Algorithms for Voting and Group Selection

A dissertation presented

by

Gregory Kehne

to

The Department of Computer Science

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Computer Science

Harvard University

Cambridge, Massachusetts

August 2023

Dissertation Advisor: Ariel Procaccia Gregory Kehne

Algorithms for Voting and Group Selection

Abstract

Many of the core challenges in social choice, new and old, stand to benefit from consideration through the computational lens. We demonstrate the particular applicability of approximation algorithms to the tasks of aggregating judgements and preferences for choosing one or more alternatives, and to the construction of representative committees and valuable teams.

The first part of this thesis presents contributions to the theory of judgement and preference aggregation in settings where only incomplete information is known. It considers the task of combining expert advice given only incomplete information about experts' competence in order to reach a maximally accurate judgement. For preferences, we consider two settings of choosing alternative(s) given only incomplete information about voters' preferences. In the first we aim to choose a single alternative when only the ranking of voters' valuations of alternatives is known. In the second we aim to select a committee of candidates that all coalitions within the electorate find at least partially acceptable, under the restriction that each voter is only evaluates a few candidates.

The second part studies sampling processes which arise in sortition, a direct democratic paradigm in which representative subsets (or panels) of a population are randomly chosen in order to represent the whole. Here the goal of choosing the panel at random is in tension with choosing a maximally representative panel; we present a new framework and approach to making this tradeoff, and further address the challenge of satisfying these aims via a process which is also readily comprehensible to observers.

The third part provides algorithmic approaches to group selection problems motivated by hiring, admissions, and team formation. We consider the task of making a collection of

simultaneous invitations to candidates who, in contrast to the previous committee selection problems, may decline to be included in the group with some probability. We present approaches to eliciting information from participants and extending sets of invitations in order to realize a group of favorable participants which is close to the target size.

Table of Contents

List of Figures

Acknowledgments

My deepest thanks to Ariel, for first teaching a captivating course on this work and then offering me the chance to try my hand at it. Thank you for your sage advice, generosity, patience, and humor; and moreover for an authentic demonstration of what makes this pursuit worthwhile. The skill, cheer, and apparent ease with which you conduct research and provide mentorship is admirable bordering on uncanny, and I only hope to untangle these mysteries in the years to come. I am also grateful for the group you head and of which I was glad to be a part. Thank you Alex, Nika, Dominik, Gerdus, Ritesh, Anson, Paul, Bailey, Dan, Jamie, Manuel, Gili, and Shirley, for your encouragement, camaraderie, and shared joy through social outings, games, and whiteboard sessions alike.

My PhD traversed two departments and institutions and a rather tumultuous time. I am fortunate to have completed this second leg with the EconCS group at Harvard, who welcomed me through lockdown from my living room in Pittsburgh and proved an even more amazing research community to join in person. My particular thanks to Yannai and Yiling for your conversation, engagement, and service on this thesis committee. I am also fortunate to have started in the CMU ACO program, and owe thanks to its administrators, organizers, and members; particularly among them Anupam, Ravi, Boris, Wes, and Ariel for your invaluable mentorship and instruction.

A huge and heartfelt thanks to all of my collaborators: Soroush Ebadian, Bailey Flanigan, Kobi Gal, Yannai Gonczarowski, Anupam Gupta, Daniel Halpern, Ellis Hershkowitz, Nicholas Hoernle, Anson Kahng, Roie Levin, Evi Micha, Dominik Peters, Ariel Procaccia, R. Ravi, Ben Schiffer, Nisarg Shah, Piotr Skowron, Jamie Tucker-Foltz, Jingyan Wang, Manuel Wüthrich, and Shirley Zhang; It has been an ongoing privilege to work with and learn from you, and to share in all the excitements of the research process.

I could not have found such joy in, let alone survived, this chapter without the love and support of my friends. I am grateful to have met and befriended some truly excellent people and cadres in grad school, among them Roie and Ellis, colleagues, confidantes, and cohabitators par excellence; CS pals Aria, Raj, Ainesh, Zoe, Pedro, Emily, Anson, Paul, and Bailey; and Ellen, David, Naama, Vijay, Goran, Nic, and Laxman; the Philosophers Katie, Viv, Will, Nedah, Gal, and Dasha; my math cohort (and honorary cohort) Matt, Anish, Junichi, Landon, Kayla, and Kerrek; Joe, Chris, Clive, Andy, and Antoine, math and human exemplars; and the big-hearted Cantabs Lily, Sohini, Dan, Jamie, Alex, Andrew, and Toly.

Thanks also to teammates and friends from prior chapters who have seen me through this one: Ephs especially including Matt, Matt, Greg, Cole, and Bijan; and my homies Kiernan, Dominic, and Gabe. I also count myself lucky to have stumbled across the motely crew of big-hearted cyclists that comprise Boston Bike Polo, a sport and community that have each made lasting impressions. Other entities deserve acknowledgment—particular thanks to the Dolly Sods Wilderness, the city of Pittsburgh, and my Trek 950.

Thank you to Marina and Grendel, my companions through this journey and onward. Crossing and joining paths has been an unanticipated joy, and I can't wait to see what the future has in store.

Finally, to my family: a special, existential sort of thanks for your love and your nourishment, your encouragement and support, for especially the best parts of who I am.

For my grandparents.

Chapter 1

Introduction

The scope of this thesis falls primarily within the field of computational social choice, and secondarily within the intersection of computer science and economics more broadly. It is motivated by deep interest in approximation algorithms, and by the belief that they form a toolkit which is at once theoretically rich and profoundly applicable to problems of practical concern. Decision-making systems in social choice contexts often incompletely incorporate algorithmic techniques, or can dramatically benefit from their introduction. Furthermore, as has been the case throughout the history of theoretical computer science, these application domains in turn yield problems which are worthwhile objects of study in their own right.

Computational social choice is an interdisciplinary field which includes the study of algorithms for preference and judgment aggregation (among others), with a vibrant collection of established models, methods, and results. Furthermore, democratic innovations and new deliberative practices are constantly emerging, and their effective development and practice entails theoretical work which draws from and adds to this field. Some of these canonical topics have made the leap to online settings; for instance betting markets like PredictIt and Metaculus may be viewed as dynamic aggregators of collective beliefs about future events. However we do not yet have a systematic understanding of online and dynamic approaches for preference aggregation, despite their rapidly increasing popularity in practice. Many nascent online communities are actively exploring formal structures for governance which are asynchronous, and aim for both responsiveness to their members and stability over time. On the other hand, online platforms such as Polis and Consider.it cater to established communities and municipalities. These platforms, like other (re)emerging direct democratic paradigms, provide tools to aggregate public beliefs and preferences and foster structured deliberation, but lack theoretical grounding.

This thesis presents several contributions to this field, and Chapter 4 and Part II in particular provide the beginnings of, and perhaps even a basis for the answer to this imperative. In what follows we will provide a brief introduction to each of the three main areas of this thesis, and a brief description of the motivation and setting of each of the constituent chapters, along with an overview of its major findings. We will conclude the overview of each part with a higher-level discussion of the questions and results and their contributions towards answering these motivating challenges.

1.1 Overview

Part I: Preference and Judgment Aggregation

Computational social choice concerns the computational aspects of tasks that arise in collective decision making. This study generally entails designing algorithms (mechanisms) which interact with agents who possess beliefs and preferences, with the aim of achieving some desired outcome. This describes the constituent subjects of voting theory, fair allocation, and judgment aggregation, among others [Bra+16]. Many problems may be cast as refinements of the question "How can we attain socially desirable outcomes, given constraints on what we can practically elicit from individuals?" For instance, in the implicit utilitarian model, voting rules work with rankings (or approval votes) rather than voters' utilities, in part because the former are much easier to reliably elicit; proxy voting, where voters either delegate to other voters [Mil69] or to learned models of their preferences [Noo+18] fit this paradigm as well.

Chapter 2: Aggregating Binary Judgments Ranked by Accuracy

One of the earliest results in computational social choice is the Condorcet jury theorem [Con85], which pertains to the inference problem of learning a ground truth based on independent and noisy jurors. More generally, one can assume that the independent experts report this unknown ground truth with known probabilities; a seminal result establishes that a log-odds scoring rule is optimal for inferring the ground truth [NP82]. If these probabilities are unknown, however, it is unclear how to proceed. Past work has studied levering auxiliary information [Bah+11; Bah+12], when it exists.

This chapter charts a different approach. We instead assume that the judgment aggregator knows the relative ranking of the accuracies of the experts, though not their exact accuracies. We draw an analogy to the information environment of implicit utilitarian voting, which motivates our choice of optimistic and pessimistic objectives with respect to the unobserved expert accuracies. We then derive natural, interesting, and efficiently computable rules with respect to these objectives. (In a nod to their efficacy in the known-probability setting, we also identify optimal scoring rules.) Testing these rules on collaborative filtering problems supports our premise: that assuming restricted information and then optimizing with respect to optimistic and pessimistic assumptions is a successful strategy for deriving robust judgment aggregation algorithms.

Chapter 2 is based on [Hal+21], which is joint work with Daniel Halpern, Dominik Peters, Ariel D. Procaccia, Nisarg Shah, and Piotr Skowron.

Chapter 3: Worst-Case Voting when the Stakes are High

The implicit utilitarian model for voting has been a successful basis for evaluating established social choice functions and developing new approaches [Ans+21a]. Voters are assumed to have utilities over candidates in an election which encode their subjective preferences, but report only ordinal preferences; the aim is to choose an outcome which maximizes aggregate utility (welfare), generally in the worst case. This chapter presents a careful analysis of the welfare regret attainable by voting rules, both in and beyond the worst case. Prior work has overwhelmingly pursued multiplicative approximations to the maximum-utility outcome; Caragiannis et al. [Car+17] introduced implicit utilitarian regret in the committee election setting, but did not treat the single-winner case.

One possible advantage of considering welfare regret is that rules minimizing it are less sensitive to cases when there is little aggregate utility to be conferred. Another is that it admits meaningful guarantees in a broader class of utilities than utilitarian voting generally considers [Azi19]. We define a randomized analog of positional scoring rules, and present a rule which is asymptotically optimal within this class as the number of alternatives increases. We then show that the instance-optimal social choice function can be efficiently computed. Next, we take a beyond-worst-case view, bounding the additive distortion of prominent voting rules as a function of the best welfare attainable in an instance. Lastly, we evaluate the additive distortion of a range of rules on real-world election data.

Chapter 3 is based on [KK22], which is joint work with Anson Kahng.

Chapter 4: Proportional Representation with Incomplete Votes

In this chapter we trade the implicit utilitarian model for voting with rankings for a setting where voters cast approval votes, but the preferences which may be elicited from any one voter are severely limited. In particular, we aim to select a committee of candidates that all coalitions within the electorate find at least partially acceptable, under the restriction that each voter is only evaluates a few candidates.

This work is motivated by platforms for online civic participation, which rely heavily on methods for condensing thousands of comments into a relevant handful, based on whether participants agree or disagree with them. These methods should guarantee fair representation of the participants, as their outcomes may affect the health of the conversation and inform impactful downstream decisions.

To that end, we draw on the literature on approval-based committee elections. Our setting is novel in that the approval votes are incomplete since participants will typically not vote on all comments. We prove that this complication renders non-adaptive algorithms impractical in terms of the amount of information they must gather. Therefore, we develop an adaptive algorithm that uses information more efficiently by presenting incoming participants with statements that appear promising based on votes by previous participants. We prove that this method satisfies commonly used notions of fair representation, even when participants only vote on a small fraction of comments. Finally, an empirical evaluation using real data shows that the proposed algorithm provides representative outcomes in practice.

Chapter 4 is based on [Hal+22], which is joint work with Daniel Halpern, Ariel D. Procaccia, Jamie Tucker-Foltz, and Manuel Wüthrich.

These first chapters forward the proposed aims of rendering methods from computational social choice more scalable and applicable in several ways. First, Chapter 2 deals with an epistemological problem; that of choosing a ground truth from amongst a finite set of alternatives, based on fallible inputs. The setting matches that of the Condorcet jury theorem but generalizes it in two important respects. The first is in the quality of the information accessible regarding the inputs, which are the experts' predictions. In the classical setting these probabilities are assumed to be known exactly, while in practical settings we should expect these to be difficult to estimate or elicit to a high degree of precision. If predictions of experts' accuracies are themselves derived from a process which does not produce scalar accuracies but instead claims about the experts' *relative* accuracies (such as a social choice function or voting rule considered in Chapter 3), then this ordinal information setting is more apt. Second, the set of objectives we use in order to determine an aggregate prediction based on experts' input is more general. The original result assumes the form of the distribution of experts' predictions and asks under which assignment of alternatives to 'true' outcomes the observed experts' predictions are most likely. By contrast the results here incorporate uncertainty about the experts' accuracies and therefore the form of the distribution. By considering both distributional and worst case views over this uncertainty in the production generating process, we produce algorithms which can incorporate incorporate finer-grained predictions when they're available on the one hand, and operate in ways which are more robust to this additional uncertainty in the problem description on the other. Finally, for a maximum likelihood estimation objective which generalizes the original jury theorem, we present algorithms which can made to run in time linear in the number of experts, up to polylogarithmic factors. This ensures the applicability of our approach for settings where the number of experts is potentially quite large, such as in ensemble learning applications.

The voting rules we introduce in Chapter 3 also build upon a (slightly less) classical setting of social choice and render it more robust to practice and applicable at scale. As discussed above, in this setting we adopt the implicit utiliitarian framework for voting and adopt a worst-case lens by which to evaluate social choice functions. As compared to multiplicative guarantees, our regret-based objective is differentially concerned with instances where only one or a few alternatives would confer large aggregate utility (welfare) to the voting body, a setting which we argue frequently arises in election and ranked-preference-aggregation settings of practical interest. Our assumptions on implicit utilities accommodate a broader class of possible voter values, while the assumption that each voter's utilities over alternatives sum to at most one may be viewed as the condition that no one voter can exert disproportionate influence over the objective. Finally, the social choice functions which we derive are well-suited to large-scale applications. On the one hand our instance-optimal rule may be viewed as an LP, but optimal solutions exhibit simple structure which admit efficient calculation. On the other, the randomized positional scoring rule we derive requires voters to provide a ranking of only their top three candidates. This makes it especially well-suited to settings when the number of candidates is large.

Finally, the work presented in Chapter 4 is not only applicable to but motivated directly by internet-based, deliberative, and large-scale settings. It ports existing approval-based multiwinner voting objectives and rules over from the offline setting where users are expected to provide full preferences over alternatives to an online adaptive setting where it is profoundly unrealistic to ask voters to provide full input over alternatives. Demonstrating the feasibility of this transition opens up numerous future directions which are discussed in more detail in Chapter 4 and Chapter 9.

Part II: Algorithms for Sortition

Sortition is a form of deliberative democracy in which a subset of a population is chosen to serve as a recommendation- or decision-making proxy for the whole. Crucially, it entails randomly selecting this subset, or panel. This randomness induces some probability that each individual in the population is chosen, which is generally seen as conferring fairness to the individuals and legitimacy to the process. These selection probabilities may be viewed as an *ex ante* individual guarantee, and the equity of these probabilities determines the selection algorithm's fairness to participants.

Selection algorithms also aim to satisfy an *ex post* group representation guarantee; informally, a requirement that the chosen panel resemble the overall population. In practice this has taken the form of descriptive representation, wherein the composition of the chosen panel must satisfy a fixed set of intersecting demographic quotas. Prior work has designed selection algorithms which produce maximally equitable individual selection probabilities, subject to such quotas [Fla+21].

Chapter 5: Fair Sortition Made Transparent

This chapter studies randomness as a design parameter of selection algorithms, alongside the competing desiderata of individual fairness and descriptive representation.The algorithms for sortition which have been recently developed and deployed select assemblies *maximally fairly*, meaning that subject to demographic quotas, they give all potential participants as equal a chance as possible of being chosen. While these fairness gains can bolster the legitimacy of citizens' assemblies and facilitate their uptake, existing algorithms remain limited by their lack of transparency.

To overcome this hurdle, in this work we focus on panel selection by uniform lottery, which is easy to realize in an observable way. By this approach, the final assembly is selected by uniformly sampling some pre-selected set of m possible assemblies. We provide theoretical guarantees on the fairness attainable via this type of uniform lottery, as compared to the existing maximally fair but opaque algorithms, for two different fairness objectives. We complement these results with experiments on real-world instances that demonstrate the viability of the uniform lottery approach as a method of selecting assemblies both fairly and transparently.

Chapter 5 is based on [FKP21], which is joint work with Bailey Flanigan and Ariel D. Procaccia.

Chapter 6: Sortition via Multi-Objective Optimization

Two of the key arguments in favor of sortition are that it provides *representation* (a random panel reflects the composition of the population) and *fairness* (everyone has a chance to participate). The model in prior work and Chapter 5 formulates representation criteria as constraints, and maximizes fairness criteria subject to them.

This chapter develops an alternative theoretical model of sortition, in which descriptive representation is described by a continuous objective rather than a fixed set of constraints. Our measure is (informally) characterized by how similar a given proportion of a panel is to each individual in the population. It is incomparable to but arguably richer than the traditional constraint-based formalization, in part because it accommodates a nontrivial notion of the *ex ante* representativeness of a selection algorithm.

We investigate the extent to which the two competing objectives of selection fairness and representativeness can be simultaneously maximized. Towards answering this question, we introduce the notion of a representation metric on the space of individuals, and assume that the cost of an individual for a panel is determined by the q-th closest representative; the representation of a (random) panel is measured by the ratio between the (expected) sum of costs of the optimal panel for the individuals and that of the given panel.

We find markedly different behavior in both theory and practice depending on whether or not q includes a majority of the panel. For $k/2 < q \leqslant k - \Omega(k)$, where k is the panel size, we show that uniform random selection is indeed representative by establishing a constant lower bound on this ratio. By contrast, for $q \leqslant k/2$, no random selection algorithm that is almost fair can give such a guarantee. We therefore consider relaxed fairness guarantees, and develop a new random selection algorithms that shed light on this tradeoff between representation and fairness.

Chapter 6 is based on [Eba+22b], which is joint work with Soroush Ebadian, Evi Micha, Ariel D. Procaccia, and Nisarg Shah.

The relevance of this second collection of work to improving the scalability and applicability of deliberative and innovative forms of social choice is straightforward. The resurgence of sortition is a paradigmatic instance: it has experienced a modern resurgence which has precipitated and been helped along by substantive improvements in the theory and algorithms which enable its practice. Prominent among these efforts is the constrained optimization framework identified and addressed by Flanigan et al. [Fla+21], upon which the work in Chapter 5 directly builds. This work was inspired by the method pursued by organizers of the Michigan Citizens' Panel on COVID-19, who together with the authors of [Fla+21] orchestrated a drawing for a citizens' assembly in a manner corresponding to the notion of a quantized panel distribution in this chapter. (For a more detailed description, see Appendix A of [Of 20]). The beyond-worst-case results in this chapter are indexed to various measures of the complexity of the representational quotas, and serve to establish that the benefits to be had from this latest generation of sortition algorithms need not come at the cost of inscrutability in the process of their execution. Especially in light of the fact that these panel selection algorithms necessarily employ randomness as part of a potentially high-stakes political process, we are optimistic that these results can positively influence the practice of sortition moving forward and its adoption into more collective decision-making contexts.

By contrast the family of formulations of sortition proposed and studied in Chapter 6 is more speculative and less closely informed by current practice. It aims to contribute to the goal of broadening the applicability of sortition as a decision-making paradigm, with the desired outcome of ultimately increasing the scope of its practice. In its abstract form, the proposed formulation is to view the competing desiderata of equity in individual selection probabilities on the one hand and representation of underlying group on the other as two distinct objectives. If practitioners commit to a fairness notion which expresses equity in individual selection probabilities and a metric for measuring the extent to which a panel is representative (or misrepresentative) of the underlying population on the other, then the algorithmic question becomes, to what extent can a panel selection algorithm attain both of these goals? What combinations of objective values are attainable? Ideally this question could even be answered on an instance-by-instance basis. This is general enough to capture the more practical model used in Chapter 5; it corresponds to choosing a misrepresentation measure of the extent to which demographic constraints are violated, and then picking one end of the resulting Pareto frontier. Indeed, in follow-up work Ebadian and Micha [EM23] study this framework with a core-based representation metric. The dominant paradigm for sortition takes representation to mean the intersection of individual quotas of categorical feature attributes. By introducing this reformulation, we hope to extend the scope of the practice to settings where direct democracy by lot is desirable, but where representation takes other, even instance- or subject-specific forms.

Part III: Group Selection

In contrast to the representation-based criteria for committee selection in Chapter 4 and the subset selection tasks which underpin sortition in Part II, there are settings in which the objectives for group evaluation are relatively more straightforward, but the process for selecting groups is more complex. These selection problems are motivated by hiring, admissions, and team formation. We consider the task of making a collection of simultaneous invitations to candidates who, in contrast to the previous settings, may decline to be included in the group with some probability. For instance in academic recruitment settings such as faculty hiring and PhD admissions, committees might aim to maximize the overall quality of recruited candidates, but there is uncertainty about whether any given candidate would accept an offer if presented with one.

As such, the models and algorithms in this part may be viewed as theoretical contributions to problems which arise from the perspective of a participants in a labor market without central planning or coordination, in which the pursuit of global properties like efficiency or stability is inapt. By considering simple settings and setting aside matters of strategic competition between agents or hiring entities, these chapters demonstrate that the algorithmic task of reducing uncertainty alongside other rewards can become technically involved.

Chapter 7: Strategyproof Mean Estimation from Multiple-Choice Questions

Given n values possessed by n agents, we study the problem of estimating the mean by truthfully eliciting agents' answers to multiple-choice questions about their values. In particular, in a group selection or hiring context, we view these values as probabilities of offer acceptance. Our task is then to elicit a maximally accurate estimate of the number of offers which will be accepted, so as to plan for a larger selected group or augment the selection with a second round of offers.

We consider two natural candidates for estimation error: mean squared error (MSE) and mean absolute error (MAE). We design a randomized estimator which is asymptotically optimal for both measures in the worst case. In the case where prior distributions over the agents' values are known, we give an optimal, polynomial-time algorithm for MSE, and show that the task of computing an optimal estimate for MAE is $\#\mathcal{P}$-hard. Finally, we demonstrate empirically that knowledge of prior distributions gives a significant edge.

Chapter 7 is based on [KKP20], which is joint work with Anson Kahng and Ariel D. Procaccia.

Chapter 8: Recruitment Strategies that Take a Chance

Previous theoretical work group selection with uncertain offer acceptance has considered algorithms that make offers sequentially and are subject to a hard budget constraint. In this chapter, we argue that these modeling choices may be inconsistent with the practice of academic recruitment. Instead, we restrict ourselves to a single batch of offers, and we treat the target number of positions as a soft constraint, so we risk overshooting or undershooting the target.

Specifically, our objective is to select a subset of candidates that maximizes the overall expected value associated with candidates who accept, minus an expected penalty for deviating from the target. We first analyze the guarantees provided by natural greedy heuristics, showing their desirable properties despite the simplicity. Depending on the structure of the penalty function, we further develop algorithms that provide fully polynomial-time approximation schemes and constant-factor approximations to this objective. Empirical evaluation of our algorithms corroborates these theoretical results.

Chapter 8 is based on [KPW22], which is joint work with Ariel D. Procaccia and Jingyan Wang.

In contrast to the prior two parts, the work comprising this third section is more directly related to problems arising at the intersection of economics and computer science than those arising within computational social choice proper. What is common, particularly to the contributions to computational social choice in Part I, is the success of the approximation algorithms toolkit to providing average-case and worst-case guarantees. The algorithms presented here, for reducing uncertainty by eliciting truthful answers to multiple-choice questions and greedy algorithms for stochastic batch-hiring/knapsack problems, are equipped with both theoretical guarantees and exceptionally simple structure, which we anticipate could serve to improve their comprehensibility and adaptability in practical instances and applications.

Furthermore, there appears to be potential for the truthful probability elicitation and variance reduction approaches in Part III to be brought to bear to tasks in social choice. We discuss one such potential application to sortition in Chapter 9.

Part I

Preference and Judgment Aggregation

Chapter 2

Aggregating Binary Judgments Ranked by Accuracy

2.1 Introduction

Consider the task of predicting a binary ground truth $G \in \{0,1\}$ by aggregating independent binary judgments provided by n experts. This models a wide range of real-world scenarios, where the judgments can be polls predicting the outcome of an upcoming political or sports event, weather forecasts, or juror opinions of a defendant's guilt.

The judgment of expert i, denoted X_i, is assumed to be a Bernoulli random variable, which coincides with the ground truth with probability p_i; this probability is referred to as the *accuracy* of the expert. If $\boldsymbol{p} = (p_1, \ldots, p_n)$ is known, then the classical *maximum likelihood estimation* approach chooses the ground truth estimate that maximizes the likelihood of inducing the vector of expert judgments $\boldsymbol{X} = (X_1, \ldots, X_n)$, i.e., the value of $y \in \{0,1\}$ that maximizes $\mathcal{L}[\boldsymbol{X}|G = y, \boldsymbol{p}] = \prod_{i=1}^{n} p_i^{\mathbb{1}[X_i=y]} \cdot (1-p_i)^{\mathbb{1}[X_i \neq y]}$, where $\mathbb{1}$ is the indicator variable.

However, sometimes we may not know the exact values of $p_1, \ldots, p_n$; instead, we may only know a *ranking* of the expert judgments by accuracy. This may be the case when there is metadata available about the judgments that is known to be correlated with accuracy, but the exact nature of the correlation is not known. For instance, if a pollster conducts multiple polls over time, polls conducted closer to the date of the event being predicted may be considered more accurate than the ones conducted earlier; the same reasoning applies to weather forecasts. Similarly, polls conducted concurrently may be ranked by their sample sizes. Sometimes, experts may participate in a judgment contest (such as the Good Judgment Project[1]), which may show their ranking by accuracy on the leaderboard.

Motivated by such settings, we address the following question in this work:

> *How should we aggregate n binary judgments ranked by accuracy in order to predict a binary ground truth?*

Note that the n binary judgments ordered by accuracy can be represented as a bit-string of length n. Thus, we essentially study aggregation rules which take a bit-string as input and output a bit. Due to the fundamental nature of this setting, the rules designed in this work may have applications in other domains (see Section 8.6).

2.1.1 Our Contribution

Recall that the likelihood function $\mathcal{L}[\boldsymbol{X}|G = y, \boldsymbol{p}]$ depends on $p_1, \ldots, p_n$, i.e., on the accuracy of the experts. However, we are given only partial information about these values, namely, their ordering. To address this missing information, we take a worst-case viewpoint. Specifically, let $\mathcal{P}$ denote the set of all $\boldsymbol{p}$ which are consistent with the given ordering; we define the following three natural objectives that serve as proxies for the likelihood induced by a given estimate $y \in \{0, 1\}$, and design algorithms to compute the estimate optimizing these objectives.

1. *Distortion:* $\sup_{\boldsymbol{p}\in\mathcal{P}} \mathcal{L}[\boldsymbol{X}|G = 1 - y, \boldsymbol{p}]/\mathcal{L}[\boldsymbol{X}|G = y, \boldsymbol{p}]$. Note that this is worst-case ratio of the likelihood of the estimate not chosen $(1 - y)$ to the likelihood of the estimate chosen

[1] `https://goodjudgment.com/`

(y). Our aim is to minimize this objective.

2. *Optimistic likelihood:* $\sup_{\boldsymbol{p}\in\mathcal{P}} \mathcal{L}[\boldsymbol{X}|G=y,\boldsymbol{p}]$. Maximizing this objective can be thought of as a natural extension of the maximum likelihood approach, where we make an inference about $\boldsymbol{p}$ together with one about the estimate y.

3. *Pessimistic likelihood:* $\inf_{\boldsymbol{p}\in\mathcal{P}} \mathcal{L}[\boldsymbol{X}|G=y,\boldsymbol{p}]$. Maximizing this objective can be thought of as maximizing the worst-case likelihood.

In Section 2.3, we characterize the rules which optimize these objectives, and show that they can be implemented in polynomial time. In particular, the rules optimizing the first two objectives are novel and elegant. In Section 2.4, we restrict our attention to a natural family of rules, which we refer to as *scoring rules*. These rules assign monotonic weights to judgments (i.e. judgments ranked higher by accuracy receive no less weight than those ranked lower), and return the estimate with the highest total weight. We characterize the scoring rules that optimize the three aforementioned objectives among all scoring rules.

2.1.2 Related Work

Our paper contributes to a large body of work in computational social choice [Bra+16]. A central feature that separates our setting from the vast majority of papers in the area is that the judgments (or opinions, or preferences) that are being aggregated are typically assumed to be anonymous, in the sense that individuals are indistinguishable. However, it has been noted that there are important contexts where anonymity leads to bad outcomes [Lan19].

Our setting is related to judgment aggregation [End16], an area that also aggregates binary judgments. However, that literature focuses on problems arising from the aggregation of several logically related issues simultaneously, and does not typically assume a ground truth.

In statistics there is influential work on the problem of estimating the common mean of multiple normal distributions [CS74; JK96], where the unknown variance of each distribution can be seen as a measure of (in)accuracy. Our setting is more closely related to the work of [GKM11], who, like us, consider a binary ground truth (for each "item"), and binary judgments, each of which is correct with some probability that depends on the expert's unknown accuracy. The central idea that distinguishes our work from these papers is that we assume a known ranking of the experts by accuracy. This assumption also guides our choice of (worst-case) optimization objectives, which are different from the statistical estimation problems considered in previous work.

Some of our main results pertain to the distortion objective. This objective was conceived in the context of social-welfare maximization in voting settings [PR06; Bou+15; Car+17; Ans+18], but several paper have applied the idea to other problems such as matching, facility location, and even traveling salesperson [AS16; AA18; AZ18].

Our aggregation rules can be viewed as *simple games* [TZ99] where the experts are players, and winning coalitions correspond to sets of experts such that when all these experts report 1, then so does the aggregation rule. The simple games literature has also studied *linear* simple games, which correspond to games with ranked players. This literature includes characterization results for *weighted* simple games [TZ92], which correspond to what we call scoring rules.

2.2 Model

For $k \in \mathbb{N}$, let us denote $[k] = \{1, \ldots, k\}$. Let $G \in \{0, 1\}$ denote an unknown binary ground truth. Let $N = [n]$ denote a set of experts. Each expert $i \in N$ provides a binary judgment $X_i \in \{0, 1\}$, which is a Bernoulli random variable that is correct with probability p_i, i.e., $\Pr[X_i = G] = p_i$. We refer to $\boldsymbol{X} = (X_1, \ldots, X_n)$ as the *judgment profile* and $\boldsymbol{p} = (p_1, \ldots, p_n)$ as the *accuracy profile*.

In this work, we make two crucial assumptions regarding $\boldsymbol{X}$ and $\boldsymbol{p}$. First, we assume that the expert judgments (i.e. $X_1, \ldots, X_n$) are independent. Second, we assume that each expert is at least as accurate as a coin toss, i.e., $p_i \geqslant 1/2$ for each $i \in N$. For a discussion about relaxing these assumptions, see Section 8.6.

For $y \in \{0, 1\}$, the likelihood of observing $\boldsymbol{X}$ when the ground truth is $G = y$ can now be written as

$$\mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}] = \textstyle\prod_{i=1}^{n} p_i^{\mathbb{1}[X_i = y]} \cdot (1 - p_i)^{\mathbb{1}[X_i \neq y]},$$

where $\mathbb{1}$ denotes the indicator variable. If the accuracy profile $\boldsymbol{p}$ is known, then a classical approach to aggregating the expert judgments would be to return the *maximum likelihood estimate* (MLE) of the ground truth given by $\arg\max_{y \in \{0,1\}} \mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}]$.

In this work, we assume that we do not know $\boldsymbol{p}$. Instead, we are given a ranking of the experts by their accuracy, and we are interested in aggregating the expert judgments subject to this ordinal information. Without loss of generality, assume that $p_1 \geqslant p_2 \geqslant \ldots \geqslant p_n$. Thus, expert 1 is the most accurate, while expert n is the least accurate. Let $\mathcal{P}_n = \{\boldsymbol{p} : 1 \geqslant p_1 \geqslant \ldots \geqslant p_n \geqslant 1/2\}$ denote the set of feasible accuracy profiles. Note that $\mathcal{P}_n$ contains the accuracy profile $\boldsymbol{p} = (1, \ldots, 1)$, under which the likelihood of any non-unanimous judgment profile $\boldsymbol{X}$ is zero, regardless of the estimate y. This makes some of our objectives not well-defined or uninteresting. Of course, in practice, no judgment is perfectly accurate. To circumvent this hypothetical inconsistency, we define $\mathcal{P}_n^\epsilon = \{\boldsymbol{p} : 1 - \epsilon \geqslant p_1 \geqslant \ldots \geqslant p_n \geqslant 1/2\}$, analyze the aggregation rules optimizing our objectives defined with respect to $\mathcal{P}_n^\epsilon$, and then take the limit $\epsilon \to 0$. In the limit, these rules "converge", in the sense that they become fixed once ϵ is small enough. When the objective is well-defined directly with respect to $\mathcal{P}_n$, we avoid taking this longer route.

Formally, our input is the bit-string $\boldsymbol{X} \in \{0, 1\}^n$, where we refer to X_1 as the *most accurate bit* and X_n as the *least accurate*. An *aggregation function* is denoted $f : \{0, 1\}^n \to \{0, 1, \perp\}$, where $\perp$ denotes a tie.[2] We will alternatively represent a tie as the function returning $\{0, 1\}$ instead of $\perp$.

We are also interested in a natural family of aggregation functions that we refer to as *scoring rules*. A scoring rule $f_{\boldsymbol{w}}$ is parametrized by a weight vector $\boldsymbol{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n_{\geqslant 0}$, where w_i is the weight associated with the i-th most accurate bit. Given input $\boldsymbol{X}$, $f_{\boldsymbol{w}}$ returns the bit with the highest total weight, i.e., $\arg\max_{y \in \{0,1\}} \sum_{i=1}^{n} w_i \cdot \mathbb{1}[X_i = y]$. This definition is novel in our setting of binary judgments, but it is inspired by that of a prominent family of voting rules called positional scoring rules, which includes well-known rules such as plurality and Borda count.

2.3 Worst-Case Optimal Aggregation Rules

Given incomplete information about the accuracy profile $\boldsymbol{p}$, we cannot compute the MLE, since different accuracy profiles $\boldsymbol{p}$ consistent with the given ordinal information may induce different likelihoods. Our approach is to define an objective function that summarizes the likelihoods induced by all feasible $\boldsymbol{p}$ and optimize it; we consider three proposals.

2.3.1 Distortion for Binary Judgements

Informally, given an objective function and ordinal information about cardinal inputs to the function, the distortion approach selects an outcome minimizing the ratio between the optimal objective value and the objective value under the selected outcome, in the worst case over all cardinal inputs consistent with the given ordinal information. The objective we are interested in is the likelihood function $\mathcal{L}$, and we are given ordinal information about $\boldsymbol{p}$ (specifically, that $\boldsymbol{p} \in \mathcal{P}_n$). Given a judgment profile $\boldsymbol{X}$, the *distortion* of ground truth estimate $y \in \{0, 1\}$ is

[2] Allowing ties does not significantly alter most of our results; we discuss some of the implications of ties in later sections.

then defined as

$$\mathrm{dist}(y;\boldsymbol{X}) = \sup_{\boldsymbol{p}\in\mathcal{P}_n} \frac{\max\left(\mathcal{L}[\boldsymbol{X};G=0,\boldsymbol{p}],\mathcal{L}[\boldsymbol{X};G=1,\boldsymbol{p}]\right)}{\mathcal{L}[\boldsymbol{X};G=y,\boldsymbol{p}]} = \sup_{\boldsymbol{p}\in\mathcal{P}_n} \frac{\mathcal{L}[\boldsymbol{X};G=1-y,\boldsymbol{p}]}{\mathcal{L}[\boldsymbol{X};G=y,\boldsymbol{p}]}.$$

Here, the second equality is due to the fact that with $G=y$, the ratio is always 1, which can also be achieved with $G=1-y$ at $p_1=\ldots=p_n={}^1\!/_2$ (which makes the likelihoods given both possible ground truths equal). Hence, the worst case is achieved with $G=1-y$ in the numerator. Given this definition, the *distortion-optimal* estimate is $y^* \in \arg\max_{y\in\{0,1\}} \mathrm{dist}(y;\boldsymbol{X})$.

This objective requires attention to the technicality mentioned in Section 5.2. Consider $\boldsymbol{X}$ in which some two judgments disagree. Then, under $\boldsymbol{p}=(1,\ldots,1)\in\mathcal{P}_n$, we have $\mathrm{dist}(0;\boldsymbol{X})=\mathrm{dist}(1;\boldsymbol{X})=0$, making distortion undefined. Hence, we use $\mathcal{P}_n^\epsilon=\{\boldsymbol{p}: 1-\epsilon\geqslant p_1\geqslant\ldots\geqslant p_n\geqslant {}^1\!/_2\}$ to redefine the distortion as

$$\mathrm{dist}^\epsilon(y;\boldsymbol{X}) = \sup_{\boldsymbol{p}\in\mathcal{P}_n^\epsilon} \frac{\mathcal{L}[\boldsymbol{X};G=1-y,\boldsymbol{p}]}{\mathcal{L}[\boldsymbol{X};G=y,\boldsymbol{p}]}.$$

The distortion-optimal rule f_{dist} is defined as $f_{\mathrm{dist}}(\boldsymbol{X})=\lim_{\epsilon\to 0}\mathrm{argmin}_{y\in\{0,1\}}\mathrm{dist}^\epsilon(y;\boldsymbol{X})$. Interestingly, we show that the estimate y minimizing $\mathrm{dist}^\epsilon(y;\boldsymbol{X})$ is independent of ϵ, making the limit unnecessary. First, we define a quantity that we will later show to be closely related to distortion.

Definition 1. Given $\boldsymbol{X}\in\{0,1\}^n$, the *strength* $s_{\boldsymbol{X}}(y)$ of estimate y is the maximum difference between the number of occurrences of y and that of $1-y$ in any prefix of $\boldsymbol{X}$, i.e.,

$$s_{\boldsymbol{X}}(y) = \max_{k\in[n]\cup\{0\}} \textstyle\sum_{i=1}^k \left\{\mathbb{1}[X_i=y]-\mathbb{1}[X_i=1-y]\right\}.$$

Lemma 1. *For* $\epsilon\in(0,{}^1\!/_2)$*,* $n\in\mathbb{N}$*,* $\boldsymbol{X}\in\{0,1\}^n$*, and* $y\in\{0,1\}$*, we have* $\mathrm{dist}^\epsilon(y;\boldsymbol{X}) = \left(\frac{1-\epsilon}{\epsilon}\right)^{s_{\boldsymbol{X}}(1-y)}$.

Proof. Fix $y\in\{0,1\}$. Given a sequence $\boldsymbol{p}$, we say that it has a *jump* at $i\in[n-1]$ if $p_i>p_{i+1}$.

We first show that in the definition of $\mathrm{dist}^\epsilon(y;\boldsymbol{X})$, the supremum over $\boldsymbol{p}$ is achieved at an accuracy profile with at most one jump. Let $\boldsymbol{p}$ be a vector with the minimum jumps at which the supremum is achieved. Suppose for contradiction that it has at least two jumps, and let k and j be indices such that $p_k > p_{k+1} = \ldots = p_j > p_{j+1}$.

Define $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ such that $p_i^1 = p_i^2 = p_i$ for $i \in [n] \setminus \{k+1, \ldots, j\}$, $p_i^1 = p_k$ for $i \in \{k+1, \ldots, j\}$, and $p_i^2 = p_{j+1}$ for $i \in \{k+1, \ldots, j\}$. That is, in $\boldsymbol{p}^1$, we shift the block $(p_{k+1}, \ldots, p_j)$ up and make it equal to p_k, and in $\boldsymbol{p}^2$, we shift it down and make it equal to p_{j+1}.

We show that at least one of these two vectors must yields an approximation ratio no better than that at $\boldsymbol{p}$, and is therefore also a point where the supremum is achieved; this is a contradiction because they both have one fewer jump than $\boldsymbol{p}$. To see why the claim is true, let $a = p_k$, $b = p_{k+1} = \ldots = p_j$, and $c = p_{j+1}$. Thus, $a > b > c$. Denoting $S = \{k+1, \ldots, j\}$, we have that

$$\begin{aligned}\frac{\mathcal{L}[\boldsymbol{X}; G = 1-y, \boldsymbol{p}]}{\mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}]} &= \prod_{i \in [n] \setminus S} \frac{p_i^{\mathbb{1}[X_i = 1-y]} \cdot (1-p_i)^{\mathbb{1}[X_i = y]}}{p_i^{\mathbb{1}[X_i = y]} \cdot (1-p_i)^{\mathbb{1}[X_i = 1-y]}} \times \prod_{i \in S} \frac{b^{\mathbb{1}[X_i = 1-y]} \cdot (1-b)^{\mathbb{1}[X_i = y]}}{b^{\mathbb{1}[X_i = y]} \cdot (1-b)^{\mathbb{1}[X_i = 1-y]}} \\ &= \prod_{i \in [n] \setminus S} \frac{p_i^{\mathbb{1}[X_i = 1-y]} \cdot (1-p_i)^{\mathbb{1}[X_i = y]}}{p_i^{\mathbb{1}[X_i = y]} \cdot (1-p_i)^{\mathbb{1}[X_i = 1-y]}} \times \left(\frac{b}{1-b}\right)^{\sum_{i \in S}(\mathbb{1}[X_i = 1-y] - \mathbb{1}[X_i = y])}.\end{aligned}$$

In the last expression, if the exponent of $b/(1-b)$ is non-positive, then decreasing b to c does not decrease the expression, and the expression changes from the approximation ratio at $\boldsymbol{p}$ to that at $\boldsymbol{p}^2$. Similarly, if the exponent is non-negative, then increasing b to a does not decrease the expression, and the expression changes from the approximation ratio at $\boldsymbol{p}$ to that at $\boldsymbol{p}^1$. Hence, at least one of $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ achieves a ratio at least as high as $\boldsymbol{p}$, as desired.

We have established that the supremum is achieved at some $\boldsymbol{p}$ which has at most one jump. Then, there exist $a, b \in [1/2, 1-\epsilon]$ with $a > b$ and an index $k \in [n] \cup \{0\}$ such that $p_i = a$ for all $i \leqslant k$ and $p_i = b$ for all $i > k$. Note that allowing $k = 0$ and $k = n$ permits zero jumps. We show that we can let $a = 1 - \epsilon$ and $b = 1/2$ without loss of generality. The approximation ratio at this $\boldsymbol{p}$ is given by

$$\left(\frac{a}{1-a}\right)^{\sum_{i=1}^{k}(\mathbb{1}[X_i = 1-y] - \mathbb{1}[X_i = y])} \times \left(\frac{b}{1-b}\right)^{\sum_{i=k+1}^{n}(\mathbb{1}[X_i = 1-y] - \mathbb{1}[X_i = y])}.$$

The exponent of $a/(1-a)$ must be non-negative (otherwise decreasing a to b would strictly increase the approximation ratio). Hence, increasing a to $1-\epsilon$ does not decrease the approximation ratio. Similarly, we can let $b = 1/2$.

We have thus established that the supremum is achieved at $\boldsymbol{p}$ such that for some $k \in [n] \cup \{0\}$, $p_i = 1-\epsilon$ for $i \leqslant k$ and $p_i = 1/2$ for $i > k$. Thus, the distortion of y given X is

$$\operatorname{dist}^{\epsilon}(y; X) = \max_{k \in [n] \cup \{0\}} \left(\frac{1-\epsilon}{\epsilon}\right)^{\sum_{i=1}^{k}(\mathbb{1}[X_i = 1-y] - \mathbb{1}[X_i = y])},$$

which is $\left(\frac{1-\epsilon}{\epsilon}\right)^{s_{\boldsymbol{X}}(1-y)}$, as desired. □

We can immediately obtain a characterization of the distortion-optimal estimate $y^* \in \{0,1\}$ by observing that $\operatorname{argmin}_{y \in \{0,1\}} s_{\boldsymbol{X}}(1-y) = \arg\max_{y \in \{0,1\}} s_{\boldsymbol{X}}(y)$ and applying Lemma 1: The distortion-optimal estimate is the estimate with the greatest strength in $\boldsymbol{X}$.

Theorem 1. *For any $\epsilon \in (0, 1/2)$, $n \in \mathbb{N}$, and $\boldsymbol{X} \in \{0,1\}^n$,*

$$f_{\text{dist}}(\boldsymbol{X}) = \operatorname*{argmin}_{y \in \{0,1\}} \operatorname{dist}^{\epsilon}(y; \boldsymbol{X}) = \arg\max_{y \in \{0,1\}} s_{\boldsymbol{X}}(y),$$

where f_{dist} is the distortion-optimal rule. Further, this can be computed in linear time.

Note that in case of both estimates having equal strength, the result also implies that their distortion will be equal.

A notable property of f_{dist} is that if more than $n/3$ most accurate judgments or more than $2n/3$ least accurate judgments are identical, then that will be the output of f_{dist}, regardless of the remaining judgments.

2.3.2 Other Objectives

We now turn our attention to two other objectives, namely maximization of optimistic and pessimistic likelihoods. Recall that the reason we cannot directly compute the MLE $\arg\max_{y \in \{0,1\}} \mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}]$ is because we do not know the exact accuracy profile $\boldsymbol{p}$. Instead, we know that $\boldsymbol{p} \in \mathcal{P}_n$. Given this, we define the optimistic and pessimistic likelihoods by taking the best case and the worst case over the choice of $\boldsymbol{p}$, respectively.

The *optimistic likelihood* $\mathcal{L}_\uparrow$ of observing $\boldsymbol{X}$ when the ground truth is $G = y$ is

$$\mathcal{L}_\uparrow[\boldsymbol{X}; G = y] = \sup_{\boldsymbol{p} \in \mathcal{P}_n} \mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}].$$

The *optimistic MLE* rule which maximizes this objective, denoted $f_{\mathrm{MLE}\uparrow}$, is given by

$$f_{\mathrm{MLE}\uparrow}(\boldsymbol{X}) = \operatorname*{arg\,max}_{y \in \{0,1\}} \mathcal{L}_\uparrow[\boldsymbol{X}; G = y].$$

We can view $f_{\mathrm{MLE}\uparrow}$ as simply performing a joint maximum likelihood estimation over $(y, \boldsymbol{p}) \in \{0,1\} \times \mathcal{P}_n$, and returning the y component of the resulting estimate.[3] On the other hand, we can view $f_{\mathrm{MLE}\downarrow}$ as inspired by worst-case analysis. Further, it is easy to see that maximizing the optimistic (resp. pessimistic) likelihood of the chosen estimate is equivalent to minimizing the optimistic (resp. pessimistic) likelihood of the unchosen estimate; thus, we can also view $f_{\mathrm{MLE}\uparrow}$ and $f_{\mathrm{MLE}\downarrow}$ as minimizing the optimistic and pessimistic likelihoods of the unchosen estimate, respectively. This connection holds only because we are looking for the optimal rule within the family of all possible rules; in Section 2.4.1, we will see that when we look for the optimal rule within the family of scoring rules, we have to consider four — not two — objectives.

We begin by presenting an algorithm that calculates the optimistic likelihood of an estimate $y \in \{0,1\}$ given a judgment profile $\boldsymbol{X}$. The algorithm repeatedly identifies a prefix of $\boldsymbol{X}$ with the highest density of y, and imputes that the accuracies of judgments in that prefix are equal to this density.

