Carnegie Mellon University
Tepper School of Business

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
INDUSTRIAL ADMINISTRATION
(OPERATIONS RESEARCH)

Titled

"Equity and Efficiency in Computational Social Choice"

Presented by

Gerdus Benadè

Accepted by

John Hooker	**5/1/2019**
Co-Chair: Prof. John Hooker	**Date**
Ariel Procaccia	**5/1/2019**
Co-Chair: Prof. Ariel Procaccia	**Date**

Approved by The Dean

Robert M. Dammon	**5/6/2019**
Dean Robert M. Dammon	**Date**

Equity and efficiency in computational social choice

Gerdus Benadè

May 16, 2019

Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Ariel D. Procaccia
John Hooker
R. Ravi
Jay Sethuraman

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Algorithms, Combinatorics, and Optimization.

.

Abstract

This thesis studies problems in *computational social choice* and *fair division*.

Computational social choice asks how to aggregate individual votes and opinions into a joint decision. *Participatory budgeting* enables the allocation of public funds by collecting and aggregating individual preferences over proposed infrastructure projects; it has already had a sizable real-world impact. We analytically compare four preference elicitation methods through the lens of *implicit utilitarian voting*, and find that threshold approval votes are qualitatively superior. This conclusion is supported by experiments using data from real participatory budgeting elections. We also conduct a human subject experiment on Amazon Mechanical Turk to study the cognitive burden that different elicitation formats impose on voters.

Under implicit utilitarian voting we attempt to maximize a utilitarian objective in the presence of uncertainty about voter utility functions. Next, we take a very different approach, and assume votes are noisy estimates of an unknown ground truth. We build on previous work which replaced structural assumptions on the noise with a worst-case approach, and minimize the expected error with respect to a set of feasibly true rankings. We derive mostly sharp analytical bounds on the expected error and find that our approach has useful practical properties.

Fair division problems involve allocating goods to heterogeneous agents. Motivated by the problem of a food bank allocating donations to their beneficiaries without knowledge of future arrivals, we study the online allocation of indivisible items. Our goal is to design allocation algorithms that minimize the maximum *envy*, defined as the maximum difference between any agent's overall value for items allocated to another agent and to herself. An algorithm has *vanishing envy* if the ratio of envy over time goes to zero as time goes to infinity. We find a polynomial-time, deterministic algorithm that achieves vanishing envy, and show the rate at which envy vanishes is asymptotically optimal.

Finally, we consider the problem of gerrymandering. We start with an impartial protocol and derive a notion of fairness which provides guidance about what to expect from an impartial districting. Specifically, we propose that a party should win a number of districts equal to the midpoint between what they win in their best and worst districtings. We show that this notion of fairness has close ties to proportionality yet, in contrast to proportionality, there always exists a districting satisfying our notion of fairness.

Acknowledgments

I am deeply grateful to my advisor, Ariel Procaccia, for his support and guidance throughout this PhD. Working with Ariel is a pleasure and an inspiration, and I've learned a lot from his approach to both research and life.

I have been fortunate to work with many other great people and I am thankful to all of them. My discussions with John Hooker spanned a wide set of topics and were always thought-provoking. Working with R. Ravi was a privilege. I've enjoyed collaborating with Kobi Gal, Alex Psomas, Aleks Kazachkov, Nam Ho-Nguyen, Paul Gölz, Anson Kahng and Mingda Qioa, and have learned something from every one of them. Before coming to CMU, I was advised by Jan van Vuuren, whose guidance set me on this path.

This journey would have been a lot less bearable without the support of the Tepper community. Thanks to Lawrence and Laila, without whom I have no doubt chaos would reign. Thank you Aleks, Stelios, Christian, Siddharth, Thiago, Yang, Amin, Arash, Mehmet and Thomas for the laughs and late night research discussions. Michael, Nam, Dabeen, Ryo and Dennis deserve special mention, we faced the crucible together.

None of this would have been possible without my family. My parents instilled in me a curiosity and love for learning that has guided me throughout my life, no words can do justice to the debt I owe them. Finally, I must thank my wife Mariet. Thank you for your unwavering support and for allowing me to chase my dreams. I am who and where I am because of you.

Contents

Chapter 1

Introduction

Over the past decade we have seen many traditional problems in economics have a significant real world impact. Game theoretic equilibria are used by the US Coast Guard to protect important ports, and by national parks to protect against poachers. Matching algorithms are used to allocate organs to transplant patients and children to schools. Auction theory underlies the massive ad auction industry. In this dissertation I will focus on applications of two other areas in economics: *computational social choice* and *dynamic fair division*.

As algorithms start affecting more parts of society, there is a growing awareness that making responsible decisions in dynamic and complex environments requires novel frameworks that balance notions of equality and fairness with efficiency. For example, a growing part in the machine learning community works on how to eliminate forms of bias from machine learning algorithms. In the problems we study fairness is an explicit objective, we construct algorithms to allocate goods or divide resources subject to specific notions of fairness. As an example, imagine the problem faced by a food bank who receives infrequent donations and must allocate those donations to heterogeneous beneficiaries while having incomplete information about future arrivals. An objective for the food bank is to treat all its beneficiaries equally. Another problem we study is political redistricting, or how to partition a state into districts so that each party wins their fair number of districts.

Our social structures of democracy and collaborative decision making also give rise to interesting questions about how to design interactions to elicit truthful and informative opinions from participants. The field of computational social choice studies how to aggregate these opinions or preferences into outcomes or decisions. A compelling modern application of social choice is participatory budgeting in which citizens vote over proposed infrastructure projects and these votes are aggregated to form a budget. We ask how this setting affects the voting format (what a vote looks like) and the voting rule (what we do with the votes).

Several of these problems are of a political nature. Computational social choice provides many of the tools we need to adapt our structures of governance to a world where we are more connected and (hopefully) better informed than every before, and where casting a vote can happen with the click of a button.

1.1 Outline of work

As mentioned, this dissertation consists of two main sections, dealing with problems in computational social choice and fair division respectively.

Computational social choice Social choice theory studies how to aggregate individual opinions and preferences into collective decisions. Modern social choice theory was kick-started by Arrow's celebrated impossibility result [10]. In this approach properties are identified that good voting rules must exhibit (axioms), and it is asked whether voting rules exist that simultaneously satisfy subsets of these axioms. Research in computational social choice leverages tools like approximation algorithms and complexity theory to shed new light on traditional social choice problems, often by avoiding the axiomatic approach in favor of assumptions about distance functions, voter utilities or statistical noise models.

One stream of research, labeled *implicit utilitarian voting* [44] assumes that the vote a voter casts is consistent with his utility function and asks to what extent the social welfare maximizing outcome can be approximated using only these proxies (the votes) instead of the voters' actual underlying utility functions.

In chapter 2, we study perhaps the most exciting application of computational social choice, participatory budgeting, through this lens of implicit utilitarian voting. In the participatory budgeting framework a city decides how to spend its budget after allowing the residents of the city to vote over a set of alternatives. Each alternative has a cost and the objective is to maximize social welfare subject to a budget constraint. We make an assumption that voter utilities are additive and analytically compare four preference elicitation methods — knapsack votes, rankings by value or value for money, and threshold approval votes — and find that threshold approval votes are qualitatively superior. This conclusion is supported by computational experiments using data from real world-participatory budgeting instances. This problem is challenging because asking voters to report their exact utility functions would be too burdensome, so instead we have to make do with easy-to-cast proxies. We study the cognitive burden associated with each input format through an human subject experiment conducted on Amazon Mechanical Turk. In chapter 3, we extend our results on rankings by value to a more general setting with subadditive utility functions where, instead of returning a set of alternatives satisfying the budget, the task is to return a ranking of the alternatives. The work in these chapters appeared in [20, 21, 24] and is joint work with Ariel D. Procaccia, Nisarg Shah, Swaprava Nath, Mingda Qiao and Ya'akov Gal.

A second stream of research in computational social choice views votes as estimators of some objective ground truth under some noisy process. Instead of maximizing social welfare as before, the objective is to recover this unknown ground truth. This is often done by making structural assumptions about the noise model which allows you return a maximum likelihood estimator.

In chapter 4, we tackle this problem while avoiding assumptions about noise models, symmetric noise and large sample sizes which are common in the literature. Following the worst-case approach in Procaccia et al. [109], our only assumption is that the average voter is at bounded distance from the ground truth under some distance metric. This assumption leads to a space of feasible solutions, each of which has the potential to be the ground truth. We deviate from previous work by minimizing the *average* error with respect to the set of feasible ground truth rankings instead of the worst-case error. We derive (mostly sharp) analytical bounds on the expected error and establish the practical benefits of our approach through experiments. This chapter is based on [19] and is joint work with Anson Kahng and Ariel D. Procaccia.

Fair division In fair division problems, a set of agents must be assigned a set of divisible or indivisible goods. In contrast to the traditional assignment problems, the objective is not some measure of cost or profit, instead, the aim is to ensure that the allocation is *fair* with respect to the agents' heterogenous preferences, under a suitable notion of fairness. The classical example of fair division is the problem of dividing a cake (a continuous good) between different agents. A popular notion of fairness is *envy-freeness*, where one agent is said to envy another if they value the bundle of goods that the other agent received higher than their own bundle. In problems with indivisible goods, envy-free allocations need not exist (imagine a single indivisible good and two agents who have positive value for that good) so we must make do with minimizing envy.

In chapter 5, motivated by the application of distributing donations received at food banks, we study envy in a dynamic setting where indivisible goods arrive over time. Our goal is to design allocation algorithms that minimize the maximum *envy*, defined as the maximum difference between any agent's overall value for items allocated to another agent and to herself. We say that an algorithm has *vanishing envy* if the ratio of envy over time goes to zero as time goes to infinity. We design a polynomial-time, deterministic algorithm that achieves vanishing envy, and show the rate at which envy vanishes is asymptotically optimal. We also derive tight (in the number of items) bounds for a more general setting where items arrive in batches. This chapter is based on [22] and is joint with Aleksandr Kazachkov, Alexandros Psomas and Ariel D. Procaccia.

Finally, in chapter 6, we study the problem of political redistricting from the perspective of fair division. Political redistricting involves partitioning a state into districts, each of which elects a representative to Congress. Despite the existence of guidelines specifying, for example, that congressional districts must be contiguous and contain the same number of people, the redistricting process is controversial and fraught with opportunity for abuse. Federal courts have recently redrawn districtings found to be unconstitutional, and many different metrics have been proposed for evaluating whether a districting is partisan.

We come up with our own notion of fairness in the context of redistricting which assigns every party a target number of districts derived from an impartial protocol. This target reduces to proportionality in the absence of geographic constraints, but is more sensitive to the distribution of voters in a state. We find that, when you model voters as points on plane, it is always possible to find a districting in which every party wins their target number of districts. A case study on Pennsylvania show some of the implications of this result. We conclude with a novel exact model of the redistricting problem based on a modeling the contiguity constraints recursively. The work in this chapter is joint with Ariel D. Procaccia, John Hooker and Margot Stewart.

Excluded work For the sake of brevity and coherence not all of the author's research appears in this dissertation. The excluded work includes work on deriving bounds from the branching dual of a discrete optimization problem [18] and stratification in the context of sortition [23].

Chapter 2

Participatory budgeting

2.1 Introduction

A central societal question is how to consolidate diverse preferences and opinions into reasonable, collective decisions. Classical voting theory takes an axiomatic approach which identifies desirable properties that the aggregation method should satisfy, and studies the (non-)existence and structure of such rules. A celebrated example of this is Arrow's impossibility result [10]. By contrast, the field of *computational social choice* [35] typically attempts to identify an appealing objective function and design aggregation rules to optimize this objective.

One of the best-studied problems in computational social choice deals with aggregating individual preferences over alternatives — expressed as *rankings* — into a collective choice of a *subset* of alternatives [44, 108, 119]. Nascent social choice applications, though, have given rise to the harder, richer problem of *budgeted social choice* [91], where alternatives have associated costs, and the selected subset is subject to a budget constraint.

Our interest in budgeted social choice stems from the striking real-world impact of the *participatory budgeting* paradigm [39], which allows local governments to allocate public funds by eliciting and aggregating the preferences of residents over potential projects. Indeed, in just a few years, the *Participatory Budgeting Project*[1] has helped allocate more than $300 million dollars to more than 1 600 local projects, primarily in the US and Canada (including New York City, Chicago, Boston, and San Francisco).

[1]See `http://www.participatorybudgeting.org`.

Paraticipatory budgeting has also attracted attention globally. A 2007 study by the World Bank [116] reports instances of participatory budgeting in locations as diverse as Guatemala, Peru, Romania and South Africa. In Europe, the push for participatory budgeting is arguably led by Madrid and Paris: Madrid spent €24 million through participatory budgets in 2016 and Paris €100 million[78, 87]. Notably, a participatory budgeting application is also included in the *Decide Madrid* open-source tool for civic engagement, providing a framework to simplify the hosting and management of participatory budgeting elections around the world.

In the first formal analysis of this paradigm, Goel et al. [2016] — who have facilitated several participatory budgeting elections as part of the Stanford Crowdsourced Democracy Team[2] — propose and evaluate two participatory budgeting approaches. In the first approach, the *input format* — the way in which each voter's preferences are elicited — is *knapsack votes*: Each voter reports his individual solution to the knapsack problem, that is, the set of projects that maximizes his overall value (assuming an additive valuation function), subject to the budget constraint. The second component of the approach is the *aggregation rule*; in this case, each voter is seen as approving all the projects in his knapsack, and then projects are ordered by the number of approval votes and greedily selected for execution, until the budget runs out. The second approach uses *value-for-money comparisons* as input format — it asks voters to compare pairs of projects by the ratio between value and cost. These comparisons are aggregated using variants of classic voting rules, including the Borda count rule and the Kemeny rule.

In a sense, Goel et al. [2016] take a bottom-up approach: They define novel, intuitive input formats that encourage voters to take cost — not just value — into account, and justify them after the fact. By contrast, we wish to take a top-down approach, by specifying an overarching optimization goal, and using it to compare different methods for participatory budgeting.

2.1.1 Our Approach and Results

Following Goel et al. [2016], we assume that voters have additive utility functions and vote over a set of alternatives, each with a known cost. Our goal is to choose a subset of alternatives which maximize *(utilitarian) social welfare* subject to a budget constraint.

[2]See `http://voxpopuli.stanford.edu`.

This reduces to a knapsack problem when we have access to the utility functions; the problem is challenging precisely because we do not. Rather, we have access to votes, in a certain input format, which are *consistent* with the utility functions. This goal — maximizing social welfare based on votes that serve as proxies for latent utility functions — has been studied for more than a decade [6, 8, 9, 31, 41, 106]; it has recently been termed *implicit utilitarian voting* [44].

Absent complete information about the utility functions, clearly social welfare cannot be perfectly maximized. Procaccia and Rosenschein [2006] introduced the notion of *distortion* to quantify how far a given aggregation rule is from achieving this goal. Roughly speaking, given a vote profile (a set of n votes) and an outcome, the distortion is the worst-case ratio between the social welfare of the optimal outcome, and the social welfare of the given outcome, where the worst case is taken with respect to all utility profiles that are consistent with the given votes.

Previous work on implicit utilitarian voting assumes that each voter expresses his preferences by ranking the alternatives in order of decreasing utility. By contrast, the main insight underlying our work is that

> *... the implicit utilitarian voting framework allows us to decouple the input format and aggregation rule, thereby enabling an analytical comparison of different input formats in terms of their potential for providing good solutions to the participatory budgeting problem.*

This decoupling is achieved by associating each input format with the distortion of the *optimal* (randomized) aggregation rule, that is, the rule that minimizes distortion on every vote profile. Intuitively, the distortion associated with an input format measures how useful the information contained in the votes is for achieving social welfare maximization (lower distortion is better).

In Section 2.3, we apply this approach to compare four input formats. The first is knapsack votes, which (disappointingly) has distortion *linear* in the number of alternatives, the same distortion that one can achieve in the complete absence of information. Next, we analyze two closely related input formats: rankings by value, and rankings by value for money, which ask voters to rank the alternatives by their value and by the ratio of their value and cost, respectively. We find that for both of these input formats the distortion grows no faster than the square root of the number of alternatives, which matches a lower bound up to logarithmic factors. Finally, we examine a novel input format, which we call *threshold approval votes*: each voter is asked to approve each alternative whose value for him is above a threshold that we choose. We find tight bounds showing that the distortion of threshold approval votes is essentially *logarithmic* in the number of items. To summarize, our theoretical results show striking separations between different input formats, with threshold approval votes coming out well on top.

It is worth noting that these results may also be interpreted as approximation ratios to the optimal solution of the classical knapsack problem, where we are given only partial information about voter utilities (a vote profile, in some format) and an adversary selects both the vote profile and a utility profile consistent with the votes, which is used to evaluate our performance.

While our theoretical results in Section 2.3 bound the distortion, i.e., the *worst-case ratio* of the optimal social welfare to the social welfare achieved over all instances, it may be possible to provide much stronger performance guarantees on any specific instance. In Section 2.4, we design algorithms to compute the distortion-minimizing subset of alternatives (when considering deterministic aggregation rules), and distribution over subsets of alternatives (when considering randomized aggregation rules) for a *specific* instance. We observe that the running times of these distortion-minimizing rules scale gracefully to practical sizes.

In Section 2.5 we use these algorithms to compare different approaches to participatory budgeting using the *average-case ratio* of the optimal social welfare, and the social welfare achieved by our aggregation rules. Specifically, we experimentally evaluate approaches that use the input formats we study in conjunction with their respective optimal aggregation rules, which minimize the distortion on each profile, and compare them to two approaches currently employed in practice. (Note that these rules are not guaranteed to achieve the optimal performance in our experiments as we measure performance using the *average-case* ratio of the optimal to the achieved social welfare rather than the (worst-case) distortion. Nonetheless, such rules perform extremely well.) We use data from two real-world participatory budgeting elections held in Boston in 2015 and 2016. The experiments indicate that the use of aggregation rules that minimize distortion on every input profile significantly outperforms the currently deployed approaches, and among the input formats we study, threshold approval votes remain superior, even in practice.

However, if efficiency was our only concern we would simply elicit voters' full preferences. Of course, this is not possible and place a high cognitive burden on voters [40]. The second dimension on which our input formats must be evaluated is *usability* – how easy to learn, understand and use is an input format. In Section 2.6 we conduct a user study through Amazon Mechanical Turk to compare different input formats along this axis.

2.1.2 Related Work

Let us first describe the theoretical results of Goel et al. [2016] in slightly greater detail. Most relevant to our work is a theorem that asserts that knapsack voting (i.e., knapsack votes as the input format, coupled with greedy approval-based aggregation) actually maximizes social welfare. However, the result strongly relies on their *overlap utility model*, where the utility of a voter for a subset of alternatives is (roughly speaking) the size of the intersection between this subset and his own knapsack vote. In a sense, the viewpoint underlying this model is the opposite of ours, as a voter's utility is derived from his vote, instead of the other way around. One criticism of this model is that even if certain alternatives do not fit into a voter's individual knapsack solution due to the budget constraint, the voter could (and usually will) have some utility for them. Goel et al. [2016] also provide strategyproofness results for knapsack voting, which similarly rely on the overlap utility model. Finally, they interpret their methods as maximum likelihood estimators [54, 135] under certain noise models. In addition to these theoretic results, Goel et al. also perform an empirical analysis of voter behaviour. One experiment provides timing data for knapsack votes, k-approval votes and pairwise value-for-money comparisons, where it is noted that "the knapsack interface is not much more time consuming than the k-approval interface." Our user study performs similar and perhaps more extensive experiments.

As our work applies the implicit utilitarian voting approach [31, 44] to a problem in the budgeted social choice framework [91], it is naturally related to both lines of work. Lu and Boutilier [91] introduce the budgeted social choice framework, in which the goal is to collectively select a set of alternatives subject to a budget constraint. Their framework generalizes the participatory budgeting problem studied herein as it allows the cost of an alternative to also depend on the number of voters who derive utility from the alternative. However, their results are incomparable to ours because they assume that every voter's utility for an alternative is determined solely by the rank of the alternative in the voter's preference order — specifically, that the utilities of all voters follow a common underlying *positional scoring* *rule* — which is a common assumption in the literature on resource allocation [16, 32]. This makes the elicitation problem trivial because eliciting ordinal preferences (i.e., rankings by value) is assumed to accurately reveal the underlying cardinal utilities. By contrast, we do not impose such a restriction on the utilities, and compare the rankings-by-value input format with three other input formats.

Previous work on implicit utilitarian voting focuses exclusively on the rankings-by-value input format. Boutilier et al. [2015] study the problem of selecting a single winning alternative, and provide an upper and lower bounds on the distortion achieved by the optimal aggregation rule. Their setting is a special case of the participatory budgeting problem where the cost of each alternative equals the entire budget. Consequently, their lower bound applies to our more general setting, and our upper bound for the rankings-by-value input format generalizes theirs (up to a logarithmic factor). Caragiannis et al. [2016] extend the results of Boutilier et al. [2015] to the case where a subset of alternatives of a given size k is to be selected (only for the rankings-by-value input format); this is again a special case of the participatory budgeting problem where the cost of each alternative is B/k. However, our results are incomparable to theirs because we assume additive utility functions — following previous work on participatory budgeting [75] — whereas Caragiannis et al. assume that a voter's utility for a subset of alternatives is his *maximum* utility for any alternative in the subset.

The core idea behind implicit utilitarian voting — approximating utilitarian social welfare given ordinal information — has also been studied in mechanism design. Filos-Ratsikas et al. [69] present algorithms for finding matchings in weighted graphs given ordinal comparisons among the edges by their weight; Krysta et al. [85] apply this notion to the house allocation problem; and, Chakrabarty and Swamy [47] study this notion in a general mechanism design setting, but with the restriction borrowed from Lu and Boutilier [91] that the utilities of all agents are determined by a common positional scoring rule.

A line of research on resource allocation focuses on maximizing other forms of welfare such as the egalitarian welfare or the Nash welfare [see, e.g., 97]. Maximizing the Nash welfare has the benefit that it is invariant to scaling an agent's utility function, and thus does not require normalizing the utilities. In addition, it is known to satisfy non-trivial fairness guarantees in domains that are similar to or generalize participatory budgeting [57, 66]. It remains to be seen whether maximizing the Nash welfare subject to votes that only partially reveal the underlying utilities can preserve such guarantees.

2.2 The Model

Let $[k] \triangleq \{1, \ldots, k\}$ denote the set of k smallest positive integers. Let $N \triangleq [n]$ be the set of *voters*, and A be the set of m *alternatives*. The *cost* of alternative a is denoted c_a, and the budget B is normalized to 1. For $S \subseteq A$, let $c(S) \triangleq \sum_{a \in S} c_a$. Define $\mathcal{F}_c \triangleq \{S \subseteq A : c(S) \leq 1 \ \wedge \ c(T) > 1, \ \forall S \subsetneq T \subseteq A\}$ as the inclusion-maximal budget-feasible subsets of A.

We assume that each voter has a *utility function* $v_i : A \to \mathbb{R}_+ \cup \{0\}$, where $v_i(a)$ is the utility that voter i has for alternative a, and that these utilities are additive, i.e., the utility of voter i for a set $S \subseteq A$ is defined as $v_i(S) = \sum_{a \in S} v_i(a)$. Finally, to ensure fairness among voters, we make the standard assumption [41, 31] that $v_i(A) = 1$ for all voters $i \in N$. We call the vector $\vec{v} = \{v_1, \ldots, v_n\}$ of voter utility functions the *utility profile*. Given the utility profile, the (utilitarian) *social welfare* of an alternative $a \in A$ is defined as $\mathrm{sw}(a, \vec{v}) \triangleq \sum_{i \in N} v_i(a)$; for a set $S \subseteq A$, let $\mathrm{sw}(S, \vec{v}) \triangleq \sum_{a \in S} \mathrm{sw}(a, \vec{v})$.

The utility function of a voter i is only accessible through his vote ρ_i, which is induced by v_i. The vector $\vec{\rho} \triangleq \{\rho_1, \ldots, \rho_n\}$ is called the *input profile*. Let $\vec{v} \vartriangleright \vec{\rho}$ denote that utility profile $\vec{v}$ is *consistent* with input profile $\vec{\rho}$. We study four specific formats for input votes. Below, we describe each input format along with a sample question that may be asked to the voters to elicit votes in that format. The voters can be induced to think of their utilities for the different alternatives (i.e., projects) in a normalized fashion by asking them to (mentally) divide a constant sum of points — say, 1000 points — among the alternatives based on how much they like each alternative.

- The *knapsack vote* $\kappa_i \subseteq A$ of voter $i \in N$ represents a feasible subset of alternatives with the highest value for the voter. We have $v_i \vartriangleright \kappa_i$ if and only if $c(\kappa_i) \leq 1$ and $v_i(\kappa_i) \geq v_i(S)$ for all $S \in \mathcal{F}_c$. If the total budget is \$100,000, the voters may be asked: "Select the best set of projects according to you subject to a total budget of \$100,000."

- The *rankings-by-value* and the *rankings-by-value-for-money* input formats ask voter $i \in N$ to rank the alternatives by decreasing value for him, and by decreasing ratio of value for him to cost, respectively. Formally, let $\mathcal{L} \triangleq \mathcal{L}(A)$ denote the set of rankings over the alternatives.

For a ranking $\sigma \in \mathcal{L}$, let $\sigma(a)$ denote the position of alternative a in σ, and $a \succ_\sigma b$ denote $\sigma(a) < \sigma(b)$, i.e., that a is preferred to b under σ. Then, we say that utility function v_i is consistent with the ranking by value (resp. value for money) of voter $i \in N$, denoted σ_i, if and only if $v_i(a) \geq v_i(b)$ (resp. $v_i(a)/c_a \geq v_i(b)/c_b$) for all $a \succ_{\sigma_i} b$. To elicit such votes, the voters may be asked: "If you had to divide 1000 points among the projects based on how much you like them, rank the projects in the decreasing order of the number of points they would receive (divided by the cost)."

- For a threshold t, the *threshold approval vote* τ_i of voter $i \in N$ consists of the set of alternatives whose value for him is at least t, i.e., $v_i \triangleright \tau_i$ if and only if $\tau_i = \{a \in A : v_i(a) \geq t\}$. To elicit threshold approval votes with a threshold $t = 1/10$, the voters may be asked: "If you had to divide 1000 points among the projects based on how much you like them, select all the projects that would receive at least 100 points."

For the purposes of our user study in Section 2.6, we will also consider k-*approval votes*, which are the most widely used input format in practice, for example, 4-approval votes were used in Boston, MA, in 2015, and 5-approval votes in Greensboro, NC, in 2016.

- A k-*approval* vote of voter i is a binary vector $\alpha_i \in \{0,1\}^m$ with $\sum_{a \in A} \alpha_i(a) \leq k$. This represents the voter's k most preferred alternatives. We say that α_i is consistent with utility function v_i if, for all $a, a' \in A$, $\alpha_i(a) > \alpha_i(a')$ implies $v_i(a) \geq v_i(a')$.

In our setting, a (randomized) *aggregation rule* f for an input format maps each input profile $\vec{\rho}$ in that format to a distribution over $\mathcal{F}_c$. The rule is *deterministic* if it returns a particular set in $\mathcal{F}_c$ with probability 1.

In the implicit utilitarianism framework, the ultimate goal is to maximize the (utilitarian) social welfare. Procaccia and Rosenschein [2006] use the notion of *distortion* to quantify how far an aggregation rule f is from achieving this goal. The distortion of f on a vote profile $\vec{\rho}$ is given by

$$\operatorname{dist}(f, \vec{\rho}) \triangleq \sup_{\vec{v} : \vec{v} \triangleright \vec{\rho}} \frac{\max_{T \in \mathcal{F}_c} \operatorname{sw}(T, \vec{v})}{\mathbb{E}[\operatorname{sw}(f(\vec{\rho}), \vec{v})]}.$$

The (overall) distortion of a rule f is given by $\operatorname{dist}(f) \triangleq \max_{\vec{\rho}} \operatorname{dist}(f, \vec{\rho})$. The optimal (randomized) aggregation rule f^*, which we term the *distortion-minimizing aggregation rule*, selects the distribution minimizing distortion on each input profile individually, that is,

$$f^*(\vec{\rho}) \triangleq \arg\min_{p \in \Delta(\mathcal{F}_c)} \sup_{\vec{v}:\vec{v} \triangleright \vec{\rho}} \frac{\max_{T \in \mathcal{F}_c} \mathrm{sw}(T, \vec{v})}{\mathbb{E}[\mathrm{sw}(p, \vec{v})]},$$

where $\Delta(\mathcal{F}_c)$ is the set of distributions over $\mathcal{F}_c$. Needless to say, f^* achieves the best possible overall distortion. Similarly, the *deterministic* distortion-minimizing aggregation rule $f^*_{\mathtt{det}}$ is given by

$$f^*_{\mathtt{det}}(\vec{\rho}) \triangleq \arg\min_{S \in \mathcal{F}_c} \sup_{\vec{v}:\vec{v} \triangleright \vec{\rho}} \frac{\max_{T \in \mathcal{F}_c} \mathrm{sw}(T, \vec{v})}{\mathrm{sw}(S, \vec{v})}.$$

Finally, we say that the distortion associated with an input format (i.e., elicitation method) is the overall distortion of the (randomized) distortion-minimizing aggregation rule for that format; this, in a sense, quantifies the effectiveness of the input format in achieving social welfare maximization. In a setting where deterministic rules must be used, we say that the distortion associated with deterministic aggregation of votes in an input format is the overall distortion of the deterministic distortion-minimizing aggregation rule for that format. Observe that we always mention deterministic aggregation explicitly, and the "distortion associated with an input format" allows randomized aggregation by default.

2.3 Theoretical Results

In Section 2.3.1, we present theoretical results for the distortion associated with different input formats when no constraints are imposed on the aggregation rule, i.e., when randomized aggregation rules are allowed. Subsequently, in Section 2.3.2, we study the distortion associated with deterministic aggregation under these input formats.

2.3.1 Randomized Aggregation Rules

We begin by making a simple observation that holds for (randomized) aggregation of votes in any input format.

Observation 2.1. *The distortion of any input format is at most m.*

Proof of Observation 2.1. Consider the rule that selects a single alternative uniformly at random; this is clearly budget-feasible. Due to the normalization of utility functions, the expected welfare achieved by this rule is $(1/m) \cdot \sum_{i \in N} \sum_{a \in A} v_i(a) = n/m$. On the other hand, the maximum welfare that any subset of alternatives can achieve is at most n. Hence, the distortion of this rule, which does not require any input, is at most m. $\square$

Knapsack Votes.

We now present our analysis for knapsack votes — an input format advocated by Goel et al. [2016].

Theorem 2.2. *For $n \geq m$, the distortion of knapsack votes is $\Omega(m)$.*

Proof of Theorem 2.2. Consider the case where every alternative has cost 1 (i.e., equal to the budget). Consider the input profile $\vec{\kappa}$, in which voters are partitioned into m subsets $\{N_a\}_{a \in A}$ of roughly equal size; specifically, let $n_a = |N_a|$ and enforce $\lfloor n/m \rfloor \leq n_a \leq \lceil n/m \rceil$ for all $a \in A$. For every $a \in A$ and $i \in N_a$, let $\kappa_i = \{a\}$.

Consider a randomized aggregation rule f. There must exist an alternative $a^* \in A$ such that $\Pr[f(\vec{\kappa}) = \{a^*\}] \leq 1/m$. Now, construct a utility profile $\vec{v}$ such that i) for all $i \in N_{a^*}$, we have $v_i(a^*) = 1$, and $v_i(a) = 0$ for $a \in A \setminus \{a^*\}$; and ii) for all $a \in A \setminus \{a^*\}$ and $i \in N_a$, we have $v_i(a) = v_i(a^*) = 1/2$, and $v_i(b) = 0$ for $b \in A \setminus \{a, a^*\}$.

Note that $\vec{v}$ is consistent with the input profile $\vec{\kappa}$, i.e., $\vec{v} \triangleright \vec{\kappa}$. Moreover, it holds that $\mathrm{sw}(a^*, \vec{v}) \geq n/2$, whereas $\mathrm{sw}(a, \vec{v}) \leq n_a \leq n/m+1$ for $a \in A \setminus \{a^*\}$. It follows that

$$\mathrm{dist}(f) \geq \mathrm{dist}(f, \vec{\kappa}) \geq \frac{n/2}{\frac{1}{m} \cdot n + \frac{m-1}{m} \cdot (\frac{n}{m} + 1)} \geq \frac{m}{6},$$

as desired. $\square$

In light of Observation 2.1, this result indicates that the distortion associated with knapsack votes is asymptotically indistinguishable from the distortion one can achieve with absolutely no information about voter preferences, suggesting that knapsack votes may not be an appropriate input format if the goal is to maximize social welfare. Our aim now is to find input formats that achieve better results when viewed through the implicit utilitarianism lens.

Rankings by Value and by Value for Money.

Goel et al. [2016] also advocate the use of comparisons between alternatives based on value for money, which, like knapsack votes, encourage voters to consider the trade-off between value and cost. We study *rankings* by value for money as an input format; observe that such rankings convey more information than specific pairwise comparisons.

In addition, we also study rankings by value, which are prevalent in the existing literature on implicit utilitarian voting [6, 8, 9, 31, 41, 106]. Rankings by value convey more information than k-approval votes, in which each voter submits the set of top k alternatives by their value — this is the input format of choice for most real-world participatory budgeting elections [75].

Boutilier et al. [2015] prove a lower bound of $\Omega(\sqrt{m})$ on distortion in the special case of our setting where all alternatives have cost 1, the input format is rankings by value, and $n \geq \sqrt{m}$. This result carries over to our more general setting, not only with rankings by value, but also with rankings by value for money, as both input formats coincide in case of equal costs. Our goal is to establish an almost matching upper bound.

We start from a mechanism of Boutilier et al. [2015] that has distortion $\mathcal{O}(\sqrt{m \log m})$ in their setting. It carefully balances between high-value and low-value alternatives (where value is approximately inferred from the positions of the alternatives in the input rankings). In our more general participatory budgeting problem, it is crucial to also take into account the costs, and find the perfect balance between selecting many low-cost alternatives and fewer high-cost ones. We modify the mechanism of Boutilier et al. precisely to achieve this goal. Specifically, we partition the alternatives into $\mathcal{O}(\log m)$ buckets based on their costs, and differentiate between alternatives within a bucket based on their (inferred) value. Our mechanism for rankings by value for money requires more careful treatment as values are obfuscated in value-for-money comparisons.

At first glance our setting seems much more difficult, distortion-wise, than the simple setting of Boutilier et al. [2015]. But ultimately we obtain only a slightly weaker upper bound on the distortion associated with both rankings by value and by value for money. In other words, to our surprise, incorporating costs and a budget constraint comes at almost no cost to social welfare maximization.

Theorem 2.3. *The distortion associated with rankings by value and rankings by value for money is* $\mathcal{O}(\sqrt{m}\ \log m)$.

Proof of Theorem 2.3. We first present the proof for rankings by value for money as it is trickier, and later describe how an almost identical proof works for rankings by value.

Let us begin by introducing additional notation. For a ranking σ and an alternative $a \in A$, let $\sigma(a)$ denote the position of a in σ. For a preference profile $\vec{\sigma}$ with n votes, let the *harmonic score* of a in $\vec{\sigma}$ be defined as $\text{sc}(a, \vec{\sigma}) \triangleq \sum_{j=1}^{n} 1/\sigma_j(a)$. Finally, given a set of alternatives $S \subseteq A$, let $\sigma|_S$ (resp. $\vec{\sigma}|_S$) denote the ranking (resp. preference profile) obtained by restricting σ (resp. $\vec{\sigma}$) to the alternatives in S.

For ease of exposition assume m is a power of 2. Let $\vec{\sigma}$ denote the input profile consisting of voter preferences in the form of rankings by value for money. Let $\vec{v}$ denote the underlying utility profile consistent with $\vec{\sigma}$. Let $S^* \triangleq \arg\max_{S \in \mathcal{F}_c} \text{sw}(S, \vec{v})$ be the budget-feasible set of alternatives maximizing the social welfare.

Let $\ell_0 = 0$ and $u_0 = 1/m$. For $i \in [\log m]$, set $\ell_i = 2^{i-1}/m$ and $u_i = 2^i/m$. Let us partition the alternatives into $\log m + 1$ buckets based on their costs: $S_0 \triangleq \{a \in A : c_a \leq u_0\}$ and $S_i \triangleq \{a \in A : \ell_i < c_a \leq u_i\}$ for $i \in [\log m]$. Note that for $i \in \{0\} \cup [\log m]$, selecting at most $1/u_i$ alternatives from S_i is guaranteed to be budget-feasible.

Next, let us further partition the buckets into two parts: for $i \in \{0\} \cup [\log m]$, let S_i^+ consist of the $\sqrt{m} \cdot (1/u_i)$ alternatives from S_i with the largest harmonic scores in the reduced profile $\vec{\sigma}|_{S_i}$, and $S_i^- \triangleq S_i \setminus S_i^+$. When $|S_i| \leq \sqrt{m} \cdot (1/u_i)$, note that $S_i^+ = S_i$ and $S_i^- = \emptyset$. Note that $S_0^+ = S_0$. Let $S^+ \triangleq \cup_{i=0}^{\log m} S_i^+$ and $S^- \triangleq A \setminus S^+$.

We are now ready to define our randomized aggregation rule, which randomizes over two separate mechanisms.

- *Mechanism A:* Select a bucket S_i uniformly at random, and select a $(1/u_i)$-size subset of S_i^+ uniformly at random.
- *Mechanism B:* Select a single alternative uniformly at random.

Our aggregation rule executes each mechanism with an equal probability 1/2. We now show that this rule achieves distortion that is $\mathcal{O}(\sqrt{m} \log m)$.

First, note that mechanism A selects each bucket S_i with probability $1/(\log m + 1)$, and when S_i is selected, it selects each alternative in S_i^+ with probability at least $1/\sqrt{m}$. (This is because the mechanism selects $1/u_i$ alternatives at random from S_i^+, which has at most $\sqrt{m} \cdot (1/u_i)$ alternatives.) Hence, the mechanism selects each alternative in S^+ (and therefore, each alternative in $S^* \cap S^+$) with probability at least $1/(\sqrt{m}(\log m + 1))$. In other words, the expected social welfare achieved under mechanism A is $\mathcal{O}(\sqrt{m}\log m)$ approximation of $\mathrm{sw}(S^* \cap S^+, \vec{v})$.

Finally, to complete the proof, we show that the expected welfare achieved under mechanism B is an $\mathcal{O}(\sqrt{m}\log m)$ approximation of $\mathrm{sw}(S^* \cap S^-, \vec{v})$. Let us first bound $\mathrm{sw}(S^* \cap S^-, \vec{v})$. Recall that $S_0^- = \emptyset$. Hence,

$$\mathrm{sw}(S^* \cap S^-, \vec{v}) = \sum_{i=1}^{\log m} \mathrm{sw}(S^* \cap S_i^-, \vec{v}).$$

Fix $i \in [\log m]$ and $a \in S_i^-$. One can easily check that

$$\sum_{b \in S_i} \mathrm{sc}(b, \vec{\sigma}|_{S_i}) = n \cdot H_{|S_i|} \leq n \cdot H_m,$$

where H_k is the k^{th} harmonic number. Because S_i^+ consists of the $\sqrt{m}/u_i$ alternatives in S_i with the largest harmonic scores, we have

$$\mathrm{sc}(a, \vec{\sigma}|_{S_i}) \leq \frac{n \cdot H_m}{\sqrt{m} \cdot (1/u_i)} = \frac{n \cdot (1 + \log m)}{\sqrt{m} \cdot m/2^i}. \tag{2.1}$$

Next, we connect this bound on the harmonic score of a to a bound on its social welfare. For simplicity, let us denote $\vec{\gamma} \triangleq \vec{\sigma}|_{S_i}$. Due to our definition of the partitions, we have

$$c_a \leq 2 \cdot c_b, \forall b \in S_i. \tag{2.2}$$

Further, fix a voter $j \in [n]$. For each alternative b such that $b \succ_{\gamma_j} a$, we also have $v_j(b)/c_b \geq v_j(a)/c_a$. Substituting Equation (2.2), we get

$$v_j(a) \leq 2 v_j(b), \forall j \in [n], b \in S_i \text{ s.t. } b \succ_{\gamma_j} a. \tag{2.3}$$

Taking a sum over all $b \in S_i$ with $b \succ_{\gamma_j} a$, and using the fact that the values of each voter j sum to 1, we get $v_j(a) \leq 2/\gamma_j(a)$ for $j \in [n]$, and taking a further sum over $j \in [n]$, we get

$$\mathrm{sw}(a, \vec{v}) \leq 2 \cdot \mathrm{sc}(a, \vec{\sigma}|_{S_i}). \tag{2.4}$$

Combining this with Equation (2.1), we get

$$\mathrm{sw}(a, \vec{v}) \leq \frac{2 \cdot n \cdot (1 + \log m)}{\sqrt{m} \cdot m/2^i}, \forall a \in S_i^-.$$

Note that S^* can contain at most $2/u_i = m/2^{i-1}$ alternatives from S_i while respecting the budget constraint. Hence,

$$\begin{aligned}\mathrm{sw}(S^* \cap S^-, \vec{v}) = \sum_{i=1}^{\log m} \mathrm{sw}(S^* \cap S_i^-, \vec{v}) &\leq \frac{(m/2^{i-1}) \cdot 2 \cdot n \cdot (1 + \log m)}{\sqrt{m} \cdot m/2^i} \\ &= 4 \cdot n \cdot (1 + \log m)/\sqrt{m}. \qquad (2.5)\end{aligned}$$

Because the utilities sum to 1 for each voter, the expected social welfare achieved under mechanism B is $(1/m) \cdot \sum_{i \in N} \sum_{a \in A} v_i(a) = n/m$, which is an $\mathcal{O}(\sqrt{m} \log m)$ approximation of $\mathrm{sw}(S^* \cap S^-, \vec{v})$ due to Equation (2.5).

This completes the proof of $\mathcal{O}(\sqrt{m} \log m)$ distortion associated with rankings by value for money. The proof for rankings by value is almost identical. In fact, one can make two simplifications.

First, the factor of 2 from Equation (2.3), and therefore from Equation (2.4) disappears because the rankings already dictate comparison by value. This leads to an improvement in Equation (2.5) by a factor of 2.

Second, Equation (2.3) not only holds for $b \in S_i$ such that $b \succ_{\gamma_j} a$, but holds more generally for $b \in A$ such that $b \succ_{\sigma_j} a$. Hence, there is no longer a need to compute the harmonic scores on the restricted profile $\vec{\sigma}|_{S_i}$; one can simply work with the original input profile $\vec{\sigma}$. □

Threshold Approval Votes.

Approval voting — where voters can choose to approve any subset of alternatives, and a most widely approved alternative wins — is well studied in social choice theory [33]. In our utilitarian setting we reinterpret this input format as *threshold approval votes*, where the principal sets a threshold t, and each voter $i \in N$ approves every alternative a for which $v_i(a) \geq t$.

We first investigate *deterministic threshold approval votes*, in which the threshold is selected deterministically, but find that it does not help us (significantly) improve over the distortion we can already obtain using rankings by value or by value for money. Specifically, for a fixed threshold, we are always able to construct cases in which alternatives have significantly different welfares, but either no alternative is approved or an extremely large set of alternatives are approved, providing the rule little information to distinguish between the alternatives, and yielding high distortion.

Theorem 2.4. *The distortion associated with deterministic threshold approval votes is* $\Omega(\sqrt{m})$.

Proof of Theorem 2.4. Imagine the case where $c_a = 1$ for all alternatives $a \in A$. Recall that the budget is 1. Let f denote a randomized aggregation rule. (While we study deterministic and randomized threshold selection, we still allow randomized aggregation rules. Section 2.3.2 studies the case where the aggregation rule has to be deterministic.) It must return a single alternative, possibly chosen in a randomized fashion. We construct our adversarial input profile based on whether $t \leq 1/\sqrt{m}$. Let $A \triangleq \{a_1, \ldots, a_m\}$.

Suppose $t \leq 1/\sqrt{m}$. Fix a set of alternatives $S \subseteq A$ such that $|S| = \sqrt{m}/2 + 1$ (assume for ease of exposition $\sqrt{m}$ is an even integer). Construct the input profile $\vec{\tau}$ such that $\tau_i = S$ for all $i \in N$. Now, there must exist $a^* \in S$ such that $\Pr[f(\vec{\tau}) = \{a^*\}] \leq 1/(\sqrt{m}/2+1)$. Construct the underlying utility profile $\vec{v}$ such that for each voter $i \in N$, $v_i(a^*) = 1/2$, $v_i(a) = 1/\sqrt{m}$ for $a \in S \setminus \{a^*\}$, and $v_i(a) = 0$ for $a \in A \setminus S$. Note that this is consistent with the input profile given that $t \leq 1/\sqrt{m}$. Further, $\mathrm{sw}(a^*, \vec{v}) = n/2$ whereas $\mathrm{sw}(a, \vec{v}) \leq n/\sqrt{m}$ for all $a \in A \setminus \{a^*\}$. Hence,

$$\mathbb{E}[\mathrm{sw}(f(\vec{\tau}), \vec{v})] \leq \frac{1}{\sqrt{m}/2+1} \cdot \frac{n}{2} + \frac{\sqrt{m}/2}{\sqrt{m}/2+1} \cdot \frac{n}{\sqrt{m}} = \mathcal{O}\left(\frac{n}{\sqrt{m}}\right).$$

Because the optimal social welfare is $\Theta(n)$, we have that $\mathrm{dist}(f) = \Omega(\sqrt{m})$, as required.

Now suppose that $t > 1/\sqrt{m}$. Construct an input profile $\vec{\tau}$ in which $\tau_i = \emptyset$ for every voter $i \in N$. In this case, there exists an alternative $a^* \in A$ such that $\Pr[f(\vec{\tau}) = a^*] \leq 1/m$. Let us construct the underlying utility profile $\vec{v}$ as follows. For every voter $i \in N$, let $v_i(a^*) = 1/\sqrt{m}$, and $v_i(a) = (1-1/\sqrt{m})/m$ for all $a \in A \setminus \{a^*\}$. Note that this is consistent with the input profile given that $t > 1/\sqrt{m}$. Clearly, the optimal social welfare is achieved by $\mathrm{sw}(a^*, \vec{v}) = n/\sqrt{m}$. In contrast, we have

$$\mathbb{E}[\mathrm{sw}(f(\vec{\tau}), \vec{v})] \leq \frac{1}{m} \cdot \frac{n}{\sqrt{m}} + \left(1 - \frac{1}{\sqrt{m}}\right) \cdot \frac{1-1/\sqrt{m}}{m} = \mathcal{O}\left(\frac{n}{m}\right).$$

Hence, we again have $\mathrm{dist}(f) = \Omega(\sqrt{m})$, as desired. □

For specific ranges of the threshold, it is possible to derive stronger lower bounds. However, the $\Omega(\sqrt{m})$ lower bound of Theorem 2.4 is sufficient to establish a clear asymptotic separation between the power of deterministic and randomized threshold approval votes.

Under *randomized threshold approval votes*, we can select the threshold in a randomized fashion. Technically, this is a distribution over input formats, one for each value of the threshold. Before we define the (overall) distortion of a rule that randomizes over input formats, let us recall the definition of the overall distortion of a rule for a fixed input format:

$$\operatorname{dist}(f) \triangleq \max_{\vec{\rho}} \sup_{\vec{v}:\vec{v} \triangleright \vec{\rho}} \frac{\max_{T \in \mathcal{F}_c} \operatorname{sw}(T, \vec{v})}{\mathbb{E}[\operatorname{sw}(f(\vec{\rho}), \vec{v})]} = \sup_{\vec{v}} \frac{\max_{T \in \mathcal{F}_c} \operatorname{sw}(T, \vec{v})}{\mathbb{E}[\operatorname{sw}(f(\vec{\rho}), \vec{v})]}.$$

Here, $\vec{\rho}(\vec{v})$ denotes the input profile induced by utility profile $\vec{v}$. In the case of randomized threshold approval votes, rule f specifies a distribution D over the threshold t, as well as the aggregation of input profile $\vec{\rho}(\vec{v}, t)$ induced by utility profile $\vec{v}$ and a given choice of threshold t. We define the the (overall) distortion of rule f as

$$\operatorname{dist}(f) \triangleq \sup_{\vec{v}} \ \mathbb{E}_{t \sim D} \ \frac{\max_{T \in \mathcal{F}_c} \operatorname{sw}(T, \vec{v})}{\mathbb{E}[\operatorname{sw}(f(\vec{\rho}(\vec{v}, t)), \vec{v})]}.$$

Interestingly, observe that due to the expectation over threshold t, which affects the induced input profile $\vec{\rho}(\vec{v}, t)$, we can no longer decompose the maximum over $\vec{v}$ into a maximum over $\vec{\rho}$ followed by a maximum over $\vec{v}$ such that $\vec{v} \triangleright \vec{\rho}$, in contrast to the case of a fixed input format.

This flexibility of randomizing the threshold value allows us to dramatically reduce the distortion.

Theorem 2.5. *The distortion associated with randomized threshold approval votes is $\mathcal{O}(\log^2 m)$.*

Proof of Theorem 2.5. For ease of exposition, assume m is a power of 2. Let $I_0 \triangleq [0, 1/m^2]$, and $I_j \triangleq (2^{j-1}/m^2, 2^j/m^2]$, $\ell_j = 2^{j-1}/m^2$, and $u_j = 2^j/m^2$ for $j = 1, \ldots, 2\log m$.

Let $\vec{v}$ denote a utility profile that is consistent with the input profile. For $a \in A$ and $j \in \{0, \ldots, 2\log m\}$, define $n_j^a \triangleq |\{i \in N : v_i(a) \in I_j\}|$ to be the number of voters whose utility for a falls in the interval I_j. We now bound the social welfare of a in terms of the numbers n_j^a. Specifically,

$$\operatorname{sw}(a, \vec{v}) = \sum_{i \in N} v_i(a) \leq \sum_{j=0}^{2\log m} \sum_{i \in N} \mathbb{I}\{v_i(a) \in I_j\} \cdot u_j = \sum_{j=0}^{2\log m} n_j^a \cdot u_j,$$

where $\mathbb{I}$ indicates the indicator variable. A similar argument also yields a lower bound, and after substituting $\ell_0 = 0$, $u_0 = 1/m^2$, and $n_0^a \leq n$, we get

$$\sum_{j=1}^{2\log m} n_j^a \cdot \ell_j \leq \operatorname{sw}(a, \vec{v}) \leq \frac{n}{m^2} + \sum_{j=1}^{2\log m} n_j^a \cdot u_j. \tag{2.6}$$

Next, divide the alternatives into $1+2\log m$ buckets based on their costs, with bucket $S_j \triangleq \{a \in A : c_a \in I_j\}$. Note that selecting at most $1/u_j$ alternatives from S_j is guaranteed to satisfy the budget constraint.

