Subset Selection Via Implicit Utilitarian Voting

Ioannis Caragiannis   CARAGIAN@CEID.UPATRAS.GR
University of Patras, Greece

Swaprava Nath   SWAPRAVN@CS.CMU.EDU
Carnegie Mellon University, USA

Ariel D. Procaccia   ARIELPRO@CS.CMU.EDU
Carnegie Mellon University, USA

Nisarg Shah   NISARG@G.HARVARD.EDU
Harvard University, USA

Abstract

How should one aggregate ordinal preferences expressed by voters into a measurably superior social choice? A well-established approach — which we refer to as *implicit utilitarian voting* — assumes that voters have latent utility functions that induce the reported rankings, and seeks voting rules that approximately maximize utilitarian social welfare. We extend this approach to the design of rules that select a subset of alternatives. We derive analytical bounds on the performance of optimal (deterministic as well as randomized) rules in terms of two measures, distortion and regret. Empirical results show that regret-based rules are more compelling than distortion-based rules, leading us to focus on developing a scalable implementation for the optimal (deterministic) regret-based rule. Our methods underlie the design and implementation of RoboVote.org, a not-for-profit website that helps users make group decisions via AI-driven voting methods.

1. Introduction

We are interested in the classic social choice problem of aggregating the preferences of a set of voters — represented as *rankings* over a set of alternatives — into a collective decision. Traditional social choice theory typically takes a *normative* approach, by specifying desirable axioms that the aggregation method (also known as a *voting rule*) should satisfy (Arrow, 1951). In contrast, researchers in *computational social choice* (Brandt, Conitzer, Endriss, Lang, & Procaccia, 2016) often advocate *quantitative* approaches to the same problem. The high-level idea is to identify a compelling objective function, and design voting rules that optimize this function.

Here we focus on a specific objective function: *utilitarian social welfare*. Specifically, we assume that each voter assigns a utility to each possible outcome, and the socially optimal outcome maximizes the sum of utilities. This sounds simple enough at first glance, but there is a major obstacle we must overcome: voters' preferences are expressed as *ordinal* preferences (rankings), rather than *cardinal* preferences (utilities). While this reduces the cognitive load on voters, and makes preference elicitation much easier, it does seem to be at odds with the utilitarian viewpoint.

Procaccia and Rosenschein (2006) reconcile these differences via an approach that we refer to as *implicit utilitarian voting*.[1] They propose that voters have latent utility functions, and report rankings that are consistent with these utilities, that is, the voters rank the alternatives by their utility. The performance of a voting rule — which can only access the submitted rankings, not the implicit utility functions — can then be quantified via a measure called *distortion*: the worst-case (over utility functions consistent with the reported profile of rankings) ratio between the social welfare of the optimal (welfare-maximizing) alternative, and the social welfare of the alternative selected by the voting rule. While Procaccia and Rosenschein focus on analyzing the distortion of existing voting rules, Boutilier, Caragiannis, Haber, Lu, Procaccia, and Sheffet (2015) design voting rules that minimize distortion. In particular, they bound the worst-case distortion, and show that the distortion-minimizing (randomized) voting rule can be implemented in polynomial time.

The work of Boutilier et al. (2015) provides a good understanding of optimized aggregation of rankings from the utilitarian viewpoint — but only when a single alternative is selected by the voting rule. Indeed, this understanding does not extend to common applications that require selection of a subset of alternatives, such as choosing a committee, or selecting restaurants for the next four group lunches. Our goal is therefore to

> ...build on the utilitarian approach to design optimal voting rules for selecting a subset of alternatives, and understand the guarantees they provide, as well as their performance in practice.

We make four main contributions. First, on a conceptual level, we introduce the additive notion of *regret* into the implicit utilitarian voting setting, as an alternative to the multiplicative notion of distortion. Second, in Section 3, we derive worst-case bounds on the distortion and regret of optimal deterministic and randomized voting rules. Third, in Section 4, we compare the *worst-case-optimal* deterministic voting rules with respect to distortion and regret — denoted $f^*_{\texttt{dist}}$ and $f^*_{\texttt{reg}}$, respectively — with a slew of well-known voting rules, in terms of *average-case* distortion and regret, using experiments on synthetic and real data. We find that $f^*_{\texttt{reg}}$ outperforms all other rules on average, even when measuring distortion! Fourth, in Section 5, we develop a scalable implementation for $f^*_{\texttt{reg}}$ (which, we show, is $\mathcal{NP}$-hard to compute).

1.1 Direct Real-World Implications

Research in computational social choice has frequently been justified by potential applications in multiagent systems. But recently researchers have begun to realize that, arguably, the most exciting products of this research are computer programs that help *humans* make decisions via AI-driven algorithms. One example is *Spliddit.org*, a fair division website (Goldman & Procaccia, 2014). In the voting space, existing examples include *Whale* (`whale3.noiraudes.net/whale3/`) and *Pnyx* (`pnyx.dss.in.tum.de`) — but these websites generally adopt the axiomatic viewpoint.

In November 2016, a new social choice website, *RoboVote.org*, was launched; some of us have worked on its design and implementation. The novelty of this not-for-profit website is that it relies on optimization-based approaches. For the case of objective votes — when

---

1. Cf. *utilitarian voting*, which has sporadically been used to refer to both approval voting and range voting.

a ground truth ranking of the alternatives exists (e.g., the order of different stocks by the relative change in their prices tomorrow) — RoboVote implements voting rules that pinpoint the most likely best alternative (Young, 1988), or the set most likely to contain it (Procaccia, Reddi, & Shah, 2012). For the case of subjective votes — the classic setting which is the focus of this paper, with applications to everyday scenarios such as a group of friends selecting a movie to watch or a restaurant to go to — we use the results of Boutilier et al. (2015) to select a single alternative. But, previously, the extension to subset selection was unavailable — this is precisely the motivation for the work described herein. Based on the results of Sections 4 and 5, we have implemented the deterministic regret minimization rule on RoboVote.

Interestingly, while computing the outcome of this rule requires an intricate algorithm as it is a computationally intractable problem (see Section 5), the optimality guarantee provided by the outcome is rather easy to convey to the users. The quote below illustrates how RoboVote explains the choice made by the rule on a particular instance.

> *"Assuming that the rankings submitted by voters are consistent with the values they place on the alternatives, the group of winning alternatives returned by RoboVote is guaranteed to maximize the total value up to an error of 0.2, measured on a scale in which the total value placed by all voters on all alternatives is 1. No other group of alternatives can give a better guarantee."*

1.2 Related Work

In addition to the aforementioned papers (Procaccia & Rosenschein, 2006; Boutilier et al., 2015), several other papers employ the notion of distortion to quantify how close one can get to maximizing utilitarian social welfare when only ordinal preferences are available (Caragiannis & Procaccia, 2011; Anshelevich, Bhardwaj, & Postl, 2015; Anshelevich & Postl, 2016; Anshelevich & Sekar, 2016). In particular, Anshelevich et al. (2015) study the same setting as Boutilier et al. (2015), but in addition assume the preferences of voters are consistent with distances in a metric space. Approximating utilitarian social welfare given ordinal information has also been studied in mechanism design. Filos-Ratsikas, Frederiksen, and Zhang (2014) apply this notion for finding matchings in graphs; Krysta, Manlove, Rastegari, and Zhang (2014) apply it to the house allocation problem; and Chakrabarty and Swamy (2014) study approximation of welfare under the assumption that an agent's utility for an alternative depends only on the rank of the alternative in the agent's preference order, i.e., when the utilities of all agents are dictated by a common underlying *positional scoring rule*. We refer the reader to the paper by Boutilier et al. (2015, Section 1.2) for a thorough discussion of work (in philosophy, economics, and social choice theory) related to implicit utilitarian voting more broadly.

There is quite a bit of work in computational social choice on voting rules that select subsets of alternatives. Typically it is assumed that ordinal preferences are translated into a position-based score for each alternative (in contrast to our work). Just to give a few examples, under the Chamberlin-Courant method, each voter assigns a score to a set equal to the highest score of any alternative in the set, and the (computationally hard) objective is to choose a subset of size k that maximizes the sum of scores (Chamberlin & Courant, 1983; Procaccia, Rosenschein, & Zohar, 2008). Skowron, Faliszewski, and Lang (2015) generalize

the way in which the score of a voter for a subset of alternatives is computed. The *budgeted social choice* framework of Lu and Boutilier (2011a) is more general in that the number of alternatives to be selected is not fixed; rather, each alternative has a cost that must be paid to add it to the selection.

2. The Model

Let $[t] = \{1, \ldots, t\}$. Let A be the set of *alternatives*, and denote $m = |A|$. Let $N = [n]$ be the set of *voters*. Let $\mathcal{L} = \mathcal{L}(A)$ denote the set of rankings over the alternatives. Each voter $i \in [n]$ submits a ranking $\sigma_i \in \mathcal{L}$ over the alternatives, which can alternatively be seen as a permutation of A. Therefore, $\sigma_i(a)$ is the position in which voter i ranks alternative a (1 is best, m is worst). Moreover, $a \succ_{\sigma_i} b$ denotes that voter i prefers alternative a to alternative b. The collection of voters' (submitted) rankings is called the *preference profile*, and denoted by $\vec{\sigma} \in \mathcal{L}^n$.

We assume the rankings are induced by comparisons between the voters' underlying utilities. For $i \in N$ and $a \in A$, let $u_i(a) \in [0, 1]$ be the utility of voter i for alternative a. As in previous papers (Boutilier et al., 2015; Caragiannis & Procaccia, 2011), we assume that the utilities are normalized such that $\sum_{a \in A} u_i(a) = 1$ for all $i \in N$. The collection of voter utilities, denoted $\vec{u}$, is called the *utility profile*. We say that utility profile $\vec{u}$ is consistent with preference profile $\vec{\sigma}$ — denoted $\vec{u} \triangleright \vec{\sigma}$ — if for all $a, b \in A$ and $i \in N$, $a \succ_{\sigma_i} b$ implies $u_i(a) \geqslant u_i(b)$.

Next we need to define the utility of a voter for a set of alternatives. For $S \subseteq A$, we define $u_i(S) = \max_{a \in S} u_i(a)$, that is, each voter derives utility for his favorite alternative in the set; this is in the same spirit as previous papers on set selection (Chamberlin & Courant, 1983; Monroe, 1995; Procaccia et al., 2008; Lu & Boutilier, 2011a). While there exist domains in which other definitions — for example, the case where each voter derives utility from *all* the alternatives in the set — could more accurately describe voter utilities, our definition has an additional advantage: it alleviates the "tyranny of the majority" problem associated with utilitarian welfare maximization. For instance, if a group of five voters want to select a set of five alternatives, our definition ensures that the top choice of each voter is necessarily included in the set, whereas the alternative definition could lead to a set that comprises solely of alternatives that are desired by three of the five voters. Finally, the (utilitarian) *social welfare* of S given the utility profile $\vec{u}$ is $\mathtt{sw}(S, \vec{u}) = \sum_{i=1}^{n} u_i(S)$.

We are interested in voting rules that, given a preference profile, select a subset of given cardinality k.[2] Therefore, it will be useful to denote $\mathcal{A}_k = \{S \subseteq A : |S| = k\}$. In order to unify notation, we directly define a *randomized* voting rule as a function $f : \mathcal{L}^n \to \Delta(\mathcal{A}_k)$, that is, the rule is allowed to select alternatives randomly, and formally $f(\vec{\sigma})$ is a *probability distribution* over $\mathcal{A}_k$. A *deterministic* voting rule simply gives probability 1 to a specific subset.

A voting rule can only access the preference profile $\vec{\sigma}$, yet the goal is to maximize social welfare with respect to the latent utility function $\vec{u} \triangleright \vec{\sigma}$. We study two notions that quantify how well a rule achieves this goal: distortion and regret.

---

2. Formally, this is a special case of *social choice correspondences* with fixed output cardinality (Campbell & Kelly, 1996).

The *distortion* (Procaccia & Rosenschein, 2006) of a (randomized) voting rule f on a preference profile $\vec{\sigma}$ is

$$\mathtt{dist}(f, \vec{\sigma}) = \sup_{\vec{u} \rhd \vec{\sigma}} \frac{\max_{S \in \mathcal{A}_k} \mathtt{sw}(S, \vec{u})}{\mathbb{E}[\mathtt{sw}(f(\vec{\sigma}), \vec{u})]}.$$

In words, it is the worst-case — over utility profiles consistent with the given preference profile — ratio between the social welfare of the best subset, and the expected social welfare of the subset selected by the voting rule. We define the distortion of a voting rule f by taking the worst case over preference profiles: $\mathtt{dist}(f) = \max_{\vec{\sigma} \in \mathcal{L}^n} \mathtt{dist}(f, \vec{\sigma})$.

The second measure is *regret*. While it has not been studied as part of the agenda of implicit utilitarian voting, it has been explored in other social choice settings, especially partial preferences (Lu & Boutilier, 2011b); similar measures have been extensively studied in decision theory and machine learning (Blum & Mansour, 2007; Bubeck & Cesa-Bianchi, 2012). The *regret* of a (randomized) voting rule f on a preference profile $\vec{\sigma}$ is given by

$$\mathtt{reg}(f, \vec{\sigma}) = \frac{1}{n} \cdot \sup_{\vec{u} \rhd \vec{\sigma}} \left(\max_{S \in \mathcal{A}_k} \mathtt{sw}(S, \vec{u}) - \mathbb{E}[\mathtt{sw}(f(\vec{\sigma}), \vec{u})] \right).$$

As before, define the regret of a rule f to be $\mathtt{reg}(f) = \max_{\vec{\sigma} \in \mathcal{L}^n} \mathtt{reg}(f, \vec{\sigma})$. We divide by n because the total (worst-case) regret of any voting rule f is provably linear in n (so this is *per vote* regret). Note that distortion is a multiplicative measure of loss, whereas regret is its additive version.