Theorem 2. *Algorithm 1 calculates the optimistic likelihood $\mathcal{L}_\uparrow[\boldsymbol{X}; G = y]$ in polynomial time. Thus, the aggregation rule $f_{\mathrm{MLE}\uparrow}$ can be implemented in polynomial time.*

It is clear that Algorithm 1 runs in polynomial time. Here we prove the claim that for any

[3] This is also equivalent to computing the maximum a posteriori estimate (MAP) when we are given a uniform prior over $\boldsymbol{p}$.

Algorithm 1 OPT-LIKELIHOOD

Require: Judgment profile $\boldsymbol{X} \in \{0,1\}^n$, $y \in \{0,1\}$
Ensure: Optimistic likelihood $\mathcal{L}_{\uparrow}[\boldsymbol{X}; G = y]$
if $n = 1$ **then**
  return $1^{\mathbb{1}[X_1=y]} \cdot (1/2)^{\mathbb{1}[X_1 \neq y]}$ // Find the prefix of $\boldsymbol{X}$ with the highest density of y
$i \leftarrow \max_i \, (1/i) \cdot \sum_{j=1}^{i} \mathbb{1}[X_j = y]$, breaking ties in favor of larger indices
$d \leftarrow (1/i) \cdot \sum_{j=1}^{i} \mathbb{1}[X_j = y]$
$r \leftarrow \max\{d, 1/2\}$
$L \leftarrow$ OPT-LIKELIHOOD$((X_{i+1}, \dots, X_n), y)$
return $\left(r^d (1-r)^{1-d}\right)^i \cdot L$

$\boldsymbol{X}$ and answer y, the probability vector

$$\boldsymbol{p}^* := \arg\max_{\boldsymbol{p} \in \mathcal{P}_n} \mathcal{L}[\boldsymbol{X} | G = y, \boldsymbol{p}]$$

can be found iteratively by identifying the index i which maximizes the density $r(\boldsymbol{X}^i)$ of y over all prefixes $\boldsymbol{X}^i := (X_1, \dots, X_i)$, taking $p_1^*, \dots p_i^* = \max\{r(\boldsymbol{X}^i), 1/2\}$, and recursing on the suffix $\overline{\boldsymbol{X}}^i := (X_{i+1}, \dots, X_n)$. We show that $\boldsymbol{p}^*$ takes this form for $\boldsymbol{p} \in \mathcal{P}_n$ and only remark that the analogous structure emerges in maximizing over the larger set $\mathcal{Q}_n := \{\boldsymbol{p} : 1 \geqslant p_1 \geqslant \dots \geqslant p_n \geqslant 0\}$.

We begin with two observations. First, that for a given run of experts which share a p_i, the maximizing value of these p_i is the density of the answer y in this run:

Lemma 2. *If $p_1 = p_2 = \dots p_n$ then $\boldsymbol{p} = \arg\max_{\boldsymbol{p} \in \mathcal{Q}_n} \mathcal{L}_{\uparrow}[\boldsymbol{X} | G = y]$ is given by $r(\boldsymbol{X}) = \frac{\|\mathbb{1}[X_i = y]\|_1}{n}$.*

Proof. Given that $p := p_1 = p_2 = \dots p_n$, the likelihood is

$$\mathcal{L}[\boldsymbol{X} | G = y, \boldsymbol{p}] = p^{\|\mathbb{1}[X_i = y]\|_1} (1-p)^{\|\mathbb{1}[X_i \neq y]\|_1}.$$

Taking derivatives shows that this is concave with respect to p, with unique maximum over $[0,1]$ at $p = r(\boldsymbol{X})$. □

Next, if multiple vectors of experts have likelihoods maximized by compatible probability vectors, then the concatenation of these probability vectors maximizes the likelihood of the concatenated expert vectors:

Lemma 3. *If* $\boldsymbol{p}_1 = \arg\max_{\boldsymbol{p}\in\mathcal{P}_{n_1}} \mathcal{L}[\boldsymbol{X}_1|G=y,\boldsymbol{p}], \ldots, \boldsymbol{p}_k = \arg\max_{\boldsymbol{p}\in\mathcal{P}_{n_k}} \mathcal{L}[\boldsymbol{X}_k|G=y,\boldsymbol{p}]$ *and* $p \geqslant p'$ *for all* $p \in \boldsymbol{p}_j$, $p' \in \boldsymbol{p}_{j+1}$ *and for all* j, *then*

$$\boldsymbol{p} = \arg\max_{\boldsymbol{p}\in\mathcal{P}_n} \mathcal{L}[\boldsymbol{X}|G=y,\boldsymbol{p}],$$

where $\boldsymbol{p} := (\boldsymbol{p}_1|\ldots|\boldsymbol{p}_k)$ *and* $\boldsymbol{X} := (\boldsymbol{X}_1|\ldots|\boldsymbol{X}_k)$ *and* $n := \sum_{i=1}^{i} n_i$.

Proof. In short, this is because $\mathcal{P}_n \subset \mathcal{P}_{n_1} \times \ldots \times \mathcal{P}_{n_k}$ and because

$$\mathcal{L}[(\boldsymbol{X}_1|\ldots|\boldsymbol{X}_k)|G=y,(\boldsymbol{p}_1|\ldots|\boldsymbol{p}_k)] = \mathcal{L}[\boldsymbol{X}_1|G=y,\boldsymbol{p}_1] \times \ldots \times \mathcal{L}[\boldsymbol{X}_k|G=y,\boldsymbol{p}_k]. \tag{2.1}$$

Since the $\boldsymbol{p}_i$ maximize the (nonnegative) terms $\mathcal{L}[\boldsymbol{X}_i|G=y,\boldsymbol{p}_i]$ individually, they maximize the product over the larger space $\mathcal{P}_{n_1} \times \ldots \times \mathcal{P}_{n_k}$. The compatibility of the $\boldsymbol{p}_i$ implies that in fact $\boldsymbol{p} \in \mathcal{P}_n$; therefore it maximizes $\mathcal{L}[\boldsymbol{X}|G=y,\boldsymbol{p}]$ over $\mathcal{P}_n$ also. □

Next, it will be useful to view each $\boldsymbol{p}$ as inducing a decomposition of $\boldsymbol{X}$ into chunks $\boldsymbol{X} = (C_1|\ldots|C_k)$ within which all p_i are equal. Let $q_1 \ldots q_k$ be the values of the p_i within each of the chunks, and let $r_1 \ldots r_k$ be the chunk densities $r_j := r(C_j)$.

We now tackle a special case of Theorem 2.

Lemma 4. *If* $r(\boldsymbol{X}^i)$ *is maximized by* $i = n$ *then the optimal* $\boldsymbol{p}^*$ *is* $p_1^*, \ldots p_n^* = \max\{r(\boldsymbol{X}^n), 1/2\}$.

Proof. To see this, let $\boldsymbol{p}^* = \arg\max_{\boldsymbol{p}\in\mathcal{P}_n}$. If $r := r(\boldsymbol{X}^i)$ then there are two cases to consider. First if $r > 1/2$ then, since $r_1 \leqslant r$ and $\sum_{j=1}^{k} \frac{|C_j|}{n} r_j = r$, it follows that either $r_1 = \ldots = r_k = r$ or there is some j for which $r_j \leqslant r_{j+1}$. But since $\boldsymbol{p}^*$ is decreasing, $q_j \geqslant q_{j+1}$. Therefore by Lemma 2 either $\mathcal{L}[C_j|G=y,(q_j,\ldots,q_j)]$ or $\mathcal{L}[C_{j+1}|G=y,(q_{j+1},\ldots,q_{j+1})]$ can be increased by either lowering q_j towards r_j or raising q_{j+1} towards r_{j+1}. In either case $q_1,\ldots,q_k$ remain nonincreasing and by Equation (2.1) the likelihood $\mathcal{L}[\boldsymbol{X}|G=y,\boldsymbol{p}^*]$ increases also, contradicting the optimality of $\boldsymbol{p}^*$. Therefore $r_1 = \ldots = r_k = r$, in which case by Lemma 2 and Lemma 3 $\mathcal{L}[\boldsymbol{X}|G=y,\boldsymbol{p}^*]$ is maximized by $p_1^* = \ldots = p_n^* = r$.

Finally consider $r \leqslant 1/2$. Since r is the maximum density, $r_1 \leqslant 1/2$ also. And since $\boldsymbol{p} \in \mathcal{P}_n$ we have that $q_1, \ldots q_k \geqslant 1/2$. If $q_1 > 1/2$ then since $r_1 \leqslant 1/2$ we have by Lemma 2 that $\mathcal{L}[C_1|G = y, (q_1, \ldots, q_1)]$ is increased by lowering q_1 to the value of q_2. By Equation (2.1) this increases $\mathcal{L}[\boldsymbol{X}|G = y, \boldsymbol{p}^*]$, contradicting the optimality of $\boldsymbol{p}^*$. Therefore $q_1 = 1/2$, and so $q_1, \ldots q_k = 1/2$, as desired. □

We finally prove Theorem 2 by induction.

Proof of Theorem 2. For the base case take $n = 1$. By Lemma 4 it is of the desired form.

Next suppose that for all $\boldsymbol{X}'$ such that $|\boldsymbol{X}'| < n$, the optimal probability vector has the desired form. Given some $\boldsymbol{X}$ of length n, let i be the index maximizing the prefix density $r(\boldsymbol{X}^i)$ of y in $\boldsymbol{X}$, and call this maximum density r. Splitting it into $\boldsymbol{X} = (\boldsymbol{X}^i|\overline{\boldsymbol{X}}^i)$, by Lemma 4 we have that $\boldsymbol{X}^i$ has likelihood maximized by $\boldsymbol{p}^*$ for which $p_1^* = \ldots = p_i^* = r$. By the inductive hypothesis, we also know that the $\overline{\boldsymbol{p}}^*$ maximizing likelihood for $\overline{\boldsymbol{X}}^i$) is of the desired form. Therefore $(\boldsymbol{p}^*|\overline{\boldsymbol{p}}^*)$ is of the desired form also.

It remains to show that $(\boldsymbol{p}^*|\overline{\boldsymbol{p}}^*)$ is nonincreasing and maximizes $\mathcal{L}[X|G = y, p]$. Clearly for all $\boldsymbol{p} \in \boldsymbol{p}^*$ we have $p = r$ the maximum prefix density for $\boldsymbol{X}$. If r' is the maximum prefix density for $\overline{\boldsymbol{X}}^i$, then by hypothesis $p' \leqslant r'$ for all $p' \in \overline{\boldsymbol{p}}^*$. But $r \geqslant r'$, since otherwise $\boldsymbol{X}^i$ together with the prefix of $\overline{\boldsymbol{X}}^i$ witnessing r' would be a prefix of $\boldsymbol{X}$ with a density larger than r, a contradiction. Therefore $(\boldsymbol{p}^*|\overline{\boldsymbol{p}}^*)$ is nonincreasing, and so by Lemma 3 it also maximizes $\mathcal{L}[X|G = y, \boldsymbol{p}]$. □

A more careful analysis can show that the worst-case runtime of Algorithm 1 is in fact $\tilde{\Theta}(n^{4/3})$. A different approach which draws from the problem in computational geometry of finding convex hulls can improve upon this, requiring only $\tilde{\Theta}(n)$ operations.

We conclude by illustrating Algorithm 1 by an example.

Example 1. Let us consider running Opt-Likelihood with $\boldsymbol{X} = (0, 1, 1, 1, 0, 1, 1, 0, 0, 1)$ and $y = 1$.

The first iteration selects $i = 4$ (i.e. prefix $(0, 1, 1, 1)$) because the density of $y = 1$ in this prefix is $3/4$, and this is the highest density in any prefix. This leads to $d = r = 3/4$. The second iteration selects $i = 3$ (i.e. prefix $(0, 1, 1)$), leading to $d = r = 2/3$. The final iteration selects the remaining string, and sets $d = 1/3$ but $r = 1/2$.

Thus, $\boldsymbol{p} = (3/4, 3/4, 3/4, 3/4, 2/3, 2/3, 2/3, 1/3, 1/3, 1/3)$ is the accuracy profile leading to the optimistic likelihood of

$$\left(\frac{3}{4}^{\frac{3}{4}} \cdot \frac{1}{4}^{\frac{1}{4}}\right)^4 \cdot \left(\frac{2}{3}^{\frac{2}{3}} \cdot \frac{1}{3}^{\frac{1}{3}}\right)^3 \cdot \left(\frac{1}{2}^{\frac{1}{3}} \cdot \frac{1}{2}^{\frac{2}{3}}\right)^3. \qquad \square$$

We now turn our attention to maximizing the pessimistic likelihood $\mathcal{L}_\downarrow$. If we define it as $\mathcal{L}_\downarrow[\boldsymbol{X}; G = y] = \inf_{\boldsymbol{p} \in \mathcal{P}_n} \mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}]$, then we run into the issue discussed in Section 5.2: the pessimistic likelihood of any non-unanimous $\boldsymbol{X}$ becomes 0 under both values of y due to the accuracy profile $\boldsymbol{p} = (1, \ldots, 1) \in \mathcal{P}_n$, leading to unnecessary ties. Hence, we again consider $\mathcal{P}_n^\epsilon$ instead of $\mathcal{P}_n$, define $\mathcal{L}_\downarrow^\epsilon[\boldsymbol{X}; G = y] = \inf_{\boldsymbol{p} \in \mathcal{P}_n^\epsilon} \mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}]$, and define the *pessimistic MLE* rule, denoted $f_{\text{MLE}\downarrow}$, as $f_{\text{MLE}\downarrow}(\boldsymbol{X}) = \lim_{\epsilon \to 0} \arg\max_{y \in \{0,1\}} \mathcal{L}_\downarrow^\epsilon[\boldsymbol{X}; G = y]$. Unlike in the case of distortion, the choice of y does not turn out to be independent of ϵ, but as we see in the proof of Theorem 3, the rule converges once $\epsilon < 2^{-n}$.

The next result identifies $f_{\text{MLE}\downarrow}$ analytically. This is possible because the accuracy profile resulting in the pessimistic likelihood always consists of only $1 - \epsilon$ and $1/2$. This is in contrast to the one leading to the optimistic likelihood, which, as Example 1 demonstrates, can be more complex.

Theorem 3. *The pessimistic MLE rule $f_{\text{MLE}\downarrow}$, given a judgment profile $\boldsymbol{X}$, outputs the majority judgment; if tied, it outputs the opposite of the least accurate judgment (i.e. $1 - X_n$).*

Proof. Fix $\epsilon \in (0, 1/2)$. First, we demonstrate that if $\boldsymbol{p}^\star \in \operatorname{argmin}_{\boldsymbol{p} \in \mathcal{P}_n^\epsilon} \mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}]$, then $p_i^\star \in \{1 - \epsilon, 1/2\}$ for all $i \in [n]$. Suppose for contradiction that this is not the case. Let $S = \{k, \ldots, j\}$ be a maximal contiguous block of indices of $\boldsymbol{p}^\star$ with some value $a \notin \{1 - \epsilon, 1/2\}$. Note that the contribution of this block to $\mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}^\star]$ is equal to

$a^{\sum_{i \in S} \mathbb{1}[X_i = y]} \cdot (1-a)^{\sum_{i \in S} \mathbb{1}[X_i \neq y]}$. It is easy to check that this is a convex function of a in $[0,1]$.

Define $p^\star_0 = 1 - \epsilon$ and $p^\star_{n+1} = 1/2$. Then, we have that $p^\star_{k-1} > a > p^\star_{j+1}$ and $a \notin \{1-\epsilon, 1/2\}$. Hence, it is feasible to increase or decrease a slightly. Since one of these two operations must reduce the likelihood, we have a contradiction.

Thus, in computing $\mathcal{L}^\epsilon_\downarrow[\boldsymbol{X}; G = y]$, it is sufficient to minimize over $\boldsymbol{p}$ which consist of only $1 - \epsilon$ and $1/2$. Hence, we have that $\mathcal{L}^\epsilon_\downarrow[\boldsymbol{X}; G = y]$ is equal to

$$\min_{k \in [n] \cup \{0\}} (1-\epsilon)^{\sum_{i=1}^{k} \mathbb{1}[X_i = y]} \cdot \epsilon^{\sum_{i=1}^{k} \mathbb{1}[X_i \neq y]} \cdot (1/2)^{n-k}.$$

Note that the estimate y^* maximizing this objective can be found in linear time. We now study the value of y^* as $\epsilon \to 0$.

When $\epsilon \leqslant 2^{-n}$, the ϵ term in the equation dominates; that is, the pessimistic likelihood is achieved by maximizing the exponent of ϵ, and subject to that, maximizing the exponent of $1/2$. Thus, if y appears more often then $1 - y$, then $\mathcal{L}^\epsilon_\downarrow[\boldsymbol{X}; G = 1 - y]$ will have a higher exponent of ϵ (and therefore, will be lower) than $\mathcal{L}^\epsilon_\downarrow[\boldsymbol{X}; G = y]$. In this case, the rule will return y. Finally, suppose that both 0 and 1 appear exactly $n/2$ times. If $\boldsymbol{X}$ ends with k occurrences of y, then we will have $\mathcal{L}^\epsilon_\downarrow[\boldsymbol{X}; G = y] = (1-\epsilon)^{n/2-k} \cdot (\epsilon)^{n/2} \cdot (1/2)^k$ whereas $\mathcal{L}^\epsilon_\downarrow[\boldsymbol{X}; G = 1 - y] = (1-\epsilon)^{n/2} \cdot (\epsilon)^{n/2}$, leading the rule to return $1 - y$, as desired. $\square$

2.4 Optimal Scoring Rules

We now turn our attention to a natural class of aggregation rules, scoring rules. Specifically, we are interested in how well we can optimize certain objectives when we are restricted to this class of functions. There are two clear ambiguities about scoring rules that we must take into account. First, it is often unlikely that a scoring rule can be *instance optimal* for a given objective and rules can often be incomparable, one rule achieving a better objective value on a judgment string while worse on another. To handle this, we change our goal slightly and instead choose scoring rules that are optimal in the worst case. Next, is the issue of ties. In this case we'll take a pessimistic view (in line with our worst case objective goals) and say that when a scoring rule outputs a tie, the value of the objective in this instance will be the worse of the two outcomes.

Theorem 4. *For any $\epsilon \in (0, 1/2)$ and $n \in \mathbb{N}$, the scoring rule given by $\boldsymbol{w}^* = (1, \ldots, 1, 0, \ldots, 0)$ with exactly $2\lfloor n/3 \rfloor + 1$ ones minimizes the worst case distortion $\max_{\boldsymbol{X} \in \{0,1\}^n} \mathrm{dist}^\epsilon(f_{\boldsymbol{w}}(\boldsymbol{X}); \boldsymbol{X})$ over all possible scoring rules parametrized by $\boldsymbol{w} \in \mathbb{R}^n_{\geqslant 0}$.*

Proof. Fix $\epsilon \in (0, 1/2)$ and $n \in \mathbb{N}$. Recall that minimizing the distortion, $\mathrm{dist}^\epsilon(y; \boldsymbol{X})$ is equivalent to minimizing the strength of the unchosen judgment, $s_{\boldsymbol{X}}(1-y)$.

First, we'll show that no rule f (scoring or otherwise) can guarantee $s_{\boldsymbol{X}}(1 - f(\boldsymbol{X})) < \lfloor n/3 \rfloor$ for all $\boldsymbol{X} \in \{0,1\}^n$. This will imply $\max_{\boldsymbol{X} \in \{0,1\}^n} s_{\boldsymbol{X}}(1 - f_{\boldsymbol{w}}(\boldsymbol{X})) \geqslant \lfloor n/3 \rfloor$ for all $\boldsymbol{w} \in \mathbb{R}^n_{\geqslant 0}$. To see this, we construct $\boldsymbol{X} \in \{0,1\}^n$ such that both $s_{\boldsymbol{X}}(0)$ and $s_{\boldsymbol{X}}(1)$ are at least $\lfloor n/3 \rfloor$. Consider the judgment profile $\boldsymbol{X}^s = (1, \ldots, 1, 0, \ldots, 0)$ with $\lfloor n/3 \rfloor$ ones and $n - \lfloor n/3 \rfloor$ zeros. On the prefix of the first $\lfloor n/3 \rfloor$ judgments, there are $\lfloor n/3 \rfloor$ more 1s than there are 0s. Thus, the strength of 1 is at least $\lfloor n/3 \rfloor$. On the other hand, on the entire profile, there are $(n - \lfloor n/3 \rfloor) - \lfloor n/3 \rfloor \geqslant n - \frac{2n}{3} \geqslant \lfloor n/3 \rfloor$ more 0s then 1s. Hence, the strength of 0 is at least $\lfloor n/3 \rfloor$, as desired.

Next, we show that $s_{\boldsymbol{X}}(1 - f_{\boldsymbol{w}^*}(\boldsymbol{X})) \leqslant \lfloor n/3 \rfloor$ for all judgment profiles $\boldsymbol{X} \in \{0,1\}^n$. Qualitatively, $f_{\boldsymbol{w}^*}$ simply picks the majority bit of the first $2\lfloor n/3 \rfloor + 1$ bits. Note that since $2\lfloor n/3 \rfloor + 1$ is odd, there is always a majority bit and thus $f_{\boldsymbol{w}^*}$ will never output a tie.

Let $\boldsymbol{X} \in \{0,1\}^n$ and, without loss of generality, suppose $f_{\boldsymbol{w}^*}(\boldsymbol{X}) = 1$. We show that $s_{\boldsymbol{X}}(0) \leqslant \lfloor n/3 \rfloor$. Since $f_{\boldsymbol{w}^*}$ chose 1, there cannot be a majority of 0s in the first $2\lfloor n/3 \rfloor + 1$ bits. Hence, 0 occurs at most $\lfloor n/3 \rfloor$ times in this prefix. This implies that for $k \leqslant 2\lfloor n/3 \rfloor + 1$, $\sum_{i=1}^k \{\mathbb{1}[X_i = 0] - \mathbb{1}[X_i = 1]\} \leqslant \lfloor n/3 \rfloor$. Next, since 1 has a majority among the first $2\lfloor n/3 \rfloor + 1$

bits, $\sum_{i=1}^{2\lfloor n/3 \rfloor+1} \{\mathbb{1}[X_i = 0] - \mathbb{1}[X_i = 1]\} \leqslant -1$. or $k > 2\lfloor n/3 \rfloor + 1$,

$$\begin{aligned} &\textstyle\sum_{i=1}^{k} \{\mathbb{1}[X_i = 0] - \mathbb{1}[X_i = 1]\} \\ &\leqslant \textstyle\sum_{i=2\lfloor n/3 \rfloor+2}^{k} \{\mathbb{1}[X_i = 0] - \mathbb{1}[X_i = 1]\} - 1 \\ &\leqslant k - (2\lfloor n/3 \rfloor + 1) - 1 \\ &\leqslant n - (2\lfloor n/3 \rfloor + 1) - 1 \\ &= n - (3\lfloor n/3 \rfloor + 2) + \lfloor n/3 \rfloor \\ &\leqslant n - n + \lfloor n/3 \rfloor = \lfloor n/3 \rfloor. \end{aligned}$$

So, for all $k \in \{0\} \cup [n]$, $\sum_{i=1}^{k} \{\mathbb{1}[X_i = 0] - \mathbb{1}[X_i = 1]\} \leqslant \lfloor n/3 \rfloor$, and hence $s_{\boldsymbol{X}}(0) \leqslant \lfloor n/3 \rfloor$ as desired. □

2.4.1 Other Objectives

We also investigate optimal scoring rules with respect to the optimistic and pessimistic MLE rules $f_{\mathrm{MLE}\uparrow}$ and $f_{\mathrm{MLE}\downarrow}$. Section 2.3.2 posits that $f_{\mathrm{MLE}\uparrow}$ and $f_{\mathrm{MLE}\downarrow}$ can equivalently be viewed as either maximizing their respective likelihoods for the chosen estimate or as minimizing their respective likelihoods for the unchosen estimate. However when it comes to identifying the worst-case optimal scoring rule across judgment profiles, this equivalence ceases to hold. Thus, we need to derive an optimal scoring rule for each case.

Definition 2. We define optimal scores $\boldsymbol{w}_\circ^\uparrow, \boldsymbol{w}_\times^\uparrow, \boldsymbol{w}_\circ^\downarrow, \boldsymbol{w}_\times^\downarrow$ as

- $\boldsymbol{w}_\circ^\uparrow \in \arg\max_{\boldsymbol{w} \in \mathbb{R}^n_{\geqslant 0}} \min_{\boldsymbol{X}} \mathcal{L}_\uparrow[\boldsymbol{X}; G = f_{\boldsymbol{w}}(\boldsymbol{X})]$
- $\boldsymbol{w}_\times^\uparrow \in \arg\min_{\boldsymbol{w} \in \mathbb{R}^n_{\geqslant 0}} \max_{\boldsymbol{X}} \mathcal{L}_\uparrow[\boldsymbol{X}; G = 1 - f_{\boldsymbol{w}}(\boldsymbol{X})]$
- $\boldsymbol{w}_\circ^\downarrow \in \arg\max_{\boldsymbol{w} \in \mathbb{R}^n_{\geqslant 0}} \min_{\boldsymbol{X}} \mathcal{L}_\downarrow[\boldsymbol{X}; G = f_{\boldsymbol{w}}(\boldsymbol{X})]$
- $\boldsymbol{w}_\times^\downarrow \in \arg\min_{\boldsymbol{w} \in \mathbb{R}^n_{\geqslant 0}} \max_{\boldsymbol{X}} \mathcal{L}_\downarrow[\boldsymbol{X}; G = 1 - f_{\boldsymbol{w}}(\boldsymbol{X})]$.

For example, $\boldsymbol{w}_\circ^\uparrow$ is the 'worst-case MLE-optimal scoring rule' in that it maximizes the optimistic likelihood of its chosen answer in the worst case. For this rule it suffices to always choose the most accurate expert's judgment:

Theorem 5. *The score* $\boldsymbol{w}_\circ^\uparrow = (1, 0, \ldots, 0)$ *is optimal.*

Proof. Consider the even and odd length alternating judgment vectors $\boldsymbol{X}^e := (0, 1, 0, 1, \ldots, 0, 1)$ and $\boldsymbol{X}^o := (0, 1, 0, 1, \ldots, 0, 1, 0)$. For any fixed $\epsilon > 0$, observe that by Theorem 2 for both $\boldsymbol{X}^{alt} = \boldsymbol{X}^e, \boldsymbol{X}^o$ we have that $\mathcal{L}_\uparrow[\boldsymbol{X}^{alt}; G = 1] = 2^{-n}$, and $\mathcal{L}_\uparrow[\boldsymbol{X}^{alt}; G = 0] = (1 - \epsilon)2^{-n+1}$. These judgment vectors are noteworthy because no matter how $f_{\boldsymbol{w}}$ breaks ties, it must choose either 0 or 1 and incur either 2^{-n} or $(1 - \epsilon)2^{-n+1}$ as its maximum likelihood for $\boldsymbol{X}^{alt}$. If $Q(\boldsymbol{w}) := \min_{\boldsymbol{X}} \mathcal{L}_\uparrow[\boldsymbol{X}; G = f_w(\boldsymbol{X})]$ is the objective we seek to maximize, then for all $\boldsymbol{w}$ we therefore have that $Q(\boldsymbol{w}) \leqslant (1 - \epsilon)2^{-n+1}$.

Now consider the score $\boldsymbol{w}^* = (1, 0, \ldots, 0)$. For any judgment vector $\boldsymbol{X}$, the scoring rule chooses $f_{\boldsymbol{w}^*}(\boldsymbol{X}) = X_1$. Then taking $p = (1 - \epsilon, 1/2, \ldots, 1/2)$ we see that $\mathcal{L}_\uparrow[\boldsymbol{X}; f_{\boldsymbol{w}^*}(\boldsymbol{X})] \geqslant \mathcal{L}[\boldsymbol{X}; G = X_1, p] = (1 - \epsilon)2^{-n+1}$. Minimizing over all $\boldsymbol{X}$ yields $Q(\boldsymbol{w}^*) \geqslant (1 - \epsilon)2^{-n+1}$, and so $\boldsymbol{w}^*$ is optimal. $\square$

For the cases based on pessimistic likelihood (both maximizing it for the chosen answer and minimizing it for the unchosen answer), characterizing the optimum scoring rule is easy, since the optimum rule we identified in Theorem 3 can be represented as a scoring rule.

Theorem 6. *Scores* $\boldsymbol{w}_\circ^\downarrow = \boldsymbol{w}_\times^\downarrow = (1, \ldots, 1, 1/2)$ *are optimal for* $\epsilon \leqslant 2^{-n}$*, and coincide with the rule of Theorem 3.*

The remaining case, $\boldsymbol{w}_\times^\uparrow$, that is minimizing the optimistic likelihood of the unchosen answer, is less straightforward. Optimal scores $\boldsymbol{w}_\times^\uparrow$ for $n \leqslant 20$ are found using a linear program and cataloged in the full version of [Hal+21].

For general n, $\boldsymbol{w}_\times^\uparrow$ is unknown, but we show that there exists an optimal scoring rule that is nonincreasing, for all n.

Theorem 7. *For each n, there is a choice of $\boldsymbol{w}_\times^\uparrow$ that is nonincreasing.*

Proof. We proceed by arguing that for any score with an increasing pair of entries, the score which flips these entries to be decreasing performs at least as well with respect to the objective. By applying this repeatedly we find that there is a decreasing optimal scoring rule.

To see this, consider a score $\boldsymbol{w}$ with increasing pair of indices $i < j$ for which $w_i < w_j$, and take $\boldsymbol{w}'$ to be this score which flips these two entries; that is, $\boldsymbol{w}'_i = \boldsymbol{w}_j$, $\boldsymbol{w}'_i = \boldsymbol{w}_j$, and $\boldsymbol{w}'_k = \boldsymbol{w}_k$ for all $k \neq i, j$. Again let $Q(\boldsymbol{w}) := \max_{\boldsymbol{X}} \mathcal{L}_\uparrow[\boldsymbol{X}; G = 1 - f_w(\boldsymbol{X})]$ be the objective quantity which our optimal scoring rule $\boldsymbol{w}_\times^\uparrow$ minimizes; we argue that $Q(\boldsymbol{w}') \leqslant Q(\boldsymbol{w})$.

Let $\boldsymbol{X}'$ be any vector of expert judgments, and let $y := f_w(\boldsymbol{X}')$ denote the estimate chosen by $\boldsymbol{w}$ for judgment vector $\boldsymbol{X}'$. We will construct an $\boldsymbol{X}$ for which $\mathcal{L}_\uparrow[\boldsymbol{X}'; G = 1 - f_{w'}(\boldsymbol{X}')] \leqslant \mathcal{L}_\uparrow[\boldsymbol{X}; G = 1 - f_w(\boldsymbol{X})]$. First, if $f_{w'}(\boldsymbol{X}') = y$ then take $\boldsymbol{X} = \boldsymbol{X}'$. This is the case if $X'_i = X'_j$, since $\boldsymbol{w} \cdot \mathbb{1}[X'_k = y] = \boldsymbol{w}' \cdot \mathbb{1}[X'_k = y]$. If $X'_i = y$ and $X'_j = 1 - y$ then $\boldsymbol{w}' \cdot \mathbb{1}[X'_k = y] \geqslant \boldsymbol{w} \cdot \mathbb{1}[X'_k = y]$ and so $f_{w'}(\boldsymbol{X}') = y$ also.

Therefore when $f_{w'}(\boldsymbol{X}') = 1 - y$ we have that $X'_i = 1 - y$ and $X'_j = y$. In this case let $\boldsymbol{X}$ be $\boldsymbol{X}'$ with X_i and X_j flipped; that is, $X_i := X'_j = y$ and $X_j := X'_i = 1 - y$ and $X_k := X'_k$ otherwise. Note that $f_w(\boldsymbol{X}) = f_{w'}(\boldsymbol{X}') = 1 - y$, since the scores are identical; $\boldsymbol{w} \cdot \mathbb{1}[X_k = 1 - y] = \boldsymbol{w}' \cdot \mathbb{1}[X'_k = 1 - y]$. Let $\boldsymbol{p}'$ be the probability vector which witnesses $\mathcal{L}_\uparrow[\boldsymbol{X}'; G = 1 - f_{w'}(\boldsymbol{X}')] = \mathcal{L}_\uparrow[\boldsymbol{X}'; G = y]$, so that it equals

$$(1 - p'_i) \cdot p'_j \cdot \prod_{X'_k = 1-y; k \neq i,j} p'_k \prod_{X'_k = y; k \neq i,j} (1 - p'_k).$$

Then since it is a maximum we have that $\mathcal{L}_\uparrow[\boldsymbol{X}; G = 1 - f_w(\boldsymbol{X})] \geqslant \mathcal{L}[\boldsymbol{X}; G = y, \boldsymbol{p}']$, which equals

$$p'_i \cdot (1 - p'_j) \cdot \prod_{X'_k = 1-y; k \neq i,j} p'_k \prod_{X'_k = y; k \neq i,j} (1 - p'_k).$$

But since $p'_i \geqslant p'_j$ and $(1 - p'_j) \geqslant (1 - p'_i)$, this is itself greater than or equal to $\mathcal{L}_\uparrow[\boldsymbol{X}'; G = y]$.

Therefore for every $\boldsymbol{X}'$ there is some $\boldsymbol{X}$ such that

$$\mathcal{L}_{\uparrow}[\boldsymbol{X}'; G = 1 - f_{w'}(\boldsymbol{X}')] \leqslant \mathcal{L}_{\uparrow}[\boldsymbol{X}; G = 1 - f_{w}(\boldsymbol{X})].$$

Taking the max over all $\boldsymbol{X}$ then yields $Q(\boldsymbol{w}') \leqslant Q(\boldsymbol{w})$, as desired.

Since every score can be made decreasing by applying finitely many such flips, we have that for every score $\boldsymbol{w}$ there is a decreasing score $\boldsymbol{w}^*$ for which $Q(\boldsymbol{w}^*) \leqslant Q(\boldsymbol{w})$. Therefore there is an optimal score which is decreasing. □

2.5 Discussion

Our setting boils down to the design of Boolean functions that take a string of bits as input and output a single bit — with the twist that the order of bits matters, in that earlier bits are given greater importance. We view this as a fundamental problem, and there are many ways to approach it. In addition to the objectives and algorithms described in Sections 2.3 and 2.4, three additional approaches are presented in [Hal+21]: axiomatic, Bayesian, and randomized.

One might ask whether the assumption that $p_i \geqslant 1/2$ for all $i \in N$ can be relaxed. If the identities of experts with $p_i < 1/2$ are known, we can simply flip their judgments and reverse their order (as the flipped judgment of the least accurate expert is now the most accurate). Interestingly, our problem now becomes that of aggregating *two* strings of judgments, ordered by accuracy, into a single bit. This problem is potentially richer than ours because there is no information on the relative accuracy of experts associated with two different strings. An even more general setup simply provides a *partial* order of the experts by accuracy; this is especially challenging because it generalizes both our original problem (total order) and the well-studied setting where there is no information about the accuracy of experts (empty order).

Another natural variant of our setting is one where, instead of binary judgments, experts provide real-valued judgments in, say, $[0, 1]$, and the goal is to aggregate them to return a single real number in $[0, 1]$. Interestingly, given a binary aggregation rule f from our work, one can compute the greatest value $x \in [0, 1]$ such that converting expert judgments to binary depending on whether they are at least x and feeding them to f gives output 1; this is well-defined when f satisfies a natural monotonicity condition.

Chapter 3

Worst-Case Voting when the Stakes are High

3.1 Introduction

Distortion is a widely-used metric that captures the worst-case loss in efficiency of a social choice function (SCF) [Ans+21a]. It is defined in the *implicit utilitarian* model where voters have cardinal utilities for alternatives but only report ordinal information, e.g., (partial) rankings, to the social choice function, which then outputs a distribution over winning alternatives.

Distortion evaluates SCFs according to their worst-case performance over all implicit utilities and corresponding induced rankings, where performance is measured in terms of *(utilitarian) social welfare*, i.e. the sum of all agents' utilities. Specifically, the distortion of a rule is the maximum ratio between the social welfare of the optimal alternative and the expected social welfare given by the rule.

While utilitarian social welfare is a defensible basis on which to evaluate social choice functions [Bou+12], distortion is not always the best tool for the job. In particular, we might prefer a social choice function which delivers poor multiplicative guarantees on instances where *no alternative confers significant social welfare*, so long as it performs well on instances where the potential gains are large. For example, a $1/\sqrt{m}$-approximation is a much more tolerable loss when the maximum attainable social welfare is $O(\log n)$ (as for a symmetric profile with n alternatives) than when it is fully $\Omega(n)$.

Indeed, the canonical instance which demonstrates a $\Omega(\sqrt{m})$ lower bound on distortion for randomized social choice functions [Bou+12] allots at most a $1/\sqrt{m}$ proportion of the total utility to any alternative. In practice—for example, in political contests—we often expect that there are alternatives which confer *much* larger social welfare than the average alternative.

To address these concerns we instead study the *additive distortion* of randomized social choice functions, which may be viewed as their worst-case expected regret [Car+17]. The additive distortion of a social choice function is the *difference* between the maximum social welfare attainable and the expected social welfare that f delivers, in the worst case over all implicit utilities. Different profiles in the implicit utilitarian model can have vastly different maximum attainable social welfare, and we posit that, in evaluating social choice functions, additive distortion appropriately prioritizes the instances in which the most utility can be gained or lost. More concretely, consider a fixed profile of ordinal votes. Multiplicative distortion hedges against bad performance in the case of consistent utilities which assign low total welfare for all candidates, which harms its performance for consistent utilities that yield high-welfare candidates. Additive distortion, on the other hand, prioritizes good performance for this latter case. In other words, multiplicative distortion hedges against bad performance in situations where every candidate has low total welfare, which harms its performance when there exist high-welfare candidates. Additive distortion, on the other hand, prioritizes good performance in these exact situations.

In its introduction to the social choice setting, distortion was compared to the distortion of metric embeddings [PR06]; this additive distortion is similarly analogous [LS93].

Although we advocate for additive distortion primarily on the above grounds, another advantage is that it remains a meaningful worst-case metric under weaker assumptions about voters' utilities. Past work on distortion in the (non-metric) implicit utilitarian model has made the assumption that all voters' utilities are *unit-sum* [PR06; CP11; Bou+12; Car+17; Ben+21]. This is not a coincidence: with potentially apathetic voters whose utilities are instead *unit-capped*, one can show that choosing an alternative uniformly at random (incurring distortion m) is optimal, and that the distortion of any deterministic rule is infinite. However, the assumption that all participating voters' total utility is *equal* is unreasonable in many settings, and we instead uniformly cap the sum of voters' utilities at one [Azi19]. As we will show in Section 3.3, additive distortion provides a discerning metric by which to evaluate SCFs in this broader context.

In this chapter we aim to answer the following questions:

> **Question 1:** *What is the best additive distortion attainable for randomized social choice functions?*
>
> **Question 2:** *How well do prominent social choice functions perform with respect to additive distortion, both in theory and in practice?*

Our Results

In the pursuit of randomized SCFs with low additive distortion, we define a natural class of rules which we call *randomized scoring rules*, which are the natural randomized analog of positional scoring rules. A randomized scoring rule (RSR) first computes aggregate scores based on a scoring vector (as scoring rules do), and then chooses each alternative with probability proportional its score. Like scoring rules, RSRs are both intuitive and easy to compute. The two most prominent RSRs—Randomized Dictatorship, and the harmonic rule of Boutilier et al. [Bou+12]—are nearly distortion-optimal in the normalized utility and metric settings, respectively. When considered together with our results, we argue that RSRs merit wider attention in the study of distortion.

In Section 3.3 we address Question 1. We establish that Randomized Dictatorship (RD) has additive distortion $\frac{1}{2}(1-1/m)\cdot n$, and lower bound the best additive distortion obtainable by any randomized social choice function. We then present the *Best-or-Bust* (BoB) rule, which has distortion at most $\frac{11}{27}\cdot n$ and asymptotically minimizes additive distortion within the class of all randomized scoring rules. In particular, this establishes an asymptotic separation between deterministic and randomized voting rules with respect to additive distortion, even as m becomes large. We also show that the obstructions to minimizing additive distortion are information-theoretic rather than computational by presenting an instance-optimal randomized social choice function which can be computed efficiently.

In Section 3.4 we present an alternative metric for prioritizing the worst-case performance on instances with high attainable social welfare, which we call *promise distortion*. This is a beyond-worst-case guarantee that some alternative confers social welfare at least $\alpha \cdot n$, for some $\alpha \in [0,1]$. We analyze the extent to which multiplicative promise distortion circumvents the $\Omega(\sqrt{m})$ lower bound of Boutilier et al. [Bou+12], relate it to additive distortion, and provide an analysis of some social choice functions with respect to both additive and multiplicative promise distortion.

We answer Question 2 in Section 3.4 and Section 7.5. In Section 3.4 we analyze a range of prominent social choice functions through the lens of additive distortion, providing upper and lower bounds on their worst-case performance.

In Section 7.5, we evaluate the performance of our asymptotically optimal positional scoring rule against other scoring rules commonly used in practice, optimal randomized and deterministic algorithms for additive distortion, and an optimal randomized algorithm for (multiplicative) distortion. We observe that the optimal algorithm for multiplicative distortion is no longer optimal for additive distortion, and that the Plurality RSR performs the best on profiles encountered in practice, which suggests that, in practice, votes are far from worst-case instances.

3.1.1 Related Work

Distortion was first introduced by Procaccia and Rosenschein [PR06] in the context of deterministic single-winner social choice functions and normalized utilities. In a later paper, Caragiannis and Procaccia [CP11] proved that the Plurality rule has a distortion of $O(m^2)$, and further work demonstrated that this is the best possible distortion of any deterministic voting rule [Car+17].

Beyond deterministic social choice functions, Boutilier et al. [Bou+12] initiated the study of average-case analysis of randomized social choice functions under distributional assumptions about utilities. They also showed an $\Omega(\sqrt{m})$ lower bound on the distortion of any randomized rule in the worst case, and introduced a pair of voting rules with distortion $O(\sqrt{m} \cdot \log^* m)$ and $O(\sqrt{m \cdot \log m})$, the latter of which makes use of the harmonic scoring vector. Caragiannis et al. [Car+17] introduced regret to the implicit utilitarian model of voting; the regret that they study is equivalent to additive distortion in their unit-sum utility setting. They study choosing a k-subset of alternatives when social welfare is linear in the winners. For $k = 1$ and deterministic rules, their straightforward claims apply to additive distortion also; for randomized rules their results imply a $\frac{1}{4} \cdot n$ lower bound on additive distortion and a rule with at most $\frac{1}{2}\left(1 - \frac{1}{m^2}\right) \cdot n$ additive distortion. We show better upper and lower bounds for randomized rules.

Multiplicative distortion has also received attention in the metric setting. There voters and alternatives sit in a metric space, distances are costs, and one generally aims to minimize the social cost of a chosen alternative, given only voters' rankings. Anshelevich et al. [Ans+18] first studied metric distortion, demonstrating that the Copeland rule has a distortion of 5, in stark contrast to the bounds of the unit-sum utility setting. They also conjectured that the deterministic lower bound of 3 is tight, and many papers made progress toward this conjecture [SE17; GKM17; MW19; Kem20a] before its ultimate proof by Gkatzelis, Halpern, and Shah [GHS20]. Here again randomized rules do better: Anshelevich and Postl [AP17] showed that Randomized Dictatorship has distortion at most $3 - 2/n$ and gave a lower bound of 2 on the distortion of all randomized rules in the metric setting. Kempe [Kem20b] and Gkatzelis, Halpern, and Shah [GHS20] each present rules attaining $3 - 2/m$, and Anshelevich and Postl [AP17] and Fain et al. [Fai+19] study variants of the randomized dictatorship mechanism. Lastly, Seddighin, Latifian, and Ghodsi [SLG21] studies distortion when some voters may abstain. Unfortunately additive distortion is uninteresting here because there is no (dis)utility normalization—additive distortion is made arbitrarily large by rescaling an instance. For a comprehensive survey of works concerning multiplicative distortion, see [Ans+21a; Ans+21b].

Finally, we study a class of SCFs which are the randomized analog of positional scoring rules. Young [You75] characterized deterministic scoring functions (with rounds of tiebreaking) as the SCFs which are anonymous, neutral, and consistent, and Xia and Conitzer [XC08] provide a striking deterministic generalization of scoring rules. Walsh and Xia [WX12] and Bentert and Skowron [BS20] present schemes which may be viewed as randomized generalizations of scoring rules, where deterministic rules are applied to profiles formed by subsampling voters and alternatives, respectively.

3.2 Setting and Definitions

Consider voters $N = [n]$ and alternatives A, with $|A| = m$. Each voter $i \in N$ has a ranking σ_i over A which is a strict total order; we say that $a \succ_i b$ for alternatives $a, b \in A$ if $\sigma_i(a) < \sigma_i(b)$. The collection of rankings $\sigma = (\sigma_i)_{i \in N}$ is a *profile*; let $\Sigma := \mathcal{S}_A^n$ denote the collection of all profiles.

Voters have implicit utilities $u_i \in \mathbb{R}_+^A$ which are consistent with their rankings; that is, if

$a \succ_i b$ then $u_i(a) \geqslant u_i(b)$. We say that $u \triangleright \sigma$ for a collection of utilities u if u_i is consistent with σ_i for all voters i. Weakening the standard unit-sum implicit utility assumption, we assume:

Assumption 8. The total utility of each voter is unit-capped at $\sum_{a\in A} u_i(a) \leqslant 1$ for all voters i.

Given a profile σ, a deterministic *social choice function* $f : \Sigma \to A$ chooses an alternative to be the winner for this profile. Similarly, a randomized social choice function $f : \Sigma \to \Delta_A$ returns a probability distribution over winners, where Δ_A is the probability simplex over A; at election time, a winner is drawn randomly from the probability distribution $f(\sigma) \in \Delta_A$. Here SCFs are randomized unless otherwise stated.

Perhaps the most prominent class of deterministic SCFs are *scoring functions*, or *(positional) scoring rules* (SRs). Each SR f^s is given by a scoring vector $s \in \mathbb{R}^m$. It first assigns to each alternative $a \in A$ the aggregate score $S_a := \sum_i s_{\sigma_i^{-1}(a)}$, which is the score associated with each voter i's ranking of a, summed over all voters. The alternative with the maximum score is then chosen. Scoring functions can handle ties either by returning the set of alternatives with maximal scores, or by using additional scoring vectors to iteratively break ties.

As outlined above, the multiplicative distortion of a randomized SCF f is the worst-case ratio

$$\mathrm{dist}(f) := \max_{\sigma} \max_{u \triangleright \sigma} \frac{\max_{a^* \in A} \mathrm{sw}(a^*)}{\mathbb{E}_{a\sim f(\sigma)}[\mathrm{sw}(a)]},$$

over all profiles σ and utility profiles u consistent with σ, where $\mathrm{sw}(a)$ denotes the social welfare of a: $\mathrm{sw}(a) := \sum_{i\in N} u_i(a)$. Additive distortion is the difference, rather than the ratio:

$$\mathrm{dist}^+(f) := \max_{\sigma} \max_{u \triangleright \sigma} \left(\max_{a^* \in A} \mathrm{sw}(a^*) - \mathbb{E}_{a\sim f(\sigma)}[\mathrm{sw}(a)] \right).$$

For beyond-worst-case distortion, we will use this notion of a utility promise:

Definition 3. The utility profile u satisfies an *α-promise* on its maximum social welfare if there exists some alternative $a \in A$ for which $\mathrm{sw}(a) \geqslant \alpha \cdot n$.