Let $S^* \triangleq \arg\max_{S\in\mathcal{F}_c} \mathrm{sw}(S,\vec{v})$ be the feasible set of alternatives maximizing the social welfare. For $j,k \in \{0,\ldots,2\log m\}$, let $n^*_{j,k} \triangleq \sum_{a\in S^*\cap S_k} n^a_j$. Using Equation (2.6), we have

$$\sum_{j=1}^{2\log m} n^*_{j,k} \cdot \ell_j \le \mathrm{sw}(S^* \cap S_k, \vec{v}) \le |S^* \cap S_k| \cdot \frac{n}{m^2} + \sum_{j=1}^{2\log m} n^*_{j,k} \cdot u_j. \tag{2.7}$$

We now construct three different mechanisms; our final mechanism will randomize between them.

Mechanism A: Pick a pair (j,k) uniformly at random from the set $T \triangleq \{(j,k) : j,k \in [2\log m]\}$. Then, set the threshold to ℓ_j, and using the resulting input profile, greedily select the $1/u_k$ alternatives from S_k with the largest number of approval votes (or select S_k if $|S_k| \le 1/u_k$). Let $B_{j,k}$ denote the set of selected alternatives for the pair (j,k). Because we have $j > 0$ and $k > 0$,

$$\mathrm{sw}(B_{j,k},\vec{v}) \ge \sum_{a\in B_{j,k}} \left(\sum_{p=j}^{2\log m} n^a_p\right) \cdot \ell_j \ge \frac{1}{4} \cdot \left(\sum_{p=j}^{2\log m} n^*_{p,k}\right) \cdot u_j \ge \frac{1}{4} \cdot n^*_{j,k} \cdot u_j, \tag{2.8}$$

where, in the first transition, we bound the welfare from below by only considering utilities that are at least ℓ_j, and the second transition holds because $u_j = 2\ell_j$, $|S^* \cap S_k| \le 2|B_{j,k}|$, and $B_{j,k}$ consists of greedily-selected alternatives with the highest number of approval votes. Thus, the expected social welfare achieved by mechanism A is

$$\begin{aligned}
\frac{1}{(2\log m)^2} \sum_{j=1}^{2\log m} \sum_{k=1}^{2\log m} \mathrm{sw}(B_{j,k},\vec{v}) &\ge \frac{1}{4\cdot(2\log m)^2} \sum_{j=1}^{2\log m} \sum_{k=1}^{2\log m} n^*_{j,k} \cdot u_j \\
&\ge \frac{1}{16\log^2 m} \left(\mathrm{sw}(S^*\backslash S_0,\vec{v}) - |S^*\backslash S_0| \cdot \frac{n}{m^2}\right) \\
&\ge \frac{1}{16\log^2 m} \left(\mathrm{sw}(S^*\backslash S_0,\vec{v}) - \frac{n}{m}\right),
\end{aligned}$$

where the first transition follows from Equation (2.8), and the second transition follows from Equation (2.7).

Mechanism B: Select all the alternatives in S_0. Because each alternative in S_0 has cost at most $1/m^2$, this is clearly budget-feasible. The social welfare achieved by this mechanism is $\mathrm{sw}(S_0, \vec{v}) \geq \mathrm{sw}(S^* \cap S_0, \vec{v})$.

Mechanism C: Select a single alternative uniformly at random from A. This is also budget-feasible, and due to normalization of values, its expected social welfare is n/m.

Our final mechanism executes mechanism A with probability $16 \log^2 m/(2 + 16 \log^2 m)$, and mechanisms B and C each with probability $1/(2 + 16 \log^2 m)$. It is easy to see that its expected social welfare is at least $\mathrm{sw}(S^*, \vec{v})/(2 + 16 \log^2 m)$. Hence, its distortion is $\mathcal{O}(\log^2 m)$. □

We also show that at least logarithmic distortion is inevitable even when using randomized threshold approval votes.

Theorem 2.6. *The distortion associated with randomized threshold approval votes is $\Omega(\log m/ \log \log m)$.*

Proof of Theorem 2.6. Imagine the case where $c_a = 1$ for all $a \in A$. Recall that the budget is 1. Let f denote a rule that elicits randomized threshold approval votes and aggregates them to return a distribution over A (as only a single project can be executed at a time). Note that f is not simply the aggregation rule, but the elicitation method and the aggregation rule combined.

Divide the interval $(1/m, 1]$ into $\lceil \log m/ \log (2 \log m) \rceil$ sub-intervals: For $j \in [\lceil \log m/ \log (2 \log m) \rceil]$, let

$$I_j \triangleq \left(\frac{(2 \log m)^{j-1}}{m}, \min \left\{ \frac{(2 \log m)^j}{m}, 1 \right\} \right],$$

note that the minimum in the upper bound only affects the last interval. Let u_j and ℓ_j denote the upper and lower end points of I_j and observe that $u_j \leq 2 \log m \cdot \ell_j$ for all $j \in [\lceil \log m/ \log (2 \log m) \rceil]$.

Let t denote the threshold picked randomly by f. There must exist $k \in [\lceil \log m/ \log (2 \log m) \rceil]$ such that $\Pr[t \in I_k] \leq \log (2 \log m)/ \log m$. Fix a subset $S \subseteq A$ of size $\log m$, and let $V = u_k/2 + (\log m - 1) \cdot \ell_k$. Construct a (partial) utility profile $\vec{v}$ such that for each voter $i \in N$, $v_i(a) \in I_k$ for $a \in S$, $\sum_{a \in S} v_i(a) = V$, and $v_i(a) = (1 - V)/(m - \log m)$ for $a \in A \setminus S$. First, this is feasible because

$$V = \frac{u_k}{2} + (\log m - 1) \cdot \ell_k \leq \frac{1}{2} + \frac{\log m - 1}{2 \log m} \leq 1.$$

Second, this partial description completely dictates the induced input profile when $t \notin I_k$. Because f can only distinguish between alternatives in S when $t \in I_k$, there must exist $a^* \in S$ such that $\Pr[f \text{ returns } a^* | t \notin I_k] \leq 1/\log m$. Suppose the underlying utility profile $\vec{v}$ satisfies, for each voter $i \in N$, $v_i(a^*) = u_k/2$ and $v_i(a) = \ell_k$ for $a \in S \setminus \{a^*\}$. Observe that this is consistent with the partial description provided before.

In this case, the optimal social welfare is given by $\text{sw}(a^*, \vec{v}) = n \cdot u_k/2$, whereas $\text{sw}(a, \vec{v}) \leq n \cdot \ell_k$ for all $a \in A \setminus \{a^*\}$. The latter holds because $\ell_k > (1-V)/(m - \log m)$. The expected social welfare achieved by f under $\vec{v}$ is at most

$$\begin{aligned}\Pr[t \in I_k] \cdot \frac{n \cdot u_k}{2} &+ \Pr[t \notin I_k] \left(\frac{1}{\log m} \cdot \frac{n \cdot u_k}{2} + \frac{\log m - 1}{\log m} \cdot n \cdot \ell_k \right) \\ &\leq \frac{\log (2 \log m) + 2}{\log m} \cdot \frac{n \cdot u_k}{2},\end{aligned}$$

where the final transition holds because $u_k \leq 2 \log m \cdot \ell_k$. Thus, the distortion achieved by f is $\Omega(\log m / \log \log m)$, as desired. □

Our proof of Theorem 2.6 establishes a lower bound of $\Omega(\log m / \log \log m)$ on the distortion associated with randomized threshold approval votes by only using the special case of the participatory budgeting problem in which $c_a = 1$ for each $a \in A$, i.e., exactly one alternative needs to be selected. This is exactly the setting studied by Boutilier et al. [2015]. On the other hand, Theorem 2.5 establishes a slightly weaker upper bound of $\mathcal{O}(\log^2 m)$ for the general participatory budgeting problem. For the restricted setting of Boutlier et al. [2015], one can improve the general $\mathcal{O}(\log^2 m)$ upper bound to $\mathcal{O}(\log m)$, thus leaving a very narrow gap from the $\Omega(\log m / \log \log m)$ lower bound. This proof is similar to the proof of Theorem 2.5, whose $\mathcal{O}(\log^2 m)$ bound is the result of a randomization over $\mathcal{O}(\log m)$ partitions of the alternatives based on their cost and $\mathcal{O}(\log m)$ possible values of the threshold. When costs are identical there is no need to partition based on cost, reducing the partitions by a logarithmic factor.

Theorem 2.7. *If $c_a = 1$ for all $a \in A$, the distortion associated with randomized threshold approval votes is $\mathcal{O}(\log m)$.*

Proof of Theorem 2.7. For $j \in [\log m]$, let $\ell_j = 2^{j-1}/m$ and $u_j = 2 \cdot \ell_j$. Consider the rule which chooses $j \in [\log m]$ uniformly at random, elicits approval votes with threshold $t = \ell_j$, and returns an alternative with the greatest number of approval votes. We show that the distortion of this rule is $\mathcal{O}(\log m)$.

Let $\vec{v}$ denote the underlying utility profile, and $a^* \triangleq \arg\max_{a \in A} \mathrm{sw}(a, \vec{v})$ be the welfare-maximizing alternative. If there exists $j \in [\log m]$ such that our rule returns a^* when it sets the threshold $t = \ell_j$ (which happens with probability $1/\log m$), we immediately obtain $\mathcal{O}(\log m)$ distortion. Let us assume that our rule never returns a^*. For $a \in A$ and $j \in [\log m]$, let n_j^a denote the number of approval votes a receives when the threshold $t = \ell_j$, and let $a_j \in A$ be the alternative returned by our rule when $t = \ell_j$. Because our rule returns an alternative with the greatest number of approval votes, we have

$$\forall j \in [\log m], \ \sum_{k=j}^{\log m} n_k^{a_j} \geq \sum_{k=j}^{\log m} n_k^{a^*} \geq n_j^{a^*}. \tag{2.9}$$

Now, the expected social welfare achieved by our rule is at least

$$\begin{aligned} \sum_{j=1}^{\log m} \Pr[t = \ell_j] \cdot \mathrm{sw}(a_j, \vec{v}) &\geq \frac{1}{\log m} \sum_{j=1}^{\log m} \ell_j \left(\sum_{k=j}^{\log m} n_k^{a_j} \right) \\ &\geq \frac{1}{2 \log m} \sum_{j=1}^{\log m} u_j \cdot n_j^{a^*} \geq \frac{1}{2 \log m} \cdot \mathrm{sw}(a^*, \vec{v}), \end{aligned}$$

where the first transition follows from Equation (2.9), and the second transition holds because $\ell_j = u_j/2$. Hence, the distortion of our rule is $\mathcal{O}(\log m)$, as desired. $\square$

2.3.2 Deterministic Aggregation Rules

We next study the distortion that can be achieved under different input formats if we are forced to use a deterministic aggregation rule. Recall that the distortion associated with *deterministic aggregation* of votes under an input format is the least distortion a deterministic aggregation rule for that format can achieve. Specifically, we study the distortion associated with deterministic aggregation of knapsack votes, rankings by value and value for money, and deterministic threshold approval votes. We omit randomized threshold approval votes as the inherent randomization involved in elicitation makes the use of deterministic aggregation rules hard to justify.

We find that rankings by value achieve $\Theta(m^2)$ distortion, which is significantly better than the distortion of knapsack votes (exponential in m) and that of rankings by value for money (unbounded). This separation between rankings by value and value for money in this setting stands in stark contrast to the setting with randomized aggregation rules, where both input formats admit similar distortion. One important fact, however, does not change with the use of deterministic aggregation rules: using threshold approval votes still performs at least as well as using any of the other input formats considered here. Specifically, we show that setting the threshold to be $t = 1/m$ results in $\mathcal{O}(m^2)$ distortion. The choice of the threshold is crucial as, for example, setting a slightly higher threshold $t > 1/(m-1)$ results in unbounded distortion.

Knapsack Votes.

Our first result is an exponential lower bound on the distortion associated with knapsack votes when the aggregation rule is deterministic. While our construction requires the number of voters to be extremely large compared to the number of alternatives, we remark that this is precisely the case in real participatory budgeting elections, in which a large number of citizens vote over much fewer projects.

Theorem 2.8. *For sufficiently large n, the distortion associated with deterministic aggregation of knapsack votes is $\Omega(2^m/\sqrt{m})$.*

Proof of Theorem 2.8. Imagine a case where every alternative has cost $2/m$ (recall that the budget is 1). It follows that no more than $\lfloor m/2 \rfloor$ alternatives may be selected while respecting the budget constraints. Let $S_1, \ldots, S_{\binom{m}{\lfloor m/2 \rfloor}}$ denote the $\binom{m}{\lfloor m/2 \rfloor}$ subsets of A of size $\lfloor m/2 \rfloor$.

Assume $n \geq \binom{m}{\lfloor m/2 \rfloor}$. Partition voters into $\binom{m}{\lfloor m/2 \rfloor}$ sets $N_1, \ldots, N_{\binom{m}{\lfloor m/2 \rfloor}}$, each consisting of roughly $n/\binom{m}{\lfloor m/2 \rfloor}$ voters; specifically, ensure that $\lfloor n/\binom{m}{\lfloor m/2 \rfloor} \rfloor \leq n_i \leq \lceil n/\binom{m}{\lfloor m/2 \rfloor} \rceil$, where $n_i \triangleq |N_i|$, for all $i \in [\binom{m}{\lfloor m/2 \rfloor}]$. Construct an input profile of knapsack votes $\vec{\kappa}$, where $\kappa_i \triangleq S_k$ for all $k \in [\binom{m}{\lfloor m/2 \rfloor}]$ and $i \in N_k$.

Let f denote a deterministic aggregation rule. We can safely assume that $|f(\vec{\kappa})| = \lfloor m/2 \rfloor$ as otherwise we can add alternatives to $f(\vec{\kappa})$, which can only improve the distortion. Let $f(\vec{\kappa}) \triangleq S_{k^*}$.

Construct a utility profile $\vec{v}$ consistent with the input profile $\vec{\kappa}$ as follows. Fix $b \in S_{k^*}$, and for all $i \in N_{k^*}$, let $v_i(b) = 1$ and $v_i(a) = 0$ for all $a \in A \setminus \{b\}$. Note that these valuations are consistent with the votes of voters in N_{k^*}.

Next, fix $a^* \in A \setminus S_{k^*}$. Our goal is to make a^* an attractive alternative that $f(\vec{\kappa})$ missed. Note that a^* appears in about half of the $\lfloor m/2 \rfloor$-sized subsets of A. For all $k \in [\binom{m}{\lfloor m/2 \rfloor}]$ such that $a^* \in S_k$, and all voters $i \in N_k$, let $v_i(a^*) = 1$ and $v_i(a) = 0$ for all $a \in A \setminus \{a^*\}$. This ensures $\mathrm{sw}(a^*, \vec{v}) \geq n \cdot \frac{\lfloor m/2 \rfloor}{m} \geq n/3$ (for $m \geq 2$).

For $k \in [\binom{m}{\lfloor m/2 \rfloor}] \setminus \{k^*\}$ such that $a^* \notin S_k$, and all voters $i \in N_k$, let $v_i(a') = 1$ for some $a' \in S_k \setminus S_{k^*}$, and $v_i(a) = 0$ for all $a \in A \setminus \{a'\}$.

Observe that all voters who do not belong to N_{k^*} assign zero utility to all the alternatives in S_{k^*}, yielding $\mathrm{sw}(f(\vec{\kappa}), \vec{v}) \leq n_{k^*} \leq n/\binom{m}{\lfloor m/2 \rfloor} + 1$. By assumption, $n \geq \binom{m}{\lfloor m/2 \rfloor}$, so we have

$$\mathrm{dist}(f, \vec{v}) \geq \frac{n/3}{n/\binom{m}{\lfloor m/2 \rfloor} + 1} = \frac{1}{6} \cdot \binom{m}{\lfloor m/2 \rfloor} = \Omega\left(\frac{2^m}{\sqrt{m}}\right),$$

as required. □

We next show that an almost matching upper bound can be achieved by the natural "plurality knapsack" rule that selects the subset of alternatives submitted by the largest number of voters.

Theorem 2.9. *The distortion associated with deterministic aggregation of knapsack votes is $\mathcal{O}(m \cdot 2^m)$.*

Proof of Theorem 2.9. Let $\vec{v}$ denote the underlying utility profile, and let $S^* \subseteq A$ be the set of alternatives reported by the largest number of voters. Due to the pigeonhole principle, it must be reported by at least $n/2^m$ voters. Further, each voter i who reports S^* must have $v_i(S^*) \geq 1/m$ because there must exist $a \in A$ such that $v_i(a) \geq 1/m$, and $v_i(S^*) \geq v_i(a)$.

Hence, we have $\mathrm{sw}(S^*, \vec{v}) \geq (n/2^m) \cdot 1/m$, whereas the maximum welfare any set of alternatives can achieve is at most n. Hence, the distortion of the proposed rule is at most $m \cdot 2^m$. □

Rankings by Value and by Value for Money.

While rankings by value and by value for money have similar distortion in case of randomized aggregation rules, deterministic aggregation rules lead to a clear separation between the distortion of the two input formats.

We first show that deterministic aggregation of rankings by value for money cannot offer bounded distortion. Our counterexample exploits the uncertainty in values induced when alternatives have vastly different costs.

Theorem 2.10. *The distortion associated with deterministic aggregation of rankings by value for money is unbounded.*

Proof of Theorem 2.10. Fix $a, b \in A$. Let $c_a \triangleq \epsilon > 0$, and $c_k = 1$ for all $k \in A \setminus \{a\}$. Recall that the budget is 1. Hence, every deterministic aggregation rule must select a single alternative.

Construct an input profile $\vec{\sigma}$ in which each input ranking has alternatives a and b in positions 1 and 2, respectively. Let f be a deterministic aggregation rule.

If $f(\vec{\sigma}) \in A \setminus \{a\}$, the utility profile $\vec{v}$ in which every voter has utility 1 for a, and 0 for every alternative in $A \setminus \{a\}$ ensures $\text{dist}(f) \geq \text{dist}(f, \vec{v}) = \infty$.

If $f(\vec{\sigma}) = a$, the utility profile $\vec{v}$ in which every voter has utility ϵ for a, $1 - \epsilon$ for b, and 0 for every alternative in $A \setminus \{a, b\}$ ensures that $\text{dist}(f) \geq \text{dist}(f, \vec{v}) = (1 - \epsilon)/\epsilon$.

Hence, in either case, $\text{dist}(f) \geq (1 - \epsilon)/\epsilon$. Because ϵ can be arbitrarily small, the distortion is unbounded. □

We now turn our attention to rankings by value. Caragiannis et al. [2016] study deterministic aggregation of rankings by value in the special case of our setting where the cost of each alternative equals the entire budget, and establish a lower bound of $\Omega(m^2)$ on the distortion, which carries over to our more general setting.

Theorem 2.11 (Caragiannis et al. [44])**.** *For $n \geq m - 1$, the distortion associated with deterministic aggregation of rankings by value is $\Omega(m^2)$.*

Caragiannis et al. [2016] also show that selecting the plurality winner — the alternative that is ranked first by the largest number of voters — results in distortion at most m^2. We show that this holds true even in our more general setting, giving us an asymptotically tight bound on the distortion.

Theorem 2.12. *The distortion associated with deterministic aggregation of rankings by value is $\mathcal{O}(m^2)$.*

Proof of Theorem 2.12. Due to the pigeonhole principle, the plurality winner, say $a \in A$, must be ranked first by at least n/m voters, each of which must have utility at least $1/m$ for a. Hence, the social welfare of a is at least n/m^2, while the maximum social welfare that any set of alternatives can achieve is at most n, yielding a distortion of at most m^2.

□

Threshold Approval Votes.

We now turn our attention to threshold approval votes. As mentioned earlier, our use of deterministic aggregation rules makes randomized threshold selection less motivated; we thus focus on deterministic threshold approval votes. First, we show that for some choices of the threshold, the distortion can be unbounded.

Theorem 2.13. *For a fixed threshold $t > 1/(m-1)$, the distortion associated with deterministic aggregation of deterministic threshold approval votes is unbounded.*

Proof of Theorem 2.13. Suppose $c_a = 1$ for each $a \in A$. Recall that the budget is 1. Let f denote a deterministic aggregation rule for threshold approval votes. Suppose the rule receives an input profile $\vec{\tau}$ in which no voter approves any alternative. Without loss of generality, let $f(\vec{\tau}) = a^*$.

We construct an underlying utility profile such that for each voter $i \in N$, $v_i(a) = 1/(m-1)$ for $a \in A \backslash \{a^*\}$, and $v_i(a^*) = 0$. Note that this is consistent with $\vec{\tau}$. Now, the optimal social welfare is $n \cdot 1/(m-1)$, whereas the welfare achieved by f is zero, yielding an unbounded distortion. □

We next show that slightly reducing the threshold to $1/m$ reduces the distortion to $\mathcal{O}(m^2)$, which is at least as good as the distortion associated with any other input format.

Theorem 2.14. *For the fixed threshold $t = 1/m$, the distortion associated with deterministic aggregation of deterministic threshold approval votes is $\mathcal{O}(m^2)$.*

Proof of Theorem 2.14. Let $\vec{\tau}$ denote an input profile, and $\vec{v}$ the underlying utility profile. Let $S^* \in \mathcal{F}_c$ denote the feasible set of alternatives with the highest number of total approvals. The set $S \in \mathcal{F}_c$ is returned by the following algorithm: label the alternatives in order of the number of approvals received to cost, where a_1 has the greatest ratio. Return whichever of $\{a_1, \ldots, a_{k-1}\}$ and $\{a_k\}$ has more approvals, with k chosen so that $\{a_1, \ldots, a_{k-1}\} \in \mathcal{F}_c$ and $\{a_1, \ldots, a_k\} \notin \mathcal{F}_c$. Let P^* and P denote the total number of approvals received by alternatives in S^* and S, respectively.

Consider a knapsack problem where the value of an alternative is the number of approvals it receives under $\vec{\tau}$. Then, P^* is the optimal knapsack solution, whereas P is the solution quality achieved by the greedy algorithm. Using the fact that this algorithm achieves a 2-approximation of the (unbounded) knapsack problem [58], we have

$$P \geq (1/2) \cdot P^*.$$

We can now establish an upper bound on the distortion of our rule. Let T be the feasible set of alternatives maximizing the social welfare. Then, T achieves at most P^* total approvals under $\vec{\tau}$. Each approval of an alternative in T by a voter can contribute at most 1 to the welfare of T, and each non-approval of an alternative in T by a voter can contribute at most $1/m$ to the welfare of T. Hence, we have

$$\mathrm{sw}(T, \vec{v}) \leq P^* \cdot 1 + (n \cdot m - P^*) \cdot (1/m).$$

Using a similar line of argument, we also have

$$\mathrm{sw}(S, \vec{v}) \geq P \cdot (1/m).$$

Hence, the distortion of f is at most

$$\begin{aligned} \frac{P^* + (n \cdot m - P^*)/m}{P/m} &\leq 2 \cdot \frac{1 + (n \cdot m/P^* - 1)/m}{1/m} \\ &= 2 \cdot \left(m + \frac{n \cdot m}{n/m} - 1 \right) = \mathcal{O}(m^2), \end{aligned}$$

where the first transition follows from $P \geq P^*/2$. For the second transition, note that with the threshold being $1/m$, each voter must approve at least 1 alternative. Hence, there must exist an alternative with at least n/m approvals, implying that $P^* \geq n/m$.

□

2.4 Computing Worst-Case Optimal Aggregation Rules

Our theoretical results focus on the best worst-case (over all input profiles) distortion we can achieve using different input formats. However, specific input profiles may admit distortion much better than this worst case bound. In practice, we are more interested in the deterministic or randomized aggregation rule that, on each input profile, returns the feasible set of alternatives or a distribution thereover which minimizes distortion, thus achieving the optimal distortion on *each* input profile individually. The optimal deterministic aggregation rule is given by

$$f^*(\vec{\rho}) \triangleq \arg\min_{S\in\mathcal{F}_c} \max_{\vec{v}\rhd\vec{\rho}} \frac{\max_{T\in\mathcal{F}_c} \mathrm{sw}(T,\vec{v})}{\mathrm{sw}(S,\vec{v})}, \ \forall\vec{\rho},$$

and the optimal randomized aggregation rule is given by

$$\bar{f}^*(\vec{\rho}) \triangleq \arg\min_{p\in\Delta(\mathcal{F}_c)} \max_{\vec{v}\rhd\vec{\rho}} \frac{\max_{T\in\mathcal{F}_c} \mathrm{sw}(T,\vec{v})}{\mathbb{E}[\mathrm{sw}(p,\vec{v})]}, \ \forall\vec{\rho},$$

where $\Delta(X)$ denotes the set of distributions over the elements of X.

While these profile-wise optimal aggregation rules dominate all other aggregation rules, they may be computationally difficult to implement, because they optimize a non-linear objective function (a ratio) over a complicated space.

We now turn our attention to designing practical generic algorithms for computing the deterministic and randomized profile-wise optimal aggregation rules for the input formats we study. Throughout this section, we assume that it is practically feasible to explicitly enumerate the collection of inclusion-maximal feasible sets of alternatives $\mathcal{F}_c$. This assumption is justified given that real-world participatory budgeting problems typically involve up to 20 alternatives [75].

2.4.1 Deterministic Rules

Let $V(\vec{\rho}) \triangleq \{\vec{v} : \vec{v} \rhd \vec{\rho}\}$ denote the set of utility profiles consistent with input profile $\vec{\rho}$. We are interested in computing

$$\arg\min_{S\in\mathcal{F}_c} \max_{\vec{v}\in V(\vec{\rho})} \frac{\max_{T\in\mathcal{F}_c} \mathrm{sw}(T,\vec{v})}{\mathrm{sw}(S,\vec{v})} = \arg\min_{S\in\mathcal{F}_c} \max_{T\in\mathcal{F}_c} \left\{ \max_{\vec{v}\in V(\vec{\rho})} \frac{\mathrm{sw}(T,\vec{v})}{\mathrm{sw}(S,\vec{v})} \right\}.$$

An algorithm is self-evident: compute

$$d(\vec{\rho},S,T) \triangleq \max_{\vec{v}\in V(\vec{\rho})} \frac{\mathrm{sw}(T,\vec{v})}{\mathrm{sw}(S,\vec{v})}$$

for every pair $S, T \in \mathcal{F}_c$, and return $\arg\min_{S \in \mathcal{F}_c} \max_{T \in \mathcal{F}_c} d(\vec{\rho}, S, T)$.

For the input methods we study in this paper, we can describe the set of consistent utility profiles $V(\vec{\rho})$ using linear constraints. Observe that $V(\vec{\rho}) = V(\rho_1) \times \cdots \times V(\rho_n)$, where $V(\rho_i) \triangleq \{v \geq 0 : v \triangleright \rho_i\}$ is the set of m-dimensional utility functions consistent with voter i's input ρ_i. It is therefore sufficient to describe each $V(\rho_i)$ using linear constraints.

For a ranking by value σ_i, we use:

$$V(\sigma_i) = \left\{ v_i \in \mathbb{R}_+^m \;\middle|\; \begin{array}{l} \sum_{a \in A} v_i(a) = 1, \\ v_i(\sigma_i^{-1}(k)) \geq v_i(\sigma_i^{-1}(k+1)), \;\; \forall k \in [m-1] \end{array} \right\}.$$

For a ranking by value for money σ_i, we use:

$$V(\sigma_i) = \left\{ v_i \in \mathbb{R}_+^m \;\middle|\; \begin{array}{l} sum_{a \in A} v_i(a) = 1, \\ \frac{v_i(\sigma_i^{-1}(k))}{c_{\sigma_i^{-1}(k)}} \geq \frac{v_i(\sigma_i^{-1}(k+1))}{c_{\sigma_i^{-1}(k+1)}}, \forall k \in [m-1] \end{array} \right\}.$$

For a knapsack vote κ_i, we use:

$$V(\kappa_i) = \left\{ v_i \in \mathbb{R}_+^m \;\middle|\; \begin{array}{l} \sum_{a \in A} v_i(a) = 1 \\ \sum_{a \in \kappa_i} v_i(a) \geq \sum_{a \in S} v_i(a), \;\; \forall S \in \mathcal{F}_c \end{array} \right\}.$$

For a threshold approval vote τ_i elicited using threshold t, we use:

$$V(\tau_i) = \left\{ v_i \in \mathbb{R}_+^m \;\middle|\; \begin{array}{l} \sum_{a \in A} v_i(a) = 1, \\ v_i(a) \geq t, \;\; \forall a \in \tau_i, \\ v_i(a) \leq t, \;\; \forall a \in A \setminus \tau_i \end{array} \right\}.$$

Note that the polytope for knapsack votes has exponentially many constraints, while the other polytopes have a polynomial number. When necessary, heuristics may be devised to approximate $V(\kappa_i)$, however, in our experiments with real data we only encountered instances with $m \leq 20$, where it was possible to enumerate the constraints in $V(\kappa_i)$.

This polytope $V(\rho_i)$ is the only part of our generic algorithm that is dependent on the input format. Generically, let $A(\vec{\rho})\ \vec{v} \leq b(\vec{\rho})$ be the set of linear constraints describing $V(\vec{\rho})$.

Our next goal is to use this characterization of $V(\vec{\rho})$ to compute $d(\vec{\rho}, S, T)$ for specific $S, T \in \mathcal{F}_c$. Recall that $\mathrm{sw}(S, \vec{v}) \triangleq \sum_{a \in A} x^S(a) \sum_{i \in [n]} v_i(a)$, where x^S is the characteristic vector for the set of alternatives S, and that

$d(\vec{\rho}, S, T) \triangleq \max\{\frac{\mathrm{sw}(T,\vec{v})}{\mathrm{sw}(S,\vec{v})} : A(\vec{\rho})\,\vec{v} \leq b(\vec{\rho})\}$. This is a standard *linear-fractional program*, which can be converted to a linear program $LP(\vec{\rho}, S, T)$ using the famous Charnes-Cooper transformation [49].

The complete algorithm for resolving the deterministic optimal aggregation rule on an input profile $\vec{\rho}$ is given as Algorithm 1.

Algorithm 1: Computing the worst-case optimal deterministic rule

Data: Input profile $\vec{\rho}$
Result: A set $S \in \mathcal{F}_c$ yielding the least distortion

```
dist[S] = 0, ∀S ∈ F_c
for S ∈ F_c do
    for T ∈ F_c, T ≠ S do
        dist[S] = max(dist[S], LP(ρ, S, T))
return arg min_{S∈F_c} dist[S]
```

2.4.2 Randomized Rules

Using a similar line of argument as before, observe that the optimal randomized aggregation rule returns the following distribution p over feasible sets of alternatives:

$$\operatorname*{arg\,min}_{p \in \Delta(\mathcal{F}_c)} \max_{T \in \mathcal{F}_c} \max_{\vec{v} \in V(\vec{\rho})} \frac{\mathrm{sw}(T, \vec{v})}{\sum_{S \in \mathcal{F}_c} p(S) \cdot \mathrm{sw}(S, \vec{v})}.$$

We introduce an additional continuous variable z representing the optimal distortion, and reformulate the problem as

$$\begin{aligned} \min_{p,z} \quad & z \\ s.t. \quad & \mathrm{sw}(T, \vec{v}) - z \cdot \sum_{S \in \mathcal{F}_c} p(S) \cdot \mathrm{sw}(S, \vec{v}) \leq 0, \ \forall T \in \mathcal{F}_c, \vec{v} \in V(\vec{\rho}) \\ & p \in \Delta(\mathcal{F}_c). \end{aligned} \tag{2.10}$$

At this point, it is possible to handle the constraints in (2.10) by formulating the problem in terms of the vertices of the polytope $V(\vec{\rho})$. Instead, we turn to a constraint-generation approach.

Our algorithm performs a binary search on z, the optimal distortion. For a fixed value of z, say $\tilde{z}$, an iterative two-stage procedure determines whether there exists a distribution p whose distortion on the input profile $\vec{\rho}$ is at most $\tilde{z}$. If such a distribution p exists, then $\tilde{z}$ serves as an upper bound on the smallest distortion; otherwise, it serves as a lower bound. After adjusting the bounds on the optimal distortion, the value of $\tilde{z}$ is updated as in traditional binary search.

We now describe the iterative two-stage procedure that ascertains the existence of a distribution p with distortion at most $\tilde{z}$. This procedure alternately finds a distribution satisfying a limited subset of the constraints in (2.10), then attempts to add omitted constraints from (2.10) which are violated by the current distribution. At iteration t, a set of constraints defined by $\mathcal{C}_{t-1}$ have been added and we check the feasibility of

$$\left.\begin{array}{l}\mathrm{sw}(T,\vec{v}) - \tilde{z}\cdot\sum_{S\in\mathcal{F}_c} p^t(S)\cdot\mathrm{sw}(S,\vec{v}) \le 0,\ \forall(\vec{v},T)\in\mathcal{C}_{t-1}\\ p^t\in\Delta(\mathcal{F}_c)\end{array}\right\}\ \mathrm{CF}(\tilde{z},\mathcal{C}_{t-1}).$$

If no feasible distribution p^t exists, $\tilde{z}$ is the new lower bound on the optimal distortion, and we proceed to the next step in our binary search over z. If a feasible p^t exists, we check whether it violates any constraint from (2.10) by solving the following linear program (which serves as an oracle) for every $T \in \mathcal{F}_c$:

$$\left.\begin{array}{ll}\max & \mathrm{sw}(T,\vec{v}) - z\cdot\displaystyle\sum_{S\in\mathcal{F}_c} p^t(S)\cdot\mathrm{sw}(S,\vec{v})\\ & s.t.\ \ \vec{v}\in V(\vec{\rho})\end{array}\right\}\mathrm{LP}(T,z,p^t,\vec{\rho}).$$

If the objective value of $\mathrm{LP}(T,z,p^t,\vec{\rho})$ exceeds 0, a violated constraint is found and $(\vec{v}^*,T)$ is added to $\mathcal{C}_{t-1}$ to form $\mathcal{C}_t$, where $\vec{v}^*$ is the optimal solution to $\mathrm{LP}(T,z,p^t,\vec{\rho})$. The algorithm then returns to solving $\mathrm{CF}(\tilde{z},\mathcal{C}_t)$. If no violated constraints are found, the current distribution p^t indeed has distortion at most $\tilde{z}$, and establishes an upper bound on the optimal distortion.

This complete procedure is summarized in Algorithm 2. A finite number of violated constraints can be added for each $\tilde{z}$, so we may conclude that Algorithm 2 will terminate.

2.4.3 Scaleability of computing distortion-minimizing sets

We evaluate the practicality of this approach by comparing the running times of computing deterministic voting rules, averaged over 10 trials, on data from

Algorithm 2: Computing the optimal randomized aggregation rule

Data: Input profile $\vec{\rho}$, tolerance TOL, $\mathcal{F}_c$

Result: A probability distribution in $\Delta(\mathcal{F}_c)$, the optimal distortion

```
z⁻ = 1, z⁺ = 100, z̃ = (z⁻ + z⁺)/2
while z⁻ − z⁺ > TOL do
    C_0 = ∅, t = 0
    robustFeasibleFlag ← false
    while robustFeasibleFlag is false do
        robustFeasibleFlag ← true
        t ← t + 1
        if CF(z̃, C_{t−1}) is feasible then
            p^t ← optimal solution of CF(z̃, C_{t−1})
            for T ∈ F_c do
                C_t = C_{t−1}
                if optimum of LP(T, z, p^t, ρ⃗) exceeds 0 then
                    v⃗* ← optimal solution of LP(T, z, p^t, ρ⃗)
                    C_t ← C_t ∪ (v⃗*, T)
                    robustFeasibleFlag ← false
            if robustFeasibleFlag then
                z⁺ = z̃
        else
            z⁻ = z̃
    z̃ = (z⁺ + z⁻)/2
return p^t, z⁺
```

participatory budgeting elections held in Boston in 2016. Voters were asked to choose from 10 alternatives; 4,430 votes were cast.

Our discussions with officials from several cities have revealed a hesitance to use randomized voting rules, so we are particularly interested in the performance of the determinstic rules. (Our computational results in the following section also show that deterministic worst case rules typically perform better in terms of average welfare ratio than randomized rules.)

Figure 2.1 summarizes the average time to compute the deterministic worst-case optimal set of alternatives on a log-log scale. The experiments

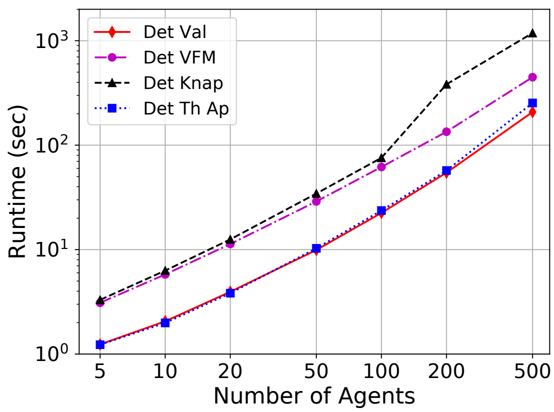


Figure 2.1: Average running time of the deterministic voting rules on the Boston 2016 dataset.

were run on an 8-core Intel(R) Xeon(R) CPU with 2.27GHz processor speed and 50GB memory. We observe that the running time scales gracefully with the number of agents. When sampling 500 voters, computing the deterministic distortion minimizing set for threshold approval votes and rankings by value takes less than 5 minutes, indicating the practicality of these methods for the participatory budgeting elections at the scale of those in Boston, MA [75]. We also note that, due to the once-off nature of participatory budgeting elections, it is conceivable to use an aggregation algorithm which takes several days or even weeks to compute the optimal set of alternatives.

2.5 Empirical Results

Our theoretical results in Section 2.3 characterize how well we can optimize distortion on an observed input profile. Recall that distortion is the *worst-case* ratio of the optimal social welfare to the social welfare achieved, where the worst case is taken over all utility profiles consistent with the observed input profile. In practice we care about this ratio according to the *actual* underlying utility profile. Thus, a distortion-minimizing aggregation rule is not guaranteed to be optimal in practice. This is why an empirical study is called for.

In this section, we compare the performance of different approaches to participatory budgeting, where the performance is measured by the *average-*

case ratio of the optimal and achieved social welfare, and the average is taken over utility profiles drawn to be consistent with input profiles from two real-world participatory budgeting elections.

Datasets: We use data from participatory budgeting elections held in 2015 and 2016 in Boston, Massachusetts. Both elections offered voters 10 alternatives. The 2015 dataset contains 2600 4-approval votes (voters were asked to approve their four most preferred alternatives) and the 2016 dataset contains 4430 knapsack votes.

For each dataset, we conduct three independent trials. In each trial, we create r sub-profiles, each consisting of n voters drawn at random from the population. For each sub-profile, we draw k random utility profiles $\vec{v}$ consistent with the sub-profile, and use these to analyze the performance of different approaches. We use the real costs of the projects throughout. The choices of parameters (r, n, k) for the three trials are $(5, 10, 10)$, $(8, 7, 10)$, and $(10, 5, 10)$. We choose this experimental design to yield sufficiently many samples to verify statistical significance of the results while completing in a reasonable amount of time.

Approaches: We use the utility profile $\vec{v}$ drawn to create an input profile in four input formats we study. For each format, we use the deterministic as well as randomized distortion-minimizing aggregation rule. The non-trivial algorithms we devise for these rules are presented in Section 2.4. These eight approaches are referred to using the type of aggregation rule used ("Det" or "Ran"), and the type of input format ("Knap", "Val", "VFM", or "Th Ap").

Unlike the other input formats, threshold approval votes are technically a family of input formats, one for each value of the threshold. While randomizing over the threshold is required to minimize the distortion (the *worst-case* ratio of the optimal and achieved social welfare), as is our goal in the theoretical results of Section 2.3, minimizing the *expected* ratio of the two can be achieved by a deterministic threshold. In our experiments, we learn the optimal threshold value based on a holdout set that is not subsequently used. This learning approach is practical as it only uses observed input votes rather than underlying actual utilities. This choice likely gives threshold approval votes an edge — but arguably it is an advantage this input format would also enjoy in practice.

In addition to our eight approaches, we also test two approaches used in real-world elections [75]: greedy 4-approval ("Gr 4-Ap"), and greedy knapsack ("Gr Knap"). The former elicits 4-approval votes, and greedily selects

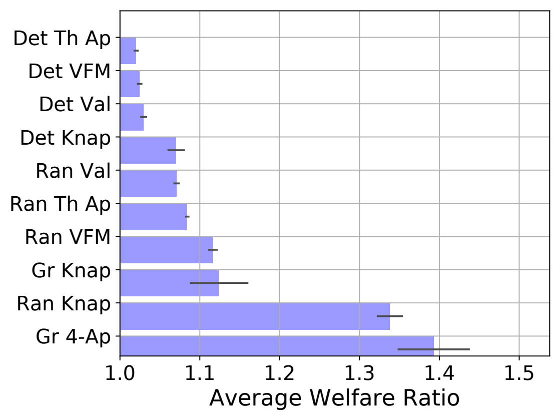


Figure 2.2: Average welfare ratio of different approaches to participatory budgeting based on data from Boston 2015 and 2016 elections (lower is better).

the most widely-approved alternatives until the budget is depleted. The latter is almost identical, except for interpreting a knapsack vote as an approval for each alternative in the knapsack.

As the performance measure for the ten approaches, we use the average ratio of the optimal and the achieved social welfare according to the actual utility profile used to induce the input profiles — termed *average welfare ratio* — where the average is taken across the entire experiment.

Results: Figure 2.2 shows the average welfare ratio of the different approaches with 95% confidence intervals, sorted from best to worst. The differences in performance between all pairs of rules — except between Det Knap and Ran Val, and between Ran VFM and Gr Knap — are statistically significant [83] at a 95% confidence level.

A few comments are in order. First, deterministic distortion-minimizing aggregation rules generally outperform their randomized counterparts. This is not entirely unexpected. While randomized rules do achieve better *distortion*, there always exists a deterministic rule minimizing the *average welfare ratio* objective; although, it is not necessarily the deterministic distortion-minimizing aggregation rule.

Second, approaches based on deterministic rules are able to limit the loss in social welfare due to incomplete information about voters' utility functions to only 2%–3%. Among these approaches, the one using threshold approval votes incurs the minimum loss.

Third, knapsack votes consistently lead to higher distortion than alternative input formats. This, together with the poor theoretical guarantees for knapsack votes, suggests that it may not be worthwhile to ask voters to solve their personal $\mathcal{NP}$-hard knapsack problems before casting vote.

2.5.1 Is It Useful to Learn the Threshold?

In our experiments, when using threshold approval votes, we use the threshold that achieves the best performance on a holdout/training set, to evaluate the performance of threshold approval votes on the test set.

After sampling a random utility profile consistent with an input vote in the training set, we generate threshold approval votes for all the thresholds in $[0, 1]$ at intervals of 0.05, and compute the average distortion (per threshold) across multiple samples. (This step will not be required in practice once sufficiently many real votes are elicited).

We select the threshold value that achieves the least average distortion. Importantly, note that we use distortion — which is only a function of the input profile — rather than the average distortion to select the optimal threshold value. Hence, this method is robust, and does not use any knowledge of the distribution of utility profiles that we later use in evaluating performance.

This optimal threshold value when evaluating the performance (average distortion) of threshold approval votes, in conjunction with both the deterministic and the randomized distortion-minimizing aggregation rules.

While threshold approval votes with deterministic aggregation rule achieves excellent performance with this method of threshold selection, it is not immediately clear whether the threshold selection was useful. Indeed, learning a threshold is only useful if the optimal threshold value remains reasonably consistent across the instances. We now investigate the usefulness of threshold selection in multiple ways.

First, Figure 2.3 shows the average distortion achieved by different values of the threshold on the training instances, when used in conjunction with the deterministic and the randomized distortion-minimizing aggregation rules. Recall that the final threshold value we select is the one that minimizes this measure. For every threshold value on the x-axis, the error bars indicate the range that contains the distortion on 95% of the training instances. We do not plot threshold values above 0.4 as the distortion is non-decreasing beyond this point.

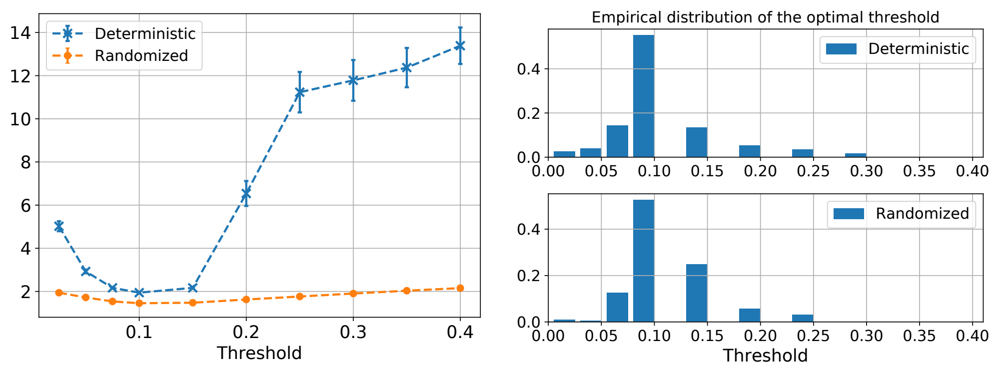


Figure 2.3: (Left) Average distortion achieved by different threshold values in threshold approval votes; (Right) Empirical distribution of the optimal threshold under deterministic and randomized aggregation.

We observe that the thresholds values that lead to the smallest average distortion are exactly those with the smallest variation across instances. Interestingly, the average distortion of different values of the threshold is wildly different under the deterministic aggregation rule, but rather similar under the randomized aggregation rule. This effect perhaps manifests itself in the improved performance of threshold approval votes with deterministic aggregation than with randomized aggregation in all of our experiments; see Figure 2.2.

Next, we measure the usefulness of training the threshold value in a different way. In Figure 2.3 (right) we plot the empirical distribution of the optimal threshold value, i.e., for each threshold value, we plot the percentage of training instances in which that value led to the smallest distortion. For both deterministic and randomized aggregation rules, the distribution of the optimal threshold value is (quite strongly) centered at 0.1. In fact, the optimal threshold value was in $[0.075, 0.15]$ in more than 80% of the training instances.

The consistency with which a single threshold value (0.1) remains the optimal value suggests that learning this value from the holdout set is very likely to be valuable.

Finally, we note that the datasets we used contain votes over 10 alternatives. That is, $m = 10$. Interestingly, this makes the empirically optimal threshold value $1/m$, which is precisely the value for which we achieve the best performance in the worst case in our theoretical results (see Theorem 2.14).

2.6 User study

We conduct experiments based on data collected from more than 1200 voters on Amazon Mechanical Turk. Voters were asked to vote over items to take to a desert island. In the first of our two studies, voters were asked to cast a vote in a single input format, or report utilities. In the second, they were asked to both cast a vote and report utilities, as well as to answer several questions about their subjective experience.

To evaluate whether the different input formats lead to outcomes with high social welfare, we aggregate a sample of votes and evaluate the outcome on a random sample of submitted utility profiles. The aggregation happens by finding the *distortion-minimizing* budget-feasible subset of alternatives as described in Section 2.4. A key insight behind the experimental design is that we can measure the social welfare of an outcome selected by one group of voters using the utilities submitted by a different group, because the *average* utility of each item would be consistent across the groups by the law of large numbers (the effect of which is present at the scale of our experiments).

We find that for most input formats distortion-minimizing aggregation leads to outcomes that are quite close to the welfare-maximizing outcome, even without access to the underlying utility profile. Moreover, we can see significant differences between different input formats, and some really shine. Most impressively, our results indicate that the k-approval and ranking by value for money input formats lead to outcomes that are essentially optimal.

Turning to usability, we consider two types of indicators. Objective indicators, which are computed from data, include consistency and response time. Subjective indicators, which are based on ratings reported by voters, include ease of use, likability, and expressiveness.

Consistency refers to the relation between a voter's utility function and her vote. For example, if we use k-approval as the input format, we expect a voter to approve the k alternatives for which she has the highest utility. If other alternatives are approved, it means that the voter may have misunderstood the instructions, or the cognitive burden imposed by the task was too high to perform it accurately. We find that k-approval by far leads to the highest degree of consistency, followed by threshold approval and knapsack. For response time, we find that k-approval again excels in terms of both time to learn and time to vote. By contrast, ranking by value for money does badly in both objective measures.

Finally, the subjective usability indicators generally favor ranking by

Item	Cost	Utility	Item	Cost	Utility
Mirror	10	5.8	Compass	5	9.4
Top coat	20	2.3	Raincoat	10	5.4
Water	3	29.3	First aid kit	10	14.9
Map	8	9.5	Pistol	30	6.7
Pocket knife	5	14.8	Sunglasses	25	1.9

Table 2.1: The 10 items used in the experiment, their costs, and voters' reported mean utilities. The budget is $65.

value for money, especially in terms of how expressive it is perceived to be by voters. By contrast, k-approval is seen as the least expressive input format.

2.6.1 Experimental Setup

We recruited more than 1200 voters on Amazon Mechanical Turk for our experiments, and asked them to evaluate a hypothetical scenario. Voters were told that they are stranded on a desert island, there is a set of items which may increase their chances of survival, each item has a cost, and there is a budget of $65. The list of items is shown in Table 2.1, along with voters' reported mean utility for each item.

This abstract task is inspired by studies of group decision making [79], and asks for a choice from a set of items, as in participatory budgeting. It was selected to eliminate biases based on voters' locations. For example, if we were to confront voters with a more traditional participatory budgeting setting in which one potential project involves upgrading a park, one may expect voters' utilities to vary drastically based on the health of their city's existing park system. This effect would be missing in real-world participatory budgeting elections, in which voters are typically residents of the same city.

In our experiments, voters are asked to report their preferences over the items in one of the five input formats described in Section 2.2 and/or report their numerical utilities for the different items. Votes in each input format (and utilities) are elicited using a dedicated user interface. Figure 2.4 shows the user interface for knapsack, in which voters use checkboxes to select items. Below, we describe how votes are elicited through each interface.

- *Knapsack vote*: Voters are shown the interface of Figure 2.4. The task is: “You need to select which items to take based on your carrying capacity of 65 pounds.”

- *Ranking by value*: Voters are shown the list of items in a drag-and-drop interface. The task is: “Rank the items from the most important to the least important according to your best judgment.”

- *Ranking by value for money*: Voters are shown the list of items and their weights in a drag-and-drop interface. The task is: “If you had to divide 100 points among the items based on how much you like them, rank the items in the decreasing order of the number of points they would receive divided by the cost.”

- *Threshold approval*: Voters are shown a list of items with checkboxes. The task is: “If you had to divide 100 points among the items based on how much you like them, select all the items that would receive at least 10 points.”

- 5-*approval*: Voters are shown a list of items with checkboxes. The task is: “You need to select up to 5 items from a list of 10 items according to your best judgment.”

- *Utilities*: Voters are shown a list of items and sliders that control the number of points given to each. The task is: “You need to distribute 100 points among 10 items. The more points you assign to an item, the more important you think the item is to your survival.”

We conducted two studies, which we refer to as A and B. In study A, 720 voters were recruited; each voter was randomly assigned one of the above input formats and cast a single vote in this format. This yields 120 votes in each format. The dataset from this study is used in the experiments detailed in Sections 2.6.2 and 2.6.3.

