Truthful and Near-Optimal Mechanisms for Welfare Maximization in Participatory Budgeting

UMANG BHASKAR, Tata Institute of Fundamental Research
VARSHA DANI, University of New Mexico

Participatory budgeting is an exciting new challenge in social choice theory, where members of a community directly vote and determine how a fixed budget should be distributed among candidates. Participatory budgeting mechanisms need to balance between many different considerations, including efficiency, complexity of information elicited from agents, and manipulability. We consider the utilitarian social welfare of mechanisms for participatory budgeting, measured by the distortion. We show that for a particular input format called threshold approval voting, if the thresholds for agents are chosen independently, there is a mechanism with nearly optimal distortion when the number of voters is large. Threshold mechanisms are potentially manipulable, but place low informational burden on voters.

We then consider truthful mechanisms. For the widely-studied class of ordinal mechanisms which elicit the rankings of candidates from each agent, we show that truthfulness essentially imposes no additional loss of welfare. We give truthful mechanisms with distortion $O(\sqrt{m \log m})$ when all alternatives have the same cost, and $O(\sqrt{m} \log m)$ distortion in the general case, where m is the number of candidates. These results nearly match known lower bounds on distortion for ordinal mechanisms that ignore strategic behaviour. Further, we show that for a natural class of truthful mechanisms, our first upper bound is tight. Lastly, for the case where agents decide between two candidates, we give tight bounds on the distortion for truthful mechanisms, both randomized as well as deterministic.

1 INTRODUCTION

How should a group of agents, presented with a set of candidates, collectively decide which candidates to select? This problem of deciding how to aggregate the preferences of multiple rational agents is the fundamental challenge in social choice theory, and has implications for diverse fields, including government formation, recommendation systems, scheduling, and resource allocation.

A recent and appealing application of preference aggregation is participatory budgeting. This is an example of direct democracy, and enables voters and those directly affected to decide how budgets available to their local government should be spent. In the US, funds worth more than \$250 million have been allocated via participatory budgeting in more than 440 community projects. Cities as diverse as Porto Alegre in Brazil to Chicago in the US use participatory budgeting to fund projects. Researchers in computational social choice have contributed significantly to this effort, both in the design and theoretical analysis of mechanisms, as well as building systems for participatory budgeting (see, e.g., [Caragiannis et al., 2017, Goel et al., 2016]).

The classical approach for designing mechanisms for preference aggregation is axiomatic: We identify properties that are intuitively fair and appealing, and then design mechanisms that satisfy these properties. Typical properties that are studied include Pareto-optimality, truthfulness, and monotonicity[1]. Many of these results require that the agents possess, in addition to ordinal preferences over the candidates, cardinal utilities for the candidates. This is particularly true for randomized mechanisms — a prominent example being the characterization of truthful randomized mechanisms [Gibbard, 1977].

[1]Informally, a mechanism is Pareto-optimal if no other outcome increases the welfare of all the agents, it is truthful if every agent maximizes its utility by reporting preferences truthfully, and monotone if increasing the preference by an agent for a candidate does not decrease the probability that the candidate is selected.

Manuscript submitted for review to ACM Economics & Computation 2017 (EC '17).

In many applications, the utilities of different agents have a common measure and may be compared to each other. This is a standard assumption, e.g., in mechanism design, where preferences of an agent can be expressed as money. In transportation systems, the time spent in transit is often used as a measure of utility, and the aggregate time spent as a measure of efficiency (e.g., [Mclean, 2016]). When interpersonal comparisons of utility are meaningful, the aggregate utility, or utilitarian welfare, is a commonly used measure to design and evaluate mechanisms. Good utilitarian welfare does not substitute for other properties, but a mechanism with bad utilitarian welfare arguably has little use in most practical applications.

Our work focuses on the utilitarian welfare of mechanisms. We use the concept of distortion, defined by Procaccia and Rosenschein [2006] approximately as the ratio of the maximum welfare obtained by a budget-feasible set, to the welfare obtained by the mechanism, in the worst case over all possible inputs[2]. The definition naturally extends to randomized mechanisms by considering the expected welfare obtained by the mechanism. Distortion is a particularly appealing measure since it is similar to the approximation ratio studied in theoretical computer science as a measure of the efficiency of algorithms. While the approximation ratio measures the loss in efficiency due to computational complexity, the distortion measures the loss due to other constraints such as truthfulness, incomplete information obtained from the agents, or computational complexity.

In fact, if our objective is solely to maximize the utilitarian welfare (equivalently, to minimize the distortion), this is easily achieved once we elicit the cardinal utility each agent obtains from the candidates. However, this approach is very problematic for many reasons. Firstly, agents in many settings are strategic, and may report their utilities falsely to the mechanism if doing so would increase their actual utility. Strategic voting in elections is a significant problem, when candidates supporting a less-popular third candidates may actually vote for one of the other candidates, to prevent their least desirable candidate getting elected. Secondly, even assuming that agents are truthful, the elicitation of cardinal utilities is a complex task: the human agent needs to be explained the scale being used, and must convert the implicit utilities for all the candidates to explicit values on this scale. The problem of utility elicitation is itself an active area of research (e.g., see [Chajewska et al., 2000, Wakker and Deneffe, 1996].

Thus, our objective in this paper is to design mechanisms for preference aggregation, and particularly for participatory budgeting, that maximize utilitarian welfare in the presence of these constraints — truthfulness and the complexity of information elicited from the agents.

1.1 Our Contribution

Our first result is a randomized mechanism that obtains distortion close to 1, when the number of agents is large. In this mechanism, we elicit from each agent the subset of candidates with utility above a given threshold. This particular format for preference elicitation (or input format) is studied by Benade et al. [2017]; for the mechanism they study, however, they give upper and lower bounds of $O(\log^2 m)$ and $\Omega(\log m/\log\log m)$ on the distortion where m is the number of candidates. We show that a subtle modification to the mechanism — when the threshold for the agents is i.i.d., rather than identical — allows us to beat the previous lower bound and give a near-optimal mechanism when the number of agents is large. The informational load on each agent is the same as the earlier mechanism.

In fact, our mechanism can be modified to further reduce the informational load on agents to a binary input. Instead of eliciting all candidates above a given threshold, we present each agent with a threshold and a single candidate, and ask if the agent's utility for the candidate is above

[2]Note that the distortion of any mechanism is at least 1, and the closer it is to 1 the better.

the threshold. We give a mechanism that, even with this severely limited information, obtains distortion close to 1, again when the number of agents is large.

Threshold mechanisms, despite our near-optimal results, suffer from two weaknesses: they depend on an explicit expression of cardinal utilities, and they are not truthful. While much of the literature in social choice theory relies on the existence of implicit cardinal utilities, especially when considering randomized mechanisms, mechanisms that require expression of these cardinal utilities are few. We therefore next consider the extensively-studied class of ordinal mechanisms, where agents order the candidates according to their preference[3]. For these mechanisms, we consider the effect of imposing truthfulness on the efficiency, measured as in previous work by the distortion.

For ordinal mechanisms, we show that insisting on truthfulness imposes essentially no loss on the distortion. Prior work by Boutilier et al. [2015] shows that any ordinal mechanism has distortion $\Omega(\sqrt{m})$, even when a single candidate is to be selected and ignoring strategic behaviour. We show that for the k-selection problem, when a subset of candidates of size k is to be selected, there is a truthful ordinal mechanism with distortion $O(\sqrt{m \log m})$. Given the strong characterizations of truthful mechanisms, we find this to be quite surprising [Gibbard, 1977]. We further extend this mechanism to the participatory budgeting problem, when candidates have costs and a budget-feasible subset is to be selected. For this, we give a truthful ordinal $O(\sqrt{m} \log m)$ mechanism. Thus, while measuring worst-case efficiency, truthfulness comes nearly for free. We note that Benade et al. [2017] give an ordinal mechanism with $O(\sqrt{m} \log m)$ distortion ignoring strategic behaviour. Our result shows this upper bound can be obtained even by truthful mechanisms.

Procaccia [2010] studies how well truthful ordinal mechanisms can approximate mechanisms based on positional scoring rules[4], such as Borda, measured by the approximation ratio. Our result for the 1-selection problem has an interesting corollary: our mechanism, based on the harmonic scoring function [Boutilier et al., 2015], is in fact a universal approximation for mechanisms based on scoring rules. That is, our mechanism does not require the scoring rule as input. However in expectation, the candidate selected by the mechanism has score that is within an $O(\sqrt{m \log m})$ factor of the candidate with maximum score, for every scoring rule. This follows by viewing the positional score of a candidate as the utility every agent has for the candidate. Procaccia shows that for truthful mechanisms and the Borda scoring rule, $\Omega(\sqrt{m})$ is a lower bound on the approximation ratio.

In fact, we show that for a class of truthful mechanisms that are unilaterally neutral between candidates, i.e. are not biased in favour of particular candidates in a strong sense, the bound of $O(\sqrt{m \log m})$ we obtain for the 1-selection problem above is tight. The class of truthful unilaterally-neutral mechanisms includes all truthful mechanisms with distortion $o(m)$ previously studied that we are aware of. To obtain this lower bound, we extend a characterization of truthful ordinal mechanisms to unilaterally neutral mechanisms. We then give a series of instances, each of which gives a lower bound on one parameter of the characterization, and shows that only mechanisms based on the harmonic scoring functions of Boutilier et al. [2015] can obtain $O(\sqrt{m \log m})$ distortion. Finally, we show that even such mechanisms have lower bound $\Omega(\sqrt{m \log m})$ on the distortion.