3.2.1 Randomized Scoring Rules

Towards the goal of minimizing additive distortion, we find it compelling to define the following class of SCFs:

Definition 4. A *randomized scoring rule* (RSR) is an SCF given by a scoring vector $s \in \mathbb{R}_+^m - \mathbf{0}$. The aggregate scores S_a are calculated in the same way as for scoring rules, and then each alternative is chosen to be the winner with probability proportional to its total score. Let $\mathcal{RSR}$ denote the class of all such rules.

Just as the prominent rules Plurality, Borda Count, and Veto belong to the class of deterministic SRs, $\mathcal{RSR}$ also contains noteworthy rules. One is the harmonic scoring vector-based rule of Boutilier et al. [Bou+12] mentioned above, which is nearly optimal for multiplicative distortion. It is given by $s = (1 + H_m/m, 1/2 + H_m/m, \ldots, 1/m + H_m/m)$, where H_m is the m^{th} harmonic number. Another is Randomized Dictatorship, given by $s = (1, 0, \ldots, 0)$. Remarkably, RD incurs $O(3 - 2/n)$ multiplicative distortion in the metric setting, which is also nearly optimal [AP17].

In principle, there are many ways in which an aggregate score vector S can be converted to a probability distribution over A. Let us call $P : \mathbb{R}_+^m - \mathbf{0} \to \Delta_A$ a *probabilizer*, and focus on neutral probabilizers, i.e., the P which commute with all permutations of A. Then a *generalized RSR* consists of a pair (s, P) of scoring vector and neutral probabilizer; given σ it first computes S according to s, then samples from the distribution $P(S)$. Let $\mathcal{RSR}^*$ denote the class of all such SCFs. This is indeed a generalization, since any RSR given by s is a generalized RSR with the probabilizer $P(S)_a := S_a/\|S\|_1$ for all a, where $\|S\|_1 := \sum_{a \in A} S_a$. Note that $\mathcal{RSR}^*$ also contains all (otherwise deterministic) scoring rules that break ties uniformly at random. For a given scoring vector s the scoring rule is given by (s, P), where P returns the uniform distribution over $\arg\max_a S_a$. In fact, $\mathcal{RSR}^*$ also generalizes the "favorite only" rules which have received recent attention for metric distortion; in addition to RD these include the "proportional to squares" mechanism studied by Anshelevich and Postl [AP17] and the Random Oligarchy mechanism of Fain et al. [Fai+19].

3.3 Additive Distortion

We begin by proving a structural lemma which establishes that, for worst-case additive distortion, voter utilities may be assumed to be normalized without loss of generality. That is, even when voters have uniformly capped (instead of normalized) utilities, the worst case instances for additive distortion are when all voters have utilities summing to 1.

Lemma 5. *For each SCF f, the utility profile that witnesses the maximum of $\mathrm{dist}^+(f)$ is normalized, i.e., $\sum_a u_i(a) = 1$ for all voters $i \in [n]$.*

Proof. We begin by decomposing additive distortion.

$$\begin{aligned}
\mathrm{dist}^+(f) &= \max_{\sigma} \mathrm{dist}^+(f, \sigma) \\
&= \max_{\sigma} \max_{a^* \in A} \mathrm{dist}^+(f, \sigma, a^*) \\
&= \max_{\sigma} \max_{a^* \in A} \sum_{i \in [n]} dist_i^+(f, \sigma, a^*),
\end{aligned}$$

where

$$\mathrm{dist}^+(f, \sigma, a^*) := \max_{u \triangleright \sigma} \left[\mathrm{sw}(a^*) - \sum_{a \in A} \Pr[f(\sigma) = a] \; \mathrm{sw}(a) \right]$$

denotes the worst-case distortion given a profile σ and a randomized SCF f with respect to alternative a^* and

$$\mathrm{dist}_i^+(f, \sigma, a^*) := \max_{u_i \triangleright \sigma_i} \left[u_i(a^*) - \sum_{a \in A} \Pr[f(\sigma) = a] \; u_i(a) \right]$$

denotes the contribution of voter i toward $\mathrm{dist}^+(f, \sigma, a^*)$.

We now show that the utilities that maximize $\mathrm{dist}_i^+(f, \sigma, a^*)$ for every voter i are of the following form. Let k_i be the position of a^* in σ_i. Then, let $u_i(a) = 1/k_i$ if $\sigma_i(a) \leqslant k_i$ and $u_i(a) = 0$ otherwise. Note that this utility profile satisfies $\sum_a u_i(a) = 1$.

Now, let $f(a)$ represent the probability that f chooses alternative a. Furthermore, let k be the position of a^* in σ_i, and define $q := \sum_{a:a \succ_i a^*} f(a)$, $s := f(a^*)$, and $r := \sum_{a:a^* \succ_i a} f(a)$. In other words, q represents the total probability mass assigned by f to alternatives that appear before a^* in σ_i, r represents the probability assigned by f to a^*, and s represents the probability assigned by f to alternatives that appear after a^* in σ_i. By definition, we have that $q + r + s = 1$. Now, note that we can denote the contribution of voter i to the total additive distortion as

$$\begin{aligned}
\mathrm{dist}_i^+(f, \sigma, a^*) &:= \max_{u_i \triangleright \sigma_i} \left[u_i(a^*) - \sum_{a \in A} \Pr[f(\sigma) = a]\, u_i(a) \right] \\
&= \max_{u_i \triangleright \sigma_i} \Bigg[q \left(u_i(a^*) - (1/q) \sum_{a:a \succ_i a^*} u_i(a) f(a) \right) + r\left(u_i(a^*) - (1/r) u_i(a^*) f(a^*)\right) \\
&\qquad + s \left(u_i(a^*) - (1/s) \sum_{a:a^* \succ_i a} u_i(a) f(a) \right) \Bigg] \\
&= \max_{u_i \triangleright \sigma_i} \left[q \left(u_i(a^*) - (1/q) \sum_{a:a \succ_i a^*} u_i(a) f(a) \right) + s \left(u_i(a^*) - (1/s) \sum_{a:a^* \succ_i a} u_i(a) f(a) \right) \right]
\end{aligned}$$

because $f(a^*) = r$ by definition.

We will now argue that our choice of u_i maximizes this expression. For the first term, note that $u_i(a^*) - (1/q) \sum_{a:a \succ_i a^*} u_i(a) f(a) \leqslant 0$ because $(1/q) \sum_{a:a \succ_i a^*} u_i(a) f(a)$ captures the average utility of alternatives that are preferred to a^* in σ_i. Our choice of utilities results in $u_i(a^*) - (1/q) \sum_{a:a \succ_i a^*} u_i(a) f(a) = 0$, which is the maximum value this term can achieve. As for the second term, note that our choice of utilities independently maximizes $u_i(a^*)$ subject to utility constraints imposed by σ_i and simultaneously minimizes $\sum_{a:a^* \succ_i a} u_i(a) f(a)$. Because total additive distortion is additive over voters and our choice of utilities maximizes each voter's individual contribution toward additive distortion, our choice of utilities is maximizes additive distortion subject to the constraint that a^* is the true best alternative. $\square$

With this lemma in hand, we next show that, in the worst case, additive distortion can inevitably be quite large.

Claim 9. For all SCFs f and $m \geqslant 3$, $dist^+(f) \geqslant \frac{5}{18} \cdot n$.

Proof. We assume that $n = 3k$ for some positive integer k, take $m = 3$, and let the alternatives be a_1, a_2, and a_3. Consider the profile in which $n/3$ voters believe $a_1 \succ a_2 \succ a_3$, $n/3$ voters believe $a_2 \succ a_3 \succ a_1$, and $n/3$ voters believe $a_3 \succ a_1 \succ a_2$. Let p_i be the probability that f chooses a_i, and without loss of generality assume that $p_1 \geqslant p_2 \geqslant p_3$.

Now, let the first $n/3$ voters have utilities $u(a_1) = u(a_2) = u(a_3) = 1/3$; the second $n/3$ voters have utilities $u(a_2) = u(a_3) = 1/2$ and $u(a_1) = 0$; and the last $n/3$ voters have utilities $u(a_3) = 1$ and $u(a_1) = u(a_2) = 0$.

Therefore, we have

$$\begin{aligned} \mathrm{dist}^+(f, \sigma) &\geqslant \max_{a^* \in A} \mathrm{sw}(a^*) - \mathbb{E}_{a \sim f(\sigma)}[\mathrm{sw}(a)] \\ &= \frac{11}{18} \cdot n - \left(\frac{1}{9} \cdot p_1 + \frac{5}{18} \cdot p_2 + \frac{11}{18} \cdot p_3 \right) n \\ &\geqslant \frac{5}{18} \cdot n. && \text{(because } p_1 \geqslant p_2 \geqslant p_3\text{)} \end{aligned}$$

Note that this construction straightforwardly extends to any other $m > 3$. □

For deterministic rules, these symmetric instances offer even stronger lower bounds. The following claim was shown by Caragiannis et al. [Car+17] in a more general setting of choosing k winners out of m alternatives; for completeness, we reproduce the example for the single-winner setting below.

Claim 10 (Theorem 1 in [Car+17])**.** For all deterministic SCFs f and $m \geqslant 2$, $\mathrm{dist}^+(f) \geqslant \frac{1}{2} \cdot n$.

Proof. Let $m = 2$ and consider the profile σ with voters equally divided between $a_1 \succ a_2$ and $a_2 \succ a_1$. Suppose that f chooses a_2. If the first group has utilities $u(a_1) = 1$, $u(a_2) = 0$ and the second has $u(a_1) = 1/2$, $u(a_2) = 1/2$, then we have

$$\mathrm{dist}^+(f, \sigma) \geqslant \mathrm{sw}(a_1) - \mathrm{sw}(a_2) = \frac{1}{2} \cdot n.$$

This again extends to $m \geqslant 3$; for $m = 3$ the instance demonstrating Claim 9 also gives $\mathrm{dist}^+(f) \geqslant \frac{1}{2} \cdot n$. □

3.3.1 Two Alternatives

As a warm-up, we begin with the case when there are $m = 2$ alternatives. Here we may compute the optimal randomized SCF directly.

Claim 11. For $m = 2$ alternatives, the optimal SCF chooses each $a \in A$ with probability proportional to the number of voters ranking a first.

Proof. If there are two alternatives then the profile σ may be parameterized by a single variable λ which denotes the proportion of $[n]$ ranking alternative a_1 before alternative a_2. Therefore in this setting each randomized social choice function is given by some $p : [0,1] \to [0,1]$, where $p(\lambda)$ is the probability that the rule chooses a_1 given λ.

To begin, let us fix λ and derive the optimal p. For fixed λ and p, and observing that by assumption 8, $\mathrm{sw}(a_2) = n - \mathrm{sw}(a_1)$, the additive distortion is given by

$$\begin{aligned}\mathrm{dist}^+(\lambda, p) &= \max_{u \rhd \lambda} \ (\max(\mathrm{sw}(a_1), n - \mathrm{sw}(a_1)) \\ &\qquad - (p \cdot \mathrm{sw}(a_1) + (1-p) \cdot (n - \mathrm{sw}(a_1)))) \\ &= n \cdot \max_{u \rhd \lambda} \ \big(\max(w,\ 1-w) \\ &\qquad - (p \cdot w + (1-p) \cdot (1-w))\big),\end{aligned}$$

where $w := \mathrm{sw}(a_1)/n$ for convenience.

There are two cases to consider: when a_1 is the utility maximizer, and when a_2 is. Letting d_1^+ and d_2^+ denote the distortions in these cases, and simplifying, we have

$$\begin{aligned}d_1^+(\lambda, p) &= n \cdot \max_{u \rhd \lambda} \ \big((1-p)(2w-1)\big), \\ d_2^+(\lambda, p) &= n \cdot \max_{u \rhd \lambda} \ \big(p(1-2w)\big).\end{aligned}$$

In the first case, the u maximizing the expression for a given λ puts maximal utility on a_1, and so $w = (\lambda + 1)/2$. In the second, the maximizing u puts maximal weight on a_2, and so $w = \lambda/2$.

Therefore

$$d_1^+(\lambda, p) = n \cdot (1-p)\lambda,$$

$$d_2^+(\lambda, p) = n \cdot p(1-\lambda),$$

and so

$$d^+(\lambda, p) = \max(d_1^+(\lambda, p),$$

$$d_2^+(\lambda, p)) = n \cdot \max((1-p)\lambda,\ p(1-\lambda)).$$

Since the first term is monotonically decreasing in p and the second is monotonically increasing, this is minimized when they are equal, giving $p = \lambda$. $\square$

Note that since this is the optimal SCF, choosing an equally divided profile of voters yields a lower bound of $\mathrm{dist}^+(f) \geqslant 1/4$ for all SCFs f, recovering that of Caragiannis et al. [Car+17].

It is also noteworthy that this rule is in $\mathcal{RSR}$:

Observation 12. For $m = 2$ the optimal randomized rule belongs to $\mathcal{RSR}$, given by scoring vector $s^* = (1, 0)$.

For more than two alternatives, the problem of identifying optimal SCFs or even optimal RSRs becomes difficult.

3.3.2 Plurality and RD

When there are two alternatives, it is intuitive that the best deterministic rule should choose the alternative most frequently ranked first. In the class of deterministic rules, it turns out that this is always the best possible, as shown by Caragiannis et al. [Car+17] in the general setting of choosing k winners out of m alternatives.

Theorem 13 (Theorem 1 in [Car+17]). *Plurality is an optimal deterministic SCF, with additive distortion* $\frac{1}{2} \cdot n$.

Proof. Claim 10 establishes a lower bound of $\frac{1}{2} \cdot n$ for all deterministic SCFs, so it remains to show that the additive distortion incurred by Plurality is never worse.

For a given instance let a^* be the social welfare maximizing alternative, and suppose without loss of generality that Plurality chooses $a \neq a^*$. There are three types of voters to consider: voters ranking a first, voters ranking a^* first, and voters ranking something else first. Denote these voter types A, B, and C, and suppose that an α, β, γ proportion of the voters are of each type, respectively. Note that $\alpha + \beta + \gamma = 1$ and that $\alpha \geqslant \beta$, since Plurality chose a over a^*. Also observe that $u_A(a^*) \leqslant u_A(a)$ for all voters of type A, since they prefer a to a^*, and that $u_C(a^*) \leqslant 1/2$ for similar reasons. Therefore the contributions to the additive distortion from each group are

$$\begin{aligned} d_A^+ &= \alpha\, n \cdot \max_{u_A} \left[u_A(a^*) - u_A(a)\right] = 0 \\ d_B^+ &= \beta\, n \cdot \max_{u_B} \left[u_B(a^*) - u_B(a)\right] = \beta\, n \\ d_C^+ &\leqslant \gamma\, n \cdot \max_{u_C} \left[u_C(a^*) - u_C(a)\right] = \frac{1}{2}\gamma\, n, \end{aligned}$$

with inequality for d_C^+ because a^* may not be ranked second for all type C voters. These together yield

$$\mathrm{dist}^+(\mathit{Plurality}) = d_A^+ + d_B^+ + d_C^+ \leqslant (\beta + \gamma/2)n \leqslant \frac{1}{2}n,$$

since $\beta \leqslant \alpha$. ☐

The randomized analog to Plurality is Randomized Dictatorship, and Section 3.3.1 revealed that RD is the optimal SCF in the two alternative setting, attaining additive distortion $\frac{1}{4} \cdot n$ and significantly outperforming Plurality. One might reasonably hope that RD continues to significantly outperform Plurality for $m \geqslant 3$. However, we show that this is not the case:

Theorem 14. *RD has additive distortion* $\frac{1}{2}\left(1 - \frac{1}{m}\right) \cdot n$.

Proof. To start, let β_a be the proportion of voters ranking a first, and let a^* be the social welfare maximizing alternative. Note that RD chooses an alternative according to the distribution β.

Since additive distortion is linear in the voters, it is maximized when it is maximized for all voters individually, and each voter i contributes $u_i(a^*) - \mathbb{E}_{a\sim\beta}[u_i(a)]$.

For an voter ranking a^* first, we may therefore assume that they value $u(a^*) = 1$ and all other alternatives confer no utility. For voters i ranking $a \neq a^*$ first, their contribution to the distortion may be rewritten as

$$u(a^*)(1 - \beta_{a^*}) - \beta_a u_i(a) - \sum_{a' \neq a, a^*} \beta_{a'} u_i(a').$$

Given that this voter values a above a^*, the ranking and utility vector which simultaneously maximizes this first term and minimizes the others assigns $u_i(a) = u_i(a^*) = 1/2$, and $u_i(a') = 0$ for all other a'. Combining these assumptions and simplifying, we obtain

$$\begin{aligned}
\text{dist}^+(RD) &= \max_\beta \left(\text{sw}(a^*) - \sum_a n \cdot \beta_a \ \text{sw}(a) \right) \\
&= \max_\beta \ n \cdot \left(\frac{1}{2} \left(1 - \sum_a \beta_a^2 \right) \right) = \frac{1}{2} \left(1 - \frac{1}{m} \right) \cdot n
\end{aligned}$$

where the last inequality follows from the $\ell_1 - \ell_2$ inequality, which implies that the maximum is obtained when $\beta_a = 1/m$ for all $a \in A$. □

In fact, we must incorporate more than just voters' first choices in order to asymptotically improve upon $\frac{1}{2} \cdot n$. In the spirit of Gross, Anshelevich, and Xia [GAX17], who give a lower bound of $3 - 2/m$ on the distortion of favorite-only mechanisms in the metric setting, the proof of Theorem 14 can be modified in order to show that:

Claim 15. All generalized RSRs $(s, P) \in \mathcal{RSR}^*$ with $s = (1, 0, \dots, 0)$ have additive distortion at least $\frac{1}{2}\left(1 - \frac{1}{m}\right) \cdot n$.

Proof. Consider the tight instance given in the proof of Theorem 14, where an equal number of voters rank each alternative first, and all voters not ranking a^* first rank a^* second. Therefore the scores derived from $s = (1, 0, \dots, 0)$ are equal; $S_a = S_{a'}$ for all $a, a' \in A$. (Note that non-neutral "favorite-only" mechanisms do only worse than neutral mechanisms on this instance). Given this uniform aggregate score vector S, every neutral mechanism f returns the uniform distribution over A. This bounds its additive distortion as

$$\mathrm{dist}^{+}(f) \geqslant \frac{m+1}{2m} - \frac{1}{2m}\left(\frac{m-1}{m} + \frac{m+1}{m}\right) = \frac{1}{2}\left(1 - \frac{1}{m}\right) \cdot n,$$

as desired. □

Since RD is optimal within the class of favorite-only mechanisms, we continue the search for better rules among RSRs which score beyond voters' first choices.

3.3.3 An Asymptotically Optimal rule in $\mathcal{RSR}$

After the success in Section 3.3.1, we might hope to derive optimal RSRs for $m \geqslant 3$ directly. Unfortunately, the natural formulations of finding such optimal RSRs are nonconvex max-min optimization problems which we have been unable to solve. In order to render this problem tractable, we let a^* denote the alternative which maximizes social welfare, and we ignore the social welfare derived by choosing any alternative besides a^*. This provides an upper bound on the additive distortion of a given rule. We call this the *best-or-bust bound*, and we will use it repeatedly:

$$\mathrm{dist}^{+}(f) \leqslant \mathrm{sw}(a^*)(1 - \Pr[f(\sigma) = a^*]). \tag{3.1}$$

Informally speaking, this bound is apt because in the worst case and for large m, the non-a^* alternatives may evenly divide the remaining utility of the voters. In this case, the social welfare attained by choosing an alternative other than a^* is approximately $\frac{n-\mathrm{sw}(a^*)}{m}$, and so (3.1) is asymptotically tight for $\mathcal{RSR}$ as m becomes large.

We formulate the problem of finding the optimal RSR under (3.1) in (3.8) below, and prove that the scoring vector which optimizes this problem is $s^* = (25/33, 7/33, 1/33, 0, \ldots, 0)$ for all $m \geqslant 3$. Since it is the RSR which minimizes the upper bound (3.1), we call this the *Best-or-Bust* (BoB) rule.

This in turn implies the following theorem:

Theorem 16. *For all* $m \geqslant 3$, $\mathrm{dist}^+(BoB) \leqslant \frac{11}{27} \cdot n$. *It is furthermore a* $\left(1 - \frac{16}{27}\frac{1}{m-1}\right)^{-1} \leqslant \left(1 + \frac{1}{m-1}\right)$*-approximation to the optimal RSR for all* $m \geqslant 3$.

We now set about formulating the problem of finding the RSR which minimizes the right-hand side of eq. (3.1). For a given choice of $\alpha \in [0,1]$ and scoring vector $s = (s_1, \ldots, s_m)$ for which $\|s\|_1 = 1$, we may parameterize the solutions according to the optimum social welfare $\alpha \cdot n$ attainable. Let a^* be the alternative for which $\mathrm{sw}(a^*) = \alpha \cdot n$; we will then consider the worst-case probability that the RSR f^s selects a^*.

To this end, let x_i denote the proportion of voters $[n]$ who rank a^* i^{th}. Note that since rankings are assumed to be complete, $\|x\|_1 = 1$. Since f^s is a randomized scoring rule, and the probability of f^s choosing a^* is less than 1, in the worst case a^* has maximum utility possible given its vector of ranking proportions x. Therefore we may assume that $\mathrm{sw}(a^*) = n \cdot \sum_i \frac{1}{i} x_i$.

We may then identify the worst-case best-or-bust bound attained by s for given α by solving the linear program

$$D^+(s, \alpha) := \max \; \alpha - \alpha \sum_i s_i x_i \tag{3.2}$$

$$s.t. \qquad \sum_i \frac{1}{i} x_i = \alpha, \qquad x \in \Delta_{[m]}. \tag{3.3}$$

The objective (3.2) is (up to scaling by n) equal to the best-or-bust bound, since $\mathrm{sw}(a^*) = \alpha \cdot n$ and we have that $\mathrm{sw}(a^*) \Pr[f^s(\sigma) = a^*] = \frac{\alpha \cdot n}{\sum_i s_i} \sum_i s_i x_i = \alpha \cdot n \sum_i s_i x_i$, since $\|s\|_1 = 1$ by assumption. By optimizing over α as well, we may similarly characterize $\mathrm{dist}^+(f^s)$ as the optimal value of a quadratic program:

$$D^+(s) := \max \; D^+(s, \alpha) \tag{3.4}$$

$$s.t. \qquad 0 \leqslant \alpha \leqslant 1, \tag{3.5}$$

where eq. (3.5) captures that $\mathrm{sw}(a^*) \leqslant n$, since each voter's utilities are normalized to 1.

We might then hope to derive the optimal RSR directly, by solving $s^* := \arg\min_s D^+(s)$.

This takes the following form:

$$s^* := \arg\min D^+(s) \tag{3.6}$$
$$s.t. \qquad s \in \Delta_{[m]}.$$

Finally note that $\alpha = \alpha \sum_i x_i$; therefore the constraints (3.3) imply (3.5). We may also replace α with $\sum_i \frac{1}{i} x_i$. Taken together, these let us rewrite (3.4) as follows:

$$D^+(s) := \max \left(\sum_i \frac{1}{i} x_i\right)\left(\sum_i (1-s_i) x_i\right) \tag{3.7}$$
$$s.t. \quad x \in \Delta_{[m]}.$$

The general problem for which we hope to find optimal s^* is then

$$\mathcal{D}^+ := \min_s \max_x \left(\sum_i \frac{1}{i} x_i\right)\left(\sum_i (1-s_i) x_i\right) \tag{3.8}$$
$$s.t. \quad s, x \in \Delta_{[m]}.$$

Two Alternatives and the Harmonic RSR

As noted in Observation 12, the optimal SCF when $m = 2$ is the RSR given by $s = (1, 0)$. On the other hand, for $m = 2$ the optimal scoring vector for the formulation (3.8) is $s_2^* = (2/3, 1/3)$. This illustrates that the formulation above (and the best-or-bust bound) indeed only asymptotically capture the problem of identifying the optimal RSR for each m.

Incidentally, this $s_2^* = (2/3, 1/3)$ coincides with the harmonic scoring vector for $m = 2$. However this coincidence does not continue even for $m = 3$. Indeed the harmonic RSR incurs an additive distortion of at least $(1 - H_m^{-1}) \cdot n$, which is witnessed by the profile in which all voters value the same alternative with utility 1. Since $H_m^{-1} = o(1)$, this incurs additive distortion which is asymptotically the worst possible.

Three Or More Alternatives

One might expect that for each m there is a distinct randomized scoring rule with scoring vector s^*_m which optimizes (3.8). However it turns out that the same scoring vector s^* which (together with a suffix of trailing zeros) optimizes (3.8) for all $m \geqslant 3$ simultaneously.

Lemma 6. *For all $m \geqslant 3$, the unique optimal solution to* (3.8) *is the scoring vector $s^* = (25/33, 7/33, 1/33, 0, \ldots, 0)$, obtaining the optimum objective value $\frac{11}{27} \approx 0.407$.*

Proof. Given the scoring vector $s^* = (25/33, 7/33, 1/33, 0, \ldots, 0)$, we argue that $D^+(s^*) \leqslant 11/27$ for all $m > 3$. Recall that the objective of (3.7) is

$$\max_x \left(\sum_i \frac{1}{i} x_i\right)\left(\sum_i (1 - s_i) x_i\right).$$

For any x with $x_i > 0$ for any $i \geqslant 4$, the vector $x' := (x_1, x_2, x_3, x'_4, 0, \ldots, 0)$ for $x'_4 := \sum_{i \geqslant 4} x_i$ will only increase this objective, since the first factor in this product increases from x to x', and the second factor remains unchanged. Therefore we may assume without loss of generality that $m = 4$.

From here, we next show that we may in fact restrict our attention to x of the form $x = (x_1, x_2, x_3, 0, \ldots, 0)$. We will take a somewhat convoluted approach, which nevertheless seems to be the best available. To see this, consider the explicit form which the objective of (3.8) takes for s^* and $m = 4$. Substituting $x_4 = 1 - (x_1, +x_2 + x_3)$, it takes the form

$$O_4(x) := -\frac{1}{396}(3 + 9x_1 + 3x_2 + x_3)(-33 + 25x_1 + 7x_2 + x_3).$$

We may ask what $x \in \Delta_4$ maximizes this objective, which may be formulated as the following program:

$$\max \; O_4(x) \tag{3.9}$$

$$s.t. \quad x_1 + x_2 + x_3 \leqslant 1 \quad (\lambda) \tag{3.10}$$

$$x_1, x_2, x_3 \geqslant 0. \quad (\mu_i) \tag{3.11}$$

We would like to establish that $x_4 = 0$, which is equivalent to constraint (3.10) being tight at optimality.

Since any active constraints are independent and linear, any x^* maximizing (3.9) must satisfy the KKT conditions (note that it must obtain a maximum, since it is a continuous function on a compact set). The KKT conditions of (3.9) are the linear inequalities

$$-\frac{\partial O_4(x)}{\partial x_i} - \mu_i + \lambda = 0 \qquad i \in \{1, 2, 3\}$$

for $\mu_i \geqslant 0$, with strict equality for constraint i if $x_i^* > 0$. If (3.10) is loose (i.e. $x_4 > 0$) then $\lambda = 0$ by complementary slackness; since $\mu_i \geqslant 0$, this becomes

$$\begin{aligned}
\frac{1}{198}\left(111 - 225x_1 - 69x_2 - 17x_3\right) \leqslant 0 \qquad & (C_1)\\
\frac{1}{198}\left(39 - 69x_1 - 21x_2 - 5x_3\right) \leqslant 0 \qquad & (C_2)\\
\frac{1}{198}\left(15 - 17x_1 - 5x_2 - xx_3\right) \leqslant 0 \qquad & (C_3),
\end{aligned}$$

with constraint (C_i) tight if $x_i > 0$ at optimality. By checking the (x_1, x_2, x_3) which are feasible for both (3.9) and these linear constraints, we may determine that the only candidate optima of (3.9) for which $x_4 > 0$ have $x_1 = x_2 = 0$ or $x_1 = x_3 = 0$ or $x_2 = x_3 = 0$. Calculating the optima for these single-variable cases, we may confirm that all feasible candidate optima have objective at most $1/4 < 11/27$. Since we know that $11/27$ is attainable for this problem, it follows that we may assume that $x_4 = 0$ at optimality.

By showing that without loss of generality the x maximizing the inner problem (3.7) are of the form $x = (x_1, x_2, x_3, 0, \ldots, 0)$, we now know by Lemma 7 that s^* attains an objective of at most $11/27$ for (3.8) for any $m \geqslant 4$.

To confirm that s^* is indeed the optimal solution for all $m \geqslant 4$, we must finally argue that for any given $m \geqslant 4$ no s can do better—i.e. attain a strictly smaller value of $D^+(s)$. Informally this is because for any s obtaining some objective d on (3.8), its prefixes must obtain d or better on the lower-dimensional problems. This is because the maximizing x may place all of its weight on some prefix of the coordinates. Formally, suppose that s is an optimal solution to (3.8) for some $m \geqslant 4$. Let $s_{:3} := (s_1, s_2, s_3, 0, \ldots, 0)$, and let $\overline{s_{:3}}$ denote $s_{:3}/\|s_{:3}\|_1$ (and if $\|s_{:3}\|_1 = 0$ then s does truly terribly, and isn't worth worrying about). It is clear that for fixed $x = (x_1, x_2, x_3, 0, \ldots, 0)$ the objective of (3.8) is weakly greater at $s_{:3}$ than it is at $\overline{s_{:3}}$; since x is feasible to begin with and $\overline{s_{:3}}$ is feasible for (3.8) when $m = 3$, it then follows from Lemma 7 that the objective value of (3.8) at s is at least $11/27$. □

This candidate optimizer s^* of (3.8) was first identified via computer-assisted search. We now prove that it is optimal.

The proof proceeds in two stages. We begin by restricting the inner problem (3.7) to a new problem $\overline{D}^+(s)$; this gives a corresponding relaxation of the outer problem (3.8). We then argue that, for $m = 3$, if $\overline{D}^+(s) \leqslant \frac{11}{27}$ then $s = s^*$. Given this s^*, we demonstrate that the objective does not increase when we move from the restricted inner problem to the general inner problem (3.7):

Lemma 7. *For $m = 3$, the unique optimal solution to (3.8) is the scoring vector $s^* = (25/33, 7/33, 1/33)$, obtaining objective value $\frac{11}{27}$.*

Proof. Consider the restricted inner problem $\overline{D}^+(s)$ given by

$$\overline{D}^+(s) := \max \left(\sum_i \frac{1}{i} x_i\right)\left(\sum_i (1 - s_i) x_i\right) \tag{3.12}$$

$$s.t. \qquad x = \begin{cases} (\beta, 1-\beta, 0) & \text{for } \beta \in [0,1] \quad \text{or} \\ (\gamma, 0, 1-\gamma) & \text{for } \gamma \in [0,1] \end{cases}$$

this is a restriction of (3.7) since the x which are feasible for (3.12) are also feasible for (3.7).

Writing $s = (s_1, s_2, 1 - s_1 - s_2)$, the objective of (3.12) for each form of x is

$$\max_{\beta \in [0,1]} \frac{1}{2}(\beta + 1)\left(1 - (s_1 - s_2)\beta - s_2\right), \tag{3.13}$$

$$\max_{\gamma \in [0,1]} \frac{1}{3}(2\gamma + 1)\left(s_1 + s_2 + \gamma(1 - 2s_1 - s_2)\right). \tag{3.14}$$

If $\overline{D}^{+}(s) \leqslant 11/27$ then both (3.13) $\leqslant 11/27$ and (3.14) $\leqslant 11/27$. Taking derivatives, we find that the β which maximizes the argument of (3.13) in terms of s_1 and s_2 is either $\beta = 0$, $\beta = 1$, or $\beta = \frac{1-s_1}{2(s_1-s_2)}$ (provided that $0 \leqslant \frac{1-s_1}{2(s_1-s_2)} \leqslant 1$). Requiring that $\overline{D}^{+}(s) \leqslant 11/27$ in each of these cases yields the inequalities

$$1 - s_1 \leqslant 11/27, \tag{3.15}$$

$$\frac{1-s_2}{2} \leqslant 11/27, \tag{3.16}$$

$$\frac{(1+s_1-2s_2)^2}{8(s_1-s_2)} \leqslant 11/27. \tag{3.17}$$

These first two constraints imply that $\frac{1-s_1}{2(s_1-s_2)} \geqslant 0$; we therefore require (3.17) when $\frac{1-s_1}{2(s_1-s_2)} \leqslant 1$.

The second partial formulation (3.14) for the critical values $\gamma = 0, \gamma = 1$, and $\gamma = \frac{1+s_2}{4(2s_1+s_2-1)}$ similarly yields the inequalities

$$\frac{s_1+s_2}{3} \leqslant 11/27 \tag{3.18}$$

$$\frac{(4s_1+3s_2-1)^2}{24(2s_1+s_2-1)} \leqslant 11/27 \tag{3.19}$$

as well as (3.15) again. The inequality (3.18) is trivial, since we assume that $w \geqslant 0$ and $\sum_i s_i = 1$. We again require (3.19) so long as $0 \leqslant \frac{1+s_2}{4(2s_1+s_2-1)} \leqslant 1$; fortunately $0 \leqslant \frac{1+s_2}{4(2s_1+s_2-1)}$ follows from (3.15), and $\frac{1+s_2}{4(2s_1+s_2-1)} \leqslant 1$ follows from (3.15) and (3.16) together.

There are now two cases to consider. First, we argue that the inequalities $\frac{1-s_1}{2(s_1-s_2)} \geqslant 1$ and (3.19) are not simultaneously satisfiable; we may therefore assume that $\frac{1-s_1}{2(s_1-s_2)} \leqslant 1$ and consequently that (3.17) holds. The first inequality is the half-plane

$$3s_1 - 2s_2 \leqslant 1$$

(since (3.15) implies that $s_1 > s_2$), and so in order to argue incompatibility it suffices to consider the optimization problem, where the constraint is a rearrangement of (3.19):

$$\begin{aligned} &\min_s \; 3s_1 - 2s_2 \\ \text{s.t.} \quad &(4s_1 + 3s_2 - 1)^2 - \frac{88}{9}(2s_1 + s_2 - 1) \leqslant 0. \end{aligned}$$

A minimum objective value greater than one will demonstrate unsatisfiability. By the change of variable $x = 4s_1 + 3s_2 - 1$ and $y = 3s_1 - 2s_2$, this becomes

$$\begin{aligned} &\min \; y \\ \text{s.t.} \quad &\frac{153}{176}x^2 - \frac{7}{2}x + 5 \leqslant y, \end{aligned}$$

which has a minimum value of $226/153 > 1$.

We have shown that (3.17) and (3.19) must hold if the restricted problem (3.12) is to have a solution with objective at most 11/27. We finally argue that s^* is the unique s satisfying both (3.17) and (3.19). As above, (3.15) implies that the denominators of (3.17) and (3.19) are positive, and we may therefore rewrite them as

$$(1 + s_1 - 2s_2)^2 - \frac{88}{27}(s_1 - s_2) \leqslant 0 \tag{3.20}$$

$$(4s_1 + 3s_2 - 1)^2 - \frac{88}{9}(2s_1 + s_2 - 1) \leqslant 0. \tag{3.21}$$

We will show that (3.20) and (3.21) are simultaneously satisfied at the single point $(s_1^*, s_2^*) = (25/33, 7/33)$ by way of the line ℓ given by $2s_1 + 7s_2 = 3$, which separates their feasible regions. Through the changes of variables $r = -s_1 + 2s_2$, $t = 2s_1 + s_2$, and $u = 4s_1 + 3s_2$, $v = -3s_1 + 4s_2$, (3.20) and (3.21) become

$$\frac{3}{88}(45 - 2r + 45r^2) \leqslant t \tag{3.22}$$

$$\frac{1}{176}(-2425 + 1418u - 225u^2) \geqslant v, \tag{3.23}$$

while ℓ takes the form $12r + 11t = 15$ and $29u + 22v = 75$.

First, if (3.20) and $2s_1 + 7s_2 \leqslant 3$ then equivalently (3.22) and $t \leqslant 15/11 - 12/11 \cdot s$. These imply that

$$\frac{3}{88}(45 - 2r + 45r^2) \leqslant 15/11 - 12/11 \cdot r$$
$$\frac{15}{88}(1 + 3r)^2 \leqslant 0$$
$$r = -1/3 \qquad t = 19/11$$
$$s_1 = 25/33 \qquad s_2 = 7/33.$$

Similarly, if (3.21) and $2s_1 + 7s_2 \geqslant 3$ then equivalently (3.23) and $v \geqslant 75/22 - 29/22 \cdot u$. These imply that

$$\frac{1}{176}(-2425 + 1418u - 225u^2) \geqslant 75/22 - 29/22 \cdot u$$
$$-\frac{25}{176}(11 - 3u)^2 \geqslant 0$$
$$u = 11/3 \qquad v = -47/33$$
$$s_1 = 25/33 \qquad s_2 = 7/33.$$

This demonstrates that the feasible regions of (3.20) and (3.21) are contained within each of the two half-planes bordered by ℓ. Since both (3.20) and (3.21) intersect with ℓ only at s^*, we conclude that s^* is the unique point satisfying both (3.20) and (3.21).

Given that s^* is the unique minimizer of (3.12), it remains to argue that it is the unique optimizer of (3.8) for $m = 3$. Writing $x = (x_1, x_2, 1 - x_1 - x_2)$, the objective of $D^+(s^*)$ takes the explicit form

$$\begin{aligned} D^+(s^*) &= \max_x \ \frac{1}{99}(2 + 4x_1 + x_2)(16 - 12x_1 - 3x_2) \\ &= \max_x \ \frac{11}{27} - \frac{1}{33}\left(4x_1 + x_2 - \frac{5}{3}\right)^2, \end{aligned}$$

which demonstrates that no x obtains a larger objective on s^* even when we relax the inner problem from (3.12) to (3.7) when $m = 3$. $\square$

We finally show that this s^* does not incur a larger objective even for $m > 3$, and that for fixed $m > 3$ no other s can do better; this demonstrates that s^* optimizes (3.8) for all $m \geqslant 3$ simultaneously, proving Lemma 6. The proof of Theorem 16 then follows:

Proof of Theorem 16. Let f^* denote this RSR with score s^*. We note first that f^* has additive distortion $d^+(f^*) \leqslant \frac{11}{27} \cdot n$ for all $m \geqslant 3$, by the best-or-bust bound eq. (3.1). The additive distortion of a rule f is given by

$$\mathrm{dist}^+(f) = \max_{\sigma} \max_{u \triangleright \sigma} \left(\mathrm{sw}(a^*) - \sum_{a \in A} \Pr[f(\sigma) = a] \; \mathrm{sw}(a) \right),$$

while we have argued above that the objective of (3.7) takes the form

$$D^+(s) = \max_{\sigma} \max_{u \triangleright \sigma} \big(\mathrm{sw}(a^*) - \Pr[f^s(\sigma) = a^*] \; \mathrm{sw}(a^*) \big) \tag{3.24}$$

for f^s an RSR, and where a^* denotes the social-welfare-maximizing alternative given u. Therefore it is clear that for any fixed RSR f^s the objective (3.24) is larger than the additive distortion of the rule. In particular, this is also true for the rule f^* which minimizes (3.24) over all s.

Next we argue that all RSRs f^s satisfy $\mathrm{dist}^+(f^s) \geqslant \left(1 - \frac{16}{27}\frac{1}{m-1}\right) \frac{11}{27} \cdot n$. To see this, fix m and choose any scoring vector s with corresponding rule f^s. For every (rational) x as in (3.7), fix some alternative a^* and consider a profile (u, σ) for which

- an x_i proportion of voters rank a^* in position i,
- for each i, an equal number of voters who rank a^* in position i rank each alternative $a \neq a^*$ in each position $i' \neq i$, and
- u maximizes $\mathrm{sw}(a^*)$ subject to consistency with σ.

Since utilities may tie, this third condition is well-defined and determines all utilities. It also guarantees that a^* is a (not necessarily unique) maximizer of social welfare for (u, σ). Finally, for the purposes of additive distortion it is without loss of generality to assume u of this form given a^* and σ, since increasing $\mathrm{sw}(a^*)$ for fixed σ only increases the maximized inner term $\mathrm{sw}(a^*) - \mathbb{E}_{a\sim f(\sigma)}[\mathrm{sw}(a)]$.

Let $\Pr[a] := \Pr[f^s(\sigma) = a]$ for convenience; then for each x we have

$$d^+(f^s, \sigma) \geqslant \mathrm{sw}(a^*) - \sum_a \Pr[a]\ \mathrm{sw}(a),$$

and regrouping and taking the maximum over all profiles σ derived from x,

$$\begin{aligned}
d^+(f^s) &\geqslant \max_x \left(\mathrm{sw}(a^*)(1 - \Pr[a^*]) - \sum_{a \neq a^*} \Pr[a]\ \mathrm{sw}(a) \right) \\
&= \max_x \left(\mathrm{sw}(a^*)(1 - \Pr[a^*]) - (1 - \Pr[a^*])\, \frac{n - \mathrm{sw}(a^*)}{m-1} \right) && (3.25) \\
&= \max_x \left(\mathrm{sw}(a^*)(1 - \Pr[a^*]) \left(1 + \frac{1}{m-1}\right) - (1 - \Pr[a^*])\, \frac{n}{m-1} \right), && (3.26)
\end{aligned}$$

where (3.25) follows from the construction of σ and u given x, so that the probability of $f^s(\sigma) = a$ is equal for all $a \neq a^*$, together with the fact that the total utility is n. This first term is precisely the objective of (3.8), and so by Lemma 6 we have

$$\begin{aligned}
&\geqslant \max_x \left(\left(1 + \frac{1}{m-1}\right) \frac{11}{27} \cdot n - (1 - \Pr[a^*])\, \frac{n}{m-1} \right) \\
&\geqslant \left(1 + \frac{1}{m-1}\right) \frac{11}{27} \cdot n - \frac{n}{m-1} \\
&= \left(1 - \frac{16}{27} \frac{1}{m-1}\right) \frac{11}{27} \cdot n.
\end{aligned}$$

Since this last term is an upper bound on the additive distortion of the rule for s^*, and since this holds for all positional scoring rules, we may take the right-hand side to be the optimal PSR for m, and then invert to obtain that f^* is a $\left(1 + \frac{1}{m-1}\right)$-approximation to the optimal RSR, for all $m \geqslant 3$. □

3.3.4 An Additive Distortion Instance-Optimal SCF

Although the RSR derived in Section 3.3.3 is asymptotically optimal within $\mathcal{RSR}$, we do not anticipate that it is optimal among all SCFs, even asymptotically. In pursuit of better rules, we turn to instance-optimal SCFs.

The instance-optimal SCF from the perspective of additive distortion, for any given profile σ, mimics the minimizer of $\text{dist}^+(f, \sigma)$ over all SCFs f (which for fixed σ are probability distributions over A). In particular,

$$\begin{aligned}\text{AddOpt}(\sigma) &:= \arg\min_f \text{dist}^+(f, \sigma) \\ &= \min_{p \in \Delta_A} \max_{u \rhd \sigma} \left(\max_{a^* \in A} \text{sw}(a^*) - \mathbb{E}_{a \sim p}[\text{sw}(a)] \right).\end{aligned}$$

We make use of Lemma 5 to show the following, which we empirically test in Section 7.5:

Theorem 17. *For any profile σ, Algorithm 2 computes the distribution over A which minimizes (expected) additive distortion in polynomial time.*

Proof. We first leverage the approach in Lemma 5 in order to drastically simplify the optimization problem by restricting the space of worst-case utilities we must consider.

The worst-case additive distortion on a profile σ can be written as

$$\begin{aligned}\text{dist}^+(\sigma) &= \min_f \text{dist}^+(f, \sigma) \\ &= \min_f \max_{a^* \in A} \text{dist}^+(f, \sigma, a^*).\end{aligned}$$

In other words, this can be then decomposed into independently finding the best f for m different optimization problems each corresponding to the case that a particular $a^* \in A$ is the true best alternative and then taking the minimum over these f solutions. However, we showed in Lemma 5 that no matter the choice of f, for each a^* we can explicitly write down the utilities consistent with σ that maximize $\text{sw}(a^*) - \mathbb{E}_{a \sim f(\sigma)}[\text{sw}(a)]$ for any f, which drastically simplifies the problem of finding instance-optimal solutions.

In particular, given a profile σ, for each $a \in A$, compute u^a as defined in Lemma 5 as follows. For each voter i, let k_i be the position of a in σ_i. Then, for all $a' \in A$, let $u_i^a(a') = 1/k_i$ if $\sigma_i(a') \leqslant k_i$ and $u_i^a(a') = 0$ otherwise. Next, compute $\text{sw}(a, a') := \text{sw}_{u^a}(a')$ for all $a, a' \in A$. Now, letting $w_a := (\text{sw}(a, a_1), \ldots, \text{sw}(a, a_n))$, solve the following linear program to minimize

Algorithm 2 AdditiveOptimal

Require: Ranking $\sigma \in \mathcal{S}_A^n$
Ensure: Distribution $p^* \in \Delta_m$ minimizing $\text{dist}^+(p, \sigma)$
1: **for** $a, b \in A$ **do**
2: $\quad w_a^b \leftarrow \sum_i 1/\sigma_i^{-1} \cdot \mathbb{1}\{b \succcurlyeq_i a\}$
3: $w_a \leftarrow (w_a^b)_{b \in A}$ for each $a \in A$
4: $p^* \leftarrow \arg\min_p \{D : w_a^a - p^T w_a \leqslant D \; \forall a \in A, p \in \Delta_A\}$
5: **return** p^*

distortion over all vectors of probabilities p, which correspond to social choice functions f:

$$\begin{aligned} & \min D \\ & \text{sw}(a, a) - p^T w_a \leqslant D && \forall a \in A \\ & \sum_{a \in A} p_a = 1 \\ & p_a \geqslant 0 && \forall a \in A \end{aligned}$$

□

We note that it is possible to extend the dual-based approach of Boutilier et al. [Bou+12] in order to derive a large linear program that calculates instance-optimal additive distortion. However this approach is much simpler, thanks in part to the structural observation made in Lemma 5 about worst-case utilities.

3.4 Distortion With a Promise

We began by motivating additive distortion based on the observation that traditional distortion may not be the best metric when the maximum social welfare attainable is *potentially* quite large. For a given profile σ, additive distortion provides a soft sort of guarantee with respect to maximum attainable welfare, in the following sense: for $u, u' \rhd \sigma$ where the maximum attainable welfare is higher under u than u', additive distortion measures the extent to which a SCF provides simultaneous guarantees for both utility profiles *simultaneously*, requiring additively better guarantees for u.

In this section we instead suppose we are *promised* that there exists an alternative with high social welfare, and ask about distortion subject to this promise. We define α-promise distortion as the distortion over all profiles (σ, u) for which u satisfies the α-promise of Definition 3:

Definition 5. For $\alpha \in [0,1]$, the *α-promise distortion* of a rule f is given by

$$\mathrm{dist}_\alpha(f) := \max_{\sigma} \max_{\substack{u \triangleright \sigma \\ u \in U_\alpha}} \mathrm{dist}(f, \sigma),$$

where U_α is the collection of preference profiles u satisfying the α-promise.