In study B, an additional 500 voters were recruited, and engaged in a two-stage process. In the first stage, half of the voters were asked to vote using one of the five input formats (randomly assigned). In the second stage, these voters were asked to specify their utility for each item. After each step, the voters were asked to rate how easy they found the activity, and how much they liked the user interface. To control for ordering effects, the other half of the voters were asked to perform the two stages in the reverse order (i.e., specify utilities in the first stage, and vote in a given input format in

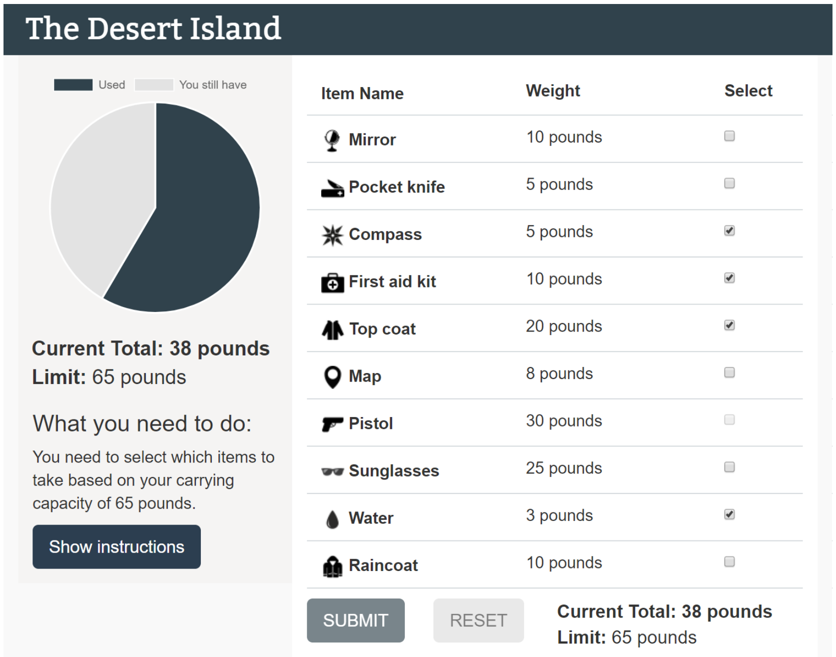


Figure 2.4: Screenshot of the knapsack interface.

the second stage). The dataset from this study is used for the experiments detailed in Sections 2.6.3 and 2.6.3.

In both studies, participation is contingent on voters reading a short tutorial, passing a pre-task quiz which verifies voters' comprehension of the interface, and passing a post-task quiz which asks voters questions about their votes to ensure that the votes received at least some consideration. For example, the post-task quizzes in the ranking by value and ranking by value for money formats ask voters whether top coat was positioned higher than water in their ranking. Voters were paid 20 cents for completing the tutorial and the pre-task quiz, and a bonus of 10 cents for completing the post-task quiz.

2.6.2 Efficiency

Given the underlying utility functions $\vec{v}$ of the voters, our goal is to choose an optimal (welfare-maximizing) budget-feasible set of alternatives:

$$S^* \in \arg\max\{\mathrm{sw}(S, \vec{v}) : S \in \mathcal{F}_c\},$$

where

$$\mathcal{F}_c \triangleq \{S \subseteq A : c(S) \leq B\}$$

is the collection of all budget-feasible sets of alternatives. When choosing a suboptimal set $S \in \mathcal{F}_c$, we face an efficiency loss defined as

$$\mathrm{EL}(S, \vec{v}) \triangleq 1 - \frac{\mathrm{sw}(S, \vec{v})}{\max_{T \in \mathcal{F}_c} \mathrm{sw}(T, \vec{v})}.$$

In words, an efficiency loss of 0.05 (5%) means that the set chosen achieves 95% of the optimal welfare.

When votes are cast in an input format, we have access only to the votes $\vec{\rho}$, and not to the utility profile $\vec{v}$. While $\vec{\rho}$ provides partial information regarding $\vec{v}$ (specifically, that $\vec{v} \rhd \vec{\rho}$), some efficiency loss is inevitable.

As before, we advocate for using distortion-minimizing aggregation rules, which in the notation introduced here means finding the deterministic aggregation rule f^* which returns

$$f^*(\vec{\rho}) \in \operatorname*{arg\,min}_{S \in \mathcal{F}_c} \sup_{\vec{v} : \vec{v} \rhd \vec{\rho}} \mathrm{EL}(S, \vec{v}).$$

In our efficiency experiment, we want to evaluate and compare the efficiency loss of the distortion-minimizing set chosen based on votes in each input format. Instead of evaluating the efficiency loss in the worst case, we want to evaluate it using the underlying utility profile. Specifically, we take the dataset from study A, sample 60 voters for each input format, compute the distortion-minimizing set for the corresponding vote profile, and evaluate its efficiency loss using the utility profile of another sample of 60 voters who were asked to submit their utility functions.

A crucial insight behind this experiment, which is necessary for its validity, is that the *average* utility of an item, according to the utility profile of the second set of voters, closely approximates its average utility according to the first set of voters. This is intuitively true by the law of large numbers, and is confirmed by our experiments in Section 2.6.3. For this reason, we can accurately estimate the social welfare of a subset of items with respect to the first set of voters, *without* asking these voters to report both utilities and votes.

Figure 2.5 reports the average efficiency loss (in percent) across 1 000 repetitions of this experiment. The Mann-Whitney U test found a statistically significant difference in performance (at the $p = 0.05$ level) between every

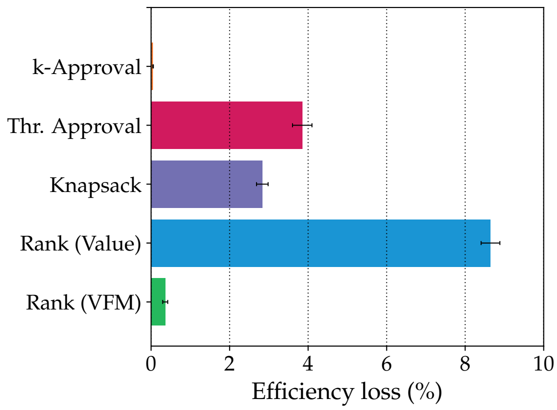


Figure 2.5: The average efficiency loss for each input format. Lower is better.

pair of input formats except between k-approval and ranking by value for money. Both k-approval and ranking by value for money perform incredibly well and achieve social welfare within 0.5% of optimal, suggesting that they capture sufficient information about voter preferences to allow computation of near-efficient outcomes. The worst performance is demonstrated by ranking by value, which incurs an 8% efficiency loss on average.

2.6.3 Usability

For an input format to be viable for deployment in participatory budgeting elections, we expect it to allow voters to accurately and quickly express their preferences, while also being easy to understand and use. To that end, we measure the usability of an input format through both *objective* and *subjective* indicators. While the objective indicators of usability are computed from data, the subjective indicators are self-reported by the voters.

We focus on two objective indicators. First, we want to ensure that the votes cast by voters in an input format are consistent with the utility functions expressed by the (same or different) voters. We call this *consistency*. Second, we record the amount of time it takes for voters to complete the tutorial and cast their vote, which is an indicator of the cognitive burden. We call this *response time*.

We additionally ask voters about their experience of casting a vote in their assigned input format, and record three subjective indicators of usability:

how easy it is to cast a vote, how much they like the user interface, and how well the input format allows them to express their preferences.

Objective Indicators

As noted earlier, we measure two objective indicators of usability: consistency between votes and utility functions, and time taken by voters.

Consistency. Intuitively, consistency measures whether voters' reported utility functions induce their votes cast in a given input format. If an input format allows voters to accurately express their preferences, we may expect a high level of consistency. We measure consistency in two forms.

For *internal consistency*, we call a voter consistent if the voter's reported utility function is consistent with that *same* voter's vote in the assigned input format (i.e., the utility function induces the vote, up to any ties). For each input format, we report the percentage of consistent voters.

Recall that we chose the desert island setting with the assumption that it minimizes the effect of voters' contextual background. If this assumption holds, we should expect consistency even between the votes and the utility functions reported by different sets of voters. We refer to this as *external consistency*. For this, we use data from study A, and from the first stage of study B. For each input format, the submitted votes form a vote profile $\vec{\rho}$, and the submitted utilities functions form a utility profile $\vec{v}$. We measure the fraction of votes induced by $\vec{v}$ that match with votes in $\vec{\rho}$. Formally, to account for ties, we create a bipartite graph with votes from $\vec{\rho}$ on one side and utility functions from $\vec{v}$ on the other, and add an edge between vote ρ_i and utility function v_j when $v_j \triangleright \rho_i$. We report the percentage of matched votes, or the cardinality of the maximum matching divided by 170 (the number of vertices on each side).

The results are provided in Figure 2.6. k-approval is comfortably the best in terms of both internal and external consistency (both above 50%). We find the internal consistency of knapsack to be surprisingly high: more than a third of the voters can report exact solutions to their personal knapsack problem, the computational hardness of the knapsack problem and the sheer number of budget-feasible subsets of alternatives notwithstanding. Ranking by value for money and ranking by value perform poorly in both forms of consistency. It is tempting to claim that this is due to the space of possible rankings being exponentially large, but as noted above, knapsack votes

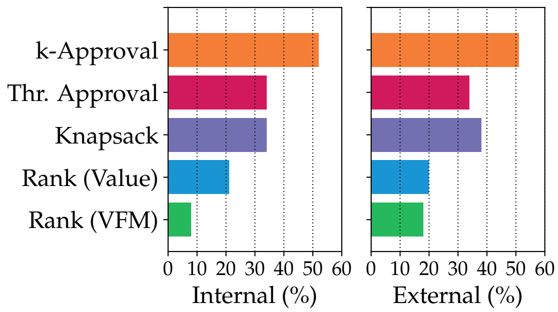


Figure 2.6: Internal and external consistency of different input formats. Higher is better.

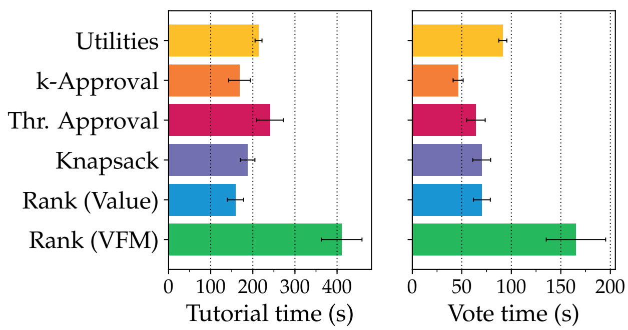


Figure 2.7: Average time taken (in seconds) to complete the pre-task tutorial and to cast a vote in each input format. Lower is better.

perform well despite this obstacle.

Finally, we remark that the high degree of similarity between internal and external consistency for each input format is yet another strong indication that the utility profile of one set of voters serves as a good substitute for the utility profile of another set of voters, which is a foundational assumption for the validity of our between-user study.

Response time. The response time to complete a task is recognized as a proxy for the objective difficulty (or cognitive load) associated with the task [112]. For each input format, we report, in Figure 2.7, the average amount of time it took to learn how to vote in the format (complete the tutorial and pass the quiz) and to cast a vote in the format.

In terms of the difficulty of learning an input format, k-approval and ranking by value are the easiest (the difference between them is not statistically significant), followed by knapsack and threshold approval. Ranking by value for money is the most difficult by a wide margin.

In terms of the time taken to cast a vote, k-approval is also by far the fastest input format, at 45 seconds on average. Knapsack and ranking by value take about 70 seconds, while ranking by value for money is again the slowest by a wide margin, at almost 3 minutes.

For reference, we also report how long it takes for voters to submit their utility functions. At 96 seconds, this is slower than every input format except ranking by value for money. This largely supports the belief that it is taxing for voters to report their exact utility functions.

Summary. The objective indicators of usability overwhelmingly point to k-approval. It is distinctively the best at allowing voters to quickly learn the format and cast a vote, and results in votes that are by far the most consistent with the voters' utility functions. By contrast, ranking by value for money performs miserably. It takes voters more than three times longer to vote using this format than under k-approval, and the resulting votes have little in common with the voters' utility functions.

Subjective Indicators

In addition to computing objective indicators of usability, we asked 500 voters in study B to report their experiences with different input formats, and measured various subjective indicators of usability. When we say below that a result is statistically significant, we are referring to the Mann-Whitney and Wilcoxon signed-rank tests at the $p < 0.05$ level.

Ease of use. We asked voters to report how easy they found the voting task on a scale of 0 to 5 (5 being the easiest). The perceived (subjective) difficulty is reported in Figure 2.8(a). Ranking by value for money is significantly worse than every other input format, while the differences between the other input formats are not statistically significant.

User interface. We also asked voters to report how much they liked the user interface on a scale of 0 to 5 (5 being the most liked). As seen in

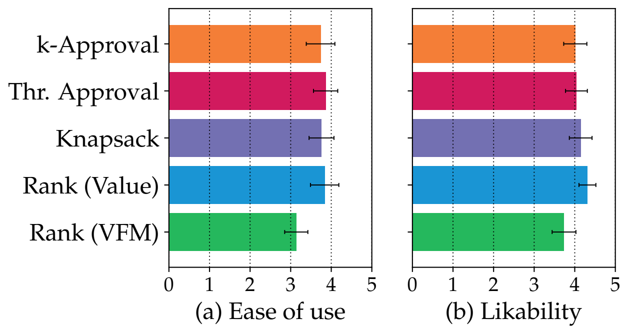


Figure 2.8: How easy to use each input format is, and how liked its user interface is, based on the subjective reports of the voters on a scale of 0 to 5, 5 being the best.

Figure 2.8(b), ranking by value and knapsack are the most liked interfaces, followed by k-approval and threshold approval (with no significant difference between each pair). Ranking by value for money was again the least liked.

We believe that this is somewhat correlated with the inherent difficulty of an input format because our choice of user interface was standard in most cases. However, the results are subject to change with design of better user interfaces.

Perceived Expressiveness. We asked voters to report how well their assigned input format captured their preferences on a scale of 0 to 5. As seen in Figure 2.9, ranking by value is reported to be much more expressive than any other input format (by a statistically significant margin), while k-approval and threshold approval votes are the least expressive. Although voters dislike using ranking by value for money, they still feel that it captures their preferences well.

Summary. In terms of subjective indicators, ranking by value seems the most preferred input format: voters feel that it best captures their preferences, and no other input format is more easy to use or liked (in a statistically significant manner). Ranking by value for money is again the most difficult to use and least liked, although voters feel it captures their preferences fairly well.

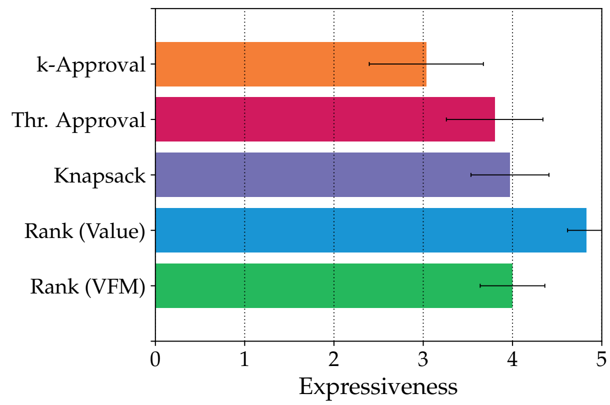


Figure 2.9: Voters' perceived expressiveness of different input formats. Higher is better.

2.6.4 Discussion of user study

Our results shed light on the efficiency and usability of five input formats used in participatory budgeting. Somewhat surprisingly, the most popular — and arguably the simplest — input format, k-approval, outperforms every other input format in terms of efficiency (welfare loss) and objective indicators of usability (consistency of votes and response time). In terms of the subjective indicators, no input format is statistically easier to use than k-approval, while the user interfaces of ranking by value and knapsack are only somewhat more liked than that of k-approval.

The results for the third subjective indicator, namely expressiveness, are the only ones that prevent k-approval from being dominant across the board. Indeed, voters feel that k-approval is the worst in capturing their preferences, while ranking by value is the best. Our efficiency experiments reveal that, in fact, the exact opposite is true: k-approval contains information that leads to the most efficient outcomes, while ranking by value leads to the least efficient ones. This highlights the distinction between what voters feel is important when casting a vote, and what is needed to enable efficient aggregation.

Ranking by value performs well in terms of subjective indicators of usability, and somewhat worse in terms of the objective indicators. However, it is especially concerning that it leads to outcomes that have relatively low social welfare.

Knapsack performs reasonably on all indicators, including surprisingly good response times, which corroborate the results of [75]. Based on our discussions with practitioners in Europe, it seems that the fact that this input format encourages voters to directly reason about the government's budget constraints is also seen as an advantage, which could potentially outweigh some of the disadvantages shown by our results.

The subjective and objective indicators agree that voters find ranking by value for money to be difficult to use and that these votes rarely reflect voters' true utility functions (but mysteriously lead to efficient outcomes). This cautions strongly against the use of ranking by value for money, although it is less clear what the implications are for pairwise value for money comparisons as advocated by [75].

Finally, we acknowledge several limitations of our study, and point to directions for future work. First, our efficiency results use the distortion-minimizing aggregation method for each input format. While this provides a consistent choice across input formats, it would be interesting to use more realistic (e.g., greedy) aggregation methods to better understand the efficiency loss in practice. Second, our results are closely tied to our choice of user interfaces for eliciting voter preferences. Arguably, a better user interface can lead to increased measures of usability, including votes that are more consistent with voters' utility functions, which in turn can lead to greater efficiency. Hence, the design of improved, more intuitive user interfaces is an important direction for future research.

Next, in all of our experiments, except in the measurement of internal consistency, we only used data generated by asking voters to vote in a single input format. This choice was based on the assumption that asking voters to vote in multiple formats would not only be tiring, but can also affect the votes themselves. This was partially confirmed by our measurements of internal consistency. We observed that if we ask voters to report their utility functions and cast their votes using an input format, voters are generally far less consistent when utility functions are reported first. However, there is a need for more thorough experiments to identify and understand the effects of asking voters to report their preferences in multiple forms.

We note that our desert island setting uses 10 items (alternatives), while real participatory budgeting elections may require voters to compare more items. We limited the number of items to allow voters to accurately report their utility functions, which was necessary to measure consistency and efficiency loss. An important direction for future work is to study voter behavior when evaluating more than 10 items, which may require indirectly measuring consistency and efficiency loss without access to the utility functions.

More broadly, while our desert island setting provides a good abstraction of participatory budgeting and reduces the effect of voters' contextual background, it makes the voters a bit too homogeneous. In our setting, it is likely that *all* voters have similar preferences. By contrast, in participatory budgeting, it is likely that voter preferences are clustered based on factors such as personal interests and geographical location. Studying the structure of voter preferences and its effect on the choice of efficient outcomes in real participatory budgeting elections is perhaps the most compelling direction for future research.

2.7 Discussion

Our theoretical results indicate that threshold approval votes should receive serious consideration as the input format of choice for participatory budgeting. However, our user study does not point to one format being clearly superior, although it does lead to useful conclusions, for example, rankings by value for money is extremely difficult to use. We expect that further human experiments will be required to determine which input formats strike thebest balance between being user-friendly and leading to efficient outcomes.

Whatever the best approach to participatory budgeting is, now is the time to identify it, before various heuristics become hopelessly ingrained. We believe that this is a grand challenge for computational social choice, especially at a point in the field's evolution where it is gaining real-world relevance by helping people make decisions in practice.

Chapter 3

Low-distortion rankings

3.1 Introduction

Classic social choice theory typically approaches the design of voting rules from an axiomatic viewpoint, that is, researchers formulate attractive properties, and ask whether there are voting rules that satisfy them. By contrast, research in *computational social choice* [35] is often guided by optimization, in that researchers specify quantitative measures of the desirability of different alternatives, and construct voting rules that optimize them.

One such approach is known as *implicit utilitarian voting* [31, 46, 106]. In a nutshell, the idea is that each voter i has a utility function u_i that assigns a value to each alternative. However, these utility functions are *implicit*, in the sense they cannot be communicated by the voters (because they are unknown or difficult to pin down). Instead, voters report rankings of the alternatives that are *consistent* with the underlying utility functions, that is, each voter sorts the alternatives in non-increasing order of utility. The goal is to choose an alternative a that maximizes *(utilitarian) social welfare* — the sum of utilities $\sum_i u_i(a)$ — using the reported rankings as a proxy for the latent utility functions.

From that viewpoint, the best voting rule is the one that minimizes a measure called *distortion*, defined by Procaccia and Rosenschein [106] as the ratio between the social welfare of the best alternative, and the social welfare of the alternative selected by the rule, in the worst case over all utility functions that are consistent with the observed rankings. Put another way, this is the approximation ratio to the welfare-maximizing solution, and the need for approximation stems from lack of information about the true utilities.

In recent years, implicit utilitarian voting has emerged as a practical approach for helping groups of people make joint decisions. In particular, optimal voting rules, based on the implementation of Caragiannis et al. [46], are deployed on the not-for-profit website `RoboVote.org` for the case where the desired output is a single alternative or a subset of winning alternatives.

However, RoboVote also has a third output type, a ranking of the alternatives, and for this case the website currently does *not* take the same approach — instead it uses the well-known Kemeny rule [55, 59]. Indeed, it is unclear how to even view the problem of returning a ranking through the lens of implicit utilitarian voting — if a voter has a utility for each alternative, what is his utility for a ranking of the alternatives? One could assume that a voter i has a *weight* $w_{i,j}$ for each *position* j, so his utility for the ranking $(a_1, a_2, \ldots, a_m)$ would be $\sum_{j=1}^{m} w_{i,j} u_i(a_j)$; but any particular choice of weights would be *ad hoc*.

3.1.1 Our Approach and Results

The insight underlying our approach is that the worst-case perspective also extends to the choice of weights. That is, when we measure the social welfare of an output ranking given reported input rankings, we consider the worst case over *both* utility functions and weights.

Of course, this is a very conservative approach, and one might worry that it would lead to massive distortion. But our main theoretical result is that the distortion of optimal voting rules is asymptotically identical to the case where only a single alternative is selected (and there are no weights whatsoever), up to a polylogarithmic factor — in both cases it is $\tilde{\Theta}(\sqrt{m})$, where m is the number of alternatives.

In fact, we establish a significantly stronger result, as we allow voters to have *combinatorial* utility functions over subsets of alternatives, and we measure the utility of a voter for a ranking as the weighted sum of his utilities for prefixes of that ranking; the foregoing distortion bound holds when the utility functions are monotonic and subadditive. We find it striking that it is possible to formulate the problem in such generality with no tangible increase in distortion.

Our computational results demonstrate that it is practical to compute deterministic distortion-minimizing rankings for instances with up to 10 alternatives. This constraint on the instance size is not unreasonable, as 98.3% of RoboVote instances have 10 or fewer alternatives. For larger instances we test several heuristics and find that the Borda and Kemeny rules typically lead to low distortion and near-optimal social welfare.

3.1.2 Related Work

Generally speaking, the implicit utilitarian voting literature can be partitioned into two complementary strands of research. One does not constrain the structure of voters' utility functions [20, 31, 41, 46, 106]. The other (which is more recent) assumes that utility functions are derived from an underlying metric space, naturally leading to smaller distortion [7, 9, 67, 76, 77]. Our setup is consistent with the former line of work.

On a technical level, two of the foregoing papers are most closely related to ours. The first is by Boutilier et al. [31], who study the distortion minimization problem when the output is a distribution over winning alternatives. They prove an upper bound of $\mathcal{O}\left(\sqrt{m} \cdot \log^* m\right)$ on the distortion of optimal voting rules, and a lower bound of $\Omega(\sqrt{m})$. Their setting coincides with ours when $w_{i,1} = 1$ for each voter i, because in that case social welfare depends only on the utility of each voter for the top-ranked alternative. Achieving low distortion is much more difficult in our setting, and, in particular, their lower bound directly carries over (whereas their upper bound clearly does not).

The second paper, by Benadè et al. [20], studies distortion-minimizing rules for the participatory budgeting problem, where each alternative has a cost, and the goal is to choose a subset of alternatives that satisfies a budget constraint. Voters are assumed to have *additive* utility functions. Their results are incomparable to ours — their problem is "harder" in that they have to deal with (known) costs and budget constraints, but "easier" in that they choose a single subset, whereas we, in a sense, choose m nested subsets (the m prefixes of our ranking), which are weighted according to *unknown* weights. Furthermore, our results hold for richer (subadditive) combinatorial utility functions.

3.2 The Model

Our setting involves a set of voters $[n] = \{1, \ldots, n\}$, and a set of alternatives $[m] = \{1, \ldots, m\}$. We are interested in the set $\mathcal{S}_m$ of rankings, or permutations, over $[m]$. We think of a ranking $\tau \in S_m$ as a function from positions to alternatives, i.e., $\tau(j)$ is the alternative in position j in τ, and $\tau^{-1}(j)$ is the position in which τ places alternative j.

The preferences of each voter i are represented as a ranking $\sigma_i \in \mathcal{S}_m$. A *preference profile* is a vector $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ of the rankings of all voters.

A (randomized) *social choice function* is a function $f : (\mathcal{S}_m)^n \to \Delta([m])$, which takes a preference profile as input, and returns a distribution over winning alternatives. In this paper we focus on (randomized) *social welfare functions*, whose range is instead $\Delta(\mathcal{S}_m)$, i.e., they also take a preference profile as input, but return a distribution over rankings.

A novel component of our model is that we assume that each voter $i \in [n]$ is associated with a *combinatorial* utility function $u_i : 2^{[m]} \to \mathbb{R}_+$ and a *weight vector* $w_i \in \mathbb{R}^m_+$. Following previous work [41, 31, 46, 20], both are assumed to be *normalized*, that is, for all $i \in [n]$, $u_i(\emptyset) = 0$ and $\sum_{j=1}^m u_i(\{j\}) = \sum_{j=1}^m w_{i,j} = 1$. Moreover, our results make use of the following properties of utility functions:

- *Monotonicity:* $u_i(S) \leq u_i(T)$ for all $S \subseteq T \subseteq [m]$.
- *Subadditivity:* $u_i(S) + u_i(T) \geq u_i(S \cup T)$ for all $S, T \subseteq [m]$.

The utility of voter i for a ranking $\tau \in \mathcal{S}_m$ is given by the weighted sum of his utilities for the prefixes of τ, that is,

$$u_i(\tau) = \sum_{j=1}^{m} w_{i,j} \cdot u_i(\{\tau(1), \tau(2), \ldots, \tau(j)\}).$$

We remark that even *additive* utility functions are able to capture the simpler setting discussed in Section 3.1, which can be formalized by assigning each voter a utility function $u'_i : [m] \to \mathbb{R}_+$ and weights such that $w'_{i,j} \geq w'_{i,j+1}$ for all $j \in [m-1]$, and letting $u'_i(\tau) = \sum_{j=1}^m w'_{i,j} \cdot u'_i(\tau(j))$.

We assume that each voter reports a ranking that is *consistent* with his utility function, which, in our general formulation with combinatorial utilities, we take to mean that voter i reports σ_i only if

$$u_i(\{\sigma_i(1)\}) \geq u_i(\{\sigma_i(2)\}) \geq \cdots \geq u_i(\{\sigma_i(m)\}).$$

We denote this notion of consistency by $u_i \triangleright \sigma_i$, and, when σ_i is consistent with u_i for all $i \in [n]$, $\vec{u} \triangleright \vec{\sigma}$.

Our goal is to optimize (utilitarian) *social welfare*, that is, the sum of utilities voters have for the output ranking. Formally,

$$\mathrm{sw}(\tau) \triangleq \sum_{i=1}^{n} u_i(\tau).$$

However, since we only observe the given preference profile, we cannot directly optimize social welfare. To measure how far a social welfare function is from maximizing this objective, we adapt the concept of *distortion* [106]. Formally, the distortion of a social welfare function f on a preference profile $\vec{\sigma}$ is

$$\mathsf{dist}(f, \vec{\sigma}) \triangleq \max_{\vec{u}\colon\ \vec{u} \triangleright \vec{\sigma}} \max_{\vec{w}} \frac{\max_{\tau \in \mathcal{S}_m} \mathrm{sw}(\tau)}{\mathbb{E}_{\mu \sim f(\vec{\sigma})}[\mathrm{sw}(\mu)]}.$$

In words, distortion measures the ratio between the social welfare of the welfare-maximizing ranking, and the expected social welfare of the distribution over rankings produced by f, in the worst case over all possible weights $\vec{w} = (w_{i,j})_{i \in [n], j \in [m]}$, and all possible utility profiles that are consistent with the given preference profile. Finally, the distortion of f is the worst case distortion over all possible preference profiles: $\mathsf{dist}(f) \triangleq \max_{\vec{\sigma}} \mathsf{dist}(f, \vec{\sigma})$.

3.3 Distortion Bound

In this section we establish a tight (up to polylogarithmic factors) bound on the distortion of optimal social welfare functions. As noted in Section 3.1.2, Boutilier et al. [31] prove a lower bound of $\Omega(\sqrt{m})$ on the distortion of optimal social *choice* functions, which carries over to our setting. Therefore, to show that optimal social welfare functions have distortion $\tilde{\Theta}(\sqrt{m})$, it is sufficient to prove the following theorem, which is our main result.

Theorem 3.1. *Under the monotonicity and subadditivity assumptions, there exists a randomized social welfare function with distortion* $\mathcal{O}\left(\sqrt{m}\ln^{3/2} m\right)$.

The construction of our social welfare function relies on the *harmonic scoring function* [31], defined as follows. Recall that $\sigma_i^{-1}(j)$ denotes the position of alternative j in the ranking of voter i. The harmonic score of alternative j is $\mathsf{score}(j) \triangleq \sum_{i=1}^{n} 1/\sigma_i^{-1}(j)$.

We will make use of the following two properties of the harmonic scoring function.

Lemma 3.2. *For any $m \geq 2$, $\sum_{j=1}^{m} \mathsf{score}(j) \leq 3n \ln m$.*

Proof of Lemma 3.2. By definition,

$$\sum_{j=1}^{m} \mathsf{score}(j) = \sum_{j=1}^{m} \sum_{i=1}^{n} 1/\sigma_i^{-1}(j) = \sum_{i=1}^{n} \sum_{j=1}^{m} 1/j$$
$$\leq n(\ln m + 1) \leq 3n \ln m.$$

□

Lemma 3.3. *Under the subadditivity assumption, for any $S \subseteq [m]$ it holds that $\sum_{i=1}^{n} u_i(S) \leq \sum_{j \in S} \mathsf{score}(j)$.*

Proof of Lemma 3.3. For any voter $i \in [n]$ and alternative $a \in [m]$,

$$1 = \sum_{j=1}^{m} u_i(\sigma_i(j)) \geq \sum_{j=1}^{\sigma_i^{-1}(a)} u_i(\sigma_i(j)) \geq \sigma_i^{-1}(a) \cdot u_i(\{a\}).$$

Thus, $u_i(\{a\}) \leq 1/\sigma_i^{-1}(a)$. Moreover, by the subadditivity of u_i, $u_i(S) \leq \sum_{j \in S} u_i(\{j\})$. It follows that

$$\sum_{i=1}^{n} u_i(S) \leq \sum_{i=1}^{n} \sum_{j \in S} u_i(\{j\}) \leq \sum_{j \in S} \sum_{i=1}^{n} 1/\sigma_i^{-1}(j)$$
$$= \sum_{j \in S} \mathsf{score}(j).$$

□

We require one other lemma that is quite technical; its proof is relegated to the appendix.[1] In the lemma, and throughout the theorem's proof, we denote by $T \xleftarrow{k} S$ the experiment of drawing a subset T of size k from S uniformly at random.

[1] An anonymous appendix containing the proof of Lemma 3.4 is available via `http://www.filedropper.com/442appendix`. Please note that this is allowed according to the IJCAI'18 CFP, see `https://www.ijcai-18.org/FAQ/#q6`.

Lemma 3.4. *Suppose $A \subseteq B \cap C$ and $k \le |B| \le |C|$. Function $g : 2^A \to \mathbb{R}_+$ satisfies the monotonicity and subadditivity conditions. Then*

$$|B| \cdot \mathop{\mathbb{E}}_{T \xleftarrow{k} B}[g(T \cap A)] \le 4|C| \cdot \mathop{\mathbb{E}}_{T \xleftarrow{k} C}[g(T \cap A)].$$

We are now ready to prove the theorem.

Proof of Theorem 3.1. We construct a randomized social welfare function that, given a preference profile $\vec{\sigma}$, proceeds as follows.

- Sort the alternatives into a ranking ν with $\mathsf{score}(\nu(1)) \ge \mathsf{score}(\nu(2)) \ge \cdots \ge \mathsf{score}(\nu(m))$.
- Let $t_{\max} = \lceil \log_2 m \rceil$ and $\alpha = \sqrt{m \ln m}$. Draw t uniformly at random from $[t_{\max}]$ and set $m'_t = \min\left(\lfloor 2^t \alpha \rfloor, m\right)$.
- With probability $1/2$, return a uniformly random permutation of $[m]$. Otherwise, shuffle the first m'_t elements of ν uniformly at random, and return the resulting ordering.

The rest of the proof analyzes the distortion of the foregoing function. By the monotonicity of utility functions, the social welfare of every ranking $\tau \in \mathcal{S}_m$ is at least

$$\begin{aligned}
\mathrm{sw}(s) &= \sum_{i=1}^{n} \sum_{j=1}^{m} w_{i,j} \cdot u_i(\{\tau(1), \tau(2), \ldots, \tau(j)\}) \\
&\ge \sum_{i=1}^{n} \sum_{j=1}^{m} w_{i,j} \cdot u_i(\{\tau(1)\}) = \sum_{i=1}^{n} u_i(\{\tau(1)\}),
\end{aligned}$$

where the last transition follows from $\sum_{j=1}^{m} w_{i,j} = 1$.

If the mechanism decides to return a random permutation τ, $\tau(1)$ is uniformly distributed in $[m]$, and thus the expected social welfare is at least

$$\frac{1}{m} \sum_{\tau(1)=1}^{m} \sum_{i=1}^{n} u_i(\{\tau(1)\}) = \frac{1}{m} \sum_{i=1}^{n} \sum_{j=1}^{m} u_i(\{j\}) = \frac{n}{m}.$$

On the other hand, consider the case where the mechanism randomly shuffles the first m'_t elements in ν. Let $m_t \triangleq \min\{2^t, m\}$, and define $R_t \triangleq$

$\{\nu(1), \nu(2), \ldots, \nu(m_t')\}$. The resulting expected social welfare is at least

$$\sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} R_t} [u_i(T)].$$

Let SOL_t denote the expected social welfare conditioning on the value of t. Then the above discussion implies that

$$\mathsf{SOL}_t \geq \frac{n}{2m} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} R_t} [u_i(T)]. \tag{3.1}$$

Let μ^* denote the welfare-maximizing ranking. Let $S_t \triangleq \{\mu^*(1), \mu^*(2), \ldots, \mu^*(m_t)\}$, and

$$\mathsf{OPT}_t \triangleq \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t} [u_i(T)].$$

In the following, we show that

$$\sum_{t=1}^{t_{\max}} \mathsf{OPT}_t \geq \frac{\mathrm{sw}(\mu^*)}{2}, \tag{3.2}$$

and for any $t \in [t_{\max}]$,

$$\mathsf{SOL}_t \geq \frac{\mathsf{OPT}_t}{12\sqrt{m \ln m}}. \tag{3.3}$$

Inequalities (3.2) and (3.3) directly imply that the expected social welfare obtained by the mechanism is at least

$$\begin{aligned} \frac{1}{t_{\max}} \sum_{t=1}^{t_{\max}} \mathsf{SOL}_t &\geq \frac{1}{\lceil \log_2 m \rceil} \sum_{t=1}^{t_{\max}} \frac{\mathsf{OPT}_t}{12\sqrt{m \ln m}} \\ &\geq \frac{\mathrm{sw}(\mu^*)}{\mathcal{O}\left(\sqrt{m} \ln^{3/2} m\right)}, \end{aligned}$$

which concludes the proof.

Proof of Equation (3.2). Note that for any $t \in [t_{\max}]$ and $j \in [m_t/2, m_t]$,

$$
\begin{aligned}
\mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}[u_i(T)] &\geq \mathop{\mathbb{E}}_{T \xleftarrow{m_t/2} S_t}[u_i(T)] \geq \mathop{\mathbb{E}}_{T \xleftarrow{m_t} S_t}[u_i(T)] \cdot \frac{1}{2} \\
&= u_i(S_t) \cdot \frac{1}{2} \\
&\geq u_i(\{\mu^*(1), \mu^*(2), \ldots, \mu^*(j)\}) \cdot \frac{1}{2},
\end{aligned} \tag{3.4}
$$

where the first transition follows from the monotonicity of u_i, the second from its subadditivity, the third from $|S_t| = m_t$, and the last again from monotonicity. Therefore,

$$
\begin{aligned}
\sum_{t=1}^{t_{\max}} \mathsf{OPT}_t &\geq \sum_{t=1}^{t_{\max}} \sum_{i=1}^{n} \sum_{j=m_t/2}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}[u_i(T)] \\
&\geq \frac{1}{2} \sum_{i=1}^{n} \sum_{t=1}^{t_{\max}} \sum_{j=m_t/2}^{m_t} w_{i,j} \cdot u_i(\{\mu^*(1), \ldots, \mu^*(j)\}) \\
&\geq \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{i,j} \cdot u_i(\{\mu^*(1), \ldots, \mu^*(j)\}) \\
&= \frac{\mathrm{sw}(\mu^*)}{2},
\end{aligned}
$$

where the second inequality follows from Equation (3.4), and the third holds because $m_1/2 = 1$ and $m_{t_{\max}} = m$.

Proof of Equation (3.3). Let $S_t^+ = S_t \cap R_t$ and $S_t^- = S_t \setminus R_t$. The subadditivity of u_i implies that for any $T \subseteq S_t$, $u_i(T) \leq u_i(T \cap S_t^+) + u_i(T \cap S_t^-)$. Thus, we can derive an upper bound on OPT_t as follows:

$$
\begin{aligned}
\mathsf{OPT}_t &= \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}[u_i(T)] \\
&\leq \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}\left[u_i(T \cap S_t^+) + u_i(T \cap S_t^-)\right] \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}\left[u_i(T \cap S_t^+)\right]
\end{aligned}
$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{m_t} w_{i,j}\cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}\left[u_i(T\cap S_t^-)\right]. \tag{3.5}$$

We establish upper bounds on the two terms on the right hand side of Equation (3.5) separately. For the first term, note that $S_t^+ \subseteq S_t \cap R_t$ and $|S_t| = m_t \le m_t' = |R_t|$. For any $i \in [n]$ and $j \in [m_t]$, applying Lemma 3.4 with $g = u_i$, $k = j$, $A = S_t^+$, $B = S_t$ and $C = R_t$ gives

$$|S_t| \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}\left[u_i(T\cap S_t^+)\right] \le 4|R_t| \mathop{\mathbb{E}}_{T \xleftarrow{j} R_t}\left[u_i(T\cap S_t^+)\right].$$

It follows that

$$\begin{aligned}\mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}\left[u_i(T\cap S_t^+)\right] &\le \frac{4m_t'}{m_t}\mathop{\mathbb{E}}_{T \xleftarrow{j} R_t}\left[u_i(T\cap S_t^+)\right]\\ &\le 4\alpha \mathop{\mathbb{E}}_{T \xleftarrow{j} R_t}\left[u_i(T\cap S_t^+)\right].\end{aligned}$$

Summation over i and j yields

$$\begin{aligned}&\sum_{i=1}^{n}\sum_{j=1}^{m_t} w_{i,j}\cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}\left[u_i(T\cap S_t^+)\right]\\ &\quad\le 4\alpha \sum_{i=1}^{n}\sum_{j=1}^{m_t} w_{i,j}\cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} R_t}\left[u_i(T\cap S_t^+)\right] \qquad (3.6)\\ &\quad\le 8\alpha\cdot\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m_t} w_{i,j}\cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} R_t}\left[u_i(T)\right].\end{aligned}$$

We next bound the second term on the right hand side of Equation (3.5). Note that

$$\begin{aligned}&\sum_{i=1}^{n}\sum_{j=1}^{m_t} w_{i,j}\cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t}\left[u_i(T\cap S_t^-)\right]\\ &\quad\le \sum_{i=1}^{n}\sum_{j=1}^{m_t} w_{i,j}\cdot u_i(S_t^-) \le \sum_{i=1}^{n} u_i(S_t^-) \le \sum_{j\in S_t^-}\mathsf{score}(j).\end{aligned}$$

Here the first step is due to the monotonicity of u_i, the second step holds since $\sum_{j=1}^{m_t} w_{i,j} \le \sum_{j=1}^{m} w_{i,j} = 1$, while last step applies Lemma 3.3. For each alternative $a \in S_t^-$, it follows from Lemma 3.2 that

$$3n \ln m \geq \sum_{j=1}^{m} \mathsf{score}(j) \geq \sum_{j=1}^{m'_t} \mathsf{score}(\nu(j)) \geq m'_t \cdot \mathsf{score}(a),$$

so $\mathsf{score}(a) \leq 3n \ln m / m'_t$ for any $a \in S_t^-$. Therefore, we have

$$\sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t} \left[u_i(T \cap S_t^-)\right] \leq 3n \ln m \cdot \frac{|S_t^-|}{m'_t}.$$

Recall that $m'_t = \min\left(\lfloor 2^t \alpha \rfloor, m\right)$, and $S_t^- = S_t \setminus R_t = S_t \setminus \{\nu(1), \ldots, \nu(m'_t)\}$. If $m'_t = m$, we have $S_t^- = \emptyset$ and $|S_t^-|/m'_t = 0$. When $m'_t < m$, it holds that $m'_t = \lfloor 2^t \alpha \rfloor \geq 2^{t-1}\alpha$ and $m_t = 2^t$. Thus, $|S_t^-|/m'_t \leq m_t/m'_t \leq 2/\alpha$. In either case,

$$\begin{aligned}
\sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t} \left[u_i(T \cap S_t^-)\right] &\leq 3n \ln m \cdot \frac{|S_t^-|}{m'_t} \\
&\leq 3n \ln m \cdot \frac{2}{\alpha} \\
&= \frac{n}{2m} \cdot 12\alpha. \qquad (3.7)
\end{aligned}$$

Putting everything together, we have that

$$\begin{aligned}
\mathsf{OPT}_t &\leq \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t} \left[u_i(T \cap S_t^+)\right] \\
&\quad + \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} S_t} \left[u_i(T \cap S_t^-)\right] \\
&\leq 8\alpha \cdot \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m_t} w_{i,j} \cdot \mathop{\mathbb{E}}_{T \xleftarrow{j} R_t} \left[u_i(T)\right] + 12\alpha \cdot \frac{n}{2m} \\
&\leq 12\alpha \cdot \mathsf{SOL}_t = 12\sqrt{m \ln m} \cdot \mathsf{SOL}_t,
\end{aligned}$$

where the first inequality follows from Equation (3.5), the second from (3.6) and (3.7), and the third from (3.1). This proves Equation (3.3) and completes the proof of the theorem. $\square$

We remark that some restriction on the combinatorial structure of the valuation functions, beyond monotonicity, is necessary to achieve sublinear distortion. Indeed, in the following example we construct non-subadditive utility functions such that any social welfare function must have distortion $\Omega(m)$.[2]

Example 3.5. *Consider the following utility function: for two distinct alternatives $a, b \in [m]$,*

$$u_{a,b}(S) = \begin{cases} 1, & \{a,b\} \subseteq S, \\ |S|/m, & \textit{otherwise.} \end{cases}$$

Note that $u_{a,b}$ is monotonic. Moreover, the function is consistent with any ranking of $[m]$, so we can assume that a voter with utility function $u_{a,b}$ reports the same ranking σ regardless of a and b.

Let there be a single voter with weight vector $(0, 1, 0, 0, \ldots)$. The utility of a ranking τ is given by $u_{a,b}(\{\tau(1), \tau(2)\})$. In order to achieve a utility of 1 (rather than $2/m$), it is necessary to place a and b in the top two slots. Any randomized welfare function has two alternatives that, given σ as input, are placed in the first two positions with probability at most $2/m(m-1)$. By choosing these two alternatives to be a and b, we can guarantee that the function achieves expected social welfare at most

$$\frac{2}{m(m-1)} \cdot 1 + \left(1 - \frac{2}{m(m-1)}\right) \cdot \frac{2}{m},$$

whereas the optimum is 1. The ratio is $\Omega(m)$.

3.3.1 Proof of Lemma 3.4

Before proving Lemma 3.4, we introduce two lemmas that establish some properties of monotonic and subadditive functions. The following lemma states that, given a monotonic utility function defined on set S, when a subset of size k is drawn from S uniformly at random, the expected utility of the resulting subset is non-decreasing in k.

[2] Ranking the alternatives uniformly at random achieves distortion $\mathcal{O}(m)$. Thus, in such cases we cannot significantly outperform a random guess.

Lemma 3.6. *Suppose function $g : 2^S \to \mathbb{R}_+$ satisfies the monotonicity condition. Define $(a_k)_{k=0}^{|S|}$ as $a_k \triangleq \mathbb{E}_{T \xleftarrow{k} S}[g(T)]$. Then,*

$$a_0 \le a_1 \le \cdots \le a_{|S|}.$$

Proof of Lemma 3.6. Fix integer k between 1 and $|S|$. Suppose we draw $X \xleftarrow{k} S$, choose x from X uniformly at random, and then let $Y = X \setminus \{x\}$. Note that X and Y are uniformly (yet not independently) distributed in all subsets of S of size k and $k-1$, respectively. Since $Y \subseteq X$, the monotonicity of g implies that $g(Y) \le g(X)$. Taking the expectation yields $a_{k-1} \le a_k$. $\square$

We now show that the sequence $(a_k)_{k=0}^{|S|}$, defined in Lemma 3.6, is subadditive, assuming the subadditivity of function g.

Lemma 3.7. *Suppose function $g : 2^S \to \mathbb{R}_+$ satisfies the subadditivity condition. Let $a_k = \mathbb{E}_{T \xleftarrow{k} S}[g(T)]$. Then for any integers $n, m \ge 0$ that satisfy $n + m \le |S|$,*

$$a_n + a_m \ge a_{n+m}.$$

Proof of Lemma 3.7. Draw $X \xleftarrow{n} S$ and $Y \xleftarrow{m} S \setminus X$. Clearly, X and Y are uniformly distributed among all subsets of S with n elements and m elements, respectively. Moreover, $X \cup Y$ is also a uniformly random subset of size $n+m$. For each realization of X and Y, it follows from the subadditivity of g that

$$g(X) + g(Y) \ge g(X \cup Y).$$

Taking the expectation over the randomness in (X, Y) yields

$$a_n + a_m \ge a_{n+m}.$$

$\square$

Proof of Lemma 3.4. Fix set A, integer k, and function $g : 2^A \to \mathbb{R}_+$ that satisfies the monotonicity and subadditivity conditions. For $n \ge \max(k, |A|)$, define $f(n)$ as

$$f(n) \triangleq n \cdot \mathop{\mathbb{E}}_{T \xleftarrow{k} S_n} [g(T \cap A)],$$

where S_n is a superset of A with n elements.[3] It suffices to prove that $f(n)$ is *approximately non-decreasing* in the sense that for any $n_1 \le n_2$, $f(n_1) \le 4f(n_2)$.

[3] Any such set S_n gives the same definition of $f(n)$.

Define $a_j \triangleq \mathbb{E}_{T \xleftarrow{j} A}[g(T)]$ for $0 \leq j \leq |A|$, and let $a_j = a_{|A|}$ for $j > |A|$.

Sequence $(a_j)_{j=0}^{\infty}$ is monotonic and subadditive. Lemma 3.6 implies that the finite sequence $(a_0, a_1, \ldots, a_{|A|})$ is non-decreasing. Since $a_j = a_{|A|}$ for any $j > |A|$, the complete sequence $(a_j)_{j=0}^{\infty}$ is also non-decreasing. Moreover, we claim that $2a_k \geq a_{2k}$ for any $1 \leq k \leq |A|$. In fact, if $2k \leq |A|$, the inequality directly follows from Lemma 3.7. If $2k > |A|$, by Lemmas 3.6 and 3.7,

$$2a_k \geq a_k + a_{|A|-k} \geq a_{|A|} = a_{2k}.$$

Approximation of $(a_j/j)_{j=1}^{|A|}$. For each $j \in [|A|]$, define b_j as $b_j \triangleq a_{j'}/j'$, where $j' = 2^{\lceil \log_2 j \rceil}$. Moreover, let $b_{|A|+1} = 0$. We show that $(b_j)_{j=1}^{|A|}$ is non-increasing and approximates $(a_j/j)_{j=1}^{|A|}$.

By construction,

$$(b_j)_{j=1}^{|A|} = \left(\frac{a_1}{1}, \frac{a_2}{2}, \frac{a_4}{4}, \frac{a_4}{4}, \frac{a_8}{8}, \ldots\right).$$

Since $2a_k \geq a_{2k}$ for any $k \in [|A|]$, we have

$$\frac{a_1}{1} \geq \frac{a_2}{2} \geq \frac{a_4}{4} \geq \cdots$$

This proves that $(b_j)_{j=1}^{|A|}$ is non-increasing.

Let $j' = 2^{\lceil \log_2 j \rceil}$. Since $j \leq j' \leq 2j$, it follows from the monotonicity and subadditivity of $\{a_j\}$ that $a_j \leq a_{j'} \leq a_{2j} \leq 2a_j$. Therefore,

$$\frac{a_j}{2j} \leq \frac{a_{j'}}{j'} \leq \frac{2a_j}{j},$$

i.e., b_j approximates a_j/j up to a factor of 2.

Compute $f(n)$. Fix $n \geq \max(k, |A|)$. Recall that

$$f(n) = n \cdot \mathop{\mathbb{E}}_{T \xleftarrow{k} S_n} [g(T \cap A)].$$

For $0 \leq j \leq |A|$, let $\mathcal{E}_j$ denote the event that $|T \cap A| = j$. When conditioning on $\mathcal{E}_j$, $T \cap A$ is uniformly distributed among all subsets of size j in A, i.e., $\mathbb{E}[g(T \cap A)|\mathcal{E}_j] = a_j$. Moreover,

$$\Pr[\mathcal{E}_j] = \frac{\binom{|A|}{j}\binom{n-|A|}{k-j}}{\binom{n}{k}}.$$

By the law of total expectation,

$$
\begin{aligned}
f(n) &= n \sum_{j=0}^{|A|} \Pr[\mathcal{E}_j] \cdot \mathbb{E}[g(T \cap A) | \mathcal{E}_j] \\
&= n \sum_{j=1}^{|A|} \frac{\binom{|A|}{j}\binom{n-|A|}{k-j}}{\binom{n}{k}} \cdot a_j \\
&= n \sum_{j=1}^{|A|} \frac{\frac{|A|}{j}\binom{|A|-1}{j-1}\binom{n-|A|}{k-j}}{\frac{n}{k}\binom{n-1}{k-1}} \cdot a_j \\
&= k|A| \sum_{j=1}^{|A|} \frac{\binom{|A|-1}{j-1}\binom{n-|A|}{k-j}}{\binom{n-1}{k-1}} \cdot \frac{a_j}{j},
\end{aligned}
$$

where the second equality holds because $a_0 = g(\emptyset) = 0$, and the third because

$$
\binom{n}{m} = \frac{n}{m}\binom{n-1}{m-1}
$$

for $1 \le m \le n$.

