

# Distortion in Social Choice Problems: The First 15 Years and Beyond

Elliot Anshelevich<sup>1</sup>, Aris Filos-Ratsikas<sup>2</sup>, Nisarg Shah<sup>3</sup> and Alexandros A. Voudouris<sup>4</sup>

<sup>1</sup>Computer Science Department, Rensselaer Polytechnic Institute

<sup>2</sup>Department of Computer Science, University of Liverpool

<sup>3</sup>Department of Computer Science, University of Toronto

<sup>4</sup>School of Computer Science and Electronic Engineering, University of Essex  
eanshel@cs.rpi.edu, aris.filos-ratsikas@liverpool.ac.uk, nisarg@cs.toronto.edu,  
alexandros.voudouris@essex.ac.uk

## Abstract

The notion of *distortion* in social choice problems has been defined to measure the loss in efficiency—typically measured by the utilitarian social welfare, the sum of utilities of the participating agents—due to having access only to limited information about the preferences of the agents. We survey the most significant results of the literature on distortion from the past 15 years, and highlight important open problems and the most promising avenues of ongoing and future work.

## 1 Introduction

Social choice theory is concerned with aggregating the preferences of individuals into a desirable collective decision, and has many applications such as choosing an electoral candidate, a public policy, the recipient of an award, or something as simple as choosing the most appropriate time for a meeting. Founded in the classic theory of Von Neumann and Morgenstern [1953], these preferences are typically assumed to be captured by *utility functions*, which assign numerical values to the different options, indicating the intensity by which an individual prefers one possible outcome to another. While the existence of such a *cardinal* utility structure is rarely disputed, the predominant approach in social choice theory is to elicit more limited preference information from the participants; in particular, they are usually required to provide *ordinal* preference rankings over the different outcomes. This is primarily due to cognitive reasons, as it is much more conceivable to come up with a ranking based on comparisons rather than a numerical utility structure.

The inevitable loss of information due to the restricted expressiveness of the preferences makes it rather challenging to optimize objectives of a cardinal nature. A natural such objective is the maximization of the (*utilitarian*) *social welfare*, defined as the sum of the individual utilities for the chosen outcome. How can we make the right choice when we do not have access to the utilities themselves? As the following example demonstrates, there are cases in which we cannot.

**Example 1.** Consider a simple scenario, with two outcomes  $a$  and  $b$ . The ordinal preferences of the participants are such that half of them prefer  $a$  to  $b$ , and the remaining half prefer  $b$

to  $a$ . In such a case, based only on these ordinal preferences, it is impossible to distinguish between the two outcomes, and thus either of them could be selected. So, let us assume that the winner is  $a$ . The unknown underlying utilities of the participants for the outcomes might be such that those that prefer  $a$  to  $b$  are mostly *indifferent* between the two, and have utility about  $1/2$  for each of them, whereas the remaining participants *strongly prefer*  $b$ , and have utility 1 for  $b$  and 0 for  $a$ . Consequently, the total utility of the participants for the winner  $a$  is only about  $n/4$  and their utility for  $b$  is  $3n/4$ , where  $n$  is the number of participants. This shows that  $a$  is clearly not the optimal choice, and in fact only achieves a social welfare that is a multiplicative factor of 3 away from the optimal.

While Example 1 makes it obvious that selecting the outcome that maximizes the social welfare is not always possible, it does not conclusively answer the question of *how far away from the optimal outcome our choice can be*. Driven by the principles of worst-case analysis and approximation algorithms, a large part of the computational social choice literature [Brandt *et al.*, 2016] has studied questions about computing the best outcome for many fundamental social choice problems. In particular, in 2006 Procaccia and Rosenschein [2006] defined the notion of the *distortion* to measure the worst-case deterioration of an aggregate cardinal objective, such as utilitarian social welfare, in social choice settings, due to having access to preferences of limited expressiveness, particularly ordinal rankings. The associated research agenda posed the following fundamental question:

*Which social choice functions achieve the best possible distortion for a given cardinal objective?*

Over the past 15 years, this type of question has been studied extensively in the context of many different social choice problems, such as *general single-winner elections*, *elections with metric preferences*, *committee selection*, *participatory budgeting*, and *matching*, thus giving rise to a rich and vibrant literature at the intersection of economics, computation, and artificial intelligence. In this survey, we highlight the most significant of these results, some current open problems, as well as the most prominent ongoing and future directions.

## 2 Normalized Social Choice

The first works on the distortion of social choice rules following Procaccia and Rosenschein [2006] studied the setting of

single-winner elections with normalized utilities. In this setting, there is a set  $N$  of  $n$  agents and a set  $A$  of  $m$  alternatives. Every agent  $i \in N$  has a value  $v_{ij} \in \mathbb{R}_{\geq 0}$  for every alternative  $j \in A$ ; let  $\mathbf{v} = (v_{ij})_{i \in N, j \in A}$ . These values are usually assumed to be *unit-sum* normalized so that  $\sum_{j \in A} v_{ij} = 1$  for every agent  $i \in N$ . We refer the reader to the note of Aziz [2020] for justifications of such normalization assumptions. Given these values, the goal is to compute the alternative that maximizes the *social welfare*, which for any alternative  $j \in A$  is defined as the total value of the agents for  $j$ , i.e.,  $\text{SW}(j|\mathbf{v}) = \sum_{i \in N} v_{ij}$ . It is important to note that while this setting is referred to as an “election”, and the associated terminology is used, it can broadly capture most decision-making scenarios, where a possible outcome corresponds to an alternative and a participant corresponds to an agent.