3. Worst-Case Bounds

In this section we provide bounds on worst-case distortion and regret, for both deterministic and randomized voting rules. Boutilier et al. (2015) show that for selecting a single winner ($k = 1$), we can achieve $O(\sqrt{m} \cdot \log^* m)$ distortion using a randomized rule, where $\log^* m$ is the iterated logarithm of m (the number of alternatives). This bound is asymptotically almost tight: they also show that the worst-case distortion is always $\Omega(\sqrt{m})$.

For a large k, though, one can hope for a better bound. Clearly, when $k = m$ there is only one voting rule (which selects every alternative), and its distortion is 1. More generally, the voting rule f that selects a subset from $\mathcal{A}_k$ uniformly at random picks each alternative in A with probability k/m, and thus achieves social welfare of $n \cdot k/m$ in expectation. As social welfare cannot exceed n, we have $\mathtt{dist}(f) \leqslant m/k$. However, since we can already achieve $O(\sqrt{m} \cdot \log^* m)$ distortion for $k = 1$, a bound of m/k provides an improvement only for $k = \Omega(\sqrt{m}/\log^* m)$. Can we achieve better distortion for smaller values of k as well? It is not even clear whether the optimal worst-case distortion should monotonically decrease in k, because as our flexibility grows with k, so does the flexibility of the welfare-maximizing solution. In fact, a part of our main result shows that the worst-case distortion remains $\Omega(\sqrt{m})$ for all values of k up to $\Theta(\sqrt{m})$.

Theorem 1. *Let $m = |A|$, and let k be the number of alternatives to be selected.*

1. **Distortion, deterministic rules:** *There exists a deterministic voting rule f^* with $\texttt{dist}(f^*) \leqslant 1 + m\ (m-k)/k$. Moreover, for every deterministic voting rule f,*

$$\texttt{dist}(f) \geqslant \begin{cases} 1 + \frac{m(m-3k)}{6k} & \text{if } k \leqslant \frac{m}{9}, \\ 1 + m & \text{if } \frac{m}{9} < k \leqslant \frac{m}{2}, \\ 1 + \frac{m(m-k)}{k} & \text{otherwise.} \end{cases}$$

   These bounds are tight up to a constant factor of 8.

2. **Distortion, randomized rules:** *There exists a randomized voting rule f^* such that*

$$\texttt{dist}(f^*) \leqslant \begin{cases} 2\sqrt{m \cdot H_m} & \text{if } k \leqslant \frac{2 \cdot m \cdot H_m}{m + H_m}, \\ 4\sqrt{m \cdot k} & \text{if } \frac{2 \cdot m \cdot H_m}{m + H_m} < k \leqslant \left(\frac{m}{4}\right)^{\frac{1}{3}}, \\ \frac{m}{k} & \text{otherwise,} \end{cases}$$

   where $H_m = \Theta(\log m)$ is the mth harmonic number. Moreover, for every randomized voting rule f,

$$\texttt{dist}(f) \geqslant \begin{cases} \frac{\sqrt{m}}{2} & \text{if } k \leqslant \frac{m \cdot (\sqrt{m}-1)}{m-1} \approx \sqrt{m}, \\ \frac{m}{k+m/k} & \text{otherwise.} \end{cases}$$

   These bounds are tight up to a factor of $6.35 \cdot m^{1/6}$.

3. **Regret, deterministic rules:** *There exists a deterministic voting rule f^* such that*

$$\texttt{reg}(f^*) \leqslant \begin{cases} \frac{1}{2} & \text{if } k \leqslant \frac{m}{2}, \\ 1 - \frac{k}{m} & \text{otherwise,} \end{cases}$$

   and this upper bound is completely tight.

4. **Regret, randomized rules:** *There exists a randomized voting rule f^* such that*

$$\texttt{reg}(f^*) \leqslant \frac{1}{2} \cdot \left(1 - \frac{k^2}{m^2}\right).$$

   Moreover, for every randomized voting rule f,

$$\texttt{reg}(f) \geqslant \begin{cases} \frac{1}{4} & \text{if } k \leqslant m/2 \\ \frac{1}{2} \cdot \frac{k}{m}\left(1 - \frac{k}{m}\right) & \text{otherwise.} \end{cases}$$

   These bounds are tight up to a constant factor of 2.

All the upper bounds above can be achieved via polynomial-time algorithms.

The bounds presented above are simplified forms of the exact bounds that we derive. Figure 1 shows our exact bounds for $m = 100$.[3]

Before we dive into the proof, we simplify the formulae for the distortion and regret of deterministic voting rules.

---

3. The second upper bound in part 2 of Theorem 1 (which increases with k) does not play a role unless m is very large.
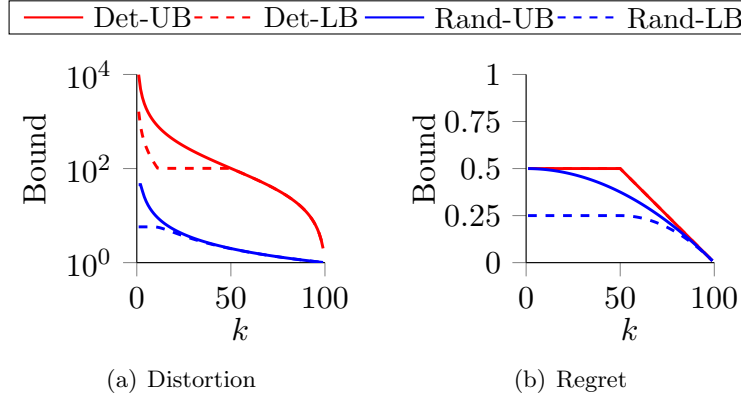
(a) Distortion (b) Regret

Fig. 1: The upper and lower bounds on worst-case distortion and regret for $m = 100$.

Definition 1 (Comparing Sets). *Given a ranking $\sigma \in \mathcal{L}$ and an alternative $a \in A$, recall that $\sigma(a)$ denotes the position of a in σ. More generally, for a set $S \subseteq A$ let $\sigma(S) = \min_{a \in S} \sigma(a)$. For sets $S, T \subseteq A$, we say $T \succ_\sigma S$ if $\sigma(T) < \sigma(S)$, i.e., if there exists an alternative in T that is preferred to every alternative in S in σ.*

Definition 2 (Plurality Score). *The plurality score of an alternative $a \in A$ in a preference profile $\vec{\sigma}$ is the number of votes in which a appears first, i.e.,*

$$\texttt{plu}(a, \vec{\sigma}) = \sum_{i=1}^{n} \mathbb{I}[\sigma_i(1) = a],$$

where $\mathbb{I}$ is the indicator function. More generally, we define the plurality score of a set $S \subseteq A$ to be the number of votes in which an alternative in S is ranked first, i.e.,

$$\texttt{plu}(S, \vec{\sigma}) = \sum_{i=1}^{n} \mathbb{I}[\sigma_i(S) = 1] = \sum_{a \in S} \texttt{plu}(a, \vec{\sigma}).$$

Lemma 1. *For a deterministic voting rule f and a preference profile $\vec{\sigma}$, the regret of f on $\vec{\sigma}$ is given by*

$$\texttt{reg}(f, \vec{\sigma}) = \max_{S \in \mathcal{A}_k} \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{\mathbb{I}[S \succ_{\sigma_i} f(\vec{\sigma})]}{\sigma_i(S)}, \tag{1}$$

and the distortion of f on $\vec{\sigma}$ is given by

$$\texttt{dist}(f, \vec{\sigma}) = 1 + m \cdot \frac{n \cdot \texttt{reg}(f, \vec{\sigma})}{\texttt{plu}(f(\vec{\sigma}), \vec{\sigma})}. \tag{2}$$

Proof. First, note that $\texttt{reg}(f, \vec{\sigma})$ and $\texttt{dist}(f, \vec{\sigma})$ can be rewritten as follows.

$$\texttt{reg}(f, \vec{\sigma}) = \frac{1}{n} \cdot \sup_{\vec{u} \rhd \vec{\sigma}} \left[\max_{S \in \mathcal{A}_k} \texttt{sw}(S, \vec{u}) - \texttt{sw}(f(\vec{\sigma}), \vec{u}) \right] = \frac{1}{n} \cdot \left[\max_{S \in \mathcal{A}_k} \sup_{\vec{u} \rhd \vec{\sigma}} \texttt{sw}(S, \vec{u}) - \texttt{sw}(f(\vec{\sigma}), \vec{u}) \right], \tag{3}$$

129

and similarly,

$$\text{dist}(f, \vec{\sigma}) = \sup_{\vec{u} \rhd \vec{\sigma}} \frac{\max_{S \in \mathcal{A}_k} \text{sw}(S, \vec{u})}{\text{sw}(f(\vec{\sigma}), \vec{u})} = \max_{S \in \mathcal{A}_k} \sup_{\vec{u} \rhd \vec{\sigma}} \frac{\text{sw}(S, \vec{u})}{\text{sw}(f(\vec{\sigma}), \vec{u})}. \tag{4}$$

If $S = f(\vec{\sigma})$, then the regret term is 0 and the distortion term is 1. Fix $S \in \mathcal{A}_k \setminus \{f(\vec{\sigma})\}$. To maximize the regret (resp. distortion) term, we want to find the utility profile $\vec{u}$ that maximizes the difference (resp. ratio) of $\text{sw}(S, \vec{u})$ and $\text{sw}(f(\vec{\sigma}), \vec{u})$ subject to $\vec{u} \rhd \vec{\sigma}$. For each voter i for which $S \succ_i f(\vec{\sigma})$, we can maximize the discrepancy between $u_i(S)$ and $u_i(f(\vec{\sigma}))$, for both the regret and the distortion objectives, by letting $u_i(f(\vec{\sigma})) = 0$, and maximizing $u_i(S)$ by setting the utility of the first $\sigma_i(S)$ alternatives as $1/\sigma_i(S)$. For each voter i for which $f(\vec{\sigma}) \succ_i S$, we have $u_i(S) \leqslant u_i(f(\vec{\sigma}))$. The best we can hope for is to make the two utilities equal. While this is sufficient to optimize the regret objective, optimizing distortion further requires the two utilities to be the smallest possible for such voters. Hence, for $\sigma_i(f(\vec{\sigma})) > 1$, we need both utilities to be 0, and for $\sigma_i(f(\vec{\sigma})) = 1$, we need both to be $1/m$, which is the smallest possible value of $u_i(f(\vec{\sigma}))$ in this case. This intuition leads us to the following explicit utility profile $\vec{u}^*$. For each voter i,

1. If $\sigma_i(f(\vec{\sigma})) = 1$, let $u_i(a) = 1/m$ for all $a \in A$.

2. If $S \succ_i f(\vec{\sigma})$, let $u_i(a) = 1/\sigma_i(S)$ if $\sigma_i(a) \in [\sigma_i(S)]$, and $u_i(a) = 0$ otherwise.

3. If $S \not\succ_i f(\vec{\sigma})$ and $\sigma_i(f(\vec{\sigma})) \neq 1$, let $u_i(a) = 1$ if $\sigma_i(a) = 1$, and $u_i(a) = 0$ otherwise.

First, it is easy to check that $\vec{u}^* \rhd \vec{\sigma}$. Also, note that this utility profile maximizes $u_i(S) - u_i(f(\vec{\sigma}))$ subject to $u_i \rhd \sigma_i$, for each voter i in each of the three cases above. Hence, it maximizes the regret term $\text{sw}(S, \vec{u}) - \text{sw}(f(\vec{\sigma}), \vec{u})$. Further, we have

$$\text{sw}(f(\vec{\sigma}), \vec{u}^*) = \frac{1}{m} \cdot \text{plu}(f(\vec{\sigma}), \vec{\sigma}),$$

$$\text{sw}(S, \vec{u}^*) = \frac{1}{m} \cdot \text{plu}(f(\vec{\sigma}), \vec{\sigma}) + \sum_{i=1}^{n} \frac{\mathbb{I}[S \succ_i f(\vec{\sigma})]}{\sigma_i(S)}.$$

This immediately gives us Equation (1) for the regret of f on $\vec{\sigma}$. Now, the distortion of f on $\vec{\sigma}$ under the utility profile $\vec{u}^*$ is

$$\frac{\text{sw}(S, \vec{u}^*)}{\text{sw}(f(\vec{\sigma}), \vec{u}^*)} = 1 + m \cdot \frac{\sum_{i=1}^{n} \mathbb{I}[S \succ_i f(\vec{\sigma})]/\sigma_i(S)}{\text{plu}(f(\vec{\sigma}), \vec{\sigma})} = 1 + m \cdot \frac{n \cdot \text{reg}(f, \vec{\sigma})}{\text{plu}(f(\vec{\sigma}), \vec{\sigma})} \geqslant 1. \tag{5}$$

Finally, take a utility profile $\vec{u}$ satisfying $\vec{u} \rhd \vec{\sigma}$. We want to show that the distortion under $\vec{u}$, i.e., $\text{sw}(S, \vec{u})/\text{sw}(f(\vec{\sigma}), \vec{u})$ is no more than the distortion under $\vec{u}^*$. Note that $f(\vec{\sigma})$ has the least possible welfare under $\vec{u}^*$. Hence, $\text{sw}(f(\vec{\sigma}), \vec{u}) \geqslant \text{sw}(f(\vec{\sigma}), \vec{u}^*)$. To achieve a greater distortion, we must have $\text{sw}(S, \vec{u}) > \text{sw}(S, \vec{u}^*)$, i.e., the voters must assign a greater utility to S under $\vec{u}$ than under $\vec{u}^*$.