Let

$$
p_{n,j} = \frac{\binom{|A|-1}{j-1}\binom{n-|A|}{k-j}}{\binom{n-1}{k-1}}
$$

and $q_{n,j} = \sum_{l=1}^{j} p_{n,l}$. Recall that $b_{|A|+1} = 0$. Using summation by parts, we have

$$
\begin{aligned}
f(n) &= k|A| \sum_{j=1}^{|A|} p_{n,j} \cdot \frac{a_j}{j} \\
&\le 2k|A| \sum_{j=1}^{|A|} p_{n,j} \cdot b_j \\
&= 2k|A| \sum_{j=1}^{|A|} (b_j - b_{j+1})\, q_{n,j}.
\end{aligned}
$$

Similarly, we have

$$
f(n) \ge \frac{1}{2} k|A| \sum_{j=1}^{|A|} (b_j - b_{j+1})\, q_{n,j}.
$$

In the following, we show that for any $j \in [|A|]$, $(q_{n,j})_{n=\max(k,|A|)}^{\infty}$ is non-decreasing in n. This completes the proof of the lemma, as for any $\max(k,|A|) \le n_1 \le n_2$, we have

$$\begin{aligned} f(n_1) &\le 2k|A| \sum_{j=1}^{|A|} (b_j - b_{j+1})\, q_{n_1,j} \\ &\le 2k|A| \sum_{j=1}^{|A|} (b_j - b_{j+1})\, q_{n_2,j} \\ &\le 4f(n_2). \end{aligned}$$

Monotonicity of $(q_{n,j})$. Fix set A, $j \in [|A|]$ and k. Recall that

$$p_{n,j} = \frac{\binom{|A|-1}{j-1}\binom{n-|A|}{k-j}}{\binom{n-1}{k-1}}$$

and $q_{n,j} = \sum_{l=1}^{j} p_{n,l}$. For every $n \ge \max(k,|A|)$, define random variable $T_n \xleftarrow{k-1} [n-1]$, i.e., T_n is a random subset of size $k-1$ drawn from $[n-1]$. It can be verified that

$$p_{n,j} = \Pr[|T_n \cap [|A|-1]| = j-1],$$

and thus,

$$q_{n,j} = \Pr[|T_n \cap [|A|-1]| < j].$$

To show that $(q_{n,j})_{n=\max(k,|A|)}^{\infty}$ is non-decreasing in n, we consider the following experiment:

- Draw $X \xleftarrow{k-1} [n]$.
- Let $Y = X$ if $n \notin X$; otherwise, let $Y = X \setminus \{n\} \cup \{x\}$, where x is drawn uniformly from $[n-1] \setminus X$.

By construction, the marginal distributions of X and Y are identical to those of T_{n+1} and T_n, respectively. Moreover, as Y is either equal to X, or obtained from X by replacing n with a smaller element, we have

$$|X \cap [|A|-1]| \le |Y \cap [|A|-1]|.$$

Therefore,

$$\mathbb{I}\left[|X \cap [|A|-1]| < j\right] \geq \mathbb{I}\left[|Y \cap [|A|-1]| < j\right],$$

where $\mathbb{I}[\cdot]$ denotes the indicator function.

Taking the expectation over the randomness in (X, Y) yields

$$\Pr[|T_{n+1} \cap [|A|-1]| < j] \geq \Pr[|T_n \cap [|A|-1]| < j],$$

i.e., $q_{n+1,j} \geq q_{n,j}$. This proves the monotonicity of $(q_{n,j})$, and thus completes the proof of the lemma. $\square$

3.4 Empirical Results

For our computational experiments we focus on deterministic social welfare functions, and on additive utility functions — but we generalize the position-weighted model slightly. Let voter $i \in [n]$ have ranking σ_i consistent with the utility *matrix* $U^i \in \mathbb{R}^{m \times m}$, where U^i_{ap} is the utility voter i has for alternative a appearing in position p. As before, voter i's preferences impose constraints on U^i. Specifically, higher ranked alternatives have utility at least as large as lower ranked alternatives, for any specific position, that is, $U^i_{\sigma_i(p),j} \geq U^i_{\sigma_i(p+1),j}$ for all $p \in [m-1]$, $j \in [m]$, and $U^i_{a,p} \geq U^i_{a,p+1}$ for all $a \in [m]$, $p \in [m-1]$. Utilities are normalized to have $\sum_{a\in[m]}\sum_{p\in[m]} U^i_{ap} = 1$. The *utility profile* is $\vec{U} = (U^1, \ldots, U^n)$.

Let us represent a ranking τ by a permutation matrix $X(\tau) \in \Pi_m$. The social welfare of a ranking τ is $\sum_{i\in[n]}\langle U^i, X(\tau)\rangle$, where $\langle A, B\rangle = \sum_{ij} A_{ij}B_{ij}$ is the *Frobenius inner product*. We can now write the mathematical program that finds the (deterministic) ranking X with minimum distortion z given an input profile $\vec{\sigma}$ as

$$\begin{aligned} \min_{z,\tau\in\mathcal{S}_m} \quad & z \\ z \quad & \geq \frac{\sum_{i=1}^n \langle U^i, X(\rho)\rangle}{\sum_{i=1}^n \langle U^i, X(\tau)\rangle} \quad \forall \vec{U} \triangleright \vec{\sigma}\,,\ \rho \in \mathcal{S}_m \end{aligned} \tag{3.8}$$

This formulation has intractably many constraints in Equation (3.8). But these constraints may be omitted and added as needed, by solving the sub-problem

$$\min_{\vec{U}\triangleright\vec{\sigma},\rho\in\mathcal{S}_m} \qquad \bar{z}\cdot\sum_{i=1}^n \langle U^i, X(\bar{\tau})\rangle - \sum_{i=1}^n \langle U^i, X(\rho)\rangle$$

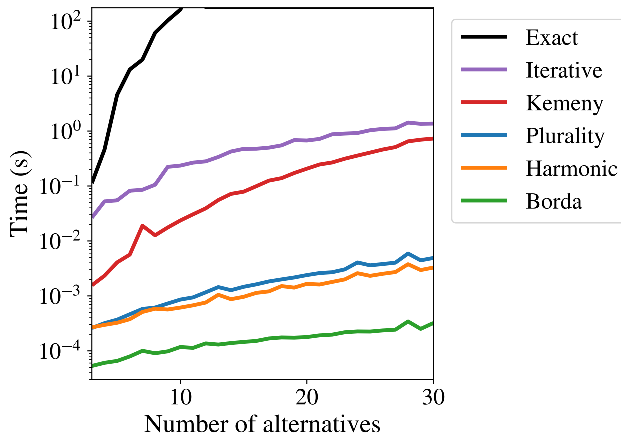


Figure 3.1: Runtime (in seconds) for increasing instance size, on an a machine with an Intel Core i5-4200U CPU and 8 GB RAM.

where $\bar{z}, \bar{\tau}$ are the current optimal solutions to the master problem. A violated cut is found if the objective function value of the subproblem is strictly less than 0. The procedure terminates with the optimal z, τ when no violated cuts are found.

The subproblem is nonconvex even when the integrality constraints are relaxed, and, therefore, finding violated cuts is computationally expensive. Nevertheless, Figure 3.1 shows that it is currently practical to compute distortion-minimizing rankings exactly for instances with up to 10 alternatives within a couple of minutes. We expect that this will be sufficient for the vast majority of instances seen in practice. Indeed, 98.3% of the instances submitted to RoboVote (as of January 19, 2018) have 10 or fewer alternatives.

For larger instances, we evaluate the performance of the following, more scalable, heuristics:

1. *Kemeny*: Return the ranking that minimizes the total number of disagreements on pairs of alternatives with the input profile.

2. *Borda*: Rank alternatives by their Borda scores, defined as $\sum_{i=1}^{n}(m - \sigma_i^{-1}(a))$.

3. *Plurality*: Rank alternatives based on the number of times they are ranked first. Break ties by considering subsequent positions.

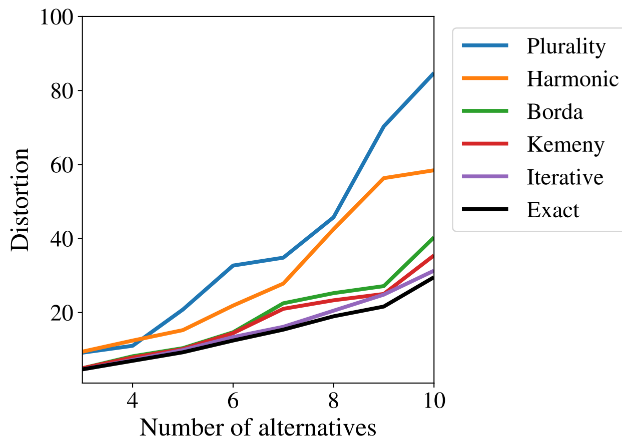


Figure 3.2: Average distortion of heuristic and exact methods.

4. *Harmonic*: Return a ranking according to Theorem 3.1.

5. *Iterative*: Iteratively find and remove the alternative that minimizes distortion for the problem of returning a single alternative with maximum social welfare.

We evaluate these heuristics on instances with $n = 10$ and $m \in \{3, \ldots, 30\}$. Every alternative a is assigned a *quality* c_a, and $u_i(\{a\})$ is drawn from a truncated normal distribution around c_a. Vector $u_i = (u_i(\{a\}))_{a \in [m]}$ induces σ_i. Position weights w_i are drawn uniformly at random in $[0, 1]$, and ordered. Voter i's utility matrix $U^i = w_i u_i$ is normalized.

Every social welfare function f only has access to $\vec{\sigma}$ and is evaluated on two metrics: the distortion of the returned ranking $\rho = f(\vec{\sigma})$, and the *social welfare ratio* $\max_{\tau \in \mathcal{S}_m} \mathrm{sw}(\tau, \vec{U}) / \mathrm{sw}(\rho, \vec{U})$. Note that the latter measure estimates the *average case* with respect to utility profiles.

The distortion and social welfare ratios of the proposed heuristics are shown in Figures 3.2 and 3.3. Distortion is reported for $m \leq 10$, where it is possible to compare to the optimal distortion, and 100 repetitions; social welfare for $m \leq 30$ and 200 repetitions.

The distortion of Borda, Kemeny and especially Iterative compares well with the optimal distortion. Kemeny and Borda also lead to very high efficiency, with average social welfare within 1% of optimal.

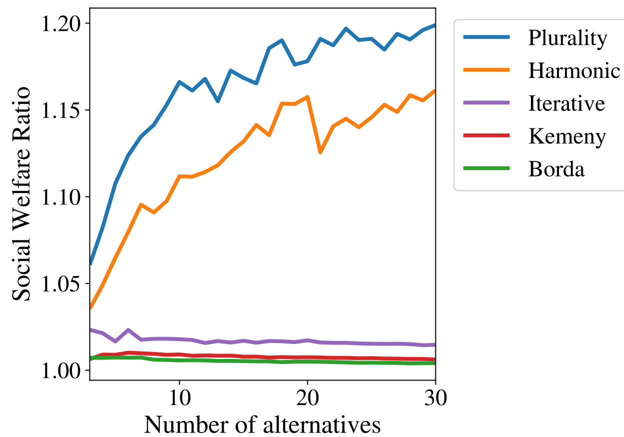


Figure 3.3: Average social welfare ratio of different heuristics.

3.5 Discussion

Much like previous papers on implicit utilitarian voting [46, 20], there is a certain gap between the theoretical and empirical results, in the sense that the theoretical distortion bound of Theorem 3.1 holds for *randomized* social welfare functions, whereas the empirical results of Section 3.4 hold for deterministic functions. The value of theoretical distortion bounds is that they tell us whether rankings inherently provides useful information for optimizing social welfare. The fact that the bound is essentially no worse than for the case of a single winner means that the implicit utilitarian voting approach does extend to the design of social welfare functions.

On a practical level, our empirical results suggest that classic methods like the Kemeny rule (which is currently deployed on RoboVote) and Borda count provide near-optimal performance from the viewpoint of implicit utilitarian voting. Alternatively, it is possible to compute the distortion-minimizing social welfare function if instances are restricted to at most ten alternatives. Although almost all instances arising from small-group decisions (of the type made on RoboVote) are of that size, some high-stakes decisions, such as ranking applicants for a job or candidates for a PhD program, involve a much larger number of alternatives, and motivate the development of faster algorithms.

Chapter 4

Aggregating noisy estimates of a ground truth

4.1 Introduction

The field of *computational social choice* [35] has been undergoing a transformation, as rigorous approaches to voting and resource allocation, previously thought to be purely theoretical, are being applied to group decision making and social computing in practice [52]. RoboVote.org, a not-for-profit social choice website launched in November 2016, gives a compelling (and unquestionably recent) example. Its short-term goal is to facilitate effective group decision making by providing free access to optimization-based voting rules. In the long term, Procaccia [2016] argue that RoboVote and similar applications of computational social choice can change the public's perception of democracy.

RoboVote distinguishes between two types of social choice tasks: aggregation of *subjective preferences*, which was the topic of chapter 2, and aggregation of *objective opinions*. Examples of the former task include a group of friends deciding where to go to dinner or which movie to watch; family members selecting a vacation spot; and faculty members choosing between faculty candidates. In all of these cases, there is no single correct choice — the goal is to choose an outcome that makes the participants as happy as possible overall.

By contrast, the latter task involves situations where some alternatives are objectively better than others, i.e., there is a true *ranking* of the alternatives by quality, but voters can only obtain noisy estimates thereof. The goal is to aggregate these noisy opinions, which are themselves rankings of the alternatives, and uncover the true ranking. For example, consider a group of engineers deciding which product prototype to develop based on an objective metric, such as projected market share. Each prototype, if selected for development (and, ultimately, production), would achieve a particular market share, so a true ranking of the alternatives certainly exists. Other examples include a group of investors deciding which company to invest in, based on projected revenue; and employees of a movie studio selecting a movie script for production, based on projected box office earnings.

In this chapter, we focus on the second setting — aggregating objective opinions. This is a problem that boasts centuries of research: it dates back to the work of the Marquis de Condorcet, published in 1785, in which he proposed a random noise model that governs how voters make mistakes when estimating the true ranking. He further suggested — albeit in a way that took 203 years to decipher [135] — that a voting rule should be a *maximum likelihood estimator (MLE)*, that is, it should select an outcome that is most likely to coincide with the true ranking, given the observed votes and the *known* structure of the random noise model. Condorcet's approach is the foundation of a significant body of modern work [54, 56, 65, 64, 133, 132, 90, 108, 12, 13, 14, 93, 42, 45, 131].

While the MLE approach is conceptually appealing, it is also fragile. Indeed, it advocates rules that are tailor-made for one specific noise model, which is unlikely to accurately represent real-world errors [93]. Recent work [45, 42] circumvents this problem by designing voting rules that are robust to large families of noise models, at the price of theoretical guarantees that only kick in when the number of voters is large — a reasonable assumption in crowdsourcing settings. However, here we are most interested in helping small groups of people make decisions — on RoboVote, typical instances have 4–10 voters — so this approach is a nonstarter.

4.1.1 The Worst-Case Approach

In recent work, Procaccia et al. [109] have taken another step towards robustness (we will argue shortly that it is perhaps a step too far). Instead of positing a random noise model, they essentially remove all assumptions about the errors made by voters. To be specific, first fix a distance metric d on the space of rankings. For example, the *Kendall tau (KT)* distance between two rankings is the number of pairs of alternatives on which they disagree. We are given a vote profile and an upper bound t on the average distance between the input votes and the true ranking. This induces a set of *feasible* true rankings — those that are within average distance t from the votes. The *worst-case optimal voting rule* returns the ranking that minimizes the maximum distance (according to d) to any feasible true ranking. If this minimax distance is k, then we can guarantee that our output ranking is within distance k from the true ranking. The most pertinent theoretical results of Procaccia et al. are that for any distance metric d, one can always recover a ranking that is at distance at most $2t$ from the true ranking, i.e., $k \leq 2t$; and that for the four most popular distance metrics used in the social choice literature (including the KT distance), there is a tight lower bound of (roughly) $k \geq 2t$.

Arguably the more compelling results of Procaccia et al. [109], though, are empirical. In the case of objective opinions, the measure used to evaluate a voting rule is almost indisputable: the distance (according to the distance metric of interest, say KT) between the output ranking and the *actual* true ranking. And, indeed, according to this measure, the worst-case approach significantly outperforms previous approaches — including those based on random noise models — on real data [93]; we elaborate on this dataset later.

Based on the foregoing empirical results, the algorithms deployed on RoboVote for aggregating objective opinions implement the worst-case approach. Specifically, given an upper bound t on the average KT distance between the input votes and the true ranking,[1] the algorithm computes the set of feasible true rankings (by enumerating the solutions to an integer program), and selects a ranking that minimizes the KT distance to any ranking in that set (by solving another integer program).

RoboVote also supports two additional output types: single winning alternative, and a subset of alternatives. When the user requests a single alternative as the output, the algorithm computes the set of feasible true rankings as before, and returns the alternative that minimizes the maximum position in any feasible true ranking, that is, the alternative that is guaranteed to be as close to the top as possible. Computing a subset is similar, with the exception that the *loss* of a subset with respect to a specific feasible

[1]This value is set by minimizing the average distance between any input vote and the remaining votes. This choice guarantees a nonempty set of feasible true rankings, and performs extremely well in experiments.

true ranking is determined based on the *top-ranked* alternative in the subset; the algorithm selects the subset that minimizes the maximum loss over all feasible true rankings. In other words, if this loss is s then any feasible true ranking has an alternative in the subset among its top s alternatives.

4.1.2 Our Approach and Results

To recap, the worst-case approach to aggregating objective opinions has proven quite successful. Nevertheless, it is very conservative, and it seems likely that better results can be achieved in practice by modifying it. We therefore take a more "optimistic" angle by carefully injecting some randomness into the worst-case approach.

In more detail, we refer to the worst-case approach as "worst case" because the errors made by voters are arbitrary, but there is actually another crucial aspect that makes it conservative: the optimization objective — minimizing the *maximum* distance to any feasible true ranking when the output is a ranking, and minimizing the maximum position or loss in any feasible true ranking when the output is a single alternative or a subset of alternatives, respectively. We propose to modify these objective functions, by replacing (in both cases) the word "maximum" with the word "average". Equivalently, we assume a uniform prior over the set of all rankings, which induces a uniform posterior over the set of feasible true rankings, and replace the word "maximum" with the word "expected".[2] Note that this model is fundamentally different from assuming that the *votes* are random: as we mentioned earlier, it is arguable whether real-world votes can be captured by any particular random noise model, not to mention a uniform distribution.[3] By contrast, we make no structural assumptions about the noise, and, in fact, we do not make any new assumptions about the world; we merely modify the optimization objective with respect to the same set of feasible true rankings.

In Section 4.3, we study the case where the output is a ranking. We find that for any distance metric, if the average distance between the vote profile and the true ranking is at most t, then we can recover a ranking whose average distance to the set of feasible true rankings is also t. We also establish essentially matching lower bounds for the four distance metrics studied by [109]. Note that our relaxed goal allows us to improve their bound from

[2]Our positive results actually work for any distribution; see Section 4.6.

[3]That said, some social choice papers do analyze uniformly random vote profiles [125, 103] — a model known as *impartial culture*.

$2t$ to t, which, in our view, is a *qualitative* improvement, as now we can guarantee performance that is at least as good as the average voter. While we would like to outperform the average voter, this is a worst-case (over noisy votes) guarantee, and, as we shall see, in practice we indeed achieve excellent performance.

In Section 4.4, we explore the case where the output is a subset of alternatives (including the all-important case of a single winning alternative). This problem was not studied by [109], in part because their model does not admit nontrivial analytical solutions (as we explain in detail later) — but it is just as important in practice, if not even more so (see Section 4.1.1). We find significant gaps between the guarantees achievable under different distance metrics. Our main technical result concerns the practically significant KT distance and the closely related *footrule* distance: If the average distance between the vote profile and the true ranking is at most t, we can pinpoint a subset of alternatives of size z, whose average loss — that is, the average position of the subset's top-ranked alternative in the set of feasible true rankings (smaller position is closer to the top) — is $O(\sqrt{t/z})$. We also prove a lower bound of $\Omega(\sqrt{t}/z)$, which is tight for a constant subset size z (note that z is now outside of the square root). For the *maximum displacement* distance, we have asymptotically matching upper and lower bounds of $\Theta(t/z)$. Interestingly, for the *Cayley* distance and $z = 1$, we prove a lower bound of $\Omega(\sqrt{m})$, showing that there is no hope of obtaining positive results that depend only on t.

In Section 4.5, we present empirical results from real data. Our key finding is that our methods are robust to overestimates of the true average level of noise in the vote profile — significantly more so than the methods of [109], which are currently deployed on RoboVote. We believe that this conclusion is meaningful for real-world implementation.

4.2 Preliminaries

Let A be a set of *alternatives* with $|A| = m$. Let $\mathcal{L}(A)$ be the set of possible rankings of A, which we think of as permutations $\sigma : A \to [m]$, where $[m] = \{1, \ldots, m\}$. That is, $\sigma(a)$ gives the position of $a \in A$ in σ, with $\sigma^{-1}(1)$ being the highest-ranked alternative, and $\sigma^{-1}(m)$ being the lowest-ranked alternative. A ranking σ induces a strict total order $\succ_\sigma$, such that $a \succ_\sigma b$ if and only if $\sigma(a) < \sigma(b)$. A *vote profile* $\pi = (\sigma_1, \ldots, \sigma_n) \in \mathcal{L}(A)^n$ consists of

n votes, where σ_i is the vote of voter i.

We next introduce notations that will simplify the creation of vote profiles. For a subset of alternatives $A_1 \subseteq A$, let σ^{A_1} be an arbitrary ranking of A_1. For a partition A_1, A_2 of A, $A_1 \succ A_2$ is a *partial order* of A which specifies that every alternative in A_1 is preferred to any alternative in A_2. Similarly, $A_1 \succ \sigma^{A_2}$ is a partial ordering where the alternatives in A_1 are preferred to those in A_2 and the order of the alternatives in A_2 is specified to coincide with σ^{A_2}. An *extension* of a partial order $\mathcal{P}$ is any ranking $\sigma \in \mathcal{L}(A)$ satisfying the partial order. Denote by $\mathcal{F}(\mathcal{P})$ the set of possible extensions of $\mathcal{P}$. For example, $|\mathcal{F}(A_1 \succ A_2)| = |A_1|! \cdot |A_2|!$ and $|\mathcal{F}(A_1 \succ \sigma^{A_2})| = |A_1|!$.

Distance metrics on permutations play an important role in the paper. We pay special attention to the following well-known distance metrics:

- The *Kendall tau (KT)* distance, denoted d_{KT}, measures the number of pairs of alternatives on which the two rankings disagree:

  $$d_{KT}(\sigma, \sigma') \triangleq |\{(a,b) \in A^2 |\ a \succ_\sigma b \text{ and } b \succ_{\sigma'} a\}|.$$

  Equivalently, the KT distance between σ and σ' is the number of swaps between *adjacent* alternatives required to transform one ranking into the other. Some like to think of it as the “bubble sort” distance.

- The *footrule* distance, denoted d_{FR}, measures the total displacement of alternatives between two rankings:

  $$d_{FR}(\sigma, \sigma') \triangleq \sum_{a \in A} |\sigma(a) - \sigma'(a)|.$$

- The *maximum displacement* distance, denoted d_{MD}, is the largest absolute displacement of any alternative between two rankings:

  $$d_{MD}(\sigma, \sigma') \triangleq \max_{a \in A} |\sigma(a) - \sigma'(a)|.$$

- The *Cayley* distance, denoted d_{CY}, measures the number of pairwise swaps required to transform one ranking into the other. In contrast to the KT distance, the swapped alternatives need not be adjacent.

We also require the following definitions that apply to any distance metric d. For a ranking $\sigma \in \mathcal{L}(A)$ and a set of rankings $S \subseteq \mathcal{L}(A)$, define the average distance between σ and S in the obvious way,

$$d(\sigma, S) \triangleq \frac{1}{|S|} \sum_{\sigma' \in S} d(\sigma, \sigma').$$

Similarly, define the average distance between two sets of rankings $S, T \subseteq \mathcal{L}(A)$ as

$$d(S, T) \triangleq \frac{1}{|S| \cdot |T|} \sum_{\sigma \in S} \sum_{\sigma' \in T} d(\sigma, \sigma').$$

Finally, let $d^{\downarrow}(k)$ be the largest distance allowed under the distance metric d which is at most k, i.e.,

$$d^{\downarrow}(k) \triangleq \max\{s \leq k : \exists \sigma, \sigma' \in \mathcal{L}(A) \text{ s.t. } d(\sigma, \sigma') = s\}.$$

4.3 Returning the Right Ranking, in Theory

We first tackle the setting where our goal is to return an accurate ranking. We assume that there is an objective ground truth ranking σ^*, and that n voters submit a *vote profile* π of noisy estimates of this true ranking. As in the work of Procaccia et al. [109], an individual vote is allowed to deviate from the ground truth in any way, but we expect that the average error is bounded, that is, the average distance between the vote profile and the ground truth is no more than some parameter t. Formally, for a distance metric d on $\mathcal{L}(A)$, we are guaranteed that

$$d(\pi, \sigma^*) = \frac{1}{n} \sum_{\sigma \in \pi} d(\sigma, \sigma^*) \leq t.$$

There are several approaches for obtaining good estimates for this upper bound t; we return to this point later.

A combinatorial structure that plays a central role in our analysis is the "ball" of feasible ground truth rankings,

$$\mathcal{B}_t(\pi) \triangleq \{\sigma \in \mathcal{L}(A) : d(\pi, \sigma) \leq t\}.$$

If this ball were a singleton (or empty), our task would be easy. But it typically contains multiple feasible ground truths, as the following example shows.

Example 4.1. *Suppose that $A = \{a, b, c\}$ and the vote profile consists of 5 votes, $\pi = \{(a \succ b \succ c), (a \succ b \succ c), (b \succ c \succ a), (c \succ a \succ b), (a \succ c \succ b)\}$. For each distance metric, let the bound on average error equal half of the maximum distance allowed by the distance metric; in other words, $t_{KT} = 1.5, t_{FR} = 2, t_{MD} = 1$ and $t_{CY} = 1$. The set of feasible ground truths for the vote profile π under the respective distance metrics may be found in Table 4.1.*

Table 4.1: The set of feasible ground truths in Example 4.1 for various distance metrics.

d	t	$\mathcal{B}_t(\pi)$
KT	1.5	$\{(a \succ b \succ c), (c \succ a \succ b), (a \succ c \succ b)\}$
FR	2	$\left\{ \begin{array}{c} (a \succ b \succ c), \\ (a \succ c \succ b) \end{array} \right\}$
MD	1	
CY	1	

Procaccia et al. [109] advocate a conservative approach — they choose a ranking that minimizes the *maximum* distance to any feasible ground truth. By contrast, we are concerned with the *average* distance to the set of feasible ground truths. In other words, we assume that each of the feasible ground truths is equally likely, and our goal is to find a ranking that has a small expected distance to the set of feasible ground truths $\mathcal{B}_t(\pi)$.

Our first result is that is it always possible to find a ranking $\sigma \in \pi$ that is close to $\mathcal{B}_t(\pi)$.

Theorem 4.2. *Given a profile π of n noisy rankings with average distance at most t from the ground truth according to some distance metric d, there always exists a ranking within average distance t from the set of feasible ground truths $\mathcal{B}_t(\pi)$ according to the same metric.*

Proof. For any $\sigma \in \mathcal{B}_t(\pi)$, $d(\sigma, \pi) \le t$. It follows from the definitions that

$$\begin{aligned} d(\pi, \mathcal{B}_t(\pi)) &= \frac{1}{n \cdot |\mathcal{B}_t(\pi)|} \sum_{\sigma' \in \pi} \sum_{\sigma \in \mathcal{B}_t(\pi)} d(\sigma, \sigma') = \frac{1}{|\mathcal{B}_t(\pi)|} \sum_{\sigma \in \mathcal{B}_t(\pi)} \frac{1}{n} \sum_{\sigma' \in \pi} d(\sigma, \sigma') \\ &= \frac{1}{|\mathcal{B}_t(\pi)|} \sum_{\sigma \in \mathcal{B}_t(\pi)} d(\sigma, \pi) \le t. \end{aligned}$$

To conclude the proof, observe that if the average distance from π to $\mathcal{B}_t(\pi)$ is no more than t, then there certainly exists $\sigma'' \in \pi$ with $d(\sigma'', \mathcal{B}_t(\pi)) \leq t$. $\square$

This result holds for any distance metric. Interestingly, it also generalizes to any probability distribution over $\mathcal{B}_t(\pi)$, not just the uniform distribution (see Section 4.6 for additional discussion of this point).

We next derive essentially matching lower bounds for the four common distance metrics introduced in Section 4.2.

Theorem 4.3. *For $d \in \{d_{KT}, d_{FR}, d_{MD}, d_{CY}\}$, there exists a profile π of n noisy rankings with average distance at most t from the ground truth, such that for any ranking, its average distance (according to d) from $\mathcal{B}_t(\pi)$ is at least $d^{\downarrow}(2t)/2$.*

The proof of this theorem relies heavily on constructions that appear in Procaccia et al. [109] and the following four technical lemmas established by Procaccia et al. [109, Theorem 5].

Lemma 4.4. *For $d = d_{KT}$ and $t \leq (m/12)^2$, there exists a partition of A into A_1, A_2, A_3, A_4, and a vote profile consisting of $n/2$ copies of each of the rankings*

$$\begin{aligned} \sigma &= \sigma^{A_1} \succ \sigma^{A_2} \succ \sigma^{A_3} \succ \sigma^{A_4} \\ \sigma' &= \sigma^{A_1}_{rev} \succ \sigma^{A_2}_{rev} \succ \sigma^{A_3}_{rev} \succ \sigma^{A_4}, \end{aligned}$$

for which $\mathcal{B}_t(\pi) = \mathcal{F}(A_1 \succ A_2 \succ A_3 \succ \sigma^{A_4})$ and $\lfloor 2t \rfloor = \sum_{i=1}^{3} \binom{m_i}{2}$, where $m_i \triangleq |A_i|$ for $i \in [4]$.

Lemma 4.5. *For $d = d_{FR}$ and $t \leq (m/8)^2$, there exists a partition of A into A_1, A_2, A_3, A_4, and A_5, and a vote profile $\pi \in \mathcal{L}(A)^n$ consisting of $n/2$ copies of each of the following rankings,*

$$\begin{aligned} \sigma &= \sigma^{A_1} \succ \sigma^{A_2} \succ \sigma^{A_3} \succ \sigma^{A_4} \succ \sigma^{A_5} \\ \sigma' &= \sigma^{A_1}_{rev} \succ \sigma^{A_2}_{rev} \succ \sigma^{A_3}_{rev} \succ \sigma^{A_4}_{rev} \succ \sigma^{A_5}, \end{aligned}$$

for which

$$\mathcal{B}_t(\pi) = \left\{ \rho \in \mathcal{L}(A) \;\middle|\; \begin{array}{ll} \rho(a_i^j), \rho(a_i^{2m_i+1-j})\} = \{\sigma(a_i^j), \sigma(a_i^{2m_i+1-j})\} & \\ & \textit{for } i \in [4], j \in [2m_i], \\ \rho(a_5^j) = \sigma(a_5^j) = \sigma'(a_5^j) & \textit{for } j \in [m_5] \end{array} \right\},$$

where $2m_i = |A_i|$ *for* $i \in [4]$, $m_5 = |A_5|$, *and*

$$d^{\downarrow}_{FR}(2t) = \sum_{i=1}^{4} \left\lfloor \frac{(2m_i)^2}{2} \right\rfloor .$$

Lemma 4.6. *For* $d = d_{CY}$ *and* t *such that* $2\lfloor 2t \rfloor \leq m$*, there exists a vote profile* π *consisting of* $n/2$ *copies of each of the following rankings,*

$$\begin{aligned} \sigma &= (a_1 \succ \cdots \succ a_{2\lfloor 2t \rfloor} \succ a_{2\lfloor 2t \rfloor + 1} \succ \cdots \succ a_m) \\ \sigma' &= (a_{2\lfloor 2t \rfloor} \succ \cdots \succ a_1 \succ a_{2\lfloor 2t \rfloor + 1} \succ \cdots \succ a_m), \end{aligned}$$

for which

$$\mathcal{B}_t(\pi) = \left\{ \rho \in \mathcal{L}(A) \;\middle|\; \begin{array}{l} \{\rho(a_i), \rho(a_{2\lfloor 2t \rfloor + 1 - i})\} = \{i, 2\lfloor 2t \rfloor + 1 - i\} \textit{ for } i \in [\lfloor 2t \rfloor], \\ \rho(a_i) = i \textit{ for } i > 2\lfloor 2t \rfloor \end{array} \right\}.$$

We will need a similar result for maximum displacement.

Lemma 4.7. *For* $d = d_{MD}$ *and* t *such that* $2\lfloor 2t \rfloor \leq m$*, there exists a vote profile* π *consisting of* $n/2$ *copies of each of the following rankings,*

$$\begin{aligned} \sigma &= (a_1 \succ \cdots \succ a_{\lfloor 2t \rfloor}) \succ (a_{\lfloor 2t \rfloor + 1} \succ \cdots \succ a_{2\lfloor 2t \rfloor}) \succ \sigma^{A'} \\ \sigma' &= (a_{\lfloor 2t \rfloor + 1} \succ \cdots \succ a_{2\lfloor 2t \rfloor}) \succ (a_1 \succ \cdots \succ a_{\lfloor 2t \rfloor}) \succ \sigma^{A'}, \end{aligned}$$

where $A' = A \setminus \{a_1, \ldots, a_{2\lfloor 2t \rfloor}\}$*, for which* $\mathcal{B}_t(\pi) = \{\sigma, \sigma'\}$.

Proof. It is easy to see that $\sigma \in \mathcal{B}_t(\pi)$ and $\sigma' \in \mathcal{B}_t(\pi)$, as $d(\sigma, \sigma') = \lfloor 2t \rfloor$. We therefore need to show that $\mathcal{B}_t(\pi)$ does not contain any other rankings.

Let $\rho \in \mathcal{B}_t(\pi)$, and consider its first-ranked alternative, $a = \rho^{-1}(1)$. It holds that $\sigma(a) \geq \lfloor 2t \rfloor + 1$ or $\sigma'(a) \geq \lfloor 2t \rfloor + 1$, because the two rankings place disjoint subsets of alternatives in the first $\lfloor 2t \rfloor$ positions. Suppose first that the former inequality holds; then

$$d(\rho, \sigma) \geq \sigma(a) - \rho(a) \geq \lfloor 2t \rfloor .$$

If $\rho \neq \sigma'$ then $d(\rho, \sigma') \geq 1$, and therefore

$$d(\rho, \mathcal{B}_t(\pi)) = \frac{d(\rho, \sigma) + d(\rho, \sigma')}{2} \geq \frac{\lfloor 2t \rfloor + 1}{2} > t.$$

It follows that $\rho = \sigma'$. Similarly, if the latter inequality holds, then $\rho = \sigma$. $\square$

We are now in a position to prove Theorem 4.3.

Proof of Theorem 4.3. We address each distance metric separately.

The Kendall tau distance. Let π and $\mathcal{B}_t(\pi)$ have the structure specified in Lemma 4.4. For all $\rho \in \mathcal{L}(A)$ and $i \in [3]$, and every pair of alternatives $a \in A_i$, $b \in A_i \setminus \{a\}$, we can divide the rankings in $\mathcal{B}_t(\pi)$ into pairs that are identical except for swapping a and b. Note that for each pair, one ranking agrees with ρ on a and b, and one does not. Therefore,

$$d(\rho, \mathcal{B}_t(\pi)) \geq \frac{\sum_{i=1}^{3} \binom{m_i}{2}}{2} = \frac{\lfloor 2t \rfloor}{2} \geq \frac{d^{\downarrow}(2t)}{2}.$$

The footrule distance. Let π and $\mathcal{B}_t(\pi)$ have the structure specified in Lemma 4.5. For all $\rho \in \mathcal{L}(A)$ and $i \in [4]$, and for every alternative $a_i^j \in A_i$, we can divide the rankings in $\mathcal{B}_t(\pi)$ into pairs that are identical except for swapping a_i^j and $a_i^{2m_i+1-j}$. Note that for each such pair σ and σ', $|\sigma(a_i^j) - \sigma'(a_i^j)| = 2m_i + 1 - 2j$, and using the triangle inequality,

$$|\rho(a_i^j) - \sigma(a_i^j)| + |\rho(a_i^j) - \sigma'(a_i^j)| \geq 2m_i + 1 - 2j.$$

Furthermore, by the structure of $\mathcal{B}_t(\pi)$, we know that

$$\sum_{j=1}^{2m_i} 2m_i + 1 - 2j = \left\lfloor \frac{(2m_i)^2}{2} \right\rfloor.$$

By summing over all $j \in [2m_i]$ and $i \in [4]$, we get

$$d(\rho, \mathcal{B}_t(\pi)) \geq \frac{\sum_{i=1}^{4} \sum_{j=1}^{2m_i} 2m_i + 1 - 2j}{2} = \frac{\sum_{i=1}^{4} \left\lfloor \frac{(2m_i)^2}{2} \right\rfloor}{2} = \frac{d^{\downarrow}(2t)}{2}.$$

The Cayley distance. Let π and $\mathcal{B}_t(\pi)$ have the structure specified in Lemma 4.6. For all $\rho \in \mathcal{L}(A)$, and every pair of alternatives $\{a_i, a_{2\lfloor 2t \rfloor + 1 - i}\}$ for $i \in [\lfloor 2t \rfloor]$, we can divide the rankings in $\mathcal{B}_t(\pi)$ into pairs τ_i and τ_i' that are identical except for swapping a and b. Note that for each pair, one ranking agrees with ρ on a and b, and one does not. Since each swap places at most two alternatives in their correct positions, each of the $\lfloor 2t \rfloor$ pairs adds at least $1/2$ to $d(\rho, \mathcal{B}_t(\pi))$ because $d(\rho, \tau_i) + d(\rho, \tau_i') \geq 1$. Overall we have

$$d(\rho, \mathcal{B}_t(\pi)) \geq \frac{\lfloor 2t \rfloor}{2} \geq \frac{d^{\downarrow}(2t)}{2}.$$

The maximum displacement distance. Let π and $\mathcal{B}_t(\pi)$ have the structure specified in Lemma 4.7. Consider any ranking $\rho \in \mathcal{L}(A)$. Let $a \in A$ be the alternative ranked first in ρ, i.e., $a = \rho^{-1}(1)$. If $a \in \{a_1, \ldots, a_{\lfloor 2t \rfloor}\}$, then $d(\rho, \sigma') \geq \lfloor 2t \rfloor$. Similarly, if $a \in \{a_{2t+1}, \ldots, a_{2\lfloor 2t \rfloor}\}$ then $d(\rho, \sigma) \geq \lfloor 2t \rfloor$. Therefore,

$$d(\rho, \mathcal{B}_t(\pi)) = \frac{d(\rho, \sigma) + d(\rho, \sigma')}{2} \geq \frac{\lfloor 2t \rfloor}{2} \geq \frac{d^{\downarrow}(2t)}{2}.$$

$\square$

4.4 Returning the Right Alternatives, in Theory

In the previous section, we derived bounds on the expected distance of the ranking closest to the set of feasible ground truth rankings. In practice, we may not be interested in eliciting a complete ranking of alternatives, but rather in selecting a subset of the alternatives (often a single alternative) on which to focus attention, time, or effort.

In this section, we bound the average position of the *best* alternative in a subset of alternatives, where the average is taken over the set of feasible ground truths as before. This type of utility function, where the utility of a set is defined by its highest utility member, is consistent with quite a few previous papers that deal with selecting subsets of alternatives in different social choice settings [48, 96, 107, 91, 108, 46]. For example, when selecting a menu of movies to show on a three hour flight, the utility of passengers depends on their most preferred alternative. From a technical viewpoint, this choice has the advantage of giving bounds that improve as the subset size increases, which matches our intuition. Of course, in the important special case where the subset is a singleton, all reasonable definitions coincide.

Formally, let $Z \subseteq A$ be a subset of alternatives; the *loss of Z in σ* is $\ell(Z, \sigma) \triangleq \min_{a \in Z} \sigma(a)$, and therefore the *average loss of Z in $\mathcal{B}_t(\pi)$* is

$$\ell(Z, \mathcal{B}_t(\pi)) \triangleq \frac{1}{|\mathcal{B}_t(\pi)|} \sum_{\sigma \in \mathcal{B}_t(\pi)} \ell(Z, \sigma).$$

For given average error t and subset size z, we are interested in bounding

$$\max_{\pi \in \mathcal{L}(A)^n} \min_{Z \subseteq A \text{ s.t. } |Z| = z} \ell(Z, \mathcal{B}_t(\pi)).$$

In words, we wish to bound the the average loss of the best Z (of size z) in $\mathcal{B}_t(\pi)$, in the worst case over vote profiles.

Let us return to Example 4.1. For the footrule, maximum displacement, and Cayley distance metrics, it is clear from Table 4.1 that selecting $\{a\}$ when $z = 1$ guarantees average loss 1, as $\mathcal{B}_t(\pi)$ only contains rankings that place a first. For the KT distance, the set $\{a\}$ has average loss 4/3, and the set $\{a, c\}$ has average loss 1.

We now turn to the technical results, starting with some lemmas that are independent of specific distance metrics. Throughout this section, we will rely on the following lemma, which is the discrete analogue of selecting a set of z numbers uniformly at random in an interval and studying their order statistics. No doubt someone has proved it in the past, but we include our (cute, if we may say so ourselves) proof, as we will need to reuse specific equations.

Lemma 4.8. *When choosing z elements $Y_1, \ldots, Y_z$ uniformly at random without replacement from the set $[k]$, $\mathbb{E}[\min_{i\in[z]} Y_i] = \frac{k+1}{z+1}$.*

Proof. Let $Y_{min} = \min_{i\in[z]} Y_i$ be the minimum value of the z numbers chosen uniformly at random from $[k]$ without replacement. It holds that

$$\Pr[Y_{min} = y] = \frac{\binom{k-y}{z-1}}{\binom{k}{z}},$$

and therefore

$$\mathbb{E}[Y_{min}] = \sum_{y=1}^{k} y \frac{\binom{k-y}{z-1}}{\binom{k}{z}} = \frac{1}{\binom{k}{z}} \sum_{y=1}^{k} y \binom{k-y}{z-1} = \frac{1}{\binom{k}{z}} \sum_{y=1}^{k-z+1} y \binom{k-y}{z-1}. \tag{4.1}$$

We claim that

$$\sum_{y=1}^{k-z+1} y \binom{k-y}{z-1} = \binom{k+1}{z+1}. \tag{4.2}$$

Indeed, the left hand side can be interpreted as follows: for each choice of $y \in [k-z+1]$, elements $\{1, \ldots, y\}$ form a committee of size y. We have y possibilities for choosing the head of the committee. Then we choose $z-1$ clerks among the elements $\{y+1, \ldots, k\}$. We can interpret the right hand side of Equation (4.2) in the same way. To see how, choose $z+1$ elements from $[k+1]$, and sort them in increasing order to obtain $s_1, \ldots, s_{z+1}$. Now

s_1 is the head of the committee, $y = s_2 - 1$ is the number of committee members, and $s_3 - 1, \ldots, s_{z+1} - 1$ are the clerks.

Plugging Equation (4.2) into Equation (4.1), we get

$$\mathbb{E}[Y_{min}] = \frac{\binom{k+1}{z+1}}{\binom{k}{z}} = \frac{k+1}{z+1}.$$

$\square$

Our strategy for proving upper bounds also relies on the following lemma, which relates the performance of randomized rules on the worst ranking in $\mathcal{B}_t(\pi)$, to the performance of deterministic rules on average, and is reminiscent of Yao's Minimax Principle [134]. This lemma actually holds for any distribution over ground truth rankings, as we discuss in Section 4.6.

Lemma 4.9. *Suppose that for a given $\mathcal{B}_t(\pi)$, there exists a distribution D over subsets of A of size z such that*

$$\max_{\sigma \in \mathcal{B}_t(\pi)} \mathbb{E}_{Z \sim D} \left[\ell(Z, \sigma)\right] = k.$$

Then there exists $Z^* \subseteq A$ of size z whose average loss in $\mathcal{B}_t(\pi)$ is at most k.

Proof. Let U be the uniform distribution over rankings in $\mathcal{B}_t(\pi)$. Then clearly

$$\mathbb{E}_{Z \sim D, \sigma \sim U} \left[\ell(Z, \sigma)\right] \leq k,$$

as this inequality holds pointwise for all $\sigma \in \mathcal{B}_t(\pi)$. It follows there must exist at least one Z^* such that

$$\ell(Z^*, \mathcal{B}_t(\pi)) = \mathbb{E}_{\sigma \sim U} \left[\ell(Z^*, \sigma)\right] \leq k,$$

that is, the average loss of Z^* in $\mathcal{B}_t(\pi)$ is at most k. $\square$

Finally, we require a simple lemma of Procaccia et al. [109].

Lemma 4.10. *Given a profile π of n noisy rankings with average distance at most t from the ground truth according to a distance metric d, there exists $\sigma \in \mathcal{L}(A)$ such that for all $\tau \in \mathcal{B}_t(\pi)$, $d(\sigma, \tau) \leq 2t$.*

4.4.1 The KT and Footrule Distances

We first focus on the KT distance and the footrule distance. The KT distance is by far the most important distance metric over permutations, both in theory, and in practice (see Section 4.1.1). We study it together with the footrule distance because the two distances are closely related, as the following lemma, due to Diaconis and Graham [61], shows.

Lemma 4.11. *For all* $\sigma, \sigma' \in \mathcal{L}(A)$*,* $d_{KT}(\sigma,\sigma') \leq d_{FR}(\sigma,\sigma') \leq 2d_{KT}(\sigma,\sigma')$*.*

Despite this close connection between the two metrics, it is important to note that it does not allow us to automatically transform a bound on the loss for one into a bound for the other.

The next upper bound is, in our view, our most significant theoretical result. It is formulated for the footrule distance, but, as we show shortly, also holds for the KT distance.

Theorem 4.12. *For* $d = d_{FR}$*, given a profile* π *of* n *noisy rankings with average distance at most* t *from the ground truth, and a number* $z \in [m]$*, there always exists a subset of size* z *whose average loss in the set of feasible ground truths* $\mathcal{B}_t(\pi)$ *is at most* $O(\sqrt{t/z})$*.*

At some point in the proof, we will rely on the following (almost trivial) lemma.

Lemma 4.13. *Given two positive sequences of* k *real numbers,* P*, and* Q*, such that* P *is non-decreasing,* Q *is strictly decreasing and* $\sum_{i=1}^{k} P_i = C$*, the sequence* P *that maximizes* $S = \sum_{i=1}^{n} P_i Q_i$ *is constant, i.e.,* $P_i = C/k$ *for all* $i \in [k]$*.*

Proof. Assume for contradiction that P maximizes S and contains consecutive elements such that $P_j < P_{j+1}$. Now moving mass from P_{j+1} and distributing it to all lower positions in the sequence will strictly increase S. Concretely, if $P_{j+1} = P_j + \varepsilon$, we can subtract $j\varepsilon/(j+1)$ from P_{j+1} and add $\varepsilon/(j+1)$ to P_i for all $i \in [j]$. Because Q is strictly decreasing, this increases S by

$$\left(\sum_{i=1}^{j} \frac{Q_i \varepsilon}{j+1}\right) - \frac{Q_{j+1} j \varepsilon}{j+1} > \left(\sum_{i=1}^{j} \frac{Q_j \varepsilon}{j+1}\right) - \frac{Q_{j+1} j \varepsilon}{j+1} = \frac{j\varepsilon}{j+1}(Q_j - Q_{j+1}) > 0,$$

contradicting the assumption that P maximizes S. We may conclude that P is constant. □

Proof of Theorem 4.12. By Lemma 4.9, it is sufficient to construct a randomized rule that has expected loss at most $O(\sqrt{t/z})$ on *any* ranking in $\mathcal{B}_t(\pi)$. To this end, let $\sigma \in \mathcal{L}(A)$ such that $d(\sigma, \tau) \leq 2t$ for any $\tau \in \mathcal{B}_t(\pi)$; its existence is guaranteed by Lemma 4.10. Let $k = \sqrt{tz}$, and assume for ease of exposition that k is an integer. For $y = 1, \ldots, k$, let $a_y = \sigma^{-1}(y)$. Our randomized rule simply selects z alternatives uniformly at random from the top k alternatives in σ, that is, from the set $T \triangleq \{a_1, \ldots, a_k\}$. So, fixing some $\tau \in \mathcal{B}_t(\pi)$, we need to show that choosing z elements uniformly at random from the worst-case positions occupied by T in τ has expected minimum position at most $O(\sqrt{t/z})$.

Let Y^{σ}_{min} be the minimum position in σ of a random subset of size z from T. By Lemma 4.8 and Equation (4.1), we have

$$\mathbb{E}[Y^{\sigma}_{min}] = \sum_{y=1}^{k} y \frac{\binom{k-y}{z-1}}{\binom{k}{z}} = \frac{k+1}{z+1}.$$

However, we are interested in the positions of these elements in $\tau \in \mathcal{B}_t(\pi)$, not σ. Instead of appearing in position y, alternative a_y appears in position $p_y \triangleq \tau(a_y)$. Therefore, the expected minimum position in τ is

$$\mathbb{E}[Y^{\tau}_{min}] = \sum_{y=1}^{k} p_y \frac{\binom{k-y}{z-1}}{\binom{k}{z}}.$$

We wish to upper bound $\mathbb{E}[Y^{\tau}_{min}]$. Equivalently, because $\mathbb{E}[Y^{\sigma}_{min}]$ is fixed and independent of τ, it is sufficient to maximize the expression

$$\begin{aligned}\mathbb{E}[Y^{\tau}_{min}] - \mathbb{E}[Y^{\sigma}_{min}] &= \sum_{y=1}^{k} p_y \frac{\binom{k-y}{z-1}}{\binom{k}{z}} - \sum_{y=1}^{k} y \frac{\binom{k-y}{z-1}}{\binom{k}{z}} \\ &= \sum_{y=1}^{k} (p_y - y) \frac{\binom{k-y}{z-1}}{\binom{k}{z}}.\end{aligned} \tag{4.3}$$

Let us now assume that $p_y < p_{y+1}$ for all $y \in [k-1]$, that is, τ and σ agree on the order of the alternatives in T; we will remove this assumption later. Since the original positions of the alternatives in T were $\{1, \ldots, k\}$ it follows that $p_y \geq y$ for all $y \in [k]$. Moreover, because

$$\frac{\binom{k-y}{z-1}}{\binom{k}{z}} > \frac{\binom{k-(y+1)}{z-1}}{\binom{k}{z}},$$

the sequence of probabilities

$$Q = \left\{ \frac{\binom{k-y}{z-1}}{\binom{k}{z}} \right\}_{y \in [k]}$$

is strictly decreasing in y. Additionally, the sequence $P = \{p_y - y\}_{y \in [k]}$ is non-decreasing, because $p_{y+1} > p_y$, coupled with the fact that both values are integers, implies that $p_{y+1} \geq p_y + 1$.