Since α-promise multiplicative distortion and additive distortion both address the high-stakes setting, our first result interrelates the two:

Claim 18. For any randomized SCF f,

- If $\mathrm{dist}^+(f) \leqslant \beta \cdot n$, then $\mathrm{dist}_\alpha(f) \leqslant \frac{\alpha}{\alpha - \beta}$.
- If $\mathrm{dist}_\alpha(f) \leqslant \gamma$, then $\mathrm{dist}^+(f) \leqslant \max(\alpha \cdot n,\, n - n/\gamma)$.

In the promise setting, we might also hope to circumvent the relatively low-welfare $\Omega(\sqrt{m})$ lower bound given in [Bou+12]. Indeed, the lower bound instance in [Bou+12] translates directly into a lower bound on distortion with an α-promise:

Theorem 19. *For any randomized SCF f,*

$$\mathrm{dist}_\alpha(f) = \Omega(\min\{\sqrt{m}, {}^1\!/\!_\alpha\}).$$

A slight modification of the *Stable Lottery Rule* f_{SLR} introduced by Ebadian et al. [Eba+22a] yields a matching upper bound for all $\alpha \geqslant 1/\sqrt{m}$. In particular, the modified rule samples alternatives from the stable lotteries of Cheng et al. [Che+20], which are distributions over committees of size $2/\alpha$.

Theorem 20. *There is an SCF ℓ_α with* $\mathrm{dist}_\alpha(\ell_\alpha) = O\left(\frac{1}{\alpha}\right)$.

We take as our point of departure the analysis that Ebadian et al. [Eba+22a] give for their Stable Lottery Rule f_{SLR}. For fixed $k \in \mathbb{N}$, a *stable lottery* is a distribution $\mathcal{X}$ over committees

$X \subseteq A$, of size $|X| = k$ for which the expected number of agents preferring any fixed alternative a^* to any $a \in X$ is small. In particular, for a fixed preference profile σ, the lottery $\mathcal{X}$ is *stable* if for all $a^* \in A$, $\Pr_{i \in N} \Pr_{X \sim \mathcal{X}}[a^* \succ_i X] \leqslant \frac{1}{k}$, where $a^* \succ_i X$ denotes that i ranks a^* ahead of all $a \in X$. Such stable lotteries are shown to exist for all σ and k in Theorem 1 of Cheng et al. [Che+20].

We will now define the rule ℓ_α.

Definition 6. For α with $1/\alpha \in \mathbb{N}$, the randomized SCF ℓ_α identifies some lottery $\mathcal{X}$ over committees of size $k = 2/\alpha$ which is *stable* for the input profile σ. It then samples a committee $X \sim \mathcal{X}$ and finally returns an alternative $a \sim X$ drawn uniformly at random.

Although we do not emphasize the efficient computability of ℓ_α in Theorem 20, Cheng et al. [Che+20] show that stable lotteries sufficient for our purposes can be calculated in polynomial time, and so ℓ_α can be efficiently implemented.

Proof of Theorem 20. To begin, assume that $1/\alpha \in \{2, 3, 4, \ldots\}$; all $\alpha \in [0, 1]$ are within a constant factor of such an $\alpha' \leqslant \alpha$, and in such cases an α-guarantee implies an α'-guarantee, and so this is without loss of generality.

As usual, let a^* denote the social-welfare-maximizing alternative. For a fixed committee X, let V_X denote the voters $i \in N$ for which $a^* \succ_i X$, and let $\overline{V}_X = N \setminus V_X$ denote its complement. By the definition of social welfare, for all X we have that

$$sw(a^*) = \sum_{i \in N} u_i(a^*) = \sum_{i \in V_X} u_i(a^*) + \sum_{i \in \overline{V}_X} u_i(a^*),$$

and so

$$sw(a^*) = \mathbb{E}_{X \sim \mathcal{X}} \left[\sum_{i \in V_X} u_i(a^*) + \sum_{i \in \overline{V}_X} u_i(a^*) \right]. \tag{3.27}$$

Since $u_i(a^*) \leqslant 1$, by the stability of $\mathcal{X}$ we have that

$$\mathbb{E}_{X \sim \mathcal{X}} \left[\sum_{i \in V_X} u_i(a^*) \right] \leqslant \mathbb{E}_{X \sim \mathcal{X}} \left[|V_X| \right] \leqslant \frac{n}{k}. \tag{3.28}$$

Using the α-promise that $sw(a^*) \geqslant \alpha \cdot n = 2\frac{n}{k}$, we find that

$$\begin{aligned}
\frac{1}{2} \cdot sw(a^*) &\leqslant sw(a^*) - \frac{n}{k} && (3.29)\\
&\leqslant sw(a^*) - \mathbb{E}_{X \sim \mathcal{X}}\left[\sum_{i \in V_X} u_i(a^*)\right] && (3.30)\\
&= \mathbb{E}_{X \sim \mathcal{X}}\left[\sum_{i \in \overline{V}_X} u_i(a^*)\right] && (3.31)\\
&\leqslant \mathbb{E}_{X \sim \mathcal{X}}\left[\sum_{i \in \overline{V}_X} k \cdot \mathbb{E}_{a \sim X} u_i(a)\right] && (3.32)\\
&\leqslant k \cdot \mathbb{E}_{X \sim \mathcal{X}}\left[\sum_{i \in N} \mathbb{E}_{a \sim X} u_i(a)\right] && (3.33)\\
&= \frac{2}{\alpha} \cdot \mathbb{E}_{a \sim \ell_\alpha(\sigma)}\left[sw(a)\right]. && (3.34)
\end{aligned}$$

Here (3.30) follows from (3.28), (3.31) follows from (3.27), (3.32) follows because if $i \in \overline{V}_X$ then i has at least as much utility for some $a \in X$ as for a^*, and this a has probability $1/k$ of being sampled from X, and (3.34) follows from the definitions of k and ℓ_α.

In conclusion, we have shown that $sw(a^*) \leqslant \frac{4}{\alpha}\mathbb{E}_{a \sim \ell_\alpha(\sigma)}\left[sw(a)\right]$, and therefore $\text{dist}_\alpha(\ell_\alpha) = O(\frac{1}{\alpha})$. □

3.4.1 Additive Distortion With a Promise

We now turn to *α-promise additive distortion*, which is defined analogously to Definition 5. In this subsection, we are focused on the *robustness* of each rule, where we ask how the additive distortion guarantees degrade with the promise α. Intuitively, this asks "How well do these rules perform when the winner is clear?" We consider $\alpha \geqslant 1/2$; for all $\alpha < 1/2$ we know the additive distortion is at most α.

We begin with three deterministic scoring rules:

- The *Plurality Rule* (f_{Plur}) is a deterministic scoring rule with score vector $s = (1, 0, \ldots, 0)$.

- The *Harmonic Rule* (f_{Harm}) is a deterministic positional scoring rule with score vector $s = (1, 1/2, \ldots, 1/m)$.
- The *Borda Rule* (f_{Borda}) is a deterministic positional scoring rule with score vector $s = (m-1, m-2, \ldots, 0)$.

We begin by showing that Plurality and the Harmonic Rule are robust for $\alpha \geqslant 3/4$, but once $\alpha < 3/4$ their additive distortion becomes as bad as the worst case:

Claim 21. For the Plurality Rule (f_{Plur}),

$$\mathrm{dist}^{+}_{\alpha}\left(f_{Plur}\right) = \begin{cases} 0 & \text{for } \alpha \geqslant 3/4 \\ 1/2 & \text{for } \alpha < 3/4. \end{cases}$$

Proof. Suppose for now that $\mathrm{sw}(a^*) = \alpha$ exactly. We start by examining the first case. When $\alpha \geqslant 3/4$, this means that at least $n/2$ voters rank a^* first, and therefore, if we break ties in our favor, a^* is the winner under Plurality and the additive distortion is 0.

In the second case, consider a profile in which a $\frac{1}{2} - \epsilon$ fraction of voters rank a^* first (and value it at 1) and the other $\frac{1}{2} + \epsilon$ fraction of voters ranks an alternative $a \neq a^*$ first. This means that a is the winner under Plurality.

Now, let V_a be the set of voters who rank $a \neq a^*$ first. Of these voters, let a β fraction of them rank $a \succ a^* \succ \ldots$, where they value a^* and a at $1/2$. The other $1 - \beta$ fraction of voters in V_a are indifferent between all alternatives and value each of them at $1/m$. It is easy to verify that setting

$$\beta = \frac{\alpha - \frac{1}{2} - \frac{1}{2m}}{\frac{1}{2} - \frac{1}{m}}$$

suffices to ensure that a^* confers $\alpha \cdot n$ utility:

$$\frac{1}{2} + \frac{1}{2} \cdot \beta + \left(\frac{1}{2} - \beta\right) \cdot \frac{1}{m} = \alpha.$$

Plugging this value of β into the expression for additive distortion yields

$$\begin{aligned}
\mathrm{dist}_{\alpha}^{+} &= \alpha - \left(\frac{1}{2} \cdot \beta + \left(\frac{1}{2} - \beta\right) \frac{1}{m}\right) \\
&= \alpha - \left(\frac{1}{2} \cdot \frac{\alpha - \frac{1}{2} - \frac{1}{2m}}{\frac{1}{2} - \frac{1}{m}} + \left(\frac{1}{2} - \frac{\alpha - \frac{1}{2} - \frac{1}{2m}}{\frac{1}{2} - \frac{1}{m}}\right) \frac{1}{m}\right) \\
&= 1/2. && \text{(after simplifying)}
\end{aligned}$$

It remains only to relax our original assumption that $\mathrm{sw}(a^*) = \alpha$ exactly. Since the above expression is decreasing monotonically in $\mathrm{sw}(a^*)$, for $\mathrm{sw}(a^*) \geqslant 1/2$ we may assume that $\mathrm{sw}(a^*) = \alpha$ as above.

Finally, by Theorem 13, this bound is tight. □

Claim 22. For the Harmonic rule (f_{Harm}),

$$\mathrm{dist}_{\alpha}^{+}(f_{Harm}) \begin{cases} = 0 & \text{for } \alpha \geqslant 3/4 \\ \geqslant 1/2 & \text{for } \alpha < 3/4. \end{cases}$$

Proof. Suppose that $\mathrm{sw}(a^*) = \alpha$.

In the first case, we must show that for any $\alpha \geqslant 3/4$, a^* will be selected by f_{Harm}. Let y_j represent the proportion of voters who rank a^* in position j. Note that the maximum score of any other alternative $a \neq a^*$ is $(1 - y_1) + \frac{1}{2} \cdot y_1$ because each voter who does not rank a^* first can rank a first, and each voter who ranks a^* first can rank a second. Furthermore, let $y_1^{\downarrow}(\alpha)$ represent the minimum value of y_1 such that the α promise holds, i.e., it must be the case that $\sum_j y_j \cdot \frac{1}{j} \geqslant \alpha$ because the maximum utility a^* can receive in position j is exactly $1/j$. Note that this value of $y_1^{\downarrow}(\alpha)$ also maximizes a's score. However, for all $\alpha \geqslant 3/4$, we can see that $y_1^{\downarrow}(\alpha) \geqslant 1/2$, meaning that $s(a) \leqslant \left(1 - \frac{1}{2}\right) + \frac{1}{2} \cdot \frac{1}{2} \leqslant \frac{3}{4}$, whereas $s(a^*) = \sum_j y_j \cdot \frac{1}{j} \geqslant \alpha$ due to the α promise guarantee. Therefore, for all $\alpha \in [3/4, 1]$, f_{Harm} will select a^* (breaking ties in our favor when $\alpha = 3/4$).

For the case of $2/3 \leqslant \alpha < 3/4$, consider the profile in which a $2(\alpha - 1/2)$ fraction of voters rank a^* first (with utility 1) and another alternative a in second (with utility 0), and a $2(1 - \alpha)$

fraction of voters rank a^* second (with utility 1/2) and a first (with utility 1/2). Note that $u(a^*) = 2(\alpha - 1/2) \cdot n + \frac{1}{2} \cdot 2(1-\alpha) \cdot n = \alpha \cdot n$ and $u(a) = \frac{2(1-\alpha)}{2} \cdot n = (1-\alpha) \cdot n$.

With respect to scores, $s_{Harm}(a) = \frac{1}{2} \cdot 2(\alpha - 1/2) + 2(1-\alpha) = 3/2 - \alpha$, and $s_{Harm}(a^*) = 2(\alpha - 1/2) + \frac{1}{2} \cdot 2(1-\alpha) = \alpha$. For all $\alpha \in [2/3, 3/4)$, we indeed see that $s_{Harm}(a) > s_{Harm}(a^*)$ and therefore a is chosen, yielding an additive distortion of $\alpha - (1-\alpha) = 2\alpha - 1$.

For the case of $1/2 \leqslant \alpha < 2/3$, consider the profile in which an α fraction of voters rank a^* first (with utility 1) and a second (with utility 0), and the remaining $1-\alpha$ fraction of voters ranks a first (with utility $\frac{1}{m-1}$) and a^* last (with utility 0). Note that $u(a^*) = \alpha \cdot n$ and $u(a) = \frac{1-\alpha}{m-1} \cdot n$.

Turning to scores, $s_{Harm}(a) = \alpha \cdot \frac{1}{2} + (1-\alpha) = 1 - \frac{\alpha}{2}$, whereas $s_{Harm}(a^*) = \alpha + \frac{1-\alpha}{m}$. It is easy to verify that $s_{Harm}(a) > s_{Harm}(a^*)$ for all $\alpha \in [1/2, 2/3)$, so a will be selected, yielding an additive distortion of $\alpha - \frac{1-\alpha}{m-1}$.

Regardless of m, for $\alpha \in [1/2, 3/4]$ the two-alternative profile above with a utility promise of $3/4 \cdot n$ gives an additive distortion of $1/2 \cdot n$. □

Claim 23. For the Borda rule (f_{Borda}),

$$\operatorname{dist}^+_\alpha(f_{Borda}) \begin{cases} = 0 & \text{for } \alpha \geqslant \frac{m-1}{m} \\ \geqslant \frac{m-1}{m} - \frac{1}{m^2} & \text{for } \alpha < \frac{m-1}{m}. \end{cases}$$

Proof. Suppose that $\mathrm{sw}(a^*) = \alpha$.

For the purposes of this proof, let $\alpha \cdot n$ be the exact social welfare conferred by the social-welfare-maximizing alternative. The stated bound for α-promise additive distortion follows by taking the maximum over all social welfare values above the promise.

When $\alpha \geqslant \frac{m-1}{m}$, the profile that minimizes the score of a^* consists of a $\frac{m-1}{m}$ fraction of voters who rank a^* first and assign it utility 1 and a $\frac{1}{m}$ fraction of voters who rank a^* last and assign it utility 0. The best any other alternative $a \neq a^*$ can do is be ranked second for a $\frac{m-1}{m}$

fraction of voters and first for a $\frac{1}{m}$ fraction of voters.

Therefore, $s_{Borda}(a^*) \geqslant \frac{m-1}{m} \cdot (m-1) + \frac{1}{m} \cdot (m-2)$, and $s_{Borda}(a) \leqslant \frac{m-1}{m} \cdot (m-2) + \frac{1}{m} \cdot (m-1)$. It is easy to verify that $s_{Borda}(a^*) > s_{Borda}(a)$.

In the second case, note that for $\alpha \in [1/2, (m-1)/m)$, one profile that achieves this α guarantee consists of an α fraction of voters who rank a^* first and assign it utility 1 and a $1-\alpha$ fraction of voters who rank a^* last and assign it utility 0. In this profile, $s_{Borda}(a^*) = \alpha \cdot (m-1)$. However, consider an alternative $a \neq a^*$ that is ranked second whenever a^* is ranked first and ranked first whenever a^* is ranked last. We can see that $s_{Borda}(a) = \alpha \cdot (m-2) + (1-\alpha) \cdot (m-1)$, which is greater than $s_{Borda}(a^*)$ for all $\alpha < \frac{m-1}{m}$, so the Borda rule chooses a.

Note that a may be valued at $1/m$ every time it is ranked first and at 0 every time it is ranked last. Therefore, the additive distortion is $\alpha - (1-\alpha) \cdot \frac{1}{m}$, which approaches α as m increases. □

Plurality and the Harmonic Rule are robust for $\alpha \geqslant 3/4$, which is the largest possible interval of α on which any SCF can guarantee an α-promise additive distortion of 0. For smaller α the situation for the Borda Rule is much worse. In particular, Borda ceases to be robust as soon as α dips below $\frac{m-1}{m}$. Lastly, we consider Randomized Dictatorship:

Claim 24. For Randomized Dictatorship,

$$\begin{aligned} &\mathrm{dist}^+_\alpha(RD) \\ &= \begin{cases} 2\alpha(1-\alpha) - \frac{2(1-\alpha)^2}{m-1} & \text{for } \alpha \geqslant \frac{1}{2}\left(1+\frac{1}{m}\right) \\ \frac{1}{2}\left(1-\frac{1}{m}\right) & \text{for } \alpha < \frac{1}{2}\left(1+\frac{1}{m}\right). \end{cases} \end{aligned}$$

Proof. To begin we assume that $\mathrm{sw}(a^*) = \alpha$ exactly. In order to maximize additive distortion, by now-familiar arguments may we assume without loss of generality that all agents ranking a^* first value it at utility 1, and all not ranking a^* first value it second at utility $1/2$. Letting β_a

denote the proportion of agents ranking $a \in A$ first, we have

$$\mathrm{sw}(a^*) = \frac{1}{2}(1 + \beta_{a^*}), \qquad \mathrm{sw}(a) = \frac{\beta_a}{2} \qquad \text{for } a \neq a^*.$$

Omitting the factor of n and noting that $\beta_{a^*} = 2\alpha - 1$ (for $\alpha \geqslant 1/2$), we have

$$\begin{aligned}
\mathrm{dist}^+_\alpha(RD) &= \max_\beta \left(\alpha - \sum_a \beta_a \,\mathrm{sw}(a) \right) \\
&= \max_\beta \left(\alpha - \left(\beta_{a^*}\alpha + \sum_{a \neq a^*} \beta_a \,\mathrm{sw}(a) \right) \right) \\
&= \alpha - \left(\beta_{a^*}\alpha + \frac{m-1}{2} \left(\frac{1 - \beta_{a^*}}{m-1} \right)^2 \right) \\
&= 2\alpha(1-\alpha) - \frac{2(1-\alpha)^2}{m-1}.
\end{aligned}$$

It remains only to relax our original assumption that $\mathrm{sw}(a^*) = \alpha$ exactly. Therefore to find the α-promise additive distortion, we maximize over all $\alpha' \geqslant \alpha$ and obtain the stated bound. □

In particular, as we might expect for randomized rules, additive distortion decays smoothly towards 0 as $\alpha \to 1$.

3.5 Experiments

We evaluated the performance of various SCFs on four datasets of election data from PrefLib [MW13]: *Vermont* consists of data from public office elections in 2014 (15 different races, with 3 to 6 candidates and 532 to 1960 voters per race); *Glasgow* consists of data from the 2007 Glasgow City Council elections (21 wards, with 8 to 13 candidates and 5199 to 12744 voters per ward); *Debian* consists of votes for the Debian logo (8 elections, with 4 to 9 alternatives and 142 to 504 voters per election); and *APA* consists of election data from the American Psychological Association between 1998 and 2009 (12 elections, with 5 alternatives and 13318 and 20239 voters).

We also considered seven SCFs. Four of them are randomized scoring rules: *Randomized Dictatorship* has score vector $s = (1, 0, \dots)$ [AS98]; *RSR Borda* has score vector $s = (m-1, m-2, \dots, 0)$; *RSR Harmonic* has score vector $s = (1, 1/2, \dots, 1/m)$; and *BoB* has score vector $s = (25/33, 7/33, 1/33, 0, \dots)$. The other three are instance-optimal rules: *Det Add OPT* is the deterministic rule that minimizes additive distortion [Car+17]; *Mult OPT* is the randomized rule that minimizes multiplicative distortion [Bou+12]; and *Add OPT* is the randomized rule that minimizes additive distortion based on Theorem 17.

Notably, all data was presented as a complete ranking that allowed ties between alternatives. Therefore in computing the rules, we split weight equally in the RSRs (i.e., if k alternatives were tied, they split the total score that the rule allocates over those k positions) and enforced the constraint that the implicit utility assigned to all tied alternatives is equal.

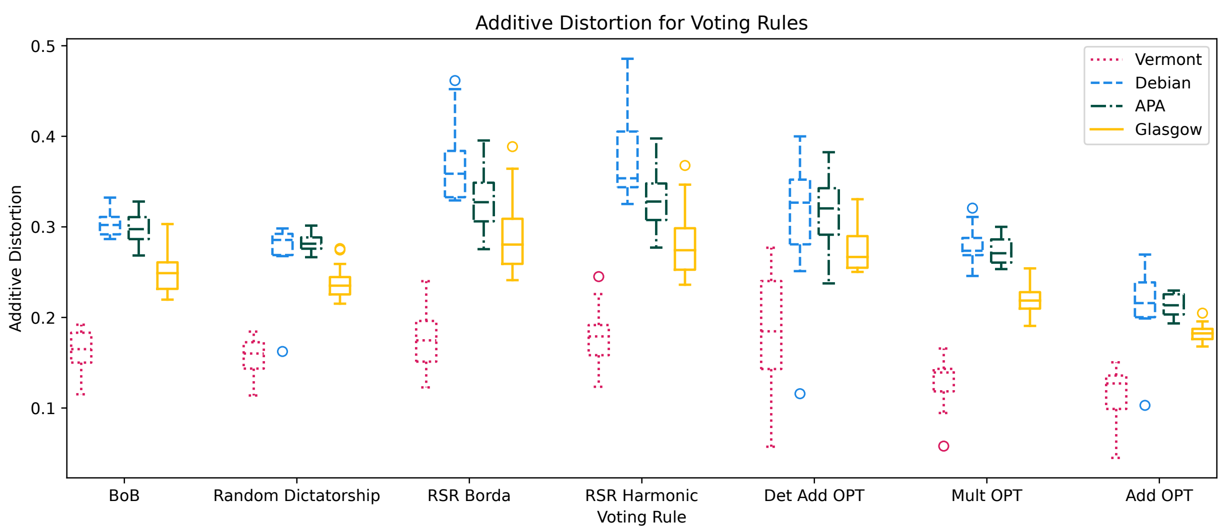


Figure 3.1: Additive distortion of voting rules on the Vermont, Glasgow, Debian, and APA datasets, normalized by n.

The additive distortions of each voting rule for each dataset are depicted in Figure 3.1. BoB generally outperforms Det Add OPT on all datasets, meaning that it results in lower additive distortion than *any* deterministic rule, which is why we compare its performance to the other randomized scoring rules RD, RSR Borda, and RSR Harmonic. We find that RD consistently outperforms BoB on the four datasets, while RSR Borda and RSR Harmonic both do worse. This is surprising, since Theorem 16 demonstrates that BoB is asymptotically worst-case optimal among the class of all RSRs. This suggests that real-life instances may not resemble worst-case additive distortion instances, and that more "imbalanced" randomized positional scoring rules (with more precipitous drop-offs in scores after the first position) result in lower additive distortion in practice.

Notably, Caragiannis et al. [Car+17] performed experiments in which Det Add OPT performed the best of the (deterministic) rules that they tested; the fact that both BoB and RD outperform Det Add OPT in terms of worst-case additive distortion is surprising and encouraging.

Additionally, there is a separation between the performance of Add OPT and Mult OPT (particularly for the Debian and APA datasets), which suggests that existing distortion-optimal rules do not optimize for additive distortion. Despite this separation, Mult OPT often outperforms the randomized positional scoring rules we implemented.

Furthermore, note that Add OPT significantly outperforms all rules on all elections. Encouragingly, calculating Add OPT is extremely efficient due to Theorem 17, and we expect that this approach is scalable to much larger elections. In comparison, Mult OPT took on the order of thousands of times longer than the others we tested.

3.6 Discussion

There are many exciting directions for future inquiry. Most immediately, it would be nice to close the gap between our upper and lower bounds of $\frac{5}{18} \cdot n$ and $\frac{11}{27} \cdot n$ for randomized rules. It would also be interesting to explore the additive distortion guarantees of more rules (especially randomized rules) in the α-promise setting. We believe that is also worth further exploring the class of rules $\mathcal{RSR}^*$, since it features rules that perform remarkably well with respect to additive and multiplicative distortion in a range of settings. Finally, it would be interesting to characterize the instances on which multiplicative and additive distortion come apart; this could help to determine which distortion is the right fit in various settings.

Chapter 4

Proportional Representation with Incomplete Votes

4.1 Introduction

A recent surge of interest in empowering citizens through online civic participation has spurred the development of a number of platforms [SL15; Ito+20; Shi+19; Fis+19; Ara+17; IKZ09]. A particularly successful example is *Polis* [Sma+21],[1] an open-source "system for gathering, analyzing, and understanding what large groups of people think in their own words." It has been widely used by local and national government agencies around the world. Most notably, it is the basis of vTaiwan, a system commissioned by the government of Taiwan, whose participatory process — involving thousands of ordinary citizens — has led to new regulation of ride-sharing services and financial technology. A similar (albeit commercial) system called *Remesh*[2] allows users to "save resources by democratizing insights in live, flexible conversations

[1] `https://pol.is`

[2] `https://www.remesh.ai`

with up to 1,000 people at the same time."[3]

The key idea underlying both systems is simple and broadly applicable: Participants can submit free-text comments about the discussion topic at hand and choose to agree or disagree with others' comments presented to them by the platform. An essential part of the process is the aggregation of these opinions toward an "understanding of what large groups of people think." Polis, for instance, displays a list of comments that received the most support among participants to whom they were shown. But this aggregation method may fail to represent minority groups, even those that are very large: if 51% of participants agree with one set of comments, while 49% of participants agree with another set of comments, only comments from the first set will appear on this list. Polis has recognized this problem and sought to mitigate it by employing a second, more elaborate procedure [Sma+21].[4] While this procedure has produced interesting results in practice, it does not guarantee summarizations that are representative of the discussion in a rigorous sense.

In this paper, we reexamine opinion aggregation in systems such as Polis and Remesh through the lens of computational social choice [Bra+16]. We observe that *selecting a subset of comments based on agreements and disagreements is equivalent to electing a committee based on approval votes.* From this viewpoint, the primary aggregation method used by Polis corresponds to classical approval voting (AV). There is substantial work — starting with the paper of Aziz et al. [Azi+17] — on *approval-based committee elections* that seeks to avoid the shortcomings of AV by guaranteeing that the selected committee satisfies fairness notions. To define one such notion (which is not satisfied by AV), note that if the size of the committee is k and the number of voters is n, a subset of n/k voters is large enough to demand a seat on the committee if they agree on at least one candidate. This intuition is captured by a property

[3] See also `consider.it`, `citizens.is`, `make.org`, `kialo.com`.

[4] sThe idea is to find clusters of participants with similar opinions and then ensure that each cluster is represented by comments that distinguish it from the others.

called *justified representation (JR)*, which guarantees that every such subset of voters has an approved candidate on the committee.

There is a major gap, however, between the literature on approval-based committee elections and the reality of systems like Polis and Remesh: these systems only have access to partial votes. For example, in the discussion facilitated by Polis around ridesharing regulation in Taiwan, 197 comments were submitted, but each participant only cast a vote on 10.57 comments on average — roughly 5% of all comments. Our main conceptual insight is that we can overcome the partial-information gap via statistical estimation and adaptive querying (i.e., by deciding which comments to show to incoming users based on previous votes).

Our approach and results

In our model, each voter (user) can be asked to express their opinion (approval/disapproval) about at most t candidates (comments). More formally, a *query* asks a randomly-chosen voter for their approval votes on a subset of candidates S of size $|S| \leqslant t$. Note that this query model is consistent with how Polis works, where participants express their agreement or disagreement with the comments shown to them by the system. We can view the response to such a query, i.e., the approval votes of a single voter, as noisy information about the profile of the entire population of voters (restricted to these t candidates). Therefore, we refer to these real-world queries as *noisy queries*.

Before we discuss this setting, we consider a simplified setting in Section 4.3, where queries yield the profile of the entire population of voters on the t candidates in the query. While such *exact queries* are not realistic, they provide an abstraction that is easier to study and allows us to derive lower bounds on the number of queries required to achieve JR (which apply also to the noisy-query setting, since it is strictly harder). We start by studying the required number of queries of *non-adaptive* algorithms, which decide on their queries before any votes are cast. While non-adaptive algorithms may be preferable in some cases (e.g., because no voter can influence what alternatives are shown to other voters or because computation can be performed offline), we show that they are impractical because they must ask at least $\Omega(m^{11})$ queries (and hence voters) to achieve JR, where m is the number of candidates.

Therefore, we focus on *adaptive* algorithms in the rest of the paper. In Section 4.3.2 we adapt a local-search algorithm of Aziz et al. [Azi+18] to the case of exact queries and show that it can achieve JR (and even stronger properties) with $\mathcal{O}(mk^2 \log k)$ queries.

In Section 4.4, we move on to the realistic, noisy-query model, where a query corresponds to a single voter. Since we need to estimate the answer to each exact query using multiple noisy queries to control uncertainty, the query complexity of the adaptive algorithm for the same guarantees increases to $\mathcal{O}\left(mk^6 \log k \log m\right)$. By applying martingale theory, we develop an extension of this algorithm that allows the reuse of votes in a statistically sound way.

In Section 4.5 we show empirically (on real datasets from Polis and Reddit) that this extension allows us to find committees satisfying (approximate) JR (and stronger properties) despite access to little information (i.e., few voters, each voting on only a small fraction of the comments).

4.2 Preliminaries

We begin by introducing the standard approval-based committee selection setting [Azi+17]. For $s \in \mathbb{N}$, we use the notation $[s] = \{1, \ldots, s\}$. We have a set $N = [n]$ of n voters and a set C of m candidates. Each voter $i \in N$ approves a set of candidates $A_i \subseteq C$. We refer to the vector $\mathbf{A} = (A_1, \ldots, A_n)$ as an *approval profile*. The goal is to choose a *committee* $W \subseteq C$ of size $k \leqslant m$. The value k is called the *target committee size*. We refer to an algorithm that takes as input the profile and candidates and outputs a committee of size k as a *k-committee-selection algorithm*.

Notions of representation

We say that a group of voters $V \subseteq N$ is *ℓ-large* if $|V| \geqslant \ell \cdot \frac{n}{k}$; V is *ℓ-cohesive* if $|\bigcap_{i \in V} A_i| \geqslant \ell$. Aziz et al. [Azi+17] introduced the following two fairness notions:

Definition 7 (Justified Representation (JR))**.** A committee W provides JR if for every 1-large, 1-cohesive group of voters V, there exists a voter $i \in V$ who approves a member of W, i.e., $|A_i \cap W| \geqslant 1$.

Definition 8 (Extended Justified Representation (EJR))**.** A committee W provides EJR if for every $\ell \in [k]$ and every ℓ-large, ℓ-cohesive group of voters V, there exists a voter $i \in V$ who approves at least ℓ members of W, i.e., $|A_i \cap W| \geqslant \ell$.

We also study the following approximate version of EJR:

Definition 9 (α-Extended Justified Representation (α-EJR))**.** A committee W provides α-EJR if for every $\ell \in [k]$ and every $\frac{\ell}{\alpha}$-large, ℓ-cohesive group of voters V, there exists a voter $i \in V$ who approves at least ℓ members of W, i.e., $|A_i \cap W| \geqslant \ell$.

Fernández et al. [Fer+17] proposed another notion of representation called the *average satisfaction* of a group of voters V for a committee W, defined as $\text{avs}_W(V) = \frac{1}{|V|} \sum_{i \in V} |A_i \cap W|$. Related to this quantity, we define the following property:

Definition 10 (α-Optimal Average Satisfaction (α-OAS))**.** A committee W provides *α-OAS* if for every $\lambda \in [0, k]$ and every $\frac{\lambda+1}{\alpha}$-large, λ-cohesive group of voters V, we have $\text{avs}_W(V) \geqslant \lambda$.

This property measures how close a committee is to the maximum average satisfaction that can be guaranteed to hold for all elections. To see this, note that the condition above is equivalent to the following condition: for every $\ell \in [\frac{1}{\alpha}, \frac{k+1}{\alpha}]$ and every ℓ-large, $(\alpha\ell - 1)$-cohesive group of voters V, we have $\text{avs}_W(V) \geqslant \alpha\ell - 1$. This implies a *proportionality guarantee* [Sko21] of $g(\ell, k) = \alpha\ell - 1$. Since there is no selection rule that satisfies a proportionality guarantee with $g(\ell, k) > \ell - 1$ for all elections [Azi+18; Sko21], $\alpha = 1$ is the best we can hope for, so we refer to 1-OAS simply as OAS.

Proportional approval voting

Proportional Approval Voting (PAV) is a widely-studied committee selection algorithm: given an approval profile $\mathbf{A}$ and a committee size k, it returns a committee W of size k maximizing the *PAV score*, defined as

$$\text{PAV-SC}(W) := \frac{1}{n}\sum_{i\in N}\sum_{j=1}^{|A_i\cap W|}\frac{1}{j}.$$

PAV satisfies EJR and OAS [Fer+17; Azi+18], but is NP-hard to compute [Azi+15]. Consequently, Aziz et al. [Azi+18] propose a local search approximation of PAV (LS-PAV), which continues to satisfy EJR and OAS, but, unlike PAV, runs in polynomial time. As we shall see, LS-PAV is a useful basis for algorithms in our query model.

4.3 Exact Queries

In the exact-query setting, the response R to a query Q consists of a proportion p_S for every subset $S \subseteq Q$, where p_S is the proportion of voters who only approve the candidates in S among the queried candidates Q, i.e.,

$$p_S := \frac{1}{n}\sum_{i\in N}\mathbb{I}[A_i \cap Q = S],$$

where $\mathbb{I}$ is the indicator function. We refer to an algorithm that makes queries of size t, receives this type of response, and outputs a committee of size k as a (k,t)*-committee selection algorithm with exact queries*. We say an algorithm is *adaptive* if the queries it chooses depend on responses from previous queries. Note that we allow all of our algorithms to be randomized. In the following, we ask how many queries are needed to guarantee the notions of representation introduced in Section 4.2.

4.3.1 Nonadaptive Algorithms

In this section, we take m to be large (many comments will be submitted to the system), while we think of k and t as small constants (since we wish to select only a few comments and voters have limited time). Since we are primarily interested in *lower* bounds on the query complexity of non-adaptive algorithms, we consider only JR, the weakest fairness criterion.

An initial observation is that, if $t \geqslant k$, JR can always be guaranteed with $O(m^k)$ queries, as simply querying every set of k candidates provides all the information necessary to run PAV. For $k = 1$, this bound is tight, as voters could all approve only a single candidate, which will take a linear number of queries to find. Our first result is a tight quadratic lower bound for $k = 2$.

Theorem 25. *For any constants k and t such that $k \geqslant 2$, and any $\varepsilon > 0$, any non-adaptive (k, t)-committee selection algorithm that makes fewer than $\Omega(m^2)$ queries satisfies JR with probability at most ε.*

This result provides a separation between the non-adaptive and the adaptive settings: In Section 4.3.2, we discuss an adaptive (k, t)-committee selection algorithm guaranteeing JR with only $O(m)$ queries for any k and t such that $k < t$.

Theorem 25 follows from a more general result that we present formally in Appendix 4.7.1. Here, we illustrate the argument by considering the special case where $t = k = 2$ and $\varepsilon = \frac{5}{6}$: Consider an adversary that picks a random set of 3 candidates, call them 1, 2, and 3, and answers queries according to the approval matrix visualized in Figure 4.1(a): half of the voters approve only candidate 1, and the other half of the voters approve only candidates 2 and 3. To satisfy JR, the algorithm needs to include candidate 1 in the committee. However, if the algorithm never queries $\{1, 2\}$, $\{1, 3\}$, or $\{2, 3\}$, it receives no information that can distinguish candidates 1, 2, and 3 from each other, so it can do no better than selecting a random pair from these three candidates, which will succeed with probability $\frac{2}{3}$. This problematic case will occur frequently if the number of queries is not very large, say $\frac{1}{18} \cdot \binom{m}{2}$: Since there are $\binom{m}{2}$ pairs of candidates, the probability that the algorithm queries any randomly selected pair of candidates is at most $\frac{1}{18}$. By the union bound, the probability that the algorithm queries any of $\{1, 2\}$, $\{1, 3\}$, or $\{2, 3\}$ is at most $3 \cdot \frac{1}{18} = \frac{1}{6}$. To summarize, for the algorithm to succeed, it either needs to get lucky during the querying phase, which happens with probability at most $\frac{1}{6}$, or during the selection phase, which happens with probability at most $\frac{2}{3}$. By the union bound, the algorithm succeeds with probability at most $\frac{1}{6} + \frac{2}{3} = \frac{5}{6}$.

A natural follow-up question is whether the $O(m^k)$ upper bound is tight for larger k. Interestingly, this is *not* the case for $k \geqslant 3$.

Theorem 26. *For any $t \geqslant \frac{2}{3}k$, there exists a (k, t)-committee selection algorithm guaranteeing JR with $O(m^t)$ exact queries.*

Proof. Consider Algorithm 3, described below.

Algorithm 3 (k, t)-non-adaptive (for $t \geqslant \frac{2}{3}k$)

Query every set of candidates of size t
for $i \leftarrow 1, 2, \ldots, t$ **do**
  $c_i \leftarrow$ approval winner among voters not approving $c_1, c_2, \ldots, c_{i-1}$
for $i \leftarrow t + 1, \ldots, \lfloor \frac{3}{2}t \rfloor$ **do**
  $c_i \leftarrow$ arbitrary default candidate
  for $c \in C$ **do**
    $A \leftarrow$ set of voters approving of c but not any of $\{c_1, c_2, \ldots, c_{t-1}\}$
    $B \leftarrow$ set of voters approving of c but not any of $\{c_1, c_2, \ldots, c_{i-1}\} \setminus \{c_{\lfloor t/2 \rfloor}, \ldots, c_{t-1}\}$
    if $|A| \geqslant \frac{n}{k}$ and $|B| \geqslant \frac{n}{k}$ **then**
      $c_i \leftarrow c$
return $\{c_1, c_2, \ldots, c_k\}$

Provided that $t \geqslant \frac{2}{3}k$, it is straightforward to verify that the **if** condition can be checked using only information about sets of voters of size t. Thus, Algorithm 3 is indeed a non-adaptive (k, t)-committee selection algorithm with exact queries.

For each $i \in \{1, 2, \ldots, \lfloor \frac{3}{2}t \rfloor\}$, we say that a voter is *satisfied on round i* if it approves of c_i, but none of the previously selected candidates $c_1, c_2, \ldots, c_{i-1}$, and we say that a voter is *satisfied by round* i if it was satisfied on some round $j \leqslant i$. We prove that the final committee satisfies JR by counting the fraction of voters that are satisfied on each round. Indeed, JR is equivalent to the property that there is no 1-cohesive $1/k$-fraction of voters that is left unsatisfied by the k^{th} round.

The case where $k \leqslant t$ is easy: on each of the first k rounds, either we satisfy a $1/k$ fraction of voters, or there is no 1-cohesive set of $\frac{n}{k}$ unsatisfied voters. Thus, by round k, either all voters are satisfied, or the remaining set of unsatisfied voters has no 1-cohesive set of size $\frac{n}{k}$.

Now suppose that $k > t$. For each i, let x_i denote the fraction of voters that are satisfied on round i. Note that the sequence of x_i are weakly decreasing for $i \leqslant t$. Again, if any $x_i < \frac{1}{k}$, it means that there is no 1-cohesive $\frac{1}{k}$-fraction of unsatisfied voters after round i, so JR is already satisfied. So assume each $x_i \geqslant \frac{1}{k}$ for all $i \leqslant t$. Further, if on any round $i > t$ we fail to find a candidate c making the **if** condition true, we claim that JR is already satisfied. For if JR were not satisfied, then there would be some candidate c approved by a $\frac{1}{k}$-fraction of voters S who approve of no previous candidates. Clearly, we would then have $S \subseteq A$ and $S \subseteq B$, so A and B both contain at least $\frac{1}{k}$ fractions of voters.

Thus, we may assume that, for each $i > t$, candidate c_i satisfies the **if** statement on round i. Consider an arbitrary round i. Let A and B denote the respective values of the variables on the iteration of the inner loop where c_i was set to its ultimate value. Observe that the candidates enumerated in the definitions of A and B cover all previously selected candidates. This means that $A \cap B$ is precisely the set of voters approving c_i and not any of the previous candidates; in other words, $A \cap B$ is the set of voters satisfied on round i. On the other hand, since candidates $c_1, c_2, \ldots, c_{\lfloor t/2 \rfloor}$ are enumerated in the definitions of both sets, it follows that $A \cup B$ is a set of voters approving c_i but not any of $c_1, c_2, \ldots, c_{\lfloor t/2 \rfloor}$. This means that $A \cup B$ can contain at most an $x_{\lfloor t/2 \rfloor}$ fraction of voters, for otherwise candidate c_i should have been selected earlier, on round $\lfloor t/2 \rfloor$. Thus, we may lower bound the fraction of voters satisfied on round i as

$$\frac{1}{n}\left(|A \cap B|\right) = \frac{1}{n}\left(|A| + |B| - |A \cup B|\right) \geqslant \frac{1}{n}\left(\frac{n}{k} + \frac{n}{k} - nx_{\lfloor t/2 \rfloor}\right) = \frac{2}{k} - x_{\lfloor t/2 \rfloor}.$$

Summing over each of the first k rounds, the number of satisfied voters is

$$\begin{aligned}
\sum_{i=1}^{k}(\#\text{ satisfied voters on round } i) &\geqslant \sum_{i=1}^{t} x_i + \sum_{i=t+1}^{k}\left(\frac{2}{k} - x_{\lfloor t/2 \rfloor}\right) \\
&= \sum_{i=1}^{t} x_i - \sum_{i=t+1}^{k} x_{\lfloor t/2 \rfloor} + (k-t)\frac{2}{k} \\
&= \sum_{i=1}^{k-t}\left(x_i - x_{\lfloor t/2 \rfloor}\right) + \sum_{i=k-t+1}^{k} x_i + (k-t)\frac{2}{k}.
\end{aligned}$$

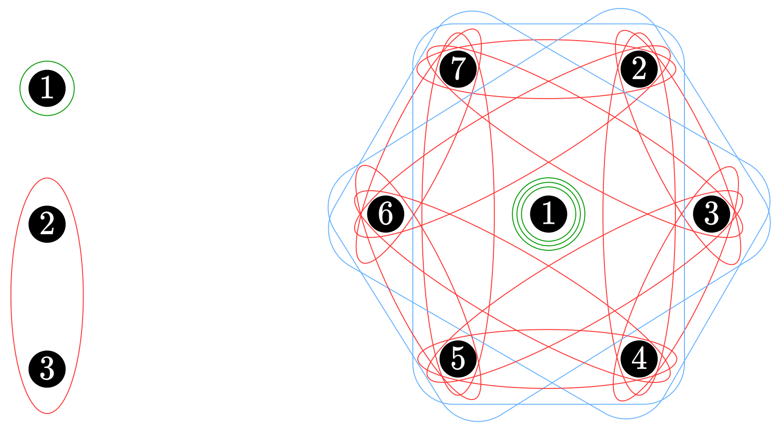
Since the x_i are decreasing and $k - t \leqslant \lfloor t/2 \rfloor$ when $t \geqslant 2k/3$, this first term is nonnegative, and $x_i \geqslant 1/k$ for $i \leqslant t$ in this case. Therefore we have

$$\begin{aligned}
\sum_{i=1}^{k}(\#\text{ satisfied voters on round } i) &\geqslant \sum_{i=k-t+1}^{k} x_i + (k-t)\frac{2}{k} \\
&\geqslant (t-(k-t))\frac{1}{k} + (k-t)\frac{2}{k} \\
&= \frac{2(k-t) + (2t-k)}{k} \\
&= 1.
\end{aligned}$$

Since all voters are satisfied by round k, the final committee satisfies JR. □

However, the exponent does have a dependence on k. In particular, we find that guaranteeing JR requires $\Omega(m^3)$ queries starting at $k = 6$. The adversary employs an analogous strategy, now picking 7 random candidates and imposing the approval matrix depicted in Figure 4.1(b). Satisfying JR requires that the algorithm include candidate 1, which is indistinguishable from the other six candidates unless the algorithm makes $\Omega(m^3)$ queries, since every candidate is approved by $\frac{6}{18}$ of the voters and every pair of candidates is approved by $\frac{2}{18}$ of the voters.

In Appendix 4.7.1, we describe a computational search we conducted to find similar instances for larger values of k. The best lower bound obtained is as follows.



(a) $k = 2$ (b) $k = 6$

Figure 4.1: Adversarial approval matrices. Each region represents a disjoint, equally-sized set of voters who approve only the candidates within the region. In (a), queries of size $t \geqslant 2$ are needed to distinguish the candidates; in (b), we need $t \geqslant 3$.

Theorem 27. *For any $\varepsilon > 0$, there exists a target committee size k with $k = \Theta(\log 1/\varepsilon)$ such that for all t, any non-adaptive (k, t)-committee selection algorithm with exact queries that makes fewer than $\Omega(m^{11})$ queries satisfies JR with probability at most ε.*

This theorem definitively establishes the impracticality of non-adaptive committee selection algorithms. We therefore turn our attention to adaptive algorithms.

4.3.2 An Efficient Adaptive Algorithm

In this section, we propose an adaptive algorithm based on LS-PAV [Azi+18], and we show that it achieves EJR and OAS with a practically-feasible number of queries.

For convenience, we introduce the following notation: For a committee W and candidates $c \in W$ and $c' \notin W$, let

$$\Delta(W, c', c) := \text{PAV-SC}(W \cup \{c'\} \setminus \{c\}) - \text{PAV-SC}(W)$$

Algorithm 4 (k,t)-α-PAV

```
Choose W ∈ (C choose k), c ∈ W, and c' ∉ W arbitrarily
γ ← ∞
while γ ⩾ 1/(αk) do
    W ← W ∪ {c'} \ {c}
    Choose 𝒬 = {Q_i}_i, with |Q_i| = t, s.t. W ⊆ ⋂𝒬
     and C ⊆ ⋃𝒬
    c' ← arg max_{x∉W} Δ(W, x) // (using 𝒬)
    c ← arg max_{x∈W} Δ(W, c', x) // (using 𝒬)
    γ ← Δ(W, c')
return W
```

denote the difference in PAV score obtained by replacing c with c' in W. Additionally, let

$$\Delta(W,c) := \text{PAV-SC}(W \cup \{c\}) - \text{PAV-SC}(W)$$

denote the marginal gain in PAV score by adding c to W.

LS-PAV starts with an arbitrary committee W and repeatedly replaces a committee member $c \in W$ with a candidate $c' \notin W$, provided the improvement to the PAV score satisfies $\Delta(W,c',c) \geqslant \frac{1}{k^2}$. Aziz et al. [Azi+18] show that after at most $\mathcal{O}(k^2 \log k)$ swaps, no such swap pairs c, c' remain, at which point W satisfies OAS and EJR.

We first observe that LS-PAV can be implemented using exact queries: For any set of candidates S, PAV-SC(S) can be computed using any query $Q \supseteq S$, as it is sufficient to know the proportion of voters that approve each subset of S. Hence, for any W, $c \in W$, and $c' \notin W$, $\Delta(W,c',c)$ can be computed using a query Q that contains both W and c'. Using $\left\lceil \frac{m-k}{t-k} \right\rceil$ queries of size t, we can cover all $m-k$ candidates that are not in W, which leads to an overall query complexity of $\mathcal{O}\left(mk^2 \log k\right)$.