If we had access to the values of the agents, maximizing the social welfare would be an easy task; we could simply compute the social welfare of every alternative and then select the one with the largest social welfare. However, the values are the agents’ *private information*; they are unknown to the designer and it might be the case that they are even unknown to the agents themselves. This stems from the motivation presented in the Introduction, namely that coming up with these values is a cognitively demanding task, which the agents are not often asked to partake in. Instead, we have access to the linear orderings defined by the values. Specifically, every agent  $i \in N$  is associated with a (strict) *ranking*  $\succ_i$  over the alternatives such that  $j \succ_i j'$  implies that  $v_{ij} \geq v_{ij'}$ ; let  $\succ_{\mathbf{v}} = (\succ_i)_{i \in N}$ . These rankings are given as input to a *social choice function* (or *voting rule*)  $f$ , which decides a single winning alternative  $f(\succ_{\mathbf{v}}) \in A$ . The *distortion* of  $f$  measures how far away the alternative  $f(\succ_{\mathbf{v}})$  is from the optimal alternative in the worst case. Formally, it is defined as the worst-case ratio (over all possible instances) of the maximum social welfare over the social welfare of the alternative returned by  $f$ , i.e.,

$$\text{distortion}(f) = \sup_{(N, A, \mathbf{v})} \frac{\max_{j \in A} \text{SW}(j|\mathbf{v})}{\text{SW}(f(\succ_{\mathbf{v}})|\mathbf{v})}.$$

For this setting, Procaccia and Rosenschein [2006] showed that the distortion of the well-known Borda voting rule is unbounded. Another well-known and in fact simpler voting rule is the following.

**Voting Rule (Plurality).** Elect the alternative that the most agents consider their favorite one, breaking ties arbitrarily.

Caragiannis and Procaccia [2011] showed that the Plurality rule has a distortion of  $O(m^2)$ ; this bound was later proven to be best possible among all deterministic voting rules [Caragiannis *et al.*, 2017]. To give the reader a soft introduction to the type of arguments used to bound the distortion of voting rules, we provide a sketch of the rather simple arguments used for Plurality. We remark that proving bounds on the distortion may require rather involved arguments in general depending on the setting and the type of rules considered.

**Theorem.** *The distortion of Plurality is  $O(m^2)$ , and this is best possible among deterministic voting rules.*

*Proof Sketch.* For the upper bound, observe that the winning alternative  $a$  under Plurality must appear at the first position

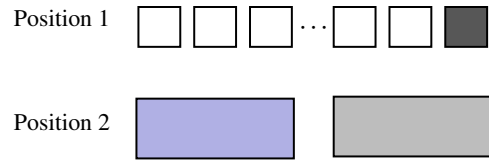


Figure 1: The instance used in the lower bound proof of the theorem.

of the rankings of least  $n/m$  agents by the pigeonhole principle. Since the values are unit-sum, these agents must have value at least  $1/m$  for  $a$ , and thus the social welfare of  $a$  is at least  $n/m^2$ . On the other hand, the maximum social welfare of any alternative is  $n$ , and the bound follows.

For the lower bound, see the instance of Figure 1, where (a) there is a set of  $m-2$  alternatives such that each of them appears at position 1 in the rankings of  $n/(m-2)$  agents, and (b) each of the remaining two alternatives appear at position 2 in the rankings of  $n/2$  agents. Any deterministic voting rule must select either one of the first  $m-2$  alternatives or one of the remaining two alternatives, and has no way of distinguishing between alternatives in each group; therefore we can assume without loss of generality that the rule selects either the alternative shaded dark gray or the alternative shaded light gray in the figure. Let the values of the agents that rank the dark gray alternative first be  $1/m$  for all alternatives. Let the values of all the remaining agents that rank the light gray alternative second be 1 for their top-ranked alternative and 0 for all others. Finally, let the values of all the agents that rank the blue alternative second be  $1/2$  for their top two alternatives and zero for the rest. Observe that these values are unit-sum. The social welfare of both the light gray and the dark gray alternatives is  $n/(m(m-2))$ , whereas the social welfare of the blue alternative is  $n/4$ , yielding a lower bound  $\Omega(m^2)$ .  $\square$

Randomization proved to be a powerful tool in the quest to achieve improved distortion bounds, overcoming the sharp impossibility results for deterministic rules. A randomized rule outputs a probability distribution over alternatives rather than a single alternative, and its distortion is defined with respect to the expected social welfare that it achieves. The distortion of randomized rules was most notably studied by Boutilier *et al.* [2015], who designed a particularly involved voting rule with distortion  $O(\sqrt{m} \cdot \log^* m)$ . They also presented the following much simpler rule that achieves a slightly inferior bound of  $O(\sqrt{m} \cdot \log m)$ :

**Voting Rule ([Boutilier *et al.*, 2015]).** With probability  $1/2$ , select an alternative uniformly at random, and with probability  $1/2$ , select every alternative  $j \in A$  with probability that is proportional to  $j$ ’s harmonic score, computed based on the harmonic scoring vector  $(1, 1/2, \dots, 1/m)$ , which assigns to every alternative a score of  $1/\ell$  for each appearance at position  $\ell$  in the rankings of the agents.