Let us revisit the three cases in the construction of $\vec{u}^*$. All the voters covered under case 2 already assign S the highest possible utility. For all the other voters, the top-ranked alternative of $f(\vec{\sigma})$ is at least as high as the top-ranked alternative of S. Hence, an increase of δ in the utility for S would require an increase of at least δ in the utility for $f(\vec{\sigma})$.

That is, for any $\delta > 0$, $\mathtt{sw}(S, \vec{u}) = \mathtt{sw}(S, \vec{u}^*) + \delta$ implies $\mathtt{sw}(f(\vec{\sigma}), \vec{u}) \geqslant \mathtt{sw}(f(\vec{\sigma}), \vec{u}^*) + \delta$. Finally,

$$\frac{\mathtt{sw}(S, \vec{u}^*)}{\mathtt{sw}(f(\vec{\sigma}), \vec{u}^*)} \geqslant 1 \Rightarrow \frac{\mathtt{sw}(S, \vec{u})}{\mathtt{sw}(f(\vec{\sigma}), \vec{u})} \leqslant \frac{\mathtt{sw}(S, \vec{u}^*) + \delta}{\mathtt{sw}(f(\vec{\sigma}), \vec{u}^*) + \delta} \leqslant \frac{\mathtt{sw}(S, \vec{u}^*)}{\mathtt{sw}(f(\vec{\sigma}), \vec{u}^*)}.$$

Hence, the worst-case distortion is indeed $1 + m \cdot \mathtt{reg}(f, \vec{\sigma})/\mathtt{plu}(f(\vec{\sigma}), \vec{\sigma})$, as required. Note that in finding the worst-case distortion, the distortion of 1 achieved with $S = f(\vec{\sigma})$ is ignored because the distortion achieved by every $S \neq f(\vec{\sigma})$ is at least 1. ∎

One interesting consequence of Lemma 1 is that selecting a set of alternatives, none of which appear at the top position in any vote, results in an unbounded distortion. Hence, the rule that optimizes distortion would never select such a set. We are now ready for the proof of our main result.

Proof of Theorem 1. Below, we provide a proof for the upper and the lower bound in each of the four cases.

Distortion, deterministic rules:

Upper bound: Inspired by the denominator $\mathtt{plu}(f(\vec{\sigma}), \vec{\sigma}) = \sum_{a \in f(\vec{\sigma})} \mathtt{plu}(a, \vec{\sigma})$ in the formula for the distortion of f on $\vec{\sigma}$ from Lemma 1, let us analyze the rule f^* that selects the k alternatives with the highest plurality scores. We show that f^* achieves the required upper bound on the optimal distortion.

Since the sum of plurality scores of all the alternatives is n, the sum of top k plurality scores is at least $n \cdot k/m$. Hence, $\mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma}) \geqslant n \cdot k/m$. Next, for $S \in \mathcal{A}_k \setminus \{f^*(\vec{\sigma})\}$, note that the number of voters i for which $S \succ_i f^*(\vec{\sigma})$ is at most $n - \mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})$. Using Lemma 1, it follows that the distortion of f^* on $\vec{\sigma}$ is at most

$$1 + m \cdot \frac{n - \mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})}{\mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})} = 1 + m \cdot \left(\frac{n}{\mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})} - 1 \right) \leqslant 1 + m \cdot \left(\frac{m}{k} - 1 \right),$$

which is the required upper bound.

Lower bound: Next, we establish three different lower bounds on the distortion of deterministic rules.

1. For $k \leqslant m/6$, $\mathtt{dist}(f) \geqslant 1 + m \cdot (m - 3k)/(6k)$ for every deterministic rule f.

2. For $k \leqslant m/2$, $\mathtt{dist}(f) \geqslant 1 + m$ for every deterministic rule f.

3. For $k \geqslant m/2$, $\mathtt{dist}(f) \geqslant 1 + m \cdot (m - k)/k$ for every deterministic rule f.

It is easy to check that for $k \leqslant m/9$, the first bound is the strongest; for $k \in [m/9, m/2]$, the second bound is the strongest; and, for $k \geqslant m/2$, the third bound is the strongest. Hence, the optimal combination of these three bounds gives us the desired result.

Let the set of alternatives be $A = \{a_1, \ldots, a_m\}$. Let us begin by proving the first bound. Fix a value of $k \leqslant m/6$, and partition A into two sets: $X = \{a_1, \ldots, a_{m-3k}\}$ and $Y = A \setminus X$. Let $Y = \{b_1, \ldots, b_{3k}\}$. Note that $m \geqslant 6k$ implies $|X| \geqslant |Y|$. Construct a ranked profile $\vec{\sigma}$ as follows.

- For each alternative $a_j \in X$ (thus $j \leqslant m - 3k$), let a_j be ranked first in the votes of voters $i \in [(j-1) \cdot n/(m-3k) + 1, j \cdot n/(m-3k)]$. That is, we partition the set of voters into $|X| = m - 3k$ contiguous blocks, and have each alternative of X ranked first in one of the blocks.

- For each alternative $b_j \in Y$ (thus $j \leqslant 3k$), let b_j be ranked second in the votes of voters $i \in [(j-1) \cdot n/3k + 1, j \cdot n/3k]$. That is, we partition the set of voters into $|Y| = 3k$ contiguous blocks, and have each alternative of Y ranked second in one of the blocks.

Since we chose the blocks of voters to be contiguous in both cases, it follows that for every $a_j \in X$, the set of voters ranking a_j first can have at most two distinct alternatives in Y in their second position. Take a deterministic rule f for selecting a set of k alternatives. Let $|f(\vec{\sigma}) \cap X| = t \leqslant k$. Then, we have $\texttt{plu}(f(\vec{\sigma}), \vec{\sigma}) = t \cdot n/(m-3k)$.

Consider the voters who rank an alternative of $f(\vec{\sigma})$ first. Let Y' denote the set of alternatives appearing in the second position in the votes of such voters. From the argument above, we have $|Y'| \leqslant 2t \leqslant 2k$. Hence, $|Y \setminus Y'| \geqslant 3k - 2k = k$. Choose an arbitrary set $S \subseteq Y \setminus Y'$ such that $|S| = k$. Now, there are $(n/3k) \cdot k = n/3$ voters that rank an alternative in S in their second position. Hence,

$$\texttt{dist}(f) \geqslant \texttt{dist}(f, \vec{\sigma}) \geqslant 1 + m \cdot \frac{\sum_{i=1}^{n} \mathbb{I}[S \succ_i f(\vec{\sigma})]/\sigma_i(S)}{\texttt{plu}(f(\vec{\sigma}), \vec{\sigma})} \geqslant 1 + m \cdot \frac{(n/3) \cdot (1/2)}{t \cdot n/(m-3k)}$$
$$\geqslant 1 + m \cdot \frac{m - 3k}{6k},$$

where the second transition uses Lemma 1, and the final transition uses $t \leqslant k$.

For the second and the third lower bound, we simply construct a profile $\vec{\sigma}$ in which each alternative in A appears first in n/m votes, and the remaining positions in the votes are filled arbitrarily. Fix a deterministic rule f with $|f(\vec{\sigma})| = k$. Note that $\texttt{plu}(f(\vec{\sigma}), \vec{\sigma}) = (n/m) \cdot k$.

If $k \leqslant m/2$, choose a set $S \subseteq A \setminus f(\vec{\sigma})$ such that $|S| = k$. Note that $S \cap f(\vec{\sigma}) = \emptyset$, and an alternative in S is ranked first in $(n/m) \cdot k$ votes. Hence, by Lemma 1,

$$\texttt{dist}(f) \geqslant \texttt{dist}(f, \vec{\sigma}) \geqslant 1 + m \cdot \frac{\sum_{i=1}^{n} \mathbb{I}[S \succ_i f(\vec{\sigma})]/\sigma_i(S)}{\texttt{plu}(f(\vec{\sigma}), \vec{\sigma})} \geqslant 1 + m \cdot \frac{(n/m) \cdot k}{(n/m) \cdot k} = 1 + m.$$

If $k > m/2$, choose $S \supseteq A \setminus f(\vec{\sigma})$ such that $|S| = k$. In this case, an alternative in $S \setminus f(\vec{\sigma})$ is ranked first in $(n/m) \cdot (m-k)$ votes. Hence, by Lemma 1,

$$\texttt{dist}(f) \geqslant \texttt{dist}(f, \vec{\sigma}) \geqslant 1 + m \cdot \frac{\sum_{i=1}^{n} \mathbb{I}[S \succ_i f(\vec{\sigma})]/\sigma_i(S)}{\texttt{plu}(f(\vec{\sigma}), \vec{\sigma})} \geqslant 1 + m \cdot \frac{(n/m) \cdot (m-k)}{(n/m) \cdot k}$$
$$= 1 + m \cdot \frac{m - k}{k},$$

as required.

Gap between upper and lower bounds: Note that the gap between the upper and the lower bounds is $\max(G_1, G_2, G_3)$, where

$$G_1 = \max_{k \in [1, \frac{m}{9}]} \frac{1 + m(m-k)/k}{1 + m(m-3k)/(6k)}, \quad G_2 = \max_{k \in [\frac{m}{9}, \frac{m}{2}]} \frac{1 + m(m-k)/k}{1 + m}, \quad \text{and } G_3 = 1.$$

For G_1, it can be checked that the ratio of the upper and lower bounds is an increasing function of k for $m \geqslant 3$. Hence, the maximum is achieved at $k = m/9$, and is equal to $1 + 8 \cdot m/(1 + m) \leqslant 8$.

For G_2, the ratio of the upper and the lower bounds is clearly a decreasing function of k. Hence, the maximum is achieved at $k = m/9$, and is equal to $G_1 \leqslant 8$.

Hence, the upper and the lower bounds are tight up to a constant factor of 8.

Distortion, randomized rules:

Upper bound: In our opinion, the proof of this part is the most non-trivial. It uses a construction that builds on the one used by Boutilier et al. (2015) for $k = 1$, but requires additional tools and introduces novel techniques. As mentioned at the beginning of this section, choosing a set uniformly at random from $\mathcal{A}_k$ (under which the marginal probability of every alternative being chosen is k/m) has distortion at most m/k. However, this approach does not work well if some alternatives are significantly better than others.

In that case, one may wish to choose the alternatives with probabilities proportional to their "quality". For $a \in A$, let us define its quality by its harmonic score $\mathtt{har}(a, \vec{\sigma}) = \sum_{i \in [n]} 1/\sigma_i(a)$. Then, we wish to choose alternative a with marginal "probability" $k \cdot \mathtt{har}(a, \vec{\sigma})/\sum_{b \in A} \mathtt{har}(b, \vec{\sigma})$. Note that this quantity may be greater than 1. Moreover, this approach fails when all sets are almost equally good. Hence, we employ a combination of the two approaches.

Fix $0 \leqslant \alpha \leqslant 1$, and for an alternative $a \in A$ define

$$p_a = \alpha \cdot \frac{k}{m} + (1 - \alpha) \cdot \frac{k \cdot \mathtt{har}(a, \vec{\sigma})}{\sum_{b \in A} \mathtt{har}(b, \vec{\sigma})}. \tag{6}$$

Using the bihierarchy extension (Budish, Che, Kojima, & Milgrom, 2013) of the Birkhoff-von Neumann theorem (Birkhoff, 1946; von Neumann, 1953), we can show that there exists a distribution over $\mathcal{A}_k$ under which the marginal probabilities of selected alternatives match those dictated by Equation (6) if and only if

$$\forall a \in A, 0 \leqslant p_a \leqslant 1 \quad \text{and} \quad \sum_{a \in A} p_a = k.$$

Note that $p_a \geqslant 0$ and $\sum_{a \in A} p_a = k$. The constraint $p_a \leqslant 1$ will be applied later to restrict the feasible values of α. For now, suppose such a distribution D exists. Consider a preference profile $\vec{\sigma}$ and a utility profile $\vec{u} \rhd \vec{\sigma}$. Let $S^* \in \arg\max_{S \in \mathcal{A}_k} \mathtt{sw}(S, \vec{u})$. Define

$$X = \sqrt{\frac{H_m}{m} \cdot \frac{\alpha}{1 - \alpha}},$$

where $H_m = \sum_{t=1}^m 1/t$ is the mth harmonic number. Note that $\sum_{a \in A} \mathtt{har}(a, \vec{\sigma}) = n \cdot H_m$. Now, consider two cases.

