leq \frac{1}{\lambda_k} \sum_{j \in A \setminus \{x\}} \sum_{i \in N} v_{iz} \\ &= \frac{m-1}{\lambda_k} \text{SW}(z|\mathbf{v}). \end{aligned} \tag{4}$$

Finally, the desired inequality (1) follows by combining inequalities (2) and (4). ∎

The next statement follows by appropriately setting the value of the parameter k in Theorem 2, and shows how mechanism k-ARV improves upon the distortion guarantees of λ-PRV using fewer value queries per agent.

Corollary 1. *We have that*

- 1-ARV *achieves distortion* $O(\sqrt{m})$ *using* $O(\log m)$ *values queries per agent;*
- $(\log m)$-ARV *achieves distortion* $O(1)$ *using* $O(\log^2 m)$ *value queries.*

Implementing *k*-ARV with comparison queries

A crucial observation is that mechanism k-ARV can actually be implemented using just *one* value query. We can ask the value of each agent for her favorite alternative, and then ask $O(k \log m)$ comparison queries that guide the binary search in computing the maximal acceptable sets. Hence, $(\log m)$-ARV achieves constant distortion using only one value query and $O(\log^2 m)$ comparison queries.

It is natural to ask whether we can avoid this single value query entirely, and rely only on comparison queries instead. Surprisingly, for unit-sum valuation functions, we show that this is indeed possible at no extra cost! More precisely, we show that we can approximate the value of an agent for her favorite alternative within a factor of $1 \pm \varepsilon$, using $O(\log^2 m)$ comparison queries.

For the sake of readability, we focus on a single agent and write u_j for her value for the alternative that she ranks at position $j \in [m]$. We take the same approach as in the proof of Theorem 2 in order to build an approximate valuation profile. Since everything in this profile is expressed in terms of the largest value u_1, we utilize the unit-sum assumption to approximately solve for u_1.

Theorem 3. *For any constant* $\varepsilon \geq 1/m$, *it is possible to compute some* u^* *such that* $(1-\varepsilon)\, u^* \leq u_1 \leq (1+\varepsilon)\, u^*$, *using* $O(\log^2 m)$ *comparison queries.*

By inspecting the proof of Theorem 2, it is easy to see that knowing the approximate valuation profile $\tilde{\mathbf{v}}$ exactly or perturbed within a multiplicative constant factor, makes no difference asymptotically. Therefore, we augment k-ARV with a pre-processing step where each maximum value v_i^* is approximated according to Theorem 3 above. For $k = \log m$, this new mechanism, which we call *modified* $(\log m)$-ARV, achieves the same distortion guarantee and asks the same number of queries (asymptotically) with $(\log m)$-ARV.

Corollary 2. *Modified* $(\log m)$-ARV *achieves distortion* $O(1)$ *using* $O(\log^2 m)$ *comparison queries per agent.*

4 Achieving $\sqrt{m}$ with Two Value Queries

Here we present a more sophisticated mechanism, which makes two value queries per agent. The first query is used to learn the value of each agent for her favorite alternative. However, we would like to avoid making a naive second query as we do with 2-PRV. Ideally, we would like to ask each agent about an alternative that is qualitatively similar to the one identified by 1-ARV. In other words, we would like to reveal for each agent the position where her value is roughly $1/\sqrt{m}$ of that for her favorite alternative. Although maintaining the same guarantee as 1-ARV, while substituting each binary search with a single query seems far-fetched, we do come very close. By utilizing the available ordinal information globally rather than per agent, our mechanism achieves distortion $O(\sqrt{m})$ with just two value queries, under the assumption that $n = \Theta(m)$. The crucial idea is that the second query for each agent depends on the number of appearances of the alternatives in the whole ordinal profile.

For any *threshold* $\tau \in \mathbb{N}$ we define a mechanism called τ-*Threshold Range Voting* (τ-TRV). As noted above, the first query for each agent is used to ask about her favorite alternative. The remaining queries are made in successive steps. During the ℓ-th step we make queries about alternatives that are ranked within the first ℓ positions by at least τ agents. These queries are made *only* if they are meaningful and possible: we never repeat a query and never ask an agent more than twice. After at most m steps, τ-TRV returns an alternative with maximum *revealed welfare*.