Boutilier *et al.* [2015] also proved a general lower bound of  $\Omega(\sqrt{m})$  on the achievable distortion of randomized rules, thus effectively settling the problem for the single-winner case. Despite this fact, as we will see in Section 5, the single-winner normalized social choice setting still offers rather fruitful ground for further work on interesting variants of the

problem.

### 2.1 Beyond the Single-Winner Setting

Many interesting multi-winner extensions have also been studied over the years. Indicatively, Caragiannis *et al.* [2017] considered the problem of selecting a subset (committee) of  $k$  alternatives that maximizes the social welfare, when the value of each agent for a committee of alternatives is defined as the maximum value she derives from the committee’s members. They showed almost tight bounds for both deterministic and randomized rules as functions of  $m$  and  $k$ . Benadè *et al.* [2017] considered the participatory budgeting problem, where we are given a budget, each alternative has an associated cost, and the goal is to choose a subset of alternatives so as to maximize the social welfare of the agents while ensuring that the total cost of the chosen alternatives does not exceed the budget. They showed several tight bounds for many different types of input given by the agents, such as knapsack votes, rankings by value, rankings by value for money, and threshold approval votes. Finally, Benadè *et al.* [2019] studied the distortion of voting rules that output rankings of alternatives rather than subsets, and showed that bounds that are qualitatively similar to those for the single-winner setting can be achieved by randomized rules.

### 3 Metric Social Choice

In the normalized social choice setting discussed in the previous section, the values of the agents for the alternatives sum up to 1 but can otherwise be highly arbitrary. However, in many applications (such as elections), voters usually prefer candidates that express opinions which are closer to their own personal beliefs, and thus the values are determined by an “ideological distance”. For example, one can envision ideological axes from “liberal” to “conservative”, or from “libertarian” to “authoritarian”, and the preferences of the agents as points in the space defined by these axes.

More generally, we can think of the agents and the alternatives as points in a (possibly high-dimensional, e.g., corresponding to multiple political issues or ideological dimensions) metric space, in which the triangle inequality is satisfied (see also [Merrill III *et al.*, 1999; Enelow and Hinich, 1984]). Instead of thinking about values and the social welfare objective, in this setting it is more natural to talk about costs and the analogous social cost objective, which we aim to minimize. In particular, the *cost* of an agent for an alternative equals their distance in the space, the agents rank the alternatives in a non-decreasing order of their costs, and the *social cost* of an alternative is the total cost of the agents for this alternative. Then, the distortion of a voting rule is defined as the ratio of the social cost of the alternative chosen by the rule over the minimum social cost, when given as input the ordinal preferences of the agents.

An example very similar to Example 1 shows that the distortion of any deterministic voting rule in the metric social choice setting is at least 3; see Figure 2. However, in contrast to the normalized setting from Section 2, the metric domain restriction makes it possible to achieve small constant bounds, even with deterministic rules. This was first shown

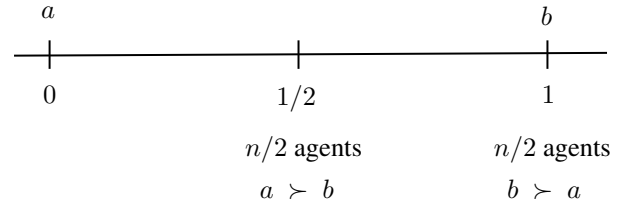


Figure 2: An example showing that no deterministic voting rule can achieve a distortion better than 3 in the metric social choice setting. Since  $n/2$  agents prefer  $a$  to  $b$  and  $n/2$  agents prefer  $b$  to  $a$ , without loss of generality, any rule may select  $a$  as the winner for a social cost of  $n/4 + n/2 = 3n/4$ . On the other hand, the optimal alternative  $b$  has social cost  $n/4$ .

by Anshelevich *et al.* [2015], who initiated the study of distortion in the metric setting and showed that the well-known Copeland rule achieves a distortion of 5.

Anshelevich *et al.* [2015] further conjectured that the lower bound of 3 provided by the example of Figure 2 must be the tight bound for deterministic rules, and in fact that it can be achieved by the well-known Ranked Pairs rule. In the years that followed, a succession of papers made significant progress towards resolving this conjecture. Skowron and Elkind [2017] showed that the distortion of every positional scoring rule is  $\Omega(\sqrt{\log m})$ , without however providing any scoring rule with distortion better than  $O(m)$ . They also showed that the distortion of the Single Transferable Vote (STV) rule is between  $\Omega(\sqrt{\log m})$  and  $O(\log m)$ . Goel *et al.* [2017] were the first to disprove that the Ranked Pairs rule matches the bound of 3, by presenting a lower bound of 5 for it; its exact distortion was later shown to be  $\Theta(\sqrt{m})$  by Kempe [2020a]. Munagala and Wang [2019] made significant progress towards resolving the conjecture by designing a novel voting rule that broke the barrier of 5 and achieved distortion 4.236. They also identified many important properties that would suffice to achieve the conjectured bound of 3. Based on these properties, the conjecture was finally positively resolved very recently by Gkatzelis *et al.* [2020], who designed the following *plurality matching* voting rule.