Lastly, we consider mechanisms when agents are presented with two candidates. A large amount of work has focused on this case when the agents' utilities are drawn from known distributions; we discuss these results in the next section. For the worst-case setting, with no prior information about agent utilities, it is easy to see that any truthful mechanism must be ordinal. We give a truthful randomized mechanism with distortion 1.5, and show this is a lower bound on any ordinal (and

[3] E.g., in order of decreasing implicit cardinal utility.

[4] A positional scoring rule is given by a vector $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ of nonnegative integer values. A candidate ranked r by an agent is awarded α_r points. The candidate with the maximum total points is selected by the mechanism.

hence truthful) mechanism as well. Interestingly, the randomized version of majority, which picks each candidate with probability proportional to the number agents that prefer it, can be shown to have distortion strictly greater than 1.5. Instead, our mechanism applies a transformation to the number of agents that prefer a candidate, to obtain the probability with which the candidate is picked. For truthful deterministic mechanisms, we show that the lower and upper bound on distortion is 3.

1.2 Related Work

Distortion as a measure of loss of welfare due to the input format is introduced by Procaccia and Rosenschein [2006], who show that no mechanism has unit distortion, even in simple instances, and for many popular ordinal mechanisms such as Borda and Veto, the distortion is unbounded. They define a related measure called misrepresentation that restricts the cardinal utilities, and obtain positive results for this measure.

Boutilier et al. [2015] consider randomized ordinal mechanisms for the unit-cost case, when a single alternative is to be selected, and show a lower bound of $\Omega(\sqrt{m})$ and an upper bound of $O(\sqrt{m}\log^* m)$ on the distortion. They give a randomized mechanism which uses a harmonic scoring function with distortion $O(\sqrt{m \log m})$. We use this scoring function in our work as well. Further, they show that the ordinal mechanism that obtains least distortion in an instance can be computed in polynomial time. Boutilier et al. also consider a distributional model for utilities, and show that for neutral distributions, the optimal ordinal mechanism is a scoring function. Finally, they develop a learning-theoretical model, and analyse the sample as well as computational complexity of optimal mechanisms.

For participatory budgeting, when candidates have costs and a budget-feasible subset of items is to be selected, Benade et al. [2017] analyse the distortion for different mechanisms. For ordinal mechanisms they show that the distortion is bounded from above by $O(\sqrt{m}\log m)$, and below by $\Omega(\sqrt{m})$. They also introduce a new input format called threshold approval voting, where a real-valued threshold is fixed and each agent reports the candidates with utility above the threshold. For this input format, the distortion is shown to be bounded by $O(\log^2 m)$ and $\Omega(\log m/\log\log m)$. If a single candidate is to be selected, a mechanism with distortion $O(\log m)$ is given.

Caragiannis et al. [2017] study the distortion of ordinal mechanisms when a size-k subset of candidates is to be selected, and the welfare of a subset for an agent is the maximum utility of a candidate in the subset. This is in contrast to our problem, when the welfare of a subset for an agent is the sum of utilities of candidates in the subset. In this setting, Caragiannis et al. provide nearly tight bounds on the distortion for both deterministic and randomized mechanisms.

The previous papers did not consider the effect of strategic misreporting by agents. Truthful ordinal mechanisms are characterized by Gibbard and Satterthwaite [Gibbard, 1973, 1977, Satterthwaite, 1975]. Procaccia [2010] studies the use of truthful randomized mechanisms to approximate mechanisms obtained from prominent scoring rules. A randomized mechanism is said to be a γ-approximation of a scoring rule if the candidate chosen by the mechanism has expected score within $1/\gamma$ fraction of the maximum score. Procaccia shows that a class of mechanisms called positional scoring rules, which includes Plurality, can be approximated within $O(\sqrt{m})$, and this is tight. Borda, in particular can be approximated to within almost a factor of 2. He shows similar results for Copeland and Lull, and a lower bound of $\Omega(m)$ for Maximin.

The mechanism used in many applications of participatory budgeting is k-approval voting, where agents select a subset of the candidates of size k, and the mechanism selects the candidates with most votes that are also feasible given the budget. In k-approval voting, voters may not consider the costs of the candidate. In an effort to better align the constraints faced by the voters with those faced by the mechanism, Goel et al. [2015, 2016] introduce knapsack voting, where each voter selects a budget-feasible subset of candidates of arbitrary size. In different models of utility, and assuming that a candidate can be allocated fractionally, the authors show that knapsack voting is truthful and welfare-maximizing. They also show that empirically, knapsack voting appears to outperform k-approval voting. When a single candidate is to be selected, and agent costs (rather than utilities) form a metric, constant upper and lower bounds on the distortion are known [Anshelevich et al., 2015, Anshelevich and Postl, 2016], including for truthful mechanisms [Feldman et al., 2016].

A number of papers consider ordinal mechanisms that obtain optimal welfare under distributional assumptions on the utility functions of agents. For two candidates, Rae [1969] considers mechanisms when each agent prefers one candidate to the other with equal and independent probability [Rae, 1969]. Rae shows that among threshold mechanisms, which select a candidate if the number of votes for the candidate exceed a threshold, simple majority maximizes the expected welfare. If a candidate is preferred with greater probability than the other, simple majority may not maximize welfare. These results were extended to more general distributions in later work [Badger, 1972, Curtis, 1972, Schofield, 1972]. Schmitz and Tröger show that if utilities are i.i.d. across agents and candidates, majority rules and variants of these rules maximize the welfare among all truthful mechanisms [Schmitz and Tröger, 2012]. This may not be true if agent utilities are correlated. They also characterize truthful mechanisms in this setting, showing in particular that truthful mechanisms must ignore the utilities agents have for candidates, and consider as input only the preference order. These results are further extended to independently (but not identically) distributed utilities [Azrieli and Kim, 2014].

If agents' utilities for candidates are i.i.d., the optimal ordinal mechanisms for more than two candidates and for three different notions of welfare — utilitarian social welfare, maximin, and maximax — are known to be scoring rules [Apesteguia et al., 2011]. This work ignores strategic behaviour. Later work by Kim et al. [2012] shows that the social welfare obtained by any ordinal ex-ante Pareto-efficient mechanism, including the previous mechanisms, is also obtained by an ordinal Bayesian incentive-compatible (BIC) mechanism [Kim et al., 2012]. This is more generally true when each agent's utilities are neutral with regard to candidates, i.e., for an agent and vector $u \in \mathbb{R}^m$, any permutation of u is equally likely to be the utility vector for the agent. Kim also shows that ordinal mechanisms may not maximize social welfare among BIC mechanisms. Specifically, he gives an example with three alternatives, where a BIC mechanism that considers agent utilities has strictly greater social welfare than any ordinal BIC mechanism.

2 NOTATION AND PRELIMINARIES

In the participatory budgeting problem, a population of agents N of size n is faced with the problem of selecting candidates from a set C of size m. We will use i, j, k for agents, and x, y, z for candidates. Each candidate x has nonnegative cost c_x. There is a fixed budget $B = 1$, and the set of candidates selected must have total cost at most 1; we say that such a set of candidates is feasible. In the unit-cost case, each candidate has cost 1, and hence at most one candidate is to be selected. In the k-selection problem for $k \in \mathbb{Z}_+$, each candidate has cost $1/k$.

Each agent has a utility function $u_i : C \to \mathbb{R}_+$ so that the sum of utilities $\sum_{x \in C} u_i(x)$ for each agent is exactly 1. This is a standard normalization assumption to ensure that all agents have equal influence. For technical reasons we will assume that if $x \neq y$, then $u_i(x) \neq u_i(y)$ for any agent i. Let $\vec{u} := (u_i)_{i \in N}$ be a vector of utility functions, or a utility profile, for the agents. For a candidate x, we define the utilitarian social welfare (or simply welfare) $\mathrm{uw}(x)$ to be $\sum_i u_i(x)$, the sum of utilities of agents for that candidate. The welfare of a set of candidates is the sum of welfare of the candidates in the set. Agent i's utility for a subset $S \subseteq C$ is $u_i(S) := \sum_{x \in S} u_i(x)$.

We use $\lambda = ((u_i)_{i\in N}, (c_x)_{x\in C})$ to refer to an instance of participatory budgeting, and Λ as the set of all possible instances. For an instance λ, let $F(\lambda)$ be the set of all feasible subsets of C. Define $S^*(\lambda)$ as the feasible subset with maximum welfare, and $\mathbf{OPT}(\lambda)$ as the welfare of this set. If the instance λ is clear, we use S^* and **OPT** to simplify notation.

Input formats and distortion. Our objective is to design mechanisms, possibly randomized, to select feasible subsets of maximum welfare. If the utility functions of the agents were available to the mechanism, this would be a simple problem[5]. However, keeping in mind that human agents find it burdensome to accurately report their utility functions, we consider mechanisms with differing input formats, that describe what information is elicited from the agents.

- *Independent threshold approval votes:* Each agent i is given a real-valued threshold T_i and returns a subset $S_i \subseteq C$. We say a utility profile $\vec{u} = (u_i)_{i\in N}$ and the subsets $\vec{S} = (S_i)_{i\in N}$ are consistent, written $\vec{u} \cong \vec{S}$, if $S_i = \{x \in C : u_i(x) \geq T_i\}$, i.e., S_i is the set of candidates with utility greater than or equal to agent i's threshold for all i.
- *Binary threshold approval votes:* Each agent i is given a real-valued threshold T_i and a candidate x_i, and returns a bit b_i. We say a utility profile $\vec{u} = (u_i)_{i\in N}$ and the vector $(\vec{b}, \vec{S}) = (b_i, S_i)_{i\in N}$ are consistent, written $\vec{u} \cong (\vec{b}, \vec{S})$, if $b_i = 1$ implies $u_i(x_i) \geq T_i$.
- *Ordinal votes:* Each agent i returns a linear order $\prec_i$ of C. Let $\vec{\prec} := (\prec_i)_{i\in N}$. We say that $\vec{\prec}$ is consistent with utility profile $\vec{u} = (u_i)_{i\in N}$, written $\vec{\prec} \cong \vec{u}$, if for each agent i, $u_i(x) > u(y)$ implies $x \prec_i y$. A mechanism with this input format is called an ordinal mechanism.