In light of these facts, let us return to Equation (4.3). We wish to maximize

$$\mathbb{E}[Y^{\tau}_{min}] - \mathbb{E}[Y^{\sigma}_{min}] = \sum_{y=1}^{k} (p_y - y) \frac{\binom{k-y}{z-1}}{\binom{k}{z}} = \sum_{y=1}^{k} P_y Q_y.$$

By Lemma 4.13, $p_y - y$ is the same for all $y \in [k]$, that is, all alternatives in T are shifted by the same amount from σ to form τ. Moreover, we have that

$$\sum_{y=1}^{k} (p_y - y) \leq d(\sigma, \tau) \leq 2t.$$

Using $k = |T| = \sqrt{zt}$, we conclude that $p_y - y \leq 2\sqrt{t/z}$ for all $y \in [k]$. Therefore, in the worst $\tau \in \mathcal{B}_t(\pi)$, we have that the alternatives in T occupy positions $2\sqrt{t/z} + 1$ to $2\sqrt{t/z} + \sqrt{tz}$ in τ. By Lemma 4.8, the expected minimum position of T in τ is

$$2\sqrt{\frac{t}{z}} + \frac{\sqrt{tz} + 1}{z + 1} = O\left(\sqrt{\frac{t}{z}}\right).$$

To complete the proof, it remains to show that our assumption that $p_y < p_{y+1}$ for all $y \in [k-1]$ is without loss of generality. To see this, note that since we are selecting uniformly at random from T, Y^{τ}_{min} only depends on the positions occupied by T in τ. Moreover, if τ does not preserve the order over T, we can find a ranking τ' that has the following properties:

1. $d(\sigma, \tau') \leq 2t$.
2. T occupies the same positions: $\{\tau(a_1), \ldots, \tau(a_k)\} = \{\tau'(a_1), \ldots, \tau'(a_k)\}$.
3. τ' preserves the order over T: $\tau'(a_y) < \tau'(a_{y+1})$ for all $y \in [k-1]$.

Now all our arguments would apply to τ', and $\mathbb{E}[Y^{\tau}_{min}] = \mathbb{E}[Y^{\tau'}_{min}]$.

In order to construct τ', suppose that $\tau(a_y) > \tau(a_{y+1})$, and consider τ'' that is identical to τ except for swapping a_y and a_{y+1}. Then

$$\begin{aligned} d(\tau'', \sigma) = d(\tau, \sigma) &+ |\tau''(a_y) - y| + |\tau''(a_{y+1}) - (y+1)| \\ &- |\tau(a_y) - y| - |\tau(a_{y+1} - (y+1)| \\ \leq d(\tau, \sigma) &\leq 2t. \end{aligned}$$

By iteratively swapping alternatives we can easily obtain the desired τ'. $\square$

We next formulate the same result for the KT distance. The proof is very similar, so instead of repeating it, we just give a proof sketch that highlights the differences.

Theorem 4.14. *For $d = d_{KT}$, given a profile π of n noisy rankings with average distance at most t from the ground truth, and a number $z \in [m]$, there always exists a subset of size z whose average loss in the set of feasible ground truths $\mathcal{B}_t(\pi)$ is at most $O(\sqrt{t/z})$.*

Proof sketch. The proof only differs from the proof of Theorem 4.14 in two places.

First, the footrule proof had the inequality

$$\sum_{y=1}^{k} (p_y - y) \leq d_{FR}(\sigma, \tau) \leq 2t.$$

In our case,

$$\sum_{y=1}^{k} (p_y - y) \leq d_{FR}(\sigma, \tau) \leq 2 \cdot d_{KT}(\sigma, \tau) \leq 4t,$$

where the second inequality follows from Lemma 4.11.

Second, if τ does not preserve the order over T, we needed to find a ranking τ' that has the following properties:

1. $d(\sigma, \tau') \leq 2t$.
2. T occupies the same positions: $\{\tau(a_1), \ldots, \tau(a_k)\} = \{\tau'(a_1), \ldots, \tau'(a_k)\}$.
3. τ' preserves the order over T: $\tau'(a_y) < \tau'(a_{y+1})$ for all $y \in [k-1]$.

To construct τ' under $d = d_{KT}$, we use the same strategy as before: Suppose that $\tau(a_y) > \tau(a_{y+1})$, and consider τ'' that is identical to τ except for swapping a_y and a_{y+1}. We claim that $d(\tau', \sigma) \leq d(\tau, \sigma) \leq 2t$. Indeed, notice that all $a \in T$ precede all $b \in A \setminus T$ in σ. Therefore, holding all else equal, switching the relative order of alternatives in T will not change the number of pairwise disagreements on alternatives $b \in T$, $b' \in A \setminus T$, nor will it change the number of pairwise disagreements on alternatives $b, b' \in A \setminus T$. It will only (strictly) decrease the number of disagreements on alternatives in T. $\square$

Our next result is a lower bound of $\Omega(\sqrt{t}/z)$ for both distance metrics. Note that here z is outside the square root, i.e., there is a gap of $\sqrt{z}$ between the upper bounds given in Theorems 4.12 and 4.14, and the lower bound. That said, the lower bound is tight for a constant z, including the important case of $z = 1$.

Theorem 4.15. *For $d \in \{d_{FR}, d_{KT}\}$, $z \in [m]$, and an even n, there exist $t = O(m^2)$ and a profile π of n noisy rankings with average distance at most t from the ground truth, such that for any subset of size z, its average loss in the set of feasible ground truths $\mathcal{B}_t(\pi)$ is at least $\Omega(\sqrt{t}/z)$.*

Proof. We first prove the theorem for the KT distance, that is, $d = d_{KT}$. For any $k \geq 1$, let $t = \binom{k}{2}/2$; equivalently, let

$$k = \frac{1 + \sqrt{1 + 16t}}{2} = \Theta\left(\sqrt{t}\right).$$

Let $\sigma = (a_1 \succ \cdots \succ a_m)$, and let $\sigma_{R(k)} = (a_k, a_{k-1}, \ldots, a_1, a_{k+1}, \ldots, a_m)$ be the ranking that reverses the first k alternatives of σ. Consider the vote profile π with $n/2$ copies of each ranking σ and $\sigma_{R(k)}$.

Let $A_k = \{a_1, \ldots, a_k\}$ and denote by σ_{-k} the ranking of $A \setminus A_k$ ordered as in σ. We claim that $\mathcal{B}_t(\pi) = \mathcal{F}(A_k \succ \sigma_{-k})$, i.e., exactly the rankings that have some permutation of A_k in the first k positions, and coincide with σ in all the other positions. Indeed, consider any $\tau \in \mathcal{L}(A)$. This ranking will disagree with exactly one of σ and $\sigma_{R(k)}$ on every pair of alternatives in A_k, so

$$d(\tau, \pi) \geq \frac{\binom{k}{2}}{2} = t.$$

It follows that if $\tau \in \mathcal{B}_t(\pi)$ then τ must agree with σ_{-k} on the remaining alternatives.

Now let Z be a subset of z alternatives. Note that for every $a \in A \backslash A_k$ and $\tau \in \mathcal{B}_t(\pi)$, $\tau(a) > k$, so it is best to choose $Z \subset A_k$. We are interested in the expected loss of Z under the uniform distribution on $\mathcal{B}_t(\pi)$, which amounts to a random permutation of A_k. This is the same as choosing z positions at random from $[k]$. By Lemma 4.8, the expected minimum position of a randomly chosen subset of size z is $\frac{k+1}{z+1}$. Since $k = \frac{1+\sqrt{1+16t}}{2}$, it holds that

$$\mathbb{E}[Y_{min}] = \frac{\frac{1+\sqrt{1+16t}}{2} + 1}{z+1} = \Omega\left(\frac{\sqrt{t}}{z}\right).$$

For $d = d_{FR}$, the construction is analogous to above, with one minor modification. For any $k \geq 1$, we let $t = \lfloor k^2/2 \rfloor/2$, because the footrule distance between σ and $\sigma_{R(k)}$ is $\lfloor k^2/2 \rfloor$, instead of $\binom{k}{2}$ as in the KT case. Now, the proof proceeds as before. $\square$

An important remark is in order. Suppose that instead of measuring the average loss of the subset Z in $\mathcal{B}_t(\pi)$, we measured the *maximum* loss in any ranking in $\mathcal{B}_t(\pi)$, in the spirit of the model of Procaccia et al. [109]. Then the results would be *qualitatively* different. To see why on an intuitive level, consider the KT distance, and suppose that the vote profile π consists of n copies of the same ranking σ. Then for any $a \in A$, $\mathcal{B}_t(\pi)$ includes a ranking σ' such that $\sigma'(a) \geq t$ (by using our "budget" of t to move a downwards in the ranking). Therefore, for $z = 1$, it is impossible to choose an alternative whose maximum position (i.e., loss) in $\mathcal{B}_t(\pi)$ is smaller than t. In contrast, Theorem 4.12 gives us an upper bound of $O(\sqrt{t})$ in our model.

4.4.2 The Maximum Displacement Distance

We now turn to the maximum displacement distance. Here the bounds are significantly worse than in the KT and footrule settings. On an intuitive level, the reason is that two rankings that are at maximum displacement distance t from each other can be drastically different, because *every* alternative can move by up to t positions. Therefore, $\mathcal{B}_t(\pi)$ under maximum displacement would typically be larger than under the distance metrics we previously considered. Indeed, this is the case in Example 4.1 if one sets $t_{MD} \geq 1.5$.

Theorem 4.16. *For $d = d_{MD}$, given a profile π of n noisy rankings with average distance at most t from the ground truth, and a number $z \in [m]$,* *there always exists a subset of size z whose average loss in the set of feasible ground truths $\mathcal{B}_t(\pi)$ is at most $O(t/z)$.*

Proof. By Lemma 4.9, it is sufficient to construct a randomized rule that has expected loss at most $O(t/z)$ on *any* ranking in $\mathcal{B}_t(\pi)$. To this end, let $\sigma \in \mathcal{L}(A)$ such that $d(\sigma, \tau) \leq 2t$ for any $\tau \in \mathcal{B}_t(\pi)$; its existence is guaranteed by Lemma 4.10. For $y = 1, \ldots, 3t$, let $a_y = \sigma^{-1}(y)$. Our randomized rule selects z alternatives uniformly at random from the top $3t$ alternatives in σ, that is, from the set $T \triangleq \{a_1, \ldots, a_{3t}\}$.

Let T' be the top t alternatives in a ranking $\tau \in \mathcal{B}_t(\pi)$. Since $d(\sigma, \tau) \leq 2t$, we know that $T' \subset T$. Moreover, for any $a_y \in T$, we have that $p_y \triangleq \tau(a_y) \leq 5t$. Assume without loss of generality that $p_y \leq p_{y+1}$ for all $y \in [3t-1]$; then we have that the vector of positions $(p_1, \ldots, p_{3t})$ is pointwise at least as small as the vector $(1, 2, \ldots, t, 5t, 5t, \ldots, 5t)$. Using Lemma 4.8 and Equation (4.1), we conclude that the minimum position in τ when selecting z alternatives uniformly at random from T, denoted Y^{τ}_{min}, satisfies

$$\begin{aligned}
\mathbb{E}[Y^{\tau}_{min}] &= \sum_{y=1}^{3t} p_y \frac{\binom{3t-y}{z-1}}{\binom{3t}{z}} = \sum_{y=1}^{t-1} p_y \frac{\binom{3t-y}{z-1}}{\binom{3t}{z}} + \sum_{y=t}^{3t} p_y \frac{\binom{3t-y}{z-1}}{\binom{3t}{z}} \\
&\leq \sum_{y=1}^{t-1} y \frac{\binom{3t-y}{z-1}}{\binom{3t}{z}} + \sum_{y=t}^{3t} 5t \frac{\binom{3t-y}{z-1}}{\binom{3t}{z}} \\
&\leq 5 \cdot \sum_{y=1}^{3t} y \frac{\binom{3t-y}{z-1}}{\binom{3t}{z}} = 5 \cdot \frac{3t+1}{z+1} = \Theta\left(\frac{t}{z}\right).
\end{aligned}$$

□

We next establish a lower bound of $\Omega(t/z)$ on the average loss achievable under the maximum displacement distance. Note that this lower bound matches the upper bound of Theorem 4.16.

Theorem 4.17. *For $d = d_{MD}$, given $k \in \mathbb{N}$ and $z \in [m]$, there exist $t = \Theta(k)$ and a vote profile π of $k!$ noisy votes at average distance at most t from the ground truth, such that for any subset of size z, its average loss in the set of feasible ground truths $\mathcal{B}_t(\pi)$ is at least $\Omega(t/z)$.*

Proof. Let $\pi = \mathcal{F}(A_k \succ \sigma^{A \setminus A_k})$, where $|A_k| = k$. For some $\tau \in \pi$, let $t = d(\tau, \pi)$. By symmetry, $\tau' \in \mathcal{B}_t(\pi)$ for all $\tau' \in \pi$.

We first claim that $t = \Omega(k)$. Indeed, t is the average distance between τ and π. Letting U be the uniform distribution over π, we have that $t = \mathbb{E}_{\tau' \sim U}[d(\tau, \tau')]$. Now consider the top-ranked alternative in τ, $a \triangleq \tau^{-1}(1)$. Because U amounts to a random permutation over A_k, it clearly holds that $\mathbb{E}_{\tau' \sim U}[\tau'(a)] = (k+1)/2$, and therefore

$$\begin{aligned} t = \mathbb{E}_{\tau' \sim U}[d(\tau, \tau')] &= \mathbb{E}_{\tau' \sim U}\left[\max_{b \in A} |\tau'(b) - \tau(b)|\right] \\ &\geq \mathbb{E}_{\tau' \sim U}[\tau'(a) - \tau(a)] = \frac{k+1}{2} - 1 = \Omega(k). \end{aligned}$$

Now, suppose that we have shown that $\mathcal{B}_t(\pi) = \pi$; we argue that the theorem follows. Let $Z \subseteq A$ be a subset of alternatives of size z. We can assume without loss of generality that $Z \subseteq A_k$, as A_k is ranked at the top of every $\tau \in \mathcal{B}_t(\pi)$. But because $\mathcal{B}_t(\pi)$ consists of all permutations of A_k, $\ell(Z, \mathcal{B}_t(\pi))$ is equal to the expected minimum position when z elements are selected uniformly at random from the positions occupied by A_k, namely $[k]$. That is, we have that

$$\ell(Z, \mathcal{B}_t(\pi)) = \frac{k+1}{z+1} = \Omega\left(\frac{t}{z}\right).$$

Therefore, it only remains to show that $\mathcal{B}_t(\pi) = \pi$. Indeed, let $\tau \notin \pi$, then there exists $a \in A_k$ such that $\tau(a) > k$. Without loss of generality assume a is unique and let $\tau(a) = k+1$. There must then be some $b \in A \setminus A_k$ with $\tau(b) \leq k$. Recall that the alternatives in $A \setminus A_k$ remain in fixed positions in π, and, again without loss of generality, suppose that $\sigma(b) = k+1$ for all $\sigma \in \pi$. We wish to show that $d(\tau, \pi) > d(\sigma, \pi)$ for all $\sigma \in \pi$.

Let τ' be τ except that a and b are swapped, so $\tau'(a) = \tau(b)$ and $\tau'(b) = \tau(a)$. Observe that $\tau' \in \pi$ since a is unique. By definition, $d(\tau', \pi) = d(\sigma, \pi)$ for all $\sigma \in \pi$. It is therefore sufficient to show that $d(\tau', \pi) < d(\tau, \pi)$.

To this end, we partition the rankings $\sigma \in \pi \setminus \{\tau'\}$ into two sets, analyze them separately, and in both cases show that $d(\tau', \sigma) \leq d(\tau, \sigma)$.

1. $\sigma(a) \leq \tau'(a)$ (see Figure 4.1): In this case, we have that $|\sigma(a) - \tau(a)| \geq |\sigma(a) - \tau'(a)|$. Also, because σ and τ' agree on the position of $b \in A \setminus A_k$, $0 = |\sigma(b) - \tau'(b)| \leq |\sigma(b) - \tau(b)|$. We conclude that $d(\tau', \sigma) \leq d(\tau, \sigma)$.

2. $\sigma(a) > \tau'(a)$ (see Figure 4.2): It again holds that $0 = |\sigma(b) - \tau'(b)| \leq |\sigma(b) - \tau(b)|$, so if $d(\tau', \sigma) > d(\tau, \sigma)$ then $d(\tau', \sigma)$ is determined by a

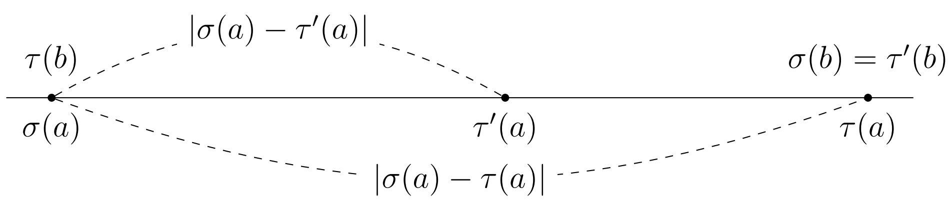


Figure 4.1: Illustration of Case 1 of the proof of Theorem 4.17.

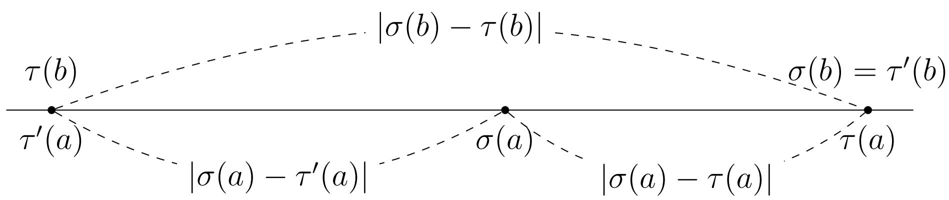


Figure 4.2: Illustration of Case 2 of the proof of Theorem 4.17.

(i.e., a has the maximum displacement). Assume for contradiction that $d(\tau', \sigma) > d(\tau, \sigma)$. It follows that

$$\begin{aligned} d(\tau', \sigma) &= |\sigma(a) - \tau'(a)| \\ &\leq |\sigma(a) - \tau'(a)| + |\sigma(a) - \tau(a)| = |\sigma(b) - \tau(b)| \leq d(\tau, \sigma), \end{aligned}$$

a contradiction. We may conclude that $d(\tau', \sigma) \leq d(\tau, \sigma)$.

Since $d(\tau', \sigma) \leq d(\tau, \sigma)$ for all $\sigma \in \pi \setminus \{\tau'\}$, and $d(\tau', \tau') = 0 < d(\tau, \tau')$ we may conclude that $d(\tau, \pi) > d(\tau', \pi) = t$. It follows that $\mathcal{B}_t(\pi) = \pi$, thereby completing the proof. $\square$

4.4.3 The Cayley Distance

In the previous sections, we have seen that our bounds are very different for different distance metrics. Still, all those bounds depended on t. By contrast, we establish a lower bound of $\Omega(\sqrt{m})$ on the average loss of any subset with $z = 1$ (i.e., the average position of any alternative) under the Cayley distance. We view this as a striking negative result: Even if the votes are extremely accurate, i.e., t is very small, the ball $\mathcal{B}_t(\pi)$ could be such that the average position of any alternative is as large as $\Omega(\sqrt{m})$.

Theorem 4.18. *For $d = d_{CY}$ and every $k \in [\sqrt{m}/3]$, there exists $t = \Theta(k)$ and a vote profile π with*

$$n = k!\binom{\sqrt{m}}{k}^2$$

noisy rankings at average distance at most t from the ground truth, such that for any single alternative, its average position in the set of feasible ground truths $\mathcal{B}_t(\pi)$ is at least $\Omega(\sqrt{m})$.

The theorem's proof appears in the full version of the paper. Note that the delicate construction is specific to the case of $z = 1$. It remains open whether the theorem still holds when, say, $z = 2$, and, more generally, how the bound decreases as a function of z.

Proof. Suppose for ease of exposition that $\sqrt{m} \in \mathbb{Z}$. Let $\sigma = (a_1 \succ a_2 \succ \ldots \succ a_m)$ be a ranking and let $L = \{1, 2, \ldots, \sqrt{m}\}$, $M = \{\sqrt{m}+1, \ldots, m - \sqrt{m}\}$ and $R = \{m - \sqrt{m}+1, \ldots, m\}$. Define the ranking σ_{ij} for $i \in L, j \in R$ to have $\sigma_{ij}(a_i) = \sigma(a_j)$ and $\sigma_{ij}(a_j) = \sigma(a_i)$ while $\sigma_{ij}(a_c) = \sigma(a_c)$ for all $c \in [m] \setminus \{i, j\}$. In other words, σ_{ij} is exactly σ with element $i \in L$ and element $j \in R$ swapped.

Construct a vote in π by selecting $S \subseteq L$, $T \subseteq R$ with $|S| = |T| = k$, then selecting a perfect matching $M : S \to T$, and finally swapping each a_i for $i \in S$ with a_j for $j = M(i)$. We have such a vote for every choice of S and T, and every perfect matching between them. This results in a vote profile of cardinality

$$n = |\pi| = k!\binom{\sqrt{m}}{k}^2.$$

Let $t = k + 1 - \frac{2k}{m}$. By construction $d(\tau, \sigma) = k$ for all $\tau \in \pi$. It follows that $d(\pi, \sigma) = k \leq t$, and therefore $\sigma \in \mathcal{B}_t(\pi)$.

We next claim that

$$d(\sigma_{ij}, \pi) \leq k + 1 - \frac{2k}{m} = t.$$

It suffices to consider two classes of rankings $\tau \in \pi$. First, if $\tau(a_i) = j = \sigma_{ij}(a_i)$ and $\tau(a_j) = i = \sigma_{ij}(a_j)$, then $d(\sigma_{ij}, \tau) \leq k - 1$, since reversing the other $k - 1$ pairwise swaps changes τ into σ_{ij}. There are

$$\hat{n} = \binom{\sqrt{m}-1}{k-1}^2 \cdot (k-1)!$$

such rankings in π. Second, for all other $\tau \in \pi$, we have $d(\sigma_{ij}, \tau) \leq k+1$, since it is always possible to reverse the k pairwise exchanges that changed σ into $\tau \in \pi$, and then perform one additional exchange to put a_i and a_j into the correct positions. It follows that for all $i \in L, j \in R$,

$$\begin{aligned} d_{CY}(\sigma_{ij}, \pi) &\leq \frac{1}{|\pi|}(k-1)\hat{n} + \frac{1}{|\pi|}(k+1)(|\pi| - \hat{n}) \\ &= (k+1) + \frac{(k-1)\hat{n} - (k+1)\hat{n}}{|\pi|} = (k+1) - \frac{2\hat{n}}{|\pi|} \\ &= (k+1) - 2 \cdot \frac{\binom{\sqrt{m}-1}{k-1}^2 \cdot (k-1)!}{|\pi|} = k+1-\frac{2k}{m}. \end{aligned}$$

We conclude that $\{\sigma\} \cup \{\sigma_{ij} : i \in L, j \in R\} \subseteq \mathcal{B}_t(\pi)$.

We next show that this, in fact, fully describes $\mathcal{B}_t(\pi)$. To show this, we must use the Hamming distance, denoted d_{HM}, which is defined as the number of positions at which two rankings of the same length differ. In particular, we use the relationship $d_{CY}(\tau, \tau') \geq \frac{1}{2} d_{HM}(\tau, \tau')$ between the Cayley and Hamming distance metrics for all $\tau, \tau' \in \mathcal{L}(A)$. This is a direct result of the fact that a single swap can place at most two alternatives in their correct positions.

For an arbitrary $\tau' \in \mathcal{L}(A)$ we can decompose the Hamming distance metric as

$$\begin{aligned} d_{HM}(\tau', \pi) &= \frac{1}{|\pi|}\sum_{\tau \in \pi} d_{HM}(\tau', \tau) = \frac{1}{|\pi|}\sum_{\tau \in \pi}\sum_{i \in [m]} \mathbb{I}[\tau(a_i) \neq \tau'(a_i)] \\ &= \sum_{i \in [m]} \frac{1}{|\pi|} \sum_{\tau \in \pi} \mathbb{I}[\tau(a_i) \neq \tau'(a_i)] = \sum_{i \in [m]} q_i(\pi, \tau'), \end{aligned} \tag{4.4}$$

where

$$q_i(\pi, \tau') \triangleq \frac{1}{|\pi|} \sum_{\tau \in \pi} \mathbb{I}[\tau(a_i) \neq \tau'(a_i)]$$

is the average *penalty* that a_i incurs in τ' with respect to π under the Hamming distance metric.

Consider $q_i(\pi, \tau')$ for $i \in L$. If $\tau'(a_i) = i$, then $q_i(\pi, \tau') = k/\sqrt{m}$ since a_i is swapped with an alternative in the right endpoint in a $k/\sqrt{m}$ fraction of the rankings in π. If $\tau'(a_i) \in (L \setminus \{i\}) \cup M$, then a penalty is incurred in every $\tau \in \pi$, so $q_i(\pi, \tau') = 1$. If $\tau'(a_i) \in R$, then $q_i(\pi, \tau') = 1 - (k/\sqrt{m})(1/\sqrt{m}) =$

$1 - k/m$. The analysis for $q_i(\pi, \tau')$, $i \in R$ is identical. For $q_i(\pi, \tau')$, $i \in M$, observe that $\tau(a_i) = i$ for all $\tau \in \pi$, so $q_i(\pi, \tau') = 0$ if $\tau'(a_i) = i$ and 1 otherwise.

It is clear from the decomposition and above discussion that $\tau' = \sigma$ is the unique ranking minimizing $d_{HM}(\tau', \pi)$. We partition the rankings $\tau' \in \mathcal{L}(A)$ according to their Hamming distance from σ and analyze which rankings can appear in $\mathcal{B}_t(\pi)$.

1. $d_{HM}(\tau', \sigma) = 1$: The Hamming distance metric does not allow rankings at distance 1 from each other.

2. $d_{HM}(\tau', \sigma) = 2$: We have shown that $\sigma_{ij} \in \mathcal{B}_t(\pi)$. If $\tau' \notin \{\sigma_{ij} : i \in L, j \in R\}$, then $d(\tau', \tau) = k+1$ for all $\tau \in \pi$ and thus $\tau' \notin \mathcal{B}_t(\pi)$. This is because the Cayley distance between σ and any $\tau \in \pi$ is exactly k due to the k pairwise disjoint swaps described above, and τ' involves an additional swap that is not allowed when transforming σ into $\tau \in \pi$.

3. $d_{HM}(\tau', \sigma) \geq 3$: For every ranking $\tau' \in \mathcal{L}(A)$ at Hamming distance at least 3 from σ, it holds that $\tau'(a_i) \neq i$ for at least three values of i, and therefore at least three of the penalties in Equation (4.4) are not minimal, meaning that they are at least $1 - k/m$. Moreover, the minimal penalty for $i \in L \cup R$ is $k/\sqrt{m}$. It follows that

$$\begin{aligned} d_{CY}(\tau', \pi) &\geq \frac{1}{2} d_{HM}(\tau', \pi) \\ &\geq \frac{1}{2}\left[\frac{k}{\sqrt{m}}(2\sqrt{m} - 3) + 3\left(1 - \frac{k}{m}\right)\right] \\ &= k + \frac{3}{2} - \frac{3k}{2m} - \frac{3k}{2\sqrt{m}} \\ &= k + 1 - \frac{2k}{m} + \left(\frac{1}{2} + \frac{k}{2m} - \frac{3k}{2\sqrt{m}}\right) \\ &\geq k + 1 - \frac{2k}{m} + \left(\frac{1}{2} + \frac{k}{2m} - \frac{1}{2}\right) \\ &= k + 1 - \frac{2k}{m} + \frac{k}{2m} \;>\; k + 1 - \frac{2k}{m}, \end{aligned}$$

where the fifth transition follows from the assumption that $k \leq \sqrt{m}/3$.

We conclude that $\mathcal{B}_t(\pi) = \{\sigma\} \cup \{\sigma_{ij} : i \in L, j \in R\}$ and thus that $|\mathcal{B}_t(\pi)| = m + 1$.

To complete the proof, we show that every alternative has average position at least $\Omega(\sqrt{m})$ in $\mathcal{B}_t(\pi)$. For every a_i with $i \in L$, a_i appears in position $j \in R$ in $\sqrt{m}$ of the $m+1$ rankings in $\mathcal{B}_t(\pi)$. Therefore the average loss of a_i over $\mathcal{B}_t(\pi)$ is at least

$$\frac{m+1-\sqrt{m}}{m+1} \cdot 1 + \frac{\sqrt{m}}{m+1} \cdot \frac{m}{2} = \Omega(\sqrt{m}).$$

For $i \in M$, alternative a_i never appears in position smaller than $\sqrt{m}+1$ in $\mathcal{B}_t(\pi)$ and clearly has average position $\Omega(\sqrt{m})$. Finally, for $j \in R$, alternative a_j appears in position j in at least $m+1-\sqrt{m}$ of the rankings in $\mathcal{B}_t(\pi)$, and also has average position at least $\Omega(\sqrt{m})$.

$\square$

4.5 Making the right decisions, in practice

We have two related goals in practice, to recover a ranking that is close to the ground truth, and identify a subset of alternatives with small loss in the ground truth. We compare the optimal rules that minimize the *average* distance or loss on $\mathcal{B}_t(\pi)$, denoted AVG^d, which we developed, to those that minimize the *maximum* distance or loss, denoted MAX^d, which were developed by Procaccia et al. [109]. Importantly, at least for the case where the output is a ranking, Procaccia et al. [109] have compared their methods against a slew of previously studied methods — including MLE rules for famous random noise models like the one due to Mallows [92] — and found theirs to be superior. In addition, their methods are the ones currently used in practice, on RoboVote. Therefore we focus on comparing our methods to theirs.

Datasets. Like Procaccia et al. [109], we make use of two real-world datasets collected by Mao et al. [93]. In both of these datasets — *dots* and *puzzle* — the ground truth rankings are known, and data was collected via Amazon Mechanical Turk. Dataset *dots* was obtained by asking workers to rank four images containing different numbers of dots in increasing order. Dataset *puzzle* was obtained by asking workers to rank four different states of a puzzle according to the minimal number of moves necessary to reach the goal state. Each dataset consists of four different noise levels, corresponding to levels of difficulty, represented using a single noise parameter. In *dots*, higher noise corresponds to smaller differences between the number of dots in the images, whereas in *puzzle*, higher noise entails ranking states that are all a constant number of steps further from the goal state. Overall the two datasets contain thousands of votes — 6367, to be precise.

Experimental design When recovering complete rankings, the evaluation metric is the distance of the returned ranking to the *actual* (known) ground truth. We reiterate that, although MAX^d is designed to minimize the maximum distance to any feasible ground truth given an input profile π and an estimate of the average noise t, that is, it is designed for the worst case, it is known to work well in practice [109]. Similarly, AVG^d is designed to optimize the average distance to the set of feasible ground truths; our experiments will determine whether this is a useful proxy for minimizing the distance to an unknown ground truth.

When selecting a subset of alternatives, the evaluation metric is the loss of that subset in the *actual* ground truth. As discussed above, the current implementation of RoboVote uses the rule MAX^d that returns the set of alternatives that minimizes the maximum loss in any feasible true ranking. As in the complete ranking setting, the rule AVG^d returns the set of alternatives that minimizes the average loss over the feasible true rankings.

It is important to emphasize that in both these settings, MAX^d and AVG^d optimize an objective over the set of feasible ground truths, but are evaluated on the actual known ground truth. It is therefore impossible to predict in advance which of the methods will perform best.

Our theoretical results assume that an upper bound t on the average error is given to us, and our guarantees depend on this bound. In practice, though, t has to be estimated. For example, the current RoboVote implementation uses $t_{RV} = \min_{\sigma \in \pi} d(\sigma, \pi)/|\pi|$, or the minimum average distance from one ranking in π to all other rankings in π.

In our experiments, we wish to study the impact of the choice of t on the performance of AVG^d and MAX^d. A natural choice is $t^* \triangleq d(\pi, \sigma^*)$, where π is the vote profile and σ^* is the actual ground truth. That is, t^* is the average distance between the vote profile and the actual ground truth. In principle it is an especially good choice because it induces the smallest ball $\mathcal{B}_t(\pi)$ that contains the actual ground truth. However, it is also an impractical choice, because one cannot compute this value without knowing the ground truth.

We also consider

$$t_{KEM} \triangleq \min_{\sigma \in \mathcal{L}(A)} d(\sigma, \pi)$$

(named after the *Kemeny* rule) — the minimum possible distance between the vote profile and any ranking.

In order to synchronize results across different profiles, we let $\hat{t}$ be the estimate of t that we feed into the methods, and define

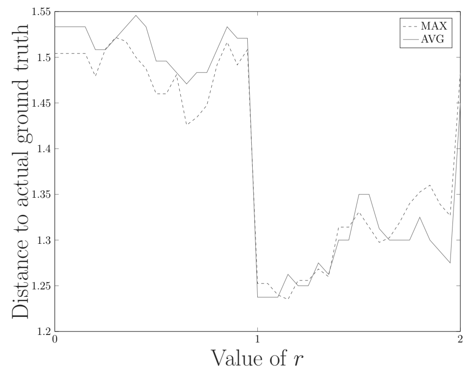
$$r = \frac{\hat{t} - t_{KEM}}{t^* - t_{KEM}}.$$

Note that because t_{KEM} is the minimum value that allows for a nonempty set of feasible ground truths, we know that $t^* - t_{KEM} \geq 0$. For any profile, $r = 0$ implies that $\hat{t} = t_{KEM}$, $r < 1$ implies that $\hat{t} < t^*$, $r = 1$ implies that $\hat{t} = t^*$, and $r > 1$ implies that $\hat{t} > t^*$. In our experiments, as in the work of Procaccia et al. [109], we use $r \in [0, 2]$.

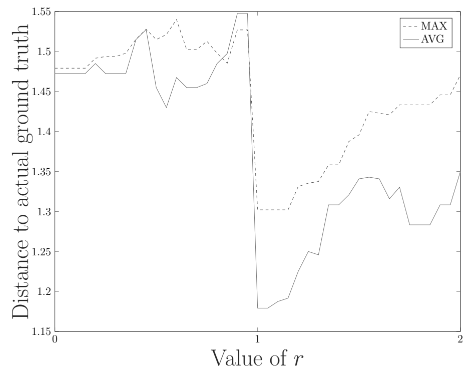
Results and interpretation Our results for three output types — ranking, subset with $z = 1$ (single winner), and subset with $z = 2$ — can be found in Figures 4.3, 4.4, and 4.5, respectively. Each has two subfigures, for the KT distance, and the Cayley distance. All Figures show r on the x axis. In Figure 4.3, the y axis shows the distance between the output ranking and the actual ground truth. In Figures 4.4 and 4.5, the y axis shows the loss of the selected subset on the actual ground truth. All figures are based on the dots dataset with the highest noise level (4). The results for the puzzle dataset are similar (albeit not as crisp), and the results for different noise levels are quite similar. The results differ across distance functions, but the conclusions below apply to all four, not just the two that are shown here.

It is interesting to note that, while in Figure 4.3 the accuracy of each distance metric is measured using that metric (i.e., KT is measured with KT and Cayley with Cayley), in the other two figures the two distances are measured in the exact same way: based on position or loss in the ground truth. Despite the dismal theoretical results for Cayley (Theorem 4.18), its performance in practice is comparable to KT.

More importantly, we see that although MAX^d and AVG^d perform similarly on low values of r, AVG^d significantly outperforms MAX^d on medium and high values of r, and especially when $r > 1$, that is, $\hat{t} > t^*$. This is true in all cases (including the two distance metrics that are not shown), except

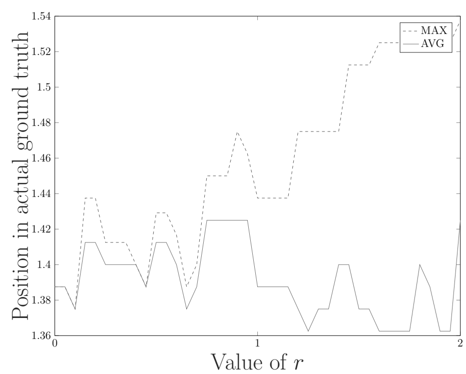


(a) Kendall tau

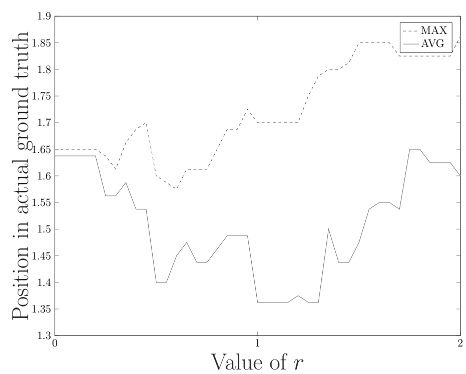


(b) Cayley

Figure 4.3: Dots dataset (noise level 4), ranking output.



(a) Kendall tau

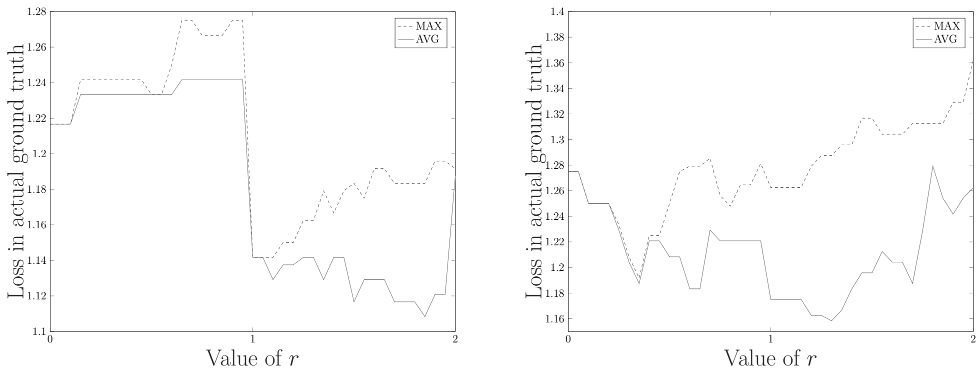


(b) Cayley

Figure 4.4: Dots dataset (noise level 4), subset output with $z = 1$.

for the ranking output type under the KT distance (Figure 4.3a) and the footrule distance (in the full version of the paper), where the performance of the two methods is almost identical across the board (values of r, datasets, and noise levels).

These results match our intuition. As r increases, so does $\hat{t}$, and the set $\mathcal{B}_{\hat{t}}(\pi)$ grows larger. When this set is large, the conservatism of MAX^d becomes a liability, as it minimizes the maximum distance with respect to rankings that are unlikely to coincide with the actual ground truth. By contrast, AVG^d is more robust: It takes the new rankings into account, but does not allow them to dictate its output.



(a) Kendall tau (b) Cayley

Figure 4.5: Dots dataset (noise level 4), subset output with $z = 2$.

The practical implication is clear. Because we do not have a way of divining t^*, which is often the most effective choice in practice, we resort to relatively crude estimates, such as the deployed choice of t_{RV} discussed above. Moreover, underestimating t^* is often risky, as the results show, because the ball $\mathcal{B}_{\hat{t}}(\pi)$ does not contain the actual ground truth when $\hat{t} < t^*$. Therefore in practice we try to aim for estimates such that $\hat{t} > t^*$, and robustness to the value of $\hat{t}$ is crucial. In this sense AVG^d is a better choice than MAX^d.

4.6 Discussion

We wrap up with a brief discussion of several key points.

Non-uniform distributions All of our upper bound results, namely Theorems 4.2, 4.12, 4.14, and 4.16, apply to any distribution over $\mathcal{B}_t(\pi)$, not just the uniform distribution (when replacing "average" distance/loss with "expected" distance/loss). To see why this is true for the latter three theorems, note that their proofs construct a randomized rule and leverage Lemma 4.9, which easily extends to any distribution. While this is a nice point to make, we do not believe that non-uniform distributions are especially well motivated — where would such a distribution come from? By contrast, the uniform distribution represents an agnostic viewpoint.

Computational complexity We have not paid much attention to computational complexity. In our experiments there are only four alternatives, so we can easily compute $\mathcal{B}_t(\pi)$ by enumeration. For real-world instances, integer programming is used, as we briefly discussed in Section 4.1.1. While those implementations are for rules that minimize the *maximum* distance or loss over $\mathcal{B}_t(\pi)$ [109], they can be easily modified to minimize the average distance or loss. Therefore, at least for the purposes of applications like RoboVote, computational complexity is *not* an obstacle.

Real-world implications As noted in Section 4.5, our empirical results suggest that minimizing the average distance or loss has a significant advantage in practice over minimizing the maximum distance or loss. We are therefore planning to continue refining our methods, and ultimately deploy them on RoboVote, where they will influence the way thousands of people around the world make group decisions.

Chapter 5

Dynamic fair division of indivisible goods

5.1 Introduction

We consider the setting of fairly allocating indivisible goods to agents who have additive valuations. Our goal is to have a strong mathematical guarantee of the interpersonal *fairness* of the resulting allocation. An *allocation* is a partition of the goods into bundles, such that each agent is assigned one bundle. Such assignment problems are not only frequently encountered in a myriad of operations research domains, but also often require the imposition of additional fairness conditions, such as in organ transplantation [27] and nurse shift scheduling [95]. A variety of fairness constraints are used in practice, but arguably the gold standard of fairness is *envy freeness*, which requires that each agent is at least as happy with her own allocation as the allocation of any other agent.

Envy-free solutions indeed always exist in well-studied fair division settings that involve *divisible* goods or a numéraire, such as cake cutting [34, 104] and rent division [71, 123]. When items are divisible, one strategy for finding a fair allocation is the *competitive equilibrium from equal incomes* (CEEI) solution of Varian [127]. In the equilibrium allocation, agents use assigned (equal) budgets to purchase their preferred bundles of goods at virtual prices, and the market clears (all goods are allocated). This solution is envy free [70] and coincides with the solution that maximizes the *Nash social welfare* [11], that is, the solution which maximizes the product of agent utilities.

By contrast, with *indivisible* goods, envy is clearly unavoidable in general — consider a single (indivisible) item that is desired by two agents. That is why recent papers [43, 89] focus on the relaxed notion of *envy-freeness up to one good* (EF1), in which envy may exist, but for any bundle that an agent prefers over her own, there exists a single good whose removal eliminates that envy. With indivisible goods, the approximate-CEEI solution [37] and the solution which maximizes the Nash social welfare are both EF1. In fact, EF1 is quite easy to guarantee, e.g., by allocating the items in a round-robin fashion where each agent in her turn picks her favorite item among those that are still available.

Our point of departure is that we allow items to arrive *online*, that is, we must choose how to allocate an item immediately and irrevocably at the moment it arrives, without knowing the values of items that will arrive in the future. This setup mirrors common decision-making scenarios in humanitarian logistics. A paradigmatic example is that of food banks [5], which receive food donations, and deliver them to nonprofit organizations such as food pantries and soup kitchens. Indeed, items are often perishable, which is why allocation decisions must be made quickly, and donated items are typically leftovers, leading to lack of information about items that will arrive in the future.

As noted, in the static setting there exists an EF1 solution for any number of items, but this requires complete information about values upfront. In contrast, in the online setting, one would expect the maximum envy to increase with the number of items. Nevertheless, we can hope to control the rate at which envy grows over time. Specifically, we aim to design algorithms with *vanishing envy* — algorithms that lead to envy growing sublinearly in the number of items allocated. Our primary research question is:

> *Are there online allocation algorithms with vanishing envy, and, if so, at what rate does envy vanish?*

5.1.1 Our Results

We mainly focus on the *full information* version of the problem, where the algorithm sees the agent valuations for an item before assigning that item. In Section 5.3.1, we study randomized algorithms. The most natural candidate is the "random allocation" algorithm: allocate each item to an agent chosen uniformly at random. We analyze this algorithm against an adaptive adversary that chooses the agent values for an arriving item after seeing the (realised) allocations of all the previous items. We show that the optimal strategy for an adaptive adversary (against random allocation) is, in fact, nonadaptive. This enables us to use standard concentration inequalities for bounding the overall envy. Our first result, Theorem 5.3, asserts that this algorithm has vanishing envy.

One may hope that it would be possible to do better than allocating blindly. Surprisingly, we show in Theorem 5.14 that the random allocation algorithm causes envy to vanish at an asymptotically optimal rate (up to logarithmic factors). However, despite its theoretical optimality, the random allocation algorithm is intuitively unappealing. We therefore turn our attention to deterministic algorithms in Section 5.3.2, and discover that natural, greedy schemes like allocating items to minimize envy fail miserably.

Nevertheless, as we prove in Theorem 5.7, there exists a deterministic polynomial-time algorithm with the same envy bound as the random allocation algorithm (up to logarithmic factors). The former algorithm is the result of derandomizing the latter with the method of *pessimistic estimators* [111]. Specifically, we define a potential function which is essentially a penalty function exponential in each of the pairwise envies, and show that allocating each item so as to minimize this potential function leads to asymptotically optimal envy.

Having completed the picture for the setting in which one item arrives at a time, we proceed to study a more general model in Section 5.4. Suppose, as before, that items with adversarially chosen values arrive over time. Instead of arriving one by one, as assumed in Section 5.3, items now arrive in *batches*. To motivate this, note that in the food bank setting it is reasonable to wait until the end of the day before allocating all food donations that arrived that day. When a batch arrives, the algorithm learns the values of all the items in the batch for all agents, and must allocate these items immediately and irrevocably, before the next batch arrives.

Since the allocation algorithm is less myopic in this setting, it is natural to expect stronger performance guarantees than when items arrive one by one. For example, in the extreme case where all items arrive simultaneously, there exist algorithms that are EF1, giving a constant bound on the envy. To realize this intuition, we leverage a result from the literature on continuous cake cutting, which allows us to show, for each batch, the existence of a fractional allocation that is entirely envy free and can be written as a convex combination of integral allocations with constant pairwise envy. We then use ideas from the derandomization employed in Section 5.3.2 to give, in Theorem 5.17, a deterministic algorithm for the batch setting, and find a nearly matching lower bound (Theorem 5.27).

In Section 5.5, we study a *partial information model*, a natural variant of the full information model where item values are only revealed to the algorithm *after* that item has been allocated. The randomized upper bound of Theorem 5.3 carries over, because the algorithm ignores the values shown to it. So does the lower bound of Theorem 5.14, because it holds *even* against more powerful algorithms. However, the game between the algorithm and the adversary is now an extensive-form game of *incomplete* information, where randomization may help. This turns out to be the case, and we show in Theorem 5.29 that deterministic algorithms cannot have vanishing envy in this setting.

5.1.2 Related Work

Conceptually our paper is related to the growing literature on *online* or *dynamic* fair division [4, 5, 84, 100, 128]. In particular, motivated by applications to the food bank domain, Aleksandrov et al. [5] introduce and analyze a closely related setting where indivisible items arrive online. However, they generally assume that all values are binary, i.e., each agent "likes" or "dislikes" every item. They introduce two simple mechanisms, Like and Balanced Like; the former allocates the current item uniformly at random among agents who like it, whereas the latter allocates the current item uniformly at random among agents who like it and have so far received the fewest items. The analysis of these mechanisms focuses on properties such as strategyproofness, envy-freeness, and impact on welfare. Most relevant to us is the observation that Balanced Like is EF1. This also highlights the technical differences between our setting and theirs, because, as noted above, with general values EF1 is impossible.

This work is also related to the vast body of work on *online learning* [36]. In the quintessential setting, *experts learning* (with full-information feedback), at each time step, the algorithm chooses to follow the advice of one of several experts. Then, the value of each expert is revealed, and the algorithm gains the value of the expert whose advice it chose to follow. The algorithm's regret is the difference between the total value accumulated by the best expert in hindsight and the value it itself has accumulated; a *no-regret learning algorithm* has the property that the ratio between regret and time goes to zero (*vanishing regret* may have been a more accurate term). Similarly, we are also interested in the difference in value accumulated over time. However, to the best of our knowledge the two problems are technically unrelated. To appreciate the difference, note that in our setting the values of the current item to all agents are known to the algorithm. But if the values of the different experts were known in the expert learning setting, the problem would be trivial — the algorithm would simply choose the expert with maximum value. Nevertheless, some of our notation was chosen to be consistent with that used in the online learning literature.

Finally, we can make a technical connection to the literature on *vector balancing games* [120]. At each time step, the adversary picks a vector and the algorithm chooses to either add or subtract this from a running partial sum vector. In one version of this game, the goal of the algorithm is to minimize the maximum entry of the partial sum vector, while the adversary wishes to maximize that quantity. When there are only two agents and items arrive one by one, our setting can be reduced to a version of vector balancing games equipped with a weaker adversary. This means that the upper bound of Spencer [120] applies to our setting (and matches our results). Conversely, our lower bound for the two agent setting matches the lower bound from that paper, indicating that the ostensibly weaker adversary that we consider — restricted to picking values from just one orthant — has roughly the same strength as the stronger adversary of Spencer [120]; consequently, our lower bound is significantly more involved. For more than two agents, the two problems appear unrelated, and, moreover, the batch setting has no equivalent in the vector balancing games literature.

5.2 Model

We consider a set $[n] \triangleq \{1, \ldots, n\}$ of agents, and a set of T items. Each agent $i \in [n]$ assigns a (normalized) value $v_{it} \in [0, 1]$ to each item $t \in [T]$; for a bundle of items S, the value of agent i is $v_i(S) \triangleq \sum_{t \in S} v_{it}$. The values are chosen by an adaptive adversary. An *allocation* is a partition of the items into bundles $A_1, \ldots, A_n$, where A_i is assigned to agent $i \in [n]$. The allocation is said to be *envy free* if $v_i(A_i) \geq v_i(A_j)$ for all $i, j \in [n]$. The allocation is EF1 when $v_i(A_i) \geq v_i(A_j) - \max_{t \in A_j} v_{it}$ for all $i, j \in [n]$.

When items arrive one at a time, an item t arrives at every step, where $t = 1, \ldots, T$. The allocation decisions made by an algorithm at each step induce an allocation $A_1, \ldots, A_n$ at the end of step T. For $i, j \in [n]$, let

$$\textsc{Envy}_T^{ij} \triangleq \max\{v_i(A_j) - v_i(A_i), 0\}$$

be the envy of agent i towards j. This measure of cumulative envy increases by v_{it} if item t is allocated to j (which places it in A_j), and decreases by v_{it} if item t is allocated to i (which places it in A_i). Moreover, let

$$\textsc{Envy}_T \triangleq \max_{i,j\in[n]} \textsc{Envy}_T^{ij}$$

be the maximum envy. We say that an algorithm has *vanishing envy* when $\textsc{Envy}_T \in o(T)$, equivalently, when $\lim_{T\to\infty} \textsc{Envy}_T/T = 0$.

We later consider the setting in which items arrive instead in batches of size $m \geq 1$ at a time, in which all m items need to be assigned when they arrive. We will use $\textsc{Envy}_{T,m}$ to denote the envy after the allocation of T items arriving in batches of size m, with the convention that $\textsc{Envy}_{T,1} = \textsc{Envy}_T$.