**Voting Rule** ([Gkatzelis *et al.*, 2020]). For every alternative  $j \in A$ , construct the integral domination graph of  $j$ , which is a bipartite graph with the nodes of both sides corresponding to agents, and each edge  $(x, y)$  between agents  $x$  and  $y$  indicating that  $x$  prefers alternative  $j$  to the favorite alternative of  $y$ . Return an alternative whose integral domination graph admits a perfect matching; such an alternative is guaranteed to exist.

These results have impressive and counter-intuitive implications. They state that, as long as agents and alternatives correspond to points in a metric space, it is always possible to choose a close-to-optimal alternative based only on the ordinal rankings of the agents, without even the need to know their actual locations in the space.

### 3.1 Randomized Rules

Randomized voting rules have also been considered in the metric setting by Anshelevich and Postl [2017], who showed that randomization is able to break the barrier of 3 suggested by the example of Figure 2, albeit only by a small amount which vanishes as the number of agents becomes large. In particular, they showed that *Random Dictatorship* (which selects an agent uniformly at random and outputs her favorite alternative) has distortion  $3 - 2/n$ , and complemented this with a lower bound of 2 on the distortion of any randomized rule, which holds even when the metric space is a simple line. Kempe [2020b] presented a randomized rule that has distortion  $3 - 2/m$ , which is slightly better than that of Random Dictatorship when  $n$  is large but  $m$  is small. However, the following is still a major open question.

**Open problem 1.** *What is the best possible distortion for randomized social choice rules in the metric setting?*

Progress has been made for some special cases, such as the line metric, for which Feldman *et al.* [2016] presented a randomized rule, the *Spike mechanism*, with a tight distortion bound of 2. Fain *et al.* [2017] also considered restricted spaces called median graphs, and showed that it is possible to achieve a distortion of almost 1.218, via a *sequential deliberation* rule. Finally, Anshelevich and Postl [2017] also introduced the notion of  $\alpha$ -*decisiveness* to measure the preference strength of an agent for her top choice over her second choice, and showed upper and lower bounds parameterized by this quantity.

### 3.2 Limited Ordinal Information

While the bulk of papers on distortion in metric social choice settings have focused on rules that in general take as input complete rankings, there has also been work on rules that require incomplete ordinal information, motivated by the fact that some rules like Plurality and Random Dictatorship essentially only need to know the favorite alternatives of the agents. In this context, Gross *et al.* [2017] showed a lower bound of  $3 - 2/\lfloor m/k \rfloor$  on the distortion when the agents submit top- $k$  rankings for  $k < m/2$ , and showed that a very simple randomized rule, which asks random agents for their favorite alternatives until two of them agree, has many nice axiomatic properties and is almost optimal in terms of distortion when the number of alternatives is not too large.

Fain *et al.* [2019] studied voting rules that require the agents to compare only a constant number of pairs of alternatives (constant sample complexity) and presented the *Random Referee* rule, which achieves a constant *squared distortion*, a measure of the variance of the distortion of randomized rules. Furthermore, for any constant  $k$ , they showed that no rule that elicits only top- $k$  rankings can have squared distortion that is sublinear in  $m$ , and presented the *Random Oligarchy* rule with distortion almost matching the lower bound of  $3 - 2/m$  for  $k = 1$ . As mentioned above, this small gap was later resolved by Kempe [2020b]. He mainly focused on deterministic rules and showed that the distortion of rules that take as input the alternatives that the agents rank in a given set of  $k$  positions is  $\Theta(m/k)$ . Moreover, he showed that the distortion

of rules that require the agents to submit  $b$  bits of information about their rankings is  $\Omega(m/b)$ .

### 3.3 Other Settings

Several papers have also studied variations of the main metric setting by making different assumptions about the origin of the alternatives, the behavior of the agents, and the objective. In particular, Cheng *et al.* [2017] considered a probabilistic model in which there are only two alternatives that are independently drawn from the population of agents. They showed that the distortion of the majority rule is between 1.5 and 2 for arbitrary metric spaces, and improves to 1.716 for the line metric. In follow-up work, Cheng *et al.* [2018] extended their analysis to the case where multiple alternatives are randomly chosen from the set of agents, and showed bounds on the distortion of positional scoring rules. Borodin *et al.* [2019] studied *primary voting systems* wherein there are political parties and agents affiliated with those vote to choose the party representatives, who then compete in the general elections. Their main results were that every voting rule is guaranteed to perform almost as well under the primary system as under the direct system (in which all alternatives are candidates in the election), but not vice versa.

Ghods *et al.* [2021] considered settings with two alternatives, in which some agents are allowed to abstain from the election process in case they are pretty much indifferent between the alternatives or the distance from their favorite alternative is sufficiently large. They presented a probabilistic model to capture these properties and characterized the distortion of the majority rule.