We note that there is thus a difference between an instance of participatory budgeting and an input to a mechanism. While the former includes the utility function for each agent, the latter may not, depending on the input format. Further, the costs for candidates $(c_x)_{c\in C}$ will always be implicit inputs to the mechanism.

Given an input format, a mechanism μ is defined as a map, possibly randomized, from possible inputs to distributions over feasible subsets of C. We consider as our primary measure of efficiency of a mechanism its distortion [Benade et al., 2017, Procaccia and Rosenschein, 2006]. To define distortion formally, consider the ordinal votes input format. Then for a mechanism μ and input $I = \left(\vec{\prec}, (c_x)_{x\in C}\right)$, the distortion is defined as the worst case ratio over all possible utility functions consistent with $\vec{\prec}$ of the maximum utility of a feasible subset, to the expected utility obtained by the mechanism.

$$\mathrm{dist}(\mu, I) := \sup_{\vec{u}\cong\vec{\prec}} \frac{\max\{\sum_{i\in N,\, x\in S} u_i(x) : \sum_{x\in S} c_x \leq 1\}}{\mathbb{E}_{S\sim\mu(\vec{\prec})} \sum_{i\in N,\, x\in S} \sum_{i\in N} u_i(x)} \,.$$

Correspondingly, the distortion of a mechanism is defined as the maximum distortion over all possible inputs.

$$\mathrm{dist}(\mu) := \sup_I \mathrm{dist}(\mu, I)\,.$$

For mechanisms where the inputs are obtained deterministically, even when the mechanism is randomized, taking the supremum over all utility functions consistent with the mechanism input is appropriate. However, in Section 3 we will consider mechanisms for the threshold approval input formats when the thresholds are chosen randomly, and hence, the inputs themselves are random variables. For these mechanisms, we show that for any instance, in expectation over the input to

[5]Ignoring computational considerations. Since the problem is NP-hard even for the case of a single agent when we are given the utility function, we will mostly ignore computational considerations. Note however that the mechanisms we describe in Sections 4 and 5 run in polynomial time.

the mechanism, the welfare obtained is large. Here, taking the supremum over all utility functions consistent with the mechanism input would in effect remove the randomization and we would be doing a worst-case analysis — not just over inputs, but also over the random bits, which defeats the purpose of randomization.

For mechanisms where the input is itself randomized, we therefore propose and use the following simpler definition of distortion[6]:

$$\mathrm{dist}(\mu) := \sup_{\lambda \in \Lambda} \frac{\mathbf{OPT}(\lambda)}{\mathbb{E}_{S\sim\mu(\lambda)}[\mathrm{uw}(S)]}$$

The definition is particularly appealing because it corresponds to the approximation ratio studied for randomized algorithms. The difference is that for the approximation ratio, typically the constraint is computational complexity, whereas for us there are many constraints, including the input format, computational complexity, and truthfulness.

Truthfulness. Truthful ordinal mechanisms are our focus in Section 4, and we define truthfulness with respect to these mechanisms. We say a mechanism is truthful if in any instance, each agent obtains maximum utility in expectation by reporting the linear order $\prec_i$ consistent with its utility function. Formally, for any instance $\lambda = ((u_i)_{i\in N}, (c_x)_{x\in C})$, let $\vec{\precsim}'$ be an arbitrary vector of linear orders over C, and let $\vec{\precsim}$ be $\vec{\precsim}'$ with the ith component replaced by the linear order consistent with u_i. Then mechanism μ is truthful if:

$$\mathbb{E}_{S\sim\mu(\vec{\precsim})}[u_i(S)] \geq \mathbb{E}_{S\sim\mu(\vec{\precsim}')}[u_i(S)]$$

where we assume that the costs c_x are implicit inputs to the mechanism. Note that we do not insist that the components of $\vec{\precsim}$, $\vec{\precsim}'$ other than the ith are consistent with the utility functions of the other agents, hence we require truthfulness to be a dominant strategy.

For a nonnegative integer n, we use $[n]$ to denote the set $\{1, \ldots, n\}$. $H_n = \sum_{i\in[n]} 1/i$ is the nth Harmonic number, and $\log(n+1) \leq H_n \leq 1 + \log n$. For a set S of candidates, we define S^c as the complement $C \setminus S$.

3 INDEPENDENT THRESHOLD MECHANISMS

We start with a mechanism that, when the number of agents is large, gives nearly optimal distortion. Our mechanism uses the idea of randomized thresholds as in Benade et al. [2017]. but presents each agent with a different randomized threshold, and for large enough agents obtains a solution with distortion is close to 1. We note that if the same threshold were presented to all the agents, Benade et al. show a lower bound of $\Omega(\log m / \log\log m)$ on the distortion. Our mechanism thus places the same informational load on each agent, but obtains a significantly lower distortion.

Theorem 3.1. *Let* $\delta = m^2\sqrt{\frac{18\log(2mn)}{n}}$. *Mechanism 1 returns set* S *so that* $\mathrm{uw}(S) \geq \mathbf{OPT}(1-\delta)$ *with probability at least* $(1 - 1/n)$.

Proof. We note that for each agent i and candidate x, since the threshold T_i is drawn from the uniform distribution over $[0, 1]$,

$$\mathbb{E}[V_{i,x} | x \in S_i] = \mathbb{E}[2T_i | u_i(x) \geq T_i] = u_i(x)$$

and similarly,

[6]Benade et al. [2017] also analyse a version of independent threshold approval votes when each agent gets the same threshold. Their mechanism input is thus also randomized, and in communication with the authors it appears the distortion bounds they obtain hold under the simpler definition of distortion.

Mechanism 1 Independent Thresholds Mechanism

for each agent i **do**
  $T_i \sim \mathcal{U}[0,1]$ ▷ Randomized threshold for agent i
  $S_i \leftarrow \{x \in C : u_i(x) \geq T_i\}$ ▷ agent i returns all candidates with value above the threshold
  for each candidate $x \in C$ **do**
    if $x \in S_i$ **then**
      $V_{i,x} \leftarrow 2T_i$
    else
      $V_{i,x} \leftarrow 2T_i - 1$
$\bar{V}_x \leftarrow \frac{1}{n}\sum_i V_{i,x}$ for each candidate x
return $S \leftarrow \arg\max_{T \subseteq C : \sum_{x \in T} c_x \leq 1} \{\sum_{x \in T} \bar{V}_x\}$

$$\mathbb{E}[V_{i,x} | x \notin S_i] = \mathbb{E}[2T_i - 1 | u_i(x) < T_i] = 2\left(u_i(x) + \frac{1 - u_i(x)}{2}\right) - 1 = u_i(x)\,.$$

Hence the random variables $V_{i,x}$ each have expected value $u_i(x)$, and are independent across agents. Further, each $V_{i,x}$ lies in the interval $[-1, 2]$. Let $\bar{U}_x = \text{uw}(x)/n$. Using Hoeffding's inequality, we see that for each candidate x,

$$\mathbb{P}[|\bar{V}_x - \bar{U}_x| \geq v] \leq 2\exp\left(\frac{-2nv^2}{9}\right)$$

Choose $v = \sqrt{\frac{9\log(2mn)}{2n}}$. Then for each x, $\mathbb{P}[|\bar{V}_x - \bar{U}_x| \geq v] \leq 1/(mn)$ and using the union bound, we obtain that with probability at least $1 - 1/n$,

$$|\bar{U}_x - \bar{V}_x| \leq v = \frac{\delta}{2m^2} \tag{1}$$

holds simultaneously for all candidates. We assume this is the case for the remainder of the proof.

Noting that the sets S and S^* contain at most m candidates,

$$\mathbf{OPT} = n\sum_{x \in S^*} \bar{U}_x \leq n\sum_{x \in S^*} \bar{V}_x + nmv \leq n\sum_{x \in S} \bar{V}_x + nmv \leq \sum_{x \in S} \bar{U}_x + 2nmv = \text{uw}(S) + 2nmv$$

where the second inequality follows from the choice of the set S and the first and third come from Eqn. (1). Since $v = \delta/(2m^2)$, $\mathbf{OPT} \leq \text{uw}(S) + n\delta/m$. Since $\mathbf{OPT} \geq n/m$, we get that $\text{uw}(S) \geq \mathbf{OPT} - \delta\mathbf{OPT}$, completing the proof. □

Corollary 3.2. *Given any $\epsilon > 0$ for fixed m and sufficiently large n, the distortion of Mechanism 1 is less than $1 + \epsilon$.*

Proof. By Theorem 3.1 the distortion of Mechanism 1 is at most $\frac{n}{(n-1)(1-\delta)}$ which, for fixed m, tends to 1 as $n \to \infty$. □

An independent threshold mechanism with binary inputs

As in Mechanism 1, in Mechanism 2 we select an independent random threshold for each agent. However, instead of asking them about all the candidates, we ask them about only one candidate. Which particular candidate we ask about is effectively random because we implement it by randomly equipartitioning N into m random sets of agents, one for each candidate. We will show that when the number of agents is large enough, with high probability this is arbitrarily close to optimal.