We next present a version of LS-PAV, which we call α-PAV (Algorithm 4), that has the same query complexity as LS-PAV for finding a committee that satisfies EJR and OAS, but lower query complexity for approximate ($\alpha < 1$) α-EJR and α-OAS. Besides introducing the approximation parameter α, we make two other modifications to LS-PAV: First, Algorithm 4 terminates when there is no candidate c' such that $\Delta(W,c') \geqslant \frac{1}{k}$ (for $\alpha = 1$), while LS-PAV terminates when there is no pair c, c' such that $\Delta(W, c', c) \geqslant \frac{1}{k^2}$. As we shall see in Lemma 8, the termination condition of Algorithm 4 is weaker than that of LS-PAV, implying that it may terminate earlier. Second, instead of considering all possible swaps c, c', we only consider adding the candidate c' with the largest $\Delta(W, c')$. This modification makes the algorithm slightly simpler and more computationally efficient (by a factor of k).

Theorem 28. *For any $m \geqslant t > k$, Algorithm 4 yields a committee satisfying α-OAS and α-EJR while making at most*

$$\left\lceil \frac{m-k}{t-k} \right\rceil \frac{\alpha k^2}{(1-\alpha)k+1} H_k$$

queries, where H_k is the k^{th} harmonic number. For $\alpha = 1$, this leads to a query complexity of $\mathcal{O}\left(mk^2 \log k\right)$ while for any fixed $\alpha < 1$, this leads to a query complexity of $\mathcal{O}\left(mk \log k\right)$.

The proof of theorem 28 essentially follows from the following two lemmas, the first of which uses the notation

$$\Delta^*(W) \coloneqq \max_{c \in C} \Delta(W, c).$$

Theorem 29. *If a committee W satisfies $\Delta^*(W) < \frac{1}{\alpha k}$, then W satisfies α-EJR and α-OAS.*

Lemma 8. *For any committee W and $c \notin W$, we have that $\max_{x \in W} \Delta(W, c, x) \geqslant \frac{(k+1)\Delta(W,c)-1}{k}$. In particular, if $\Delta(W, c) \geqslant \frac{1}{\alpha k}$, then $\max_{x \in W} \Delta(W, c, x) \geqslant \frac{(1-\alpha)k+1}{\alpha k^2}$.*

Theorem 29 guarantees that when Algorithm 4 terminates the desired fairness properties are satisfied. Lemma 8 establishes that the PAV score increases over the algorithm's run. This bounds the number of swaps it performs since $\textsc{pav-sc}(W)$ is at most H_k.

Theorem 29 is a generalization of the lower bound from Lemma 1 of Skowron [Sko21]. This generalization is useful because it states that to establish EJR and OAS of any given committee W (no matter how it is derived), it is sufficient to prove that $\Delta^*(W)$ is small; hence it can be used as a certificate of satisfaction. In Section 4.7.4, we show that standard PAV and LS-PAV satisfy $\Delta^*(W) < \frac{n}{k}$, which is noteworthy in that it provides a simple proof of the known result that they satisfy EJR and OAS.

We observe that, for exact queries, an α-approximation with $\alpha < 1$ improves the query complexity by a factor of k. In the next section, we will see that such an approximation yields an even larger improvement in query complexity for noisy queries, as it also reduces the accuracy with which we need to estimate $\Delta(W, c', c)$.

4.4 Noisy Queries

We now turn to a query model that includes the noise we abstracted away in Section 4.3. In order to represent voters arriving to the platform one-by-one, we assume that each time the algorithm performs a query $Q \subseteq C$ a voter $i \in N$ is selected independently and uniformly at random[5]and then the algorithm observes their votes on the queried candidates $Q \cap A_i$. We refer to an algorithm that performs queries of size t, receives as a response the votes of a single voter, and outputs a committee of size k as a (k, t)-*committee selection algorithm with noisy queries.*

To see the connection between this query model and the previous one, note that an algorithm with noisy queries can approximate an exact query Q by estimating the values of p_S by taking the empirical proportion of repeated samples. By standard sample complexity bounds, using $\Theta\left(\log(2^t/\delta)/\varepsilon^2\right)$ queries, a noisy-query algorithm could guarantee $\pm\varepsilon$ estimates of p_S for all $S \subseteq Q$ with probability $1 - \delta$. Hence if an exact-query algorithm requires no more than $\text{poly}(m)$ queries with additive ε error, then it can be implemented using a factor of $\Theta(\log m)$ more noisy queries and yield a correct result with probability $1 - \delta$. What's more, this log factor is in some cases necessary when moving from the exact-query to the noisy-query

[5]Note that a voter-profile A_i may be queried more than once during the run of the algorithm because we sample *with replacement*. This model simplifies the statistical analysis and has a natural interpretation: Rather than thinking of a finite population of voters, we draw samples from an underlying population distribution where each profile $A_1, .., A_n$ has the same frequency (probability). Furthermore, our model approaches sampling without replacement if the size of the underlying population n is large compared to the number of queried voters, hence both models are qualitatively interchangeable.

Algorithm 5 (k, t)-noisy-α-PAV

$\ell \leftarrow \left\lceil 288 \left(\frac{\alpha k^2}{(1-\alpha)k+1}\right)^2 \log\left(\frac{8mk^4}{\delta}\right)\right\rceil$
Choose $W \in \binom{C}{k}$, $c \in W$, and $c' \notin W$ arbitrarily
$\gamma \leftarrow \infty$
while $\gamma \geqslant 1/(\alpha k) - ((1-\alpha)k+1)/(12\alpha k^2)$ **do**
    $W \leftarrow W \cup \{c'\} \setminus \{c\}$
    Choose $\mathcal{Q} = \{Q_i\}_i$, with $|Q_i| = t$, such that $W \subseteq \bigcap \mathcal{Q}$ and $C \subseteq \bigcup \mathcal{Q}$
    Ask each query $Q \in \mathcal{Q}$ to ℓ new voters
    $\widehat{\Delta}(W, x) \leftarrow$ estimate of $\Delta(W, x)$ using ℓ voters from query Q containing $W \cup \{x\}$ // $\forall x \notin W$
    $\widehat{\Delta}(W, x, y) \leftarrow$ estimate of $\Delta(W, x, y)$ using ℓ voters from Q containing $W \cup \{x\}$ // $\forall x \notin W, \forall y \in W$
    $c' \leftarrow \arg\max_{x \notin W} \widehat{\Delta}(W, x)$
    $c \leftarrow \arg\max_{x \in W} \widehat{\Delta}(W, c', x)$
    $\gamma \leftarrow \widehat{\Delta}(W, c')$
return W

setting. In Section 4.7.2, we demonstrate instances for which a non-adaptive exact-query algorithm needs only $\Theta(m)$ queries, while in order to be correct with any fixed probability δ, a non-adaptive noisy-query algorithm requires $\Omega(m \log m)$ queries.

Conversely, notice that one can use exact queries to simulate noisy queries. Indeed, p_S is exactly the probability that an incoming voter will vote yes on candidates S and no on $Q \setminus S$ in response to a query Q. An algorithm with access to exact query values can simply sample a voter response and feed it to a noisy-query algorithm. Therefore, the lower bounds on the query complexity of exact-query, non-adaptive algorithms, in particular Theorem 27, apply to noisy-query, non-adaptive algorithms as well. As the number of candidates becomes large, adaptivity is therefore necessary to attain theoretical guarantees — mirroring the approach of online platforms in practice.

A natural starting point is the exact-query adaptive algorithm, namely Algorithm 4. Indeed, it can be adapted to the noisy setting by replacing exact queries with a sufficient number of noisy queries, ℓ, to obtain high-probability bounds on Δ, yielding Algorithm 5.

The key is to choose ℓ large enough that if the termination condition is not met, i.e., we have $\widehat{\Delta}(W,c') < \frac{1}{\alpha k} - \frac{(1-\alpha)k+1}{12\alpha k^2}$, the resulting swap is guaranteed to yield a positive improvement in the PAV-score, such that the number of steps of the algorithm is bounded. With the choice of ℓ in Algorithm 5, we obtain the following theorem, whose proof can be found in Section 4.7.5.

Theorem 30. *For any $m \geqslant t > k$, with probability at least $1-\delta$, Algorithm 5 returns a committee that satisfies α-EJR and α-OAS after querying no more than*

$$578 H_k \left\lceil \frac{m-k}{t-k} \right\rceil \left(\frac{\alpha k^2}{(1-\alpha)k+1} \right)^3 \log\left(\frac{4mk^4}{\delta} \right)$$

voters. For any fixed $\delta > 0$, if $\alpha = 1$, this leads to a query complexity of $\mathcal{O}\left(mk^6 \log k \log m\right)$, and if $\alpha < 1$, this leads to a query complextiy of $\mathcal{O}\left(mk^3 \log k \log m\right)$.

While Algorithm 5 achieves good worst-case query complexity, it may be suboptimal on certain instances because of two reasons: (i) after each swap, Algorithm 5 discards all previous information so each candidate is reassessed from scratch, and (ii) it presents each candidate $c \notin W$ to the same number of voters, even though it may quickly become apparent that some candidates are more promising than others.

To address issue (i), we can use all past votes to compute bounds on Δ. A difficulty with this approach is that past voters may not have voted on all candidates in W (which is necessary to directly estimate $\Delta(W,c)$), since they may have been queried on a different committee W'. But we can nonetheless use these past votes to obtain upper and lower bounds on estimated values. To address issue (ii), we can present promising candidates to voters more often. Further, it is possible to perform swaps as soon as we are confident they yield an increase of the PAV-score of at least some value ε, rather than first querying a predetermined number of voters as in Algorithm 5.

These ideas are incorporated into Algorithm 6, called ucb-α-PAV; see Appendix 4.7.6 for a formal description of the algorithm and an analysis of its query complexity.

4.5 Experiments

Since the analysis in the theoretical sections considers worst-case approval profiles, it is possible that, in practice, we may be able to find good committees with fewer queries than required by Theorem 30. We investigate this question empirically on real data from online discussions with only a few hundred voters, each voting on only a fraction of all comments.

Datasets

Polis provides open-use data from real deliberations hosted on their platform.[6] These include, for instance, a discussion organized by the government of Taiwan, which led to the successful regulation of Uber. Since participants typically only vote on a fraction of comments, most votes are missing. To be able to simulate the proposed adaptive algorithms, we first infer these missing votes using a matrix factorization library, LensKit.[7] Importantly, we infer votes only for the purpose of the experiments; if our algorithms were executed during the discussion, they would adaptively query users about the relevant comments without relying on any inference method.

In most datasets, we observe several comments that are nearly universally approved. Since these comments make achieving EJR and OAS trivial, we remove comments approved by more than 60% of participants. This step may also be appropriate in practice to gain insights into participants' opinions beyond uncontroversial issues.

The number of queried voters L ranges from 87 to 1000 across the 13 datasets. For all datasets, we assume that each voter votes on $t = 20$ comments. Since the total number of comments m ranges from 31 to 1719 across datasets, the percentage of comments each voter votes on, t/m, ranges from 1% to 65%. The data for each conversation consists of an

[6] `https://github.com/compdemocracy/openData`

[7] `https://lenskit.org`

$L \times m$ matrix, with component $v_{ij} \in \{\texttt{agree}, \texttt{disagree}, \texttt{missing}\}$ representing the vote of participant i on comment j. For each dataset, we run the algorithms with target committee sizes $k = 5, 7, 10$. Hence, there are a total of $13 \cdot 3 = 39$ experiments (times 10 random seeds).

The second dataset we consider consists of Reddit discussions.[8] To obtain an interesting dataset, we combined voting data from two subreddits, r/politics and r/Conservative, which are arguably situated at opposite ends of the American political spectrum. More details can be found in the additional dataset-specific sections below.

Algorithms

We evaluate noisy-α-PAV (Algorithm 5) and ucb-α-PAV (Algorithm 6). Both query L voters in random order, each of whom votes on $t = 20$ comments. To enable these algorithms to swap candidates after querying only a small number of voters, we make the following modifications: For both Algorithm 5 and Algorithm 6 we treat ℓ, the number of times we ask voters about each candidate, as a parameter. In addition, for Algorithm 6, we replace the numerator in the confidence intervals err_s with a parameter θ. Both ℓ and θ were chosen based on validation on a separate dataset. We run both algorithms on all the L voters, rather than terminating as soon as we can guarantee $\Delta^*(W) < \frac{1}{\alpha k}$ (and hence EJR and OAS).

To obtain an upper bound on the attainable performance, we execute α-PAV (Algorithm 4) with access to exact queries. To obtain the best possible α, we let Algorithm 4 run as long as the swap increases the PAV score, i.e., $\Delta(W, c', c) > 0$, instead of terminating as soon as $\Delta(W, c') < 1/k$ (which would be sufficient to guarantee EJR and OAS).

To verify that the proposed algorithms do indeed take the complementarity of different candidates into account, we also compare against standard approval voting (AV) with access to all votes, which simply selects the k candidates with the most approval votes.

---

[8] `https://www.kaggle.com/datasets/josephleake/huge-collection-of-reddit-votes`

Algorithm Parameters

In practice, for both Algorithm 5 and Algorithm 6, we treat ℓ, the number of times we ask voters about each candidate, as a parameter. In addition, for Algorithm 6, we replace the numerator in the confidence intervals err_s with a parameter θ. We assessed $\ell \in \{4, 6, 8, 12, 18, 30\}$ and $\theta \in \{0.03, 0.05, 0.1, 0.2, 0.5, 1.0\}$ on artificial data (e.g., approval profiles generated by sampling each vote independently). We observed that the algorithms are not sensitive to these parameters and picked $\ell = 6$ and $\theta = 0.05$ since they appeared to yield good results based on visual inspection.

Performance Metric

As a performance metric, we use $\widehat{\alpha} := \frac{1}{k\Delta^*(W)}$, where W is the committee selected by the respective algorithm. According to Theorem 29, $\alpha > \widehat{\alpha}$, so this implies $\widehat{\alpha}$-EJR and $\widehat{\alpha}$-OAS. As discussed in Section 4.2, $\alpha = 1$ is the best that can be guaranteed across all possible approval profiles. Note that α may be larger than $\widehat{\alpha}$, hence obtaining $\widehat{\alpha} = 1$ is a sufficient, but not a necessary condition for OAS and EJR. Nevertheless, we will use $\widehat{\alpha}$ as a metric for two reasons: first, verifying whether $\alpha \geqslant 1$ (i.e., whether a committee satisfies EJR and OAS) is computationally hard [Azi+17], which makes it impractical for evaluation; and the stronger condition $\widehat{\alpha} \geqslant 1$ provides the additional benefit that EJR and OAS can easily be verified through Theorem 29. Second, one could argue that $\widehat{\alpha}$ is a meaningful quantity in its own right since it (or rather its inverse $1/\widehat{\alpha}$) measures how much voter satisfaction could be improved by adding another candidate (giving lower weight to voters who already have many approved candidates).

Polis Datasets

Table 4.1 describes the the sizes of the Polis datasets.

L	m	t	t/m
162	31	20	0.65
1000	1719	20	0.01
87	39	20	0.51
353	231	20	0.09
340	209	20	0.10
94	40	20	0.50
1000	114	20	0.18
230	83	20	0.24
258	98	20	0.20
405	94	20	0.21
278	104	20	0.19
1000	586	20	0.03

Table 4.1:: Polis datasets statistics: Number of queried voters L, number of comments m, comments per query t, and fraction of comments t/m voted on by each voter.

Polis Results

In Figure 4.2, we show the $\widehat{\alpha}$ achieved on all the Polis datasets for each of the four algorithms. Recall that higher $\widehat{\alpha}$ is better and that $\widehat{\alpha} \geqslant 1$ implies OAS and EJR. As expected, α-PAV performs best since it has access to exact queries. Note that it often achieves an α substantially larger than 1, which means that the corresponding instance allows for better representation than can be guaranteed in the worst case. AV performs surprisingly well in most experiments, but in 38% of the cases, it yields $\widehat{\alpha}$ smaller than 1 (and sometimes much smaller). We conclude that for some datasets, it is important to take the complementarity of candidates into account rather than selecting them individually. The challenge for the proposed algorithms is to do so while being sample-efficient. We see that noisy-α-PAV often fails to achieve an $\widehat{\alpha} \geqslant 1$. We know from Theorem 30 that given enough queries, noisy-α-PAV achieves $\widehat{\alpha} \geqslant 1$, so this failure is due to the low number of queried voters. By contrast, ucb-α-PAV yields $\widehat{\alpha} \geqslant 1$ in 83% of the cases, and $\widehat{\alpha} \geqslant 0.75$ in all cases, which indicates that the proposed extensions (i.e., querying promising candidates more often, swapping as soon as possible, and reusing voters) indeed lead to more efficient use of data.

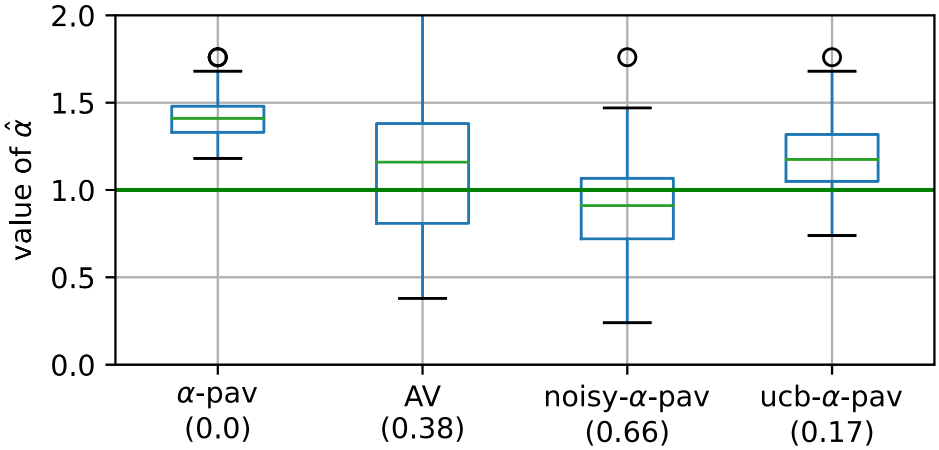


Figure 4.2: Boxplots where datapoints correspond to the 39 Polis problems ($\times 10$ random seeds). The top / bottom whiskers indicate the maximal / minimal points (except outliers, which are marked by circles), the line in the middle is the median, and the bottom and top of the boxes are the 1st and 3rd quartiles, respectively. The numbers in parenthesis are the fractions of problems where the respective algorithm yields a $\widehat{\alpha} \leqslant 1$.

Reddit Dataset

We preprocessed this dataset in the same way as the Polis datasets (including matrix factorization to infer missing votes). Although the output on Reddit differs from Polis (rankings rather than a subset of the comments), the *input* is similar. We can interpret upvotes as approvals and downvotes as disapprovals, so the data fits well with our experiments. See Section 4.6 for additional discussion on how our approach applies to social media.

Reddit Results

To illustrate why approval voting can perform poorly despite having access to the full votes, we execute the algorithms on the Reddit dataset described above. In this experiment, AV achieves only $\widehat{\alpha} = 0.68$. To understand why this happens, we show in Figure 4.3 the fraction of voters who have at least $1, \ldots, 10$ approved comments in the committee. We see that AV

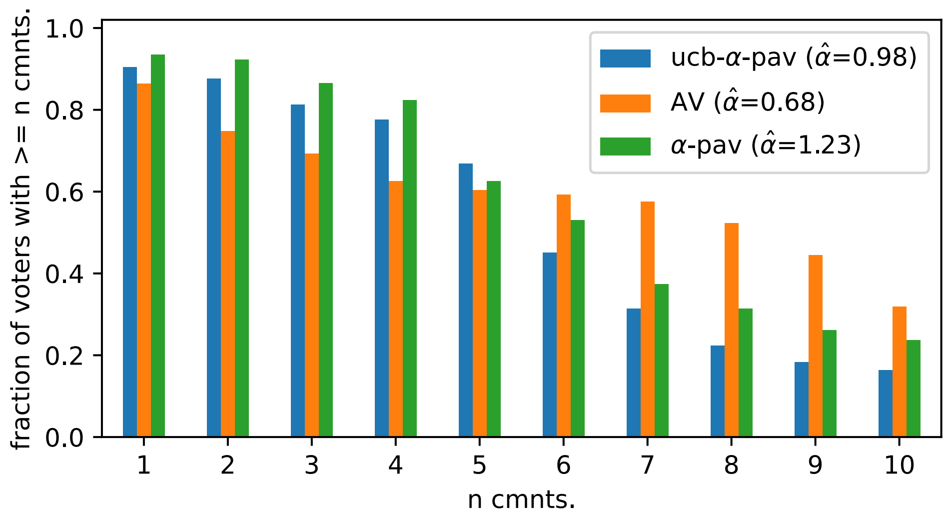


Figure 4.3: Results on Reddit dataset (with $L = 608, m = 2135, k = 10$): the fraction of voters (y-axis) that approve of at least $1, 2, .., 10$ candidates (x-axis) among the selected committee of size $k = 10$.

yields a committee where a high fraction of voters approve many candidates, e.g., about 60% of voters approve 7 or more candidates, whereas for α-PAV, this is the case for only about 40%. This comes at the cost of a high fraction of voters who are poorly represented by AV, e.g., about 25% of voters get at most one approved candidate, whereas for α-PAV, this percentage is less than 10%. This is to be expected as approval voting does not take the complementarity of candidates into account and can therefore lead to less equitable results. Finally, we observe that ucb-α-PAV achieves an $\widehat{\alpha}$ close to 1, and its approval fractions look similar to α-PAV, i.e., more equitable than AV. It is interesting that ucb-α-PAV performs well on this example, since it only has access to $t = 20$ votes for each of the $L = 608$ queried candidates, while it has to select from a large number of comments, $m = 2135$.

4.6 Discussion

The work in this chapter bridges the gap between online civic-participation systems, such as Polis, and committee-election methods by enabling them to handle incomplete votes. To deploy the proposed algorithms on such platforms, two practical issues must be considered.

First, our adaptive approach requires control over what the Polis creators call *comment routing* [Sma+21]: the algorithm that decides which comments are shown to which participants. If on a given platform a comment-routing algorithm is already in place, shared control is possible: each algorithm could determine part of the slate of comments shown to a participant, or the participants themselves can be divided between the algorithms.

Second, in our analysis, we assumed that all comments have been submitted — or all candidates are known — at the time we run our algorithms. Nevertheless, our algorithms can be extended straightforwardly to a growing set of comments, but we would inevitably lose the representation guarantees for comments that were submitted late if not enough participants could vote on them. In practice, this could be resolved by setting a comment submission deadline, which has been done previously by Polis.

An alternative to our approach would be to complete partial approval votes using collaborative filtering [RV97]. The completed approval votes can then be aggregated through any approval-based committee election rule, such as PAV. The disadvantage of this approach is that it is unlikely to lead to worst-case guarantees of the type we establish in this paper.

Finally, we emphasize that our approach may be applicable to social media more generally. For instance, as mentioned in Section 4.5, Reddit users also approve or disapprove comments through upvotes and downvotes. However, Reddit uses these inputs to produce a *ranking* of the comments, in contrast to our goal of selecting a subset. There is work on obtaining justified-representation-type guarantees for rankings [Sko+17], which could possibly be extended to the setting of incomplete votes using the techniques developed in this paper. More broadly, this article provides insights into how to fairly represent opinions of groups given incomplete information, which may be relevant for the design of more constructive online ecosystems.

4.7 Additional Proofs

4.7.1 Proofs of lower bounds for non-adaptive algorithms with exact queries

In this section, we prove Theorems 25 and 27. Both theorems can be derived as applications of the following lemma, which formalizes the properties we require of the adversarial instances shown in Figure 4.1. We only use statement (ii) in this paper, but we include statement (i) as well because we believe it may be of independent interest.

Lemma 9. *Suppose that, for some integers $0 \leqslant h \leqslant k_0 \leqslant \ell$, there exists a probability distribution $\{x_S\}_{S \subseteq [\ell]}$ over subsets of $[\ell]$ such that:*

(1) For any sets $T_1, T_2 \subseteq [\ell]$ such that $|T_1| = |T_2| \leqslant h$,

$$\sum_{S \supseteq T_1} x_S = \sum_{S \supseteq T_2} x_S.$$

(2) For some $s^* \in [\ell]$, $x_{\{s^*\}} \geqslant \frac{1}{k_0}$.

Then there exist exact query adversaries for which:

(i) For any $t \leqslant h$, any non-adaptive (k, t)-committee selection algorithm satisfies JR with probability at most $\left(\frac{k_0}{\ell}\right)^{\lfloor k/k_0 \rfloor}$.

(ii) For any $t > h$, for any $\delta > 0$, any non-adaptive (k, t)-committee selection algorithm that makes fewer than $\Omega(m^{h+1})$ queries satisfies JR with probability at most $\left(\frac{k_0}{\ell}\right)^{\lfloor k/k_0 \rfloor} + \delta$.

Proof. Given such a probability distribution $\{x_S\}_{S \subseteq [\ell]}$, we define the query adversary as follows. This adversary will be a distribution over profiles over $[m]$ candidates, for some sufficiently large m to be determined later. We denote a given non-adaptive (k,t)-committee selection algorithm by $\mathcal{A}$.

First let $k = pk_0 + r$, where p is a nonnegative integer and $0 \leqslant r < k_0$. Partition the candidates into $[m] = C_1 \cup C_2 \cup \cdots \cup C_p \cup D$ where, for each $i \in [p]$, $|C_i| = \lfloor (m-r)/p \rfloor$, and D contains the remaining candidates, of which there are at least r. For each i, the adversary will randomly select a subset of ℓ distinct candidates $S^i \coloneqq \{c_1^i, c_2^i, \ldots, c_\ell^i\} \subseteq C_i$. The adversary chooses all subsets and orderings with equal probability, independently for each C_i. The adversary will then respond to all queries according to the following approval matrix.

We partition the voters into $p+r$ distinct "parties" $P_1, \ldots, P_p$ and $Q_1, \ldots, Q_r$, and every voter is a member of exactly one party. Each party P_i contains a k_0/k proportion of voters, and voters in P_i approve only of some subset of the candidates contained in $S^i \subseteq C_i$, and none of the other candidates. For these P_i, for all $S \subseteq [\ell]$, let the fraction of voters whose approval set is exactly $\{c_s^i s.t. s \in S\}$ be equal to x_S. Every party Q_j is a $1/k$ proportion of the voters, and each voter belonging to Q_j approves only of one candidate $d_j \in D$ which is specific to Q_j.

Let us say that $\mathcal{A}$ *h-covers* a given set of candidates $S \subset [m]$ if $\mathcal{A}$ ever submits a query $T \subseteq [m]$ such that $|T \cap S| > h$. If, for any of the parties P_i with $i \in [p]$, the algorithm fails to h-cover the set S^i, then condition (1) implies that all ℓ of these candidates are completely symmetric (i.e. indistinguishable) to $\mathcal{A}$ given all of its query responses. Since each of the distinguished candidates $c_{s^*}^i$ is distributed uniformly at random among the candidates S^i, $\mathcal{A}$ selects $c_{s^*}^i$ to be part of its chosen committee with probability at most $\min(k_i/\ell, 1)$, where k_i is the number of candidates that $\mathcal{A}$ selects from P_i.

However, in order to satisfy JR, $\mathcal{A}$ must select at least r candidates from D, since there are r distinct candidates in D approved by the r parties Q_j, which are disjoint fractions of $1/k$ of the voters. In order to satisfy JR $\mathcal{A}$ must also select the distinguished candidate $c_{s^*}^i \in C_i$ for each party P_i, since condition (2) implies that for each P_i at least a $\frac{1}{k_0} \cdot \frac{k_0}{k} = \frac{1}{k}$ fraction of the voters approve only $c_{s^*}^i$ and none of the other candidates.

This already implies (i): assuming that $\mathcal{A}$ selects at least r candidates from D, then if $t \leqslant h$, it is impossible for $\mathcal{A}$ to h-cover S^i with any number of queries, and thus $\mathcal{A}$ succeeds in satisfying JR with probability at most

$$\Pr[\mathcal{A} \text{ selects } c^i_{s^*}, \text{ for all } i \in [p]] \leqslant \frac{k_1}{\ell} \cdot \frac{k_2}{\ell} \cdot \dots \cdot \frac{k_p}{\ell} = \frac{k_1 k_2 \dots k_p}{\ell^p} \leqslant \frac{k_0^p}{\ell^p} = \left(\frac{k_0}{\ell}\right)^p.$$

Here the second inequality holds due to the constraint that $k_1 + k_2 + \dots + k_p \leqslant k - r = pk_0$, since $\mathcal{A}$ must select at least r candidates from D.

To prove (ii), we must analyze the likelihood that an algorithm $\mathcal{A}$ making a small number of queries h-covers any given S^i. Let us suppose that $\mathcal{A}$ knows the partition of candidates into $C_1 \cup C_2 \cup \dots \cup C_p \cup D$, knows everything about the approval matrix except for which sets S^i were chosen within each party P_i, and is allowed to make at most cm^{h+1} queries within each party P_i, separately, where

$$c := \frac{\delta}{2^{t+\ell}\, \ell!\, p^{h+2}}.$$

Clearly, these assumptions only make the algorithm $\mathcal{A}$ stronger: an impossibility for this kind of algorithm implies the desired lower bound. Fix a party P_i. For sufficiently large m, every set $S \subseteq C$ of ℓ candidates that is h-covered by a query $T \subseteq [m]$ of size t can be decomposed into two parts: a set of size j (where $h + 1 \leqslant j \leqslant t$) that is contained in T, and a set of size $\ell - j$ that is contained in $C_i \setminus T$. Thus, the number of sets S of size ℓ within C_i that any single query can h-cover is exactly

$$\sum_{j=h+1}^{t} \binom{t}{j} \binom{\lfloor (m-r)/p \rfloor - t}{\ell - j} \leqslant 2^t \binom{m/p}{\ell - h - 1} \leqslant 2^t \left(\frac{m}{p}\right)^{\ell-(h+1)}$$

provided that $\ell - h - 1 \leqslant m/2p$, which holds for sufficiently large m.

Since we have that $\mathcal{A}$ made at most cm^{h+1} queries within party P_i by assumption, at most

$$cm^{h+1} \cdot 2^t \left(\frac{m}{p}\right)^{\ell-(h+1)} = \frac{2^t}{p^{\ell-(h+1)}} cm^\ell$$

sets of size ℓ can be h-covered. Since, within each party P_i there are a total of

$$\binom{\lfloor (m-r)/p \rfloor}{\ell} \geqslant \frac{(\lfloor (m-r)/p \rfloor - \ell)^\ell}{\ell!} \geqslant \frac{(m/(2p))^\ell}{\ell!}$$

(for sufficiently large m) sets of size ℓ in total, and each one of them is chosen to be S_i by the adversary with equal probability, the likelihood that $\mathcal{A}$ h-covered S_i is at most

$$\frac{2^t c m^\ell \ell!}{p^{\ell-(h+1)}(m/(2p))^\ell} = 2^{t+\ell}\ell! p^{h+1} c = \frac{\delta}{p}.$$

It follows from a union bound over all p of the parties P_i that the probability that $\mathcal{A}$ h-covered *any* of the S_i is at most δ. For $\mathcal{A}$ to satisfy JR, it is necessary for it to either h-cover some S_i with the initial queries or subsequently select every $c^i_{s^*}$ after having failed to h-cover any S_i. By the union bound, the probability that $\mathcal{A}$ satisfies JR is at most the sum of the probabilities of these two events, which is at most $\left(\frac{k_0}{\ell}\right)^p + \delta$. ☐

Thus, to prove lower bounds against non-adaptive algorithms with exact queries, it suffices to construct probability distributions over subsets of a finite set $[\ell]$ with certain special properties. To prove our $\Omega(m^2)$ lower bound, which holds for any $k \geqslant 2$, we generalize the construction from Figure 4.1 (a) by simply adding more candidates to the larger approval set:

Proof of Theorem 25. Given any t and $\varepsilon \in (0,1]$, let $h = 1$, $k_0 = 2$, and $\ell = \lceil 4/\varepsilon \rceil$. Consider the probability distribution over subsets of $[\ell]$ where $x_{\{1\}} = \frac{1}{2}$, $x_{\{2,3,4,\ldots,\ell\}} = \frac{1}{2}$, and all other sets have probability zero (Figure 4.1 (a) shows the special case of this distribution where $\ell = 3$). Notice that these parameters meet all the requirements of Lemma 9 with $s^* = 1$. Letting $\delta \coloneqq \varepsilon/2$, it follows that, for any $k \geqslant k_0 = 2$, any (k,t)-committee selection algorithm that makes fewer than $\Omega(m^2)$ queries satisfies JR with probability at most

$$\left(\frac{k_0}{\ell}\right)^{\lfloor k/k_0 \rfloor} + \delta \leqslant \left(\frac{k_0}{\ell}\right)^1 + \delta = \frac{2}{\ell} + \delta \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$ ☐

To prove stronger lower bounds we need to increase the h parameter. Probability distributions $\{x_S\}_{S \subseteq [\ell]}$ satisfying the hypotheses of Lemma 9 prove difficult to construct by hand for $h > 1$, so we conducted a computational search. By a straightforward averaging argument, one can see that it is without loss of generality to consider "symmetric" distributions, where for any sets $S, T \subseteq [\ell]$ of the same size that either both contain s^* or both do not contain s^*, $x_S = x_T$. Thus, it suffices to consider solutions encoded as points in the following polyhedron, which we refer to as $P(h, k_0, \ell) \subseteq \mathbb{R}^{2\ell}$. We parameterize the space by the 2ℓ variables

$$\{x_{i,j} s.t. i \in \{0,1\}, j \in \{0,1,2,\ldots,\ell-1\}\},$$

where $x_{0,j}$ encodes the value of x_S for all S of size j that do not include s^*, and $x_{1,j}$ encodes the value of x_S for all S containing s^* and j other elements from $[\ell]$. For a solution to be in $P(h, k_0, \ell)$, there are four kinds of constraints it must satisfy.

- All probabilities must be nonnegative: for all $i \in \{0,1\}$ and $j \in \{0,1,2,\ldots,\ell-1\}$,

$$x_{i,j} \geqslant 0.$$

- Probabilities must all sum to 1:

$$\sum_{i=0}^{1}\sum_{j=0}^{\ell-1}\binom{\ell-1}{j}x_{i,j} = 1.$$

- Condition (1) from Lemma 9 must be satisfied. Due to the symmetry that is baked in to the solutions we're considering, we only need to check the constraint for pairs of sets where s^* is contained in one set but not the other. This constraint is as follows: for all $t' \in [h]$,

$$\sum_{i=0}^{1}\sum_{j=t'}^{\ell-1}\binom{\ell-1-t'}{j-t'}x_{i,j} = \sum_{j=t'-1}^{\ell-1}\binom{\ell-t'}{j-t'+1}x_{1,j}.$$

- Condition (2) from Lemma 9 must be satisfied:

$$x_{1,0} \geqslant \frac{1}{k_0}.$$

Thus, there exists a probability distribution satisfying the hypotheses of Lemma 9 if and only if $P(h, k_0, \ell)$ is nonempty. The description of this polyhedron is only of polynomial-size, so we can solve it efficiently using linear programming. However, many of the coefficients are extremely large, and we eventually ran into numerical difficulties. Table 4.2 lists the tightest lower bounds we were able to obtain in terms of how large k_0 had to be for a given value of h.

h	1	2	3	4	5	6	7	8	9	10
k_0	2	6	10	16	21	30	38	49	59	72

Table 4.2:: For each positive integer h, smallest value of k_0 for which $P(h, k_0, k_0 + 1)$ is nonempty, i.e., the smallest committee size for which Lemma 9 implies that guaranteeing JR requires $\Omega(m^{h+1})$ exact queries of size $t > h$. The constraints are tight only for $h \in \{1, 2\}$.

Proof. Proof of Theorem 27 Let $\varepsilon > 0$ be given. Then let

$$k \coloneqq 72\left(\frac{\log(2/\varepsilon)}{\log(73/72)} + 1\right)$$

and $\delta \coloneqq \varepsilon/2$. As Table 4.2 claims, Lemma 9 holds for $h = 10$, $k_0 = 72$, and $\ell = 73$. Thus, for any t, any non-adaptive (k, t)-committee selection algorithm that makes fewer than $\Omega(m^{11})$ queries satisfies JR with probability at most

$$\left(\frac{72}{73}\right)^{\lfloor k/72 \rfloor} + \delta \leqslant \left(\frac{72}{73}\right)^{(k/72)-1} + \delta = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

We note that there is a gap between these results and Theorem 28. An intriguing direction for future work is to obtain matching upper and lower bounds for the query complexity of guaranteeing JR using a non-adaptive algorithm with exact queries. For $k = 1$ we need $\Theta(m)$ queries, and for $k \in \{2, 3\}$, we need $\Theta(m^2)$ queries. However, the complexity is unknown for all larger k, and we conjecture that the exponent of m grows as a polynomial function of k.

4.7.2 Family of examples for noisy vs exact queries

Fix some $k \geqslant 4$, t, and m. We will construct a family of instances on m candidates where there exists a non-adaptive exact-query algorithm which can guarantee JR using $\lceil m/t \rceil$ queries while a non-adaptive noisy-query algorithm necessarily needs $\Omega(m \log(m)/t)$ to guarantee it with any fixed probability δ. We describe the approval profile by the distribution over approval sets by sampling a voter uniformly at random. There is one special candidate a^*. This candidate a^* is approved by a $2/k$ fraction of the electorate while all other candidates b are approved by $1/(2k)$. Further, these approvals are independent in the sense that when we sample a voter, the joint distribution over approvals is as if each of these approvals were selected independently. For example, given a set $S \subseteq C$ of candidates such that S contains a^* and ℓ other candidates, the proportion of voters who approve exactly the set S is $(2/k) \cdot (1/(2k))^\ell \cdot (1 - 1/(2k))^{m-1-\ell}$.

The first observation we make is that the committees that satisfy JR are exactly those that include a^*. Notice that if a committee includes a^*, no other candidate has enough approval support ($1/k$) to have a blocking coalition to violate JR. On the other hand, if there is a committee W of size k that does not include a^*, we can compute the proportion of voters that approve a^* that do not approve of any candidates in W. This is

$$\frac{2}{k} \cdot \left(1 - \frac{1}{2k}\right)^k > \frac{2}{k}\left(1 - \frac{k}{2k}\right) = \frac{1}{k}.$$

Hence, for such a W, there would exist a sufficiently large blocking coalition for a^*.

Next, we show that there is a non-adaptive exact-query algorithm that can guarantee JR for any instance of this form (i.e., regardless of which candidate is a^*). Indeed, it simply makes $\lceil m/t \rceil$ queries that cover all candidates. From this, it can deduce candidate approval scores and ensure that the committee it chooses contains the candidate with approval score $2/k$.

Finally, let us consider a non-adaptive algorithm that makes ℓ queries and guarantees JR with probability $1 - \delta$ regardless of which candidate is a^*. Notice that such an algorithm should guarantee JR with this same probability against a distribution of instances where a^* is selected uniformly at random. Let us consider an algorithm A that maximizes the probability of selecting a JR committee against this distribution. Notice that it is without loss of generality that A is deterministic by Yao's minimax principle. We show that this can only be done if $\ell \geqslant f(m)$ where $f(m) \in \Omega(m \log m)$ (treating k, t, and δ as constants) is a function to be defined later.

Suppose for a contradiction $\ell < f(m)$. Let H be the set of candidates that appear in strictly more than $q := \frac{2tf(m)}{m\delta}$ and let L be the remaining candidates. Notice that $|H| \leqslant \delta/2 \cdot m$, as otherwise $\ell \geqslant f(m)$. We show that conditioned on $a^* \in L$, the probability A chooses a committee containing a^* is at most $\delta/2$. This implies that A's probability of success is at most $(1-\delta/2) \cdot \delta/2 + \delta/2 < \delta$.

To that end, consider an algorithm that receives extra queries such that all candidates in L are in exactly q queries. Notice that conditioned on a^* being in L, since all candidates in L are in the same number of queries, the optimal strategy to maximize the probability a^* is a committee-member is to take the k candidates in L with highest empirical approval score. Indeed, this dominates any other strategy as conditioned on any empirical approval scores, this choice of committee covers the maximum likelihood estimates of the underlying distribution.

What we finally show is that with probability at least $1 - \delta/2$, conditioned on $a^* \in L$, a^* will *not* be among the k highest approval scores. Intuitively, with reasonably high probability a^* will have empirical not too much more than it's true approval, say at most $3/k$, while, by choosing $q \in O(\log m)$, due to the noise in estimating empirical approval scores, at least k of the remaining candidates in L will have approval score this large. Indeed, ensuring $q > \frac{k^2 \log(4/\delta)}{2}$ ensures the empirical estimate of a^* is less than $3/k$ with probability at least $1 - \delta/4$ using standard Hoeffding's inequality. For the other emprical means, using tail bounds on the Binomial distribution [Ash90], the probability they are at least $3/k$ is at least $\frac{1}{\sqrt{2q}} \exp(-q\Theta(k))$. Notice we can choose $q \in \Theta(\log m)$ such that this value is at least $2(k + \log(4/\delta))/m$. For sufficiently large m, this choice of q will be above $\frac{k^2 \log(4/\delta)}{2}$, leading to a valid $\Theta(m \log m)$ function of f. Further, again applying standard Hoeffding's inequality on the remaining at least $m/2$ candidates in L shows that at least k will satisfy this, as needed.

4.7.3 Proofs for the Adaptive Exact-Query Setting

In the following, we prove Theorem 29, Lemma 8, and Theorem 28.

Proof of Theorem 29

Theorem 29 and its proof are based on the lower bound from Lemma 1 in [Sko21]. Our result is more general in two ways: (1) our statement holds for any committee W, no matter what algorithm computed it, and (2) we introduce an approximation parameter α. We begin with the following intermediate lemma:

Lemma 10. *For any committee $W \subseteq C$ and group of voters $V \subseteq N$, we have*

$$avs_W(V) \geqslant \min\left\{\left|\bigcap_{i\in V} A_i\right|, \frac{1}{n}\cdot\frac{|V|}{\Delta^*(W)} - 1\right\}.$$

Proof. As mentioned, the following proof is closely related to the proof of Lemma 1 of Skowron [Sko21]. Suppose there exist V and W such that both

$$\frac{1}{|V|}\sum_{i\in V}|W\cap A_i| < \left|\bigcap_{i\in V} A_i\right| \text{ and } \frac{1}{|V|}\sum_{i\in V}|W\cap A_i| < \frac{1}{n}\cdot\frac{|V|}{\Delta^*(W)} - 1.$$

We then have

$$\begin{aligned}\left|\bigcap_{i\in V} A_i\right| &> \frac{1}{|V|}\sum_{i\in V}|W\cap A_i| \\ &\geqslant \frac{1}{|V|}\sum_{i\in V}\left|W\cap\left(\bigcap_{j\in V} A_j\right)\right| \\ &= \left|W\cap\left(\bigcap_{j\in V} A_j\right)\right|.\end{aligned}$$

This implies $W\cap\left(\bigcap_{i\in V} A_i\right) \subsetneq \bigcap_{i\in V} A_i$, so $\overline{W}\cap\left(\bigcap_{i\in V} A_i\right) \neq \emptyset$. Hence, there is a candidate $c \in \overline{W}\cap\left(\bigcap_{i\in V} A_i\right)$ that is not on the committee W, but is approved by all voters in V. For such a candidate c, we have

$$\begin{aligned}
\Delta(W, c) &= \frac{1}{n} \sum_{i \in N : c \in A_i \setminus W} \frac{1}{|A_i \cap W| + 1} \\
&= \frac{1}{n} \sum_{i \in N : c \in A_i} \frac{1}{|A_i \cap W| + 1} && (c \notin W) \\
&\geqslant \frac{1}{n} \sum_{i \in V} \frac{1}{|A_i \cap W| + 1} && (c \in \textstyle\bigcap_{i \in V} A_i) \\
&\geqslant \frac{1}{n} |V| \frac{1}{\frac{1}{|V|} \sum_{i \in V} (|W \cap A_i| + 1)} && (\text{convexity of } 1/x) \\
&> \frac{1}{n} |V| \frac{1}{\frac{1}{n} \frac{|V|}{\Delta^*(W)} - 1 + 1} \\
&= \Delta^*(W),
\end{aligned}$$

a contradiction, as $\Delta(W, x) \leqslant \Delta^*(W)$ for all candidates x. $\square$

We are now ready to prove Theorem 29

Proof. Fix a committee W satisfying $\Delta^*(W) < \frac{1}{\alpha k}$. We begin with α-OAS. Fix a $\lambda \in [0, k]$, and a $\frac{\lambda+1}{\alpha}$-large, λ-cohesive group of voters V. By definition of λ-cohesive, $\left|\bigcap_{i \in V} A_i\right| \geqslant \lambda$. Further, we have

$$\begin{aligned}
\frac{1}{n} \cdot \frac{|V|}{\Delta^*(W)} - 1 &\geqslant \frac{1}{n} \cdot \frac{1}{\Delta^*(W)} \cdot \frac{\lambda + 1}{\alpha} \cdot \frac{n}{k} - 1 && (V \text{ is } \tfrac{\lambda+1}{\alpha}\text{-large}) \\
&= \frac{1}{\Delta^*(W)} \cdot \frac{\lambda + 1}{\alpha} \cdot \frac{1}{k} - 1 \\
&> \lambda + 1 - 1 && (\Delta^*(W) < \tfrac{1}{\alpha k}) \\
&= \lambda
\end{aligned}$$

Together, these imply that

$$\min \left\{ \left| \bigcap_{i \in V} A_i \right|, \frac{1}{n} \cdot \frac{|V|}{\Delta^*(W)} - 1 \right\} \geqslant \lambda.$$

Invoking Lemma 10, we have $\mathrm{avs}_W(V) \geqslant \lambda$, as needed.