Pierczynski and Skowron [2019] focused on approval-based preferences according to which the agents do not rank the alternatives, but instead list those that they approve. Besides showing that the distortion of the approval rule is unbounded, they also introduced the notion of *acceptability-based distortion*, which aims to capture how far away the chosen alternative is from the alternative that is approved by the most agents. For this definition of distortion, they showed bounds for well-known rules such as Plurality, Borda, Ranked Pairs, and Copeland. Jaworski and Skowron [2020] studied the distortion of multi-winner rules for the election of a committee, whose quality is determined by the extent to which the decisions it makes on binary issues are consistent to the preferences of the agents on those issues.

## 4 Other Problems and Extensions

Distortion has also been considered in many other settings that have either been studied under the general umbrella of social choice theory, or are very much related.

### 4.1 Matching and Other Graph Problems

The *one-sided matching* problem (also known as *house allocation*) is a classic problem in economics, where  $n$  agents have preferences over a set of  $n$  items, and the goal is to match each agent to a single item in order to maximize the social welfare, defined as the total value of the agents for the items they are matched to. The problem was first studied in the context of distortion by Filos-Ratsikas *et al.* [2014], who showed

that *Random Serial Dictatorship* (e.g., see [Abdulkadiroğlu and Sönmez, 1998]) has distortion  $O(\sqrt{n})$ , and furthermore, this is best possible among all rules that take as input the ordinal preferences of the agents induced by their underlying normalized values. Recently, for the same setting, Amanatidis *et al.* [2020a] showed that the best possible deterministic rule for one-sided matching has distortion  $\Theta(n^2)$ .

Anshelevich and Sekar [2016a; 2016b] and Abramowitz and Anshelevich [2018] studied the distortion in *maximum weight matching* and many other graph problems such as *maximum  $k$ -Sum*, *densest  $k$ -Subgraph*, and *maximum travelling salesperson (TSP)*. They primarily focused on the case of symmetric values that satisfy the triangle inequality and presented a general framework for designing greedy rules that achieve small constant distortion bounds. For instance, among other results, they showed that deterministic rules can achieve a distortion bound of 2 for maximum-weight matching and maximum-TSP, whereas randomized rules can achieve bounds of 1.6 and 1.88 for those two problems respectively. Anshelevich and Zhu [2018] studied a general facility assignment model with the goal of coming up with a (possibly constrained) assignment of agents to facilities. For many different optimization objectives, they showed a black-box reduction which converts a cardinal rule for the problem into an ordinal rule with small distortion. In particular, they showed that the distortion of many facility assignment and matching problems is at most 3 when the possible locations of the facilities are public knowledge, but only ordinal preferences of the agents for the facilities are known.

## 4.2 Truthfulness and Distortion

In most of the aforementioned results, the implicit assumption was that when agents are asked to submit their ordinal preference rankings, they do so honestly. In many cases, however, the agents might have incentives to misreport their preferences, if that could result in a more favorable outcome (e.g., the election of a more preferred alternative, or the allocation of a more preferable item). In such a case, we would like to design rules that do not provide incentives for such strategic behavior; these are typically called *truthful mechanisms* and are rooted in the principles of *mechanism design* (e.g., see [Nisan and others, 2007]).

For the normalized social choice setting, Bhaskar *et al.* [2018] studied multi-winner elections and participatory budgeting problems. The main takeaway from their results was that essentially the best randomized truthful mechanisms are ordinal, and are in fact asymptotically almost as good (up to logarithmic factors) as the best ordinal non-truthful rules. The single-winner setting was studied previously by Filos-Ratsikas and Miltersen [2014], albeit for a different normalization (known as *unit-range*), and later on by Caragiannis *et al.* [2018] for a related hybrid social choice setting. In the metric setting, and the facility location problem in particular, Feldman *et al.* [2016] considered different degrees of expressiveness in terms of the agents' preferences, and proved upper and lower bounds on the distortion of truthful mechanisms.

Even before Bhaskar *et al.* [2018] reached a similar conclusion, Filos-Ratsikas *et al.* [2014] observed that the best truthful mechanisms are ordinal in the one-sided matching

setting described earlier. In particular, they showed that the  $O(\sqrt{n})$  bound of random serial dictatorship (which is both ordinal and truthful) is best possible among all truthful mechanisms. Caragiannis *et al.* [2016] studied the same rule for the minimum-cost metric matching problem, and showed that its distortion is between  $\Omega(n^{0.29})$  and  $O(n)$ . This is in fact the best known upper bound for any ordinal rule in this setting, even non-truthful ones. To this end, we have the following open problem.

**Open problem 2.** *What is the distortion of the best ordinal rule for the minimum-cost metric matching problem? What if we also require truthfulness?*

## 5 Generalizations of Distortion

While there are still several interesting open questions in the classic settings described above, in this section we would like to highlight a more recent stream of work which views the distortion as a more general concept of measuring the loss of efficiency due to having access to incomplete information. We believe this opens up many intriguing avenues for future investigations. In particular, these works refine the notion of distortion to capture scenarios in which the loss of efficiency is not necessarily because the agents express ordinal preferences, but because there is only *limited access* to the *complete* cardinal information. Some works consider settings in which the provided information is in a sense “less than cardinal but more than ordinal”, whereas others explore the limits of communication between the agents and the designer, which inherently results in some information loss. Finally, some recent papers consider the loss of information that is incurred by making decisions in a distributed manner. We highlight some of the recent results below, as well as areas for future work.