Mechanism 2 Independent Thresholds and Candidates Mechanism

Let $A_1, A_2, \ldots, A_m$ be a uniformly random partition of N into m subsets of size $\lceil n/m \rceil$ or $\lfloor n/m \rfloor$.
for each candidate $x \in [m]$ **do**
  for each agent $i \in A_x$ **do**
    $T_i \sim \mathcal{U}[0, 1]$ ▷ Randomized threshold for agent i
    $b_i \leftarrow 1$ if $u_i(x) \geq T_i$, else $b_i \leftarrow 0$ ▷ agent i returns if he values candidate x above threshold T_i
    if $b_i = 1$ **then**
      $V_{i,x} \leftarrow 2T_i$
    else
      $V_{i,x} \leftarrow 2T_i - 1$
  $\bar{V}_x \leftarrow \frac{1}{|A_x|} \sum_{i \in A_x} V_{i,x}$
$S \leftarrow \arg\max_{T \subseteq C : \sum_{x \in T} c_x \leq 1} \{\sum_{x \in T} \bar{V}_x\}$
return S

Theorem 3.3. *Let* $\delta = m^{5/2}\sqrt{\frac{72\log(4mn)}{n}}$. *Mechanism 2 returns set* S *so that* $\mathrm{uw}(S) \geq \mathbf{OPT}(1 - \delta)$ *with probability at least* $(1 - 1/n)$.

The proof of Theorem 3.3 is moved to the appendix. As with Theorem 3.1, the following corollary is an immediate consequence.

Corollary 3.4. *Given* $\epsilon > 0$ *for fixed m and sufficiently large n, the distortion of Mechanism 1 is less than* $1 + \epsilon$.

4 TRUTHFUL ORDINAL MECHANISMS

We now consider ordinal mechanisms, and show that the lower bound of $\Omega(\log m)$ on the distortion of ordinal mechanisms [Boutilier et al., 2015] can in fact nearly be achieved by *truthful* ordinal mechanisms.

Recall that in an ordinal mechanism, each agent i returns a linear order $\prec_i$ of C. Let $\vec{\prec} := (\prec_i)_{i \in N}$. We say that $\vec{\prec}$ is consistent with a utility profile $\vec{u} = (u_i)_{i \in N}$, written $\vec{\prec} \cong \vec{u}$, if for each agent i, $u_i(x) > u(y)$ implies $x \prec_i y$. Given a linear order $\prec_i$ for agent i and a candidate x, $\mathrm{rk}_i(x)$ is the number of candidates that i prefers to x (including x), i.e., $|\{y : y \preceq_i x\}|$. For $S \subseteq C$, and candidate $x \in S$, we define $\mathrm{rk}_i(x|S) := |\{y \in S : y \preceq_i x\}|$ as the number of candidates in S that i prefers to x, including x. Define the score $\mathrm{score}(x|S) = \sum_{i \in N} 1/\mathrm{rk}_i(x|S)$, and $\mathrm{score}(x) := \mathrm{score}(x|C)$. As before, S^* is the feasible set of candidates with maximum welfare, and $\mathbf{OPT} = \mathrm{uw}(S^*)$.

We first give a randomized mechanism with distortion $O(\sqrt{m \log m})$ for the k-selection problem, where each candidate has cost $1/k$ for $k \in \mathbb{Z}_+$, and hence k candidates are to be chosen. Our mechanism runs the Harmonic Scoring mechanism [Boutilier et al., 2015] as the sampling subroutine, and outputs either the resulting *single candidate* with probability $1/2$, or a randomly chosen subset of size k[7]. We will use this mechanism as a subroutine later on with subsets of C as possible candidates, and hence explicitly give the set of candidates as an input.

We note that if $S = Z$ in Mechanism 3, a single candidate is chosen. Further, for any candidate i, the sum of ranks of candidates $\sum_{x \in A} 1/\mathrm{rk}_i(x) = H_m$. Hence, for any candidate x,

[7]If it is important that a subset of candidates of size k be returned, in the first case, we can always add $k - 1$ randomly chosen candidates. This affects neither the truthfulness nor the upper bound on the distortion.

Mechanism 3 k-Selection

Input: Set A of m candidates, k

1: Let i be a randomly chosen agent. Sample z at random from A with probability proportional to $1/\mathrm{rk}_i(z)$. Let $Z \leftarrow z$.
2: Y is a set of size k, sampled uniformly from A.
3: S is chosen from $\{Y, Z\}$ with equal probability
4: **return** S

$$\mathbb{P}[x \in Z] = \sum_{i \in N} \frac{1}{n} \frac{1}{H_m \mathrm{rk}_i(x)} = \frac{\mathrm{score}(x)}{nH_m} \,. \tag{2}$$

We first show that the mechanism is truthful.

Theorem 4.1. *The k-Selection Mechanism is truthful.*

Proof. Fix an agent i. In the rest of the proof, we condition on i being chosen in Step 1, since otherwise i's expected utility is independent of its input to the mechanism, and it has no reason to be untruthful. For $r \in [m]$, define $p_r = 1/(r\, H_m)$, and note that this is exactly the probability that a candidate ranked r by agent i is selected. That is, for any candidate x, if $\mathrm{rk}_i(x) = r$, then $\mathbb{P}[x \in Z] = p_r$.

Let u_i be agent i's utility function, let $\prec_i$ be the ordering consistent with u_i and $\widehat{\prec}_i$ be an arbitrary ordering not equal to $\prec_i$. We will show that i's utility is maximized in expectation if it reports $\prec_i$ to the mechanism. We index the candidates so that $x_1 \prec_i x_2 \prec_i \ldots x_m$. Let π be a permutation of $[m]$ so that x_r's rank in $\widehat{\prec}_i$ is $\pi(r)$ for $r \in [m]$.

The change in expected utilities for agent i if it reports $\widehat{\prec}_i$ is given by

$$\sum_{r=1}^{m} u_i(x_r) p_{\pi(r)} - \sum_{r=1}^{m} u_i(x_r) p_r$$

which is nonpositive by the rearrangement inequality, since both $u_i(x_r)$ and p_r are decreasing in r. Hence, agent i's expected utility is maximized by correctly reporting its true linear order. □

We now prove the bound on the social welfare.

Theorem 4.2. *The expected social welfare of S is at least* $\mathbf{OPT}/4\sqrt{m \log m}$.

Proof. Let $S^* = \{x_1^*, x_2^*, \ldots, x_k^*\}$ be the optimal set of candidates. Further, let S_1^* be the set of candidates in S^* with score at least $n\sqrt{\log m/m}$, and $S_2^* = S^* \setminus S_1^*$.

Note that for any candidate x, $\mathrm{score}(x)$ is at least $\mathrm{uw}(x)$. This is because for any agent i the sum of utilities over the candidates sums to 1, and for any candidate x there are $\mathrm{rk}_i(x)$ candidates with at least as much utility. Thus for any candidate x, $u_i(x) \leq 1/\mathrm{rk}_i(x)$. Then

$$\mathrm{uw}(x) = \sum_{i \in N} u_i(x) \leq \sum_{i \in N} \frac{1}{\mathrm{rk}_i(x)} = \mathrm{score}(x) \,. \tag{3}$$

We first show that in expectation, the social welfare of Z is at least $1/\sqrt{m \log m}$ times the social welfare of S_1^*. Intuitively, the reason that a single candidate in Z has good social welfare compared to S_1^* is that the size of S_1^* is at most $\sqrt{m \log m}$. This is because, as mentioned above, the sum of scores of candidates is nH_m, while each candidate in S_1^* by definition has score $n\sqrt{\log m/m}$. More formally, the expected social welfare of Z is:

$$\mathbb{E}[\mathrm{uw}(Z)] \geq \sum_{x \in S_1^*} \mathrm{uw}(x)\mathbb{P}[x \in Z] = \sum_{x \in S_1^*} \mathrm{uw}(x)\frac{\mathrm{score}(x)}{nH_m} \geq \frac{n\sqrt{\log m/m}}{nH_m} \sum_{x \in S_1^*} \mathrm{uw}(x) \geq \frac{\mathrm{uw}(S_1^*)}{2\sqrt{m \log m}} \,.$$

Above, the first equality follows from (2), and the second inequality by definition of S_1^*.

We now show that the same inequality holds for sets Y and S_2^* as well. By definition, the expected social welfare of Y is exactly kn/m, since it is a subset of size k chosen uniformly at random, and the welfare of a randomly chosen candidate is n/m. Further, we can obtain the following upper bound on the social welfare of S_2^* :

$$\sum_{x \in S_2^*} \mathrm{uw}(x) \leq \sum_{x \in S_2^*} \mathrm{score}(x) \leq |S_2^*| n\sqrt{\log m/m} \leq kn\sqrt{\log m/m} \,.$$

The first inequality is from (3), and the second by definition of S_2^*. Then comparing the two, we obtain

$$\sum_{x \in S_2^*} \mathrm{uw}(x) \leq kn\sqrt{\log m/m} = k\frac{n}{m}\sqrt{m \log m} = \mathbb{E}[\mathrm{uw}(Y)]\sqrt{m \log m} \,.$$

The expected welfare of $Y + Z$ is then at least $\mathbf{OPT}/2\sqrt{m \log m}$. Since $\mathbb{E}[\mathrm{uw}(S)] = (\mathbb{E}[\mathrm{uw}(Y) + \mathrm{uw}(Z)])/2$, the proof follows. □

We now adapt the k-Selection Mechanism to general costs. Our mechanism uses the Ranking-by-Value Mechanism from Benade et al. [2017], except that we use the above k-selection mechanism in place of Mechanism A by Benade et al. to recover truthfulness.