Case 1: Suppose $\mathtt{sw}(S^*, \vec{u}) \leqslant n \cdot X$. Then,

$$\mathbb{E}_{S \sim D}[\mathtt{sw}(S, \vec{u})] = \sum_{S \in \mathcal{A}_k} \Pr_D[S] \cdot \left(\sum_{i=1}^n \max_{a \in S} u_i(a) \right) \geqslant \sum_{i=1}^n \left(\sum_{S \in \mathcal{A}_k} \Pr_D[S] \cdot \frac{\sum_{a \in S} u_i(a)}{k} \right)$$

$$= \frac{1}{k} \sum_{i=1}^n \sum_{a \in A} u_i(a) \cdot \Pr_{S \sim D}[a \in S] \geqslant \frac{1}{k} \sum_{i=1}^n \sum_{a \in A} u_i(a) \cdot \alpha \cdot \frac{k}{m} = \alpha \cdot \frac{n}{m}.$$

Hence, the distortion is

$$\frac{\mathtt{sw}(S^*, \vec{u})}{\mathbb{E}_{S \sim D}[\mathtt{sw}(S, \vec{u})]} \leqslant \frac{n \cdot X}{\alpha \cdot n/m} = \frac{X \cdot m}{\alpha} = \sqrt{\frac{m \cdot H_m}{\alpha \cdot (1 - \alpha)}}.$$

Case 2: Suppose $\mathtt{sw}(S^*, \vec{u}) > n \cdot X$. Then, for each alternative $a \in S^*$, let N_a denote the subset of voters who rank a above any other alternative of S^*, i.e.,

$$N_a = \{i \in [n] : \forall b \in S^* \setminus \{a\}, \ a \succ_{\sigma_i} b, \}.$$

Let $\mathtt{sw}_{N_a}(S, \vec{u})$ denote the welfare of the voters in N_a for the set of alternatives S under the utility profile $\vec{u}$. Let T_a denote the total utility that agents in N_a have for alternative a, i.e., $T_a = \sum_{i \in N_a} u_i(a)$. For all $a \in A$, we have $\mathtt{har}(a, \vec{\sigma}) \geqslant T_a$ because $u_i(a) \leqslant 1/\sigma_i(a)$ for each voter $i \in [n]$. Because $\{N_a\}_{a \in S^*}$ is a partition of the set of voters, we have

$$\mathbb{E}_{S \sim D}[\mathtt{sw}(S, \vec{u})] = \mathbb{E}_{S \sim D}\left[\sum_{a \in S^*} \mathtt{sw}_{N_a}(S, \vec{u})\right] \geqslant \sum_{a \in S^*} T_a \cdot \mathrm{Pr}_{S \sim D}[a \in S]$$

$$\geqslant \sum_{a \in S^*} T_a \cdot (1 - \alpha) \cdot \frac{k \cdot \mathtt{har}(a, \vec{\sigma})}{\sum_{b \in A} \mathtt{har}(b, \vec{\sigma})} \geqslant \frac{(1 - \alpha) \cdot k}{n \cdot H_m} \cdot \sum_{a \in S^*} (T_a)^2$$

$$\geqslant \frac{1 - \alpha}{n \cdot H_m} \cdot \left(\sum_{a \in S^*} T_a\right)^2 = \frac{1 - \alpha}{n \cdot H_m} \cdot (\mathtt{sw}(S^*, \vec{u}))^2.$$

Here, the fourth transition uses $\mathtt{har}(a, \vec{\sigma}) \geqslant T_a$, the fifth transition uses the power-mean inequality, and the final transition uses $\mathtt{sw}(S^*, \vec{u}) = \sum_{a \in S^*} T_a$. Now, the distortion is

$$\frac{\mathtt{sw}(S^*, \vec{u})}{\mathbb{E}_{S \sim D}[\mathtt{sw}(S, \vec{u})]} \leqslant \frac{n \cdot H_m}{(1 - \alpha) \cdot \mathtt{sw}(S^*, \vec{u})} < \sqrt{\frac{m \cdot H_m}{\alpha \cdot (1 - \alpha)}},$$

where the final transition uses our assumption $\mathtt{sw}(S^*, \vec{u}) > n \cdot X$ along with the definition of X.

Combined analysis: In both cases, the distortion is at most $\sqrt{m H_m / (\alpha(1 - \alpha))}$. The final step involves choosing the optimal value of α by minimizing this quantity subject to our constraints: $p_a \leqslant 1$ for all $a \in A$. This translates to

$$\alpha \cdot \frac{k}{m} + (1 - \alpha) \cdot \frac{k \cdot \mathtt{har}(a, \vec{\sigma})}{\sum_{b \in A} \mathtt{har}(b, \vec{\sigma})} \leqslant 1, \forall a \in A.$$

Note that $\mathtt{har}(a, \vec{\sigma}) \leqslant n$, and $\sum_{b \in A} \mathtt{har}(b, \vec{\sigma}) = n \cdot H_m$. Hence, we can safely replace these constraints by the following constraint:

$$\alpha \cdot \frac{k}{m} + (1 - \alpha) \cdot \frac{k}{H_m} \leqslant 1.$$

We used Mathematica to minimizing the value of $\sqrt{m H_m / (\alpha(1 - \alpha))}$ subject to this constraint; see Figure 2 for the relevant code. The result is as follows.

$\text{OPT} = \text{FullSimplify}\left[\text{Minimize}\left[\left\{\sqrt{\frac{m*H_m}{\alpha*(1-\alpha)}}, \alpha \geqslant 0 \wedge \alpha \leqslant 1 \wedge \alpha * \frac{k}{m} + (1-\alpha) * \frac{k}{H_m} \leqslant 1\right\}, \alpha\right],\right.$
$\left. m \geqslant 2 \wedge k < m \wedge k \geqslant 1 \wedge H_m > 0 \wedge m > H_m\right];$

OPT[[1]]   (* **Optimal Distortion** *)

$$\begin{cases} \dfrac{k(m-H_m)}{\sqrt{(-k+m)(k-H_m)}} & km + (k-2m)H_m \geqslant 0 \\[2ex] 2\sqrt{mH_m} & \text{True} \end{cases}$$

OPT[[2]]   (* **Optimal α value** *)

$$\alpha \to \begin{cases} \dfrac{1}{2} & km + (k-2m)H_m < 0 \\[2ex] \dfrac{m(k-H_m)}{k(m-H_m)} & \text{True} \end{cases}$$

Fig. 2: Optimizing the upper bound on distortion using Mathematica

- For $k \leqslant 2mH_m/(m+H_m)$, the optimal distortion is achieved at $\alpha = 1/2$, and is equal to $2\sqrt{m \cdot H_m}$;

- For $k > 2mH_m/(m+H_m)$, the optimal distortion is achieved at $\alpha = m \cdot (k-H_m)/(k \cdot (m-H_m))$, and is equal to
$$\frac{k \cdot (m-H_m)}{\sqrt{(m-k) \cdot (k-H_m)}}.$$

  Note that $H_m \leqslant (7/8) \cdot k$. Further, let $k \leqslant m/2$ (as this bound will anyway be replaced by a better bound for $k > m/2$). Hence, $m - k \geqslant m/2$. Substituting these, we get that the optimal distortion in this case is at most $4\sqrt{m \cdot k}$.

Finally, recall that we have a universal upper bound of m/k, achieved by choosing from $\mathcal{A}_k$ uniformly at random. One can check that $m/k < 4\sqrt{m \cdot k}$ for $k > (m/4)^{1/3}$. Hence, combining the upper bounds from the analysis above with this universal upper bound gives us the desired result.

Lower bound: Fix a set of alternatives $T = \{a_1, \ldots, a_t\}$, where $t \geqslant k$ (the exact value of t will be determined later). Partition the set of voters into t buckets; each bucket i, denoted by N_i, consists of n/t voters. Construct a ranked profile $\vec{\sigma}$ in which for all $i \in [t]$, all voters in bucket N_i rank alternative a_i first, and the remaining alternatives arbitrarily.

Let us analyze the output of a randomized rule f on this profile. For each alternative $a \in A$, define $p_a = \Pr_{S \sim f(\vec{\sigma})}[a \in S]$. Let $X \subseteq [t]$ be the indices corresponding to the lowest k values in the sequence $(p_{a_i})_{i \in [t]}$; in other words, let X be such that $\{a_i : i \in X\}$ is the set of k alternatives from T with the lowest values of p. We now construct a utility profile $\vec{u}$ that is a modification of the one used in the lower bound for the deterministic case; instead of letting voters have high utility for alternatives that are *not selected*, we let voters have high utility for alternatives that are *selected with the lowest probabilities*.

1. For each $i \in [t] \setminus X$, every voter in bucket N_i has utility $1/m$ for each alternative.

(* Variable $x = k/t$ *)

OPT = FullSimplify $\left[\text{Maximize}\left[\left\{\frac{(1-x)(1/m)+x}{(1-x)(1/m)+x^2}, x \geqslant \frac{k}{m} \wedge x \leqslant 1\right\}, x\right],\right.$

$m > 1 \wedge k < m \wedge k \geqslant 1]$;

OPT[[1]]   (* Optimal Distortion Bound *)

$$\begin{cases} \frac{k(-1+m)+m}{(-1+k)k+m} & k(2+k) + \sqrt{k^3(4+k)} \geqslant 2m \\ \frac{(1+\sqrt{m})^2}{1+2\sqrt{m}} & \text{True} \end{cases}$$

Solve $\left[k(2+k) + \sqrt{k^3(4+k)} == 2m, k\right]$ (* Simplify the condition on k *)

$\left\{\left\{k \rightarrow \frac{-m-m^{3/2}}{-1+m}\right\}, \left\{k \rightarrow \frac{-m+m^{3/2}}{-1+m}\right\}\right\}$

Fig. 3: Optimizing the lower bound on distortion using Mathematica

2. For each $i \in X$, every voter in bucket N_i has utility 1 for the alternative it ranks first (i.e., a_i), and utility 0 for the remaining alternatives.

First, note that $\vec{u} \rhd \vec{\sigma}$. Further, under $\vec{u}$ the optimal set of k alternatives is clearly $\{a_i : i \in X\}$, and its corresponding welfare is

$$\left(n - \frac{n}{t} \cdot k\right) \cdot \frac{1}{m} + \frac{n}{t} \cdot k \cdot 1,$$

because it provides utility $1/m$ to every voter in bucket N_i for $i \in [t] \setminus X$, and utility 1 to every voter in bucket N_i for $i \in X$. In contrast, the expected welfare under f is

$$\left(n - \frac{n}{t} \cdot k\right) \cdot \frac{1}{m} + \frac{n}{t} \cdot \sum_{i \in X} p_{a_i}. \tag{7}$$

Next, note that $\sum_{a \in T} p_a \leqslant \sum_{a \in A} p_a = k$, where the last equality follows because f always returns a set of size k. Hence, the sum of the lowest k values from $\{p_a : a \in T\}$ (i.e., $\sum_{i \in X} p_{a_i}$, by the definition of X) is at most $(k/t) \cdot k$. Substituting this in Equation (7), we obtain that the worst-case distortion is bounded from below by

$$\frac{\left(n - \frac{n}{t} \cdot k\right) \cdot \frac{1}{m} + \frac{n}{t} \cdot k}{\left(n - \frac{n}{t} \cdot k\right) \cdot \frac{1}{m} + \frac{n}{t} \cdot \frac{k^2}{t}} = \frac{\left(1 - \frac{k}{t}\right) \cdot \frac{1}{m} + \frac{k}{t}}{\left(1 - \frac{k}{t}\right) \cdot \frac{1}{m} + \left(\frac{k}{t}\right)^2}.$$

Finally, we minimize this with respect to t using Mathematica; the relevant code, in which k/t is represented as variable x, is presented in Figure 3. The result is as follows.

$$\text{dist}(f) \geqslant \text{dist}(f, \vec{\sigma}) \geqslant \begin{cases} \frac{(\sqrt{m}+1)^2}{1+2\sqrt{m}} & \text{if } k \leqslant \frac{m(\sqrt{m}-1)}{m-1}, \\ \frac{m+k(m-1)}{m+k(k-1)} & \text{otherwise.} \end{cases}$$

Finally, note that

$$\frac{(\sqrt{m}+1)^2}{1+2\sqrt{m}} \geqslant \frac{(\sqrt{m}+1)^2}{2 \cdot (1+\sqrt{m})} = \frac{\sqrt{m}+1}{2} \geqslant \frac{\sqrt{m}}{2},$$

and

$$\frac{m+k(m-1)}{m+k(k-1)} = \frac{m \cdot k + m - k}{k^2 + m - k} \geqslant \frac{m \cdot k}{k^2 + m} = \frac{m}{k + m/k},$$

which are the required bounds.

Gap between upper and lower bounds: In this case, the gap between the upper and the lower bounds is $\max(G_1, G_2, G_3, G_4)$, where

$$G_1 = \max_{k \in \left[1, \frac{2mH_m}{m+H_m}\right]} \frac{2\sqrt{m \cdot H_m}}{\sqrt{m}/2} \leqslant 4\sqrt{H_m},$$

$$G_2 = \max_{k \in \left[\frac{2mH_m}{m+H_m}, \left(\frac{m}{4}\right)^{\frac{1}{3}}\right]} \frac{4\sqrt{m \cdot k}}{\sqrt{m}/2} \leqslant \max_{k \in \left[\frac{2mH_m}{m+H_m}, \left(\frac{m}{4}\right)^{\frac{1}{3}}\right]} 8 \cdot \sqrt{k} \leqslant 2^{8/3} \cdot m^{1/6},$$

$$G_3 = \max_{k \in \left[\left(\frac{m}{4}\right)^{\frac{1}{3}}, \frac{m(\sqrt{m}-1)}{m-1}\right]} \frac{m/k}{\sqrt{m}/2} \leqslant \max_{k \in \left[\left(\frac{m}{4}\right)^{\frac{1}{3}}, \frac{m(\sqrt{m}-1)}{m-1}\right]} 2 \cdot \frac{\sqrt{m}}{k} \leqslant 2 \cdot \frac{\sqrt{m}}{(m/4)^{1/3}} = 2^{5/3} \cdot m^{1/6},$$

$$G_4 = \max_{k \geqslant \frac{m(\sqrt{m}-1)}{m-1}} \frac{m/k}{m/(k+m/k)} = \max_{k \geqslant \frac{m(\sqrt{m}-1)}{m-1}} 1 + \frac{m}{k^2} = 1 + \frac{m \cdot (m-1)^2}{m^2 \cdot (\sqrt{m}-1)^2}$$

$$= 1 + \frac{(\sqrt{m}+1)^2}{m} \leqslant 5.$$

It is easy to check that $2^{8/3} \cdot m^{1/6} \leqslant 6.35 \cdot m^{1/6}$ is the highest among all four factors, for all values of m. Hence, the upper and the lower bounds are tight by a factor of at most $6.35 \cdot (m/4)^{1/6}$.