### 5.1 Beyond Ordinal Information

As discussed in the Introduction, the standard argument in favor of ordinal preferences is based on the cognitive limitations of the agents. However, one can conceivably consider settings where some *limited additional* information of a cardinal nature is available, without imposing a large cognitive overhead on the agents.

Along those lines, Amanatidis *et al.* [2020b] studied deterministic rules for single-winner social choice that have access to the ordinal preferences of the agents and can also learn some of their numerical values by appropriately asking a small number of *queries* to the agents. They showed that a single query per agent is enough to achieve distortion  $O(m)$  (improving the bound of  $\Theta(m^2)$  achieved by ordinal rules), and designed a rule that (as special cases) achieves distortion  $O(\sqrt{m})$  by making  $O(\log m)$  queries per agent (thus matching the best bound possible for ordinal *randomized* rules), and *constant* distortion by making  $O(\log^2 m)$  queries per agent; quite remarkably, these bounds hold even without any normalization assumptions. Moreover, the authors showed that it is impossible to obtain sublinear distortion with a single query per agent, and also that  $\omega(\log m)$  queries per agent are necessary to achieve constant distortion.

Amanatidis *et al.* conjectured that it is possible to achieve distortion  $O(\sqrt{m})$  by asking only a constant number of

queries per agent, and constant distortion by asking  $\Theta(\log m)$  queries per agent. Another interesting direction is to consider randomized rules, where randomization can be used in either the query phase, the election phase, or both. This line of work raises the following conceptual open question.

**Open problem 3.** *Which other settings admit interesting tradeoffs between the distortion and the number of queries?*

Follow-up works [Amanatidis *et al.*, 2020a; Ma *et al.*, 2020] identified the one-sided matching problem and several of its generalizations as such settings, and showed quantitatively similar tradeoffs between the number of queries and the distortion. It would be very interesting to see if a similar approach can be applied to the metric social choice setting, where the bounds are already small constants and significant improvements are more difficult to obtain.

In the metric setting, Abramowitz *et al.* [2019] also considered the case where the agents provide limited cardinal information on top of their ordinal preferences, by indicating preference strengths (the voters' *passion*). In particular, for each agent  $i$  and every pair of alternatives  $a$  and  $b$ , the voting rule has access to the ratio of the distances of agent  $i$  to alternatives  $a$  and  $b$ . The authors showed that there exist deterministic rules with distortion  $\sqrt{2}$  when there only two alternatives, and distortion 2 when there are more alternatives. They also considered rules that do not know the exact preference strengths, but only whether these strengths exceed a predefined threshold  $\tau$ , and showed how the distortion bounds change as functions of  $\tau$ . Perhaps the most general conceptual research direction is the following.

**Open problem 4.** *What other meaningful ways of eliciting limited cardinal information can be used to improve the distortion in metric or normalized social choice?*

We remark that the idea of eliciting cardinal information has also been present in earlier works. In particular, Benadé *et al.* [2017], and later on Bhaskar *et al.* [2018], considered (deterministic or randomized) *threshold approval* queries, in which the agents are asked to list all the alternatives that they value higher than a specified threshold.

## 5.2 Communication Complexity

Mandal *et al.* [2019] considered similar tradeoffs between information and distortion, but without assuming a priori access to the ordinal preferences of the agents. Instead, they focused on voting rules that consist of an elicitation part which collects a specific number of bits of information per agent about their values, and an aggregation part which uses this information to decide a single winning alternative. Mandal *et al.* [2019] proposed a voting rule that uses deterministic elicitation and aggregation, and achieves distortion  $d$  with communication complexity  $O(m \log(d \log m)/d)$ . Furthermore, they showed that to achieve distortion  $d$ , any voting rule using deterministic elicitation must have communication complexity  $\Omega(m/d^2)$ , whereas any voting rule using randomized elicitation must have communication complexity  $\Omega(m/d^3)$ .

In follow-up work, Mandal *et al.* [2020] improved these tradeoffs by showing tight bounds (up to logarithmic factors) on the communication complexity necessary to achieve distortion  $d$ . Specifically, by exploiting recent advances in

streaming algorithms and making a connection to the literature on communication complexity from theoretical computer science, they showed a bound of  $\Theta(m/d)$  for deterministic elicitation and a bound of  $\tilde{\Theta}(m/d^3)$  for randomized elicitation. Mandal *et al.* [2020] also considered the multi-winner setting of Caragiannis *et al.* [2017], and showed that the required communication complexity for achieving distortion  $d$  is  $\tilde{\Theta}(m/(kd))$  for deterministic elicitation and  $\tilde{\Theta}(m/(kd^3))$  for randomized elicitation, where  $k$  is the number of winners selected. While these two papers have provided an almost complete picture of the communication complexity of single- and multi-winner voting rules, one could ask similar questions for other settings.