Next, we show α-EJR. Fix an $\ell \in [k]$, and an $\frac{\ell}{\alpha}$-large, ℓ-cohesive group of voters V. As before, by the definition of ℓ-cohesive, we have $\left|\bigcap_{i \in V} A_i\right| \geqslant \ell$. Further, by the same argument as above with $\ell = \lambda + 1$,

$$\frac{1}{n} \cdot \frac{|V|}{\Delta^*(W)} - 1 > \ell - 1.$$

Together, these imply that

$$\min\left\{\left|\bigcap_{i \in V} A_i\right|, \frac{1}{n} \cdot \frac{|V|}{\Delta^*(W)} - 1\right\} > \ell - 1.$$

Invoking Lemma 10, we have $\operatorname{avs}_W(V) > \ell - 1$, and since utilities are integers, this implies that $|A_i \cap W| \geqslant \lceil \operatorname{avs}_W(V) \rceil \geqslant \ell$ for at least one voter $i \in V$, as needed. □

Proof of Lemma 8

Proof. Fix W and $c \notin W$. We will use the notation $W^+ := W \cup \{c\}$. First, we show that

$$\min_{x \in W} \Delta(W^+ \setminus \{x\}, x) \leqslant \frac{1 - \Delta(W, c)}{k}. \tag{4.1}$$

To that end, let us consider $\Delta(W^+ \setminus \{x\}, x)$ for an arbitrary $x \in W^+$. We have

$$\begin{aligned}\Delta(W^+ \setminus \{x\}, x) &= \textsc{pav-sc}(W^+) - \textsc{pav-sc}(W^+ \setminus \{x\}) \\ &= \frac{1}{n} \sum_{i \in N : x \in A_i} \frac{1}{|W^+ \cap A_i|}.\end{aligned}$$

Adding up over all $x \in W^+$, we have

$$\begin{aligned}\sum_{x \in W^+} \Delta(W^+ \setminus \{x\}, x) &= \frac{1}{n} \sum_{x \in W^+} \sum_{i \in N : x \in A_i} \frac{1}{|W^+ \cap A_i|} \\ &= \frac{1}{n} \sum_{i : W^+ \cap A_i \neq \emptyset} \frac{|W^+ \cap A_i|}{|W^+ \cap A_i|} \\ &\leqslant 1.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\sum_{x \in W^+} \Delta(W^+ \setminus \{x\}, x) &= \Delta(W, c) + \sum_{x \in W} \Delta(W^+ \setminus \{x\}, x) \\ &\geqslant \Delta(W, c) + k \cdot \min_{x \in W} \Delta(W^+ \setminus \{x\}, x).\end{aligned}$$

Combining these two inequalities, we get that

$$\Delta(W, c) + k \cdot \min_{x \in W} \Delta(W^+ \setminus \{x\}, x) \leqslant 1.$$

Rearranging yields (4.1). Finally, notice that

$$\max_{x \in W} \Delta(W, c, x) = \Delta(W, c) - \min_{x \in W} \Delta(W^+ \setminus \{x\}, x).$$

Hence, by (4.1),

$$\max_{x \in W} \Delta(W, c, x) \geqslant \Delta(W, c) - \left(\frac{1 - \Delta(W, c)}{k}\right) = \frac{(k+1)\Delta(W, c) - 1}{k},$$

as needed. □

Proof of Theorem 28

Proof. Clearly, if Algorithm 4 terminates, the resulting committee W satisfies

$$\Delta^*(W) = \Delta(W, c') < \frac{1}{\alpha k}$$

and hence α-EJR and α-OAS by Theorem 29. What remains is to bound how many steps the algorithm takes to terminate. To do this, we use the PAV score of the current committee as a potential function. In every iteration of the loop for which the algorithm does not terminate, we have

$$\Delta(W, c') \geqslant \frac{1}{\alpha k}$$

and hence, by Lemma 8, the increase in PAV score at each step will be

$$\max_{x} \Delta(W, c', x) \geqslant \frac{(1-\alpha)k + 1}{\alpha k^2}.$$

Notice that the minimum and maximum PAV score that can be possibly attained by any committee of size k are 0 (when nobody approves of any candidate) and H_k (the harmonic number which is attained when everyone approves of every candidate) respectively. Hence, there can be at most

$$\frac{\alpha k^2}{(1-\alpha)k + 1} H_k$$

steps. Since at each step we make $\left\lceil \frac{m-k}{t-k} \right\rceil$ queries, the result follows. $\square$

4.7.4 PAV and LS-PAV Yield a Committee That Satisfies $\Delta^*(W) < 1/k$

As mentioned previously, it is known that PAV and LS-PAV satisfy both EJR and OAS. Here, we show that they yield committees that satisfy $\Delta^*(W) < 1/k$, which implies EJR and OAS through Theorem 29. This is noteworthy because 1) it has a much simpler proof, 2) it implies that PAV and LS-PAV committees can be certified in a computationally efficient manner, by verifying that $\Delta^*(W) < 1/k$.

Lemma 11. *For both PAV and LS-PAV, the returned committee W satisfies $\Delta^*(W) < 1/k$.*

Proof. For the committee W computed by PAV or LS-PAV, we have that for any candidate $c \notin W$,

$$\max_{x \in W} \Delta(W, c, x) < \frac{1}{k^2}$$

since otherwise the PAV score of W could be improved by at least $1/k^2$ by adding c and removing the worst candidate. By Lemma 8, this implies that

$$\frac{(k+1)\Delta^*(W) - 1}{k} < \frac{1}{k^2}.$$

Rearranging yields $\Delta^*(W) < \frac{1}{k}$, so the result follows from Theorem 29. $\square$

4.7.5 Proof of Theorem 30

We now provide a proof of Theorem 30.

Proof. We first show that with probability at least $1-\delta$, all $\widehat{\Delta}$ estimates in the first $H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}$ iterations of the loop are within $\pm\varepsilon := \frac{(1-\alpha)k+1}{12\alpha k^2}$ of the corresponding true Δ values. Then, we show that conditioned on these accurate estimates, the algorithm satisfies the theorem properties.

To show the error bounds, we will use a straightforward application of Hoeffding's inequality. Indeed, when the corresponding ℓ voters are sampled, notice that $\widehat{\Delta}$ is simply the sample mean of independent samples with expectation of the corresponding Δ. Further, these samples are always proportions falling in $[-1, 1]$. Hence, any specific estimate will not be within $\pm\varepsilon$ with probability at most $2\exp(-\varepsilon^2\ell/2)$. Note that there are $m-k$ choices of x for $\widehat{\Delta}(W, x)$ and $(m-k) \cdot k$ choices of (x, y) pairs for $\widehat{\Delta}(W, x, y)$. Hence, there are a total of $(m-k)(k+1)$ estimates per iteration. Therefore, there are at most $H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}(m-k)(k+1)$ estimates in the first $H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}$ iterations. A union bound tells us the probability that all estimates in these iterations are within $\pm\varepsilon$ is at least

$$1 - 2\exp(-\varepsilon^2\ell/2) \cdot H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}(m-k)(k+1).$$

We simply need to show that this value is at least $1-\delta$.

To that end, recall that

$$\ell = \left\lceil 288 \left(\frac{\alpha k^2}{(1-\alpha)k+1}\right)^2 \log\left(\frac{8mk^4}{\delta}\right)\right\rceil.$$

Noting that $H_k \leqslant k$ and $\frac{\alpha k^2}{(1-\alpha)k+1} \leqslant k^2$, we have that

$$8mk^4 \geqslant 2k \cdot (2k^2) \cdot m \cdot (2k) \geqslant 2H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}(m-k)(k+1).$$

Hence,

$$\ell \geqslant \frac{2}{\varepsilon^2} \log\left(\frac{2H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}(m-k)(k+1)}{\delta}\right),$$

so we have

$$2\exp\left(-\varepsilon^2\ell/2\right) \leqslant \frac{\delta}{H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}(m-k)(k+1)},$$

as needed.

Next, condition on all of these estimates being accurate. Notice if the algorithm terminates within the first $H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}$ iterations, this means that for the returned committee, $\max_{x \notin W} \widehat{\Delta}(W, x) < \frac{1}{\alpha k} - \frac{(1-\alpha)k+1}{8\alpha k^2} = \frac{1}{\alpha k} - \varepsilon$. By our assumption about the accuracy of each $\widehat{\Delta}$, we have that $\Delta^*(W) = \max_{x \notin W} \widehat{\Delta}(W, x) < \frac{1}{\alpha k}$. Hence, by Theorem 29, W satisfies the desired properties. Further, there are $\left\lceil \frac{m-k}{t-k} \right\rceil \cdot \ell$ queries per iteration. Noting that

$$\ell \leqslant 289 \left(\frac{\alpha k^2}{(1-\alpha)k+1} \right)^2 \log \left(\frac{8mk^4}{\delta} \right)$$

to avoid the ceiling, this means the total query complexity is at most

$$H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1} \cdot \left\lceil \frac{m-k}{t-k} \right\rceil \cdot \ell \leqslant 578 H_k \left\lceil \frac{m-k}{t-k} \right\rceil \left(\frac{\alpha k^2}{(1-\alpha)k+1} \right)^3 \log \left(\frac{8mk^4}{\delta} \right)$$

as needed.

What remains to be shown is that conditioned on the accurate estimates, the algorithm terminates within $H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}$ iterations. Indeed, we show that each iterations, the PAV score of W increases by at least $\frac{(1-\alpha)k+1}{2\alpha k^2}$. As the minimum and maximum PAV scores of a committee are 0 and H_k respectively, this can occur at most $H_k \cdot \frac{2\alpha k^2}{(1-\alpha)k+1}$ times. Hence, we obtain the desired bound on the number of iterations.

To that end, note that when when we make a swap of c' for c, it must be the case that $\widehat{\Delta}(W, c) > \frac{1}{\alpha k} - \varepsilon$. Using our assumptions on $\widehat{\Delta}$ errors, this implies that $\Delta(W, c) \geqslant \frac{1}{\alpha k} - 2\varepsilon$. By Lemma 8, we have that

$$\max_{x \in W} \Delta(W, c, x) \geqslant \frac{(1-\alpha)k+1}{\alpha k^2} - \frac{k+1}{k} 2\varepsilon \geqslant \frac{(1-\alpha)k+1}{\alpha k^2} - 4\varepsilon.$$

Again, by our assumption on $\widehat{\Delta}$ errors,

$$\max_{x \in W} \widehat{\Delta}(W, c, x) \geqslant \max_{x \in W} \Delta(W, c, x) - \varepsilon \geqslant \frac{(1-\alpha)k+1}{\alpha k^2} - 5\varepsilon.$$

Finally, for the choice c' that maximizes $\widehat{\Delta}(W, c, c')$,

$$\Delta(W, c, c') \geqslant \widehat{\Delta}(W, c, c') - \varepsilon \geqslant \frac{(1-\alpha)k+1}{\alpha k^2} - 6\varepsilon = \frac{(1-\alpha)k+1}{2\alpha k^2},$$

as needed. $\square$

4.7.6 Description and Analysis of Algorithm 6

In this section, we state Algorithm 6 and analyze its complexity. The worst-case query guarantees are slightly worse than those of Algorithm 5; however, as we discuss below, there are instances where Algorithm 6 performs better, and this can additionally be seen in the experiments of Section 4.5.

Theorem 31. *For any $m \geqslant t > k$, with probability $1 - \delta$, Algorithm 6 yields a committee satisfying α-OAS and α-EJR after querying at most*

$$1152 H_k \left\lceil \frac{m-k}{t-k} \right\rceil \left(\frac{\alpha k^2}{(1-\alpha)k+1} \right)^3 \log \left(\frac{4608 k^8 m^{k+2}}{\delta} \right)$$

voters. For any fixed $\delta > 0$, if $\alpha = 1$, this leads to a query complexity of $\mathcal{O}\left(mk^7 \log k \log m\right)$, and if $\alpha < 1$, this leads to a query complexity of $\mathcal{O}\left(mk^4 \log k \log m\right)$.

We use

$$L := 1152 H_k \left\lceil \frac{m-k}{t-k} \right\rceil \left(\frac{\alpha k^2}{(1-\alpha)k+1} \right)^3 \log \left(\frac{4608 k^8 m^{k+2}}{\delta} \right)$$

to denote this upper bound (notice that it is the same as the L from the algorithm).

Importantly, this theorem states that despite the extensions we introduced in Algorithm 6, it remains theoretically sound: with a sufficient number of samples, it yields a committee satisfying OAS and EJR.

The proof follows a relatively similar structure to Theorem 30: we first show that with probability $1 - \delta$, many estimates are sufficiently accurate, and conditioned on this, the algorithm makes progress in terms of PAV score and terminates with a good committee. However, unlike Theorem 30, the samples we take are not fresh for each round, so we can not directly apply Hoeffding's inequality in the most straightforward way. Nonetheless, the proof goes through by instead treating the Δ estimates as Martingales in order to use Azuma's inequality. Due to its additional intricacy, we separate this portion into its own lemma.

Lemma 12. *With probability $1 - \delta$, at every step after querying up to L voters,*

$$\Delta(W, x) \leqslant \widetilde{\Delta}^{+}(W, x) \leqslant \Delta(W, x) + 2err_k(W, x)$$

Algorithm 6 (k,t)-ucb-α-PAV.

```
Choose W ∈ (C choose k) arbitrarily
Q ← {} // List to store queries and responses
ℓ ← 576·(αk²/((1−α)k+1))² log(4608k⁸m^(k+2)/δ) // Constant to be used later
L ← 2H_k ⌈(m−k)/(t−k)⌉ (αk²/((1−α)k+1))·ℓ // Constant to be used later
for i = 1, 2, ... do
    V_s(W,x) ← {i s.t. x ∈ Q_i and |Q_i ∩ W| ⩾ s}
        // ∀s ∈ {0} ∪ [k], ∀x ∉ W,
    V_s(W,x,y) ← {i s.t. {x,y} ⊆ Q_i and |Q_i ∩ W| ⩾ s}
        // ∀s ∈ [k], ∀x ∉ W, ∀y ∈ W
    Δ̂_s^+(W,x) ← (1/|V_s|) Σ_{i∈V_s(W,x)} 𝕀[x∈R_i]/(|R_i∩W|+1) if V_s(W,x) ≠ ∅ else ∞
      // Upper bound on estimate for Δ(W,x) using voters queried on ⩾ s of W (along with x)
    Δ̂_s^-(W,x,y) ← (1/|V_s(W,x,y)|) Σ_{i∈V_s(W,x,y)} 𝕀[x∈R_i and y∉R_i]/(|R_i∩W|+|W\Q_i|+1) − 𝕀[x∉R_i and y∈R_i]/|R_i∩W| if V_s(W,x,y) ≠ ∅
        else −∞
      // Lower bound on estimate for Δ(W,x,y) using voters queried on ⩾ s of W (and x, y)
    err_s(W,x) ← sqrt( 2 log(4L(k+1)m^(k+1)/δ) / |V_s(W,x)| )
    err_s(W,x,y) ← sqrt( 2 log(4L(k+1)m^(k+1)/δ) / |V_s(W,x,y)| )
    Δ̃^+(W,x) ← min_{s∈[k]} Δ̂_s^+(W,x) + err_s(W,x) // Best UCB-style UB on Δ(W,x) given queries
    Δ̃^-(W,x,y) ← max_{s∈[k]} Δ̂_s^-(W,x,y) − err_s(W,x,y)
        // Best UCB-style lower bound on Δ(W,x,y) given queries
    c' ← arg max_{x∉W} Δ̃^+(W,x)
    if Δ̃^+(W,c') < 1/(αk) then
        return W
    c ← arg max_{x∈W} Δ̃^-(W,c',x)
    if Δ̃^-(W,c',c) ⩾ (1/2)((1−α)k+1)/(αk²) then
        W ← (W ∪ {c'}) \ {c}
    else
        A ← {x ∈ C s.t. |{i s.t. W ∪ {x} ⊆ Q_i}| ⩾ ℓ}
            // Candidates already queried more than ℓ times with W
        S ← C\W\A // Potential candidates to query along with W
        Make query Q_i on W and t − k candidates x ∉ S with highest Δ̃^+(W,x), breaking
            ties arbitrarily
        Receive response R_i and append (i, Q_i, R_i) to Q
```

and

$$\Delta(W,x,y) - 2err_k(W,x,y) \leqslant \widetilde{\Delta}^-(W,x,y) \leqslant \Delta(W,x,y)$$

for all committees W, $x \notin W$, and $y \in W$.

Proof. We begin by considering estimates of the form $\widetilde{\Delta}^-(W,x,y) \leqslant \Delta(W,x,y)$; the rest of the estimates will follow similar arguments which we discuss later. Fix an arbitrary committee W, $x \notin W$, $y \in W$, and $s \in [k]$. We define a sequence of random variables $X_0, X_1, \dots$ where X_j is the unnormalized estimate $|V_s(W,x,y)| \cdot \widehat{\Delta}_s^-(W,x,y)$ when $|V_s(W,x,y)| = j$, i.e., when j voters have been queried on x, y and at least s candidates of W, and $X_0 = 0$. In other words, when the j^{th} voter of $V_s(W,x,y)$ is queried, X_j is X_{j-1} plus that j^{th} voters estimate for $\Delta(W,x,y)$, $\frac{\mathbb{I}[x \in R_i \text{ and } y \notin R_i]}{|R_i \cap W| + |W \setminus Q_i| + 1} - \frac{\mathbb{I}[x \notin R_i \text{ and } y \in R_i]}{|R_i \cap W|}$.

Notice that when the j^{th} voter is queried, regardless of the algorithm's choices of when to make such a query, this is simply a random voter from the population chosen independently of everything else. Hence, if their entire approval set was known, the expectation of their estimate of $\Delta(W,x,y)$ would be exactly $\Delta(W,x,y)$. When only part W intersects the query, we choose a bound that would always upper bounds the true estimate. Therefore, $\mathbb{E}[X_j \mid X_{j-1}] \geqslant X_{j+1} + \Delta(W,x,y)$.

Let $Y_0, Y_1, \dots$ be the additive errors of X_j from the true $\Delta(W,x,y)$, that is, $Y_j = X_j - j \cdot \Delta(W,x,y)$. The key observation we will make is that the sequence $Y_0, Y_1, Y_2, \dots$ is, in fact, a submartingale. Indeed, since $Y_j = X_j - j \cdot \Delta(W,x,y)$ and $Y_{j-1} = X_{j-1} - (j-1) \cdot \Delta(W,x,y)$, we have $\mathbb{E}[Y_j \mid Y_{j-1}] \geqslant 0$.

Additionally, note that an individual voter's Δ estimate is always within $[-1,1]$, so $X_j - X_{j-1} \in [-1,1]$. Using the definition of Y_j, we have that this implies $Y_j - Y_{j-1} \in [-1 - \Delta(W,x,y), 1 - \Delta(W,x,y)]$. Note that this is a range of size 2, and we can hence use (the asymmetric version of) Azuma's inequality to get that for all $\varepsilon > 0$,

$$\Pr[Y_j \leqslant -\varepsilon] = \Pr[Y_j - Y_0 \leqslant -\varepsilon] \leqslant \exp\left(-\frac{2\varepsilon}{j \cdot 2^2}\right) = \exp\left(-\frac{\varepsilon}{2j}\right).$$

Using this, we can now analyze the errors. When $|V_s(W,x,y)| = j$ for any such j,

$$\begin{aligned}
\Pr[\widehat{\Delta}_s^-(W,x,y) + err_s(W,x,y) \leqslant \Delta(W,x,y)]& \\
&= \Pr[\widehat{\Delta}_s(W,x,y) - \Delta(W,x,y) \leqslant -err_s(W,x,y)] \\
&= \Pr[j \cdot \widehat{\Delta}_s(W,x,y) - j \cdot \Delta(W,x,y) \leqslant -j \cdot err_s(W,x,y)] \\
&= \Pr[X_j - j \cdot \Delta(W,x,y) \leqslant -j \cdot err_s(W,x,y)] \\
&= \Pr[Y_j \leqslant -j \cdot err_s(W,x,y)] \\
&= \Pr\left[Y_j \leqslant -j \cdot \sqrt{\frac{2\log\left(\frac{4L(k+1)m^{k+1}}{\delta}\right)}{j}}\right] \\
&= \Pr\left[Y_j \leqslant -\sqrt{2j\log\left(\frac{4L(k+1)m^{k+1}}{\delta}\right)}\right] \\
&\leqslant \exp\left(-\frac{\left(\sqrt{2j\log\left(\frac{4L(k+1)m^{k+1}}{\delta}\right)}\right)^2}{2j}\right) \\
&= \frac{\delta}{4L(k+1)m^{k+1}}.
\end{aligned}$$

Additionally, note that when $s = k$, this is in fact a martingale (no loose upper bounding is needed), so this inequality continues to hold in other direction for $\widehat{\Delta}_k^-(W,x,y) - err_k(W,x,y) \geqslant \Delta(W,x,y)$. A symmetric argument shows

$$\Pr[\widehat{\Delta}_s^-(W,x) - err_s(W,x) \geqslant \Delta(W,x)] \leqslant \frac{\delta}{4L(k+1)m^{k+1}}$$

and

$$\Pr[\widehat{\Delta}_k^-(W,x) + err_s(W,x) \leqslant \Delta(W,x)] \leqslant \frac{\delta}{4L(k+1)m^{k+1}}.$$

for all W, x, and s.

Notice that in the first L queries, the sizes of the V_s sets are trivially upper bounded by L. Hence, we can union bound over all at most L sizes, the two choices of either upper and lower bounds, two choices of either $\Delta(W,x,y)$ or $\Delta(W,x)$ the at most $k+1$ choices of s, and at most m^{k+1} choices of W, x, and y (we are choosing $m+1$ candidates with two being special, so clearly at most choosing a sequence of $k+1$ candidates with repeats). This leads to at most $4L(k+1)m^{k+1}$ possible bad events. Hence, with probability $1-\delta$, none of these bad events happen. Conditioned on this, we have that

$$\widetilde{\Delta}^{+}(W,x) = \min_{s\in[k]} \widehat{\Delta}_s^{+}(W,x) + err_s(W,x) \geqslant \Delta(W,x)$$

and

$$\widetilde{\Delta}^{-}(W,x,y) = \max_{s\in[k]} \widehat{\Delta}_s^{-}(W,x) - err_s(W,x) \geqslant \Delta(W,x).$$

In addition, using the bounds on $\widehat{\Delta}_k$, we have

$$\begin{aligned}\widetilde{\Delta}^{+}(W,x) &= \min_{s\in\{0\}\cup[k]} \widehat{\Delta}_s^{+}(W,x) + err_s(W,x)\\ &\leqslant \widehat{\Delta}_k^{+}(W,x) + err_k(W,x)\\ &\leqslant \Delta(W,x) + 2err_k(W,x),\end{aligned}$$

and

$$\begin{aligned}\widetilde{\Delta}^{-}(W,x,y) &= \max_{s\in\cup[k]} \widehat{\Delta}_s^{-}(W,x,y) - err_s(W,x,y)\\ &\geqslant \widehat{\Delta}_k^{-}(W,x,y) - err_k(W,x,y)\\ &\geqslant \Delta(W,x,y) - 2err_k(W,x,y).\end{aligned}$$

Hence, the desired bounds are satisfied. □

We are now ready to prove the theorem.

Proof of Theorem 31. We condition on the event that the estimates after at most L voters are all accurate as in Lemma 12. Let $\ell := 576 \cdot \left(\frac{\alpha k^2}{(1-\alpha)k+1}\right)^2 \log\left(\frac{4608k^8m^{k+2}}{\delta}\right)$ as defined in the algorithm. The technical portion of this proof is to show that for any committee W, after at most $\left\lceil\frac{m-k}{t-k}\right\rceil \cdot \ell$ queries, the algorithm either makes a swap or terminates. Notice that when a swap is made, assuming the estimate is accurate, the PAV score increases by $\frac{(1-\alpha)k+1}{2\alpha k^2}$. Hence, just as in previous proofs, such a swap can only happen $2H_k \frac{\alpha k^2}{(1-\alpha)k}$ times. The choice of L implies termination will occur while estimates are still accurate. Hence, at the point that we terminate, $\Delta^*(W) < \frac{1}{\alpha k}$, so the desired properties are satisfied by Theorem 29.

What remains is to show that after $\left\lceil \frac{m-k}{t-k} \right\rceil \cdot \ell$ queries with a committee W, either a swap is made or we terminate. By our query selection strategy, after this many queries, $W \cup \{x\}$ will be contained in at least ℓ queries for all $x \notin W$. This implies that $|V_k(W,x)| \geqslant \ell$ and $|V_k(W,x,y)| \geqslant \ell$ for all such x and y. We will later show that when this happens, $err_k(W,x)$ and $err_k(W,x,y)$ are upper bounded by $\varepsilon := \frac{1}{12} \frac{(1-\alpha)k+1}{\alpha k^2}$ for all x and y. Once this upper bound of ε has been shown, the proof is very similar of a swap or termination is similar to Theorem 30. If $\widetilde{\Delta}^+(W,c') \geqslant \frac{1}{\alpha k}$, we will certainly terminate. Otherwise, if $\widetilde{\Delta}^+(W,c') < \frac{1}{\alpha k}$, this means $\Delta(W,c') < \frac{1}{\alpha k} - 2\varepsilon$. Hence, there is a candidate x such that

$$\Delta(W,c',x) \geqslant \frac{(1-\alpha)k+1}{\alpha k^2} - \frac{k+1}{k} \cdot 2\varepsilon \geqslant 12\varepsilon - 4\varepsilon = 8\varepsilon.$$

For such an x,

$$\widetilde{\Delta}^-(W,c',x) \geqslant \Delta(W,c',x) - 2\varepsilon \geqslant 6\varepsilon = \frac{1}{2}\frac{(1-\alpha)k+1}{\alpha k^2}.$$

Hence, the swap if condition must pass and a swap will be made.

Finally, let us show the necessary bound on err_k. More formally, we must show

$$\sqrt{\frac{2\log\left(\frac{4L(k+1)m^{k+1}}{\delta}\right)}{\ell}} \leqslant \varepsilon.$$

Observing that $\ell = \frac{2}{\varepsilon^2} \cdot 2\log\left(\frac{4608k^8m^{k+2}}{\delta}\right)$, it is sufficient to show that

$$\log\left(\frac{2L(k+1)m^{k+1}}{\delta}\right) \leqslant 2\log\left(\frac{4608k^8m^{k+2}}{\delta}\right)$$

To that end, we have

$$\begin{aligned}
\log\left(\frac{4L(k+1)m^{k+1}}{\delta}\right) &\leqslant \log\left(\frac{8Lkm^{k+1}}{\delta}\right) \\
&= \log\left(\frac{8\left(2H_k\left\lceil\frac{m-k}{t-k}\right\rceil\left(\frac{\alpha k^2}{(1-\alpha)k+1}\right)\cdot\ell\right)km^{k+1}}{\delta}\right) \\
&\leqslant \log\left(\frac{16\left(k\cdot m\cdot k^2\cdot\ell\right)km^{k+1}}{\delta}\right) \\
&= \log\left(\frac{16k^4m^{k+2}\cdot\ell}{\delta}\right) \\
&\leqslant \log\left(\frac{16k^4m^{k+2}\cdot\frac{2}{\varepsilon^2}\cdot\left(2\log\left(\frac{4608k^8m^{k+2}}{\delta}\right)\right)}{\delta}\right).
\end{aligned}$$

Then using that $\frac{1}{\varepsilon}\leqslant 12k^2$,

$$\begin{aligned}
&\leqslant \log\left(\frac{4608k^8m^{k+2}\cdot\left(2\log\left(\frac{4608k^8m^{k+2}}{\delta}\right)\right)}{\delta}\right) \\
&= \log\left(\frac{4608k^8m^{k+2}}{\delta}\right)+\log\left(2\log\left(\frac{4608k^8m^{k+2}}{\delta}\right)\right) \\
&\leqslant \log\left(\frac{4608k^8m^{k+2}}{\delta}\right)+\log\left(\frac{4608k^8m^{k+2}}{\delta}\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad(\log(2a)\leqslant a\text{ for all }a\in\mathbb{R}) \\
&= 2\log\left(\frac{4608k^8m^{k+2}}{\delta}\right),
\end{aligned}$$

as needed. □

Comparing Theorem 31 with Theorem 30, we see that our upper bound on the query complexity Algorithm 6 is k times worse asymptotically. However, even in the worst case, it is unclear whether these bounds are tight; the difference may instead be due to slack in our analysis.

Beyond the worst case, there are problem instances where Algorithm 6 requires fewer queries than Algorithm 5. Consider a setting with k "good" candidates supported by all voters and $m' > k$ "bad" candidates that no one supports. Note that with $\alpha = 1$, to satisfy EJR, all good candidates must be selected. In this instance, Algorithm 5 will perform $\Theta(\frac{k^4 m}{t-k}\log(m))$ queries per swap. Further, since $m' > k$, with probability at least $\frac{1}{2}$, even a randomly-selected initial committee contains no more than $\frac{k}{2}$ good candidates, so $\Omega(k)$ swaps are required. Hence, Algorithm 5 requires $\Omega(\frac{k^5 m}{t-k}\log(m))$ queries.

In contrast, Algorithm 6 does not discard votes after swaps. In particular, consider the estimate $\widehat{\Delta}_0^+(W,c)$ for a bad candidate c that uses all voters that voted on c regardless of if they voted on anyone in W. Note that it is always 0 as no voter ever approves of c. Hence, $\widetilde{\Delta}^+(W,c) \leqslant err_0(W,c)$. On the other hand, for all good candidates c, $\Delta(W,c) \geqslant 1/k$, so $\widetilde{\Delta}^+(W,c) \geqslant 1/k$ as well. Hence, once a bad candidate has been queried $\Omega(k^2 \log m)$ times, it will have a worse $\widetilde{\Delta}^+$ when compared to any good candidate. In addition, only $\Omega(k^2 \log m)$ queries are needed for a good candidate's error term to be small enough to ensure a swap (fewer for the earlier swaps). Hence, at most $O(mk^2 \log m)$ queries are needed for Algorithm 6 to terminate.

In summary, despite being slightly worse in terms of worst-case analysis, there is evidence that Algorithm 6 may work better in practice, an intuition confirmed through experimental comparison in Section 4.5.

Part II

Algorithms for Sortition

Chapter 5

Fair Sortition Made Transparent

5.1 Introduction

In a *citizens' assembly*, a panel of randomly chosen citizens is convened to deliberate and ultimately make recommendations on a policy issue. The defining aspect of citizens' assemblies is the randomness of the process, *sortition*, by which participants are chosen. In practice, the sortition process works as follows: first, volunteers are solicited via thousands of letters or phone calls, which target individuals chosen uniformly at random. Those who respond affirmatively form the *pool* of volunteers, from which a final panel will be chosen. Finally, a *selection algorithm* is used to randomly select some pre-specified number k of pool members for the panel. To ensure adequate representation of demographic groups, the chosen panel is often constrained to satisfy some upper and lower quotas on feature categories such as age, gender, and ethnicity. We call a quota-satisfying panel of size k a *feasible panel.* As this process illustrates, citizens' assemblies offer a way to involve the public in informed decision-making. This potential for civic participation has recently spurred a global resurgence in the popularity of citizens assemblies; they have been commissioned by governments and led to policy changes at the national level [Iri19; OEC20; Fla+21].

Prompted by the growing impact of citizens' assemblies, there has been a recent flurry of computer scientific research on sortition, and in particular, on the fairness of the procedure by which participants are chosen [BGP19; Fla+20a; Fla+21]. The most practicable result to date is a family of selection algorithms proposed by Flanigan et al. [Fla+21], which are distinguished from their predecessors by their use of randomness toward the goal of fairness: while previously-used algorithms selected pool members in a random but ad-hoc fashion, these new algorithms are *maximally fair*, ensuring that pool members have as equal probability as possible of being chosen for the panel, subject to the quotas.[1] To encompass the many interpretations of "as equal as possible," these algorithms permit the optimization of any fairness objective with certain convexity properties. There is now a publicly available implementation of the techniques of Flanigan et al. [Fla+21], called *Panelot*, which optimizes the egalitarian notion that no pool member has too little selection probability via the *Leximin* objective from fair division [Mou03; Fla+20b]. This algorithm has already been deployed by several groups of panel organizers, and has been used to select dozens of panels worldwide.

Fairness gains in the panel selection process can lend legitimacy to citizens' assemblies and potentially increase their adoption, but only insofar as the public trusts that these gains are truly realized. Currently, the potential for public trust in the panel selection process is limited by multiple factors. First, the latest panel selection algorithms select the final panel via behind-the-scenes computation. When panels are selected in this manner, observers cannot even verify that any given pool member has *any* chance of being chosen for the panel. A second and more fundamental hurdle is that randomness and probability, which are central to the sortition process, have been shown in many contexts to be difficult for people to understand and reason about [RB08; MP16; WML20]. Aiming to address these shortcomings, we propose and pursue the following notion of transparency in panel selection:

[1] Quotas can preclude giving individuals exactly equal probabilities: if the panel must be 1/2 men, 1/2 women but the pool is split 3/4 men, 1/4 women, then some women must be chosen more often than some men.

> *Transparency:* Observers should be able to, without reasoning in-depth about probability, (1) understand the probabilities with which each individual will be chosen for the panel *in theory*, and (2) verify that individuals are actually selected with these probabilities *in practice*.

In this work, we aim to achieve transparency and fairness simultaneously: this means advancing the defined goal of transparency, while preserving the fairness gains obtained by maximally fair selection algorithms. Although this task is reminiscent of existing AI research on trade-offs between fairness or transparency with other desirable objectives [BFT12; FKL16; Ber+18; Tra+20], to our knowledge, this is the first investigation of the trade-off between fairness and transparency.

Setting aside for a moment the goal of fairness, we consider a method of random decision-making that is already common in the public sphere: the uniform lottery. To satisfy quotas, a uniform lottery for sortition must randomize not over individuals, but over entire feasible panels. In fact, this approach has been suggested by practitioners, and was even used in 2020 to select a citizens' assembly in Michigan. The following example, which closely mirrors that real-world pilot,[2] illustrates that panel selection via uniform lottery is naturally consistent with the transparency notion we pursue.

Suppose we construct 1000 feasible panels from a pool (possibly with duplicates), numbered 000-999, and publish an (anonymized) list of which pool members are on each panel. We then inform spectators that we will choose each panel with equal probability. This satisfies criterion (1): spectators can easily understand that all panels will be chosen with the same probability of 1/1000, and can easily determine each individual's selection probability by counting the number of panels containing the individual. To satisfy criterion (2), we enact the

[2]Of By For's pilot of live panel selection via lottery can be viewed at `https://vimeo.com/458304880#t=17m59s` from 17:59 to 21:23. For a more detailed description, see Figure 3 and surrounding text of Flanigan et al. [Fla+21].

lottery by drawing each of the three digits of the final panel number individually from lottery machines. Lottery spectators can confirm that each ball is drawn with equal probability; this provides confirmation that panels are indeed being chosen with uniform probabilities, thus confirming the enactment of the proposed individual selection probabilities. In addition to its conventionality as a source of randomness, decision-making via drawing lottery balls invites an exciting spectacle, which can promote engagement with citizens' assemblies.

This simple method neatly satisfies our transparency criteria, but it has one obvious downside: a uniform lottery over an arbitrary set of feasible panels does not guarantee any measure of equal probabilities to individuals. In fact, it is not even clear that the *fairest possible* uniform lottery over m panels, where m is a number conducive to selection by physical lottery (e.g. $m = 1000$), would not be significantly less fair than maximally fair algorithms, which sample the fairest possible unconstrained distribution over panels. For example, if m is too small, there may be *no* uniform lottery which gives all individuals non-zero selection probability, even if each individual appears on some feasible panel (and so can attain a non-zero selection probability under an unconstrained distribution).

Fortunately, empirical evidence suggests that there is hope: in the 2020 pilot mentioned above, a uniform lottery over $m = 1000$ panels was found that nearly matched the fairness of the maximally fair distribution generated by Panelot. Motivated by this anecdotal evidence, we aim to understand whether such a fair uniform lottery is guaranteed to exist in general, and if it does, how to find it. We summarize this goal in the following research questions:

> *Does there exist a uniform lottery over m panels that nearly preserves the fairness of the maximally fair unconstrained distribution over panels?* And,
>
> *Algorithmically, how do we compute such a uniform lottery?*

Results and Contributions

After describing the model in Section 5.2, in Section 5.3 we prove that it is possible to round an (essentially) arbitrary distribution over panels to a uniform lottery while preserving *all* individuals' selection probabilities up to only a small bounded deviation. These results use tools from discrepancy theory and randomized rounding. Intuitively, this bounded change in selection probabilities implies bounded losses in fairness; we formalize this intuition in Section 5.4, showing that there exists in general a uniform lottery that is nearly maximally fair, with respect to multiple choices of fairness objective. Although we would ideally like to give such bounds for the *Leximin* fairness objective, due to its use practice, we cannot succinctly represent bounds for this objective because it is not scalar valued.

We therefore give bounds for *Maximin*, a closely related egalitarian objective which only considers the minimum selection probability given to any pool member [Che+07]. We discuss in Section 5.4 why bounds on loss in Maximin fairness are, in the most meaningful sense, also bounds on loss in Leximin fairness. We additionally give upper bounds on the loss in *Nash Welfare* [Mou03], a similarly well-established fairness objective that has also been implemented in panel selection tools [HG20].

Finally, in Section 5.5, we consider the algorithmic question in practice: given a maximally fair distribution over panels, can we actually *find* nearly maximally fair uniform lotteries that match our theoretical guarantees? To answer this question, we implement two standard rounding algorithms, along with near-optimal (but more computationally intensive) integer programming methods, for finding uniform lotteries. We then evaluate the performance of these algorithms in 11 real-world panel selection instances. We find that in all instances, we can compute uniform lotteries that nearly exactly preserve not only fairness with respect to both objectives, but *entire sets* of Leximin-optimal marginals, meaning that from the perspective of individuals, there is essentially no difference between using a uniform lottery versus the optimal unconstrained distribution sampled by the latest algorithms. We discuss these results, their implications, and how they can be deployed directly into the existing panel selection pipeline in Section 5.6.

5.2 Model

Panel Selection Problem

First, we formally define the task of panel selection for citizens' assemblies. Let $N = [n]$ be the *pool* of volunteers for the panel—individuals from the population who have indicated their willingness to participate in response to an invitation. Let $F = \{f_t\}_t$ denote a fixed set of *features* of interest. Each feature $f_t : N \to \Omega_t$ maps each pool member to their value of that feature, where Ω_t is the set of f_t's possible values. For example, for feature $f_t =$ "gender", we might have $\Omega_t = \{$"male","female", "non-binary"$\}$. We define individual i's *feature vector* $F(i) = (f_t(i))_t \in \prod_t \Omega_t$ to be the vector encoding their values for all features in F.

As is done in practice and in previous research [Fla+20a; Fla+21], we impose that the chosen panel P must be a subset of the pool of size k, and must be representative of the broader population with respect to the features in F. This representativeness is imposed via *quotas:* for each feature f and corresponding value $v \in \Omega$, we may have lower and upper quotas $l_{f,v}$ and $u_{f,v}$. These quotas require that the panel contain between $l_{f,v}$ and $u_{f,v}$ individuals i such that $f(i) = v$.

In terms of these parameters, we define an instance of the panel selection problem as: given (N, k, F, l, u)—a pool, panel size, set of features, and sets of lower and upper quotas—randomly select a *feasible panel*, where a feasible panel is any set of individuals P from the collection $\mathcal{K}$:

$$\mathcal{K} := \left\{ P \in \tbinom{N}{k} :\ l_{f,v} \leqslant |\{i \in P : f(i) = v\}| \leqslant u_{f,v} \text{ for all } f, v \right\}.$$

Maximally Fair Selection Algorithms

A *selection algorithm* is a procedure that solves instances of the panel selection problem. A selection algorithm's level of fairness on a given instance is determined by its *panel distribution* p, the (possibly implicit) distribution over $\mathcal{K}$ from which it draws the final panel. Because we care about fairness to individual pool members, we evaluate the fairness of p in terms of the fairness of selection probabilities, or *marginals*, that p implies for all pool members.[3] We denote the vector of marginals implied by p as π, and we will sometimes specify a panel distribution as p, π to explicitly denote this pair. We say that π is *realizable* if it is implied by some distribution p over the feasible panels $\mathcal{K}$.

Maximally fair selection algorithms are those which solve the panel selection problem by sampling a specifically chosen p: one which implies marginals π that allocate probability as fairly as possible across pool members. The fairness of p, π is measured by a *fairness objective* $\mathcal{F}$, which maps an allocation—in this case, of selection probability to pool members—to a real number measuring the allocation's fairness. Fixing an instance, a fairness objective $\mathcal{F}$, and a panel distribution p, we express the fairness of p as $\mathcal{F}(p)$. Existing maximally fair selection algorithms can maximize a wide range of fairness objectives, including those considered in this chapter.

Leximin, Maximin, and Nash Welfare

Of the three fairness objectives we consider in this chapter, Maximin and Nash Welfare (NW) have succinct formulae. For p, π they are defined as follows, where π_i is the marginal of

[3] A panel distribution p implies a unique vector of marginals π as follows: fixing p, π, a pool member i's marginal selection probability π_i is equal to the probability of drawing a panel from p containing that pool member. For a more detailed introduction to the connection between panel distributions and marginals, we refer readers to Flanigan et al. [Fla+21].

individual i:

$$\text{Maximin}(p) := \min_{i \in N} \pi_i, \qquad \text{NW}(p) := \left(\prod_i \pi_i\right)^{1/n}.$$

Intuitively, NW maximizes the geometric mean, prioritizing the marginal π_i of each individual i in proportion to π_i^{-1}. Maximin maximizes the marginal probability of the individual least likely to be selected. Finally, Leximin is a refinement of Maximin, and is defined by the following algorithm: first, optimize Maximin; then, fixing the minimum marginal as a lower bound on any marginal, maximize the second-lowest marginal; and so on.

Our task: quantize a fair panel distribution with minimal fairness loss.

We define a $1/m$-*quantized* panel distribution as a distribution over all feasible panels $\mathcal{K}$ in which all probabilities are integer multiples of $1/m$. We use $\bar{p}$ to denote a panel distribution with this property. Formally, while an (unconstrained) panel distribution p lies in $\mathcal{D} := \{p \in \mathbb{R}_+^{|\mathcal{K}|} : \|p\|_1 = 1\}$, a $1/m$-quantized panel distribution in $\bar{p}$ lies in $\overline{\mathcal{D}} := \{\bar{p} \in (\mathbb{Z}_+/m)^{|\mathcal{K}|} : \|\bar{p}\|_1 = 1\}$. Note that a $1/m$-quantized distribution $\bar{p}$ immediately translates to a physical uniform lottery of over m panels (with duplicates): if $\bar{p}$ assigns probability ℓ/m to panel P, then the corresponding physical uniform lottery would contain ℓ duplicates of P. Thus, if we can compute a $1/m$-quantized panel distribution $\bar{p}$ with fairness $\mathcal{F}(\bar{p})$, then we have designed a physical uniform lottery over m panels with that same level of fairness.

Our goal follows directly from this observation: we want to show that given an instance and desired lottery size m, we can compute a $1/m$-quantized distribution $\bar{p}$ that is nearly as fair, with respect to a fairness notion $\mathcal{F}$, as the maximally fair panel distribution in this instance $p^* \in \arg\max_{p \in \mathcal{D}} \mathcal{F}(p)$. We define the *fairness loss* in this quantization process to be the difference $\mathcal{F}(p^*) - \mathcal{F}(\bar{p})$. We are aided in this task by the existence of practical algorithms for computing p^* Flanigan et al. [Fla+21], which allows us to use p^* as an input to the quantization procedure we hope to design. For intuition, we illustrate this quantization task in

Figure 5.1, where π^*, $\bar{\pi}$ are the marginals implied by $p^*, \bar{p}$, respectively. Since the fairness of $p^*, \bar{p}$ are computed in terms of π^*, $\bar{\pi}$, it is intuitive that a quantization process that results in small *marginal discrepancy*, defined as the maximum change in any marginal $\|\pi - \bar{\pi}\|_\infty$, should also have small fairness loss. This idea motivates the upcoming section, in which we give quantization procedures with provably bounded marginal discrepancy, forming the foundation for our later bounds on fairness loss.

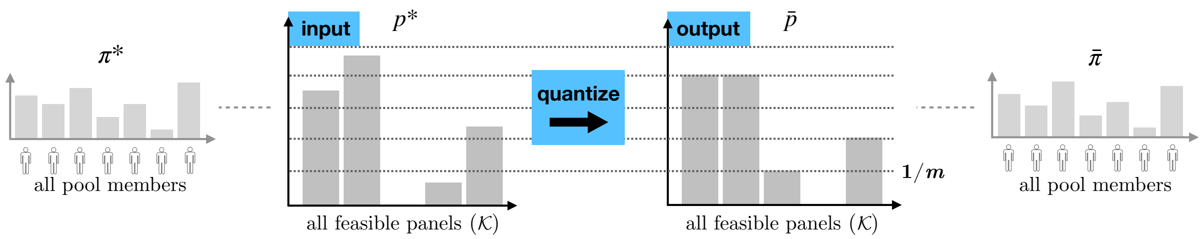


Figure 5.1: The quantization task takes as input a maximally fair panel distribution p^* (implying marginals π^*), and outputs a $1/m$-quantized panel distribution $\bar{p}$ (implying marginals $\bar{\pi}$).

5.3 Theoretical Bounds on Marginal Discrepancy

Here we prove that for a fixed panel distribution p, π, there exists a uniform lottery $\bar{p}, \bar{\pi}$ such that $\|\pi - \bar{\pi}\|_\infty$ is bounded. Preliminarily, we note that it is intuitive that bounds on this discrepancy should approach 0 as m becomes large with respect to n and k. To see why, begin by fixing some distribution p, π over panels: as m becomes large, we approach the scenario in which a uniform lottery $\bar{p}$ can assign panels arbitrary probabilities, providing increasingly close approximations to p. Since the marginals π_i are continuous with respect to p, as $\bar{p} \to p$ we have that $\bar{\pi}_i \to \pi_i$ for all i.

While this argument demonstrates convergence, it provides neither efficient algorithms nor tight bounds on the rate of convergence. In this section, our task is therefore to bound the rate of this convergence as a function of m and the other parameters of the instance. All omitted proofs of results from this section are included in Section 5.8.

5.3.1 Worst-Case Upper Bounds

Our first set of upper bounds result from rounding STANDARD LP, the LP that most directly arises from our problem. This LP is defined in terms of a panel distribution p, π, and M, an $n \times |\mathcal{K}|$ matrix describing which individuals are on which feasible panels: $M_{i,P} = 1$ if $i \in P$ and $M_{i,P} = 0$ otherwise.

STANDARD LP
$$Mp = \pi \tag{5.1}$$
$$\|p\|_1 = 1 \tag{5.2}$$
$$p \geqslant 0.$$

Here, (5.1) specifies n total constraints. Our goal is to round p to a uniform lottery $\bar{p}$ over m panels (so the entries $\bar{p}$ are multiples of $1/m$) such that (5.2) is maintained exactly, and no constraint in (5.1) is relaxed by too much, i.e., $\|Mp - M\bar{p}\|_\infty = \|\pi - \bar{\pi}\|_\infty$ remains small.

Randomized rounding is a natural first approach. Any randomized rounding scheme satisfying negative association (which includes several that respect (5.2)) yields the following bound:

Theorem 32. *For any realizable π, we may efficiently randomly generate $\bar{p}$ such that its marginals $\bar{\pi}$ satisfy*
$$\|\pi - \bar{\pi}\|_\infty = O\left(\frac{\sqrt{n \log n}}{m}\right).$$

To prove this and other results, we will make repeated use of the following generalization of Hoeffding's inequality (see e.g. Proposition 5 of [DR96]):

Lemma 13. *If $\{\xi_j\}$ are negatively associated random variables with $\xi_j \in [a_j, b_j]$ and $\xi = \sum_j \xi_j$, then*
$$\Pr\left[|\mathbb{E}[\xi] - \xi| \geqslant t\right] \leqslant 2\exp\left\{-\frac{2t^2}{\sum_j (b_j - a_j)^2}\right\}.$$

Proof of Theorem 32. Given a vector of marginals π, let p be a basic solution to $Mp = \pi$, where M is the individual-feasible panel membership matrix, so that $|\operatorname{supp}(p)| \leqslant n$.

Then, we will construct $\bar{p}$ from p by constructing x', rounding it to $\bar{x}' \in \{0,1\}^{|\mathcal{K}|}$, and then reconstructing $\bar{p}$ as described in Definition 11. To do this 0/1 rounding, here we use any randomized rounding procedure that satisfies the following properties: preservation of adding up constraint $\|\bar{x}'\|_1 = \|x'\|_1$, preservation of marginals $E[\bar{x}'_j] = x'_j$, and that $\bar{x}'_j$ are *negatively associated*, as defined in [BJ12; DR96]. These properties are satisfied via any number of randomized rounding algorithms [BJ12]. Note as in Definition 11, $\|\bar{x}'\|_1 = \|x'\|_1$ implies that $\bar{p} \in \bar{\mathcal{D}}$.