Mechanism 4 Truthful Ranking-by-Value

1: For $s \in [\log m]$, define $l_s = 2^{s-1}/m$, $u_s = 2^s/m$

2: Let $T_0 := \{x : c_x \leq 1/m\}$, and $T_s = \{x : l_s < c_x \leq u_s\}$ for $s \in [\log m]$. Let $m_s = |T_s|$, $s \in [\log m] \cup \{0\}$ ▷ $1/u_s$ candidates can be chosen from T_s within the budget.

3: Choose $r \in [\log m] \cup \{0\}$, where $\mathbb{P}[r = s] \propto \sqrt{m_s \log m_s}$

4: Run the k-Selection Mechanism with inputs T_r and $k = 1/u_r$. Let U be the set of candidates returned.

5: **return** U

For truthfulness of the mechanism, note that the sets T_s as well as r are decided independent of the reported linear orders. Given T_r, we can restrict attention to the valuations given by an agent j to candidates in T_r. The proof of Theorem 4.1 then shows that the expected utility of agent j is maximized by reporting the true linear orders for these candidates, and the proof of truthfulness follows.

Theorem 4.3. *The Truthful Ranking-by-Value Mechanism has distortion* $O\left(\sqrt{m} \log m\right)$.

Proof. For $s \in [\log m] \cup \{0\}$, let T_s^* be the set of candidates in T_s that are budget-feasible and maximize the social welfare. Then since $c_x \geq l_s$ for $x \in T_s$ and the budget is 1, $|T_s^*| \leq 1/l_s = 2/u_s$. Let $T_s' \subseteq T_s$ be the set of candidates of size $1/u_s$ with maximum social welfare. Then

$$\mathrm{uw}(T_s') \geq \frac{1}{2}\mathrm{uw}(T_s^*) \geq \frac{1}{2}\mathrm{uw}(S^* \cap T_s) \,.$$

By Theorem 4, if $r = s$, then for the set U returned by the k-selection mechanism,

$$\mathbb{E}[\mathrm{uw}(U)] \geq \frac{1}{4}\frac{\mathrm{uw}(S^* \cap T_s)}{\sqrt{m_s \log m_s}} .$$

The expected social welfare of the set U returned is thus

$$\sum_{s=0}^{\log m} \mathbb{P}[r = s]\frac{1}{4}\frac{\mathrm{uw}(S^* \cap T_s)}{\sqrt{m_s \log m_s}} = \frac{1}{4\sum_{s=0}^{\log m}\sqrt{m_s \log m_s}}\mathrm{uw}(S^*) .$$

Since $\sum_s m_s = m$ and $\sqrt{x \log x}$ is concave, the expression on the right is maximized when the m_s's are all equal, in which case the expected social welfare is $\mathbf{OPT}/\left(\sqrt{m}\log m\right)$. □

A tight lower bound for unilaterally-neutral truthful ordinal mechanisms

Recall that for an ordinal mechanism μ, the input consists of linear orders $\vec{\prec} = (\prec_i)_{i \in N}$ reported by the agents, and costs of the candidates $(c_x)_{x \in A}$ for the given instance. The latter is implicit, and is not part of the notation. We present lower bounds for the unit-cost problem, when a single candidate is to be selected.

Our proof relies on the characterization of truthful randomized mechanisms by Gibbard. We first give the necessary definitions, followed by Gibbard's characterization.

Definition 4.4 ([Gibbard, 1977]). A mechanism μ is *duple* if there exist candidates x_1, x_2 such that for any input $\vec{\prec}$, if candidate $x \notin \{x_1, x_2\}$, the probability that x is chosen is zero. In this case, we say that x_1 and x_2 are the choices for duple mechanism μ. The mechanism is *unilateral* if it depends on a single agent, i.e., there exists $i \in N$ such that for all $\vec{\prec}$, $\vec{\prec}'$, if $\prec_i = \prec_i'$, then $\mu(\vec{\prec}) = \mu(\vec{\prec}')$. In this case, we say the mechanism depends on voter i.

Gibbard's characterization requires any truthful mechanism to be localized and non-perverse. Informally, a mechanism is localized if changing the relative ordering of two candidates by an agent does not affect the probability that any other candidate is selected by the mechanism. A mechanism is non-perverse if increasing the rank of a candidate by an agent cannot decrease the probability that the candidate is selected. We will use non-perversity in our proof, and hence define it formally here.

Definition 4.5 ([Gibbard, 1977]). Consider linear orders $\vec{\prec}$, $\vec{\prec}'$ and agent i so that $\prec_j = \prec_j'$ for every agent $j \neq i$. Further, x and y are candidates so that $x \prec_i y$ and for any $z \notin \{x, y\}$, either $y \prec_i z$ or $z \prec_i x$. In $\prec_i'$, candidates x and y are switched , so that $y \prec_i' x$ and the relative order of other candidates does not change. Mechanism μ is *non-perverse* if $\mathbb{P}[\mu(\vec{\prec}') = y] \geq \mathbb{P}[\mu(\vec{\prec}) = y]$.

Theorem 4.6 ([Gibbard, 1977]). *A mechanism μ is truthful if and only if it is a probability distribution over mechanisms, each of which is localized, non-perverse, and either unilateral or duple.*

We will use the following notation. Let $\prec$ be a linear order over the candidates, and let the candidates be indexed in this order, so that $x_1 \prec x_2 \prec \cdots \prec x_m$. Let $\pi \in \mathrm{S}_m$ be a permutation of $[m]$. We define $\prec^\pi$ as the linear order where x_r has rank $\pi(r)$. We define $\vec{\prec}^\pi = (\prec_i^\pi)_{i \in N}$. Thus $\vec{\prec}^\pi$ is obtained by relabeling each candidate x_r as $x_{\pi^{-1}(r)}$.

Definition 4.7. A unilateral mechanism is *neutral* if for any $\vec{\prec}$, $\pi \in \mathrm{S}_m$, and candidate x_r, $\mathbb{P}(\mu(\vec{\prec}) = x_r) = \mathbb{P}\left(\mu(\vec{\prec}^\pi) = x_{\pi^{-1}(r)}\right)$. A truthful mechanism is *unilaterally neutral* if each unilateral mechanism in its support is neutral.

Lemma 4.8. *If μ is a neutral unilateral mechanism that depends on agent i, there exists a probability distribution $(p_i)_{i \in [m]}$ so that for any input $\vec{\prec}$, the probability that candidate x is selected is exactly $p_{\mathrm{rk}_i(x)}$.*

Proof. Fix an input $\vec{\prec}$, assume the candidates are ordered so that $x_1 \prec_i x_2 \prec_i \cdots \prec_i x_m$, and let p_i be the probability that x_i is chosen by the mechanism. Now let $\prec_i'$ be a different linear order for i, and $\pi(r) = \mathrm{rk}_i(x_r)$ in $\prec_i'$, i.e., the candidate ranked r in $\prec_i$ is ranked $\pi(r)$ in $\prec_i'$. Since μ is a unilateral mechanism, its output distribution is the same for inputs $\left(\prec_{-i}, \prec_i^{\pi}\right)$ and $\vec{\prec}^{\pi}$. Now the probability that μ chooses x_r on input $\vec{\prec}^{\pi}$ can be obtained from the definition of neutral mechanisms as:

$$\mathbb{P}\left(\mu(\vec{\prec}^{\pi}) = x_r\right) = \mathbb{P}(\mu(\vec{\prec}) = x_{\pi(r)}) = p_{\pi(r)} \,.$$

Thus, if the rank of $x(r)$ is now $\pi(r)$, the probability that it is chosen is also $p_{\pi(r)}$, as required. □

We use this to simplify notation as follows. Any unilaterally neutral truthful mechanism μ is a probability mixture of duple and neutral unilateral mechanisms. We assume that there are n neutral unilateral mechanisms in its support $\mu^1, \ldots, \mu^n$, with μ^i depending on agent i. If multiple unilateral mechanisms depend on agent i we combine them into a single distribution, and if μ does not depend on agent i, we will assume the corresponding unilateral mechanism μ^i is selected with probability zero.

$$\mu = \sum_{i=1}^{n} \alpha_i \mu^i + \sum_{j=1}^{o} \beta_j \nu^j$$

where ν^js are the duple mechanisms, and α_i, β_j are the probabilities that mechanisms μ^i, ν^j are picked. Thus if μ does not depend on agent i, we will assume α_i is zero.

Further, p_r^i is the probability that unilateral mechanism μ^i selects the candidate ranked r by agent i. By Lemma 4.8 these probabilities are well-defined. We now show that for unilateral mechanisms in the support of a truthful mechanism, the p_r^is are decreasing with the rank of the candidates.

Lemma 4.9. *Let μ^i be a neutral unilateral mechanism in the support of a truthful mechanism that depends on agent i. Let p_r^i be the probability that the candidate ranked r by agent i is selected by mechanism μ^i. Then $p_1^i \geq p_2^i \geq \cdots \geq p_m^i$.*

The proof is immediate from Gibbard's characterization that any unilateral mechanism in the support of a truthful mechanism must be non-perverse (Theorem 4.6). Then for any agent i, $p_m^i \leq 1/m$.