**Open problem 5.** *Which other settings admit interesting tradeoffs between distortion and communication?*

Settings like the one-sided matching and its generalizations seem like good candidates, and so does the metric social choice setting.

## 5.3 Distributed Settings

In all the papers discussed so far, the voting rules operate by collecting the preferences of all agents and then aggregating them into a common decision in a single step. However, in many important applications (such as presidential elections in the US) the aggregation process is *distributed*, in the sense that the agents are partitioned into *districts* and vote locally therein to choose representative alternatives, which are then aggregated into a final collective outcome. With this in mind, Filos-Ratsikas *et al.* [2020] initiated the study of the distortion in distributed settings, aiming to capture the loss of efficiency not only due to the possibly limited expressiveness of the agents' preferences, but also due to not having full access to the details of the local decisions within the districts.

Filos-Ratsikas *et al.* [2020] focused on the unit-sum single-winner setting and bounded the distortion of deterministic distributed rules, which use classic voting rules for the local district elections, and choose as overall winner the alternative that is the representative of the most districts (or has the largest total weight among the representatives in case the districts are weighted). Their results showed that the distributed nature of such mechanisms leads to considerably larger distortion compared to their centralized counterparts. Later on, Filos-Ratsikas and Voudouris [2020] studied the distortion of deterministic distributed rules for discrete and continuous facility location on the line (single-dimensional metric setting), and showed tight bounds for the social cost objective that are small constants, both in general and under truthfulness constraints. There are many interesting open questions to be answered in distributed settings.

**Open problem 6.** *What is the distortion of randomized distributed mechanisms? How do the distortion bounds look like when the goal is to choose a committee of alternatives or locate more facilities? What about general metric spaces and other settings?*

## Acknowledgements

This work was partially supported by NSF awards CCF-1527497 and CCF-2006286.

## References

- [Abdulkadiroğlu and Sönmez, 1998] Atila Abdulkadiroğlu and Tayfun Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.
- [Abramowitz and Anshelevich, 2018] Ben Abramowitz and Elliot Anshelevich. Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 894–901, 2018.
- [Abramowitz *et al.*, 2019] Ben Abramowitz, Elliot Anshelevich, and Wennan Zhu. Awareness of voter passion greatly improves the distortion of metric social choice. In *Proceedings of the The 15th Conference on Web and Internet Economics (WINE)*, pages 3–16, 2019.
- [Amanatidis *et al.*, 2020a] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. A few queries go a long way: Information-distortion tradeoffs in matching. *CoRR*, abs/2009.06543, 2020.
- [Amanatidis *et al.*, 2020b] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A. Voudouris. Peeking behind the ordinal curtain: Improving distortion via cardinal queries. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1782–1789, 2020.
- [Anshelevich and Postl, 2017] Elliot Anshelevich and John Postl. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research*, 58:797–827, 2017.
- [Anshelevich and Sekar, 2016a] Elliot Anshelevich and Shreyas Sekar. Blind, greedy, and random: Algorithms for matching and clustering using only ordinal information. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, pages 390–396, 2016.
- [Anshelevich and Sekar, 2016b] Elliot Anshelevich and Shreyas Sekar. Truthful mechanisms for matching and clustering in an ordinal world. In *Proceedings of the 12th Conference on Web and Internet Economics (WINE)*, pages 265–278, 2016.
- [Anshelevich and Zhu, 2018] Elliot Anshelevich and Wennan Zhu. 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)*, pages 3–20, 2018.
- [Anshelevich *et al.*, 2015] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. Approximating optimal social choice under metric preferences. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pages 777–783, 2015.
- [Aziz, 2020] Haris Aziz. Justifications of welfare guarantees under normalized utilities. *ACM SIGecom Exchanges*, 17(2):71–75, 2020.
- [Benadè *et al.*, 2017] Gerdus Benadè, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. Preference elicitation for participatory budgeting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 376–382, 2017.
- [Benadè *et al.*, 2019] Gerdus Benadè, Ariel D. Procaccia, and Mingda Qiao. Low-distortion social welfare functions. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1788–1795, 2019.
- [Bhaskar *et al.*, 2018] Umang Bhaskar, Varsha Dani, and Abheek Ghosh. Truthful and near-optimal mechanisms for welfare maximization in multi-winner elections. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 925–932, 2018.
- [Borodin *et al.*, 2019] Allan Borodin, Omer Lev, Nisarg Shah, and Tyrone Strangway. Primarily about primaries. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1804–1811, 2019.
- [Boutilier *et al.*, 2015] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213, 2015.
- [Brandt *et al.*, 2016] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of computational social choice*. Cambridge University Press, 2016.
- [Caragiannis and Procaccia, 2011] Ioannis Caragiannis and Ariel D. Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175(9-10):1655–1671, 2011.
- [Caragiannis *et al.*, 2016] Ioannis Caragiannis, Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, Kristoffer Arnsfelt Hansen, and Zihan Tan. Truthful facility assignment with resource augmentation: An exact analysis of serial dictatorship. In *Proceedings of the 12th International Conference on Web and Internet Economics (WINE)*, pages 236–250, 2016.
- [Caragiannis *et al.*, 2017] Ioannis Caragiannis, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152, 2017.
- [Caragiannis *et al.*, 2018] Ioannis Caragiannis, Aris Filos-Ratsikas, Swaprava Nath, and Alexandros A. Voudouris. Truthful mechanisms for ownership transfer with expert advice. *CoRR*, abs/1802.01308, 2018.
- [Cheng *et al.*, 2017] Yu Cheng, Shaddin Dughmi, and David Kempe. Of the people: Voting is more effective with representative candidates. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)*, pages 305–322, 2017.
- [Cheng *et al.*, 2018] Yu Cheng, Shaddin Dughmi, and David Kempe. On the distortion of voting with multiple representative candidates. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 973–980, 2018.