Theorem 4. *The mechanism* τ-TRV *has distortion*

$$\mathcal{D}(\tau\text{-}TRV) = \begin{cases} O(m) & \text{when } n = \omega(m^2), \text{ for } \tau = 1 \\ O(\sqrt{n}) & \text{when } n = O(m^2), \text{ for } \tau = \sqrt{n} \\ O(\sqrt{m}) & \text{when } n = \Theta(m), \text{ for } \tau = \sqrt{m}. \end{cases}$$

Proof. We prove the case $n = m$ for which the distortion is $O(\sqrt{m})$; see the full version for the full proof. Consider any instance with valuation profile $\mathbf{v}$. Let y be the alternative elected by $\sqrt{m}$-TRV, and let x be an optimal alternative. By $\text{SW}_r(z|\mathbf{v})$ we denote the total value of the agents for alternative $z \in A$, among those that have been queried for z; we refer to this quantity as the *revealed welfare* of z. By the definition of the mechanism, $\text{SW}_r(y|\mathbf{v}) = \max_{z \in A} \text{SW}_r(z|\mathbf{v})$. Since $\text{SW}(y|\mathbf{v}) \geq \text{SW}_r(y|\mathbf{v})$, to prove the theorem, it suffices to show that $\text{SW}(x|\mathbf{v}) \leq (1 + 2\sqrt{m}) \cdot \text{SW}_r(y|\mathbf{v})$. We have that

$$\text{SW}(x|\mathbf{v}) = \text{SW}_r(x|\mathbf{v}) + \text{SW}_c(x|\mathbf{v}),$$

where $\text{SW}_c(x|\mathbf{v})$ is the *concealed welfare* of x, that is, the total value of the agents for x, among those that have not been queried for x. By the discussion above, we have that

$$\text{SW}(x|\mathbf{v}) \leq \text{SW}_r(y|\mathbf{v}) + \text{SW}_c(x|\mathbf{v}).$$

Next, we focus on bounding the quantity $\text{SW}_c(x|\mathbf{v})$. Let $\mathcal{E}_\ell$ be the set of *eligible alternatives* at step ℓ, for which there are at least $\sqrt{m}$ agents who rank them in the first ℓ positions.

Let $\ell^* \in \{2, ..., m\}$ be such that $x \in \mathcal{E}_{\ell^*}$ and $x \notin \mathcal{E}_{\ell^*-1}$; in other words, ℓ^* is the smallest index for which alternative x became an eligible alternative. By definition, we have that $x \in \mathcal{E}_\ell \Rightarrow x \in \mathcal{E}_{\ell+1}$, for any $\ell \in [m-1]$. Thus, we can further partition the concealed welfare of x as

$$\text{SW}_c(x|\mathbf{v}) = \text{SW}_c^{<\ell^*}(x|\mathbf{v}) + \text{SW}_c^{\geq \ell^*}(x|\mathbf{v}),$$

where $\text{SW}_c^{<\ell^*}(x|\mathbf{v})$ is the contribution of agents who rank x at some position $\ell < \ell^*$, and $\text{SW}_c^{\geq \ell^*}(x|\mathbf{v})$ is the contribution of agents who rank x at some position $\ell \geq \ell^*$.

Note that since $x \notin \mathcal{E}_\ell$ for any $\ell \in \{2, ..., \ell^* - 1\}$, there are strictly less than $\sqrt{m}$ agents who rank x before position

ℓ^*. Therefore, we have that

$$\text{SW}_c^{<\ell^*}(x|\mathbf{v}) < \sqrt{m} \cdot \max_{i,j}\{v_{ij}\} \leq \sqrt{m} \cdot \text{SW}_r(y|\mathbf{v}).$$

Finally, we consider the quantity $\text{SW}_c^{\geq\ell^*}(x|\mathbf{v})$. Let i be an agent who is not queried for x, ranks x at some position $k \geq \ell^*$, and is queried by the mechanism for the second time at some step $\ell \in \{\ell^*, ..., m\}$. Since $x \in \mathcal{E}_{\ell^*}$, it must be the case that $k > \ell$. Hence, there exists an alternative b such that i ranks b at position ℓ and $v_{ib} \geq v_{ix}$. Let T be the set of all alternatives b, for whom there exists at least one such agent i. Since there are $n = m$ agents and any alternative $b \in T$ has to be eligible when queried for, it must be $|T| \leq \sqrt{m}$.