Regret, deterministic rules:

Upper bound: We show that the upper bound in this case is achieved by the rule f^* that selects the k alternatives with the highest plurality scores. Fix a profile $\vec{\sigma}$ and a set of alternatives $T \in \mathcal{A}_k \setminus \{f^*(\vec{\sigma})\}$. Let us calculate the worst-case regret due to T in the simplified regret formula from Lemma 1.

Recall that there are $\mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})$ votes i in which $\sigma_i(f^*(\vec{\sigma})) = 1$, and thus we cannot have $T \succ_i f^*(\vec{\sigma})$. Further, there are exactly $\mathtt{plu}(T \setminus f^*(\vec{\sigma}), \vec{\sigma})$ votes i in which $T \succ_i f^*(\vec{\sigma})$ and $\sigma_i(T) = 1$. Hence, there are at most $n - \mathtt{plu}(T \setminus f^*(\vec{\sigma}), \vec{\sigma}) - \mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})$ votes i in which $T \succ_i f^*(\vec{\sigma})$ and $\sigma_i(T) \geqslant 2$. Substituting these into the formula from Lemma 1, we get that the worst-case regret due to T is at most

$$\frac{1}{n} \cdot \left(\frac{\mathtt{plu}(T \setminus f^*(\vec{\sigma}), \vec{\sigma})}{1} + \frac{n - \mathtt{plu}(T \setminus f^*(\vec{\sigma}), \vec{\sigma}) - \mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})}{2} \right) \tag{8}$$

$$= \frac{n + \mathtt{plu}(T \setminus f^*(\vec{\sigma}), \vec{\sigma}) - \mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})}{2n}. \tag{9}$$

Next, note that $\mathtt{plu}(T \setminus f^*(\vec{\sigma}), \vec{\sigma}) \leqslant \mathtt{plu}(T, \vec{\sigma}) \leqslant \mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma})$, where the last equation follows due to the definition of f^*. Substituting this into Equation (9), we get that the regret is at most $1/2$, as desired.

For $k > m/2$, we can derive a better bound because we know $T \cap f^*(\vec{\sigma}) \neq \emptyset$. Note that $T \setminus f^*(\vec{\sigma}) \subseteq A \setminus f^*(\vec{\sigma})$. Because $f^*(\vec{\sigma})$ consists of the k alternatives with the highest plurality scores, and the plurality scores sum to n, we have

$$\mathtt{plu}(f^*(\vec{\sigma}), \vec{\sigma}) \geqslant \frac{k}{m} \cdot n. \tag{10}$$

Similarly, $A \setminus f^*(\vec{\sigma})$ consists of the $m - k$ alternatives with the lowest plurality scores. Hence, we have $\mathtt{plu}(A \setminus f^*(\vec{\sigma}), \vec{\sigma}) \leqslant (m - k) \cdot n/m$. Hence, we have

$$\mathtt{plu}(T \setminus f^*(\vec{\sigma}), \vec{\sigma}) \leqslant \mathtt{plu}(A \setminus f^*(\vec{\sigma}), \vec{\sigma}) \leqslant \frac{m - k}{m} \cdot n \tag{11}$$

Substituting Equations (10) and (11) into Equation (9), we get that the worst-case regret caused by T is at most

$$\frac{n + \frac{m-k}{m} \cdot n - \frac{k}{m} \cdot n}{2n} = 1 - \frac{k}{m}.$$

Since the choices of T and $\vec{\sigma}$ were arbitrary, we have that $\mathtt{reg}(f^*) \leqslant 1 - k/m$, as required.

Lower bound: Next, we prove the matching lower bound. For $k \leqslant m/2$, fix a set $X \subseteq A$ of $2k$ alternatives. Construct a profile $\vec{\sigma}$ in which every alternative in X appears first in $n/(2k)$ votes. In this case, for any deterministic rule f with $|f(\vec{\sigma})| = k$, one can find a set $T \subseteq X \setminus f(\vec{\sigma})$ with $|T| = k$. Note that $T \cap f(\vec{\sigma}) = \emptyset$, and there are exactly $k \cdot n/(2k) = n/2$ votes in which $T \succ_i f(\vec{\sigma})$ and $\sigma_i(T) = 1$. Hence, due to Lemma 1, we have $\mathtt{reg}(f) \geqslant \mathtt{reg}(f, \vec{\sigma}) \geqslant (1/n) \cdot (n/2) = 1/2$, as required.

For $k > m/2$, we construct a profile $\vec{\sigma}$ in which every alternative appears first in exactly n/m votes. Once again, for any deterministic rule f with $|f(\vec{\sigma})| = k$, we can choose a set $T \supseteq A \setminus f(\vec{\sigma})$ with $|T| = k$. Note that $|T \setminus f(\vec{\sigma})| = m - k$. Hence, there are $(m - k) \cdot n/m$ votes in which $T \succ_i f(\vec{\sigma})$, and $\sigma_i(T) = 1$. Thus, due to Lemma 1, we have that $\mathtt{reg}(f) \geqslant \mathtt{reg}(f, \vec{\sigma}) \geqslant (1/n) \cdot (m - k) \cdot n/m = 1 - k/m$, as required.

Regret, randomized rules:

Upper bound: We explicitly construct the randomized rule f^* that provides the required upper bound. Fix a preference profile $\vec{\sigma}$. Without loss of generality, let us relabel the set of alternatives as $A = \{a_1, a_2, \dots, a_m\}$ such that $\mathtt{plu}(a_i, \vec{\sigma}) \geqslant \mathtt{plu}(a_{i+1}, \vec{\sigma})$ for $i \in [m - 1]$. Further, define $A_i \triangleq \{a_1, a_2, \dots, a_i\}$ for $i \in [m]$. We now prove an important technical result.

Lemma 2. *There exists an integer value of $t \in [k, m]$ such that*

$$t \leqslant k + \mathit{plu}(a, \vec{\sigma}) \cdot \sum_{b \in A_t} \frac{1}{\mathit{plu}(b, \vec{\sigma})}, \quad \forall a \in A_t, \tag{12}$$

$$t \geqslant k + \mathit{plu}(a, \vec{\sigma}) \cdot \sum_{b \in A_t} \frac{1}{\mathit{plu}(b, \vec{\sigma})}, \quad \forall a \notin A_t. \tag{13}$$

Proof. Due to our ordering of the alternatives, the two conditions on t can be reduced to

$$k + \text{plu}(a_{t+1}, \vec{\sigma}) \cdot \sum_{b \in A_t} \frac{1}{\text{plu}(b, \vec{\sigma})} \leqslant t \leqslant k + \text{plu}(a_t, \vec{\sigma}) \cdot \sum_{b \in A_t} \frac{1}{\text{plu}(b, \vec{\sigma})}, \tag{14}$$

where we let $\text{plu}(a_{t+1}, \vec{\sigma}) = 0$ if $t = m$. To prove that such a value of t exists, let us consider

$$f(t) = k + \text{plu}(a_t, \vec{\sigma}) \cdot \sum_{b \in A_t} \frac{1}{\text{plu}(b, \vec{\sigma})} - t.$$

If $f(m) \geqslant 0$, one can check that $t = m$ satisfies Equation (14). Suppose $f(m) < 0$. It is easy to check that $f(k) \geqslant 0$. Further, $f(t)$ is monotonically non-increasing in t. To see this, note that

$$f(t+1) = k + \text{plu}(a_{t+1}, \vec{\sigma}) \cdot \sum_{b \in A_{t+1}} \frac{1}{\text{plu}(b, \vec{\sigma})} - (t+1)$$

$$= k + \text{plu}(a_{t+1}, \vec{\sigma}) \cdot \sum_{b \in A_t} \frac{1}{\text{plu}(b, \vec{\sigma})} + 1 - (t+1)$$

$$= k + \text{plu}(a_{t+1}, \vec{\sigma}) \cdot \sum_{b \in A_t} \frac{1}{\text{plu}(b, \vec{\sigma})} - t. \tag{15}$$

Substituting $\text{plu}(a_{t+1}, \vec{\sigma}) \leqslant \text{plu}(a_t, \vec{\sigma})$ in Equation (15), we get $f(t+1) \leqslant f(t)$.

Finally, consider the largest value of t satisfying $f(t) \geqslant 0$. We have $f(t) \geqslant 0$ and $f(t+1) < 0$. We can show that this value of t satisfies the two inequalities in Equation (14). The first inequality follows by substituting the value of $f(t+1)$ from Equation (15) into $f(t+1) < 0$, and the second inequality follows directly from $f(t) \geqslant 0$. ∎

Using Lemma 2, fix a value of $t \in [k, m]$ that satisfies Equation (13) and (12). Define

$$p_a = \begin{cases} 1 - \dfrac{t-k}{\text{plu}(a, \vec{\sigma}) \cdot \sum_{b \in A_t} 1/\text{plu}(b, \vec{\sigma})} & \text{if } a \in A_t, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $0 \leqslant p_a \leqslant 1$ for each $a \in A$ due to Equation (12), and $\sum_{a \in A} p_a = k$. The bihierarchy extension (Budish et al., 2013) of the Birkhoff-von Neumann theorem (Birkhoff, 1946; von Neumann, 1953) implies that there exists a distribution over $\mathcal{A}_k$ (which can be computed in polynomial time) under which the probability of each alternative $a \in A$ being selected is p_a. We let our rule f^* return this distribution.

Next, we bound $\text{reg}(f^*)$. Fix $T \in \mathcal{A}_k$. Using Lemma 1, the worst-case regret due to T is 1 from each voter that ranks an alternative from $T \setminus f^*(\vec{\sigma})$ first, and at most $1/2$ from each voter that ranks an alternative from $A \setminus (f^*(\vec{\sigma}) \cup T)$ first. Hence, the expected regret

is at most

$$
\frac{1}{n} \cdot \mathbb{E}\left[\sum_{a \in T \setminus f^*(\vec{\sigma})} \mathtt{plu}(a, \vec{\sigma}) + \frac{1}{2} \sum_{a \in A \setminus (f^*(\vec{\sigma}) \cup T)} \mathtt{plu}(a, \vec{\sigma})\right]
$$

$$
= \frac{1}{2n} \cdot \mathbb{E}\left[\sum_{a \in T \setminus f^*(\vec{\sigma})} \mathtt{plu}(a, \vec{\sigma})\right] + \frac{1}{2n} \cdot \mathbb{E}\left[\sum_{a \in A \setminus f^*(\vec{\sigma})} \mathtt{plu}(a, \vec{\sigma})\right]
$$

$$
= \frac{1}{2n} \cdot \sum_{a \in T}(1 - p_a) \cdot \mathtt{plu}(a, \vec{\sigma}) + \frac{1}{2} - \frac{1}{2n} \cdot \sum_{a \in A_t} p_a \cdot \mathtt{plu}(a, \vec{\sigma})
$$

$$
\leqslant \frac{k \cdot (t - k)}{2n \cdot \sum_{a \in A_t} 1/\mathtt{plu}(a, \vec{\sigma})} + \frac{1}{2} - \frac{1}{2n} \cdot \sum_{a \in A_t} \mathtt{plu}(a, \vec{\sigma}) + \frac{t \cdot (t - k)}{2n \cdot \sum_{a \in A_t} 1/\mathtt{plu}(a, \vec{\sigma})}
$$

$$
= \frac{1}{2} + \frac{t^2 - k^2}{2n \cdot \sum_{a \in A_t} 1/\mathtt{plu}(a, \vec{\sigma})} - \frac{1}{2n} \cdot \sum_{a \in A_t} \mathtt{plu}(a, \vec{\sigma})
$$

$$
\leqslant \frac{1}{2} - \frac{k^2}{2n} \cdot \sum_{a \in A_t} \mathtt{plu}(a, \vec{\sigma}) \leqslant \frac{1}{2} - \frac{k^2}{2n \cdot t \cdot (m/n)} \leqslant \frac{1}{2} \cdot \left(1 - \frac{k^2}{m^2}\right),
$$

as desired. The first two equalities follow from the definitions of f^*, T, A_t, and p_a. To see why the first inequality holds, note that

$$
(1 - p_a) \cdot \mathtt{plu}(a, \vec{\sigma}) \leqslant \frac{t - k}{\sum_{b \in A_t} 1/\mathtt{plu}(b, \vec{\sigma})}, \ \forall a \in A.
$$

For $a \notin A_t$, this reduces to Equation (13). For $a \in A_t$, this follows from the definition of p_a. The second inequality follows by the power mean inequality, the third inequality follows because the t alternatives in A_t have the highest plurality scores (hence, $\sum_{a \in A_t} \mathtt{plu}(a, \vec{\sigma}) \geqslant t \cdot n/m$), and the final inequality follows because $t \leqslant m$.