Theorem 4.10. *Let μ be a truthful ordinal mechanism for unit-cost instances that is unilaterally neutral. Then the distortion of μ is at least $\sqrt{m \log m}/12$.*

Proof. Let $\gamma = \sqrt{m \log m}$. We prove this by contradiction: Let μ be such a mechanism with distortion strictly better than $\gamma/12$. Our proof proceeds as follows. We first show that we can ignore duple mechanisms in the support of μ, since there is at least one candidate x_0 which they select with very low probability. This will be our optimal candidate. We then focus on unilateral mechanisms, and recall that there is a correspondence between unilateral mechanisms in the support of μ and agents. We define $\hat{p}_r$ as (approximately) the probability that the unilateral mechanism that is selected by μ, selects the candidate ranked r by the corresponding agent. The sum of $\hat{p}_r$ over all r should be 1. We will construct a series of instances, one for each r, to show that, in order to get distortion better than $\gamma/12$, the probabilities $\hat{p}_r$ must follow a roughly harmonic progression, i.e., $\hat{p}_r \geq c/(r \log m)$ for some constant c. However, the constant c will be large enough that summing over these lower bounds will give us something larger than 1, giving us a contradiction.

We start with some definitions that will simplify notation in the proof. For a set of agents D, we define $\alpha(D) = \sum_{i \in D} \alpha_i$ as the probability that a unilateral mechanism corresponding to an agent in D is selected. As before, p_r^i is the probability that the unilateral mechanism for agent i selects the candidate ranked r by agent i. Define $p_r(D) = \sum_{i \in D} \alpha_i p_r^i / \alpha(D)$ as the probability that an agent in D selects the candidate it ranks r, conditioned on an agent in D being selected by μ. Note that for any subset D, the sum $\sum_{r=1}^{m} p_r(D) = 1$.

We will restrict our attention to unilateral mechanisms in the support of μ that are picked with near-uniform probability. Let N' be a set of $n/2$ agents for so that $\alpha_i \leq 2/n$ for each $i \in N'$. Such a set must exist, since there can be at most $n/2$ agents for which $\alpha_i > 2/n$. Let $\hat{p}_r = p_r(N')$.

Note that for any D that is a subset of N', since $\alpha_i \leq 2/n$ for each $i \in N'$, we get that

$$\alpha(D) \leq 2|D|/n \,. \tag{4}$$

We show that there is a candidate x_0 for which the probability that it is selected by a duple mechanism is at most $2/m$. That is, if $I(x)$ is the set of indices of duple mechanisms for which x is a choice, then the probability that a duple mechanism from $I(x_0)$ is chosen is at most $4/m$. To see this, define $\beta(x) = \sum_{j \in I(x)} \beta_j$ as the probability that a duple mechanism that has x as a choice is chosen by μ. We want to show that $\beta(x_0) \leq 4/m$. Note that each duple mechanism has two choices, and hence if we sum $\beta(x)$ over all candidates, the sum is at most 2, since

$$\sum_x \beta(x) = \sum_x \sum_{j \in I(x)} \beta_j = \sum_{j=1}^{o} \beta_j \sum_{x : j \in I(x)} 1 \leq 2 \sum_{j=1}^{o} \beta_j \leq 2 \,.$$

Hence there must exist a candidate x_0 with $\beta(x_0) \leq 2/m$.

For each r, we will construct a different instance to show the lower bound $\hat{p}_r \geq c/(r \log m)$. Let us now describe the instance for a fixed r. Define $\beta_r = r\gamma/(2m)$ as a fraction less than 1, and note that then r must be at most $(2m)/\gamma$. Let $N_r \subseteq N'$ be a subset of agents of size $\beta_r n/2$, and for which $p_r(N_r)$, the probability that a candidate at rank r is selected conditioned on a mechanism in N_r being selected, is minimum among all such subsets. We note that for any larger subset of agents in N', this conditional probability will be larger than it is for N_r. Thus in particular, $\hat{p}_r \geq p_r(N_r)$.

The instance we construct has utility functions and preference orders with the following properties:

(1) All agents in N_r have rank r for candidate x_0, while all other agents place x_0 at rank m.
(2) All agents in N_r have utility approximately $1/r$ for the first r candidates, and 0 for the others. All other agents have utility approximately $1/m$ for all candidates.
(3) Restricted to agents in N_r, and for any rank $s \in [m]$, each candidate other than x_0 appears with approximately the same frequency at rank s.

For such an instance, candidate x_0 has utilitarian welfare at least $n\beta_r/(2r)$. For any other candidate the utilitarian welfare over agents in N_r is at most $1/r \cdot 1/(m-1) \cdot \beta_r n/2$, and at most $n/(m-1)$ for agents outside N_r. Replacing the value of $\beta_r = r\gamma/(2m)$, we get that the welfare for any candidate other than x_0 is at most $2n/m$.

Further, mechanism μ selects candidate x_0 with the following probabilities:

(1) a duple mechanism is selected, and selects x_0; this occurs with probability at most $(1 - \alpha(N))4/m \leq 4/m$.
(2) a unilateral mechanism from N_r is selected, and selects x_0; this occurs with probability at most $\alpha(N_r)p_r(N_r)$.
(3) a unilateral mechanism from N_r^c is selected, and selects x_0; this occurs with probability at most $\alpha(N_r^c)p_m(N_r^c)$, and since $p_m^i \leq 1/m$ for any unilateral mechanism i, this is at most $1/m$.

Thus the distortion for mechanism μ is at least

$$\frac{n\beta_r}{2r} \cdot \frac{1}{\frac{2n}{m} + \frac{n\beta_r}{2r}\left(\frac{5}{m} + \alpha(N_r)p_r(N_r)\right)} = 1/\left(\frac{r}{m\beta_r} + \frac{5}{m} + \alpha(N_r)p_r(N_r)\right) \le \gamma/12\,,$$

where the inequality is because by assumption, μ has distortion at most $\gamma/12$. Since N_r is a subset of N', we get from (4) that $\alpha(N_r)$ is at most $2|N_r|/n$, and since N_r has size $\beta_r n/2$, we get that $\alpha(N_r) \le \beta_r$. Further, substituting $\beta_r = \gamma r/(2m)$, and inverting both sides of the inequality, we get the following:

$$\frac{12}{\gamma} \le \frac{2}{\gamma} + \frac{5}{m} + \frac{r\gamma}{2m} p_r(N_r)\,.$$

For large enough m, $\frac{2}{\gamma}$ is much greater than $\frac{5}{m}$, so we ignore this second term. Some manipulation and using $\gamma = \sqrt{m \log m}$ then gives us the following lower bound for $p_r(N_r)$:

$$p_r(N_r) \ge \frac{10}{\gamma} \cdot \frac{2m}{\gamma r} = \frac{20}{r \log m}\,.$$

Thus, $\hat{p}_r \ge p_r(N_r) \ge 16/(r \log m)$. Since β_r is at most 1, this inequality only holds true for $r \le (2m)/\gamma$. But this is enough to show a contradiction, since

$$\sum_{r=1}^{2m/\gamma} \hat{p}_r \ge \frac{20}{\log m} \sum_{r=1}^{2m/\gamma} \frac{1}{r} \ge \frac{20}{\log m} \log\left(1 + \frac{2m}{\gamma}\right) \ge \frac{20}{\log m} \frac{1}{2} \log\left(\frac{m}{\log m}\right)$$

For large enough m, $m/\log m \ge \sqrt{m}$, and the expression on the right evaluates to 5. This is a contradiction, since the sum of conditional probabilities $\hat{p}_r$ must equal 1. □

5 OPTIMAL TRUTHFUL MECHANISMS FOR TWO CANDIDATES

We will now study the case where there are two candidates, and present optimal truthful mechanisms for this case. We will use a and b to denote the two candidates. If the sum of costs of both candidates is at most the budget, we should just select both of them, and this is clearly optimal. Hence we assume that each candidate has cost equal to the budget. As is in the rest of the paper, we assume that the utilities of an agent for the two candidates are not equal.

We start by showing that any truthful mechanism for two candidates, whether randomized or deterministic, must be ordinal. We note that similar results were earlier obtained in the setting where the utilities of agents for the candidates are drawn from independent distributions [Azrieli and Kim, 2014, Schmitz and Tröger, 2012].

Theorem 5.1. *Let μ be a truthful mechanism for two candidates. If utility profiles $\vec{u}, \vec{u}'$ are consistent with the same linear order $\vec{\prec}$, then $\mu(\vec{u}) = \mu(\vec{u}')$.*

Proof. We assume without loss of generality that $\vec{u}$ and $\vec{u}'$ differ in the vote of a single agent. This is justified by considering the output of the mechanism in steps, as each vote u_i is changed to u_i'. Since $\mu(\vec{u}) \neq \mu(\vec{u}')$, the output must change at some step when agent i's vote changes from u_i to u_i'. Let agent k be this pivotal agent. For a contradiction, assume $\mu(\vec{u}) \neq \mu(\vec{u}')$, and that candidate a is picked with strictly greater probability when agent k votes u_k'. Suppose that agent k has greater utility for candidate b (in both u_k and u_k', since they correspond to the same linear order). Since b is picked with greater probability when agent k votes u_k, agent k should vote u_k even when its utility function is u_k' to maximize its expected utility, contradicting the truthfulness of the mechanism. □

Given this result, we focus on ordinal mechanisms in the remainder of the section. We first consider deterministic mechanisms, and show that the MAJORITY mechanism that selects the candidate preferred by a majority of agents is truthful and has distortion 3, which is a lower bound for all truthful mechanisms.

Theorem 5.2. *majority is truthful, and has distortion 3. Any deterministic truthful mechanism has distortion at least 3.*

Our proof in the Appendix shows that majority is in fact input-optimal for ordinal inputs: for any deterministic mechanism μ with ordinal inputs, and any preference order $\vec{\precsim}$, the welfare of μ is at most that obtained by majority in the worst-case, i.e., over all utility functions consistent with $\vec{\precsim}$.