For every alternative $b \in T$, let S_b be the set of agents that are queried for b instead of x and contribute to $\text{SW}_c^{\geq\ell^*}(x|\mathbf{v})$. Clearly, since y maximizes the revealed social welfare, we have that $\sum_{i \in S_b} v_{ib} \leq \text{SW}_r(y|\mathbf{v})$. We now obtain

$$\begin{aligned}\text{SW}_c^{\geq\ell^*}(x|\mathbf{v}) &= \sum_{b\in T}\sum_{i\in S_b} v_{ix} \leq \sum_{b\in T}\sum_{i\in S_b} v_{ib} \\ &\leq |T| \cdot \text{SW}_r(y|\mathbf{v}) \leq \sqrt{m} \cdot \text{SW}_r(y|\mathbf{v}).\end{aligned}$$

The bound follows. □

Of course, the cases of Theorem 4 are neither exhaustive nor disjoint, yet they all deserve some attention. However, due to space constraints, we focused on the highlight of this section, i.e., on the fact that, when $n = \Theta(m)$, with only two value queries per agent we deterministically match the lower bound for any randomized ordinal mechanism, and beat the best known randomized ordinal mechanism of Boutilier et al. (2015) which achieves a distortion of $O(\sqrt{m} \cdot \log^* m)$.

5 Lower Bounds

In this section we explore the limitations of query-based mechanisms, by presenting lower bounds on their distortion. These bounds depend only on the number of value queries the mechanisms are allowed to ask per agent, and are unconditional on how and where these queries are made.

Before we proceed with the presentation of the results of this section, let us give a high-level idea of the constructions used in the proofs. Assuming an arbitrary mechanism (that is allowed to make a specific number of queries per agent), we first define a single ordinal preference profile which is given as input to the mechanism. Then we carefully define the cardinal information that is revealed from all possible queries the mechanism could make. This cardinal information is such that it is always possible to complete the valuation profile in a way that leads the social welfare of the optimal alternative to be much higher than that of the alternative selected by the mechanism. Since we do not know how the mechanism makes its query selection, we need to take into account every possible scenario, and therefore define many different valuation profiles to capture different cases.

We start by showing that, for unrestricted valuations, any mechanism that makes one value query per agent has linear distortion. This also shows that the straightforward mechanism 1-PRV from Section 3 is the best possible mechanism among such mechanisms.

Theorem 5. *For unrestricted valuation functions, the distortion of any mechanism that uses one value query per agent is* $\Omega(m)$.

Proof. Let $\mathcal{M}$ be an arbitrary mechanism that makes one value query per agent, and consider an instance with $m \geq 4$ alternatives and $n = m - 2$ agents, where m is an even number. We denote the set of alternatives as $A = \{a_1, ..., a_{m-2}, x, y\}$. Using the notation $[z, w]$ to denote the fact that alternatives z and w are ordered arbitrarily in the ranking of an agent, we define the ordinal profile as follows. The ranking of agent $i \leq \frac{n}{2}$ is

$$a_i \succ_i x \succ_i y \succ_i [a_1, ..., a_{i-1}, a_{i+1}, ..., a_{m-2}],$$

while the ranking of agent $i > \frac{n}{2}$ is

$$a_i \succ_i y \succ_i x \succ_i [a_1, ..., a_{i-1}, a_{i+1}, ..., a_{m-2}].$$

Depending on the positions at which $\mathcal{M}$ queries, we reveal the following cardinal information: For every query at a first position we reveal a value of m^{-1}; for every query at a second or third position we reveal a value of m^{-2}; for any other position we reveal a value of 0.