Lower bound: This proof is very similar to the proof of the deterministic case.

For $k \leqslant m/2$, fix a set of alternatives $X \subseteq A$ such that $|X| = 2k$. Construct a profile $\vec{\sigma}$ in which every alternative in X appears first in $n/(2k)$ votes. Now, let us consider a randomized rule f that returns a distribution over $\mathcal{A}_k$. Let T denote the set of k alternatives from X with the least probability of being picked in $\mathcal{S}$. Because $\sum_{a \in X} \Pr[a \in f(\vec{\sigma})] \leqslant k$, we have $\sum_{a \in T} \Pr[a \in f(\vec{\sigma})] \leqslant k/2$. Hence, $\sum_{a \in T} \Pr[a \notin f(\vec{\sigma})] \geqslant k/2$. Now, from Lemma 1, the expected regret of f due to T is at least

$$
\frac{1}{n} \cdot \mathbb{E}\left[\sum_{a \in T} \mathbb{I}[a \notin f(\vec{\sigma})] \cdot \frac{n}{2k}\right] = \frac{1}{2k} \cdot \sum_{a \in T} \Pr[a \notin f(\vec{\sigma})] \geqslant \frac{1}{4},
$$

as required.

Similarly, for $k > m/2$, once again construct a profile $\vec{\sigma}$ in which every alternative appears first in n/m votes. Consider a randomized rule f that returns a distribution over $\mathcal{A}_k$. Let T be the set of k alternatives with the least probability of being picked in $f(\vec{\sigma})$. Because $\sum_{a \in A} \Pr[a \in f(\vec{\sigma})] \leqslant k$, we have $\sum_{a \in T} \Pr[a \in f(\vec{\sigma})] \leqslant k \cdot k/m$. Hence,

$\sum_{a \in T} \Pr[a \notin f(\vec{\sigma})] \geqslant k \cdot (1 - k/m)$. Once again, from Lemma 1, the expected regret of f due to T is at least

$$\frac{1}{n} \cdot \mathbb{E}\left[\sum_{a \in T} \mathbb{I}[a \notin f(\vec{\sigma})] \cdot \frac{n}{m}\right] = \frac{1}{m} \cdot \sum_{a \in T} \Pr[a \notin f(\vec{\sigma})] \geqslant \frac{k}{m} \cdot \left(1 - \frac{k}{m}\right),$$

as required.

Gap between upper and lower bounds: Note that for $k \leqslant m/2$, the ratio between the upper and the lower bounds is

$$\frac{(1/2) \cdot (1 - k^2/m^2)}{1/4} = 2 \cdot \left(1 - \frac{k^2}{m^2}\right) \leqslant 2.$$

For $k > m/2$, the ratio between the upper and the lower bounds is

$$\frac{(1/2) \cdot (1 - k^2/m^2)}{(1/2) \cdot (k/m) \cdot (1 - k/m)}.$$

It can be checked easily that this is a decreasing function of k. Hence, the maximum ratio is achieved at $k = m/2$, and is equal to 2.

Thus, in both cases, the upper and the lower bounds in this case are tight up to a constant factor of 2.

Running time: Note that rules from our upper bounds only require calculating the plurality scores and finding a decomposition according to (the bihierarchy extension of) the Birkhoff-von Neumann theorem, both of which can be accomplished in polynomial time. ∎

4. Empirical Comparisons

In Section 3 we provided analytical results for both deterministic and randomized rules. In our view, randomized rules are especially practicable when the output distribution is sampled multiple times, or when the voters are well-informed, or when the voters are indifferent about the outcome (e.g., they are software agents). Moreover, we believe that the results for randomized rules are of substantial theoretical interest. But our work is partly driven by its direct applications in RoboVote (see Section 1.1), which does not satisfy the above conditions. This leads us to use deterministic voting rules, which is what we focus on hereinafter.

Let $f^*_{\mathtt{dist}}$ and $f^*_{\mathtt{reg}}$ be the *deterministic* rules that minimize the worst-case distortion and regret, respectively, on every given preference profile. The deterministic results of Section 3 establish upper and lower bounds on their *worst-case* distortion/regret. In this section, we evaluate their *average-case* performance on simulated as well as real data, and compare them against nine well-known voting rules: plurality, approval voting, Borda count, STV, Kemeny's rule, the maximin rule, Copeland's rule, Bucklin's rule, and Tideman's rule.[4]

We perform three experiments: (i) choosing a utility profile uniformly at random from the simplex of all utility profiles, (ii) drawing a real-world utility profile from the Jester

---

4. For the score-based rules, the k-subset is selected by picking the top k alternatives based on their scores.
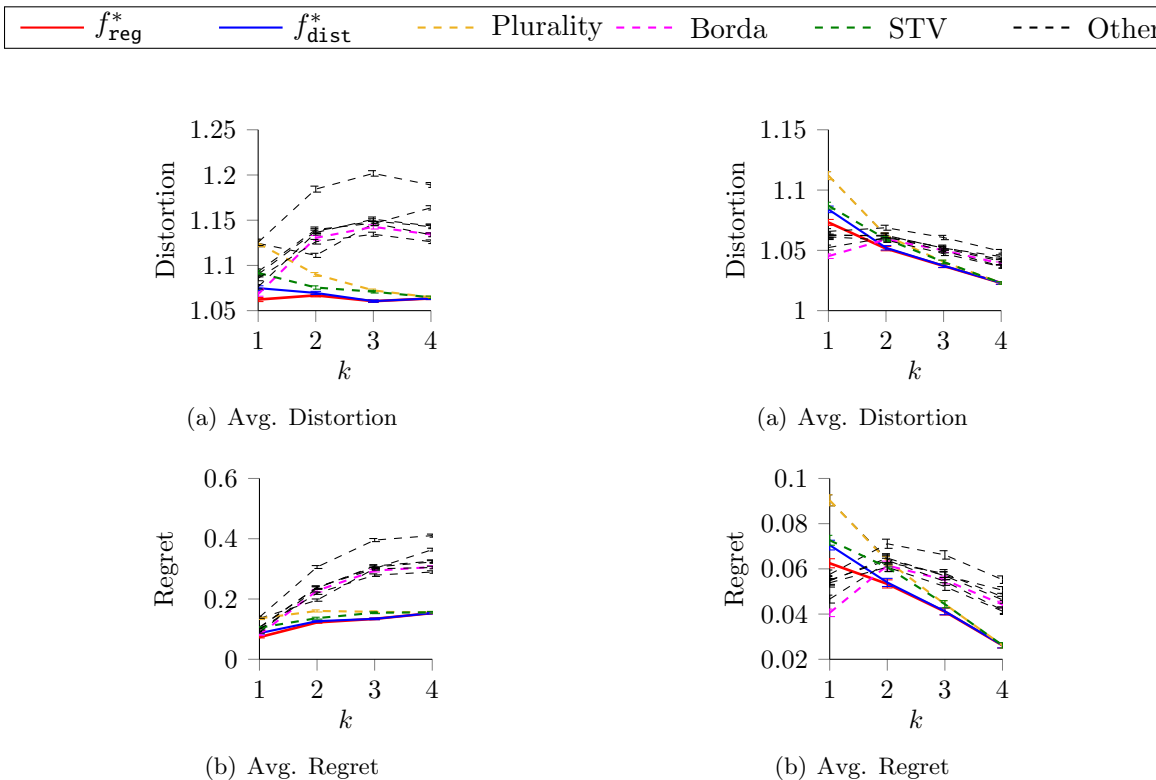
Fig. 4: Uniformly random utility profiles.

Fig. 5: Utility profiles from the Jester dataset.

datasets (Goldberg, Roeder, Gupta, & Perkins, 2001), and (iii) drawing a real-world preference profile from the PrefLib datasets (Mattei & Walsh, 2013), and choosing a consistent utility profile uniformly at random. For each experiment, we have 8 voters and 10 alternatives, and test for $k \in [4]$.[5] For each setting, we perform 10 000 random simulations, and measure both distortion and regret for the *actual utility profile*, as opposed to the worst-case utility profile. The figures show the average performance with 95% confidence intervals.

In all of our simulations, we observed that three of the classical voting rules stand out: Borda count performs well for choosing a single alternative (but not for choosing larger subsets) whereas plurality and STV perform well for choosing larger subsets (but not for choosing a single alternative). Hence, all of our graphs specifically distinguish these three rules in addition to $f^*_{\texttt{dist}}$ and $f^*_{\texttt{reg}}$.

Figure 4 shows the results for the first experiment where we choose the utility profile uniformly at random. Figure 5 shows the results for the second experiment where real-world utility profiles are drawn from one of the Jester datasets, in which more than 50 000 voters rated 150 jokes on a real-valued scale; the results from the other Jester dataset are almost identical. Finally, Figure 6 shows the results for the third experiment where real-world preference profiles are drawn from the Sushi dataset (5 000 voters ranking 100 different kinds of sushi) and the T-Shirt dataset (30 voters ranking 11 T-shirt designs) from PrefLib.

---

5. In RoboVote, we expect typical instances to have few voters and alternatives.

(a) Sushi: Avg. Distortion

(b) Sushi: Avg. Regret
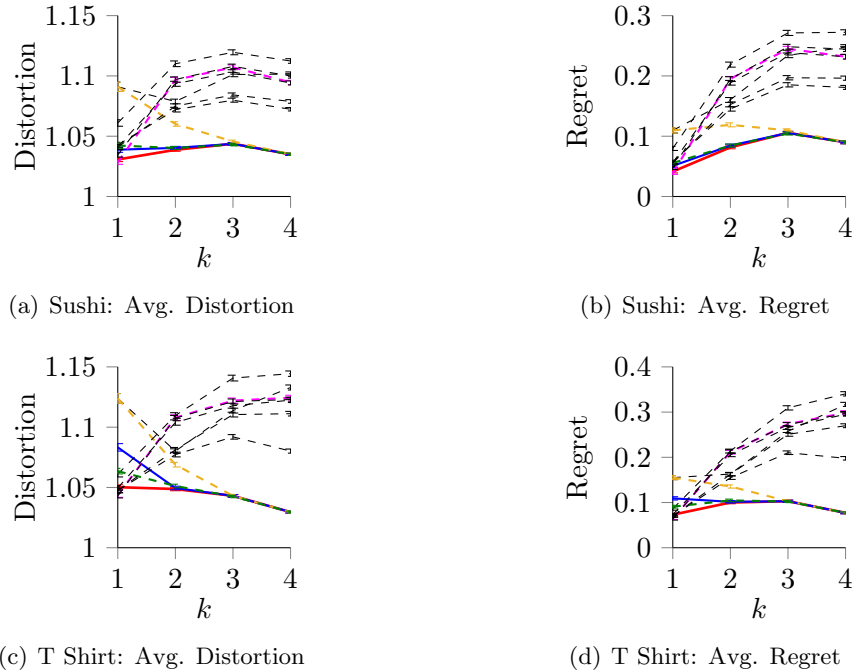
(c) T Shirt: Avg. Distortion

(d) T Shirt: Avg. Regret

Fig. 6: Preference profiles from Sushi and T-Shirt datasets, uniformly random consistent utility profiles.

Experiments on other datasets from PrefLib (AGH Course Selection, Netflix, Skate, and Web Search) yielded similar results.

Right off the bat, one can observe that the average-case distortion and regret values are much lower than their worst-case counterparts. For example, average regret is generally lower than 0.1 — compare with the tight worst-case deterministic bound of $1/2$ for $k \leqslant m/2$.

Much to our surprise, in all of our experiments, $f^*_{\mathtt{reg}}$ outperforms $f^*_{\mathtt{dist}}$ in terms of both average-case distortion (multiplicative loss) *and* regret (additive loss). While both measures of loss have been studied extensively in the literature, we are not aware of any previous work that compares the two approaches. At least in our social choice domain, the regret-based approach is clearly better on average.

Moreover, in all cases but one ($k = 1$ in the Jester experiment), $f^*_{\mathtt{reg}}$ also outperforms all the classical voting rules under consideration. We therefore conclude that, on random as well as on real-world instances, $f^*_{\mathtt{reg}}$ provides superior performance in terms of social welfare maximization.

5. Computation and Implementation

In this section, we analyze and compare the two deterministic optimal rules — $f^*_{\mathtt{dist}}$ and $f^*_{\mathtt{reg}}$ — from a computational viewpoint. Selecting optimal subsets turns out to be challenging, as both rules are $\mathcal{NP}$-hard to compute.

143

Theorem 2. *Given a preference profile $\vec{\sigma}$ and an integer k, computing a k-subset of alternatives that has the minimum distortion or the minimum regret on $\vec{\sigma}$ is $\mathcal{NP}$-hard.*

Proof. We present a polynomial-time reduction from the minimum dominating set problem, which is known to be $\mathcal{NP}$-hard (in fact, APX-hard) even in 3-regular graphs (see, e.g., Papadimitriou & Yannakakis, 1991). A set of nodes S is called a dominating set in a graph $G = (V, E)$ if any node in $V \setminus S$ is adjacent to a node in S, and is called a minimum dominating set if it is a dominating set of the minimum possible size.