We now consider randomized mechanisms. As the lower bound in the deterministic setting illustrates, majority may achieve welfare that is only a third of the optimal. The reason for this comes from the case when the votes are nearly equal, but slightly biased towards candidate a. However the agents that prefer a have nearly equal utility for the two candidates, while the agents that prefer b have zero utility for a. One way to counteract this is to retain some probability of choosing the minority candidate.

An obvious choice for the probability distribution is to bias according to the observed separation, *i.e.,* if an α fraction of the players prefer candidate a then we choose a with probability α. A simple calculation (which we omit) shows that this algorithm achieves a worst-case distortion of $1/(4\sqrt{2}-5) \approx 1/0.6569 > 3/2$, and is achieved when $\alpha = \sqrt{2}-1$ fraction of agents prefer the optimal candidate, say a; the agents that prefer a are single-minded, while those that prefer b are ambivalent. In fact we can achieve a worst-case distortion of $3/2$ by slightly altering the distribution: after querying all the agents and seeing that αn of them prefer a while the rest prefer b, choose candidate a with probability $p(\alpha) = \frac{2\alpha-\alpha^2}{1+2\alpha-2\alpha^2}$. Note that $p(\alpha) = 1 - p(1-\alpha)$, hence a does not have to be the majority candidate to apply this. We call this mechanism tempered-majority.

Theorem 5.3. *tempered-majority is truthful.*

Proof. Follows simply because $p(\alpha)$ is an increasing function of α. If an agent i votes for candidate a, this increase the probability that a is selected, and hence this vote is utility-maximizing only if i's utility for a is greater than $1/2$. □

Theorem 5.4. *tempered-majority has a worst-case distortion of 3/2.*

Proof. Let α be the fraction of agents who prefer candidate a, while a $1-\alpha$ fraction prefer b. We will show in Theorem 5.5 below that on any instance with an $(\alpha, 1-\alpha)$ split, tempered-majority achieves a distortion of $1 + 2\alpha - 2\alpha^2$. Maximizing this over $\alpha \in [0, 1]$ we see that the worst case is when $\alpha = 1/2$ where the distortion is $3/2$, as claimed. □

Theorem 5.5. *On any instance in which α fraction of the agents prefer one candidate a and $(1-\alpha)$ prefer the other, tempered-majority achieves a distortion of $1 + 2\alpha - 2\alpha^2$.*

Proof. Since the expression $1 + 2\alpha - 2\alpha^2 = 1 + 2\alpha(1-\alpha)$ is symmetric in α and $1-\alpha$, without loss of generality, $\alpha \geq 1/2$. Let $V := \mathrm{uw}(a)$ be the utilitarian welfare for candidate a and let $v = V/n$ be the average welare. Then

$$\frac{\alpha}{2} \leq v \leq \frac{1+\alpha}{2}$$

where the lower bound comes from the observation that each of the αn agents who prefer a must have utility at least $1/2$ for a, while upper bound comes from the corresponding observation about $1-\alpha$ and $1-v$.

Now consider a mechanism that chooses a with probability p and b with probability $1-p$. Then it achieves an expected welfare of $\mathbf{ALG}_p = pV + (1-p)(n-V) = (2pv + 1 - p - v)n$. Also, $\mathbf{OPT} = \max\{V, n-V\}$. Let us consider $\mathbf{ALG}_p/\mathbf{OPT}$ as a function of v on the domain $[\frac{\alpha}{2}, \frac{1+\alpha}{2}]$. Thus

$$f_p(v) := \begin{cases} \frac{2v-1}{1-v}p + 1 & \text{for } \frac{\alpha}{2} \leq v < \frac{1}{2} \\ 2p - 1 + \frac{1-p}{v} & \text{for } \frac{1}{2} \leq v \leq \frac{1+\alpha}{2} \end{cases}$$

Then f_p is increasing on $[\frac{\alpha}{2}, \frac{1}{2}]$ and decreasing on $[\frac{1}{2}, \frac{1+\alpha}{2}]$ and therefore has local minima at the endpoints of its domain. The corresponding minimum values are $1 + \frac{2(\alpha-1)}{2-\alpha}p$ and $\frac{2\alpha p+1-\alpha}{1+\alpha}$, and the smaller of these is the inverse of the distortion of any mechanism that selects candidate a with probability p.

Now consider the behaviour of $1 + \frac{2(\alpha-1)}{2-\alpha}p$ and $\frac{2\alpha p+1-\alpha}{1+\alpha}$ as p varies. We see that the former is 1 at $p = 0$ and decreases thereafter, while the latter is increasing on $[0, 1]$ with a maximum value of 1 at $p = 1$. It follows that the best choice of p is the one that makes these two local minima equal. This happens at $p = \frac{2\alpha-\alpha^2}{1+2\alpha-2\alpha^2}$, which is the value used by TEMPERED-MAJORITY. Substituting this value of p back into the expression for minimum of f_p, we see that the distortion of TEMPERED-MAJORITY on any instance where an α fraction of the agents prefer a is $1 + 2\alpha - 2\alpha^2$. □

We remark that since we explicitly optimized $p(\alpha)$ in the above proof, and since any truthful mechanism must also be ordinal (Theorem 5.1) in fact we have also proved the following lower bound:

THEOREM 5.6. *No truthful mechanism for two candidates can achieve a distortion better than* $3/2$. *Further, no ordinal mechanism has distortion better than* $1 + 2\alpha - 2\alpha^2$ *on an input where* α *fraction of candidates prefer one candidate while* $(1-\alpha)$*-fraction prefer the other.*

Similar to the deterministic case, the theorem shows that not only does TEMPERED-MAJORITY have optimal expected distortion over truthful mechanisms in the worst-case, but in fact is optimal (in terms of expected distortion) for every ordinal input. That is, let μ be any mechanism with ordinal inputs. Then for every input, the expected welfare of μ in the worst-case (over utility functions consistent with the input linear order vector) is at most the expected welfare of TEMPERED-MAJORITY. Since every truthful mechanism has ordinal inputs, on any input, TEMPERED-MAJORITY has expected welfare at least that of any truthful mechanism.

Acknowledgments. We thank Swaprava Nath for introducing us to this problem, and Tom Hayes for helpful discussions. We also thank Ariel Procaccia and Nisarg Shah for helping to clarify the distortion for mechanisms with randomized inputs.

REFERENCES

Elliot Anshelevich, Onkar Bhardwaj, and John Postl. 2015. Approximating Optimal Social Choice under Metric Preferences. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.* 777–783. http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9643

Elliot Anshelevich and John Postl. 2016. Randomized Social Choice Functions under Metric Preferences. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016.* 46–59. http://www.ijcai.org/Abstract/16/014

Jose Apesteguia, Miguel A Ballester, and Rosa Ferrer. 2011. On the justice of decision rules. *The Review of Economic Studies* 78, 1 (2011), 1–16.

Yaron Azrieli and Semin Kim. 2014. Pareto efficiency and weighted majority rules. *International Economic Review* 55, 4 (2014), 1067–1088.

Wade W Badger. 1972. Political individualism, positional preferences, and optimal decision-rules. *Probability models of collective decision making* (1972), 34–59.

Gerdus Benade, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. 2017. Preference Elicitation For Participatory Budgeting. In *Association for Advancement of Artificial Intelligence (AAAI), February 4 - 9, 2017, San Francisco, California, USA*. Forthcoming.

Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. 2015. Optimal social choice functions: A utilitarian view. *Artif. Intell.* 227 (2015), 190–213. DOI: http://dx.doi.org/10.1016/j.artint.2015.06.003

Ioannis Caragiannis, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. 2017. Subset Selection Via Implicit Utilitarian Voting. *J. Artif. Intell. Res. (JAIR)* 58 (2017), 123–152. DOI: http://dx.doi.org/10.1613/jair.5282

Urszula Chajewska, Daphne Koller, and Ronald Parr. 2000. Making rational decisions using adaptive utility elicitation. In *AAAI/IAAI*. 363–369.

Richard B Curtis. 1972. Decision rules and collective values in constitutional choice. *1972) Probability Models of Collective Decision Making, Bell and Howell Company, Columbus Ohio* (1972).

Michal Feldman, Amos Fiat, and Iddan Golomb. 2016. On Voting and Facility Location. In *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, Maastricht, The Netherlands, July 24-28, 2016*. 269–286. DOI: http://dx.doi.org/10.1145/2940716.2940725

Allan Gibbard. 1973. Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society* (1973), 587–601.

Allan Gibbard. 1977. Manipulation of schemes that mix voting with chance. *Econometrica: Journal of the Econometric Society* (1977), 665–681.

Ashish Goel, Anilesh K Krishnaswamy, Sukolsak Sakshuwong, and Tanja Aitamurto. 2015. Knapsack voting. *Collective Intelligence* (2015).

Ashish Goel, Anilesh K Krishnaswamy, Sukolsak Sakshuwong, and Tanja Aitamurto. 2016. Knapsack Voting: Voting mechanisms for Participatory Budgeting. (2016).

Wassily Hoeffding. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association* 58, 301 (1963), 13–30.

Semin Kim and others. 2012. Ordinal versus cardinal voting rules: A mechanism design approach. *The Ohio State University working paper* (2012).

Robert Mclean. 2016. Americans were stuck in traffic for 8 billion hours in 2015. *CNN Money* (2016). http://money.cnn.com/2016/03/15/news/us-commutes-traffic-cars/

Ariel D Procaccia. 2010. Can approximation circumvent Gibbard-Satterthwaite?. In *AAAI*.

Ariel D Procaccia and Jeffrey S Rosenschein. 2006. The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents*. Springer, 317–331.

Douglas W Rae. 1969. Decision-rules and individual values in constitutional choice. *American Political Science Review* 63, 01 (1969), 40–56.