We claim that $\mathcal{M}$ must query all agents at the first position, as otherwise its distortion is $\Omega(m)$. Assume that $\mathcal{M}$ does not query agent 1 for alternative a_1; this is without loss of generality due to symmetry. We now define two valuation profiles $\mathbf{v}_1$ and $\mathbf{v}_2$, which are both consistent with the ordinal profile and the revealed information, but differ on the value that agent 1 has for alternative a_1. In particular:

- In both $\mathbf{v}_1$ and $\mathbf{v}_2$, every agent $i \geq 2$ has value m^{-1} for alternative a_i, m^{-2} for alternatives x and y, and 0 for everyone else;
- In both $\mathbf{v}_1$ and $\mathbf{v}_2$, agent 1 has value m^{-2} for alternatives x and y, and 0 for every alternative a_i for $i \geq 2$. The value of agent 1 for alternative a_1 is m^{-2} in $\mathbf{v}_1$, and 1 in $\mathbf{v}_2$.

These two profiles are utilized in the following way: If $\mathcal{M}$ selects a_1, then the valuation profile is set to be $\mathbf{v}_1$, while if $\mathcal{M}$ selects some other alternative, then the valuation profile is set to be $\mathbf{v}_2$. Now, observe that

$$\text{SW}(a_i|\mathbf{v}_1) = \text{SW}(a_i|\mathbf{v}_2) = m^{-1}$$

for every $i \geq 2$, and

$$\begin{aligned}\text{SW}(x|\mathbf{v}_1) = \text{SW}(x|\mathbf{v}_2) = \text{SW}(y|\mathbf{v}_1) = \text{SW}(y|\mathbf{v}_2) \\ = (m-2) \cdot m^{-2} \leq m^{-1}.\end{aligned}$$

If $\mathcal{M}$ selects a_1, the social welfare of a_1 is $\text{SW}(a_1|\mathbf{v}_1) = m^{-2}$ and therefore any alternative a_i for $i \geq 2$ is optimal, yielding distortion m. Similarly, when $\mathcal{M}$ selects some alternative different than a_1, then a_1 is optimal with social welfare $\text{SW}(a_1|\mathbf{v}_2) = 1$, yielding distortion at least m.

Hence, $\mathcal{M}$ must query all agents at the first position in order to learn a value of m^{-1} for every alternative a_i, $i \in [n]$. We now define three valuation profiles $\mathbf{v}_3$, $\mathbf{v}_4$ and $\mathbf{v}_5$, which are consistent with the ordinal profile and this revealed information, but differ on the values that the agents have for alternatives x and y; in particular, $\mathbf{v}_4$ and $\mathbf{v}_5$ are symmetric.

- In all three profiles, every agent $i \in [n]$ has value m^{-1} for alternative a_i, and 0 for any alternative a_j such that $j \neq i$;
- In $\mathbf{v}_3$, all agents have value m^{-1} for alternatives x and y;
- In $\mathbf{v}_4$, all agents have value m^{-2} for alternative y, every agent $i > n/2$ (who ranks x after y) has value m^{-2} for x, and every agent $i \leq n/2$ (who ranks x before y) has value m^{-1} for x.
- In $\mathbf{v}_5$, all agents have value m^{-2} for alternative x, every agent $i \leq n/2$ (who ranks y after x) has value m^{-2} for y, and every agent $i > n/2$ (who ranks y before x) has value m^{-1} for y.

If $\mathcal{M}$ selects some alternative a_i for $i \in [n]$, then the valuation profile is set to be $\mathbf{v}_3$, while if $\mathcal{M}$ selects alternative y or x, then the valuation profile is set to be $\mathbf{v}_4$ or $\mathbf{v}_5$, respectively. Given this, observe that if $\mathcal{M}$ decides to select alternative a_i for some $i \in [n]$, then since

$$\text{SW}(a_i|\mathbf{v}_3) = \text{SW}(a_i|\mathbf{v}_4) = \text{SW}(a_i|\mathbf{v}_5) = m^{-1},$$

for every $i \in [n]$, and

$$\text{SW}(x|\mathbf{v}_3) = \text{SW}(y|\mathbf{v}_3) = (m-2) \cdot m^{-1} = 1 - 2m^{-1},$$

the distortion is at least $m-2$. Similarly, if $\mathcal{M}$ decides to select alternative y, then since