Let $V = \{v_1, \ldots, v_n\}$ (thus, $n = |V|$). Observe that 3-regularity of G implies $|E| = 3n/2$. Let c and d be positive integers such that c is a multiple of $12n$, $c \geqslant 96n^2 H_d$ (where H_d is the dth harmonic number), and $d = 3nc + 3c + \sum_{i=2}^{3n/2} \left(5c - \frac{c \cdot i}{3n}\right)$. Clearly, there exist such values of c and d satisfying $c = \mathcal{O}(n^2 \ln n)$ and $d = \mathcal{O}(n^3 \ln n)$, and they can be computed in polynomial time.

Now, set $k = 1 + \lfloor 3n/4 \rfloor$, and construct a profile $\vec{\sigma}$ as follows. The set of alternatives is the union of three sets:

- $A = \{a_1, \ldots, a_n\}$ (each a_i corresponds to the node $v_i \in V$);

- $B = \{b_1, \ldots, b_{3n/2}\}$ (each b_i corresponds to an edge in E);

- $D = \{f_1, \ldots, f_d\}$.

The set of voters N is the union of the following sets:

- N_e, which consists of $2c$ "edge voters" $\left\{n_e^j\right\}_{j \in \{0,1,\ldots,2c-1\}}$ for each edge $e \in E$;

- N_1, which consists of $3c$ voters;

- N_i consisting of $\frac{14c}{3} - \frac{c \cdot i}{12n}$ voters, for each $i = 2, \ldots, 3n/2$.

The reader may check that the total number of voters is exactly d. Next, the votes in $\vec{\sigma}$ are as follows:

- For each edge $e = (v_{i_1}, v_{i_2}) \in E$ with $i_1 < i_2$, voter n_e^j ranks alternatives a_{i_1} and a_{i_2} at positions 2 and 3, respectively, when j is even, and at positions 3 and 2, respectively, when j is odd. Alternatives in $A \setminus \{a_{i_1}, a_{i_2}\} \cup B$ are ranked at the bottom, in arbitrary order. Positions 1 and 4 through $d + 2$ are reserved for alternatives in D (see below for the way the alternatives in D are placed in these positions).

- Voters in N_1 rank alternative b_1 in the first position, and alternatives in $A \cup B \setminus \{b_1\}$ in the last positions, in arbitrary order. Positions 2 through $d + 1$ are reserved for alternatives in D.

- For $i = 2, \ldots, 3n/2$, voters in N_i rank alternative b_i in the second position, and alternatives in $A \cup B \setminus \{b_i\}$ in the last positions, in arbitrary order. Positions 1 and 3 through $d + 1$ are reserved for alternatives in D.

- Alternatives in D are shuffled in a cyclic fashion in the votes within the positions reserved for them. More specifically, fix an order among the votes. Then, for $i, j \in [d]$, let $t = 1 + (i + j - 2 \bmod d)$. Alternative f_i appears in the tth reserved position of the jth vote.

Given this construction, the next lemma establishes a strong relation between the k-sized set of alternatives with the minimum regret or distortion in $\vec{\sigma}$, and the minimum dominating set in graph G. Theorem 2 then follows as our reduction is polynomial time.

Lemma 3. *A k-sized set of alternatives has the minimum regret or the minimum distortion on $\vec{\sigma}$ if and only if it consists of the alternatives in A corresponding to the nodes of a minimum dominating set S^* of G and the alternatives $b_1, b_2, \ldots, b_{k-|S^*|}$.*

Proof. Let K^* denote the k-sized set of alternatives that consists of the alternatives in A corresponding to the nodes of a minimum dominating set S^* of G, and the alternatives $b_1, b_2, \ldots, b_{k-|S^*|}$. We prove an upper and a lower bound on the regret of K^* on $\vec{\sigma}$, which establishes that it is the unique k-sized set of alternatives with the minimum regret. We later provide the argument that shows it is also the unique k-sized set of alternatives with the minimum distortion.

Upper bound: Let K' be a k-sized set of alternatives that is disjoint from K^*.[6] We now show that the regret of K^* on $\vec{\sigma}$ due to K' is at most

$$\texttt{reg}(K^*, \vec{\sigma}) \leqslant T \triangleq \sum_{i=k-|S^*|+1}^{2k-|S^*|} \left(\frac{7c}{3} - \frac{c \cdot i}{24n} + 1 \right). \tag{16}$$

Since T is independent of K' itself, it would follow that the worst-case regret of K^* is also upper bounded by T.

Consider an alternative $a_i \in A \setminus K^*$. Recall that this corresponds to the node $v_i \in V$. Let v_ℓ be a node in S^* that is adjacent to v_i. From Equation 1, the contribution of alternative a_i to the regret is

- 0 in the c edge voters n_e^j corresponding to edge $e = (v_i, v_\ell)$ that have alternative a_i ranked third;

- at most $1/3$ in each of the remaining $2c$ edge voters that have alternative a_i ranked third;

- at most $1/2$ in each of the $3c$ edge voters that have alternative a_i ranked second;

- smaller than $1/d$ in any other voter (because it is ranked below all alternatives in D).

Thus, the total contribution of a_i to the regret is at most $2c/3 + 3c/2 + 1 = 13c/6 + 1$.

Due to cyclic shuffling, the contribution of an alternative in D to the regret is at most H_d. For $i \geqslant k + 1 - |S^*|$, the contribution of alternative b_i to the regret is at most $1/2$ from the voters in N_i, and at most $1/d$ in any other voter. In total, this contribution is in $[7c/3 - \frac{c \cdot i}{24n}, 7c/3 - \frac{c \cdot i}{24n} + 1]$, and therefore is higher than the contribution of any alternative in $(A \cup D) \setminus K^*$.

Hence, the k alternatives that contribute the highest regret are $b_{k+1-|S^*|}, \ldots, b_{2k-|S^*|}$, and their total contribution to the regret is at most $\sum_{i=k+1-|S^*|}^{2k-|S^*|} \left(\frac{7c}{3} - \frac{c \cdot i}{24n} + 1 \right)$ as desired.

---

6. If K' shares alternatives with K^*, such alternatives do not cause any regret in the simplified regret formula from Equation (1). Hence, to compute the worst-case regret, it is sufficient to focus on sets K' that are disjoint from K^*.

Lower bound: Next, let K be a k-sized set of alternatives that does not follow the characterization of the lemma (i.e., does not consist of exactly the alternatives in A corresponding to the nodes of a minimum dominating set S^* of G and the alternatives $b_1, \ldots, b_{k-|S^*|}$). We show that $\operatorname{reg}(K, \vec{\sigma}) > T$.

Let S be the set of nodes of G corresponding to the alternatives in $K \cap A$. Let $\mathrm{I}(S)$ denote an independent set of nodes of G (i.e., no two nodes in $\mathrm{I}(S)$ are connected to each other) such that no node in $\mathrm{I}(S)$ is adjacent to any node in S. Construct a k-sized set K' that consists of the alternatives in A corresponding to the nodes of $\mathrm{I}(S)$ (if any), and the $k - |\mathrm{I}(S)|$ alternatives from $B \setminus K$ with the smallest indices. We show that the regret of K on $\vec{\sigma}$ due to K' is more than T. We do so by considering the contribution of the alternatives of K' to the regret only from the votes in which they appear among the first three positions. For now, let us ignore the alternatives in D that appear in K; we will later the regret calculated to account for such alternatives.

Let $X = \mathbb{I}[(B \setminus K) \cap \{b_1, \ldots, b_{k-|S|-|K \cap D|}\} \neq \emptyset]$. Consider an alternative a_i corresponding to the node $v_i \in \mathrm{I}(S)$. Because the nodes in $\mathrm{I}(S)$ are mutually non-adjacent and non-adjacent to any node in S, $a_i \in K'$ is ranked above all alternatives in $K \setminus D$ in the edge voters n_e^j for $j = 0, 1, \ldots, 2c - 1$ and each edge e adjacent to node v_i. Ignoring the alternatives in $K \cap D$ (that may be ranked in the first position of these voters), the contribution of a_i to the regret is $1/2$ from each of the $3c$ edge voters that rank a_i in the second position, and $1/3$ from each of the $3c$ edge voters that rank a_i in the third position. Thus, the total contribution of the alternatives that correspond to the nodes in $\mathrm{I}(S)$ is $\frac{5c}{2}$.

Next, consider the $k - |\mathrm{I}(S)|$ alternatives from $B \setminus K$ that have the smallest indices. If K contains all of $b_1, \ldots, b_{k-|S|-|K \cap D|}$, then K' contains the alternatives b_i for $i = k - |S| - |K \cap D| + 1, \ldots, 2k - |S| - |K \cap D| - |\mathrm{I}(S)|$. Again, ignoring the alternatives in $K \cap D$ (that may be ranked in the first position in the votes of the voters in N_i), the contribution of b_i to the regret of K is $1/2$ from each of the $\frac{14c}{3} - \frac{c \cdot i}{12n}$ voters in N_i that rank b_i in the second position. If $X = 1$, we know that K does not use the alternatives of B with the smallest indices, and the contribution of the alternatives of B to the regret increases by at least $\frac{c \cdot X}{24n}$.

Finally, we consider the fact that K may include alternatives from D. Each such alternative may be ranked first by a voter, for which we have incorrectly added a regret of at most $1/2$. Hence, the actual regret may be $(1/2) \cdot |K \cap D|$ lower than our calculated regret.

Combining the entire analysis, and observing that the regret due to K' is a lower bound on the worst-case regret of K on σ, we get that

$\operatorname{reg}(K, \vec{\sigma})$

$$\geqslant \frac{5c}{2} \cdot |\mathrm{I}(S)| + \sum_{i=k-|S|-|K \cap D|+1}^{2k-|S|-|K \cap D|-|\mathrm{I}(S)|} \left(\frac{7c}{3} - \frac{c \cdot i}{24n} \right) + \frac{c \cdot X}{24n} - \frac{1}{2} \cdot |K \cap D|$$

$$\geqslant \frac{c}{6} \cdot |\mathrm{I}(S)| + \sum_{i=k-|S|-|K \cap D|+1}^{2k-|S|-|K \cap D|} \left(\frac{7c}{3} - \frac{c \cdot i}{24n} \right) + \frac{c \cdot X}{24n} - \frac{1}{2} \cdot |K \cap D|$$

$$= \sum_{i=k-|S^*|+1}^{2k-|S^*|} \left(\frac{7c}{3} - \frac{c \cdot i}{24n} + 1 \right) + \frac{c}{6} \cdot |\mathrm{I}(S)| - \frac{ck}{24n} \cdot (|S^*| - |S|) + \left(\frac{ck}{24n} - \frac{1}{2} \right) \cdot |K \cap D| - k + \frac{cX}{24n}$$

$$\geqslant \operatorname{reg}(K^*, \vec{\sigma}) + \frac{c}{6} \cdot |\mathrm{I}(S)| - \frac{c \cdot k}{24n} \cdot (|S^*| - |S|) + \left(\frac{c \cdot k}{24n} - \frac{1}{2} \right) \cdot |K \cap D| - k + \frac{c \cdot X}{24n} \tag{17}$$

Finally, we analyze the quantities on the RHS of Equation (17) to derive that

$$\texttt{reg}(K, \vec{\sigma}) - \texttt{reg}(K^*, \vec{\sigma}) > \frac{7k}{9} \cdot |K \cap D|, \tag{18}$$

which implies that $\texttt{reg}(K, \vec{\sigma}) > \texttt{reg}(K^*, \vec{\sigma})$, as required.

We now perform a case-by-case analysis to establish Equation (18). First, suppose S is a dominating set of G.

- If S is a minimum dominating set of G (i.e., $|S| = |S^*|$), we must have $X = 1$ or $|K \cap D| > 0$.

  * If $X = 1$, Equation (17) yields

  $$\texttt{reg}(K, \vec{\sigma}) - \texttt{reg}(K^*, \vec{\sigma}) \geqslant \left(\frac{c \cdot k}{24n} - \frac{1}{2} \right) \cdot |K \cap D| - k + \frac{c}{24n},$$

  which in turn implies Equation (18) because the definition of c and the fact that $k \geqslant 1$ imply $\frac{c}{24n} > k$ and $(c \cdot k)/(24n) - (1/2) > (7k)/9$.

  * If $X = 0$ and $|K \cap D| > 0$, Equation (17) yields

  $$\begin{aligned} \texttt{reg}(K, \vec{\sigma}) - \texttt{reg}(K^*, \vec{\sigma}) &\geqslant \left(\frac{c \cdot k}{24n} - \frac{1}{2} \right) \cdot |K \cap D| - k \\ &\geqslant \left(\frac{c}{24n} - \frac{3}{2} \right) \cdot k \cdot |K \cap D| > \frac{7k}{9} \cdot |K \cap D|, \end{aligned}$$

  where the last inequality follows from the definition of c.

- If $|S| > |S^*|$, Equation (17) yields

  $$\texttt{reg}(K, \vec{\sigma}) - \texttt{reg}(K^*, \vec{\sigma}) \geqslant \left(\frac{c}{24n} - 1 \right) \cdot k + \left(\frac{c \cdot k}{24n} - \frac{1}{2} \right) \cdot |K \cap D| > \frac{7k}{9} \cdot |K \cap D|,$$

  where the last inequality follows from the definition of c.

Next, suppose S is not a dominating set of G. Here, we distinguish between two cases.

- If $|S| \geqslant |S^*|$, using $|\mathrm{I}(S)| \geqslant 1$, Equation (17) yields

  $$\texttt{reg}(K, \vec{\sigma}) - \texttt{reg}(K^*, \vec{\sigma}) \geqslant \frac{c}{6} + \left(\frac{c \cdot k}{24n} - 1 \right) \cdot |K \cap D| - k.$$

  Equation (18) now follows because the definition of c implies $(c/6) > k$ and $(c \cdot k)/(24n) - 1 > (7k)/9$.