Now it remains to analyze the marginal $\bar{\pi}_i$ provided to any given individual i by $\bar{p}$. Consider the collection of $\bar{x}'_j$ for which i is contained in panel j. Then, using the negative association of these $\bar{x}'_j$s, we have that for any $t \geqslant 0$,

$$\Pr[|Mx' - M\bar{x}'| \geqslant t] = \Pr\left[\left|\mathbb{E}\left[\sum_{j \ni i} \bar{x}'_j\right] - \sum_{j \ni i} \bar{x}'_j\right| \geqslant t\right], \tag{5.3}$$

by the definition of $\bar{x}'_j$. Then by Hoeffding (Lemma 13),

$$\leqslant 2\exp\left(\frac{-2t^2}{|\{j : i \in j\}|}\right) \tag{5.4}$$

$$\leqslant 2\exp\left(\frac{-2t^2}{n}\right), \tag{5.5}$$

where here we use that $|\operatorname{supp}(p)| \leqslant n$. Then taking $t = \sqrt{\frac{1+\epsilon}{2} n \log n}$,

$$\leqslant \frac{2}{n^{1+\epsilon}}. \tag{5.6}$$

Taking a union bound over all n rows i then gives

$$\Pr\left[\|Mx' - M\bar{x}'\|_\infty \geqslant \sqrt{\frac{(1+\epsilon)}{2}} \cdot \sqrt{n \log n}\right] \leqslant \frac{2}{n^\epsilon} < 1.$$

By Lemma 17, we therefore have

$$\Pr\left[\|\pi - \bar{\pi}\|_\infty \leqslant \sqrt{\frac{1+\epsilon}{2}} \cdot \frac{\sqrt{n \log n}}{m}\right] \geqslant 1 - \frac{2}{n^\epsilon} > 0. \qquad \square$$

Note: if we are additionally guaranteed that all of the $\pi_i = \Omega(k/n)$, then a multiplicative form of Chernoff yields

$$\|\pi - \bar{\pi}\|_\infty = O\left(\sqrt{\frac{k \log n}{mn}}\right)$$

with constant probability.

Fortunately, there is potential for improvement: randomized rounding does not make full use of the fact that M is k-column sparse, due to each panel in $\mathcal{K}$ containing exactly k individuals. We use this sparsity to get a stronger bound when $n \gg k^2$, which is a practically significant parameter regime. The proof applies a dependent rounding algorithm based on a theorem of Beck and Fiala [BF81], to which a modification ensures the exact satisfaction of constraint (5.2).

Theorem 33. *For any realizable π, we may efficiently construct $\bar{p}$ such that its marginals $\bar{\pi}$ satisfy*

$$\|\pi - \bar{\pi}\|_\infty \leqslant k/m.$$

Proof. Here, we apply the rounding algorithm used by Flanigan et al. [Fla+20a] (Lemma 9, Appendix B.4.1), which builds on a notable theorem by Beck and Fiala [BF81]. Since this rounding algorithm does 0/1 rounding, we apply their algorithm to round x', as in Definition 11, to some 0/1 vector $\bar{x}'$, from which we construct $\bar{p}$. By Lemma 9 in Appendix B.4.1 in [Fla+20a], this algorithm ensures the preservation of the "adding up" constraint, that is, that $\|\bar{x}'\|_1 = \|x'\|_1$. Thus, by results (a) and (b) of Lemma 17, $\bar{p} \in \bar{D}$.

Now, it remains to show that $\|\pi - \bar{\pi}\|_\infty = \|M(p - \bar{p})\|_\infty \leqslant k/m$. Fortunately, as they prove, the rounding procedure of Flanigan et al. [Fla+20a] guarantees that when rounding x' to $\bar{x}'$, for a constraint matrix M with column sparsity k, $\|M(x' - \bar{x}')\|_\infty \leqslant k$. By Lemma 17 result (c), this immediately implies that $\|\pi - \bar{\pi}\|_\infty \leqslant k/m$. □

This bound is already meaningful in practice, where $k \ll m$ is insured by the fact that m is pre-chosen along with k prior to panel selection. Note also that k is typically on the order of

100 (Section 5.7), whereas a uniform lottery can in practice be easily made orders of magnitude larger, as each additional factor of 10 in the size of the uniform lottery requires drawing only one more ball (and there is no fairness cost to drawing a larger lottery, since increasing m allows for uniform lotteries which better approximate the unconstrained optimal distribution).

5.3.2 Beyond-Worst-Case Upper Bounds

As we will demonstrate in Section 5.3.3, we cannot hope for a better worst-case upper bound than $\text{poly}(k)/m$. We thus shift our consideration to instances which are "simple" in their feature structure, having a small number of features (Theorem 43), a limited number of unique feature vectors in the pool (Theorem 34), or multiple individuals that share each feature vector present (Theorem 44). The beyond-worst-case bounds given by Theorem 34 and Theorem 44 asymptotically dominate our worst-case bounds in Theorem 32 and Theorem 33, respectively. Moreover, Theorem 34 dominates all other upper bounds in 10 of the 11 practical instances studied in Section 5.5.

We note that while our worst-case upper bounds implied the near-preservation of *any* realizable set of marginals π, some of our beyond-worst-case results apply to only realizable π which are *anonymous*, meaning that π_i are equal for all i with equal feature vectors. We contend that any reasonable set of marginals should have this property,[4] and furthermore that the "anonymization" of any realizable π is also realizable (Claim 42); hence this restriction is insignificant. Our beyond-worst-case bounds also differ from our worst-case bounds in that they depart from the paradigm of rounding p, instead randomizing over panels that may fall outside the support of p.

The main beyond-worst-case bound we give, stated below, is parameterized by $|\mathcal{C}|$, where $\mathcal{C}$ is the set of unique feature vectors that appear in the pool. All omitted proofs and other

[4]The class of all anonymous marginals π includes the maximizers π^* of all reasonable fairness objectives, and second, this condition is satisfied by all existing selection algorithms used in practice, to our knowledge.

beyond worst-case results are stated and proven in Section 5.8.

Theorem 34. *If π is anonymous and realizable, then we may efficiently construct $\bar{p}$ such that its marginals $\bar{\pi}$ satisfy*

$$\|\pi - \bar{\pi}\|_\infty = O\left(\frac{\sqrt{|\mathcal{C}| \log |\mathcal{C}|}}{m}\right).$$

$|\mathcal{C}|$ is at most n, so this bound dominates Theorem 32. In 10 of the 11 real-world instances we study, $|\mathcal{C}|$ is also smaller than k^2 (Section 5.7), in which case this bound also dominates Theorem 33.

At a high level, our beyond-worst-case upper bounds are obtained not by directly rounding p, but instead using the structure of the sortition instance to abstract the problem into one about "types." For this bound we then solve an LP in terms of "types," round that LP, and then reconstruct a rounded panel distribution $\bar{p}, \bar{\pi}$ from the "type" solution. In particular, the *types* of individuals are the feature vectors which appear in the pool, and *types* of panels are the multisets of k feature vectors that satisfy the instance quotas. Fixing an instance, we project some p into type space by viewing it as a distribution $\mathfrak{p}$ over types of panels $\mathfrak{K}$, inducing marginals τ_c for each type individuals $c \in \mathcal{C}$.

To begin, we define the Type LP, which is analogous to eq. (5.1). We let Q be the type analog of M, so that entry Q_{cj} is the number of individuals i with $F(i) = c$ contained in panels of type $j \in \mathfrak{K}$.[5] Then,

Type LP
$$Q\,\mathfrak{p} = \tau \tag{5.7}$$
$$\|\mathfrak{p}\|_1 = 1 \tag{5.8}$$
$$\mathfrak{p} \geqslant 0.$$

We round $\mathfrak{p}$ in this LP to a panel type distribution $\bar{\mathfrak{p}}$ while preserving (5.8). All that remains, then, is to construct some $\bar{p}, \bar{\pi}$ such that p is consistent with $\bar{\mathfrak{p}}$ and $\|\pi - \bar{\pi}\|_\infty$ is small. This $\bar{p}$

[5] Completing the analogy, $\mathcal{C}, \mathfrak{K}, Q, \mathfrak{p}, \bar{\mathfrak{p}}, \tau$ are the "type" versions of $N, \mathcal{K}, M, p, \bar{p}, \pi$ from the original LP.

is in general supported by panels outside of $\mathrm{supp}(p)$, unlike the $\bar{p}$ obtained by Theorem 32. It is the anonymity of π which allows us to construct these new panels and prove that they are feasible for the instance.

Proof of Theorem 34. We begin with anonymous marginals π witnessed by some distribution p over $\mathcal{K}$. The first order of business is to project p into "type space," in order to derive a distribution over panel types. Overloading F, we let $F(P) = \mathfrak{P}$ denote the panel type of a given panel P, defined as the multiset $F(P) = \{F(i) : i \in P\}$. Then we define the distribution over panel types induced by p as $\mathfrak{p}$, where the probability of drawing panel type $\mathfrak{P}$ from $\mathfrak{p}$ is defined as $\mathfrak{p}_{\mathfrak{P}} := \sum_{P \in \mathcal{K}: F(P) = \mathfrak{P}} p_P$.

This $\mathfrak{p}$ satisfies the Panel Type LP in eq. (5.7). As an aside, note that this $\mathfrak{p}$ has support $\mathrm{supp}(\mathfrak{p}) = \{F(P) : P \in \mathrm{supp}(p)\}$. We will assume without loss of generality that $\mathfrak{p}$ is a basic solution to (5.7), so that it has at most $|\mathcal{C}|$ nonzero entries, where $\mathcal{C}$ is the set of all feature-vectors appearing in the pool, i.e., $\mathrm{supp}(p) \leqslant |\mathcal{C}|$. Since $|\,\mathrm{supp}(p)| \leqslant n$ without loss of generality, $|\,\mathrm{supp}(\mathfrak{p})| \leqslant n$ also, and so this basic $\mathfrak{p}$ may be found efficiently.

Given this distribution $\mathfrak{p}$ over panel types, we will round it to a uniform lottery $\bar{\mathfrak{p}}$ of size m over panel types $\mathfrak{K}$. Finally, we will lift this distribution over panel types $\bar{\mathfrak{p}}$ back to a distribution $\bar{p}$ over panels with the desired guarantee, and argue that this lift can be performed when the original marginals π are anonymous.

We generate $\bar{\mathfrak{p}}$, a distribution with all probabilities multiples of $1/m$, from $\mathfrak{p}$ via randomized rounding, as in Theorem 32. To produce $\bar{\mathfrak{p}}$ via a 0/1 rounding algorithm, we follow the procedure given in Definition 11, where here, $\mathfrak{p}, \bar{\mathfrak{p}}$ correspond to the $p, \bar{p}$ given in the definition. Via this definition, we construct $x, \lfloor x \rfloor, x', \bar{x}'$ analogously, so that $x = m\mathfrak{p}$, etc. By choosing a randomized rounding procedure that preserves $\|\bar{x}'\|_1 = \|x'\|_1$, by Lemma 17 we have that $\bar{\mathfrak{p}}$ is a valid distribution containing multiples of $1/m$. We again assume this rounding procedure samples $\bar{x}'_j$ which are negatively associated, and preserves that $\mathbb{E}[\bar{x}'_j] = x'_j$ for all panel types j.

Recall that type marginals $\tau_c, \bar{\tau}_c$ represent the expected number of panel spots allocated to each feature vector c by $\mathfrak{p}, \bar{\mathfrak{p}}$, respectively, and are given by $\tau = Q\mathfrak{p}$ and $\bar{\tau} = Q\bar{\mathfrak{p}}$. (Recall that Q, as described in Section 5.3, encodes the number of copies of each feature vector on each panel type.) We will next analyze the proximity of the rounded type marginals $\bar{\tau}_c$ to the original type marginals τ_c.

Proceeding via an analysis similar to that of Theorem 32, we consider the collection of random variables $\bar{x}'_j$ for which feature vector c appears on panel type j (i.e., $Q_{cj} > 0$). We note that these $\bar{x}'_j$ are again negatively associated, and thus all $Q_{cj}\bar{x}'_j$ are negatively associated, since for a fixed instance all Q_{cj} are constant.

Then for any $t \geqslant 0$,

$$\Pr[|(Qx' - Q\bar{x}')_c| \geqslant t] = \Pr\left[\left|\mathbb{E}\left[\sum_j Q_{cj}\bar{x}'_j\right] - \sum_j Q_{cj}\bar{x}'_j\right| \geqslant t\right], \tag{5.9}$$

by the definition of x_j and $\tilde{x}_j$. Then by Hoeffding (Lemma 13) with $\xi_j = Q_{cj}\tilde{x}_j$,

$$\leqslant 2\exp\left(\frac{-2t^2}{\sum_j Q_{cj}^2}\right) \tag{5.10}$$

$$\leqslant 2\exp\left(\frac{-2t^2}{|\mathcal{C}|m_c^2}\right), \tag{5.11}$$

where $m_c := \max_j Q_{cj}$, and (5.11) uses that for all c, $\sum_j Q_{cj}^2 \leqslant \sum_j m_c^2 \leqslant |\operatorname{supp}(\mathfrak{p})|m_c^2 \leqslant |\mathcal{C}|m_c^2$. Thus, taking $t_c = \alpha \cdot m_c \cdot \sqrt{|\mathcal{C}|\log|\mathcal{C}|}$,

$$\leqslant \frac{2}{|\mathcal{C}|^{2\alpha^2}}. \tag{5.12}$$

Taking $\alpha > \sqrt{\frac{1}{2}(1 + \frac{\log 2}{\log|\mathcal{C}|})}$ and union bounding over all $|C|$ feature vectors, we may therefore guarantee that with positive probability,

$$|(Qx' - Q\bar{x}')_c| \leqslant \alpha \cdot m_c\sqrt{|\mathcal{C}|\log|\mathcal{C}|}$$

for all c simultaneously. By Lemma 17, the derived $\bar{\mathfrak{p}}$ and $\bar{\tau}$ and therefore satisfy

$$|\tau_c - \bar{\tau}_c| \leqslant \alpha \cdot m_c \frac{\sqrt{|\mathcal{C}|\log|\mathcal{C}|}}{m} \tag{5.13}$$

for all c simultaneously.

Given such a $\bar{\mathfrak{p}}, \bar{\tau}$ over panel types, it remains to construct some uniform lottery $\bar{p}, \bar{\pi}$ over the panels in $\mathcal{K}$ which is consistent with $\bar{\tau}$ and satisfies the desired guarantees on $\bar{\pi}$, which are:

1. each individual appears on each panel in $\bar{p}$ at most once,[6]

2. $0 \leqslant \bar{\pi}_i \leqslant 1$ for all i, and

3. $|\pi_i - \bar{\pi}_i|$ is small for all i.

We will describe a procedure for forming $\bar{p}$ and $\text{supp}(\bar{p})$ from $\bar{\mathfrak{p}}$, and then argue that it satisfies all three of these criteria, as well as implies a valid distribution $\bar{p}$ for which all probabilities are multiples of $1/m$. At a high level, this algorithm starts with the panel types $\mathfrak{P}_j$ which form the support of $\mathfrak{p}$, and for each c in turn allocates spots in these panel types $\mathfrak{P}_j$ with feature vector c to individuals in $N_c := \{i \in [n] : F(i) = c\}$, the n_c individuals with feature vector c. Given the type marginals $\bar{\tau} = Q\bar{\mathfrak{p}}$ output by our rounding procedure, it first calculates the "ideal" number of spots $\bar{s}_i$ to allocate to each individual $i \in N_c$ across all of $\bar{p}$. It then performs the allocation in such a way that the guarantees above are satisfied. Since $\bar{\mathfrak{p}} \in (\mathbb{Z}_+/m)^{|\mathfrak{K}|}$ and this algorithm populates each $\mathfrak{P}_j$ in the support to create some $P_j \in \mathcal{K}$, it follows that the $\bar{p}$ which it ultimately produces is $\bar{p} \in (\mathbb{Z}_+/m)^{|\mathcal{K}|}$ also.

[6]We note that this is a concern because we will not simply be choosing known panels from collection $\mathcal{K}$, as we don't see the entire collection *a priori*; we will instead be *constructing* panels that must turn out to be feasible.

Algorithm 7 PANELPACKER

Require: $\bar{\mathfrak{p}} \in (\mathbb{Z}_+/m)^{|\mathfrak{K}|}$ a distribution over feasible panel types, N
Ensure: $\bar{p} \in (\mathbb{Z}_+/m)^{|\mathcal{K}|}$ a distribution over feasible panels
1: Initialize $P_j \leftarrow \emptyset$ for each $\mathfrak{P}_j \in supp(\bar{\mathfrak{p}})$, with multiplicity (i.e. for $j \in [m]$)
2: **for** $c \in \mathcal{C}$ **do**
3: Initialize spots $\bar{s}_i \in \{\lfloor m \cdot \bar{\tau}_c/n_c \rfloor, \lceil m \cdot \bar{\tau}_c/n_c \rceil\}$ for $i \in N_c$ such that $\sum_{i \in N_c} \bar{s}_i = m \cdot \bar{\tau}_c$
4: Initialize $d_i^1 \leftarrow \bar{s}_i$ for $i \in N_c$
5: **for** $j \in [m]$ **do**
6: Let I_{cj} be the first Q_{cj} many $i \in N_c$ with largest d_i^j
7: Update $P_j \leftarrow P_j \cup I_{jc}$
8: Update $d_i^{j+1} \leftarrow d_i^j - \mathbb{1}\{i \in I_{cj}\}$ for all $i \in N_c$
return $\bar{p}$ the uniform distribution over P_j

For each panel type $\mathfrak{P}_j$ in the support of $\bar{\mathfrak{p}}$, Algorithm 7 forms one panel in the support of $\bar{p}$ by, for each $c \in \mathcal{C}$, allocating each of panel type $\mathfrak{P}_j$'s Q_{cj} "spots" to individuals $i \in N_c$. It populates each panel type $\mathfrak{P}_j$ with individuals for each c independently. If Algorithm 7 succeeds at step (6) for all $c \in \mathcal{C}$, then it produces a panel $P_j \in \text{supp}(\bar{p})$. We first argue that Algorithm 7 succeeds in producing feasible panels.

Proof that Algorithm 7 succeeds. In particular, we will argue that Algorithm 7 succeeds for every iteration of step (6). Since $\sum_{i \in N_c} \bar{s}_i = \sum_{\mathfrak{P}_j \in \bar{\mathfrak{p}}} Q_{cj}$, this is equivalent to showing that it assigns all individuals $i \in N_c$ such that $d_i^{m+1} = 0$ for all i and no individual appears on any panel more than once.

In each round we have

$$d_i^j := m \cdot \bar{\pi}_i - \sum_{j' < j} \mathbb{1}\{i \in P_{j'}\}$$

the number of spots in $\bar{\mathfrak{p}}$ of type c on which i still needs to be placed at the beginning of round j in order to reach their allocation of $\bar{s}_i$ spots. (This d_i^j can be viewed as the "unsatisfied demand" of individual i at round j, according to the promised number of spots $m\bar{\pi}_i$.)

Because the $\bar{\pi}_i$ are all either $\frac{\lfloor m \cdot \bar{\tau}_c/n_c \rfloor}{m}$ or $\frac{\lceil m \cdot \bar{\tau}_c/n_c \rceil}{m}$, the initial values of d_i^0 for $i \in N_c$ are all within 1 of one another. Note that step (6) preserves this property that d_i^j remain within 1 of one another for all rounds, since at each step j it decreases some collection of maximal d_i^j by 1.

Suppose for the sake of contradiction that for some c, Algorithm 7 reaches some first step j for which a c position on panel P_j cannot be allocated to any $i \in N_c$; then there are not enough individuals with remaining "unmet demand", so $Q_{cj} > |\{i : d_i^j > 0\}|$. Since $Q_{cj} \leqslant m_c \leqslant n_c$, it must be the case that some $i \in N_c$ have already been fully assigned by this step j (meaning that for these i it is the case that $d_i^j = 0$), and so all $d_i^j \in \{0, 1\}$ because the d_i^j are within 1 of one another. But $\sum_j Q_{cj} = \sum_i d_i^0 = m \cdot \bar{\tau}_c$, while at this point

$$\sum_{j' \geqslant j} Q_{cj'} \geqslant Q_{cj} > |\{i : d_i^j > 0\}| = \sum_i d_i^j,$$

meaning that the number of unallocated positions of type c remaining at step j exceeds the remaining unmet demand of the $i \in N_c$. This implies that strictly more than $Q_{cj'}$ individuals i were given spots on panel j' at step (6) for some earlier $j' < j$. But this is impossible by the definition of Algorithm 7. Therefore Algorithm 7 must succeed in feasibly assigning individuals of each type c to panels.

Since Algorithm 7 succeeds on step (6), it successfully puts Q_{cj} individuals in N_c onto panel P_j for each j and each c. By the feasibility of $\mathfrak{P}_j$ we therefore have that $|P_j| = k$ and P_j is quota feasible, since $\mathfrak{P}_j$ is quota feasible and P_j has the exact same numbers of individuals with each feature vector as $\mathfrak{P}_j$.

Therefore Algorithm 7 terminates with a collection of quota-feasible panels, with no individual appearing on any panel more than once. $\square$

We conclude by arguing that the output of Algorithm 7 satisfies the desired guarantees.

First, it is clear that each individual i appears on each panel $P_j \in \text{supp}(p)$ at most once. This is because for each individual $i \in N_c$ for some c, i is assigned a position on P_j if and only if $i \in I_{cj}$ at step (6), and I_{cj} contains each i at most once by definition. Therefore condition (1) is satisfied.

We next show that these output $\bar{\pi}_i$ satisfy condition (2). For each i, its value of $\bar{\pi}_i$ in the distribution $\bar{p}$ output by Algorithm 7 is precisely $\bar{s}_i/m$.

Therefore clearly $\bar{\pi}_i \geqslant 0$, and since condition (1) holds we have $\sum_j \mathbb{1}\{i \in P_j\} \leqslant m$, and so $\bar{\pi}_i \leqslant 1$ also. For a more explicit proof that $\bar{\pi}_i \leqslant 1$, observe that since $\mathfrak{p}$ is a distribution,

$$\bar{\tau}_c = \sum_j \bar{\mathfrak{p}}_j Q_{cj} \leqslant \max_j Q_{cj} = m_c \leqslant n_c,$$

where the last inequality follows because all $\mathfrak{P}$ are feasible panel types, so they cannot contain more individuals $i \in N_c$ than exist in the pool. By Algorithm 7 we have $\bar{s}_i \in \{\lfloor m \cdot \bar{\tau}_c / n_c \rfloor, \lceil m \cdot \bar{\tau}_c / n_c \rceil\}$. Dividing by n_c and multiplying by m yields $\bar{s}_i \leqslant m$, and so $\bar{\pi}_i = \bar{s}_i / m \leqslant 1$. Thus (2) is satisfied.

Finally, we confirm condition (3), that the individual marginals are close. By the anonymity of π, for all i with $F(i) = c$ we have $\pi_i = \tau_c / n_c$, and by its choice of $\bar{s}_i$ and the fact that it succeeds, Algorithm 7 guarantees that $\bar{\pi}_i = \bar{s}_i / m \in (\bar{\tau}_c / n_c - 1/m, \bar{\tau}_c / n_c + 1/m)$. Since $m_c \leqslant n_c$, therefore (5.13) implies

$$|\pi_i - \bar{\pi}_i| \leqslant \frac{m_c}{n_c} \cdot \alpha \cdot \frac{\sqrt{|\mathcal{C}| \log |\mathcal{C}|}}{m} + \frac{1}{m} = O\left(\frac{\sqrt{|\mathcal{C}| \log |\mathcal{C}|}}{m}\right),$$

for all i, satisfying condition (3) and showing the claim. □

5.3.3 Lower Bounds

This method of using bounded discrepancy to derive nearly fairness-optimal uniform lotteries has its limits, since there are even sparse M and fractional x for which no integer $\bar{x}$ yields nearby $M\bar{x}$. In the worst case, we establish lower bounds by modifying those of Beck and Fiala [Spe85]:

Theorem 35. *There exist p, π for which for all uniform lotteries $\bar{p}, \bar{\pi}$,*

$$\min_{\bar{p} \in \overline{\mathcal{D}}} \|\pi - \bar{\pi}\|_\infty = \Omega\left(\frac{\sqrt{k}}{m}\right).$$

To prove this, we will make use of the following lemma:

Lemma 14. *Any k-uniform hypergraph on $[n]$ is realizable via quotas as the set of feasible panels for an instance of the panel selection problem with pool $[n]$.*

When individual membership in feasible panels is represented as $M \in \{0,1\}^{n\times|\mathcal{K}|}$, this lemma claims that any M with uniform column norms is *realizable* by an instance of the panel selection problem, meaning that there exists an instance of the panel selection problem (N, k, F, l, u) for which M is precisely the individual-panel membership matrix for the set of feasible panels.

Proof. Given a set system $\mathcal{S} \subseteq \binom{[n]}{k}$, we may construct a set of upper quotas such that the collection of feasible panels is exactly $\mathcal{S}$.

To do this, construct a binary feature f_T for each $T \notin \mathcal{S}$. For each i in $[n]$, let $f_T(i) = 1$ if and only if $i \in T$; otherwise let $f_T(i) = 0$. Finally, enforce the upper quota that for all feasible panels $P \subset [n]$,

$$\sum_{i\in P} f_T(i) \leqslant k-1,$$

for all $T \notin \mathcal{S}$—that is, no feasible panel has more than $k-1$ members belonging to any T. Clearly no $T \notin \mathcal{S}$ is a feasible panel. For $S \in \mathcal{S}$, observe that $|S| = k$, and so for all $T \notin \mathcal{S}$, we have $|S \cap T| \leqslant k-1$. Therefore all $S \in \mathcal{S}$ are feasible.

Finally, it bears noting that this is also possible to execute using lower quotas: taking $f'_T(i) = 1 - f_T(i)$, we could instead enforce for each $T \notin \mathcal{S}$ that

$$\sum_{i\in P} f'_T(i) \geqslant 1.$$

□

Proof of Theorem 35. Using Lemma 14, our aim is to identify and deploy some matrix $M \in \{0,1\}^{n\times|\mathcal{K}|}$ for which

$$\min_{\bar{x}\in\bar{\Delta}} \|M\bar{x}\|_\infty = \Omega\left(\sqrt{k}\right),$$

where $\bar{\Delta} := \{x \in \{\dots,-3,-1,1,3,\dots\}^n : \sum_i x_i = 0\}$ and all columns of M sum to k. Translating and scaling appropriately and applying Lemma 14, this will provide our desired $\Omega\left(\frac{\sqrt{k}}{m}\right)$ lower bound.

The common instances which provide lower bounds of $\Omega(\sqrt{k})$ for the Beck-Fiala problem are insufficient for our purposes in two respects. First, while they are column-sparse, they are generally not uniform in column norm. Second, they are incomparable in terms of the $\bar{x}$ which they quantify over: the Beck-Fiala problem considers minimizing $\|M\bar{x}\|_\infty$ in the more restrictive rounding setting where $\bar{x} \in \{-1, 1\}^n$, while we are concerned with $\bar{x} \in \bar{\Delta}$.

We overcome these barriers by first modifying the Walsh matrices — a family of Hadamard matrices — in order to guarantee uniform column norms, and then modifying the Beck-Fiala lower bound proof of [Spe85, Theorem 19] for arbitrary Hadamard matrices to apply to our matrices for all $\bar{x} \in (2\mathbb{Z}+1)^n$.

To begin, let H_t be the $2^t \times 2^t$ Walsh matrix, defined recursively by $H_0 = 1$ and

$$H_{t+1} = \begin{bmatrix} H_t & H_t \\ H_t & -H_t \end{bmatrix}.$$

Let $N := 2^t$ denote its dimension.[7] It is a fact that all rows (and columns) besides the first have an equal number of 1 and -1 entries. Therefore we take H_t' to be the submatrix derived by dropping the first two columns of H_t. (We remove the first column so that all remaining columns have equal sum; we remove the second so that $\bar{\Delta}$ is nonempty). Additionally, let h_i denote the rows of H_t', and h^j denote its columns. Then H_t' has the property that $\sum_i h_i^j = 0$, and in particular all columns h^j have $N/2$ 1-entries.

We have the following lemma:

Lemma 15.

$$\min_{x \in \bar{\Delta}} \|H_t' \, x\|_\infty \geqslant \frac{N-2}{\sqrt{N}},$$

where $\bar{\Delta} := \{x \in \{\ldots, -3, -1, 1, 3, \ldots\}^{N-2}\}$.

Proof. This right-hand side is $H_t' \, x = (h_1 x, \ldots, h_N x)^T$. We aim to show that there is some i

[7] Note that this N is a variable used only in this proof, and it is unrelated to the pool N and its magnitude n as elsewhere in this chapter.

for which $|h_i x|$ is large. Writing $\|H'_t x\|_2^2$ two ways, we have that

$$\begin{aligned}\sum_i (h_i x)^2 &= \|x_1 h^1 + \ldots + x_{N-2} h^{N-2}\|_2^2 \\ &= \sum_j x_j^2 \|h^j\|_2^2 + \sum_{j \neq k} x_j x_k (h^j \cdot h^k).\end{aligned}$$

The entries of H_t are all ± 1, and $h^j \cdot h^k = 0$ for $j \neq k$ (since the columns of H_t and therefore H'_t are orthogonal), so this becomes

$$\begin{aligned}&= (N-2) \sum_j x_j^2 \\ &\geqslant (N-2)^2,\end{aligned}$$

since $x_i^2 \geqslant 1$ by assumption. Therefore by averaging there is some i for which $(h_i x)^2 \geqslant \frac{(N-2)^2}{N}$, and so $|h_i x| \geqslant \frac{N-2}{\sqrt{N}}$), as desired. □

Next we translate H'_t into an instance of the panel selection problem and argue it has the desired properties. Take $M := \frac{1}{2}(H_t + 1^{N \times (N-2)})$ to be the $\{0,1\}$ matrix derived from H'_t.

The fact that M has uniform column norm $k = N/2$ directly follows from a property of Walsh matrices. Therefore we may apply Lemma 14 to argue that M is realizable as the individual-panel membership matrix for some instance of the panel selection problem, with $n = N$, $|\mathcal{K}| = N-2$, and $k = N/2$.

To conclude, consider the uniform $p = \left(\frac{1}{N-2}, \ldots, \frac{1}{N-2}\right)$, with $m = a(N-2) + (N-2)/2$ for any $a \in \mathbb{Z}_+$. In this case, each coordinate of p falls evenly between multiples of $1/m$ and must be rounded to multiples of $1/m$. Letting $x := p - \lfloor mp \rfloor / m = (1/2m, \ldots, 1/2m)$ be this vector of remainders, we must replace it with some $\bar{x} \in (\mathbb{Z}/m)^{N-2}$, while maintaining that $\sum_j \bar{x}_j = \sum_j x_j = (N-2)/2m$, so that the resulting $\bar{p} = \lfloor mp \rfloor / m + \bar{x}$ remains a distribution over panels. (Note that here negative $\bar{x}_j$ signify that the distribution mass on panel j decreases from p to $\bar{p}$.)

Explicitly, we then have

$$\|\pi - \bar{\pi}\|_\infty = \|Mp - M\bar{p}\|_\infty \tag{5.14}$$

$$= \|M(x - \bar{x})\|_\infty \tag{5.15}$$

$$= \frac{1}{2m}\|My\|_\infty, \tag{5.16}$$

where $y := 2m(\bar{x} - x)$.

$$= \frac{1}{2m}\|\frac{1}{2}H_t'y + \frac{1}{2}1^{N\times(N-2)}y\|_\infty \tag{5.17}$$

$$= \frac{1}{4m}\|H_t'y\|_\infty, \tag{5.18}$$

where $\sum_i y_i = 0$ because we require that $\bar{p}$ remain a distribution. Then since $y \in (2\mathbb{Z}+1)^{N-2}$, by Lemma 15 we have

$$\geqslant \frac{N-2}{4m\sqrt{N}} \tag{5.19}$$

$$= \Omega\left(\frac{\sqrt{k}}{m}\right), \tag{5.20}$$

since $k = N/2$.

This holds for all $y \in (2\mathbb{Z}+1)^{N-2}$. Recall that $\overline{\mathcal{D}} := \{\bar{p} \in (\mathbb{Z}_+/m)^{|\mathcal{K}|} : \|\bar{p}\|_1 = 1\}$, and so

$$\overline{\mathcal{D}} \subseteq \{p + \bar{\Delta}/2m\}.$$

Therefore (5.20) implies that

$$\min_{\bar{p}\in\overline{\mathcal{D}}} \|\pi - \bar{\pi}\|_\infty = \Omega\left(\frac{\sqrt{k}}{m}\right),$$

as desired. □

Our k-dependent upper and lower bounds are separated by a factor of $\sqrt{k}$, matching the current upper and lower bounds of the Beck-Fiala conjecture as applied to linear discrepancy (also known as the lattice approximation problem [Sri97]). The respective gaps are incomparable, however, since for a given $x \in [0,1]^n$, the former problem aims to minimize $\|M(x-\bar{x})\|_\infty$ over $\bar{x} \in \{0,1\}^n$, while we aim to do the same over a subset of the $\bar{x} \in \mathbb{Z}^n$ for which $\sum_j x_j = \sum_j \bar{x}_j$ (see Lemma 14).

5.4 Theoretical Bounds on Fairness Loss

Since the fairness of a distribution p is determined by its marginals π, it is intuitive that if uniform lotteries incur only small marginal discrepancy (per Section 5.3), then they should also incur only small fairness losses. This should hold for any fairness notion that is sufficiently "smooth" (i.e., doesn't change too quickly with changing marginals) in the vicinity of p, π.

Although our bounds from Section 5.3 apply to any reasonable initial distribution p, we are particularly concerned with bounding fairness loss from *maximally fair* initial distributions p^*. Here, we specifically consider such p^* that are optimal with respect to Maximin and NW. We note that, since there exist anonymous p^*, π^* that maximize these objectives, we can apply any upper bound from Section 5.3 to upper bound $\|\pi^* - \bar{\pi}\|_\infty$.

5.4.1 Maximin

Since Leximin is the fairness objective optimized by the maximally fair algorithm used in practice, it would be most natural to start with a p^* that is Leximin-optimal and bound fairness loss with respect to this objective. However, the fact that Leximin fairness cannot be represented by a single scalar value prevents us from formulating such an approximation guarantee. Instead, we first pursue bounds on the closely-related objective, Maximin. We argue that in the most meaningful sense, a worst-case Maximin guarantee *is* a Leximin guarantee: such a bound would show limited loss in the minimum marginal, and it is Leximin's *lexicographically first priority* to maximize the minimum marginal.

First, we show there exists some $\bar{p}, \bar{\pi}$ that gives bounded Maximin loss from p^*, π^*, the Maximin-optimal unconstrained distribution. This bound follows from Theorems 34 and 44, using the simple observation that $\bar{p}$ can decrease the lowest marginal given by p^* by no more than $\|\pi^* - \bar{\pi}\|_\infty$. Here $n_{min} := \min_c n_c$ denotes the smallest number of individuals which share any feature vector $c \in \mathcal{C}$.

Theorem 36. *By Theorem 34 and 44, for Maximin-optimal p^*, there exists a uniform lottery $\bar{p}$ that satisfies*

$$\text{Maximin}(p^*) - \text{Maximin}(\bar{p}) = \frac{1}{m} \cdot O\left(\min\left\{\sqrt{|\mathcal{C}| \log |\mathcal{C}|},\ \frac{k}{n_{min}} + 1\right\}\right).$$

Theorem 35 demonstrates that we cannot get an upper bound on Maxmin loss stronger than $O(\sqrt{k}/m)$ using a uniform bound on changes in all π_i. However, since Maximin is concerned only with the smallest π_i, it seems plausible that better upper bounds on Maximin loss could result from rounding π while tightly controlling only losses in the smallest π_i's, while giving freer reign to larger marginals. We show that this is not the case by further modifying the instances from Theorem 35 to obtain the following lower bound on the Maximin loss:

Theorem 37. *There exists a Maximin-optimal p^* such that, for all uniform lotteries $\bar{p}$,*

$$\text{Maximin}(p^*) - \text{Maximin}(\bar{p}) = \Omega\left(\frac{\sqrt{k}}{m}\right).$$

Proof. We will follow the proof of Theorem 35: first we use the Walsh matrices to construct a matrix with the desired properties, prove a modified version of Lemma 15 for it, and then appeal to Lemma 14 to argue that it corresponds to a realizable instance of the panel selection problem.

In contrast to the construction in Theorem 35, where we need only demonstrate that *some* $\bar{\pi}_i$ deviates from π_i, we must construct an instance for which (essentially) the minimum π_i necessarily *decreases*. We accomplish this by first modifying the Walsh matrices to have uniform row norm, so that π is uniform and all π_i are minimal. We then introduce a second set of "twin" individuals, each i' of which is a member of the panels which their twin i is not. This ensures that any discrepancy in $\bar{\pi} - \pi$ is witnessed in the downward direction.

To begin, again let H_t be the $2^t \times 2^t$ Walsh matrix, with $N := 2^t$ its dimension. This time we take H_t^* to be the submatrix derived by dropping the first row of H_t. By properties of Walsh matrices, all remaining rows in H_t^* have an equal number of 1 and -1 entries, (though this is no longer true of the columns).

Again letting h_i denote the rows of H_t^*, and h^j denote its columns, we have the following new version of Lemma 15, which requires the additional assumption that $\sum_j x_j = 0$:

Lemma 16.

$$\min_{x \in \Delta^*} \|H_t^* \, x\|_\infty \geqslant \sqrt{N},$$

where $\Delta^* := \{x \in \{\ldots, -3, -1, 1, 3, \ldots\}^N : \sum_j x_j = 0\}$.

Proof. This right-hand side is $H_t' \, x = (h_1 x, \ldots, h_{N-1} x)^T$. We aim to show that there is some i for which $|h_i x|$ is large. Writing $\|H_t' \, x\|_2^2$ two ways, we have that

$$\begin{aligned}
\sum_i (h_i x)^2 &= \|x_1 h^1 + \ldots + x_N h^N\|_2^2 \\
&= \sum_j x_j^2 \|h^j\|_2^2 + \sum_{j \neq k} x_j x_k (h^j \cdot h^k)
\end{aligned}$$

the entries of H_t^* are all ± 1, and $h^j \cdot h^k = -1$ for $j \neq k$ (since the columns of H_t were orthogonal), so this becomes

$$\begin{aligned}
&= (N-1) \sum_j x_j^2 - \sum_{j \neq k} x_j x_k \\
&= N \sum_j x_j^2 - \sum_j \sum_k x_j x_k \\
&= N \sum_j x_j^2 \\
&\geqslant N^2,
\end{aligned}$$

since $x_i^2 \geqslant 1$ by assumption. Therefore by averaging, there is some i for which $(h_i x)^2 \geqslant \frac{N^2}{N-1}$, and so $|h_i x| \geqslant \sqrt{N}$, as desired. □

As constructed, all rows of H_t^* have the same number of 1s, so when we transform it into some M for some instance of the panel selection problem, it will yield that the marginals π of uniform p are uniform. However we cannot yet apply Lemma 14, since the columns of the resulting M do not have constant norm; in particular, the first column will be all 1s.

In order to simultaneously correct for this and translate from ℓ_∞ to Maximin lower bounds, we introduce "twins" for each i. Letting $M^* = \frac{1}{2}(H_t^* + 1^{(N-1)\times N})$ be this $\{0,1\}$ matrix, define $\bar{M}^* := 1^{(N-1)\times N} - M^*$ to be its complement, so that $M_{ij}^* = 1 - \bar{M}_{ij}^*$ for all i, j. Finally take

$$M = \begin{bmatrix} M^* \\ \bar{M}^* \end{bmatrix}$$

and observe that this $M \in \{0,1\}^{(2N-2)\times N}$ has uniform column norm $N-1$ because of $\bar{M}^*$. We may therefore apply Lemma 14 to claim that it is the individual-panel membership matrix of some instance of the panel selection problem.

The remainder of the argument proceeds similarly to that of Lemma 15, with additional step of showing that the lower bound holds for the maximin objective. We include the full argument for completeness.

Similarly take $p = \left(\frac{1}{N}, \ldots, \frac{1}{N}\right)^T$, with $m = aN + N/2$ for any $a \in \mathbb{Z}_+$, $n = 2N-2$ (the number of individuals), and $k = N-1$. This p gives equal marginals: here $\pi_i = (Mp)_i = \frac{N-1}{2N-2} = \frac{k}{n}$ for all i. Again each coordinate of p falls evenly between multiples of $1/m$ and must be rounded to multiples of $1/m$. Letting $x := p - \lfloor mp \rfloor/m = (1/2m, \ldots, 1/2m)^T$ be this vector of remainders, we must replace it with some $\bar{x} \in (\mathbb{Z}/m)^N$, while maintaining that $\sum_j \bar{x}_j = \sum_j x_j = N/2m$, so that the resulting $\bar{p} = \lfloor mp \rfloor/m + \bar{x}$ remains a distribution over panels.

Explicitly, we then have

$$\begin{aligned} \|\pi - \bar{\pi}\|_\infty &= \|Mp - M\bar{p}\|_\infty && (5.21) \\ &= \|M(x - \bar{x})\|_\infty && (5.22) \\ &= \frac{1}{2m} \| \begin{bmatrix} M^* \\ \bar{M}^* \end{bmatrix} y \|_\infty, && (5.23) \end{aligned}$$

where $y := 2m(\bar{x} - x)$. Because ℓ_∞ is a maximum, this is

$$\geqslant \frac{1}{2m}\|M^* y\|_\infty \tag{5.24}$$
$$= \frac{1}{2m}\|\frac{1}{2}H_t^* y + \frac{1}{2}1^{(N-1)\times N} y\|_\infty \tag{5.25}$$
$$= \frac{1}{4m}\|H_t^* y\|_\infty, \tag{5.26}$$

where $\sum_i y_i = 0$ because we require that $\bar{p}$ remain a distribution. Then since $y \in (2\mathbb{Z}+1)^{N-2}$, by Lemma 15 we have

$$\geqslant \frac{\sqrt{N}}{4m} \tag{5.27}$$
$$= \Omega\left(\frac{\sqrt{k}}{m}\right), \tag{5.28}$$

since $k = N - 1$. Again since $\overline{\mathcal{D}} \subseteq \{p + \bar{\Delta}/2m\}$, we then have

$$\min_{\bar{p}\in\overline{\mathcal{D}}} \|\pi - \bar{\pi}\|_\infty = \Omega\left(\frac{\sqrt{k}}{m}\right).$$

Since π is uniform by construction (and so these p and π are optimal with respect ot Maximin), this is a lower bound on the discrepancy of each marginal which was minimal *before deviation*. It finally remains to show that this deviation happens in the downward direction, so that the minimum marginal decreases by at least this amount. Observe that by the construction of $\bar{M}^*$, for all $\bar{p}$ we have $(M^*\bar{p})_i = -(\bar{M}^*\bar{p})_i$. Therefore for any given $\bar{p}$, whichever coordinate i satisfies $|(\pi - \bar{\pi})_i| = \Omega(\sqrt{k}/m)$, there is a coordinate i' for which $(\pi - \bar{\pi})_{i'} = \Omega(\sqrt{k}/m)$. Therefore in this instance

$$\mathrm{Maximin}(p^*) - \max_{\bar{p}\in\overline{\mathcal{D}}} \mathrm{Maximin}(\bar{p}) = \Omega\left(\frac{\sqrt{k}}{m}\right),$$

as desired. □

5.4.2 Nash Welfare

As NW has also garnered interest by practitioners and is applicable in practice [HG20], we upper-bound the NW fairness loss. Unlike Maximin loss, an upper bound on NW loss does not immediately follow from one on $\|\pi - \bar{\pi}\|_\infty$, because decreases in smaller marginals have larger negative impact on the NW. As a result, the upper bound on NW resulting from Section 5.3 is slightly weaker than that on Maximin:

Theorem 38. *For* NW*-optimal p^*, there exists a uniform lottery $\bar{p}$ that satisfies*

$$\mathrm{NW}(p^*) - \mathrm{NW}(\bar{p}) = \frac{k}{m} \cdot O\left(\min\left\{\sqrt{|\mathcal{C}|\log|\mathcal{C}|},\ \frac{k}{n_{min}} + 1\right\}\right).$$

We give an overview of the proof of Theorem 38. To begin, fix a NW-optimizing panel distribution p^*, π^*. Before applying our upper bounds on marginal discrepancy from Section 5.3, we must contend with the fact that if this bounded loss is suffered by already-tiny marginals, the NW may decrease substantially or even go to 0. Thus, we first prove Lemmas 39 and 40, which together imply that no marginal in π^* is smaller than $1/n$.