$$\text{SW}(y|\mathbf{v}_4) = (m-2)m^{-2} \leq m^{-1}$$

and

$$\begin{aligned}\text{SW}(x|\mathbf{v}_4) &= \left(\frac{m}{2} - 1\right) m^{-1} + \left(\frac{m}{2} - 1\right) m^{-2} \\ &= \frac{1}{2}\left(1 - m^{-1} - 2m^{-2}\right),\end{aligned}$$

the distortion is at least $\frac{1}{2}(m-1-2m^{-1}) \geq \frac{m}{4}$ for any $m \geq 4$. The case where $\mathcal{M}$ selects x is symmetric and follows by $\mathbf{v}_5$. In any case, $\mathcal{M}$ has distortion $\Omega(m)$. □

We will now focus on mechanisms that make a number $\lambda \geq 1$ of queries per agent, and will show a weaker lower bound on the distortion which depends on λ. The construction that gives us the next theorem is more delicate and the proof is significantly more technical.

Theorem 6. *For unrestricted valuation functions, the distortion of any mechanism that uses $\lambda \geq 1$ value queries per agent is $\Omega\left(\frac{1}{\lambda+1} \cdot m^{\frac{1}{2(\lambda+1)}}\right)$.*

Using Theorem 6, we can deduce several lower bounds.

Corollary 3. *For unrestricted valuation functions, the distortion of a mechanism $\mathcal{M}$ that uses λ queries per agent is*

$$\mathcal{D}(\mathcal{M}) = \begin{cases} \Omega\left(m^{\frac{1}{2(\lambda+1)}}\right), & \textit{for any constant } \lambda \geq 1 \\ \Omega\left(\log\log m\right), & \textit{for } \lambda = O\left(\frac{\log m}{\log\log m}\right). \end{cases}$$

Next, we turn our attention to unit-sum valuation functions. Coming up with constructions that satisfy the very restricted structure of such valuation functions and at the same time capture *all* mechanisms is quite challenging. In the following, we consider mechanisms that are allowed to make only one value query per agent. For this case, we are able to show a weaker lower bound of $\Omega(\sqrt{m})$, which indicates (but does not prove) some separation between unrestricted and unit-sum valuation functions.

Theorem 7. *For unit-sum valuation functions, the distortion of any mechanism that uses only one value query per agent is $\Omega(\sqrt{m})$.*

6 Conclusions and Future Directions

We studied mechanisms for general single winner elections, and we explored the potential of improving their distortion by making a *limited* number of cardinal queries per agent. On this front, we obtained a definitive positive answer. Among other results, we showed that it is possible to achieve *constant* distortion by making $O(\log^2 m)$ value or comparison queries, while only *two* value queries are enough to guarantee distortion $O(\sqrt{m})$ when $n = \Theta(m)$, thus outperforming the best known randomized ordinal mechanism.

Quite interestingly, our positive results for value queries hold without any normalization assumption, which makes them even stronger. On top of that, by carefully inspecting the proofs of our upper bounds, one can easily observe that it is not actually necessary for the agents to be able to answer value queries *exactly*. Our arguments follow through even when the queries are *noisy*, as long as the answers are at most a (multiplicative) constant factor away from the truth. Finally, we complemented these results by showing (nearly) tight lower bounds for many interesting cases.

Possibly the most obvious open problem is to fill in the gaps between our upper and lower bounds. To this end, we make the following two conjectures.

1-Query Conjecture. *There exists a mechanism that achieves a distortion of $O(\sqrt{m})$ using one value query per agent, for unit-sum valuation functions.*

$\log m$-Queries Conjecture. *There exists a mechanism that achieves a constant distortion, using $O(\log m)$ value queries per agent, even for unrestricted valuation functions.*

We consider settling these two conjectures the most interesting problems left open in our work. Since our upper bounds for value queries do not make use of the unit-sum normalization, it is conceivable that some clever use of that extra information could possibly lead to better trade-offs.

Another natural direction is to consider randomized mechanisms. Actually, one could study two different levels of randomization: mechanisms that decide randomly what queries to make, but select the winner deterministically, and mechanisms that use randomization for both querying and making the final decision.