- If $|S| < |S^*|$, then we claim that $|\mathrm{I}(S)| \geqslant (|S^*| - |S|)/4$. Indeed, because the minimum dominating set of G has size $|S^*|$, the set of nodes S must leave $|S^*| - |S|$ nodes of G

undominated. Because G is 3-regular, any independent set $I(S)$ among these nodes should have size at least $(|S^*| - |S|)/4$. Hence,

$$\mathtt{reg}(K, \vec{\sigma}) - \mathtt{reg}(K^*, \vec{\sigma}) \geqslant \left(\frac{c}{24} - \frac{c \cdot k}{24n} \right) \cdot (|S^*| - |S|) + \left(\frac{c \cdot k}{6n} - 1 \right) \cdot |K \cap D| - k$$

$$> \frac{c}{24} - \frac{c \cdot k}{24n} - k + \left(\frac{c}{24n} - \frac{3}{2} \right) \cdot k \cdot |K \cap D|,$$

which implies Equation (18) because the definition of c implies $(c/24) - (c \cdot k)/(24n) > k$ and $c/(24n) - (3/2) > (7/9)$.

This completes the proof of Equation (18) in all cases, and hence, concludes the proof of the regret part of the lemma.

Let us now prove the distortion part of the lemma. Again, consider a k-sized set K that does not follow the characterization of the lemma (i.e., does not consist of exactly alternatives in A corresponding to the nodes of a minimum dominating set S^* of G and the alternatives $b_1, \ldots, b_{k-|S^*|}$).

If $K \cap D = \emptyset$, then we have $\mathtt{plu}(K, \vec{\sigma}) \leqslant \mathtt{plu}(K^*, \vec{\sigma})$. Hence,

$$\mathtt{dist}(K, \vec{\sigma}) = 1 + m \cdot \frac{\mathtt{reg}(K, \vec{\sigma})}{\mathtt{plu}(K, \vec{\sigma})} > 1 + m \cdot \frac{\mathtt{reg}(K^*, \vec{\sigma})}{3c} = \mathtt{dist}(K^*),$$

where the second transition uses Equation (18).

If $K \cap D \neq \emptyset$, we have $\mathtt{plu}(K, \vec{\sigma}) \leqslant 3c + |K \cap D|$, which, together with Equation (18), implies

$$\mathtt{dist}(K, \vec{\sigma}) = 1 + m \cdot \frac{\mathtt{reg}(K, \vec{\sigma})}{\mathtt{plu}(K, \vec{\sigma})} > 1 + m \cdot \frac{\mathtt{reg}(K^*, \vec{\sigma}) + \frac{7k}{9} \cdot |K \cap D|}{3c + |K \cap D|} > 1 + m \cdot \frac{\mathtt{reg}(K^*)}{3c}$$

$$= \mathtt{dist}(K^*, \vec{\sigma}).$$

The final transition holds because Equation (16) implies $7k/9 > \mathtt{reg}(K^*, \vec{\sigma})/(3c)$.∎ (Proof of Lemma 3)

This concludes the entire proof. ∎ (Proof of Theorem 2)

Given that $f_{\mathtt{reg}}^*$ outperforms $f_{\mathtt{dist}}^*$ in the experiments of Section 4, and that both rules are computationally hard, $f_{\mathtt{reg}}^*$ stands out as the clear choice for implementation in our website RoboVote. We therefore devoted our efforts to developing a scalable implementation for $f_{\mathtt{reg}}^*$.

First, let us note the simplified formula for $f_{\mathtt{reg}}^*$ that follows from Lemma 1:

$$f_{\mathtt{reg},k}^*(\vec{\sigma}) = \arg\min_{T \in \mathcal{A}_k} \max_{S \in \mathcal{A}_k} \sum_{i=1}^n \frac{\mathbb{I}[S \succ_{\sigma_i} T]}{\sigma_i(S)}. \tag{19}$$

To better understand this formula, we consider the special case of $k = 1$. In this case,

$$f_{\mathtt{reg}}^*(\vec{\sigma}) \in \arg\min_{a \in A} \max_{b \in A} \sum_{i=1}^n \frac{\mathbb{I}[b \succ_{\sigma_i} a]}{\sigma_i(b)}.$$

148

Note that this voting rule is very similar to the classical maximin rule: replacing $\sigma_i(b)$ with 1 in the denominator would yield the maximin rule. Thus, in some sense, this is a smooth version of the maximin rule, where the "victory" of b over a in voter i's vote is weighted by the strength of b in this vote (measured by $1/\sigma_i(b)$). In our view, this intuitive structure makes $f_{\texttt{reg}}^*$ even more compelling.
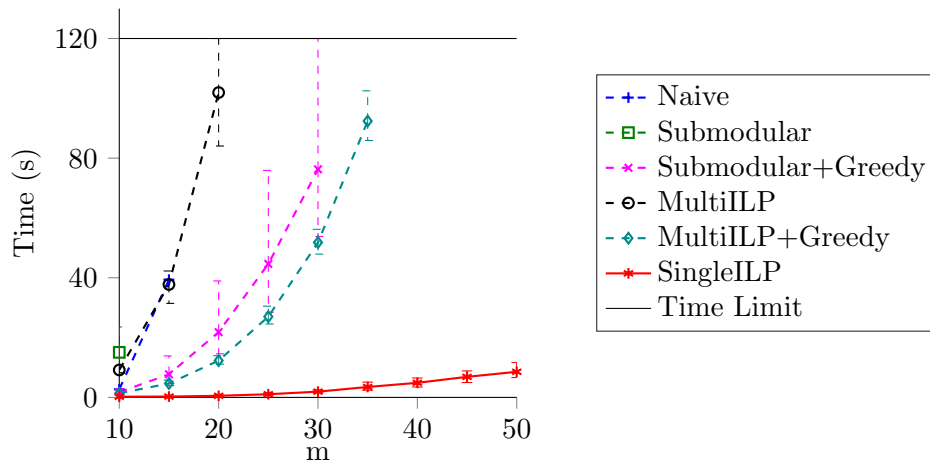

Fig. 7: Running times of six approaches to computing $f_{\texttt{reg}}^*$.

We now briefly describe six approaches we have developed for computing $f_{\texttt{reg}}^*$:

1. Naïve: This uses Equation (19), and requires $\Omega\left(n \cdot \binom{m}{k}^2\right)$ operations, which is prohibitive even for small m.

2. Submodular: The regret for set S in choosing set T, i.e., $\sum_{i \in [n]: S \succ_{\sigma_i} T} 1/\sigma_i(S)$, is submodular in S. Hence, for each $T \in \mathcal{A}_k$ we can optimize over $S \in \mathcal{A}_k$ using any algorithm for the submodular maximization subject to cardinality constraint (SMCC) problem. We use the SFO toolbox for Matlab (Krause, 2010).

3. Submodular+Greedy: This improves the previous approach by first computing a $1 - 1/e$ greedy approximation to the SMCC instance for set T, and pruning T if this is already greater than the best regret found so far.

4. MultiILP: Instead of using SMCC, for each $T \in \mathcal{A}_k$ we optimize over $S \in \mathcal{A}_k$ by solving an integer linear program (ILP) with roughly $n \cdot m$ variables and $n \cdot m^2$ constraints. Note that $\binom{m}{k}$ such ILPs need to be solved.

5. MultiILP+Greedy: This improves the MultiILP approach by using a greedy pruning procedure as before.

6. SingleILP: This approach solves a single but huge ILP with $\binom{m}{k}$ additional constraints.

Figure 7 shows the average running times of these approaches (and 95% confidence intervals) over 10 000 instances with $n = 15$, $k = 3$, and m varying from 10 to 50.[7] The

_____

7. The running time scales linearly in n, and increases with $\binom{m}{k}$.

experiments were performed on a single machine with quad-core 2.9 GHz CPU and 32 GB RAM. A time limit of 2 minutes was set because a running time greater than this would not be helpful for our website, where the results need to be delivered quickly to the users. While the greedy pruning procedure does help reduce the running time of both the Submodular and MultiILP approaches, SingleILP still computes $f_{\mathtt{reg}}^*$ much faster than any other approach, solving instances with 50 alternatives in less than 10 seconds. We have therefore implemented SingleILP on RoboVote.

6. Discussion

In this paper, we study a central problem in social choice theory that involves collectively selecting a set of k alternatives given ordinal preferences of n individuals over m alternatives. We take the viewpoint of implicit utilitarian voting, where our goal is to maximize the utilitarian social welfare according to underlying cardinal utilities given only ordinal information about these utilities available from the input preferences. Prior work by Boutilier et al. (2015) measures the loss due to incomplete information using the multiplicative notion of distortion, and analyzes this problem for the special case of $k = 1$ (i.e., selection of a single alternative). We introduce an additional loss measure, namely (additive) regret, provide a comprehensive theoretical analysis of the optimal distortion and regret for all $k \in [1, m-1]$, and empirically show that the optimal rules $f_{\mathtt{reg}}^*$ and $f_{\mathtt{dist}}^*$ derived from this framework are superior to classical social choice rules for optimizing social welfare.

We consider the empirical dominance of $f_{\mathtt{reg}}^*$ over $f_{\mathtt{dist}}^*$, in terms of both regret and (surprisingly) distortion, to be especially significant. It would be interesting to understand, on a theoretical level, why this happens. A promising starting point is to derive analytical bounds on the average-case distortion of $f_{\mathtt{reg}}^*$ and $f_{\mathtt{dist}}^*$ under uniformly random utility profiles.

We find it exciting that new theoretical questions in computational social choice are driven by concrete real-world applications. And while research in the field is often motivated by potential applications to multiagent systems, we focus on helping *people* — not software agents — make joint decisions.

Acknowledgments

References

Anshelevich, E., Bhardwaj, O., & Postl, J. (2015). Approximating optimal social choice under metric preferences. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 777–783.

Anshelevich, E., & Postl, J. (2016). Randomized social choice functions under metric preferences. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 46–59.

Anshelevich, E., & 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)*, pp. 390–396.

Arrow, K. (1951). *Social Choice and Individual Values*. Wiley.

Birkhoff, G. (1946). Three observations on linear algebra. *Universidad Nacional de Tucumán, Revista A*, *5*, 147–151.

Blum, A., & Mansour, Y. (2007). Learning, regret minimization, and equilibria. In Nisan, N., Roughgarden, T., Tardos, E., & Vazirani, V. (Eds.), *Algorithmic Game Theory*, chap. 4. Cambridge University Press.

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

Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. D. (Eds.). (2016). *Handbook of Computational Social Choice*. Cambridge University Press.

Bubeck, S., & Cesa-Bianchi, N. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, *5*(1), 1–122.

Budish, E., Che, Y.-K., Kojima, F., & Milgrom, P. (2013). Designing random allocation mechanisms: Theory and applications. *American Economic Review*, *103*(2), 585–623.

Campbell, D. E., & Kelly, J. S. (1996). Arrovian social choice correspondences. *International Economic Review*, *37*(4), 803–823.

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

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

Chamberlin, J. R., & Courant, P. N. (1983). Representative deliberations and representative decisions: Proportional representation and the Borda rule. *American Political Science Review*, *77*(3), 718–733.

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

Goldberg, K. Y., Roeder, T., Gupta, D., & Perkins, C. (2001). Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval*, *4*(2), 133–151.

Goldman, J., & Procaccia, A. D. (2014). Spliddit: Unleashing fair division algorithms. *SIGecom Exchanges*, *13*(2), 41–46.

Krause, A. (2010). SFO: A toolbox for submodular function optimization. *Journal of Machine Learning Research*, *11*, 1141–1144.

Krysta, P., Manlove, D., Rastegari, B., & Zhang, J. (2014). Size versus truthfulness in the house allocation problem. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, pp. 453–470.

Lu, T., & Boutilier, C. (2011a). Budgeted social choice: From consensus to personalized decision making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 280–286.

Lu, T., & Boutilier, C. (2011b). Robust approximation and incremental elicitation in voting protocols. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 287–293.

Mattei, N., & Walsh, T. (2013). Preflib: A library of preference data. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*, pp. 259–270.

Monroe, B. L. (1995). Fully proportional representation. *American Political Science Review*, *89*(4), 925–940.

Papadimitriou, C. H., & Yannakakis, M. (1991). Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, *43*(3), 425–440.

Procaccia, A. D., Reddi, S. J., & Shah, N. (2012). A maximum likelihood approach for selecting sets of alternatives. In *Proceedings of the 28th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 695–704.

Procaccia, A. D., & Rosenschein, J. S. (2006). The distortion of cardinal preferences in voting. In *Proceedings of the 10th International Workshop on Cooperative Information Agents (CIA)*, pp. 317–331.

Procaccia, A. D., Rosenschein, J. S., & Zohar, A. (2008). On the complexity of achieveing proportional representation. *Social Choice and Welfare*, *30*(3), 353–362.

Skowron, P., Faliszewski, P., & Lang, J. (2015). Finding a collective set of items: From proportional multirepresentation to group recommendation. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 2131–2137.

von Neumann, J. (1953). A certain zero-sum two-person game equivalent to the optimal assignment problem. In Kuhn, W., & Tucker, A. W. (Eds.), *Contributions to the Theory of Games*, Vol. 2, pp. 5–12.

Young, H. P. (1988). Condorcet's theory of voting. *The American Political Science Review*, *82*(4), 1231–1244.