In the full information setting, the agents' values for items are revealed to the algorithm prior to allocation (and thus inform the algorithm's decision); in the partial information model of Section 5.5, they are revealed after allocation (so the algorithm only knows the current pairwise envies).

5.3 Single Arrivals under Full Information

In this most basic setting exactly one item arrives at each time step, and its value is revealed to the algorithm before the allocation is made.

For intuition, we begin with an example analyzing the performance of a greedy policy that may be utilized in practical dynamic fair division settings, such as at a food bank. Namely, when a good arrives, it is allocated to the agent who needs it most, i.e., to the most envious agent. We show that this policy does not lead to vanishing envy.

Example 5.1. *Consider the algorithm that at step t allocates the item to the agent with the maximum envy (if she has positive value for the item, and otherwise, say, allocates to the agent with the highest value for the item). We claim that this algorithm leads to $\textsc{Envy}_T \in \Omega(T)$ when items arrive one by one under the full information model.*

We construct an example where each agent envies the other after the second item is allocated. For $t \geq 3$, whenever agent i has maximum envy,

Table 5.1: Blindly allocating item t to the agent with the highest envy after $t-1$ allocations leads to constant per-round envy.

t	1	2	3	4	5	$\cdots$
Value of agent 1	$\boxed{1/2}$	1	$\boxed{\epsilon}$	1	$\boxed{\epsilon}$	$\cdots$
Value of agent 2	1/2	$\boxed{1/4}$	1	$\boxed{\epsilon}$	1	$\cdots$
Envy of agent 1	$-1/2$	1/2	$1/2-\epsilon$	$3/2-\epsilon$	$3/2-2\epsilon$	$\cdots$
Envy of agent 2	1/2	1/4	5/4	$5/4-\epsilon$	$9/4-\epsilon$	$\cdots$

we present an item with value ϵ for her, and value 1 for the other agent. Table 5.1 summarizes the analysis, the agent who receive an item has her valuation highlighted.

For $t \geq 2$, the envy of each agent increases by 1 every two steps. Therefore, the maximum envy at step $2t$ is approximately t, and $\textsc{Envy}_T/T$ approaches 1/2 as T goes to infinity. □

It turns out that it is nontrivial to find a deterministic algorithm for allocating the items in a manner that achieves vanishing envy. For instance, one might consider allocating every item in a way that minimizes the current maximum envy. The next example shows that this also leads to linear (or, equivalently, constant per-round) envy.

Table 5.2: Assigning item t so as to minimize the maximum envy after t allocations does not lead to vanishing envy.

t	1	2	3	4	5	6	$\cdots$
Value of agent 1	$\boxed{1/2}$	1	$\boxed{\epsilon}$	$\boxed{1/2}$	1	$\boxed{\epsilon}$	$\cdots$
Value of agent 2	1/2	$\boxed{1/4}$	$1/4-\epsilon$	1/2	$\boxed{1/4}$	$1/4-\epsilon$	$\cdots$
Envy of agent 1	$-1/2$	1/2	$1/2-\epsilon$	$-\epsilon$	$1-\epsilon$	$1-2\epsilon$	$\cdots$
Envy of agent 2	1/2	1/4	$1/2-\epsilon$	$1-\epsilon$	$3/4-\epsilon$	$1-2\epsilon$	$\cdots$

Example 5.2. *Consider the algorithm that at step t allocates the item in a way that the maximum envy after allocation is as small as possible. We claim this algorithm leads to $\textsc{Envy}_T \in \Omega(T)$. Table 5.2 summarizes the instance which proves this bound.*

The envy of each agent increases by $\frac{1}{2} - \epsilon$ every three steps. Therefore, the maximum envy at step $6t$ is approximately t, and $\textsc{Envy}_T/T$ approaches $1/6$ as T goes to infinity. □

As a result, we begin our analysis with an algorithm achieving vanishing envy that utilizes randomness, and this ultimately leads us to a deterministic algorithm with asymptotically optimal envy guarantees.

5.3.1 Upper bound via Random Allocation

A natural randomized algorithm for the case where items arrive one by one is to allocate each item to an agent selected uniformly at random; we refer to this as the *random allocation algorithm*. We analyze the random allocation algorithm by first characterizing the adversary's optimal strategy. We prove that for an adaptive adversary who maximizes $\mathbb{E}\left[\textsc{Envy}_T\right]$, where the expectation is with respect to the randomness of the algorithm, the optimal strategy is integral, that is, all values are in $\{0, 1\}$. Using this, we show that the optimal strategy, in fact, assigns $v_{it} = 1$ for all $i \in [n], t \in [T]$. This optimal adversary strategy is nonadaptive, and therefore, since all the randomness is coming from the algorithm, the random variables for the envy between agents i and j at times t and t' are independent. Standard concentration inequalities for the envy between any pair of agents, combined with a union bound over all such pairs, gives an upper bound on the expected envy.

Theorem 5.3. *Suppose that $T \geq n \log T$, where* $\log$ *is the natural logarithm. Then the random allocation algorithm guarantees that $\mathbb{E}\left[\textsc{Envy}_T\right] \in O(\sqrt{T \log T/n})$.*

Note that the assumption of $T \geq n \log T$ is innocuous, as otherwise we can give each agent at most $\log T$ items to achieve $\textsc{Envy}_T \leq \log T$.

Proof of Theorem 5.3. Typically we would think of an extensive-form game with nodes associated with the algorithm or the adversary, and arcs corresponding to actions (allocation of the current item in the case of the algorithm, value vector in the case of the adversary). However, because we consider a fixed algorithm, it is convenient to imagine an unusual, adversary-oriented game tree.

Consider a game tree with nodes on $T + 1$ levels. Every node on level $1, \ldots, T$ has n outgoing arcs labeled $1, \ldots, n$. The leaf nodes on level $T + 1$ are labeled by the maximum envy for the corresponding path. Let Ω be the set of all paths from the root to a leaf node, so $|\Omega| = n^T$. Equivalently, Ω is the set of all possible allocations of the T items. For an allocation $\omega \in \Omega$, denote by $\omega_t \in [n]$ the agent to whom item $t \in [T]$ was allocated by ω.

A fully adaptive strategy s for the adversary is defined by labeling every internal node u with a value vector $s(u)$, where $s(u)_i$ is the value of agent i for the item corresponding to node u. The algorithm's strategy consists of selecting an outgoing edge, corresponding to an allocation of the item with valuation $s(u)$, at every node u. The adversary's strategy is allowed to depend on the allocations and valuations so far, i.e., the path from the root to u.

For a given adversary strategy s and an allocation ω, let $\textsc{Envy}^{ij}(s, \omega)$ denote the envy of agent i for agent j. Denote with

$$\textsc{Envy}(s, \omega) \triangleq \max_{i,j \in [n]} \textsc{Envy}^{ij}(s, \omega)$$

the maximum envy experienced by any agent under adversary strategy s and allocation ω. The objective of the adversary is to choose a strategy s that maximizes the expected envy $\mathbb{E}[\textsc{Envy}(s, \omega)]$, where the expectation is taken over allocating every item uniformly at random.

We consider the algorithm that allocates every item uniformly at random. This is equivalent to picking a random outgoing edge at each node u. The following two lemmas show that the adversary labels every internal node of this tree with the vector $\mathbf{1}^n$. These lemmas are inspired by the work of Sanders [114] on load balancing. The next result follows from the fact that under any allocation algorithm, for every agent's valuation of any item, it is possible to compute whether that item increases or decreases the maximum envy (in expectation). If it increases (resp. decreases) the maximum envy, the adversary benefits by increasing (resp. decreasing) the corresponding valuation to 1 (resp. to 0).

Lemma 5.4. *The adversary has an optimal adaptive strategy that labels every internal node of the game tree with a vector in $\{0, 1\}^n$.*

Proof of Lemma 5.4. Assume for the sake of contradiction that the adversary does not have an optimal strategy which assigns integral vectors to the nodes of the (adversary-centric) game tree. Let s be the optimal strategy with the smallest number of fractional values. Without loss of generality, let u be a node on layer $\ell \in [T]$ for which the value assigned to player $i \in [n]$ is fractional, i.e., $0 < s(u)_i < 1$. The values ℓ and i are fixed for the remainder of this proof. Define alternative strategies s' and s'' identical to s, except that $s'(u)_i = 1$ and $s''(u)_i = 0$. We wish to arrive at the contradiction that $\mathbb{E}\left[\textsc{Envy}(s,\omega)\right] \leq \mathbb{E}[\textsc{Envy}(s^*,\omega)]$ for $s^* = s'$ or s'', where the expectation is over the randomness of the allocation algorithm. Denote with Ω_u all paths passing through u. The envy associated with paths in $\Omega \setminus \Omega_u$ is unaffected by the move from s to s' or s'' and may be safely ignored.

When agent i is not the unique agent with maximum envy, it holds that $\textsc{Envy}(s,\omega) \leq \textsc{Envy}(s',\omega)$ and $\textsc{Envy}(s,\omega) \leq \textsc{Envy}(s'',\omega)$ as desired (recall that changing agent i's valuation for an item does not affect other agents' envy). It remains to consider the set of paths

$$\Omega_u^+ \triangleq \left\{\omega \in \Omega : \max_{j\in[n]} \textsc{Envy}^{ij}(s,\omega) > \max_{j\in[n]\setminus\{i\}} \max_{k\in[n]} \textsc{Envy}^{jk}(s,\omega)\right\},$$

in which agent i is the unique agent with maximum envy (and this envy is strictly positive). We can further partition Ω_u^+ according to which agent receives item ℓ; let $\Omega_u^{+,j}$ be the set of paths in Ω_u^+ in which agent $j \in [n]$ gets item ℓ, and for any $J \subseteq [n]$, set $\Omega_u^{+,J} \triangleq \cup_{j\in J}\Omega_u^{+,j}$. We analyze three different cases: (1) whether the player that gets item ℓ is player i, (2) a player j^* for whom player i has maximum envy, or (3) another player. Define

$$J^* \triangleq \left\{j^* \in [n] : \textsc{Envy}^{ij^*}(s,\omega) = \max_{j\in[n]} \textsc{Envy}^{ij}(s,\omega)\right\}.$$

Also, for convenience, set $f \triangleq s(u)_i$ and $J^< \triangleq [n] \setminus \{J^* \cup \{i\}\}$.

We first look at s'. The three cases are:

1. For $\omega \in \Omega_u^{+,i}$: $\textsc{Envy}(s,\omega) - (1-f) \leq \textsc{Envy}(s',\omega) \leq \textsc{Envy}(s,\omega)$.
2. For $\omega \in \Omega_u^{+,J^*}$: $\textsc{Envy}(s',\omega) = \textsc{Envy}(s,\omega) + (1-f)$.
3. For $\omega \in \Omega_u^{+,J^<}$: $\textsc{Envy}(s,\omega) \leq \textsc{Envy}(s',\omega) \leq \textsc{Envy}(s,\omega) + (1-f)$.

The only outcomes where envy can decrease when changing the adversary's strategy from s to s' are those in $\Omega_u^{+,i}$. We can compute the effect of changing s to s' on the expected maximum envy as

$$\mathbb{E}[\textsc{Envy}(s,\omega)] = \sum_{\omega\in\Omega} \Pr[\omega] \cdot \textsc{Envy}(s,\omega)$$

$$
\begin{aligned}
&= \frac{1}{n^T}\left(\sum_{\omega\in\Omega_u^{+,i}} \textsc{Envy}(s,\omega) + \sum_{\omega\in\Omega_u^{+,J^*}} \textsc{Envy}(s,\omega) + \sum_{\omega\in\Omega_u^{+,J^<}} \textsc{Envy}(s,\omega)\right)\\
&\le \frac{1}{n^T}\cdot \sum_{\omega\in\Omega_u^{+,i}} (\textsc{Envy}(s',\omega)+(1-f)) + \frac{1}{n^T}\cdot \sum_{\omega\in\Omega_u^{+,J^*}} (\textsc{Envy}(s',\omega)-(1-f))\\
&\qquad + \frac{1}{n^T}\cdot \sum_{\omega\in\Omega_u^{+,J^<}} \textsc{Envy}(s',\omega)\\
&= \mathbb{E}[\textsc{Envy}(s',\omega)] + \frac{1-f}{n^T}\left(\left|\Omega_u^{+,i}\right| - \left|\Omega_u^{+,J^*}\right|\right).
\end{aligned}
$$

If $\left|\Omega_u^{+,i}\right| \le \left|\Omega_u^{+,J^*}\right|$, it follows that $\mathbb{E}[\textsc{Envy}(s,\omega)] \le \mathbb{E}[\textsc{Envy}(s',\omega)]$. Assume therefore that $\left|\Omega_u^{+,i}\right| > \left|\Omega_u^{+,J^*}\right|$. An identical analysis for s'' shows that

1. For $\omega \in \Omega_u^{+,i}$: $\textsc{Envy}(s'',\omega) = \textsc{Envy}(s,\omega) + f$.
2. For $\omega \in \Omega_u^{+,J^*}$: $\textsc{Envy}(s,\omega) - f \le \textsc{Envy}(s'',\omega) \le \textsc{Envy}(s,\omega)$.
3. For $\omega \in \Omega_u^{+,J^<}$: $\textsc{Envy}(s,\omega) = \textsc{Envy}(s'',\omega)$.

Expanding the computation of the expected value as before shows

$$
\mathbb{E}[\textsc{Envy}(s,\omega)] \le \mathbb{E}[\textsc{Envy}(s',\omega)] + \frac{f}{n^T}\left(-\left|\Omega_u^{+,i}\right| + \left|\Omega_u^{+,J^*}\right|\right).
$$

By assumption $\left|\Omega_u^{+,i}\right| > \left|\Omega_u^{+,J^*}\right|$, so $\mathbb{E}[\textsc{Envy}(s,\omega)] \le \mathbb{E}[\textsc{Envy}(s'',\omega)]$, concluding the proof. □

While the previous result holds for any allocation strategy, the following lemma leverages specific properties of the random allocation algorithm.

Lemma 5.5. *The adversary has an optimal adaptive strategy that labels every internal node of the game tree with the vector* $\mathbf{1}^n$*.*

Proof of Lemma 5.5. By Lemma 5.4, the adversary has an optimal strategy that labels every internal node with a vector in $\{0,1\}^n$. Let s be such an optimal strategy with the smallest number of zeros, and suppose (for the sake of contradiction) that there exist internal nodes that are not labeled $\mathbf{1}^n$. Let u on layer $\ell \in [T]$ be the node closest to a leaf node for which $s(u)$ contains a 0 and $s(u') = \mathbf{1}^n$ for all descendants u' of u. Without loss of generality assume

$s(u)_i = 0$, so agent i has value 0 for item ℓ at node u. Define a strategy s' identical to s except that $s'(u)_i = 1$. Let $j(\omega) \in \arg\max_{j \in [n]} \textsc{Envy}^{ij}(s, \omega)$.

For any fixed $\omega \in \Omega$, changing s to s' only changes the envy of agent i and only for paths that go through u. In particular, if $\omega_\ell \neq i$, the envy of agent i toward agent ω_ℓ increases by 1, which only helps the adversary. By contrast, if $\omega_\ell = i$, the envy of agent i decreases by 1, toward every agent j such that $\textsc{Envy}^{ij}(s, \omega) > 0$; the maximum envy, $\textsc{Envy}(s, \omega)$, is only affected if $\textsc{Envy}(s, \omega) = \textsc{Envy}^{i, j(\omega)}(s, \omega)$.

Thus, let $\omega \in \Omega$ be an arbitrary path going through u with $\omega_\ell = i$ and satisfying $\textsc{Envy}(s, \omega) = \textsc{Envy}^{i, j(\omega)}(s, \omega) > 0$. Since agent i may not have been the unique agent having envy equal to $\textsc{Envy}(s, \omega)$, $\textsc{Envy}(s', \omega) \geq \textsc{Envy}(s, \omega) - 1$. Now consider the path ω' that is identical to ω except that $\omega_\ell = j(\omega)$. Observe that $\textsc{Envy}(s', \omega') = \textsc{Envy}(s, \omega') + 1$. Hence, any decrease in envy due to allocating item ℓ to agent i on ω is compensated for (in the calculation of expected envy) along ω'. Since ω was picked arbitrarily and the mapping $\omega \mapsto \omega'$ is injective, it follows that the expected envy under s' is at least the expected envy under s, and s' has fewer zeros than s, contradicting our assumption on s. $\square$

The fact that the adversary is adaptive naturally introduces a dependence in the change in any pairwise envy from one arrival to the next. The value of Lemma 5.5 lies is that it allows us to circumvent this dependence as though we are dealing with a nonadaptive adversary and express any pairwise envy as the sum of independent random variables.

Specifically, given this adversary strategy, define independent random variables

$$X_t^{ij} \triangleq \begin{cases} -1, & \text{with probability } 1/n, \\ 0, & \text{with probability } 1 - 2/n, \\ 1, & \text{with probability } 1/n \end{cases}$$

for all $t \in [T]$, $i, j \in [n]$. Clearly, $\textsc{Envy}_T^{ij} = \max_{i,j \in [n]} \{\sum_{t=1}^T X_t^{ij}, 0\}$. For each X_t^{ij}, $\mathbb{E}[X_t^{ij}] = 0$, $\mathbb{E}[(X_t^{ij})^2] = 2/n$ and $|X_t^{ij}| \leq 1$. We use a version of Bernstein's inequality to bound the probability of having large envy between any pair of agents i and j.

Lemma 5.6 ([26])**.** *Let $X_1, \ldots, X_T$ be independent variables with $\mathbb{E}[X_t] = 0$*

and $|X_t| \leq M$ *almost surely for all* $t \in [T]$*. Then, for all* $\lambda > 0$*,*

$$\Pr\left[\sum_{t=1}^{T} X_t > \lambda\right] \leq \exp\left(-\frac{\frac{1}{2}\lambda^2}{\sum_{t=1}^{T}\mathbb{E}\left[X_t^2\right] + \frac{1}{3}M\lambda}\right).$$

When applying this result to $\textsc{Envy}_T^{ij}$ (which equals $\sum_{t=1}^{T} X_t^{ij}$ when envy exists), it follows that

$$\begin{aligned}\Pr\left[\textsc{Envy}_T^{ij} \geq \lambda\right] &= \Pr\left[\sum_{t=1}^{T} X_t^{ij} \geq \lambda\right] \\ &\leq \exp\left(-\frac{\frac{1}{2}\lambda^2}{\frac{2T}{n} + \frac{1}{3}\lambda}\right) = \exp\left(-\frac{3n\lambda^2}{12T + 2\lambda n}\right).\end{aligned}$$

Let $\lambda = 10\sqrt{T \log T / n}$. Taking a union bound gives

$$\begin{aligned}\Pr\left[\textsc{Envy}_T \geq \lambda\right] &= \Pr\left[\exists i, j \in [n] \text{ such that } \textsc{Envy}_T^{ij} \geq \lambda\right] \\ &\leq n^2 \exp\left(-\frac{300 T \log T}{12T + 20\sqrt{nT \log T}}\right) \;\leq\; \frac{1}{T},\end{aligned}$$

where the last inequality uses the assumption that $T \geq n \log T$. Since the maximum possible envy is T, the desired bound on expected envy directly follows, completing the proof of Theorem 5.3.

□

5.3.2 Derandomization with Pessimistic Estimators

The problem of finding an allocation algorithm against an adaptive adversary induces an extensive-form game of complete information between the algorithm and the adversary. In such games, randomization does not provide any benefit to either agent, as the backward induction solution is optimal [17]. This implies that there exists a deterministic algorithm with the same envy guarantee as the random allocation algorithm, i.e., $\textsc{Envy}_T \in \tilde{O}(\sqrt{T/n})$. However, it is *a priori* unclear whether this can be achieved in polynomial time. In fact, Examples 5.1 and 5.2 showed that even though a simple randomized algorithm is optimal and there exists a deterministic algorithm with the same guarantee, natural and interpretable deterministic algorithms may not come with any useful performance guarantees.

Nevertheless, we are able to match the randomized bound of Theorem 5.3 by derandomizing the random allocation algorithm with the method of *pessimistic estimators* Raghavan [1988]. The outcome is an intuitively pleasing, polynomial-time, deterministic algorithm that at each step minimizes a potential function, which is essentially a penalty function exponential in each of the pairwise envy expressions.

Theorem 5.7. *Suppose that $T \geq n \log n$. Then there exists a polynomial-time, deterministic algorithm that achieves* $\textsc{Envy}_T \in O(\sqrt{T \log n/n})$.

The rest of this section is devoted to the proof of Theorem 5.7.

The Algorithm

We will define a potential function $\phi(t)$ that depends on n, T, the values of the first t items, and as their allocations. When item t arrives, we allocate it to the agent for which the value of $\phi(t)$ is minimized. Call this algorithm $\mathcal{A}^*$. Since our algorithm is deterministic, an adversary that wants to maximize $\textsc{Envy}_T$ does not gain from being adaptive. It therefore suffices to analyze our algorithm for an arbitrary choice of item values.

Theorem 5.7 follows from choosing $\phi(t)$ in a way that satisfies three particular properties, stated in the following three lemmas. Given $t \in [T]$, let $\mathcal{A}^t$ be the algorithm that, for all $\ell \in [t]$, allocates the item ℓ to an agent for which $\phi(\ell)$ is minimized, and the remaining items $t+1, \ldots, T$ uniformly at random. Let $\textsc{Envy}_T^{ij}(\mathcal{A}^t)$ be the envy of agent i for agent j at the end of the execution of $\mathcal{A}^t$.

Lemma 5.8. $\phi(t) \geq \sum_{i,j \in [n]} \Pr\left[\textsc{Envy}_T^{ij}(\mathcal{A}^t) > 10\sqrt{T \log n/n}\right]$.

Lemma 5.9. *For all $t \in [T-1]$, $\phi(t+1) \leq \phi(t)$.*

Lemma 5.10. *For $T \geq n \log n$, $\phi(0) < 1$.*

Proof of Theorem 5.7. Notice that $\mathcal{A}^T$ is exactly the same as the algorithm $\mathcal{A}^*$. Lemmas 5.9 and 5.10 imply that $\phi(T) < 1$. Combining with Lemma 5.8, we get that for any choice of item values, and therefore for the optimal adversary strategy,

$$\Pr\left[\exists i, j \in [n] : \textsc{Envy}_T^{ij}(\mathcal{A}^T) > 10\sqrt{\frac{T \log n}{n}}\right]$$

$$\leq \sum_{i,j\in[n]} \Pr\left[\textsc{Envy}_T^{ij}(\mathcal{A}^T) > 10\sqrt{\frac{T\log n}{n}}\right] \leq \phi(T) < 1.$$

Since $\mathcal{A}^T$ is deterministic — all items have been allocated after time T — the inequality above implies that, for the allocation of items by $\mathcal{A}^T$, there is no $i,j \in [n]$ such that $\textsc{Envy}_T^{ij} > 10\sqrt{T\log n/n}$, and we conclude that

$$\textsc{Envy}_T = \max_{i,j\in[n]} \textsc{Envy}_T^{ij} \leq 10\sqrt{\frac{T\log n}{n}} \in O\left(\sqrt{\frac{T\log n}{n}}\right).$$

□

Setup

We now define $\phi(t)$ and prove that it satisfies the desired properties from Lemmas 5.8, 5.9, and 5.10. For $i,j,k \in [n]$ and $t \in [T]$, let y_{tk}^{ij} be a helper variable for the effect on envy between agents i and j when item t goes to agent k, i.e.,

$$y_{tk}^{ij} \triangleq \begin{cases} -1, & \text{if } k = i, \\ 0, & \text{if } k \neq i, j, \\ 1, & \text{if } k = j, \end{cases}$$

and let y_t^{ij} be the same but with the dependence on k implicit (as k is exactly determined given an allocation). Denote with $f_{ij}(t) \triangleq \sum_{\ell=1}^{t} y_\ell^{ij} v_{i\ell}$ the net value agent i has for agent j's allocation with respect to her own at time t. Notice that $\textsc{Envy}_t^{ij} = \max\{f_{ij}(t), 0\}$. Let $C \triangleq (1 + (e^s + e^{-s} - 2)/n)$, where s is a damping parameter that depends only on T and n. We use $s = \sqrt{2\log\left(1 + \frac{n\log n}{T}\right)}$, and let $\lambda \triangleq 10\sqrt{T\log n/n}$ be the target maximum envy that the algorithm allows. Define the potential function at time t as $\phi(t) \triangleq \sum_{i,j\in[n]:i\neq j} \phi_{ij}(t)$, where for $i,j \in [n]$, $\phi_{ij}(t) = C^{T-t} \cdot \exp\left(s\left(f_{ij}(t) - \lambda\right)\right)$.

The Proofs

We are now in a position to prove Lemmas 5.8, 5.9 and 5.10. Lemma 5.8 follows after exponentiating and applying Markov's inequality in the style of classical proofs of concentration inequalities.

Proof of Lemma 5.8. For all $\ell \in [t]$, item ℓ has been allocated in order to minimize $\phi(\ell)$. It suffices to show that at any time $t \leq T$, for any pair of agents i, j, with $\lambda = 10\sqrt{T \log n / n}$,

$$\Pr\left[f_{ij}(t) + \sum_{\ell=t+1}^{T} X_\ell^{ij} v_{i\ell} > \lambda\right] \leq \phi_{ij}(t), \tag{5.1}$$

where X_ℓ^{ij} is a random variable that takes values -1 and 1 with probability $1/n$ each, and it takes value 0 with probability $1 - 2/n$. Notice that $\textsc{Envy}_T^{ij}(\mathcal{A}^t) = \max\{f_{ij}(t) + \sum_{\ell=t+1}^{T} X_\ell^{ij} v_{i\ell}, 0\}$; summing up over all pairs i, j proves the lemma. Equation (5.1) follows from

$$\begin{aligned}
\Pr\left[f_{ij}(t) + \sum_{\ell=t+1}^{T} X_\ell^{ij} v_{i\ell} > \lambda\right] &= \Pr\left[e^{s\left(f_{ij}(t) + \sum_{\ell=t+1}^{T} X_\ell^{ij} v_{i\ell}\right)} > e^{s\lambda}\right] \\
&\leq e^{-s\lambda} \cdot \mathbb{E}\left[e^{s\left(f_{ij}(t) + \sum_{\ell=t+1}^{T} X_\ell^{ij} v_{i\ell}\right)}\right] && \text{(Markov's ineq.)} \\
&= e^{s(f_{ij}(t) - \lambda)} \cdot \mathbb{E}\left[\prod_{\ell=t+1}^{T} e^{s X_\ell^{ij} v_{i\ell}}\right] \\
&= e^{s(f_{ij}(t) - \lambda)} \prod_{\ell=t+1}^{T} \mathbb{E}\left[e^{s X_\ell^{ij} v_{i\ell}}\right] && \text{(independence)} \\
&= e^{s(f_{ij}(t) - \lambda)} \prod_{\ell=t+1}^{T} \left(1 - \frac{2}{n} + \frac{e^{s v_{i\ell}}}{n} + \frac{e^{-s v_{i\ell}}}{n}\right) \\
&\leq e^{s(f_{ij}(t) - \lambda)} \prod_{\ell=t+1}^{T} \left(1 - \frac{2}{n} + \frac{e^{s}}{n} + \frac{e^{-s}}{n}\right) \\
&= \phi_{ij}(t).
\end{aligned}$$

The second inequality follows from the fact that $e^x + e^{-x}$ is nondecreasing for $x \geq 0$. □

For Lemma 5.9, it suffices to compute the change in potential due to a uniformly random allocation, and show that such an allocation does not increase the potential.

Proof of Lemma 5.9. Denote with $\phi_k(t+1)$ the potential function after giving item $t+1$ to agent k. We show that $(1/n)\sum_{k \in [n]} \phi_k(t+1) \leq \phi(t)$, which implies the desired result, as by definition of $\mathcal{A}^*$ we have $\phi(t+1) = \min_{k\in[n]} \phi_k(t+1)$. Recall that, for distinct $i, j, k \in [n]$, y_{tk}^{ij} takes values -1, 1, and 0 depending on whether item t was allocated to agent i, j, or k. Thus, $f_{ij}(t+1) = f_{ij}(t) + y_{t+1,k}^{ij} v_{i,t+1}$.

$$\begin{aligned}
\frac{1}{n}\sum_{k\in[n]} \phi_k(t+1) &= \frac{1}{n}\sum_{k\in[n]} \sum_{i,j\in[n]:i\neq j} e^{s\left(f_{ij}(t)+y_{t+1,k}^{ij} v_{i,t+1}-\lambda\right)} C^{T-(t+1)} \\
&= \frac{1}{n} C^{T-(t+1)} \sum_{i,j\in[n]:i\neq j} e^{s(f_{ij}(t)-\lambda)} \sum_{k\in[n]} e^{s y_{t+1,k}^{ij} v_{i,t+1}} \\
&= \frac{1}{n} C^{T-(t+1)} \sum_{i,j\in[n]:i\neq j} e^{s(f_{ij}(t)-\lambda)} \left(e^{s\cdot(1)\cdot v_{i,t+1}} + e^{s\cdot(-1)\cdot v_{i,t+1}} + \sum_{k\in[n]\setminus\{i,j\}} 1 \right) \\
&= C^{T-(t+1)} \sum_{i,j\in[n]:i\neq j} e^{s(f_{ij}(t)-\lambda)} \cdot \frac{1}{n} \cdot \left(e^{s v_{i,t+1}} + e^{-s v_{i,t+1}} + n - 2 \right) \\
&\leq C^{T-(t+1)} \sum_{i,j\in[n]:i\neq j} e^{s(f_{ij}(t)-\lambda)} \cdot C \;=\; \phi(t).
\end{aligned}$$

□

Finally, we are in a position to prove Lemma 5.10.

Proof of Lemma 5.10. We can bound $\phi(0)$ by expanding it's definition and using the fact that $f_{ij}(0) = 0$.

$$\phi(0) = \sum_{i,j\in[n]:i\neq j} \phi_{ij}(0) = \sum_{i,j\in[n]:i\neq j} C^T e^{s f_{ij}(0) - s\lambda} < n^2 C^T e^{-s\lambda} = e^{-s\lambda + 2\log n + T\log C}.$$

We want $\phi(0) < 1$ or, equivalently, $s\lambda - 2\log n - T\log C > 0$. Using $1 + x \leq e^x$ implies that

$$C = 1 + \frac{e^s + e^{-s} - 2}{n} \leq e^{(e^s + e^{-s} - 2)/n} = e^{2(\cosh(s)-1)/n}.$$

Furthermore, $\cosh(x) \leq \exp(x^2/2)$, so that $C \leq \exp(2(\exp(s^2/2) - 1)/n)$. Therefore,

$$\begin{aligned}
&s\lambda - 2\log n - T\log C \\
&\qquad\qquad \geq s\lambda - 2\log n - \frac{2T}{n}\left(e^{s^2/2} - 1\right)
\end{aligned}$$

$$= 10\sqrt{2\log\left(1+\frac{n\log n}{T}\right)\frac{T\log n}{n}} - 2\log n - \frac{2T}{n}\cdot\frac{n\log n}{T}$$
$$= \left(\frac{5}{\sqrt{2}}\sqrt{\frac{T}{n\log n}\log\left(1+\frac{n\log n}{T}\right)} - 1\right)4\log n.$$

We factored out $4\log n$ for convenience; it remains to show the parenthetical expression is positive. The function $\sqrt{x\log(1+1/x)}$ is increasing for all $x \geq 0$. Set $x = T/(n\log n)$, and note that the assumption $T \geq n\log n$ implies $x \geq 1$. Observing that $5\sqrt{\log(2)/2} > 1$ completes the proof. $\square$

5.3.3 Lower Bound

In this section, we show an adversary can guarantee $\textsc{Envy}_T \in \Omega((T/n)^{r/2})$ for any $r < 1$. It follows that the deterministic algorithm we presented in Section 5.3.1 is optimal (up to a logarithmic factor). We first prove the bound for $n = 2$, followed by the case of an arbitrary number of agents.

Lower Bound for Two Agents

Lemma 5.11. *For $n = 2$ and any $r < 1$, there exists an adversary strategy for setting item values such that any algorithm must have* $\textsc{Envy}_T \in \Omega(T^{r/2})$*.*

Proof. Proof. Label the agents L and R, and let $\{v_0 \triangleq 1, v_1, v_2, \ldots\}$ be a decreasing sequence of values that we specify later, satisfying $v_d - v_{d+1} < v_{d'} - v_{d'+1}$ for all $d' < d$. The adversary keeps track of the *state* of the game, and the current state defines its strategy for choosing the agents' valuations. The adversary strategy that implies the lower bound is illustrated in Figure 5.1. Start in state 0, which we will also refer to as L_0 and R_0, for which the adversary sets the value of the arriving item as $(1, 1)$. To the left of state 0 are states labeled $L_1, L_2, \ldots$; in state L_d, the item that arrives has value $(1, v_d)$. To the right of state 0 are states labeled $R_1, R_2, \ldots$; in state R_d, an item will arrive with value $(v_d, 1)$. Whenever the algorithm allocates an item to agent L (resp. R), which we will refer to as making an L (resp. R) step, the adversary moves one state to the left (resp. right) to determine the value of the next item.

We construct the optimal allocation algorithm against this adversary, and show that for this algorithm the envy at some time step $t \in [T]$ will be at

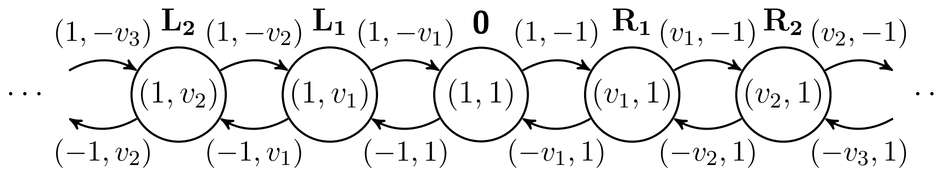


Figure 5.1: Adversary strategy for two-agent lower bound. In state L_d, an item valued $(1, v_d)$ arrives, while in state R_d, an item valued $(v_d, 1)$ arrives. The arrows indicate whether agent L or agent R is given the item in each state. The arrows are labeled by the amount envy changes after that item is allocated.

least $\Omega(T^{r/2})$ for the given $r < 1$. This immediately implies Lemma 5.11: if the envy is sufficiently large at some time step t the adversary can guarantee the same envy at time T by making all future items valued at zero by both agents.

The intuition for the adversary strategy we have defined is that it forces the algorithm to avoid entering state L_d or R_d for high d, as otherwise the envy of some agent will grow to $v_0 + v_1 + \cdots + v_d$, which will be large by our choice of $\{v_d\}$. At the same time, if an L step is taken at state L_d, followed by a later return to state L_d, the envy of R increases by at least $v_d - v_{d+1}$; we choose $\{v_d\}$ so that this increase in envy is large enough to ensure that any algorithm which spends too many time steps close to state 0 incurs a large cost.

By the pigeonhole principle, either the states to the left or to the right of state 0 are visited for at least half the time. For the rest of this section, we assume, without loss of generality, that our optimal algorithm spends time $T' \triangleq \lceil T/2 \rceil$ in the 'left' states $(L_0, L_1, \ldots)$, and that T' is an even number. We prove that the envy of agent R grows large at some time step t. We ignore any time the algorithm spends in the states R_d, $d \geq 1$. To see why this is without loss of generality, consider first a cycle spent in the right states that starts at R_0 with an item allocated to R and eventually returns to R_0. In such a cycle, an equal number of items are allocated to both agents. All of these items have value 1 to agent R, yielding a net effect of 0 on agent R's envy. (We ignore agent L completely, as our analysis is of the envy of agent R.) The other case is when the algorithm starts at R_0 but does not return to R_0. This scenario can only occur once, which means that the algorithm has already taken T' steps on the left side; the allocation of these items does not affect our proof.

Let K be an integer such that $K \leq \sqrt{T'/2}$, which we will show is without loss of generality. Denote by $\mathsf{OPT}(K)$ the set of envy-minimizing allocation algorithms that spend the T' steps in states $L_0, \ldots, L_K$ (and reach L_K). Note that the algorithm aims to minimize the maximum envy at any point in its execution. Let $\mathcal{A}^*(K)$ be the following algorithm, starting at L_0: Allocate the first K items to agent L, thus arriving at state L_K. For the next $T' - 2K$ items, alternate between allocating to agents R and L, thereby alternating between states L_{K-1} and L_K. Allocate the remaining K items to agent R. We show $\mathcal{A}^*(K)$ belongs to $\mathsf{OPT}(K)$.

Lemma 5.12. $\mathcal{A}^*(K) \in \mathsf{OPT}(K)$.

Proof of Lemma 5.12. An algorithm that starts at state 0 and spends T' steps in the left states can be described as a sequence of choices $s_t \in \{L, R\}$ for $t \in [T']$ such that $s_1 = L$, and at every $t \in [T']$, agent L has received at least as many of the first t items as agent R (to avoid entering the right states). We refer to the state at time t as the state *after* the algorithm choice s_t.

Consider any $\mathcal{A}(K) \in \mathsf{OPT}(K)$. We show that the corresponding sequence of allocations satisfy: (1) at time T' the state is L_0, so agent L receives the same number of items as agent R; and (2) there is exactly one R move at states $L_1, \ldots, L_{K-1}$. This proves the lemma, since $\mathcal{A}^*(K)$ is the only algorithm that satisfies these two conditions. We utilize the fact that the envy of an allocation sequence can be calculated from the number of L and R moves in every state: at state L_d, an L move increases the envy of agent R by v_d while an R move decreases it by v_d.

We start with the first property: suppose that the state at time T' is not 0. Let t be the last index such that $s_t = L$. Allocating $s_t = R$ instead (and $s_\ell = R$ for the remaining steps $\ell > t$) reduces the envy of agent R without entering state R_1, a contradiction.

For the second property, it suffices to show that if $s_t = L$ and $s_{t+1} = R$, then it must be that at step t the state is L_{K-1} (and therefore at step $t+1$ the state is L_K). Assume this is not the case, and we have such a t where the algorithm is in state $L_{\widehat{K}-1}$, $\widehat{K} < K$. Let ℓ be a step in which the algorithm is in state L_{K-1}, which exists by the definition of $\mathcal{A}(K)$. Assume that $\ell > t+1$ (an analogous argument can be applied to the case that $\ell < t$). We divide T' into three phases: (1) the first $t-1$ items, (2) the next $\ell-(t+1)$ items, and (3) the last $T'-\ell+2$ items and consider $s' \triangleq s_1, \ldots, s_{t-1}, s_{t+2}, \ldots, s_\ell, s_t, s_{t+1}, s_{\ell+1}, \ldots, s_{T'}$. Notice that s' is s, except the alternating allocations L then R are now made at state L_{K-1} instead of at $L_{\widehat{K}-1}$. By construction, sequence s' never goes past state L_K. We now prove that, using s', the envy decreases with respect to s at each time step after $t-1$, contradicting the assumption $\mathcal{A}(K) \in \mathsf{OPT}(K)$.

In phase (1), the envy is unchanged. For phase (2), when using $\mathcal{A}(K)$, the pair of moves s_t and s_{t+1} increases envy by $v_{\widehat{K}} - v_{\widehat{K}-1}$. Hence, in comparison, s' has that much less envy during each time step of phase (2). At the start of phase (3) in s', the alternating allocations are performed at state L_{K-1}, increasing envy (in s') by $v_{K-1} - v_K < v_{\widehat{K}} - v_{\widehat{K}+1}$. At all remaining steps in (3), the envy is smaller in s' (compared to s) by $(v_{\widehat{K}} - v_{\widehat{K}+1}) - (v_{K-1} - v_K)$. This completes the proof that $\mathcal{A}(K)$ must satisfy both properties; the lemma follows. $\square$

We analyze the envy of $\mathcal{A}^*(K)$ as a function of K before optimizing K. Agent R's maximum envy is realized at step $T'-K$, right before the sequence of R moves. $\textsc{Envy}_{T'-K}$ has two terms: the envy accumulated to reach state L_K, and the envy from alternating R and L moves between states L_K and L_{K-1}, so

$$\textsc{Envy}_{T'-K} = \sum_{d=0}^{K-1} v_d + \frac{T'-2K}{2} \cdot (v_{K-1} - v_K).$$

Given $r < 1$, define $v_d \triangleq (d+1)^r - d^r$. Notice that $\sum_{d=0}^{K-1} v_d = K^r$. This validates the initial assumption that $K \le \sqrt{T'/2}$, as otherwise $\sum_{d=0}^{K-1} v_d \ge (T'/2)^{r/2} \in \Omega(T^{r/2})$. We require the following lemma (which will be proved at the end of this section) to continue.

Lemma 5.13. $v_{K-1} - v_K \ge r(1-r)K^{r-2}$.

Applying Lemma 5.13 and distributing terms yields

$$\textsc{Envy}_{T'-K} \ge K^r - r(1-r)K^{r-1} + \frac{T'}{2} r(1-r)K^{r-2} \ge \frac{1}{2}\left(K^r + T'r(1-r)K^{r-2}\right)$$

where the second inequality uses the fact that $r(1-r) \le 1/4 < 1/2$ and assumes $K > 1$ (otherwise the envy would be linear in T'). To optimize K, noting that the second derivative of the above bound is positive for $K \leq \sqrt{T'/2}$, we find the critical point where the derivative is zero

$$\frac{\partial}{\partial K}\left(K^r + T'r(1-r)K^{r-2}\right) = rK^{r-1} - T'r(1-r)(2-r)K^{r-3} = 0$$
$$\implies K = \sqrt{T'(1-r)(2-r)}.$$

Define $C_1 \triangleq \sqrt{(1-r)(2-r)}$ and substitute into the bound on $\textsc{Envy}_{T'-K}$ to complete the proof

$$\textsc{Envy}_{T'-K} \geq \frac{1}{2}\left(C_1^r (T')^{r/2} + T'r(1-r)C_1^{r-2}(T')^{r/2-1}\right) \in \Omega(T^{r/2}).$$

□

We conclude this subsection with a proof of Lemma 5.13.

Proof of Lemma 5.13. Observe that $v_{K-1} - v_K = K^r - (K-1)^r - (K+1)^r + K^r = 2K^r - (K-1)^r - (K+1)^r$. Using Newton's generalized binomial theorem, with $(r)_k \triangleq r(r-1)\cdots(r-k+1)$, we can expand $(K+1)^r$ and $(K-1)^r$ as

$$(K+1)^r = K^r + rK^{r-1} + \frac{(r)_2}{2!}K^{r-2} + \frac{(r)_3}{3!}K^{r-3} + \frac{(r)_4}{4!}K^{r-4} + \cdots, \text{and}$$
$$(K-1)^r = K^r - rK^{r-1} + \frac{(r)_2}{2!}K^{r-2} - \frac{(r)_3}{3!}K^{r-3} + \frac{(r)_4}{4!}K^{r-4} - \cdots.$$

Combining these identities with the fact that $(r)_k$ is negative when $r < 1$ for all even k, it follows that

$$\begin{aligned} v_{K-1} - v_K &= -2\left(\frac{(r)_2}{2!}K^{r-2} + \frac{(r)_4}{4!}K^{r-4} + \frac{(r)_6}{6!}K^{r-6} + \cdots\right) \\ &= r(1-r)K^{r-2} + 2\left(\frac{|(r)_4|}{4!}K^{r-4} + \frac{|(r)_6|}{6!}K^{r-6} + \cdots\right) \\ &\geq r(1-r)K^{r-2}. \end{aligned}$$

□

Lower Bound for Any Number of Agents

Theorem 5.14. *For any $n \geq 2$ and $r < 1$, there exists an adversary strategy for setting item values such that any algorithm must have* $\textsc{Envy}_T \in \Omega((T/n)^{r/2})$.

Proof of Theorem 5.14. We augment the instance of Figure 5.1 in the following way. In addition to the first two agents, L and R, we have $n-2$ other agents. Each of these other agents will not value *any* of the items that arrive; hence, the nonzero values remain the same as before. State transitions work as follows. If the algorithm allocates an item to agent L or agent R, the transitions are the same as when $n = 2$. Otherwise, the adversary will remain in the same state.

Let T_0 be the number of items allocated to either agent L or R. We break the analysis into two cases. First, if $T_0 \in \Omega(T/n)$, then, $\textsc{Envy}_T \in \Omega((T/n)^{r/2})$ by the analysis of Lemma 5.11. Otherwise, $T_0 \in o(T/n)$ and therefore $T - T_0 \in \Theta(T)$, i.e., agents 3 through n receive many items. This implies that there exists an agent $i \in [3, n]$ that is allocated $\Omega(T/n)$ items. Without loss of generality, at least half these items were allocated in the left states, in which agent L values each item at 1, so that agent L has $\Omega(T/n)$ value for the items received by agent i. The value of agent L for her own allocation is at most $O(T_0)$, i.e., $o(T/n)$. Therefore, the envy of agent L for agent i is at least $\Theta(T/n) - o(T/n) \in \Theta(T/n)$. $\square$

5.4 Batch Arrivals under Full Information

In this section, we study the more general setting where items arrive in batches of size m, and the values of all items in a batch are revealed simultaneously. We assume m divides T for convenience.

5.4.1 Upper Bound

The upper bound of Theorem 5.3 when $m = 1$ may be interpreted as the expected distance from the origin of a random walk that remains stationary with probability $1 - 2/n$, and increases or decreases by 1, each with probability $1/n$. The "step size" of 1 is the maximum change in the envy between any pair of agents after the allocation of a single item; the number of non-stationary steps is expected to be $2T/n$. This informs our approach when items arrive in batches: It is easy to find an EF1 allocation for every batch of items (round-robin suffices). Under such an allocation the maximum change in any pairwise envy due to a single batch remains 1; however, the value of an agent's bundle is likely to change with every batch. Since there are T/m batches ("steps" in the random walk), we may expect a bound of the form $\textsc{Envy}_{T,m} \in \tilde{O}(\sqrt{T/m})$. Indeed, our main result for this setting, given in Theorem 5.17, is a deterministic algorithm that achieves this bound.

To realize this intuition, we first need to overcome a technical obstacle. Even though it is easy to find an allocation with small pairwise envy for a given batch, it is not obvious how to find allocations with low pairwise envy such that randomly outputting one of them results in an (*ex ante*) envy-free allocation. In the random walk interpretation, we need to keep the envy between agents i and j stationary in expectation, while at the same time maintaining a small step size. Note that in the one-by-one setting uniform random allocation trivially satisfies this property. When items arrive in batches, we rely on a result from the literature on the division of *divisible* goods.

Lemma 5.15 (Stromquist 122)**.** *Suppose n agents have valuation functions over the interval $[0, 1]$, such that an agent's value for a subinterval is the integral of her value density function. Then there exists an envy-free division of the interval where every agent receives one contiguous interval.*

It will be convenient to think of the n contiguous allocations as created by $n - 1$ cuts on the interval $[0, 1]$. In the context of indivisible goods with additive valuations, this result implies that, if the items are placed on a line (in any order), there exists a fractional envy-free allocation in which no agent receives more than 2 fractional items. Every item corresponds to an interval of size $1/m$, and every agent's valuation in that interval is constant and proportional to her valuation for that item. Given the solution guaranteed to exist by Lemma 5.15, every agent's allocation is between at most two cuts and therefore contains no more than 2 fractional items. Such a near-integral envy-free allocation is useful, since any integral allocation found by randomized rounding is guaranteed to have small envy *ex post*, as the following lemma shows.

Lemma 5.16. *Given m items, there exists an envy-free fractional allocation $A = A_1, \ldots, A_n$, such that every agent receives at most 2 fractional items. Furthermore, if $x_{i\ell} \in [0, 1]$ is the fraction of item ℓ allocated to agent i, then* *randomly giving each item ℓ to each agent i with probability $x_{i\ell}$ results in an integral allocation A' where for all $i, j \in [n]$, $v_i(A'_i) \geq v_i(A'_j) - 4$.*

Proof. Proof of Lemma 5.16. The first part of the statement, that there exists an envy-free allocation A in which each agent receives at most 2 fractional items, follows from the previous discussion. For the second part, notice that the worst-case scenario for an agent i is to not get either of the fractional items allocated to her in A. Furthermore, some other agent j might get both of her fractional items from A. In this scenario, the envy of agent i for agent j is maximized and is at most 4 (since the value for every item is at most 1). □

In Section 5.3, before giving a deterministic algorithm, we first analyzed the performance of the random allocation algorithm. Crucially, we showed the optimal strategy for an adaptive adversary against the random allocation algorithm is in fact nonadaptive. This allowed us to use standard concentration inequalities. Such a characterization is much trickier here. Fortunately, we can bypass this step and directly "derandomize" the algorithm that at each step outputs the (randomly-rounded) allocation of Lemma 5.16, even though we are unable to analyze its performance.

Theorem 5.17. *Suppose that $T \geq m \log n$. Then there exists a deterministic algorithm that achieves* $\textsc{Envy}_{T,m} \in O(\sqrt{T \log n / m})$.

Again, the assumption of $T \geq m \log n$ is very weak, otherwise there are at most $T/m \leq \log n$ batches, and we can use an EF1 algorithm in each to achieve $\textsc{Envy}_{T,m} \leq \log n$. The remainder of this section is devoted to the proof of Theorem 5.17, which has a very similar structure to the proof of Theorem 5.7.

The Algorithm

We define a potential function $\phi(t)$ that depends on n, T, the values of the items in the first t batches, as well as their allocations. When batch $t+1$ arrives, we first find the near-integral envy-free allocation A^{t+1} (of the items in batch $t+1$) guaranteed to exist by Lemma 5.16 (we address computation below). This fractional allocation is then rounded to an integral allocation in a way that $\phi(t+1)$ is minimized. Call this algorithm $\mathcal{A}^*$. Since our algorithm is deterministic, an adversary that wants to maximize $\textsc{Envy}_{T,m}$ does not gain from being adaptive. Therefore, there exists some optimal (for the adversary) choice of values for items 1 through T. We analyze our algorithm for an arbitrary choice of item values.

Similarly to our algorithm from Section 5.3.2, we rely on three properties of ϕ. Given $t \in [T/m]$, let $\mathcal{A}^t$ be the algorithm that rounds A^ℓ (the allocation in batch ℓ) in a way that $\phi(\ell)$ is minimized, for all $\ell \in [1, t]$, and rounds the remaining A^ℓ for $\ell = t+1, \dots, T/m$ randomly. Let $\textsc{Envy}^{ij}_{T,m}(\mathcal{A}^t)$ be the envy of agent i for agent j at the end of the execution of $\mathcal{A}^t$.

Lemma 5.18. $\phi(t) \geq \sum_{i,j \in [n]} \Pr\left[\textsc{Envy}^{ij}_T(\mathcal{A}^t) > 100\sqrt{T \log n/m}\right]$.

Lemma 5.19. *For all* $t \in [T/m - 1]$, $\phi(t+1) \leq \phi(t)$.

Lemma 5.20. *For* $T \geq m \log n$, $\phi(0) < 1$.