Mark Allen Satterthwaite. 1975. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory* 10, 2 (1975), 187–217.

Patrick W Schmitz and Thomas Tröger. 2012. The (sub-) optimality of the majority rule. *Games and Economic Behavior* 74, 2 (2012), 651–665.

Norman Schofield. 1972. Ethical decision rules for uncertain voters. *British Journal of Political Science* 2, 02 (1972), 193–207.

Peter Wakker and Daniel Deneffe. 1996. Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management science* 42, 8 (1996), 1131–1150.

APPENDIX

Proof of Theorem 3.3. Fix a candidate x and consider the following experiment: for each agent i sample $T_{i,x}$ uniformly from $[0, 1]$ and ask i whether $u_i(x) \geq T_{i,x}$. Let

$$\tau_i(x) = \begin{cases} 2T_{i,x} & \text{if } u_i(x) \geq T_{i,x} \\ 2T_{i,x} - 1 & \text{otherwise .} \end{cases}$$

We've created a new population $\mathcal{P}_x = \{\tau_i(x)\}_i$. Let $\bar{\tau}(x)$ be the population mean of $\mathcal{P}_x$. Then by Hoeffding's inequality,

$$\mathbb{P}[|\bar{\tau}(x) - \bar{U}_x| \geq \nu] \leq 2 \exp\left(-\frac{2n\nu^2}{9}\right)$$

Setting $\nu = \sqrt{\frac{9\log(4mn)}{2n}}$, we have

$$\mathbb{P}[|\bar{\tau}(x) - \bar{U}_x| \geq \nu] \leq \frac{1}{2mn}.$$

Now suppose we randomly sample k times from $\mathcal{P}_x$ without replacement. That is, we choose a random set A_x of size k, and observe the values $\{\tau_i(x)\}_{i\in A_x}$. Let $\bar{V}_x = \frac{1}{k}\sum_{i\in A_x}\tau_i(x)$ be the sample mean. Let $\bar{U}_x = \mathrm{uw}(x)/n$. Then by Hoeffding's inequality (again!) and for the same ν

$$\begin{aligned}\mathbb{P}[|\bar{V}_x - \bar{\tau}(x)| \geq \nu\sqrt{n/k}] &\leq 2\exp\left(\frac{-2k(\nu\sqrt{n/k})^2}{9}\right)\\ &\leq 2\exp\left(-\frac{2}{9}n\nu^2\right)\\ &\leq \frac{1}{2mn}\end{aligned}$$

We note that Hoeffding's inequality holds, and has been used here for non-independent random variables obtained by sampling without replacement. See Section 6 of Hoeffding's paper for details [Hoeffding, 1963].

It follows from the union bound and the triangle inequality that with probability at least $1 - \frac{1}{mn}$

$$|\bar{V}_x - \bar{U}_x| \leq \nu(1 + \sqrt{n/k})$$

Setting $k = \lfloor n/m \rfloor$ (which gives a larger bound than $k = \lceil n/m \rceil$) we have with probability at least $1 - \frac{1}{mn}$

$$|\bar{V}_x - \bar{U}_x| \leq 2\nu\sqrt{m} \leq \frac{\delta}{2m^2} \tag{5}$$

Taking a union bound over all candidates, we see that with probability at least $1 - 1/n$ (5) holds simultaneously for all x, which is analogous to (1). Thus if $S = \arg\max_{T\subseteq C:\sum_{x\in T}c_x\leq 1}\{\sum_{x\in T}\bar{V}_x\}$ and S^* is the optimal set of candidates, following along the same lines as the proof of Theorem 3.1, we get that

$$\mathrm{uw}(S) \geq (1-\delta)\mathbf{OPT}.$$

All that remains is to argue that Mechanism 2 actually simulates our experiment in which we first created the m populations $\{\mathcal{P}_x\}_x$ and then sampled from them. However note that for a fixed candidate x the only values that we need from $\mathcal{P}_x$ are those corresponding to the agents in the random set A_x, so by the principal of deferred decisions, we do not need to sample $T_{i,x}$ except for $i \in A_x$. Also, since the random sets A_x in the experiment are not required to be independent (since we are only doing a union bound over them), we can choose them by a random equipartition as in the mechanism. Finally, since this means that each i is in exactly one $A_x = A_{x(i)}$, we are only sampling one threshold $T_i = T_{i,x(i)}$. Thus Mechanism 2 simulates our experiment, completing the proof. □

Proof of Claim **??**. This can be seen by induction on S_j, for $j \in [m]$. Let x, y be the candidates ranked r and $r+1$ by i. In the base case, when $j = 1$, $\mathbb{P}[x \in Z_j] \propto 1/r$, $\mathbb{P}[y \in Z_j] \propto 1/(r+1)$ (we implicitly assume conditioning on i being chosen in Step 2), and the claim is true. In the inductive step,

$$\begin{aligned}\mathbb{P}[x \in Z_{j+1}] &= \mathbb{P}[x \in Z_j] + \mathbb{P}[x \in Z_{j+1}|x, y \notin Z_j]\mathbb{P}[x, y \notin Z_j]\\ &\quad + \mathbb{P}[x \in Z_{j+1}|x \notin Z_j, y \in Z_j]\mathbb{P}[x \notin Z_j, y \in Z_j]\,.\end{aligned} \tag{6}$$

We note the following inequalities:

$$\mathbb{P}[x \in Z_j] \geq \mathbb{P}[y \in Z_j] \qquad \text{by induction} \tag{7}$$

$$\mathbb{P}[x \in Z_{j+1} | x, y \notin Z_j] \geq \mathbb{P}[y \in Z_{j+1} | x, y \notin Z_j] \qquad \text{since } x \text{ has smaller rank than } y \tag{8}$$

$$\mathbb{P}[x \in Z_{j+1} | x \notin Z_j, y \in Z_j] = \mathbb{P}[y \in Z_{j+1} | x \notin Z_j, y \in Z_j] \qquad \text{by symmetry} \tag{9}$$

The proof then follows by applying the above inequalities and some manipulation. Firstly, applying inequality (8) to the second summand of (6) we obtain:

$$\mathbb{P}[x \in Z_{j+1} | x, y \notin Z_j]\mathbb{P}[x, y \notin Z_j] \geq \mathbb{P}[y \in Z_{j+1} | x, y \notin Z_j]\mathbb{P}[x, y \notin Z_j] \tag{10}$$

Secondly, applying inequalities (7), (9) to the first and third summands of (6), we get:

$$\begin{aligned}
&\mathbb{P}[x \in Z_j] + \mathbb{P}[x \in Z_{j+1} | x \notin Z_j, y \in Z_j]\mathbb{P}[x \notin Z_j, y \in Z_j] \\
&= \mathbb{P}[x \in Z_j] + \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j]\mathbb{P}[x \notin Z_j, y \in Z_j] \\
&= \mathbb{P}[x \in Z_j] + \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j] \left(1 - \mathbb{P}[x \in Z_j] - \mathbb{P}[x, y \notin Z_j]\right) \\
&= \mathbb{P}[x \in Z_j] \left(1 - \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j]\right) + \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j] \left(1 - \mathbb{P}[x, y \notin Z_j]\right) \\
&\geq \mathbb{P}[y \in Z_j] \left(1 - \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j]\right) + \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j] \left(1 - \mathbb{P}[x, y \notin Z_j]\right) \\
&= \mathbb{P}[y \in Z_j] + \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j] \left(1 - \mathbb{P}[y \in Z_j] - \mathbb{P}[x, y \notin Z_j]\right) \\
&= \mathbb{P}[y \in Z_j] + \mathbb{P}[y \in Z_{j+1} | y \notin Z_j, x \in Z_j]\mathbb{P}[y \notin Z_j, x \in Z_j] \,.
\end{aligned} \tag{11}$$

Above, the first equality follows from (9), and the first inequality follows from (7). Finally, adding equations (10) and (11), and observing that by (6) the right hand side thus obtained is $\mathbb{P}[y \in Z_{j+1}]$, we get the required result, completing the induction. □

Proof of Theorem 5.2. The truthfulness of MAJORITY follows simply because if the number of agents that prefer candidate a increases, so does the likelihood that a is selected. Hence a candidate that votes a over b must have greater utility for a as well.

For the upper bound on distortion, for a fixed utility profile $\vec{u}$ w.l.o.g. let a be the candidate chosen by MAJORITY, and b be optimal candidate. Let **ALG** $:= \sum_i u_i(a)$, and **OPT** $:= \sum_i u_i(b)$. Let S be the set of agents that voted for candidate a, then $|S| \geq n/2$. We obtain the following bounds on **OPT** and **ALG**:

$$\begin{aligned}
\mathbf{OPT} = \sum_i u_i(b) &= n - \sum_i u_i(a) \leq n - \sum_{i \in S} u_i(a) \\
&\leq n - |S|/2 \leq 3n/4
\end{aligned}$$

where the second inequality follows because, assuming truthfulness, if $i \in S$ then $u_i(a) \geq 1/2$. This also completes the proof, since

$$\mathbf{ALG} = \sum_i u_i(a) \geq \sum_{i \in S} u_i(a) \geq n/4 \,.$$

For the lower bound, we consider just two agents, whom we call A and B. Suppose A votes $a \prec_A b$, and B votes $b \prec_B a$. By Theorem 5.1, this is all the input to any truthful mechanism. Suppose the mechanism selects A. Then for utilities $u_A(a) = 1/2 + \epsilon$, $u_B(a) = 0$, the optimal welfare is $3/2 - \epsilon$, while the mechanism gets welfare $1/2 + \epsilon$. □