- [Enelow and Hinich, 1984] James M. Enelow and Melvin J. Hinich. *The spatial theory of voting: An introduction*. CUP Archive, 1984.
- [Fain *et al.*, 2017] Brandon Fain, Ashish Goel, Kamesh Munagala, and Sukolsak Sakshuwong. Sequential deliberation for social choice. In *Proceedings of the 13th International Conference on Web and Internet Economics (WINE)*, pages 177–190, 2017.
- [Fain *et al.*, 2019] Brandon Fain, Ashish Goel, Kamesh Munagala, and Nina Prabhu. Random dictators with a random referee: Constant sample complexity mechanisms for social choice. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1893–1900, 2019.
- [Feldman *et al.*, 2016] Michal Feldman, Amos Fiat, and Idan Golomb. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)*, pages 269–286, 2016.
- [Filos-Ratsikas and Miltersen, 2014] Aris Filos-Ratsikas and Peter Bro Miltersen. Truthful approximations to range voting. In *Proceedings of the 10th International Conference on Web and Internet Economics (WINE)*, pages 175–188, 2014.
- [Filos-Ratsikas and Voudouris, 2020] Aris Filos-Ratsikas and Alexandros A. Voudouris. Approximate mechanism design for distributed facility location. *CoRR*, abs/2007.06304, 2020.
- [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 *Proceedings of the 7th Symposium on Algorithmic Game Theory (SAGT)*, pages 1–12, 2014.
- [Filos-Ratsikas *et al.*, 2020] Aris Filos-Ratsikas, Evi Micha, and Alexandros A. Voudouris. The distortion of distributed voting. *Artificial Intelligence*, 286:103343, 2020.
- [Ghodsi *et al.*, 2021] Mohammad Ghodsi, Mohamad Lutfian, and Masoud Seddighin. On the distortion value of the elections with abstention. *Journal of Artificial Intelligence Research*, 70:567–595, 2021.
- [Gkatzelis *et al.*, 2020] Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. Resolving the optimal metric distortion conjecture. In *Proceedings of the 61st IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1427–1438, 2020.
- [Goel *et al.*, 2017] Ashish Goel, Anilesh K Krishnaswamy, and Kamesh Munagala. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)*, pages 287–304, 2017.
- [Gross *et al.*, 2017] Stephen Gross, Elliot Anshelevich, and Lirong Xia. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 544–550, 2017.
- [Jaworski and Skowron, 2020] Michal Jaworski and Piotr Skowron. Evaluating committees for representative democracies: The distortion and beyond. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 196–202, 2020.
- [Kempe, 2020a] David Kempe. An analysis framework for metric voting based on LP duality. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 2079–2086, 2020.
- [Kempe, 2020b] David Kempe. Communication, distortion, and randomness in metric voting. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 2087–2094, 2020.
- [Ma *et al.*, 2020] Thomas Ma, Vijay Menon, and Kate Larson. Improving welfare in one-sided matching using simple threshold queries. *CoRR*, abs/2011.13977, 2020.
- [Mandal *et al.*, 2019] Debmalaya Mandal, Ariel D. Procaccia, Nisarg Shah, and David P. Woodruff. Efficient and thrifty voting by any means necessary. In *Proceedings of the 32nd Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 7178–7189, 2019.
- [Mandal *et al.*, 2020] Debmalaya Mandal, Nisarg Shah, and David P. Woodruff. Optimal communication-distortion tradeoff in voting. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*, pages 795–813, 2020.
- [Merrill III *et al.*, 1999] Samuel Merrill III, Samuel Merrill, and Bernard Grofman. *A unified theory of voting: Directional and proximity spatial models*. Cambridge University Press, 1999.
- [Munagala and Wang, 2019] Kamesh Munagala and Kangning Wang. Improved metric distortion for deterministic social choice rules. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*, pages 245–262, 2019.
- [Nisan and others, 2007] Noam Nisan et al. Introduction to mechanism design (for computer scientists). *Algorithmic Game Theory*, 9:209–242, 2007.
- [Pierczynski and Skowron, 2019] Grzegorz Pierczynski and Piotr Skowron. Approval-based elections and distortion of voting rules. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 543–549, 2019.
- [Procaccia and Rosenschein, 2006] Ariel D. Procaccia and Jeffrey S. Rosenschein. The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents (CIA)*, pages 317–331, 2006.
- [Skowron and Elkind, 2017] Piotr K. Skowron and Edith Elkind. Social choice under metric preferences: Scoring rules and STV. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 706–712, 2017.
- [Von Neumann and Morgenstern, 1953] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton university press, 1953.