Proof of Theorem 5.17. Notice that $\mathcal{A}^{T/m}$ is exactly the same as the algorithm $\mathcal{A}^*$. Lemmas 5.19 and 5.20 imply that $\phi(T) < 1$. Combining this with Lemma 5.18, we get that for any item valuations,

$$\Pr\left[\exists i, j \in [n] : \textsc{Envy}^{ij}_T(\mathcal{A}^*) > 100\sqrt{\frac{T \log n}{m}}\right] \leq \sum_{i,j \in [n]} \Pr\left[\textsc{Envy}^{ij}_T(\mathcal{A}^*) > 100\sqrt{\frac{T \log n}{m}}\right] \leq \phi(T) < 1.$$

Since $\mathcal{A}^*$ is deterministic, the inequality above implies that there is no $i, j \in [n]$ such that $\textsc{Envy}^{ij}_T > 100\sqrt{T \log n/m}$, and we conclude that $\textsc{Envy}_{T,m} \leq 100\sqrt{T \log n/m} \in O(\sqrt{T \log n/m})$. $\square$

Setup

In batch t, define $A^t \triangleq A^t_1, \dots, A^t_n$ as the envy-free fractional solution of Lemma 5.16, in which no agent receives more than 2 fractional items.

Let $\hat{A}^t = \hat{A}^t_1, \dots, \hat{A}^t_n$ be an integral rounding of A^t; $\hat{A}^t$ is the actual allocation used in batch t. Define

$$\Delta^t_{ij}(\hat{A}^t) \triangleq \left(v_i(\hat{A}^t_j) - v_i(\hat{A}^t_i)\right) + \left(v_i(A^t_i) - v_i(A^t_j)\right)$$

and let $f_{ij}(t, \hat{A}^1, \dots, \hat{A}^t) \triangleq \sum_{\ell=1}^t \Delta^t_{ij}(\hat{A}^t)$. To simplify notation, we write $f_{ij}(t)$ when the allocation is clear from context. Notice that $\textsc{Envy}^{ij}_{T,m} \leq f_{ij}(T)$; this is an inequality because Δ^t_{ij} is centered to have zero mean, while

$v_i(\hat{A}_j^t) - v_i(\hat{A}_i^t)$ may have mean less than zero. Also, observe that $\hat{A}^t$ is not random. However, if we were to *randomly* round A^t to an integral allocation $\hat{B}^t$, then the resulting random variable $\Delta_{ij}^t(\hat{B}^t)$ has zero mean and satisfies $|\Delta_{ij}^t(\hat{B}^t)| \leq 4$, by Lemma 5.16.

Let $\lambda \triangleq 100\sqrt{T \log n/m}$ and $s \triangleq \frac{1}{4}\log(1 + \frac{\lambda m}{4T})$. For $i, j \in [n]$, define the potential function at time t for i with respect to j as $\phi_{ij}(t) \triangleq \exp\left(s f_{ij}(t) - s\lambda + \left(\frac{T}{m} - t\right)\left(e^{4s} - 4s - 1\right)\right)$, and define the overall potential function as $\phi(t) \triangleq \sum_{i,j\in[n]:i\neq j} \phi_{ij}(t)$.

The Proofs

We rely on the following property of bounded, centered random variables.

Lemma 5.21. *Let X be a random variable with $\mathbb{E}[X] = 0$ and $|X| \leq 4$. Then for all $v \in [0, 1]$ it holds that $\mathbb{E}\left[e^{sXv}\right] \leq \exp\left(e^{4s} - 4s - 1\right)$.*

Proof of Lemma 5.21. Taking the Taylor expansion of e^x at 0 we have:

$$\mathbb{E}\left[e^{sX}\right] = \mathbb{E}\left[1 + sXv + \sum_{k=2}^{\infty} \frac{s^k (X)^k}{k!}\right] \leq 1 + 0 + \sum_{k=2}^{\infty} \frac{s^k \mathbb{E}\left[(X)^k\right]}{k!}$$
$$\leq 1 + \sum_{k=2}^{\infty} \frac{4^k s^k}{k!} = 1 + \left(e^{4s} - 4s - 1\right) \leq \exp\left(e^{4s} - 4s - 1\right).$$

□

The proof of Lemma 5.18 is very similar to the corresponding proof in Section 5.3.2, it is restated here in the interest of completeness.

Proof of Lemma 5.18. For all $\ell \leq t$, the fractional allocation A^ℓ in the ℓ-th batch was rounded to the allocation $\hat{A}^\ell$ which minimizes $\phi(\ell)$. Recall that, since our algorithm is deterministic, it suffices to analyze an arbitrary, but fixed choice of items. Let $\hat{B}_1^\ell, \dots, \hat{B}_n^\ell$ be the (random) allocation that comes from a randomized rounding of A^ℓ, for all $\ell \in [t + 1, T/m]$.

Let $\delta_{ij}^t = v_i(A_j^t) - v_i(A_i^t)$ for all $i, j \in [n], t \in [T/m]$. Note that all $\delta_{ij}^t \leq 0$ since the allocation is envy free. Define random variables $Y_{ij}^\ell = v_i(\hat{B}_j^\ell) - v_i(\hat{B}_i^\ell) - \delta_{ij}^\ell$. These variables have zero mean and satisfy $|Y_{ij}^\ell| \leq 4$

(Lemma 5.16). It suffices to show that at any time $t \leq T$, for any pair of agents i, j, for $\lambda = 100\sqrt{T \log n/m}$,

$$\Pr\left[\textsc{Envy}_T^{ij}(\mathcal{A}) > \lambda\right] \leq \Pr\left[f_{ij}(t) + \sum_{\ell=t+1}^{T/m} Y_{ij}^{\ell} > \lambda\right] \leq \phi_{ij}(t), \tag{5.2}$$

where the first inequality results from the fact that the Y variables are centered and that $f_{ij}(t) \geq \textsc{Envy}_t^{ij}$. Summing up over all pairs i, j proves the claim. Equation 5.2 follows from

$$\begin{aligned}
\Pr&\left[f_{ij}(t) + \sum_{\ell=t+1}^{T/m} Y_{ij}^{\ell} > \lambda\right] \\
&= \Pr\left[e^{s\left(f_{ij}(t)+\sum_{\ell=t+1}^{T/m} Y_{ij}^{\ell}\right)} > e^{s\lambda}\right] \\
&\leq e^{-s\lambda} \cdot \mathbb{E}\left[e^{sf_{ij}(t)+s\sum_{\ell=t+1}^{T/m} Y_{ij}^{\ell}}\right] && \text{(Markov's ineq.)} \\
&= e^{sf_{ij}(t)}e^{-s\lambda} \cdot \mathbb{E}\left[\prod_{\ell=t+1}^{T/m} e^{sY_{ij}^{\ell}}\right] \\
&= e^{sf_{ij}(t)}e^{-s\lambda} \prod_{\ell=t+1}^{T/m} \mathbb{E}\left[e^{sY_{ij}^{\ell}}\right] && \text{(independence)} \\
&\leq e^{sf_{ij}(t)}e^{-s\lambda} \prod_{\ell=t+1}^{T/m} \exp\left(e^{4s} - 4s - 1\right) && \text{(Lemma 5.21)} \\
&= e^{sf_{ij}(t)}e^{-s\lambda}e^{\left(\frac{T}{m}-t\right)\left(e^{4s}-4s-1\right)} = \phi_{ij}(t).
\end{aligned}$$

$\square$

We are now in a position to prove Lemma 5.19.

Proof of Lemma 5.19. We prove that there exists a rounding $\hat{A}^*$ of the fractional allocation A^{t+1} of batch $t+1$ so that allocating according to $\hat{A}^*$ results in $\phi(t+1) \leq \phi(t)$. Let $x_{i\ell}^{t+1}$ be the fraction of item ℓ in batch $t+1$ allocated to agent i in A^{t+1}. We show that allocating every item ℓ to agent i with probability $x_{i\ell}^{t+1}$ makes the expected value of $\phi(t+1)$ at most $\phi(t)$. We can immediately conclude that there exists an integral allocation for which $\phi(t+1) \le \phi(t)$.

Let $\hat{B}^{t+1}$ be a possible (rounded) integral allocation, with corresponding probability $p(\hat{B}^{t+1})$, and let D be the distribution where allocation $\hat{B}^{t+1}$ appears with probability $p(\hat{B}^{t+1})$. Finally, let $\phi_{\hat{B}^{t+1}}(t+1)$ be the value of the potential function after allocating batch $t+1$ according to $\hat{B}^{t+1}$. Note that $f_{ij}(t+1, \hat{A}^1, \dots, \hat{A}^t, \hat{B}^{t+1}) = f_{ij}(t) + \Delta_{ij}^{t+1}(\hat{B}^{t+1})$.

$$\begin{aligned}
&\mathbb{E}_{\hat{B}^{t+1}\sim D}\left[\phi_{\hat{B}^{t+1}}(t+1)\right] \\
&= \sum_{\hat{B}^{t+1}} p(\hat{B}^{t+1}) \cdot \left(e^{-s\lambda} e^{(\frac{T}{m}-t-1)(e^{4s}-4s-1)} \sum_{i,j\in[n]:i\neq j} e^{sf_{ij}(t+1,\hat{A}^1,\dots,\hat{A}^t,\hat{B}^{t+1})} \right) \\
&= \sum_{\hat{B}^{t+1}} p(\hat{B}^{t+1}) \cdot \left(e^{-s\lambda} e^{(\frac{T}{m}-t-1)(e^{4s}-4s-1)} \sum_{i,j\in[n]:i\neq j} e^{sf_{ij}(t)+s\Delta_{ij}^{t+1}(\hat{B}^{t+1})} \right) \\
&= e^{-s\lambda} e^{(\frac{T}{m}-t-1)(e^{4s}-4s-1)} \sum_{i,j\in[n]:i\neq j} \left(e^{sf_{ij}(t)} \sum_{\hat{B}^{t+1}} p(\hat{B}^{t+1})\, e^{s\Delta_{ij}^{t+1}(\hat{B}^{t+1})} \right) \\
&\le e^{-s\lambda} e^{(\frac{T}{m}-t-1)(e^{4s}-4s-1)} \sum_{i,j\in[n]:i\neq j} \left(e^{sf_{ij}(t)}\, \mathbb{E}_{\hat{B}^{t+1}\sim D}\left[e^{s\Delta_{ij}^{t+1}(\hat{B}^{t+1})} \right] \right).
\end{aligned}$$

$\Delta_{ij}^{t+1}(\hat{B}^{t+1})$ is a random variable ($\hat{B}^{t+1}$ is random) that satisfies the conditions of Lemma 5.21, so

$$\begin{aligned}
\mathbb{E}_{\hat{B}^{t+1}\sim D}\left[\phi_{\hat{B}^{t+1}}(t+1)\right] &\le e^{-s\lambda} e^{(\frac{T}{m}-t-1)(e^{4s}-4s-1)} \sum_{i,j\in[n]:i\neq j} e^{sf_{ij}(t)} e^{e^{4s}-4s-1} \\
&= e^{-s\lambda} e^{(\frac{T}{m}-t)(e^{4s}-4s-1)} \sum_{i,j\in[n]:i\neq j} e^{sf_{ij}(t)} = \phi(t).
\end{aligned}$$

□

We conclude with the proof of Lemma 5.20 which is very similar to the corresponding proof in Section 5.3.2.

Proof of Lemma 5.20.

$$\phi(0) = \sum_{i,j\in[n]:i\neq j} \phi_{ij}(0)$$

$$
\begin{aligned}
&= \sum_{i,j\in[n]:i\neq j} \exp\left(sf_{ij}(0) - s\lambda + \frac{T}{m}\left(e^{4s} - 4s - 1\right)\right) \\
&< n^2 \exp\left(-s\lambda + \frac{T}{m}\left(e^{4s} - 4s - 1\right)\right) \\
&= n^2 \exp\left(-\frac{T}{m}\left(1 + 4s\left(1 + \frac{\lambda m}{4T}\right) - e^{4s}\right)\right) \\
&= n^2 \exp\left(-\frac{T}{m}\left(1 + \left(1 + \frac{\lambda m}{4T}\right)\log\left(1 + \frac{\lambda m}{4T}\right) - \left(1 + \frac{\lambda m}{4T}\right)\right)\right) \\
&= n^2 \exp\left(-\frac{T}{m}\left((1+x)\log\left(1+x\right) - x\right)\right),
\end{aligned}
$$

where $x \triangleq \frac{\lambda m}{4T}$. The function $h(x) = (1+x)\log(1+x) - x$ satisfies $h(x) \geq x^2/(2 + 2x/3)$. Therefore,

$$
\begin{aligned}
\phi(0) &< n^2 \exp\left(-\frac{T}{m}\left(\frac{(\frac{\lambda m}{4T})^2}{2 + \frac{2(\frac{\lambda m}{4T})}{3}}\right)\right) = n^2 \exp\left(-\frac{3m\lambda^2}{96T + 8\lambda m}\right) \\
&= \exp\left(2\log n - \frac{3m\lambda^2}{96T + 8\lambda m}\right).
\end{aligned}
$$

Substituting in $\lambda = 100\sqrt{T\log n/m}$ gives

$$
\phi(0) \leq \exp\left(2\log n - \frac{30000T\log n}{96T + 800\sqrt{Tm\log n}}\right),
$$

which is strictly less than 1 for $T \geq m\log n$. $\square$

Discussion

Lemma 5.16, just like Lemma 5.15, is existential and leaves unanswered the question of finding the nearly-integral envy-free allocation for every batch. We partially address this, at least from a practical point of view, by formulating a *mixed-integer program (MIP)* to compute such an allocation.

Let $x_{i\ell}$ be the fraction of item ℓ given to agent i. Binary variables $x^0_{i\ell}$ and $x^1_{i\ell}$ will sum to 0 when $x_{i\ell}$ is fractional, and sum to 1 otherwise. Lemma 5.15 implies that the following MIP is feasible:

$$\sum_{\ell=1}^{m} v_{i\ell}(x_{i\ell} - x_{j\ell}) \geq 0, \qquad \forall i, j \in [n] \tag{5.3}$$

$$\sum_{i=1}^{n} x_{i\ell} = 1, \qquad \forall \ell \in [m] \tag{5.4}$$

$$\sum_{\ell=1}^{m} (x^0_{i\ell} + x^1_{i\ell}) \geq m - 2, \qquad \forall i \in [n] \tag{5.5}$$

$$x^0_{i\ell} \leq x_{i\ell} \leq 1 - x^1_{i\ell}, \qquad \forall i \in [n], \ell \in [m] \tag{5.6}$$

$$x_{i\ell} \in [0, 1], \qquad \forall i \in [n], \ell \in [m] \tag{5.7}$$

$$x^0_{i\ell}, x^1_{i\ell} \in \{0, 1\}, \qquad \forall i \in [n], \ell \in [m]. \tag{5.8}$$

Constraint (5.3) ensures that the allocation is envy free, while Constraint (5.4) ensures every item is fully allocated. Constraint (5.6) ensures that $x^0_{i\ell}$ and $x^1_{i\ell}$ sum to 0 when $x_{i\ell}$ is fractional, and sum to 1 otherwise (using the fact that these variables are binary, by Constraint (5.8)). Constraint (5.5) guarantees at most 2 fractional items per agent. These constraints may be coupled with any objective function to find a near-integral fractional solution.

Unfortunately, solving a MIP is unlikely to be computationally efficient in general. Furthermore, known hardness results for related problems [60] suggest that producing an envy-free (or approximately envy-free) and contiguous fractional allocation in our setting might be difficult. This does not rule out a polynomial time algorithm for finding an allocation with the properties of Lemma 5.16, i.e., a fractional envy-free allocation where each agent gets at most a constant number of fractional items. The general problem is left open.

Another step that may seem problematic (from a computational viewpoint) is rounding the fractional allocation in a way that minimizes the potential function. However, since the potential function is convex in the allocation, this can be done efficiently.

5.4.2 Polynomial-Time Special Cases

In general, the allocation algorithm which achieves envy as prescribed by Theorem 5.17 requires solving an integer program. We now explore two special cases where polynomial time algorithms can be leveraged to guarantee $\textsc{Envy}_{T,m} \in O(\sqrt{T \log n/m})$. There also exists a fully polynomial time algorithm with $\textsc{Envy}_{T,m} \in O(n^2 \cdot \sqrt{T \log n/m})$, which gives the desired asymptotic bound when n is a constant.

Identical preference orderings

Random serial dictatorship (RSD), or random priority [30], is a mechanism for dividing indivisible goods where agents are ordered at random, and agents sequentially select their most preferred item until no items remain. Executing RSD independently for every batch performs well in certain settings.

Theorem 5.22. *If $n = 2$, or when all agents have the same preference ordering for every batch, executing RSD for every batch guarantees* $\textsc{Envy}_{T,m} \in O(\sqrt{T \log n/m})$ *in polynomial time.*

Proof. We first focus on the $n = 2$ case. Refer to the two agents as 1 and 2. Let A^{12} be the allocation of an arbitrary batch resulting executing RSD in the order 1, 2 (where agent 1 first selects an item, followed by agent 2, *etc.*); define A^{21} similarly. Let $v_1(X_1), v_1(X_2)$ be the value that agent 1 has for her own and agent 2's goods under allocation X.

The following lemma formalizes the idea that if agent 1's envy towards 2 increases when 2 selects first, then agent 1's envy towards 2 decreases by at least as much when 1 selects first.

Lemma 5.23. $v_1(A_1^{12}) - v_1(A_2^{12}) \geq v_1(A_2^{21}) - v_1(A_1^{21})$.

Proof. Let $\alpha = v_1(A_1^{21}) - v_1(A_2^{21})$. It is known that round-robin, or RSD, is envy free up to one good (EF1) for additive valuations, so $\alpha \geq -1$. When item valuations are known, the procedure takes time polynomial in n, T and m.

For ease of exposition assume there are $m = 2k$ items in the batch. Label the items in the order they are selected when doing round robin with the permutation $2, 1$. In other words, agent 2 selects items $1, 3, \ldots, 2k - 1$ and agent 1 selects $2, 4, \ldots, 2k$ when performing RSD with the permutation $2, 1$.

Let $v_{2,\ell}$ ($v_{1,\ell}$) be the value of agent B (A, respectively) for item ℓ. Then $v_1(A_1^{21}) = \sum_{\ell=1}^{k} v_{1,2\ell}$. Since each agent selects her most preferred item among the remaining items at each step, we observe that

$$v_{2,1} \geq v_{2,\ell}, \quad \ell = 2, \ldots, 2k$$

$$
\begin{aligned}
v_{2,3} &\geq v_{2,\ell}, \quad \ell = 4\ldots, 2k \\
&\vdots \\
v_{2,2k-1} &\geq v_{2,2k},
\end{aligned}
$$

and similarly

$$
\begin{aligned}
v_{1,2} &\geq v_{1,\ell}, \quad \ell = 3, \ldots, 2k \\
&\vdots \\
v_{1,2(k-1)} &\geq v_{1,\ell}, \quad \ell = 2k-1, 2k.
\end{aligned}
$$

Note that we have inequalities for the oddly indexed items for agent 2, and for the even ones for agent 1 since these are the items selected by the respective agents. The following lemma establishes that when agent 1 selects her r-th item under permutation $1, 2$, she is able to choose an item she values at at least $v_{1,2r-1}$ (the value of agent 2's r-th pick under permutation $2, 1$). This will directly imply that $v_1(A_1^{12}) \geq v_1(A_2^{21})$.

Lemma 5.24. *When agent 1 is about to select her r-th item under permutation $1, 2$ (after $2(r-1)$ items have been selected) , either items $\{1, \ldots, 2(r-1)\}$ or $\{1, \ldots, 2(r-1)-1, 2(r-1)+1\}$ have been selected.*

Proof. By induction on r. Base case: $r = 1$ is trivial, consider $r = 2$: For agent 1's first pick she selects item 1 or 2. If she selects 2, then agent 2 selects item 1 (her most preferred item). If agent 1 selects item 1, agent 2 selects either item 2 or 3, since $v_{2,3} \geq v_{2,\ell}$ for all $\ell = 4, \ldots, n$. In either case we obtain the required property.

Assume as induction hypothesis that whenever agent 1 is about to select her r-th item under permutation $1, 2$, either items $\{1, \ldots, 2(r-1)\}$ or $\{1, \ldots, 2(r-1)-1, 2(r-1)+1\}$ have been selected for all r up to and including s.

Suppose that when agent 1 makes her s-th pick, items $\{1, \ldots, 2(s-1)\}$ have been selected. Now agent 1 selects either $2(s-1)+1$ or $2s$, since $v_{1,2s} \geq v_{1,2s+\ell}$ for all integer $\ell > 0$. Assume agent 1 picks $2(s-1)+1$, then agent 2 will select either $2s$ or $2s+1$, since $v_{2,2s+1} \geq v_{2,2s+1+\ell}$ for $\ell \in \mathbb{Z}_+$. In either case, the induction hypothesis holds. Assume instead that agent 1 selects $2s$, leaving agent 2 to pick $2(s-1)+1$ since $v_{2,2(s-1)+1} \geq v_{2,2(s-1)+1+\ell}$ for $\ell \in \mathbb{Z}_+$. The induction hypothesis holds.

Assume now that when agent 1 is about to make her s-th pick, items $\{1, 2, \ldots, 2(s-1)-1, 2(s-1)+1\}$ have been selected. Agent 1 selects $2s$, since $v_{1,2s} \geq v_{1,2s+\ell}$ for $\ell \in \mathbb{Z}_+$. Agent 2 selects next, after items $1, \ldots, 2(s-1)+1$ have been selected. This case has been analyzed already, and we conclude that the induction hypothesis holds for agent 1's $(s+1)$-th selection. $\square$

Let v_1 be the total value that agent A has for all items in the batch, so

$$v_1(A_1^{12}) + v_1(A_2^{12}) = v_1 = v_1(A_1^{21}) + v_1(A_2^{21}). \tag{5.9}$$

It follows from the Lemma 5.24 that the item that A selects with his r-th pick under the permutation AB has value at least $v_{A,2(r-1)+1}$. This, together with (5.9), implies that $v_1(A_1^{12}) \geq v_1(A_2^{21})$ and $v_1(A_2^{12}) \leq v_1(A_1^{21})$. We conclude that

$$v_1(A_1^{12}) - v_1(A_2^{12}) \geq v_1(A_2^{21}) - v_1(A_1^{21}) = -\alpha$$

$\square$

Lemma 5.23 states that whatever increase in envy an agent may experience when placed last in the permutation is more than compensated for when that agent is placed first. The required bound for $n = 2$ may be obtained by derandomizing the selection of a permutation in every batch in a similar way as in Theorems 5.7 and 5.17.

Suppose now that all agents have the same preference ordering for the items in every batch. In this case the items allocated to a specific agent depends only on that agent's position in the ordering. Any decrease in the envy of agent i towards agent j is when placed in positions k and ℓ, respectively, is offset by an identical increase in envy when agent i is in position ℓ and agent j in position k. A result analogous to Lemma 5.23 may be established under these assumptions. Since round-robin allocations are EF1, we can derandomize as before and conclude that that $\textsc{Envy}_{T,m} \in O(\sqrt{T \log n/m})$. $\square$

Unfortunately, it is possible to construct an example with three agents where this algorithm leads to a linear growth in envy.

We remark that it is also possible to obtain the preceding result by writing bi-hierarchical envy-freeness constraints on a matrix X in the spirit of [38], where entry X_{ij} is the probability that agent i receives item j. These constraints end up being essentially totally unimodular when $n = 2$ or agents have identical preference orders for every batch, but do not have enough structure to enable positive results in general.

Constant Number of Agents

It is possible to find an envy-free fractional allocation with no more than n^2 fractional variables in polynomial time. Randomly rounding this allocation will yield an integral allocation with envy $O(n^2)$ per batch, leading to the following result after derandomization.

Theorem 5.25. *There exists a polynomial-time, deterministic algorithm that guarantees* $\textsc{Envy}_{T,m} \in O(n^2 \cdot \sqrt{T \log n/m})$. *When n is constant this reduces to* $\textsc{Envy}_{T,m} \in O(\sqrt{T/m})$.

We rely on the following lemma, which plays the role of a weaker version of Lemma 5.15, and shows how to find an 'almost integral' envy-free allocation in polynomial time. To the best of our knowledge this result first appeared in [130], and the proof we present here is due to Noga Alon.

Lemma 5.26. *Given m items and n agents, where $v_{i,j}$ is the value that agent i has for item j, there exists an envy-free fractional allocation with no more than n^2 fractional variables.*

Proof. Label a batch of m items $1, \ldots, m$ arbitrarily. Let $V_{it} = \sum_{\ell=1}^{t} v_{i\ell}$ for $i \in [n]$ and $\ell \in [m]$. After ℓ of the k items have been processed, the algorithm has maintains a fractional solution $x_{i\ell}$ for $i \in [n], \ell \in [m]$ so that:

1. For all $i, j \in [n]$, $\sum_{\ell=1}^{m} v_{i\ell} x_{i\ell} = V_{im}/n$;
2. For all $\ell \in [m]$, $\sum_{i=1}^{n} x_{i\ell} = 1$ for all $\ell \in [m]$, and $0 \leq x_{i\ell} \leq 1$.
3. The number of non-integral variables is at most $2n^2$.

If variables satisfying these properties are retained until all items have been processed, they represent a fractional allocation with no more than $O(n^2)$ fractional variables. This fractional allocation is not only envy free but also 'balanced', meaning that every agent values all n bundles identically.

The fact that there are no more than $2n^2$ fractional variables imply that at most n^2 items are allocated fractionally.

The algorithm starts with $x_{i,1} = 1/n$ for all $i \in [n]$. Assume that variables $x_{i\ell}$ have already been assigned values that satisfy the above properties for some $\ell < m$. We show how to update them and allocate item $\ell + 1$ to satisfy the properties, without changing any integral variable.

When processing item $\ell + 1$, set $x_{i,\ell+1} = 1/n$ for all $i \in [n]$. It is now possible that the third property is violated and there are more than $2n^2$ fractional variables.

If the number of fractional variables $r > 2n^2$, consider the system of linear equations consisting of the n^2 equations in (1), and the equations in (2) for those items ℓ which are fractionally assigned in the current solution. Note that there are no more than $r/2$ equations of this form. Fix all integral variables. This leaves a system of equations with r free variables and $n^2 + r/2 < r$ equations. Since there are more variables than equations, there is a line of solutions $x'_{i\ell} = x_{i\ell} + \lambda c_{i\ell}$, where (i, ℓ) runs over the indices of the free variables, some $c_{i\ell} \neq 0$ and λ is a scalar. $\lambda = 0$ is a valid solution, and we can increase λ until the first fractional variable becomes either 0 or 1. Picking this λ decreases the number of fractional variables by 1. The process can now be repeated with the new system of equations until no more than $2n^2$ variables are fractional, which concludes the processing of item $\ell + 1$. □

Using the same techniques as in the proofs of Theorems 5.7 and 5.17, we can derandomize the algorithm which randomly rounds every near-integral envy-free allocation found by Lemma 5.26 to find a deterministic algorithm with $\textsc{Envy}_{T,m} \in O(n^2 \cdot \sqrt{T \log n/m})$.

5.4.3 Lower Bound

Our last result for the batch setting is a lower bound, which is asymptotically tight in T/m, but does leave a gap in terms of the dependence on the number of agents.

Theorem 5.27. *For any $n \geq 2$ and $r < 1$, there exists an adversary strategy for setting item values such that any algorithm must have $\textsc{Envy}_T \in \Omega((\frac{T}{mn})^{r/2})$.*

Proof. Proof. The theorem follows almost directly from Theorem 5.14. Indeed, assume that in each batch there are $m - 1$ items that are worthless to all agents. In this case the batch setting reduces to the one-by-one setting, and we obtain the lower bound given by Theorem 5.14, with a total number of items equal to the number of batches, i.e., $T' = T/m$. □

5.5 Single Arrivals under Partial Information

In the full information setting, the allocation algorithm knows the value of each agent for every item. In the partial information setting, these values are only revealed after allocation. Here, as in Section 5.3, we assume items arrive one at a time.

Under partial information, the upper bound of Theorem 5.3 for the random allocation algorithm carries over directly. However, in contrast to the full information setting, where there is no distinction between deterministic and randomized algorithms, allowing randomization in this setting gives the allocation algorithm significant power. In particular, under partial information, we find that deterministic allocation mechanisms are unable to guarantee vanishing envy.

5.5.1 Randomized Algorithms

The lower bound in Theorem 5.14 shows that an adversary can ensure essentially $\textsc{Envy}(T) \in \Omega(\sqrt{T/n})$ against a randomized algorithm that knows the item valuation before making the allocation. A weaker allocation algorithm cannot improve over this. Allocating incoming items uniformly at random does not make use of item valuations and was shown to have $\textsc{Envy}(T) \in O(\sqrt{T \log T/n})$ in Theorem 5.3. The next result immediately follows from Theorems 5.3 and 5.14.

Corollary 5.28. *Uniform random allocation guarantees that* $\mathbb{E}\left[\textsc{Envy}_T\right] \in O(\sqrt{T \log T/n})$ *even in the partial information setting. On the other hand, there exists an adversary strategy for setting item values such that any algorithm must have* $\textsc{Envy}_T \in \Omega((T/n)^{r/2})$ *for any* $r < 1$.

5.5.2 Deterministic Algorithms

Consider the extensive-form game tree in the partial information case. Nodes on odd layers $2\ell - 1$ belong to the adversary, who chooses an outgoing edge corresponding to the values $v_{1\ell}, \ldots, v_{n\ell}$ for item ℓ. Nodes on the even layer 2ℓ belong to the algorithm, which selects an outgoing edge corresponding to assigning item ℓ to one of the agents.

Let u be a node on the odd layer $2\ell - 1$ and $c(u)$ its children. In the partial information setting, $c(u)$ are all in the same information set. In other words, the algorithm is unable to distinguish between being at any of the nodes $v \in c(u)$, since the value of item ℓ is hidden, and therefore selects the same allocation $a(u)$ at every node $v \in c(u)$. It is easy for the adversary to exploit this to create highly imbalanced allocations.

Theorem 5.29. *An adaptive adversary can ensure* $\textsc{Envy}_T \in \Omega(T/n)$ *against a deterministic algorithm under partial information, while an allocation algorithm can guarantee* $\textsc{Envy}_T \in O(T/n)$.

Proof. We first show the lower bound. Let u be a node of the game tree belonging to the adversary and denote with $a(u)$ the allocation of the algorithm in information set $c(u)$.

An allocation algorithm is defined by selecting an agent to allocate to in each of its information sets. Once the algorithm's strategy is fixed, it is not hard for the adversary to adapt and ensure high envy: At each node u on level $2\ell - 1$ belonging to the adversary, select the edge corresponding to the assignment of values $v_{a(u),\ell} = 0$ and $v_{i\ell} = 1$ for all $i \in [n] \setminus \{a(u)\}$. Consequently, every agent values her bundle at time T at 0, yet has value 1 for every item she did not receive. Since some agent received at least T/n items, it follows that $\textsc{Envy}(T) \geq T/n$.

For the upper bound, observe that allocating items in a round-robin manner gives every agent at most $\lceil T/n \rceil$ items. It follows that $\textsc{Envy}(T) \leq \lceil T/n \rceil$. ∎

5.6 Discussion

We conclude with a discussion of two issues that have not yet been addressed.

First, we have assumed that agents have additive valuations for bundles of items. This common assumption is typically considered strong. But for the purpose of defining envy in our online setting we consider it to be very natural. Indeed, in an online setting, the allocated items would typically be used independently of each other. Consequently, we can interpret the envy of i for j, $\sum_{t \in A_i} v_{it} - \sum_{t \in A_j} v_{it}$, as $\sum_{t=1}^{T} v_{it}(\mathbb{I}_{t \in A_i} - \mathbb{I}_{t \in A_j})$. Notice that this is a sum over per-round envy. In other words, the additivity assumption actually amounts to envy being additive over time.

Finally, we have focused with single-minded determination on a single goal — that of minimizing envy. A possible concern is that low envy, in and of itself, is not sufficient to lead to intuitively fair outcomes, as has been observed in various contexts [71, 43]. Be that as it may, even if one is interested in a combination of low envy and other properties (Pareto efficiency comes to mind), our results establish a baseline for what one could hope for, and are therefore a crucial first step in any such investigation.

Chapter 6

Political districting

6.1 Introduction

In the United States, representatives to the House of Congress and many other bodies like state legislatures, city councils etc., are elected by dividing a region, or state into disjoint geographical areas called *districts*. Each district typically elects a representative via a plurality election. A partition of the space into a set of districts is called a *districting*. We will focus on districting in the context of electing representatives to the House of Congress.

A valid congressional districting must satisfy several constraints, some prescribed at a federal level, others at the state level. These codify the principle of 'one person, one vote', by requiring that districts contain the same number of people. Most states require *contiguity* — it must possible to move from any point in a district to any other without leaving it. Many states' constitutions also contain wording to the effect that districts must be compact and/or retain 'communities of interest.' Finally, the Voter Rights Act demands minority groups be given equal opportunity to participate in the democratic process, this is often interpreted as requiring districts in which the will of a minority group determines the outcome of the election.

Despite these guidelines, what constitutes a valid redistricting is open for interpretation. For example, is a district contiguous when two distant geographical regions are joined by a strip of highway or a railroad? Does 'one person, one vote' mean that the total number of people in every district should be the same, or is it referring to the number of people of a voting age? Because of this, the process of redistricting, which happens every ten years based on the most recent census data, is a contentious issue. The most recent round of redistrictings, in 2010, saw several districtings accepted which seemed to benefit one political party over the other. The process of establishing a political advantage through redistricting is called (partisan) *gerrymandering*. Other forms of gerrymandering exist, for example racial gerrymandering, but we will focus on districtings with partisan bias.

Gerrymandering has been a part of the public discourse since at least 1812, when the infamous 'Gerry-mander' cartoon [124] appeared (see Figure 6.1), satirizing a salamander-shaped district signed off on by then governer of Massachusetts, Elbridge Gerry. However, federal courts have been reticient to participate in the discussion of what makes a districting partisan, citing the lack of an appropriate constitutional standard against which to evaluate districtings. In the wake of Gill v. Whitford [1], where a District Court used a metric called the *efficiency gap* to evaluate claims of a partisan gerrymander, there have been several rulings in similar cases, notably in Maryland, North Carolina and Pennsylvania, where the Pennsylvania Supreme Court redraw the state's districting after the original was deemed unconstitutional [2].

Figure 6.1: The infamous salamander shaped district that Elbridge Gerry signed off on, immortalised in cartoon form by Elkanah Tisdale [124]

In an attempt to avoid future gerrymanders, several states have given the power to redistrict to independent bodies. But the question remains: how should such a body evaluate a proposed districting to evaluate whether it is partisan? Given the geographical and political differences between states, is it even possible to come up with a notion of fairness in this context which can *always* be satisfied? We propose such a notion.

6.1.1 Our approach and results

We begin by discussing proportionality, the property that a party should win a fraction of the available congressional seats that best reflects its overall fraction of support. Unfortunately, we will see in section 6.2 that it is impossible to guarantee proportionality. In its stead many metrics for fairness have been proposed, each striving to capture a different aspect of idealised and impartial districtings. We briefly review several before proposing our own.

Our guiding question is an issue central to the functioning of independent districting committees: what does a fair districting look like, and how many districts may a party expect to win in such an impartial districting? We start from an impartial protocol: flip a fair coin and give the party who wins the coin flip complete control over the redistricting process. We claim that the number of districts a party wins in expectation under this protocol is a reasonable representation of what they may expect from an impartial districting. Interestingly, in the absence of geographic constraints, this notion of fairness reduces to a party's proportional share of the districts. A part of the appeal of this target lies in the fact that it remains reasonable when proportionality does not, for example, if a party with 45% of the statewide support wins 0 districts in *every* feasible districting, then their target number of districts is 0, which is far from their proportional share of the districts. We show in Section 6.3.1 that, in a model of districting where voters are points on a plane, it is always possible to find a districting in which every party wins their target number of districts (up to rounding). The result relies on an extension of the 'ham sandwich theorem' which states that given two colored sets of points on a plane, it is possible to divide the plane into convex regions each containing the same number of points of every color. This result is extended to more general models of the districting problem under mild conditions.

We believe this result may be a tool for independent redistricting committees: every proposed districting can be evaluated in light of the existence of a districting satisfying this particular notion of fairness. It may also open the door to more nuanced computational approaches to districting: the existence result says that adding fairness constraints to an optimization program that returns, say, the most compact districting does not impact the feasibility of the program. In Section 6.3.2 we heuristically study the effect that adding such constraints have on various objectives in Pennsylvania using data from the 2016 presidential elections.

Finally, in Section 6.4, we examine exact models for computing optimal districtings. We propose a recursive model which breaks symmetry by numbering precincts and letting the lowest numbered precinct in a district be its *root*. Contiguity constraints are modeled recursively and exploit the structure of the underlying graph to decrease the maximum level of recursion. Preliminary experimental results are promising.

6.1.2 Related work

Redistricting has been studied from the perspective of fair division before. Pegden et al. [101] proposed a 'I-cut-you-freeze' protocol for coming up with a valid districting as an extension of the traditional 'I-cut-you-choose' protocol for cake-cutting. In round 1 party A proposes a valid districting, and party B freezes (fixes) one of the districts in this districting. Frozen districts remain unchanged throughout the remainder of the protocol. In round 2 party B proposes a districting of the unfrozen part of the state, after which party A freezes one district. Parties alternately propose and freeze districts until every district is fixed. Pegden et al. [101] show that in the absence of geographic constraints, this protocol has the fairness property that a party with the minority of the votes will not win a majority of the districts. They also find that neither party has the power to unilaterally create a district which protects one of their incumbents. Their analysis does not extend to the case where there are geographic constraints, but conceptually the protocol shows that a competitive process between two parties can balance their conflicting interests.

Our fair target property is perhaps most similar to the geometric target of Landau et al. [86]. They propose a protocol with two political parties and an independent body. The independent body splits the state into two and ask the parties which part they would prefer to divide. As soon as the parties agree to an allocation, they may redistrict their part of the state as they wish. If the parties do not agree to an allocation, the independent body retries with a new split. Landau et al. argue that, upon completion, every party is guaranteed to win at least as many districts as the midpoint between their best and worst possible outcomes subject to the location of the dividing line.

Our fair target is similarly defined as the midpoint between a party's extreme outcomes but it is more general in that we do not require the existence of an arbitrary dividing line. Furthermore, Landau et al.'s protocol may lead to districtings where each half of the state is districted in an extremely biased way toward one party. We do not focus on a protocol, instead we guarantee that a districting satisfying our fairness target always exists. This allows us to optimize a secondary objective subject to this fairness constraint.

Finally, we propose an exact model for redistricting. The first exact method may have been the two-stage approach of Garfinkel and Nemhauser [72], in which they generate all possible districts in stage 1 and solve a set partitioning problem in stage 2 to minimize population deviations. Mehrotra et al. [94] use column-generation to solve a similar two-stage problem and incorporate the compactness of a district into the objective function of the set partioning problem. Oehrlein and Haunert [99] extend Shirabe's model for spatial unit allocation [117, 118] and formulate a multi-commodity flow network that captures contiguity and population equality as well as a form of compactness. Li et al. [88] formulate a quadratic model to essentially minimize the sum of the inter-precinct distances for precincts assigned to the same district. The model captures population equality and this flavour of compactness, but does not model contiguity constraints.

6.2 Districting metrics

The recent push towards independent committees for redistricting raises the question of how such a body is to evaluate whether a proposed districting is impartial, fair or reasonable. Intuitively, the most natural property a districting should satisfy *proportionality* — the fraction of districts a party wins should closely reflect its statewide support. Of course, some amount of rounding is inevitable, since states have relatively few districts, but unfortunately the problem with proportionality reaches much deeper than this.

Imagine that every household (alternatively precinct, census block or voter tabulation district) contains 2 people who support party A, and 1 who supports party B. No matter how these households are assigned to districts, party B will never win a district despite having 33% of the statewide support. It is not hard to make this more extreme, in the worst case the minority party may win $50 - \epsilon$% of the votes yet win 0 districts in *any* districting.

Although this example is discouraging, we may hope that real-world voter distributions do not give rise to such degenerate instances. Unfortunately, they do: In the 2016 presidential election in Massachusetts the Republican party won 32.8% of the votes compared to the Democrats' 60%, yet they won 0 of the 9 available congressional districts. Duchin et al. [63] show that this was not due to gerrymandering: there exists no districting satisfying Massachusetts' districting requirements in which the Republican party would have won any congressional seats.

Because of these types of examples, using proportionality to evaluate the fairness of a districting plan is unreasonable and we turn to other approaches for measuring fairness in districtings, each focusing on a different aspect of 'ideal' districtings, for example, impartiality, compactness, and a sensitivity to shifts in the political landscape.

6.2.1 Seats-votes curve

The seats-votes curve is a simple visualization which represents how responsive a districting is to changes in the voting population's preferences. Two important metrics are immediately visible from a seats-votes curve: the fraction of seats that a party wins with 50% of the votes, and the fraction of votes required for the party to win 50% of the congressional seats.

Seat-vote curves have been an integral part of the discussion on gerrymandering at least since Tufte [126] analysed nationwide seat-vote curves to study the relationship partisan districtings and swing-ratios. Existing districtings may be evaluated by simulating uniform (or proportional) shifts in voter preferences. The discrete nature of congressional seats naturally give rise to a piecewise constant seat-vote curve. Recent work takes a more nuanced view of district outcomes by modelling the probability that a party wins a district as a probit or logit function [50, 129], yielding smooth curves. See, for example, Nagle [98] for a recent analysis of the seat-vote curves of various districting plans proposed in Pennsylvania.

6.2.2 Efficiency gap

Stephanopoulos and McGhee [121] proposed the *efficiency gap* as a simple, numerical measure of whether a districting is partisan. The idea is that an impartial districting would cause both parties to *waste* an equal amount of votes. All the votes received by the minority party in a district is declared to be wasted; for the majority party every vote beyond the $50 + \epsilon\%$ threshold required to secure victory is wasted. Formally, the efficiency gap is defined as

$$EF(D) = \frac{|w_a(D) - w_b(D)|}{\text{total votes}}$$

where D is a districting and $w_x(D)$ is the number of wasted votes for party x in districting D.

A small efficiency gap is preferable. In practice, a threshold of 8% was proposed by Stephanopoulos and McGhee [121]. Curiously, this defines a region of acceptable districtings in the seat-vote curve which will actively reject proportional outcomes, as shown in Figure 6.2. Also notice that for states where one party receives more than 80% of the vote it is impossible to come up with a districting with an efficiency gap smaller than 8%. The efficiency gap also promotes districts where parties waste the same number of votes, in other words, districts where party support is split 75%-25%. Other shortcomings of the efficiency gap is discussed in Bernstein and Duchin [25], including the fact that it does not discourage cracking and packing nor say anything about competitiveness or compactness.

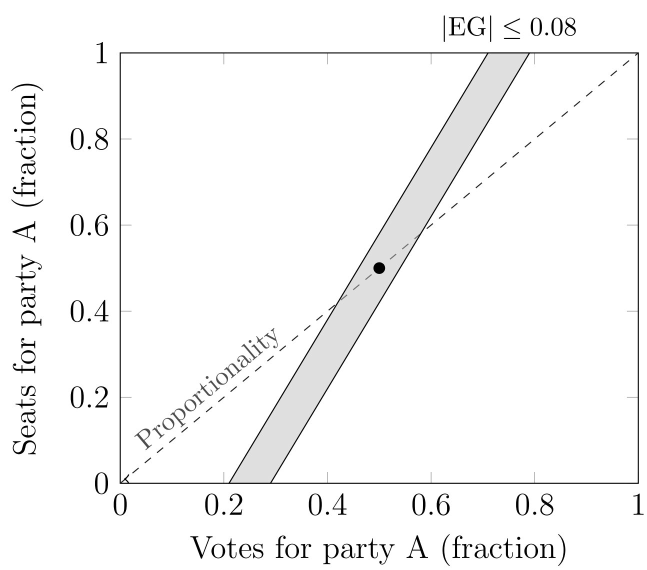


Figure 6.2: Assuming equal turnout in districts, the region of outcomes deemed acceptable with an absolute efficiency gap below 0.08 excludes many proportional outcomes.

Despite these concerns, efficiency gap analysis have played a part in several high-profile court cases, including Gill v. Whitford [1], which marked the first time in 30 years that a partisan gerrymandering case was successfully brought before a federal court [25].

6.2.3 Compactness measures

Thanks to the famous gerrymandering cartoon above, compactness is often one of the first districting criteria that jump to mind. At least 18 states have compactness as a formal requirement for congressional districts, however, there is little guidance about what constitutes a compact districting. For example, the Constitution of Illinois only says 'Legislative districts shall be compact', while Arizona requires that 'Districts shall be geographically compact and contiguous to the extent practicable.'

Because of the lack of formal guidelines, compactness have had a complicated legal history. For example, it has been said that "reapportionment is one area where appearances do matter," [3] and also that the "Constitution does not guarantee regularity of district shape" [3]. Although there exists concerns about using mathematical measures of compactness in the absence of providing end-to-end technical guidance [15], many notions of compactness have been proposed and some have impacted legal interpretations of valid districtings. Most attempt to capture some notion of dispersion or irregularity in shape, but no one measure has been widely accepted as the gold standard. Proposed measures include

- Length-width ratio: The ratio of the length to the width of the minimum bounding rectangle [80];

- Convex hull: The ratio of the area of the district to the area of its convex hull;

- Reock: The ratio of the area of the district to its minimum bounding circle [113], Gärtner [73] describes how to find such bounding circles;

- Polsby-Popper: The ratio of the area of the district to the area of the circle with the same perimeter [102, 115].

6.2.4 Distributional approaches

A recent approach is to compare the characteristics of a given districting to those that arise 'naturally', given the political geography of a state. This is done through Markov Chain Monte Carlo approaches [62, 51, 82, 68]. Given a graph representing the, say, precincts of a state, define a valid districting to be a partition of the vertices into subsets, so that the subgraph induced by each subset is contiguous and satisfies the necessary compactness and population constraints.

Every node in the Markov chain will represent a valid districting. One may transition from on districting to a neighbour by performing a random 'move', for example, by reassigning a precinct which currently exists on the boundary between two districts. The idea is that over the course of millions of random steps through the space of valid districtings, we may observe the distribution of certain properties of the districting, for example, the number of seats won by a specific party. A specific districting may now be evaluated in light of this distributional information by seeing whether it is an outlier and therefore likely to be highly engineered.

To the best of my knowledge there are currently no results about the convergence of such Markov chains, but techniques to make rigorous statistical claims do exist [53]. At a more basic level, it is questionable whether the 'natural' or distributionally average districtings are desirable, for example, the optimal solution of *any* optimization procedure which finds districtings will, by its very nature, be an outlier.

Despite the lack of supporting theoretical results, this approach have strong proponents and have played a part in expert testimonies [2].

6.3 Fair target property

We now consider the case with n districts and two parties, called A and B. Let $\mathcal{D}$ be the set off all districtings that are allowable given a state's relevant laws concerning population equality, compactness, minority-majority districts, etc. Let $D_a^+ \in \mathcal{D}$ ($D_b^+ \in \mathcal{D}$) denote a districting in which party A (B) wins the largest number of districts, and denote this quantity by $a(D_a^+)$ ($b(D_b^+)$, respectively). Denote with D_a^- a districting where A wins the least number of districtings. Notice that $a(D_a^-) = n - b(D_b^+)$ when there are only two parties. D_b^- is defined similarly.

We propose that an impartial districting should strive to let each party win a target number of districts, with the respective targets given by

$$t_a = \frac{a(D_a^+) + a(D_a^-)}{2}, \qquad t_b = \frac{b(D_b^+) + b(D_b^-)}{2} = n - t_a.$$

In other words, we claim an impartial outcome is one in which a party wins a number of seats equal to the midpoint between what they win in their best and worst outcomes. Notice that a party's popularity does not influence how heavily each of their extremes are weighed, as any difference in popularity will already be reflected in the number of seats they win in their extreme outcomes.

As an example, recall the case of the 2016 presidential election in the state of Massachussettes, where the republican party won 32% of the general vote, yet no districting allowed them to win a congressional seat. Here our target number of districts for the republican party is 0 since $b(D_b^+) = 0$.

One argument for this target is the following: Imagine a procedure in which a fair coin is flipped, and whichever party wins the coin flip is given absolute power to redistrict a state as they wish, subject to a specified set of laws. Even though this procedure leads to extremely biased districtings, it is certainly impartial, as no party is inherently favoured. We propose as target the expected number of districts that a party wins under this impartial procedure.

How does this target compare to better-known measures of fairness? In the absence of geographic constraints, in other words, when voters in a state may be arbitrarily partitioned into districts, our target reduces to proportionality.

Theorem 6.1. *Suppose an supporters of party A and bn supporters of party B is to be partitioned into n districts, each with size $a+b$. In the absence of geographic constraints $t_a \to \frac{a}{a+b} \cdot n$ and $t_b \to \frac{b}{a+b} \cdot n$ as $a+b \to \infty$.*

Proof. Without loss of generality, assume party A is the minority party. Since $t_a + t_b = n$ it is sufficient to show $t_a \to \frac{a}{a+b}$ as $a+b \to \infty$. For ease of exposition assume $a+b$ is odd.

Since party A is the minority, it wins 0 districts when every district consists of a supporters of A and b supporters of B.

Party A wins the largest number of districts by having exactly $(a+b+1)/2$ supporters in as many districts as possible. It follows that party A can win at most

$$x \cdot \frac{a+b+1}{2} = an \Rightarrow x = 2 \cdot \frac{a}{a+b+1} \cdot n = 2 \cdot \frac{a}{a+b} \cdot \frac{a+b}{a+b+1} \cdot n$$

districts. We may conclude that $t_a = \frac{a}{a+b}(\frac{a+b}{a+b+1})n \to \frac{a}{a+b} \cdot n$ as $a+b \to \infty$. $\square$

Whatever deviation the factor $\frac{a+b}{a+b+1}$ causes from exact proportionality is minuscule at the scale of practical districting problems. For example, in Pennsylvania each district contains approximately 700 000 people, so $\frac{a+b}{a+b+1} = 0.9999985$. A second interpretation of this target is that it is proportionality, to the extent possible given the political geography of a state.

A benefit of this target is that it provides a layer of abstraction with a lot a flexibility to incorporate any future constitutional or structural changes to how districtings are drawn. Any new constraints that may be imposed on districtings only changes the set of feasible districtings, the definition and interpretation of the target remains unaffected.

Despite the attractive properties of our fair target, the question remains whether it is possible to find a districting in which both parties win their target number of seats (up to rounding down, since the target may be fractional).

6.3.1 Guaranteeing targets on a plane

In this section, we prove that in a specific model of elections, it is possible to guarantee that there always exists a districting in which every party wins at least their target number of districts (rounded down). The model we consider represents voters as points on a plane, each labeled according to which party they support.

Suppose a set of points $V = A \cup B$ are spread on a subset of the plane, say on $P = [0,1]^2$. Assume no three voters are colinear, and that points (voters) in A (B) support party A (B, respectively). Let n be the number of districts to partition P (and V) into. A valid districting D is a partition of the voters into districts $d_1, \dots, d_n$ each defined by the voters in the district. A district $d_i \subseteq V$ is assigned an area $A_d \subseteq P$. For $X \subseteq V, A \subseteq P$ we say that $X \in \text{int}(A)$ if $x \in \text{int}(A)$ for all $x \in X$. We require that $d \in A_d$, in other words, no points (voters) lie on the boundary of a district. For a valid districting $D = \{d_1, \dots, d_n\}$ it is possible to assign areas A_{d_i} for each $d_i \in D$

such that $A_{d_i} \cap A_{d_j} = \emptyset$ for all $i \neq j \in [n]$ and $P = \cup_{i \in [n]} A_{d_i}$. We say that party A wins district d_i when $|d_i \cap A| > |d_i \cap B|$.