Our work takes a first step towards exploring how powerful ordinal mechanisms with limited access to cardinal information can actually be. Of course, the same idea can be applied to many different contexts, such as participatory budgeting, multi-winner elections, or the metric distortion setting, which has been extensively studied over the past years. As we mentioned in the introduction, Abramowitz, Anshelevich, and Zhu (2019) already take a step in this direction in the metric setting.

Acknowledgments

G.A. was supported by the ERC Advanced Grant AMDROMA (No. 788893), the MIUR PRIN project ALGADI-

MAR, and the NWO Gravitation Project NETWORKS (No. 024.002.003). G.B. was supported by the ERC Advanced Grant ALGAME (No. 321171) and the ERC Grant ACCORD (No. 639945). A.F.-R. was supported by the Swiss National Science Foundation under contract No. 200021_165522. A.V. was supported by the ERC Grant ACCORD (No. 639945).

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 Conference on Web and Internet Economics (WINE)*.

Amanatidis, G.; Birmpas, G.; Filos-Ratsikas, A.; and Voudouris, A. A. 2019. Peeking behind the ordinal curtain: Improving distortion via cardinal queries. *CoRR* abs/1907.08165.

Anshelevich, E., and Postl, J. 2017. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research* 58:797–827.

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

Anshelevich, E., and Zhu, W. 2018. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. In *Proceedings of the 14th International Conference on Web and Internet Economics (WINE)*, 3–20.

Anshelevich, E.; Bhardwaj, O.; Elkind, E.; Postl, J.; and Skowron, P. 2018. Approximating optimal social choice under metric preferences. *Artificial Intelligence* 264:27–51.

Benade, G.; Nath, S.; Procaccia, A. D.; and Shah, N. 2017. Preference elicitation for participatory budgeting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, 376–382.

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

Boutilier, C.; Caragiannis, I.; Haber, S.; Lu, T.; Procaccia, A. D.; and Sheffet, O. 2015. Optimal social choice functions: A utilitarian view. *Artificial Intelligence* 227:190–213.

Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D., eds. 2016. *Handbook of Computational Social Choice*. Cambridge University Press.

Caragiannis, I., and Procaccia, A. D. 2011. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence* 175(9-10):1655–1671.

Caragiannis, I.; Nath, S.; Procaccia, A. D.; and Shah, N. 2017. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research* 58:123–152.

Fain, B.; Goel, A.; Munagala, K.; and Prabhu, N. 2019. Random dictators with a random referee: Constant sample complexity mechanisms for social choice. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*.

Feldman, M.; Fiat, A.; and Golomb, I. 2016. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)*, 269–286.

Filos-Ratsikas, A., and Miltersen, P. B. 2014. Truthful approximations to range voting. In *Proceedings of the 10th International Conference on Web and Internet Economics (WINE)*, 175–188.

Filos-Ratsikas, A.; Frederiksen, S. K. S.; and Zhang, J. 2014. Social welfare in one-sided matchings: Random priority and beyond. In *Proceedings of the 7th Symposium of Algorithmic Game Theory (SAGT)*, 1–12.

Filos-Ratsikas, A.; Micha, E.; and Voudouris, A. A. 2019. The distortion of distributed voting. In *Proceedings of the 12th International Symposium on Algorithmic Game Theory (SAGT)*, 312–325.

Goel, A.; Krishnaswamy, A. K.; Sakshuwong, S.; and Aitamurto, T. 2019. Knapsack voting for participatory budgeting. *ACM Transactions on Economics and Computation* 7(2):8:1–8:27.

Goel, A.; Krishnaswamy, A. K.; and Munagala, K. 2017. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)*, 287–304.

Mandal, D.; Procaccia, A. D.; Shah, N.; and Woodruff, D. P. 2019. Efficient and thrifty voting by any means necessary. In *Proceedings of the 33rd Annual Conference on Neural Information Processing Systems (NeurIPS)*.

Munagala, K., and Wang, K. 2019. Improved metric distortion for deterministic social choice rules. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*, 245–262.

Pierczynski, G., and Skowron, P. 2019. Approval-based elections and distortion of voting rules. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 543–549.

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