Assume that $|A| = an$ and $|B| = bn$. The equal population constraint requires that each district must contain $a+b$ points. We require a fairly mild assumption that $a < b-1$, in other words, that party B has a clear majority. For the state of Pennsylvania, this assumption translates into requiring the majority party to receive 36 more votes than the minority party, out of the more than 6 million votes cast.

The result relies on a generalization of the 'ham sandwich theorem'.

Theorem 6.2 ([28])**.** *Given two sets of points A, B with $|A| = an$ and $|B| = bn$ on a plane, there exists a division of the plane into convex regions so that every region contains a points from A and b points from B.*

According to Thm 6.2, we can find a districting in which B, the majority party, wins every district. In other words, there exists a districting in which A wins no districts. It follows that party A's target is half the number of districts they win in their best districting. We now show that this target is always achievable.

Theorem 6.3. *There exists a districting in which party A wins at least $\lfloor t_a \rfloor$ districts, and B at least $\lfloor t_b \rfloor$.*

Proof. If $\lfloor t_a \rfloor = 0$, then we are done since Theorem 6.2 guarantees a districting in which party A wins 0 districts. Assume $\lfloor t_a \rfloor > 0$.

Let D_A be the districting in which A wins the greatest number of seats, say k. Let $D_A^* \subset D_A$ be any $\lfloor k/2 \rfloor$ of the districts that A wins in D_A. Intuitively, we will fix these districts to ensure that party A wins the required number of districts. Practically, find a set of spanning trees $T(D_A^*) = \{T(d) : d \in D_A^*\} \subseteq P$ by treating the voters as vertices and including the necessary edges in such a way that none of the spanning trees intersect.

We now have to divide the remaining points $V' = V \setminus \cup\{d\}_{d \in D_A^*}$ into $n - \lfloor t_a \rfloor$ regions, each containing $a+b$ points, in such a way that party B wins every region. Denote with A', B' the voters supporting party A (B) in V'.

Consider $[0,1]^2$, with the points V'. If $|A'|$ is divisible by $n - \lfloor t_a \rfloor$, then Theorem 6.2 outputs a districting of V' in which party B wins every district. It remains to handle divisibility issues.

Suppose $|A'|$ is not divisible by $n - \lfloor t_a \rfloor$. Then $|A'| = a'(n - \lfloor t_a \rfloor) + r_a$ and $|B'| = b'(n - \lfloor t_a \rfloor) + r_b$, for some integer $0 < r_a, r_b < n - \lfloor 1/2 \rfloor$. Note that $r_a + r_b = n - \lfloor t_a \rfloor$ and $a + b = a' + b' + 1$.

By assumption, $a \leq b-2$, so $|B|-|A| \geq 2n$. Since party A had a majority in each of the districts in D_A^*, $|B'| - |A'| \geq |B| - |A| + \lfloor t_a \rfloor$. It follows that

$$\begin{aligned}
2n + \lfloor t_a \rfloor &\leq b'(n - \lfloor 1/2 \rfloor) + r_b - (a'(n - \lfloor t_a \rfloor) + r_a) \\
&= (b' - a')(n - \lfloor t_a \rfloor) + r_b - r_a \\
&< (b' - a')(n - \lfloor t_a \rfloor) + (n - \lfloor t_a \rfloor) \\
\Rightarrow n &< (b' - a')(n - \lfloor t_a \rfloor) \\
\frac{n}{n - \lfloor t_a \rfloor} &< (b' - a').
\end{aligned}$$

Since $\lfloor t_a \rfloor > 0$ and $b' - a' \in \mathbb{Z}$, we conclude $b' - a' \geq 2$.

By Theorem 6.2, we can now find a valid districting D_B on V' in which B wins every district. Specifically, for an arbitrary subset $X \subset B'$ with $|X| = r_b$, pretend the voters in X support party A. We now satisfy the conditions of Theorem 6.2, since $|A' \cup X| = (a'+1)n$, which yields a set of convex regions $R_1, \ldots, R_{n-\lfloor t_a \rfloor}$, each containing a set of n points $d'_1, \ldots, d'_{n-\lfloor t_a \rfloor}$ each with $(a'+1)n$ points for party A and $b'n$ points for B. Recall that $a'+1 < b'$, so party B wins all the districts in D_B.

It is possible that $R_1, \ldots, R_{n-\lfloor t_a \rfloor}$ overlap some of the spanning trees in $T(D_a^*)$. Instead of using these regions are areas, find for each $d_i, i \in \{1, \ldots, n - \lfloor t_a \rfloor\}$ an embedded spanning tree $T(d_i)$ which does not intersect any of the spanning trees in $T(D_a^*)$. Since we are only interested in an embedding in P, the edges of the spanning tree does not have to be straight, ensuring that it is always possible to find such a simultaneous embedding. Among the possible embeddings, it is possible to select a specific one, for example to encourage the spanning trees to be as far apart as possible, however, this is not required for a proof of existence.

We have n districts $D^* = D_A^* \cup \{d_1, \ldots, d_{n-\lfloor t_a \rfloor}\}$ where party A wins $\lfloor t_a \rfloor$ of the districts and party B wins $n - \lfloor t_a \rfloor \geq \lfloor t_b \rfloor$. It remains to assign each district $d \in D^*$ an area $A_d \subset P$ such that $x \in A_d \, \forall x \in d$, $\cup_{d \in D^*} A_d = P$ and $A_d \cap A_{d'} = \emptyset$ for $d, d' \in D^*$.

Denote with $T^* = T(D_a^*) \cup T(D_B)$ the embeddings of all the spanning trees on P. Let $\epsilon' = \min\{\text{dist}(x,y) : x \in T; y \in T'; T \neq T'; T, T' \in \mathcal{T}\}$, where $\text{dist}(x,y)$ is the euclidean distance between $x, y \in P$, be the smallest distance between any two spanning trees and set $\epsilon = \epsilon'/4$.

For $d \in D^*$, set $A'_d = \{x \in P : \text{dist}(x,y) < \epsilon, y \in T(D)\}$. By choice of ϵ, $A'_{d_1} \cap A'_{d_2} = \emptyset$ for all $d_1, d_2 \in D^*$ and $d \in A'_d$. We now have a set of disjoint regions in P, one for every district in D^*. All that remains is to assign $\bar{P} = P \setminus \cup_{d \in D^*} A'_d$, but this can be done arbitrarily since $\bar{P}$ does not contain any voters. For example, assign every $x \in \bar{P}$ to its closest region. $\square$

When we treat voters as points on a plane, this result shows that we can always find a districting guaranteeing that each party wins their target number of districts. Real districts are made up of precincts or voter tabulation districts, so it is natural to ask whether this guarantee extends to more realistic models. The following observation addresses this concern.

Observation 6.4. *Suppose there exists a districting D with $a(D) = z$ for every $z \in \mathbb{Z}$, $a(D_A^-) \leq z \leq a(D_A^+)$. Then there exists a districting in which party A wins at least $\lfloor t_a \rfloor$ districts and party B wins $\lfloor t_b \rfloor$.*

In other words, as long as it is possible to transition 'smoothly' between a party's most extreme districtings, there exists a districting satisfying each party's target. Due to the granularity and scale of real-world districting instances this is an innocuous assumption.

Finally, this observation can also be framed as an approximation result. For an instance I (a distribution of voters on a geography) and a districting D in the space of valid districtings $\mathcal{D}$ consisting of n districts (satisfying whatever constraints apply), let

$$\phi(I) = \max_{D \in \mathcal{D}} \min \left\{ \frac{a(D)}{t_a}, \frac{b(D)}{t_b} \right\}$$

denote the *fairness ratio* on I. This fairness ratio represents the fraction of the worst-off party's target that can be guaranteed in the best districting (with respect to this objective) in $\mathcal{D}$. One may ask what fairness ratio can be guaranteed for all instances, or what is

$$\phi = \min_{I \in \mathcal{I}} \phi(I)$$

where $\mathcal{I}$ is the space of districting instances, and $\mathcal{D}(I)$ is the set of valid districtings for instance I. Let $D_a^-(I), D_a^+(I)$ denote the districtings on instance I where party A wins the least and most number of districtings respectively.

Theorem 6.5. *Suppose for every instance $I \in \mathcal{I}$ and $z \in \mathbb{Z}$, $a(D_A^-(I)) \leq z \leq a(D_A^+(I))$ there exists a districting D with $a(D) = z$. Then $\phi \geq \frac{n-1}{n}$.*

Proof. Fix an instance $I \in \mathcal{I}$. For ease of presentation we suppress the dependence on I.

Suppose $t_a \in \mathbb{Z}$. Then $a(D_A^-) \leq t_a \leq a(D_A^+)$, and by assumption there exists a districting D with $a(D) = t_a$. Since $n = t_a + t_b$, $b(D) = t_b$ and $\phi(I, D) = 1$.

Suppose $t_a \notin \mathbb{Z}$. Then $t_a = k + \frac{1}{2}$, for some $k \in [a(D_A^-), a(D_A^+))$, $k \in \mathbb{Z}$, and $t_b = n - k - \frac{1}{2}$. By assumption there exists a districting D_1 with $a(D_1) = k$, and D_2 with $a(D_2) = k + 1$. We show that the best of these districtings is good enough to achieve the result.

- For D_1, $\frac{a(D_1)}{t_a} = \frac{k}{k+\frac{1}{2}}$ while $\frac{b(D_1)}{t_b} \geq 1$.
- For D_2, $\frac{a(D_2)}{t_a} \geq 1$ and $\frac{b(D_2)}{t_b} = \frac{n-k-1}{n-k-\frac{1}{2}} = \frac{k'}{k'+\frac{1}{2}}$

In other words, $\phi(I) = \max\{\frac{k}{k+1/2}, \frac{k'}{k'+1/2}\}$, for $k' = n - k - 1$. The fairness ratio for every instance in $\mathcal{I}$ is characterised by the value of k. Since k' ranges over $\{0, \ldots, n-1\}$ as k ranges over $\{0, \ldots, \ n-1\}$,

$$\phi \geq \min_{(n-1)/2 \leq k \leq n-1} \frac{k}{k+1/2} = \frac{\frac{n-1}{2}}{\frac{n-1}{2} + \frac{1}{2}} = \frac{n-1}{n}.$$

□

Proving a lower bound on ϕ requires fixing a model and giving a family of instances on which it is impossible to achieve a better fairness ratio. The proof of Theorem 6.5 shows us that such a family of instances should have odd n and $t_a = \frac{n-1}{2} + \frac{1}{2}$.

Observation 6.4 says that there always exists a districting in which each party wins at least their target number of districts, rounded down. This gives independent districting committees something to compare proposed districts against. Another implication of this result is that we may add fairness constraints to whatever optimization program is used to construct districtings without impacting feasibility. For example, if compactness is the desired objective, we may safely restrict our search to the most compact districting satisfying our fairness constraints.

6.3.2 Case study

We now investigate the effect of optimizing a property like compactness or efficiency gap subject to our fairness constraint. Finding a party's target requires knowing the most and least number of districts won by that party in any districting. Unfortunately, exact models for redistricting do not currently scale to real-world situations, so we turn to heuristics for this evaluation.

We use the open-source *gerrychain* package developed by the Metric Geometry and Gerrymandering Group[1]. This facilitates running a Markov chain from a starting districting. We begin by creating a graph representation of a state where every node in the graph corresponds to a Voter Tabulation District (VTD) and is associated with the properties of that VTD including its population, area, perimeter, number of democratic and republican voters in various elections, etc.

Moving from one state of the Markov chain (a valid districting) to an adjacent one can be done with one of two types of moves: an *edge flip*, which reassigns a single VTD currently on the boundary between two districts, or a *recombination* (called ReComb in the gerrchain documentation), which merges two adjacent districts before splitting them into two again by finding a cut which satisfies the population constraints.

Before accepting any move, it is verified whether the resulting districting satisfies the remaining constraints. We impose contiguity contraints, a maximum population deviation across districts of 2% as in [74], and a compactness constraint which limits the number of cut edges between districts to be within a factor of 2 of the starting districting (it's also possible to impose more explicit compactness constraints, for example on the average Polsby-Popper score or maximum perimeter).

For the results presented here a chain length of 20 000 recombination moves was used. This was found to be a sufficient number of moves to erase the impact that the starting districting may have had on subsequent districtings [74]. Note that if you perform only edge flips significantly it is suggested that you perform 10 000 000 moves [74]. We also experimented with a mix of edge flips and recombination moves and the results were not significantly different from those presented here.

At every node in the Markov chain we note the number of districts won by every party. A party's fair target number of districts is the midpoint between their best and worst districtings observed in the Markov chain. Additionally, we keep track of the following potential objectives:

[1]See `http://mggg.org` for more information about the Metric Geometry and Gerrymandering Group. The gerrychain python package is available at `https://github.com/mggg/GerryChain`.

- Polsby-Popper score: A compactness metric defined as $\sum_d 4\pi \cdot \text{area}_d / p_d^2$, where d ranges over all the districts in the districting and p_d is the perimeter of district d. More compact districtings have higher Polsby-Popper scores.

- The efficiency gap, which attempts to measure if a districting is partisan by comparing the number of wasted votes for every party (see section 6.2.2) .

- The percentage of the statewide vote that the democratic party requires to win 50% of the districts. This is calculated by performing proportional shifts in voter sentiment in every district until the required outcome is achieved.

- The number of competitive districts, defined as districts where the majority party wins less than 54% of the votes.

Pennsylvania

We evaluate the outcomes under different districtings using the 2016 presidential election data on a VTD level. Note that congressional elections are separate from presidential elections, so our results are not exactly representative of the 2016 congressional elections. However, congressional elections typically show a very high level of correlation with the presidential election, and the data for the presidential election was more readily available.

The number of seats won by the democratic party over the course of the run may be seen in Figure 6.3. As a result we set the target number of districts to be 7 for the democratic party and 11 for the republican party.

The efficiency gap and percentage votes that the democratic party required to win 50% of the districts (9) may be seen in Figure 6.4. Suppose we are minimizing the absolute efficiency gap to ensure an impartial districting. The minimum absolute efficiency gap observed was 0.012 at the districting shown in Figure 6.5. The democrats won 8 districts in this districting, which means that the republican party did not achieve their target number of districts. Imposing the fairness constraint that every party must win at least their target number of districts leads to a minimum absolute efficiency gap of 0.069, the corresponding districting is shown in Figure 6.5. Adding the fairness constraint has a significant impact on the objective, however, this is not always the case, for example, the maximum number of competitive

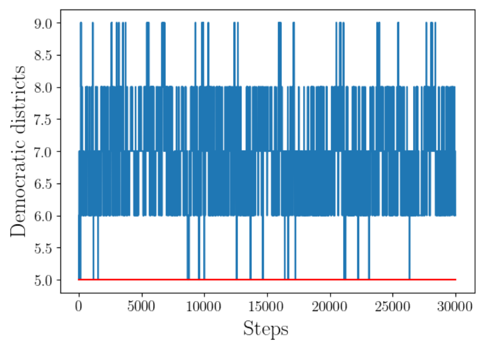


Figure 6.3: The number of districts won by Democrats in PA across the 30 000 steps in the Markov chain.

districts observed is 8, while it is possible to retain 7 competitive districts subject to the fairness constraints.

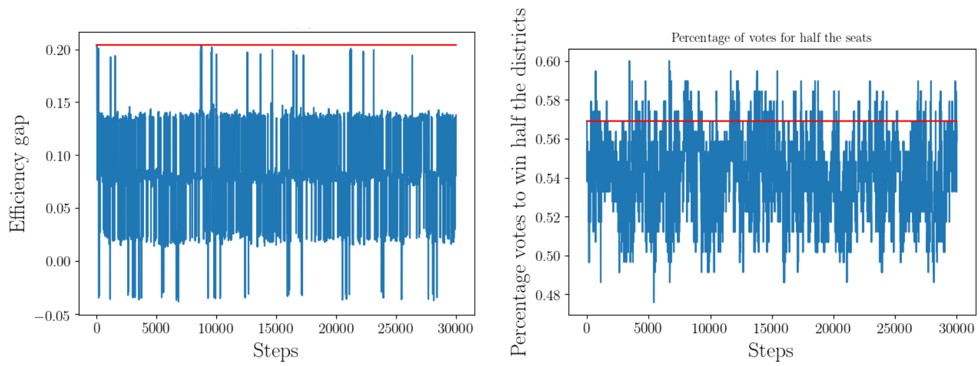


Figure 6.4: The efficiency gap (left) and percentage votes the democrats need to win 9 districts (right) across the 30 000 steps in the Markov chain.

6.4 An exact model for redistricting

In the previous section we used a heuristic to determine the largest number of districts that a party can win in any feasible districting. We now attempt to develop an integer programming model for this which avoids symmetry.

Suppose that districts are composed of *precincts* (this may in reality be VTDs or census blocks). To remove symmetry, we suppose that each district

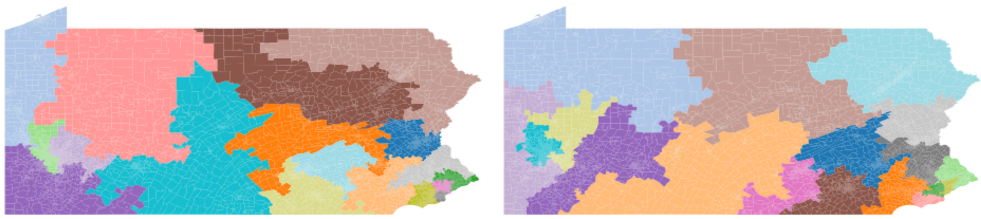


Figure 6.5: The districting which minimizes the absolute efficiency gap (left) alongside the solution subject to our fairness constraint (right) across the 30 000 steps in the Markov chain.

contains a *root*, which is the precinct in the district with the smallest index. Let binary variable $x_{ij} = 1$ for $i < j$ when precinct j is assigned to precinct i, meaning that j belongs to the district with root i. We let $x_{ii} = 1$ when i is a root. A districting must satisfy

$$\begin{aligned} &\sum_{\substack{i \\ i \le j}} x_{ij} = 1, \quad \text{all } j && (a) \\ &x_{ij} \le x_{ii}, \quad \text{all } i, j \text{ with } i < j && (b) \end{aligned} \tag{6.1}$$

where constraint (a) requires every precinct to be assigned to some district, and constraint (b) ensures precincts are only assigned to roots.

We can specify that the districting must have m districts by writing

$$\sum_i x_{ii} = m.$$

Population constraints can be enforced with

$$Lx_{ii} \le \sum_{\substack{j \\ j \ge i}} p_j x_{ij} \le U, \quad \text{all } i$$

where p_j is the population of precinct j. Bounds of $\sum_j p_j/m \cdot (1 \pm \epsilon)$ for $\epsilon = 0.02$ is often used in the literature. Similar constraints can control the number of minority voters in a district to comply with the voter rights act. This constraint can be strengthened to a convex hull formulation by rewriting it as

$$Lx_{ii} \le \sum_{\substack{j \\ j \ge i}} p_j x_{ij} \le Ux_{ii}, \quad \text{all } i. \tag{6.2}$$

This inequality is derived by writing the constraint as a disjunction

$$\begin{pmatrix} L \leq \displaystyle\sum_{j\geq i} p_j x_{ij} \leq U \\ x_{ij} \geq 0, \text{ all } j \\ x_{ii} = 1 \end{pmatrix} \vee \begin{pmatrix} \displaystyle\sum_{j\geq i} p_j x_{ij} = 0 \\ x_{ij} \geq 0, \text{ all } j \\ x_{ii} = 0 \end{pmatrix}$$

The two polyhedra have the same recession cone (namely, the origin), since their feasible sets are bounded. The standard convex hull model of this disjunction is therefore

$$\begin{aligned} &Lx_{ii} \leq \textstyle\sum_{j\geq i} p_j x_{ij}^1 \leq U x_{ii} \\ &\textstyle\sum_{j>i} p_j x_{ij}^2 = 0 \\ &x_{ij} = x_{ij}^1 + x_{ij}^2, \ \text{ all } j \end{aligned}$$

Substituting $x_{ij}^2 = x_{ij} - x_{ij}^1$, this becomes

$$\begin{aligned} &Lx_{ii} \leq \textstyle\sum_{j\geq i} p_j x_{ij}^1 \leq U x_{ii} \\ &\textstyle\sum_{j>i} p_j x_{ij}^1 = \sum_{j>i} p_j x_{ij} \end{aligned}$$

which simplifies to (6.2).

The contiguity constraints are the hardest part of the model. Existing approaches enforce contiguity with multicommodity flow formulations or by deleting edges from a spanning tree of the dual graph. We model contiguity recursively. A root is connected to itself. A precinct is connected to its root if it is adjacent to a precinct which is assigned to the same root and connected with it. Every precinct must be connected to its root for a districting to be contiguous.

Let binary variable $y_{ijd} = 1$ when precinct j is assigned to root i and they are connected in the dth level of the recursion. The constraints are

$$\begin{aligned} &y_{ii0} = x_{ii}, \ \text{ all } i && (a) \\ &y_{ik0} = 0, \ \text{ all } i, k \text{ with } i < k && (b) \\ &y_{ikd} \leq x_{ik}, \ \text{ all } i, k \text{ with } i < k, \ d = 1, \ldots, D && (c) \\ &y_{ikd} \leq \textstyle\sum_{\substack{j\in N_k \\ j>i}} y_{ij,d-1}, \ \text{ all } i, k \text{ with } i < k, \ d = 1, \ldots, D && (d) \\ &y_{ijD} \geq x_{ij}, \ \text{ all } i, j \text{ with } i \leq j && (e) \end{aligned}$$

where N_k is the set of precincts adjacent to precinct k, and D is the maximum number of precincts through which a precinct can be connected to its root.

Constraint (a) says a root is connected to itself, and constraint (b) says that these are the only precincts connected on level 0 of the recursion. Constraint (c) prevents a precinct from being connected to a root which is not his own. Constraint (d) says that if none of a precincts neighbours are connected to its root on level $d-1$ of the recursion, then that precinct can not be connected to its root by level d. Finally, constraint (e) says every precinct must be connected to its root by level D.

This model contains triply subscripted variables, but the hope is that D is relatively small compared to the number of precincts, otherwise we may have too long and thin districts. For example, the maximum possible D on an $n \times n$ grid is $2n$, so we may expect $D \in O(\sqrt{|V|})$ where V is the set of vertices of the graph (precincts).

Preprocessing may also remove some x_{ij} variables from the model along with their corresponding y-variables. For example, we may observe that two precincts are extremely far apart in the graph, or that there is no way to let them be a part of the same district without violating population constraints.

6.4.1 Modeling the Objective Function

We can add an objective to this model, for example, maximize the number of districts won by a specific party. Let binary variable $w_i = 1$ if the party A wins the district rooted at precinct i, with $w_i = 0$ if precinct i does not root a district or roots a districts which party A loses. The objective is $\max_{w,x,y} \sum_i w_i$. Then w_i might be defined by the constraint

$$Uw_i \geq \sum_{j>i} (a_j - \tfrac{1}{2}p_j)x_{ij} \tag{6.3}$$

where a_j is the number of voters for party A in precinct j and U is an upper bound on the district population. The constraint ensures that if party A has the support of more than half the voters in district i, then $w_i = 1$. This is a weak model which can be strengthened to a convex hull model.

To simplify notation, we temporarily replace $\sum_{j>i}(b_j - \frac{1}{2}p_j)x_{ij}$ with the new variable y_i. We next write the constraint as a disjunction:

$$\begin{pmatrix} y_i \geq 1 \\ w_i = 1 \end{pmatrix} \vee \begin{pmatrix} y_i \leq 0 \\ w_i = 0 \end{pmatrix}$$

This disjunction cannot be modeled as an MILP because the recession cones of these two disjuncts differ. The first recession cone is the ray $y_i \geq 1$, and the second is the ray $y_i \leq 0$. To equalize the recession cones, we require that y_i lie in the interval $[-U, U]$. This yields the disjunction

$$\begin{pmatrix} y_i \geq 1 \\ y_i \leq U \\ w_i = 1 \end{pmatrix} \vee \begin{pmatrix} y_i \leq 0 \\ y_i \geq -U \\ w_i = 0 \end{pmatrix}$$

The standard convex hull model of the disjunction is

$$\begin{aligned} & w_i \leq y_i^1 \leq U w_i \\ & -U(1-w_i) \leq y_i^2 \leq 0 \\ & y_i = y_i^1 + y_i^2 \end{aligned}$$

Substituting $y_i^2 = y_i - y_i^1$ and writing y_i^1 as z_i, we obtain

$$\begin{aligned} & w_i \leq z_i \leq U w_i \\ & y_i \leq z_i \leq y_i + U(1-w_i) \end{aligned}$$

Finally, we can substitute the definition of y_i and obtain a convex hull model for the original constraint:

$$\begin{aligned} & w_i \leq z_i \leq U w_i \\ & \sum_{j>i} (b_j - \tfrac{1}{2}p_i) x_{ij} \leq z_i \leq \sum_{j>i} (b_j - \tfrac{1}{2}p_i) x_{ij} + U(1-w_i) \end{aligned}$$

The model can also be written

$$\begin{aligned} & w_i \leq z_i \leq U w_i \\ & z_i - U(1-w_i) \leq \sum_{j>i} (b_j - \tfrac{1}{2}p_i) x_{ij} \leq z_i \end{aligned}$$

It introduces a continuous variable z_i for each tract. It may be advantageous to project out the z_i and obtain the model

$$\begin{aligned} & U w_i \geq \sum_{j>i} (a_j - \tfrac{1}{2}p_i) x_{ij} \\ & (U+1) w_i \leq U + \sum_{j>i} (a_j - \tfrac{1}{2}p_i) x_{ij}. \end{aligned} \tag{6.4}$$

6.4.2 Other valid inequalities

We can find additional valid inequalities by making simple inferences about the number of districts a party can win in the absence of geographic constraints.

Suppose a_i (b_i) is the number of voters in district i which support party A (B), and p_i the population of district i. Let m be the number of districts, and $\bar{p} = \sum_i p_i/m$ the average number of voters per district. Under exact population equality we may infer that

$$\sum_i w_i \leq \left\lfloor \frac{\sum_i a_i}{\lceil \bar{p} + 1/2 \rceil} \right\rfloor$$

since $\lceil \bar{p} + 1/2 \rceil$ votes are required to win a single districting. The same constraint from party B's perspective provides a lower bound on $\sum_i w_i$:

$$\sum_i (1 - w_i) \leq \left\lfloor \frac{\sum_i b_i}{\lceil \bar{p} + 1/2 \rceil} \right\rfloor \iff \sum_i (1 - w_i) \leq \left\lfloor \frac{\sum_i (p_i - a_i)}{\lceil \bar{p} + 1/2 \rceil} \right\rfloor .$$

These constraints can be adapted when exact population equality is not required.

6.4.3 Preliminary computational results

We test this model through some small scale experiments on $n \times n$ grid graphs for $n \in \{5, \ldots, 13\}$. Every vertex in the grid consists of one voter who is randomly assigned to either support party A or party B. In each instance the number of districts required is $m = n$ and we enforce exact population equality. For example, we must partition a 5×5 grid with 25 vertices into 5 districts each with 5 vertices. We test three versions of the problem: The first version has a dummy objective function and only attempts to find a feasible districting (Feas). The other two maximize the number of district won by party A, first using (6.3) (denoted by Max), then using the stronger constraints in (6.4) and the valid inequalities in Section 6.4.2 (Max+S+VE). Instances are given a time limit of one hour we perform ten repetitions for every instance size.

Table 6.1 reports the average time taken for the solved instances, the number of instances that did not solve to optimality within the time limit is indicated in parenthesis.

Table 6.1: Time (in seconds) to solve the recursive districting model on an $n \times n$ grid. The number of instances out of ten that timed out in one hour is shown in parentheses.

n	Feas	Max	Max+S+VE
5	0.5	4.61	2.67
6	1.10	5.98	2.24
7	1.72	244.90 (7)	221 (5)
8	8.82	396.88	101.99
9	75.49 (5)	–	3286.9 (9)
10	457.42	–	–
11	889.05	–	–
12	3247.8 (?)	–	–
13	– (10)	–	–

We see that feasible districtings can be found within an hour for graphs with up to 140 vertices. When maximizing the number of districts a party wins, graphs with up to 80 vertices can be handled within the time limit. We also observe that the stronger formulations and valid inequalities speed up solution times by a significant margin.

For context, Pennsylvania has 67 counties, 14 of which are split across districts in the most recent districting. One may imagine a model of PA which splits these 14 large counties into $3-5$ pieces each and keeps the remaining counties intact. The resulting graph will have around 100 vertices. This approach introduces its own difficulties, for example, the method is no longer exact, since there may be feasible districtings which split some of the other counties. It also becomes harder to impose strong bounds on D when different parts of the graph represent geographic areas of very different sizes.

It is conceivable that, with refinements like additional valid inequalities, such a model will be solvable, especially since redistricting only happens once every ten years. It remains to be seen to what degree optimality can be recovered when the problem is partially modeled at a county level instead of at VTD level.

Chapter 7

Conclusion

We studied computational social choice problems in which human factors, for example the cognitive load involved with eliciting utility functions and the fact that we expect certain systems to be impartial, played a central role. These new paradigms for voting and fairness not only lead to rich technical questions, but may also influence how we make decisions in the future.

Websites like RoboVote[1] and Spliddit[2] enable groups of any size to use state of the art voting and fair division mechanisms. This impact is not limited to individuals. The proliferation of participatory budgeting shows that there is a willingness to change how we make decisions as a society. We are more connected today than ever before — this creates the potential for a more direct or participative democracy in which citizens have a say in state decisions on a day–to–day basis. The mechanisms that manage such interactions will have to be designed to respect many of the same principles we considered when we studied participatory budgeting.

On a different front, increasing automation threatens jobs in many traditional industries. There is the real possibility that a large portion of the current work force will no longer be needed or able to work. How does our wage-based society change if this happens? Whatever the solution, it will have to carefully weigh fairness at an individual level with societal considerations like efficiency.

We conclude with more immediate concerns: avenues of investigation raised or left unexplored by this work.

[1] www.robovote.org

[2] www.spliddit.org

Participatory Budgeting. In discussions with officials in charge of designing the Barcelona participatory budgeting systems, one of the main concerns of the distortion-based approach in Chapter 2 was that distortion-minimizing aggregation methods are complex and it may be hard to explain the outcomes to the public, leading to a perceived lack of transparency. Current greedy aggregation methods lean to far in the other direction and may lead to extremely poor outcomes. An important avenue of research is finding voting rules that lead to outcomes with provably high social welfare that can also be easily explained.

After publication of our initial paper, Bhaskar et al. [29] studied threshold approval voting and observed that, for a large enough set of voters, randomizing the threshold independently for every voter leads to distortion approaching 1. Intuitively, the probability that a voter approves an alternative is equal to his utility for it, so the expected number of approvals an alternative receives equals its social welfare. Bhaskar et al. [29] also find that a variation of the voting rule used in Theorem 2.3, which used a harmonic scoring rule, is truthful and retains distortion $O(\sqrt{m \log m})$.

Though these results are theoretically interesting, we expect manipulating an election at this scale to be a daunting task. Instead of demanding truthfulness, a more practical alternative is to strive for outcomes that satisfy a version of proportional fairness. An outcome is said to be in the *core* of a participatory budgeting problem if there is no coalition of voters who can benefit if they are allowed to unilaterally spend their proportional fraction of the budget. This approach does not circumvent any of the elicitation concerns raised in Chapter 2, but it does provide a compelling fairness guarantee which may be easier to motivate to public officials. It is currently unknown whether the core of a participatory budgeting problem is always non-empty.

Fair division of indivisible goods. In Chapter 5 we studied envy in a setting where items arrive dynamically over time, and are irrevocably assigned to agents upon arrival. Subsequently, He et al. [81] studied a version of the problem in which it is allowed to reassign previously assigned items when necessary. They bound the number of assignments need maintain an EF1 solution after a batch of arrivals.

Another issue that Chapter 5 highlighted is that envy-minimizing algorithms are not necessarily efficient. Recall that uniform random assignment was essentially optimal in terms of envy. Consider an example where one agent values every item at 1, and every other agent values all items at 0, it is easy to see that random allocation may lead to outcomes with social welfare that is a factor of n (where n is the number of agents) less than optimal. We can measure efficiency in terms of approximate Pareto-optimality, a solution is α-Pareto optimal if multiplying the utility point by $1/\alpha$ gives an infeasible utility point (a point outside the Pareto frontier). Psomas and Zeng [110] study this trade-off between efficiency and fairness in a model where the adversary picks a distribution over instances. They show that it is impossible to find an algorithm which simultaneously has vanishing envy and leads to allocations in which agents' utilities are $(\frac{1}{n} + \epsilon)$-Pareto optimal.

One interpretation of this trade-off between efficiency and fairness is that it is the cost (in terms of social welfare) of considering envy as an objective. envy. One may also ask what is the monetary cost of requiring truthful mechanisms to be envy-free in a setting where you are allowed to allocate both money and goods to agents. This line of research may highlight interesting connections between revenue maximizing auctions and fair division problems.

Political redistricting. Although political redistricting is one of the oldest problems studied in the integer programming literature, existing models either do not capture the full complexity of redistricting, or do not scale to real-world sizes. The development of scalable algorithms for redistricting will enable legislatures to transparently optimize for properties that they deem to be desirable.

Perpendicular to drawing districtings, is evaluating existing or proposed districtings for bias. One very important characteristic of an impartial districting is that it is sensitive to voters changing their political affiliation. For example, if the statewide support for a party shifts from 45% to 51%, then this should be reflected in the outcome of the congressional elections. Our fairness property (and the Markov chain-based approaches) evaluates a districting based on actual votes cast in a specific election, but fails to say anything about how good the districting would be if voters shift in some direction. It is important to model uncertainty about voter preferences if we wish to study a districting in a wide variety of possible outcomes. A simple way to do this has already been used to construct seats-votes curves. The way it works is to look at the votes cast in a single election, and assign a party a probability of winning a district based on their fraction of support instead of a binary victory or loss. It is possible that an approach similar to that of robust optimization is more useful for modeling the uncertainty about voter preferences, and it remains to be explored what theoretical guarantees we can provide for districtings in the presence of this uncertainty.

Bibliography

[1] *Gill v. Whitford*, volume F. Supp 3d, 2016 WL-6837229, 15-cv-421-bbc. W.D. Wisc.

[2] *League of Women Voters et al. v. Commonwealth of Pennsylvania et al.*, volume No. 159 MM.

[3] *Bush v. Vera*, volume 517 U.S. 952.

[4] M. Aleksandrov and T. Walsh. Pure Nash equilibria in online fair division. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 42–48, 2017.

[5] M. Aleksandrov, H. Aziz, S. Gaspers, and T. Walsh. Online fair division: Analysing a food bank problem. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2540–2546, 2015.

[6] E. Anshelevich and J. Postl. Randomized social choice functions under metric preferences. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 46–52, 2016.

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

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

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

[10] K. J. Arrow. *Social Choice and Individual Values*. Wiley, 1951.

[11] K. J. Arrow and M.D. Intriligator, editors. *Handbook of Mathematical Economics*. North-Holland, 1982.

[12] H. Azari Soufiani, D. C. Parkes, and L. Xia. Random utility theory for social choice. In *Proceedings of the 26th Annual Conference on Neural Information Processing Systems (NIPS)*, pages 126–134, 2012.

[13] H. Azari Soufiani, D. C. Parkes, and L. Xia. Preference elicitation for general random utility models. In *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 596–605, 2013.

[14] H. Azari Soufiani, D. C. Parkes, and L. Xia. Computing parametric ranking models via rank-breaking. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, pages 360–368, 2014.

[15] R. Barnes and J. Solomon. Gerrymandering and compactness: Implementation flexibility and abuse. *arXiv preprint arXiv:1803.02857*, 2018.

[16] D. Baumeister, S. Bouveret, J. Lang, N. Nguyen, T. Nguyen, J. Rothe, and A. Saffidine. Positional scoring-based allocation of indivisible goods. *Journal of Autonomous Agents and Multi-Agent Systems*, 31 (3):628–655, 2017.

[17] S. Ben-David, A. Borodin, R. Karp, G. Tardos, and A. Wigderson. On the power of randomization in on-line algorithms. *Algorithmica*, 11(1): 2–14, 1994.

[18] J. G. Benadé and J. N. Hooker. Optimization bounds from the branching dual. *Informs Journal on Computing*, 2018.

[19] J. G. Benadè, A. Kahng, and A. D. Procaccia. Making right decisions based on wrong opinions. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*, pages 267–284, 2017.

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

[21] J. G. Benadè, N. Itzhak, N. Shah, A. D. Procaccia, and Y. Gal. Efficiency and usability of participatory budgeting methods. *Unpublished*, 2018.

[22] J. G. Benadè, A. Kazachkov, A. D. Procaccia, and C.-A. Psomas. How to make envy vanish over time. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*, pages 593–610, 2018.

[23] J. G. Benadè, P. Gölz, A. D. Procaccia, and C.-A. Psomas. No stratification without representation. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*, 2019.

[24] J. G. Benadè, A. D. Procaccia, and M. Qiao. Low-distortion social welfare functions. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019.

[25] M. Bernstein and M. Duchin. A formula goes to court: Partisan gerrymandering and the efficiency gap. *Notices of the AMS*, 64(9):1020–1024, 2017.

[26] S. Bernstein. *The Theory of Probabilities*. Gastehizdat Publishing House, 1946.

[27] D. Bertsimas, V. Farias, and N. Trichakis. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research*, 61(1):73–87, 2013.

[28] S. Bespamyatnikh, D. Kirkpatrick, and J. Snoeyink. Generalizing ham sandwich cuts to equitable subdivisions. *Discrete & Computational Geometry*, 24(4):605–622, 2000.

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

[30] A. Bogomolnaia and H. Moulin. A new solution to the random assignment problem. *Journal of Economic Theory*, 100:295–328, 2001.

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

[32] S. Bouveret and J. Lang. A general elicitation-free protocol for allocating indivisible goods. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 73–78, 2011.

[33] S. J. Brams and P. C. Fishburn. *Approval Voting*. Springer, 2nd edition, 2007.

[34] S. J. Brams and A. D. Taylor. *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge University Press, 1996.

[35] F. Brandt, V. Conitzer, U. Endress, J. Lang, and A. D. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

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

[37] E. Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103, 2011.

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

[39] Y. Cabannes. Participatory budgeting: A significant contribution to participatory democracy. *Environment and Urbanization*, 16(1):27–46, 2004.

[40] C. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, 2011.

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

[42] I. Caragiannis, A. D. Procaccia, and N. Shah. Modal ranking: A uniquely robust voting rule. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*, pages 616–622, 2014.

[43] I. Caragiannis, D. Kurokawa, H. Moulin, A. D. Procaccia, N. Shah, and J. Wang. The unreasonable fairness of maximum Nash product. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*, 2016. Forthcoming.

[44] I. Caragiannis, S. Nath, A. D. Procaccia, and N. Shah. Subset selection via implicit utilitarian voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 151–157, 2016.

[45] I. Caragiannis, A. D. Procaccia, and N. Shah. When do noisy votes reveal the truth? *ACM Transactions on Economics and Computation*, 4(3): article 15, 2016.

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

[47] D. Chakrabarty and C. Swamy. Welfare maximization and truthfulness in mechanism design with ordinal preferences. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science (ITCS)*, pages 105–120, 2014.

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

[49] A. Charnes and W. W. Cooper. Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9(3–4):181–186, 1962.

[50] J. Chen and D. Cottrell. Evaluating partisan gains from congressional gerrymandering: Using computer simulations to estimate the effect of gerrymandering in the us house. *Electoral Studies*, 44:329–340, 2016.

[51] J. Chen and J. Rodden. Cutting through the thicket: Redistricting simulations and the detection of partisan gerrymanders. *Election Law Journal*, 14(4):331–345, 2015.

[52] Y. Chen, A. Ghosh, M. Kearns, T. Roughgarden, and J. Wortman Vaughan. Mathematical foundations for social computing. *Communications of the ACM*, 59(12):102–108, 2016.

[53] M. Chikina, A. Frieze, and W. Pegden. Assessing significance in a markov chain without mixing. *Proceedings of the National Academy of Sciences*, 114(11):2860–2864, 2017.

[54] V. Conitzer and T. Sandholm. Common voting rules as maximum likelihood estimators. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 145–152, 2005.

[55] V. Conitzer, A. Davenport, and H. Kalagnanam. Improved bounds for computing Kemeny rankings. In *Proceedings of the 21st AAAI Conference on Artificial Intelligence (AAAI)*, pages 620–626, 2006.

[56] V. Conitzer, M. Rognlie, and L. Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 109–115, 2009.

[57] V. Conitzer, R. Freeman, and N. Shah. Fair public decision making. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*, pages 629–646, 2017.

[58] G. B. Dantzig. Discrete-variable extremum problems. *Operations Research*, 5(2):266–288, 1957.

[59] A. Davenport and J. Kalagnanam. A computational study of the Kemeny rule for preference aggregation. In *Proceedings of the 19th AAAI Conference on Artificial Intelligence (AAAI)*, pages 697–702, 2004.

[60] X. Deng, Q. Qi, and A. Saberi. Algorithmic solutions for envy-free cake cutting. *Operations Research*, 60(6):1461–1476, 2012.

[61] P. Diaconis and R. L. Graham. Spearman's footrule as a measure of disarray. *Journal of the Royal Statistical Society: Series B*, 39(2):262–268, 1977.

[62] M. Duchin. Gerrymandering metrics: How to measure? what's the baseline? *arXiv preprint arXiv:1801.02064*, 2018.

[63] M. Duchin, T. Gladkova, E. Henninger-Voss, B. Klingensmith, H. Newman, and H. Wheelen. Locating the representational baseline: Republicans in massachusetts. *arXiv preprint arXiv:1810.09051*, 2018.

[64] E. Elkind and N. Shah. Electing the most probable without eliminating the irrational: Voting over intransitive domains. In *Proceedings of the 30th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 182–191, 2014.

[65] E. Elkind, P. Faliszewski, and A. Slinko. Good rationalizations of voting rules. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*, pages 774–779, 2010.

[66] B. Fain, K. Munagala, and N. Shah. Fair allocation of indivisible public goods. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*, 2018. Forthcoming.

[67] M. Feldman, A. Fiat, and I. Golumb. On voting and facility location. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*, 2016. Forthcoming.

[68] B. Fifield, M. Higgins, K. Imai, and A. Tarr. A new automated redistricting simulator using markov chain monte carlo. *Work. Pap., Princeton Univ., Princeton, NJ*, 2015.

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

[70] D. Foley. Resource allocation and the public sector. *Yale Economics Essays*, 7:45–98, 1967.

[71] Y. Gal, M. Mash, A. D. Procaccia, and Y. Zick. Which is the fairest (rent division) of them all? In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*, pages 67–84, 2016.

[72] R. S. Garfinkel and G. L. Nemhauser. Optimal political districting by implicit enumeration techniques. *Management Science*, 16(8):B–495, 1970.

[73] B. Gärtner. Fast and robust smallest enclosing balls. In *European Symposium on Algorithms*, pages 325–338, 1999.

[74] Metric Geometry and Gerrymandering Group. Comparison of districting plans for the virginia house of delegates. Technical report, 2018.

[75] A. Goel, A. K. Krishnaswamy, S. Sakshuwong, and T. Aitamurto. Knapsack voting for participatory budgeting. Manuscript, 2016.

[76] A. Goel, A. K. Krishnaswamy, and K. Munagala. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*, pages 287–304, 2017.

[77] S. Gross, E. Anshelevich, and L. Xia. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 544–550, 2017.

[78] B. Gutiérrez. Madrid as a democracy lab, 2017. URL `https://www.opendemocracy.net/democraciaabierta/bernardo-guti-rrez/madrid-as-democracy-lab`. Accessed: 2018-05-02.

[79] J. Hall and W. H. Watson. The effects of a normative intervention on group decision-making performance. *Human Relations*, 23(4):299–317, 1970.

[80] C. Harris. A scientific method of districting. *Behavioral Science*, 9(3): 219–225, 1964.

[81] J. He, C.-A. Psomas, A. D. Procaccia, and D. Zeng. Achieving a fairer future by changing the past. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 2019.

[82] G. Herschlag, R. Ravier, and J. Mattingly. Evaluating partisan gerrymandering in wisconsin. *arXiv preprint arXiv:1709.01596*, 2017.

[83] V. E. Johnson. Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences*, 110(48):19313–19317, 2013.

[84] I. Kash, A. D. Procaccia, and N. Shah. No agent left behind: Dynamic fair division of multiple resources. *Journal of Artificial Intelligence Research*, 51:579–603, 2014.

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

[86] Z. Landau, O. Reid, and I. Yershov. A fair division solution to the problem of redistricting. *Social Choice and Welfare*, 32(3):479–492, 2009.

[87] J. Legendre, H Madénian, and P. L. Scully. Participatory budgeting in Paris, France, 2017. URL `http://participedia.net/en/cases/participatory-budgeting-paris-france`. Accessed: 2018-05-02.

[88] Z. Li, R. Wang, and Y. Wang. A quadratic programming model for political districting problem. In *Proceedings of the first international symposium on optimization and system biology (OSB)*, 2007.

[89] R. J. Lipton, E. Markakis, E. Mossel, and A. Saberi. On approximately fair allocations of indivisible goods. In *Proceedings of the 6th ACM Conference on Economics and Computation (EC)*, pages 125–131, 2004.

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

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

[92] C. L. Mallows. Non-null ranking models. *Biometrika*, 44:114–130, 1957.

[93] A. Mao, A. D. Procaccia, and Y. Chen. Better human computation through principled voting. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1142–1148, 2013.

[94] A. Mehrotra, E. L. Johnson, and G. L. Nemhauser. An optimization based heuristic for political districting. *Management Science*, 44(8): 1100–1114, 1998.

[95] Holmes E Miller, William P Pierskalla, and Gustave J Rath. Nurse scheduling using mathematical programming. *Operations Research*, 24 (5):857–870, 1976.

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

[97] H. Moulin. *Fair Division and Collective Welfare*. MIT Press, 2003.

[98] J. Nagle. What criteria should be used for redistricting reform? *Election Law Journal: Rules, Politics, and Policy*, 18(1):63–77, 2019.

[99] J. Oehrlein and J. Haunert. A cutting-plane method for contiguity-constrained spatial aggregation. *Journal of Spatial Information Science*, 2017(15):89–120, 2017.

[100] D. C. Parkes, A. D. Procaccia, and N. Shah. Beyond Dominant Resource Fairness: Extensions, limitations, and indivisibilities. *ACM Transactions on Economics and Computation*, 3(1): article 3, 2015.

[101] W. Pegden, A. D. Procaccia, and D. Yu. A partisan districting protocol with provably nonpartisan outcomes. *arXiv preprint arXiv:1710.08781*, 2017.

[102] D. Polsby and R. Popper. The third criterion: Compactness as a procedural safeguard against partisan gerrymandering. *Yale Law & Policy Rev.*, 9:301, 1991.

[103] G. Pritchard and M. Wilson. Asymptotics of the minimum manipulating coalition size for positional voting rules under impartial culture behaviour. *Mathematical Social Sciences*, 58(1):35–57.

[104] A. D. Procaccia. Cake cutting algorithms. In F. Brandt, V. Conitzer, U. Endress, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 13. Cambridge University Press, 2016.

[105] A. D. Procaccia. Science can restore America's faith in democracy, 2016. URL `https://www.wired.com/2016/12/science-can-restore-americas-faith-democracy/`.

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

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

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

[109] A. D. Procaccia, N. Shah, and Y. Zick. Voting rules as error-correcting codes. *Artificial Intelligence*, 231:1–16, 2016.

[110] C.-A. Psomas and D. Zeng. Fairness-efficiency tradeoffs in dynamic fair division. *Unpublished manuscript.*, 2019.

[111] P. Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *Journal of Computer and System Sciences*, 37(2):130–143, 1988.

[112] M. Rauterberg. A method of a quantitative measurement of cognitive complexity. In G. van der Veer, M. Tauber, S. Bagnara, and M. Antalovits, editors, *Human-Computer Interaction: Tasks and Organisation*, pages 295–307. CUD, 1992.

[113] E. Reock. A note: Measuring compactness as a requirement of legislative apportionment. *Midwest Journal of Political Science*, 5(1):70–74, 1961.

[114] P. Sanders. On the competitive analysis of randomized static load balancing. In *Proceedings of the 1st Workshop on Randomized Parallel Algorithms (RANDOM)*, 1996.

[115] J. Schwartzberg. Reapportionment, gerrymanders, and the notion of compactness. *Minn. L. Rev.*, 50:443, 1965.

[116] A. Shah. *Participatory Budgeting.* The World Bank, 2007.

[117] T. Shirabe. A model of contiguity for spatial unit allocation. *Geographical Analysis*, 37(1):2–16, 2005.

[118] T. Shirabe. Districting modeling with exact contiguity constraints. *Environment and Planning B: Planning and Design*, 36(6):1053–1066, 2009.

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

[120] J. Spencer. Balancing games. *Journal of Combinatorial Theory Series B*, 23(1):68–74, 1977.

[121] N. O. Stephanopoulos and E. M. McGhee. Partisan gerrymandering and the efficiency gap. *University of Chicago Law Review*, 82:831, 2015.

[122] W. Stromquist. How to cut a cake fairly. *American Mathematical Monthly*, 87(8):640–644, 1980.

[123] F. E. Su. Rental harmony: Sperner's lemma in fair division. *American Mathematical Monthly*, 106(10):930–942, 1999.

[124] E. Tisdale. The gerry-mander. *The Boston Gazette*, March 26, 1812.

[125] I. Tsetlin, M. Regenwetter, and B. Grofman. The impartial culture maximizes the probability of majority cycles. *Social Choice and Welfare*, 21:387–398, 2003.

[126] Edward R Tufte. The relationship between seats and votes in two-party systems. *American Political Science Review*, 67(2):540–554, 1973.

[127] H. Varian. Equity, envy and efficiency. *Journal of Economic Theory*, 9:63–91, 1974.

[128] T. Walsh. Online cake cutting. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*, pages 292–305, 2011.

[129] G. Warrington. Quantifying gerrymandering using the vote distribution. *Election Law Journal*, 17(1):39–57, 2018.

[130] S. J. Wilson. Fair division using linear programming. Working paper, Department of Mathematics, Iowa State University, 1998.

[131] L. Xia. Bayesian estimators as voting rules. In *Proceedings of the 32nd Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, 2016.

[132] L. Xia and V. Conitzer. A maximum likelihood approach towards aggregating partial orders. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 446–451, 2011.

[133] L. Xia, V. Conitzer, and J. Lang. Aggregating preferences in multi-issue domains by using maximum likelihood estimators. In *Proceedings of the 9th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 399–408, 2010.

[134] A. C. Yao. Probabilistic computations: Towards a unified measure of complexity. In *Proceedings of the 17th Symposium on Foundations of Computer Science (FOCS)*, pages 222–227, 1977.

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