Theorem 39. *For* NW*-optimal p^* over a support of panels* $\mathrm{supp}(p^*)$*, there exists a constant $\lambda \in \mathbb{R}^+$ such that, for all $P \in \mathrm{supp}(p^*)$, $\sum_{i \in P} 1/\pi_i^* = \lambda$.*

Proof. We can write the problem of finding the NW optimizing distribution over a fixed panel support $\mathcal{P} \subseteq \mathcal{K}$ as below on the left, where $NW^n(p)$ is equal to the product of the π_i, the marginals implied by the panel distribution p (in contrast, in Section 5.2, we let $NW(p)$ be the geometric mean—here we we take the n^{th} power). On the right, we've rewritten the program in standard form, where we set $f(p) = -NW^n(p)$, $h(p) = p_1 + p_2 + \cdots + p_{|\mathcal{P}|} - 1$, and $g_j(p) = -p_j$. Observe that, $\forall j \in [|\mathcal{P}|]$, $\nabla h(p) = \mathbf{1}$ and $\nabla g_j(p) = -e_j$, where e_j is the vector of 0s with a 1 at index j.

$$\begin{array}{ll} \max\limits_p NW^n(p) & \min\limits_p f(p) \\ \|p\|_1 = 1 & h(p) = 0 \\ p_j \geqslant 0\ \forall j \in [|\mathcal{P}|] & g_j(p) \leqslant 0\ \forall j \in [|\mathcal{P}|] \end{array}$$

Now, let p^* be an optimal solution to this program, and $\mathrm{supp}(p^*)$ be its support, i.e., the set of panels to which p^* assigns nonzero probability. Then, since the objective and constraints of the above program are continuously differentiable over their entire support (and thus at p^*), by the KKT condition Stationarity, there exist some constants λ and μ_j for all $j \in [|\operatorname{supp}(p^*)|]$ (where $\mathbf{0}$ is the zero vector) such that

$$\nabla f(p^*) + \lambda \nabla h(p^*) + \sum_{j \in [|\operatorname{supp}(p^*)|]} \mu_j \nabla g(p^*) = \mathbf{0} \implies (\nabla f(p^*))_j = \mu_j - \lambda$$

By dual feasibility and primal feasibility respectively, we have that $\mu_j, p_j \geqslant 0$ for all $j \in [|\operatorname{supp}(p^*)|]$; by complementary slackness, we have that $\sum_{j \in [|\operatorname{supp}(p^*)|]} \mu_j p_j^* = 0$. Thus, for all j, either $p_j^* = 0$, or $p_j^* > 0$ and $\mu_j = 0$. We have restricted $\operatorname{supp}(p^*)$ to panels j in which $p_j^* > 0$, so we conclude that $\mu_j = 0$. It follows that

$$\frac{\partial NW^n(p^*)}{\partial p_j^*} = -\left(\nabla f(p^*)\right)_j = -(\mu_j - \lambda) = \lambda \quad \forall j \in \operatorname{supp}(p^*)$$

Finally, we can conclude the proof by expressing this partial derivative for fixed p_j (which as shown, has a constant value across all j in the support) in terms of the marginals π. We obtain that for all j in $\operatorname{supp}(p^*)$,

$$\lambda = \frac{\partial NW^n(p^*)}{\partial p_j^*} = \sum_{i \in N} \frac{NW^n(p^*)}{\pi_i^*} \frac{\partial \pi_i^*}{\partial p_j^*} = \sum_{i \in P_j} \frac{NW^n(p^*)}{\pi_i^*} = NW^n(p^*) \left(\sum_{i \in P_j} \frac{1}{\pi_i^*} \right)$$

where P_j is the j^{th} panel in $\operatorname{supp}(p^*)$. The second equality is by the product rule for derivatives, where each term of the resulting sum is equal to the derivative of π_i^* with respect to p_j^* multiplied by NW/π_i^*, the NW holding out the marginal of individual i. The third equality is by the fact that if $i \in P_j$, then $\partial \pi_i^* / \partial p_j^* = 1$; otherwise $\partial \pi_i^* / \partial p_j^* = 0$. □

Theorem 40. *For* NW*-optimal p^*, π^*, we have that $\pi_i^* \geqslant 1/n$ for all $i \in N$.*

Proof. Let $X[P \ni i]$ be the indicator that a panel P contains individual i. Then,

$$\mathbb{E}_{P \sim p^*}\left[\sum_{i \in P} \frac{1}{\pi_i^*}\right] = \mathbb{E}_{P \sim p^*}\left[\sum_{i \in N} \frac{X[P \ni i]}{\pi_i^*}\right] = \sum_{i \in N} \frac{\mathbb{E}_{P \sim p^*}[X[P \ni i]]}{\pi_i^*} = \sum_{i \in N} \frac{\pi_i^*}{\pi_i^*} = n$$

By Theorem 39, we also have that $\mathbb{E}\left[\sum_{i \in P} \frac{1}{\pi_i^*}\right] = \lambda / NW^n(p^*)$, and thus $\lambda/NW^n(p^*) = n$. It follows that for all panels P, $\sum_{i \in P} \frac{1}{\pi_i^*} = \lambda/NW^n(p^*) = n$ and therefore $\pi_i^* \geqslant 1/n \; \forall i \in N$; otherwise, we would have some panel P for which $\sum_{i \in P} \frac{1}{\pi_i} > n$, a contradiction. □

Theorem 39 follows from the fact that the partial derivative of NW with respect to the probability it assigns a given panel must be the same as that with respect to any other panel at p^* (otherwise, mass in the distribution could be shifted to increase the NW). Theorem 40 then follows by the additional observation that $\mathbb{E}_{P \sim p^*}\left[\sum_{i \in P} 1/\pi_i^*\right] = n$.

Finally Theorem 41 follows from the fact that Theorem 40 limits the potential multiplicative, and therefore additive, impact on the NW of decreasing any marginal by $\|\pi - \bar{\pi}\|_\infty$:

Theorem 41. *For* NW-*optimal* p^*, π^*, *there exists a uniform lottery* $\bar{p}, \bar{\pi}$ *that satisfies* $\mathrm{NW}(p^*) - \mathrm{NW}(\bar{p}) \leqslant k \, \|\pi^* - \bar{\pi}\|_\infty$.

Proof. Let π^*_{min} be the smallest marginal of any individual implied by the Nash-optimal distribution over panels p^*, i.e., $\pi^*_{min} = \min_{i \in N} \pi^*_i$. Then, to upper-bound the loss in NW, we assume an unattainable worst case that between p^*, π^* and a given uniform lottery $\bar{p}, \bar{\pi}$, *all* individuals probabilities suffer the largest loss of any marginal, $\|\pi^* - \bar{\pi}\|_\infty$, and that this loss manifests multiplicatively as badly as if all agents had original marginal probability π^*_{min}. This first gives the multiplicative bound:

$$NW(\overline{p}^*) \geqslant NW(p^*) \left(\frac{\pi^*_{min} - \|\pi^* - \bar{\pi}\|_\infty}{\pi^*_{min}} \right) = NW(p^*) \left(1 - \frac{\|\pi^* - \bar{\pi}\|_\infty}{\pi^*_{min}} \right).$$

Rearranging the above conclusion and then applying the facts that $NW(p^*) \leqslant k/n$ (trivially) and $\pi^*_{min} \geqslant 1/n$ (Theorem 40), we get the desired additive bound:

$$NW(p^*) - NW(\overline{p}) \leqslant NW(p^*) \cdot \frac{\|\pi^* - \bar{\pi}\|_\infty}{\pi^*_{min}} \leqslant \frac{k}{n} \cdot \frac{\|\pi^* - \bar{\pi}\|_\infty}{1/n} \leqslant k \, \|\pi^* - \bar{\pi}\|_\infty \qquad \square$$

As the NW-optimal marginals π^* are anonymous, we can apply the upper bounds given by Theorem 34 and Theorem 44 to show the existence of a $\bar{p}, \bar{\pi}$ satisfying the claim of the theorem.

5.5 Practical Algorithms for Computing Fair Uniform Lotteries

Algorithms

First, we implement versions of two existing rounding algorithms, which are implicit in our worst-case upper bounds.[8] The first is Pipage rounding [Gan+06], or Pipage, a randomized rounding scheme satisfying negative association [DJR07]. The second is Beck-Fiala, the dependent rounding scheme used in the proof of Theorem 33. To benchmark these algorithms against the highest level of fairness they could possibly achieve, we use integer programming (IP) to compute the fairest possible uniform lotteries over $supp(p^*)$, the panels over which p^* randomizes.[9] We define IP-Maximin and IP-NW to find uniform lotteries over $supp(p^*)$ maximizing Maximin and NW, respectively. We remark that the performance of these IPs is still subject to our theoretical upper and lower bounds. We provide implementation details in Section 5.9.1.

One question is whether we should prefer the IPs or the rounding algorithms for real-world applications. Although IP-Maximin appears to find good solutions at practicable speeds, IP-NW converges to optimality prohibitively slowly in some instances (see Section 5.9.2 for runtimes). At the same time, we find that our simpler rounding algorithms give near-optimal uniform lotteries with respect to both fairness objectives. Also in favor of simpler rounding algorithms, many randomized rounding procedures (including pipage rounding) have the advantage that they exactly preserve marginals over the combined steps of randomly rounding to a uniform lottery and then randomly sampling it—a guarantee that is much more challenging

[8]We do not implement the algorithm implicit in Theorem 34 because our results already present sufficient alternatives for finding excellent uniform lotteries in practice.

[9]Note that these lotteries are not necessarily universally optimal, as they can randomize over only $supp(p^*)$; conceivably, one could find a fairer uniform lottery by also randomizing over panels not in $supp(p^*)$. However, Pipage and Beck-Fiala are also restricted in this way, and thus must be weakly dominated by the IP.

to achieve with IPs.

Uniform lotteries nearly exactly preserve Maximin, Nash Welfare fairness

We first measure the fairness of uniform lotteries produced by these algorithms in 11 real-world panel selection instances from 7 different organizations worldwide (instance details in Section 5.7). In all experiments, we generate a lottery of size $m = 1000$. This is fairly small; it requires drawing only 3 balls from lottery machines, and in one instance we have that $m < n$. We nevertheless see excellent performance of all algorithms, and note that this performance will only improve with larger m.

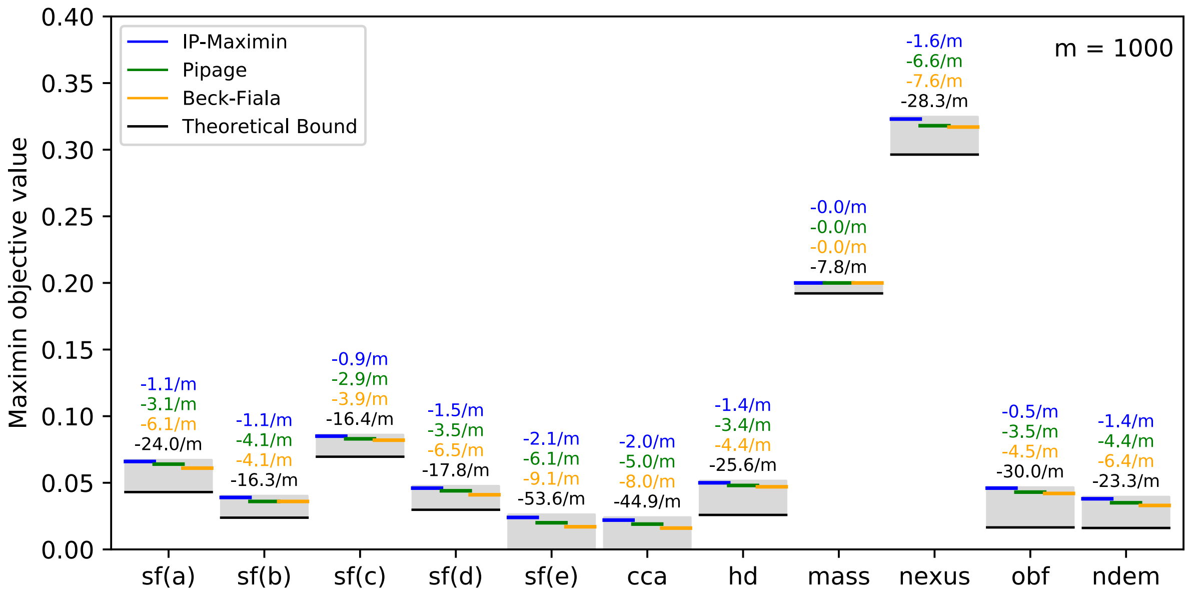


Figure 5.2: $m = 1000$. Shaded regions extend from Maximin(p^*), the fairness of the optimal unconstrained distribution, down to the minimum fairness implied by the tightest theoretical upper bound in that instance (in all instances but "obf" Theorem 34 is tightest). Each algorithm or bound's loss relative to Maximin(p^*) is written above in the corresponding color. We show a representative run of PIPAGE, a randomized algorithm.

Section 5.5 shows the Maximin fairness of the uniform lottery computed by PIPAGE, BECK-FIALA, and IP-MAXIMIN for each instance. For intuition, recall that the level of Maximin fairness given by any lottery is exactly the minimum marginal assigned to any individual by that lottery. The upper edges of the gray boxes in Section 5.5 correspond to the optimal fairness attained by an unconstrained distribution p^*. These experiments reveal that the cost of transparency to Maximin-fairness is practically non-existent: across instances, the quantized distributions computed by IP-MAXIMIN decrease the minimum marginal by at most $2.1/m$, amounting to a loss of no more than 0.0021 in the minimum marginal probability in any instance. Visually, we can see that this loss is negligible relative to the original magnitude of even the smallest marginals given by p^*. Surprisingly, though PIPAGE and BECK-FIALA do not aim to optimize any fairness objective, they achieve only slightly larger losses in Maximin fairness, with PIPAGE outperforming BECK-FIALA. Finally, the heights of the gray boxes indicate that our theoretical bounds are often meaningful in practice, giving lower bounds on Maximin fairness well above zero in nine out of eleven instances. We note these bounds only tighten with larger m. We present similarly encouraging results on NW loss in Section 5.9.3.

Uniform lotteries nearly preserve all Leximin marginals

We still remain one step away from practice: our examination of Maximin does not address whether uniform lotteries can attain the finer-tuned fairness properties of the Leximin-optimal distributions currently used in practice. Fortunately, our results from Section 5.3 imply the existence of a quantized $\bar{p}$ that closely approximates *all marginals* given by the Leximin-optimal distribution p^*, π^*. We evaluate the extent to which PIPAGE and BECK-FIALA preserve these marginals in Figure 5.3. They are benchmarked against a new IP, IP-MARGINALS, which computes the uniform lottery over *supp*(p^*) minimizing $\|\pi^* - \bar{\pi}\|_\infty$.

Figure 5.3 demonstrates that in the instance "sf(a)", all algorithms produce marginals that deviate negligibly from those given by π^*. Analogous results on remaining instances appear in Section 5.9.4 and show similar results. As was the case for Maximin, we see that our theoretical

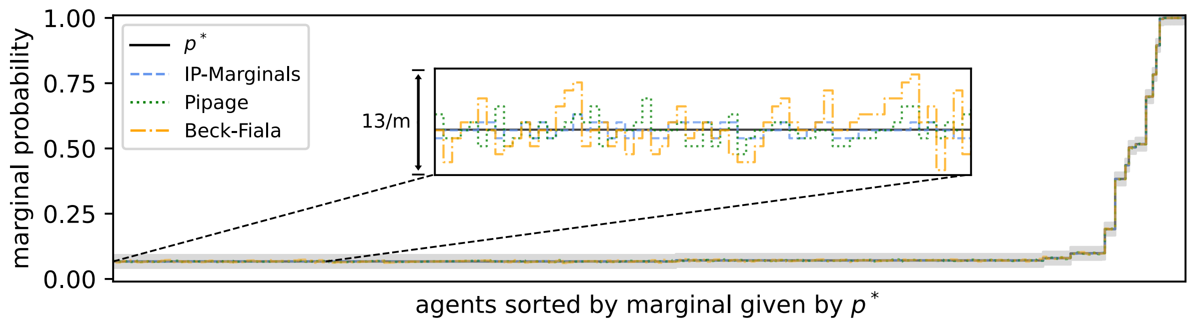


Figure 5.3: Instance = sf(a), $m = 1000$. Line plot shows the Leximin-optimal marginals π^* (implied by panel distribution p^*), along with marginals given by all algorithms sorted according to π^*. Note that each x coordinate then corresponds to an individual. The zoomed box shows the magnitude of marginal discrepancy around π^*. The surrounding shaded region shows the tightest theoretical bound on the marginal discrepancy, in this case from Theorem 34, around the optimal marginals. We show a representative run of Pipage, a randomized algorithm.

bounds are meaningful, but that we can consistently outperform them in real-world instances.

5.6 Discussion

Our aim was to show that uniform lotteries can preserve fairness, and our results ultimately suggest this, along with something stronger: that in practical instances, uniform lotteries can reliably almost exactly replicate *the entire set of marginals* given by the optimal unconstrained panel distribution. Our rounding algorithms can thus be plugged directly into the existing panel selection pipeline with essentially no impact on individuals' selection probabilities, thus enabling translation of the output of Panelot (and other maximally fair algorithms) to a nearly maximally fair *and* transparent panel selection procedure. We note that our methods are not just compatible with ball-drawing lotteries, but any form of uniform physical randomness (e.g. dice, wheel-spinning, etc.).

Although we achieve our stated notion of transparency, a limitation of this notion is that it focuses on the final stage of the panel selection process. A more holistic notion of transparency might require that onlookers can verify that the panel is not being intentionally stacked with certain individuals. This work does not fully enable such verification: although onlookers can now observe individuals' marginals, they still cannot verify that these marginals are *actually maximally fair* without verifying the underlying optimization algorithms. In particular, in the common case where quotas require even maximally fair panel distributions to select certain individuals with probability near one, onlookers cannot distinguish those from unfair distributions engineered such that one or more pool members are chosen with probability near one.

In research on economics, fair division, and other areas of AI, randomness is often proposed as a tool to make real-world systems fairer [Gri04; Bud+13; FSV20]. Nonetheless, in practice, these systems (with a few exceptions, such as school choice [New21]) remain stubbornly deterministic. Among the hurdles to bringing the theoretical benefits of randomness into practice is that allocation mechanisms fare best when they can be readily understood, and that randomness can be perceived as undesirable or suspect. Sortition is a rather unique paradigm at the heart of this tension: it relies centrally on randomness, while in the public sphere it is attaining increasing political influence. It is therefore a uniquely high-impact domain in which to study how to combine the benefits of randomness, such as fairness, with transparency. We hope that this work and its potential for impact will inspire the investigation of fairness-transparency tradeoffs in other AI applications.

5.7 Panel Selection Datasets

We examine data from the following 11 real-world sortition panel selection instances, generously provided to us by several groups that specialize in organizing citizens' assemblies. Section 5.7 shows the instance short-names we use throughout the chapter, and which organization was responsible for each panel. The final two columns compare the values of our theoretical upper bounds on the marginal discrepancy, illustrating that in all instances except "obf", the bound from Theorem 34 is tighter. Finally, we give some metadata about each instance, which is required for calculating the values of our theoretical upper bounds.

In particular, n = number of pool members, k = number panel members, $\mathcal{C}$ = set of distinct realized feature-vectors in the pool. Precise constants used for computing exact the upper bounds are derived in Section 5.8: the Theorem 33 bound is exactly k/m, the Theorem 34 bound is exactly

$$\frac{\sqrt{\frac{1}{2}(1+\frac{\ln 2}{\ln|\mathcal{C}|})}\cdot\sqrt{|\mathcal{C}|\ln(|\mathcal{C}|)}+1}{m},$$

and the Theorem 44 bound is exactly $\frac{2k/n_{min}+1}{m}$. In all instances, $n_{min}=1$.

Instance	Organization	n	k	$\lvert\mathcal{C}\rvert$	Thm 33	Thm 34	Thm 44
sf(a)	Sortition Foundation	312	35	182	$35/m$	$24.2/m$	$71/m$
sf(b)	Sortition Foundation	250	20	92	$20/m$	$16.5/m$	$41/m$
sf(c)	Sortition Foundation	161	44	92	$44/m$	$16.5/m$	$89/m$
sf(d)	Sortition Foundation	404	40	108	$40/m$	$18.0/m$	$81/m$
sf(e)	Sortition Foundation	1727	110	762	$110/m$	$53.8/m$	$221/m$
cca	Ctr. for Climate Ass.	825	75	554	$75/m$	$45.1/m$	$151/m$
hd	Healthy Democracy	239	30	202	$30/m$	$25.6/m$	$61/m$
mass	MASS LBP	70	24	25	$24/m$	$8.0/m$	$49/m$
nexus	Nexus	342	170	242	$170/m$	$28.4/m$	$341/m$
obf	Of By For	321	30	294	$30/m$	$31.6/m$	$61/m$
ndem	New Democracy	398	40	173	$40/m$	$23.5/m$	$81/m$

Table 5.1:: Instance parameters and resulting theoretical bounds

5.8 Additional Beyond-Worst-Case Upper Bounds

5.8.1 General Rounding Procedure

Throughout this section, we repeatedly face the task of rounding the entries of some distribution p to some vector $\bar{p}$ that must also be a valid distribution (i.e., have entries in $[0, 1]$ such that $\|\bar{p}\|_1 = 1$), and have entries that are integer multiples of $1/m$. However, many of the standard rounding procedures we apply, such as randomized rounding and discrepancy-based dependent rounding, only give guarantees for rounding probabilities to 0/1 vectors, rather than to multiples of $1/m$. Thus, in several proofs (Theorem 32, Theorem 33, Theorem 34, Theorem 44), we apply these canonical rounding methods to a *modified version* of our original vector p, called x'. After constructing x', we round it to a 0/1 vector $\bar{x}'$, from which we finally compute $\bar{p}$. We more precisely define this general rounding procedure, and characterize some of its useful properties, below.

Definition 11 (Procedure for using 0/1 rounding procedure to round p to $\bar{p}$)**.** Let p be a distribution, represented as a vector. Let x be the vector p with entries scaled by m, so that $x_j := m \cdot p_j$. Then, define the vector $\lfloor x \rfloor$, which we can think of as the "integer components" of each entry of x, i.e., $\lfloor x \rfloor_j := \lfloor m \cdot p_j \rfloor$. Finally, we define x' as the "decimal components" of the entries of x, so that $x' := x - \lfloor x \rfloor$. We will round x' to a 0/1 vector.

Then, construct $\bar{p}$ from p as follows:

1. Construct the vector x' as above.
2. Round x' to some 0/1 vector $\bar{x}'$ via a given rounding procedure such that $\|\bar{x}\|_1 = \|x'\|_1$.
3. Set $\bar{p}$ such that
$$\bar{p} := \frac{\lfloor x \rfloor + \bar{x}'}{m}.$$

At a high level, this rounding procedure can be thought of as scaling up the vector we want to round by m, holding this scaled vector's integer components aside and rounding its decimal components, and then adding the integer components back in and scaling back down by m.

Now, we show that this rounding procedure produces a $\bar{p}$ with the properties we want—(a) it has entries that are multiples of $1/m$ and (b) it is a valid distribution—as well as an additional property (c), which helps translate guarantees on existing roundinig schemes to guarantees in our setting.

Lemma 17. *Suppose we are given a 0/1 rounding scheme which, given $x' \in [0,1]^{|\mathcal{K}|}$ and constraint matrix M, produces some $\bar{x}'$ which satisfies*

- $\bar{x}' \in \{0,1\}^{|\mathcal{K}|}$,
- $\|\bar{x}'\|_1 = \|x'\|_1$, *and*
- $|(M(x' - \bar{x}'))_i| \leqslant g(i)$ *for each row i.*

Then given some distribution $p \in \mathbb{R}_+^{|\mathcal{K}|}$ and $m \in \mathbb{N}$, the procedure in Definition 11, using such a 0/1 rounding scheme, produces $\bar{p}$ such that

(a) $\bar{p} \in (\mathbb{Z}_+/m)^{|\mathcal{K}|}$,

(b) $\bar{p}$ *is a distribution, and*

(c) $|(M(p - \bar{p}))_i| \leqslant \frac{g(i)}{m}$ *for each row i.*

Proof. We prove each property separately:

(a) holds: $\bar{p}$ contains multiples of $1/m$, since in the general procedure (Definition 11), its entries are set to the sum of two integers divided by m.

(b) holds: $\bar{p}$ is a valid distribution: all entries of $\bar{p}$ must be non-negative, and we have that $\|\bar{p}\|_1 = \|p\|_1 = 1$, as shown below.

$$\|\bar{p}\|_1 = \left\|\frac{\lfloor x \rfloor + \bar{x}'}{m}\right\|_1 = \left\|\frac{\lfloor x \rfloor}{m}\right\|_1 + \left\|\frac{\bar{x}'}{m}\right\|_1 = \left\|\frac{\lfloor x \rfloor}{m}\right\|_1 + \left\|\frac{x'}{m}\right\|_1 = \|p\|_1$$

(c) holds: Fix some i and the corresponding row of $(M(p - \bar{p}))$, referred to as $(M(p - \bar{p}))_i$. Then,

$$|(M(p - \bar{p}))_i| = \left|\left(M\left(\frac{\lfloor x \rfloor + x'}{m} - \frac{\lfloor x \rfloor + \bar{x}'}{m}\right)\right)_i\right| = \frac{|(M(x' - \bar{x}'))_i|}{m} \leqslant \frac{g(i)}{m}$$

□

5.8.2 Additional Beyond-Worst-Case Upper Bounds

Since some of our beyond-worst-case upper bounds apply to anonymous realizable π, it is reasonable to ask how prevalent anonymous realizable π are, for arbitrary instances of sortition. Fortunately, we have the following claim:

Claim 42. For any instance of the panel selection problem and any realizable π, let π' be the "anonymized" marginals obtained by setting π'_i to the average $\pi_{i'}$ across all i' with the same feature vector as i. Then π' is realizable also.

Proof of Claim 42. Let π^* denote the "anonymization" of π, and take

$$\Pi := \left\{ \pi' : \text{ realizable, and for all } c, \sum_{i:F(i)=c} \pi'_i = \sum_{i:F(i)=c} \pi_i \right\}.$$

We will show that $\pi^* \in \Pi$.

We argue by way of contradiction. Let $\widehat{\pi}$ denote the "most anonymized" $\pi' \in \Pi$, in the sense that

$$\widehat{\pi} = \arg\min_{\pi' \in \Pi} \max_c \left(\max_{i:F(i)=c} \pi'_i - \min_{i:F(i)=c} \pi'_i \right).$$

Let i and i' be some pair of individuals with $F(i) = F(i')$ witnessing this maximum diameter, and let p be a distribution with marginals $\widehat{\pi}$. For each such pair, we will argue that p may be modified so that $\widehat{\pi}_i = \widehat{\pi}_{i'}$ while leaving all other marginals unchanged. By iteratively applying this to all such pairs, we will contradict the minimality of $\widehat{\pi}$.

To start, observe that by assumption $\widehat{\pi}_i > \widehat{\pi}_{i'}$. Let p' be the distribution over feasible panels which is the same as p, except that i and i' switch places in any panel on which either of them appear. All such panel replacements yield feasible panels, since they have the same feature vector c. Finally take $p_{new} = (p + p')/2$. As promised, this distribution has the property that $\pi_i = \pi_{i'}$ and all other marginals are unchanged. □

As a belated warm-up to the beyond-worst-case guarantees, we address the case when there is only one feature of interest, so that $F = \{f\}$. It turns out that we can obtain strong guarantees for this special case without using the machinery deployed in the proof of Theorem 34. We place no constraints on the size of the set of feature values Ω, nor do we require that π is anonymous.

Theorem 43. *If π is realizable and $|F| = 1$, then we may efficiently identify $\bar{p}$ such that its marginals $\bar{\pi}$ satisfy*

$$\|\pi - \bar{\pi}\|_\infty < \frac{2}{m}.$$

Proof of Theorem 43. Given marginals π, let p be a distribution over feasible panels $\mathcal{K}$ which witnesses π. The first step of this rounding is to consider the marginals τ_v of each feature value v: $\tau_v = \sum_{i:f(i)=v} \pi_i$. Note that $\sum_v \tau_v = \sum_i \pi_i = k$. Since there is only one feature, all feasible panels P satisfy

$$l_v \leqslant |\{i \in P : f(i) = v\}| \leqslant u_v, \tag{5.29}$$

and taking the expectation of this over p gives

$$l_v \leqslant \mathbb{E}_p[|\{i \in P : f(i) = v\}|] \leqslant u_v \tag{5.30}$$

$$l_v \leqslant \tau_v \leqslant u_v. \tag{5.31}$$

Therefore $l_v \leqslant \lfloor \tau_v \rfloor$ and $u_v \geqslant \lceil \tau_v \rceil$. We will construct a new distribution $\bar{p}$ over panels P which satisfy $\lfloor \tau_v \rfloor \leqslant |\{i \in P : f(i) = v\}| \leqslant \lceil \tau_v \rceil$ for all features v, and are therefore guaranteed to be feasible.

We will construct feasible panels via the following scheme. Consider the interval $[0, km] \subset \mathbb{R}$ as representing the km spots to be allocated across the m panels which will comprise our lottery, and let $s_t := [t-1, t)$ denote spot t. Next observe that $m \sum_i \pi_i = km$, and so $m\pi_i$ may be viewed as the expected number of spots which p would give to i.

First group the π_i by feature value to form $\tau_v = \sum_{i:f(i)=v} \pi_i$, and then pack them into $[0, km]$, so that individuals with common feature values have contiguous sections; let S_i denote the portion of $[0, km]$ allocated to i, so that $|S_i| = \pi_i$. We will choose an individual $I(t)$ for each spot s_t, and then assemble the m panels that comprise $\bar{p}$ by taking

$$P_r := \{I(t) : t = wm + r \text{ for } w \in \{0, \ldots, k-1\}\}, \tag{5.32}$$

for $r \in \{1, \ldots, m\}$.

How to choose which individual will get the spot t for each t? If $S_i \supseteq s_t$ then $I(t) = i$. Otherwise, s_t is split between two or more individuals, possibly with different feature values, in which case we call it *contested*. Observe that no matter how these contested s_t are allocated (no matter the choice of $I(t)$ for split t), it will be the case that $|\pi_i - \bar{\pi}_i| < 2/m$, since there is at most one contested s_t at each endpoint of the interval S_i.

It remains to argue that the panels chosen in (5.32) are feasible; in particular that $\lfloor \tau_v \rfloor \leqslant \bar{\tau}_v \leqslant \lceil \tau_v \rceil$ for all v. By construction, each panel P_r has some number of spots which will necessarily be allocated to an individual with feature value v, and some number of spots which are contested and may or may not be allocated to an individual with feature vector v. For each value v, there are at most two spots in all of $[0, km]$ which are *type contested* in this way. If some panel P_r contains at most one type-contested spot for type v, then no matter which way it is allocated, $|\{i \in P : f(i) = v\}| - \tau_v| < 1$, and so P_r is feasible with respect to v. In the worst case, for some given v both of the spots which are type-contested by v appear on the same panel P_r. In order to ensure that $|\{i \in P : f(i) = v\}| - \tau_v| < 1$, it must be the case that exactly one of these two spots is allocated to some i for which $f(i) = v$. Fortunately this constraint is easily satisfiable, even in the case when a given panel P_r contains both of the type-contested spots for multiple features v.

Therefore the $\bar{p}$ as constructed by (5.32) is supported by panels which are not only feasible but respect quotas which are maximally tight, given that the input p, π was realizable. Finally since each i contests at most two spots, we have that

$$\|\pi - \bar{\pi}\|_\infty < \frac{2}{m}. \qquad \square$$

Theorem 44. *Given realizable anonymous π, we may efficiently identify $\bar{p}, \bar{\pi}$ such that*

$$\|\pi - \bar{\pi}\|_\infty = O\left(\frac{1}{m} \max\left\{\frac{k}{n_{min}}, 1\right\}\right),$$

where $n_{min} := \min_c n_c$ is the minimum number of individuals in the pool which share any one feature vector.

Proof. We proceed as in the proof of Theorem 34, but apply a different rounding to the panel type LP to obtain $\bar{\mathfrak{p}}$. To begin, p, π projects to some $\mathfrak{p}, \tau$. Without loss of generality assume that it is a basic solution to the Type LP (5.8).

We will construct $\bar{\mathfrak{p}}$ from $\mathfrak{p}$ by applying 0/1 rounding as in Definition 11.

Note that the constraint matrix Q in (5.7) has the property that for all columns q^j, $\|q^j\|_1 = k$. As a special case of [Doe07, Theorem 6], applied to x' and the panel type LP, there exists an $\bar{x}' \in \{0,1\}^{|\mathfrak{K}|}$ such that

$$\|Q(x' - \bar{x}')\|_\infty < 2k.$$

and for which $\|\bar{x}'\|_1 = \|x\|_1$. (This follows from a generalization of the Beck-Fiala algorithm which both respects hard constraints and applies to arbitary matrices Q with bounded column norms, and is therefore also algorithmic.)

Applying Lemma 17, we then have

$$\|\tau - \bar{\tau}\|_\infty < \frac{2k}{m}.$$

Given that such a $\bar{\mathfrak{p}}, \bar{\tau}$ exists, it remains to generate $\bar{p}$ and $\bar{\pi}$ in such a way as to give the desired bound on the discrepany in individual marginals. We proceed in a manner identical to the proof of Theorem 34.

Again we have that $\bar{\tau} \geqslant 0$ and $\bar{\tau} = \sum_j Q_{cj}\bar{\mathfrak{p}}_j \leqslant m_c \leqslant n_c$, where $m_c = \max_j Q_{cj}$ and n_c is the number of individuals i for which $F(i) = c$, since $\bar{\mathfrak{p}}$ is a distribution over feasible panel types j. Therefore dividing $\bar{\tau}$ amongst the $\bar{\pi}_i$ as equally as possible for each c gives $\bar{\pi}_i \in [0, 1]$.

By the anonymity of π, for all i with $F(i) = c$, $\pi_i = \tau_c/n_c$, and dividing the spots in $\bar{\mathfrak{p}}$ for feature vector c as equally as possible amongst the n_c individuals gives $\bar{\pi}_i \in \{\bar{\tau}_c/n_c \pm \frac{1}{m}\}$. This equal division of spots in order to form $\bar{p}$ from $\bar{\mathfrak{p}}$ is feasible by the same Algorithm 7 as in the proof of Theorem 34. Therefore the resulting $\bar{p}, \bar{\pi}$ satisfies

$$\begin{aligned}\|\pi - \bar{\pi}\|_\infty &= \max_c |\tau_c/n_c - \bar{\pi}| \\ &< \frac{1}{n_c}\|\tau - \bar{\tau}\|_\infty + \frac{1}{m} \\ &< \frac{2k}{n_{min} \cdot m} + \frac{1}{m}.\end{aligned}$$

□

5.9 Algorithm Descriptions and Evaluations

5.9.1 Algorithms for calculating optimal panel distributions

In this chapter, we calculate optimal panel distributions across instances with respect to Maximin, NW, and Leximin objectives. To do this, we build on publicly-available code [HG20], which implements the column generation techniques from [Fla+21].

Rounding algorithms

At a high level, the task solved by the Pipage and Beck-Fiala rounding algorithms in Section 5.5 can be thought of as rounding an input panel distribution p to some uniform lottery $\bar{p}$ by rounding the Standard LP described in Section 5.3. However, neither of these rounding methods are used to directly round p; rather, they are used to round a modified version p', which transforms the task from rounding entries of p to multiples of $1/m$ to the task of rounding entries of p' to 0/1. The details of this transformation are described in the proof of Theorem 33 in Section 5.8.

- PIPAGE

  We round p' exactly according to the Pipage Rounding algorithm specified in Gandhi *et al* [Gan+06]. We note that their algorithm is specified for the task of rounding bipartite graphs; we apply their methods by formulating our rounding problem as a star graph, where each of the $|\mathcal{K}|$ vertices surrounding the central vertex corresponds to a feasible panel P. Each edge from the central vertex i to a surrounding vertex P has a weight (which will ultimately be rounded to 0/1) equal to $x_{i,P} = p'_P$, the probability of drawing panel P from the modified version of the initial distribution p'. Gandhi *et al*'s degree preservation property guarantees the satisfaction of our adding up constraint $\|p'\| = \|\bar{p'}\|$.

- BECK-FIALA

  Our Beck-Fiala implementation is identical to the deterministic implementation specified in the proof of Lemma 9, Appendix B.4.1 of [Fla+20a]. For details on the mapping of their setting to ours, see the proof of Theorem 33 in Section 5.8.

Integer Programs

- IP-MAXIMIN

  The below integer program computes a lottery $\bar{p} \in (\mathbb{Z}^+/m)^{|\mathcal{K}|}$, where the variables are y, the lower bound on any marginal probability; $\bar{p}$, the uniform lottery; and $\bar{\pi}$, the implied vector of marginals. The first constraint, along with the objective, result in the maximization of the minimum marginal. The second constraint imposes the relationship between the panel distribution $\bar{p}$ and the marginals $\bar{\pi}$. The third constraint imposes that the resulting panel distribution x will be a uniform lottery. The fourth and fifth constraints impose that $\bar{p}$ is a valid distribution.

$$\begin{aligned}
\text{Maximize } & y \\
\text{s.t. } & \bar{\pi}_i \geqslant y && \forall i \in N \\
& \sum_{\substack{P \in \mathcal{K}, \\ P \ni i}} \bar{p}_P = \bar{\pi}_i && \forall i \in N \\
& m\ \bar{p}_P \in \mathbb{Z}^+ && \forall P \in \mathcal{K} \\
& \sum_{P \in \mathcal{K}} \bar{p}_P = 1 \\
& \bar{p}_P \geqslant 0 && \forall P \in \mathcal{K}
\end{aligned}$$

- IP-NW

  This integer program is essentially the same as IP-Maximin, except that instead of maximizing the lower bound on the marginals, it maximizes the geometric mean of the marginals by equivalently maximizing the sum of their logarithms.

$$\begin{aligned}
\text{Maximize } & \sum_{i \in N} \log(\bar{\pi}_i) \\
\text{s.t. } & \sum_{\substack{P \in \mathcal{K}, \\ P \ni i}} \bar{p}_P = \bar{\pi}_i && \forall i \in N \\
& m\ \bar{p}_P \in \mathbb{Z}^+ && \forall P \in \mathcal{K} \\
& \sum_{P \in \mathcal{K}} \bar{p}_P = 1 \\
& \bar{p}_P \geqslant 0 && \forall P \in \mathcal{K}
\end{aligned}$$

- IP-Marginals

  This IP takes as input some panel distribution p, π to be rounded, and minimizes the largest discrepancy of any resulting $\bar{\pi}_i$ from the corresponding π_i. Again, several of the constraints and variables are common with IP-Maximin.

$$
\begin{aligned}
&\text{Minimize } z \\
&\text{s.t. } |\pi_i - \bar{\pi}_i| \leqslant z && \forall i \in N \\
&\sum_{\substack{P \in \mathcal{K}, \\ P \ni i}} \bar{p}_P = \bar{\pi}_i && \forall i \in N \\
&m\ \bar{p}_P \in \mathbb{Z}^+ && \forall P \in \mathcal{K} \\
&\sum_{P \in \mathcal{K}} \bar{p}_P = 1 \\
&\bar{p}_P \geqslant 0 && \forall P \in \mathcal{K}
\end{aligned}
$$

5.9.2 Implementation Notes and Algorithm Runtimes

Our experiments were implemented in Python and run on a 13-inch MacBook Air (2018) with a 1.6 GHz Intel Core i5 processor.

Runtimes of Pipage, Beck-Fiala, and IP-NW on rounding an unconstrained distribution are given in Section 5.9.2. We optimized IP-NW with Gurobi using its built-in piecewise linear approximation of logarithms (given that IP-NW is nonlinear) with the parameter controlling the error in the piecewise approximation set to FuncPieceError=0.0001. This worked quite well in most instances, getting within $1/m$ of optimal fairness on 10 out of 11 instances.

IP-Maximin and IP-Marginals were run in Gurobi and struggled to converge completely (even after many hours), but showed good performance after a short time. The results in the paper show their solutions after 30 minutes of run-time.

* indicates capped at 7200s (2 hours). Time is measured in seconds. All times given (except those that timed out) represent the average over 3 runs.

Instance	Pipage	Beck-Fiala	IP-NW
sf(a)	1.5	1.6	17.1
sf(b)	1.3	1.3	27.8
sf(c)	1.0	1.1	33.1
sf(d)	2.1	2.3	40.6
sf(e)	17.0	28.3	7245*
cca	4.4	6.4	7207*
hd	1.5	1.7	120.1
mass	0.4	0.4	3.4
nexus	2.8	3.2	21.1
obf	2.3	2.4	22.3
ndem	2.2	2.6	34.8

Table 5.2:: Runtimes for Pipage, Beck-Fiala, and IP-NW

5.9.3 Discussion of Nash Welfare Fairness Preservation

Figure 5.4 presents the algorithm performance results analogous to Section 5.5, but for NW. We see, first that there is some algorithm in every instance that achieves within $0.1/m$ of $NW(p^*)$, where p^* is the NW optimizing unconstrained distribution. This indicates that the cost of transparency to NW in practice is essentially 0. We note that in a few instances, IP-NW, which should theoretically dominate all other algorithms, is outperformed by either Pipage or Beck-Fiala. As we discuss in Section 5.9.2, this is due to small errors in the integer optimization errors.

We find that our theoretical upper bounds on NW loss are less useful than those on the Maximin loss, because they are multiplied by an additional factor of k, while the value of the NW objective falls within a similar range to the Maximin objective. We note, however, that these bounds would be useful for larger m: currently, the maximum possible losses implied by the bounds fall between $191/m = 0.191$ and $5922/m = 5.922$. If we increased m by a factor of 100 to $m = 100,000$ (this would mean drawing 5 lottery balls instead of 3), then our bounds would be nearly tight to optimal in multiple instances (e.g., in "sf(a)", this would yield a loss of 0.008), and would be meaningful in all instances.

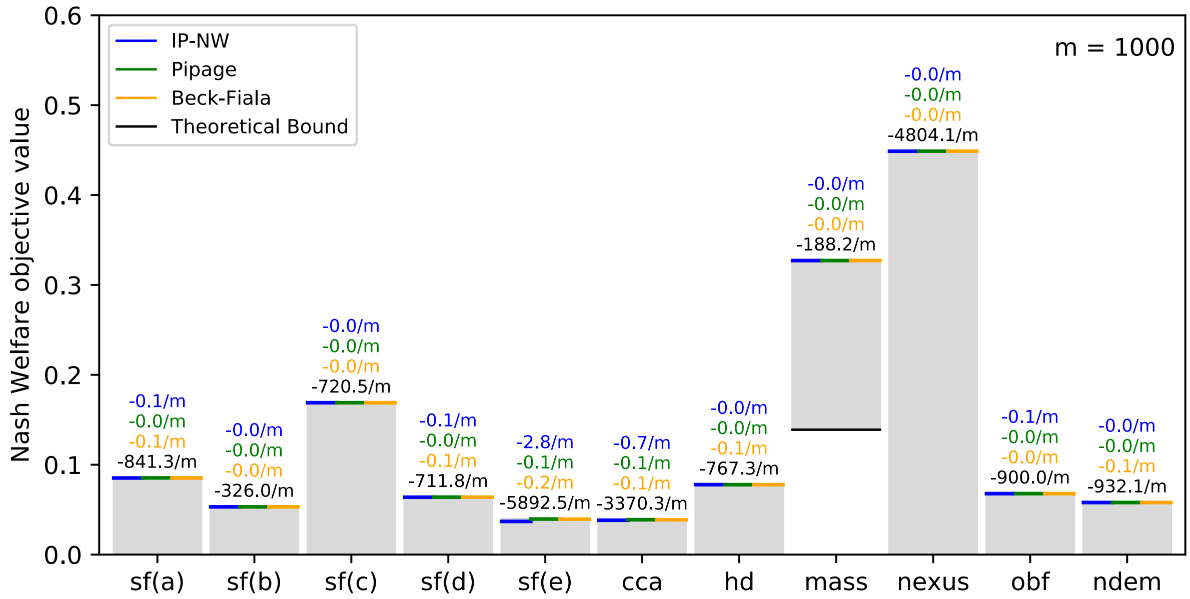


Figure 5.4: $m = 1000$. Shaded regions extend from $NW(p^*)$, the fairness of the optimal unconstrained distribution, down to the minimum fairness implied by the tightest theoretical upper bound in that instance (in all instances but "obf" Theorem 34 is tightest). Each algorithm or bound's loss relative to $NW(p^*)$ is written above in the corresponding color. We show a representative run of Pipage, a randomized algorithm.

5.9.4 Discussion of Leximin Preservation

Figures 5.5 to 5.14 give the corresponding analysis from Figure 5.3 for all other instances. In all instances, the conclusions we draw are essentially the same as those drawn from Figure 5.3: in all instances, all algorithms almost exactly preserve the Leximin-optimal marginals. Our theoretical bounds are meaningful, but we consistently outperform them in practice.

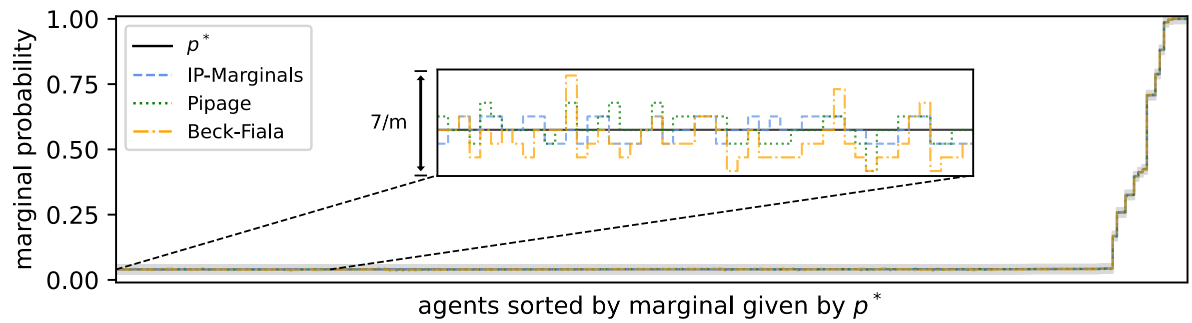


Figure 5.5: sf(b)

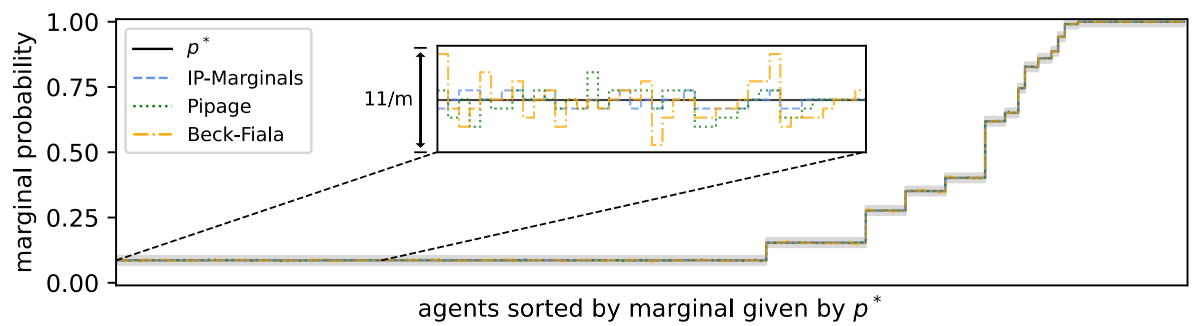


Figure 5.6: sf(c)

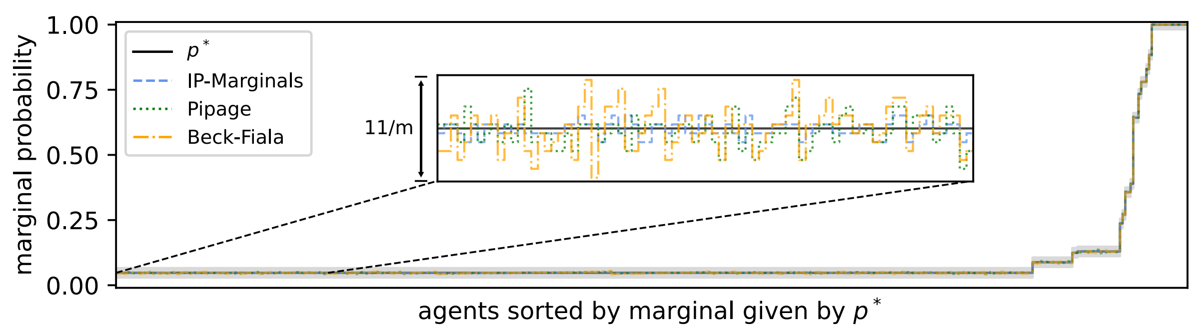


Figure 5.7: sf(d)

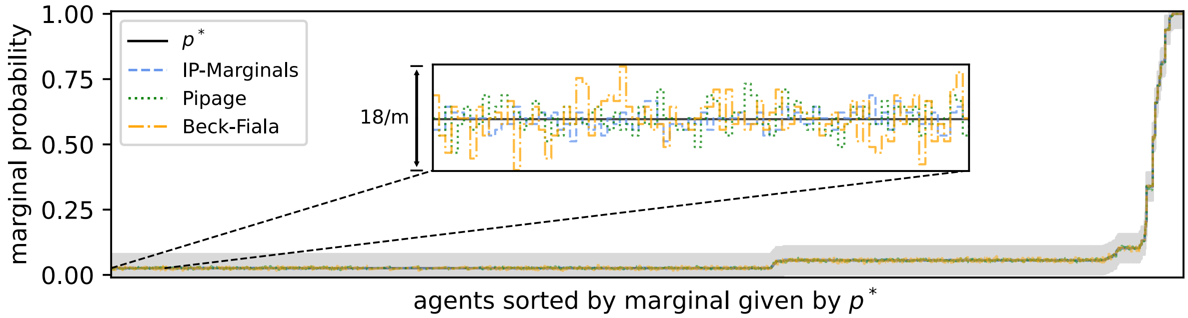


Figure 5.8: sf(e)

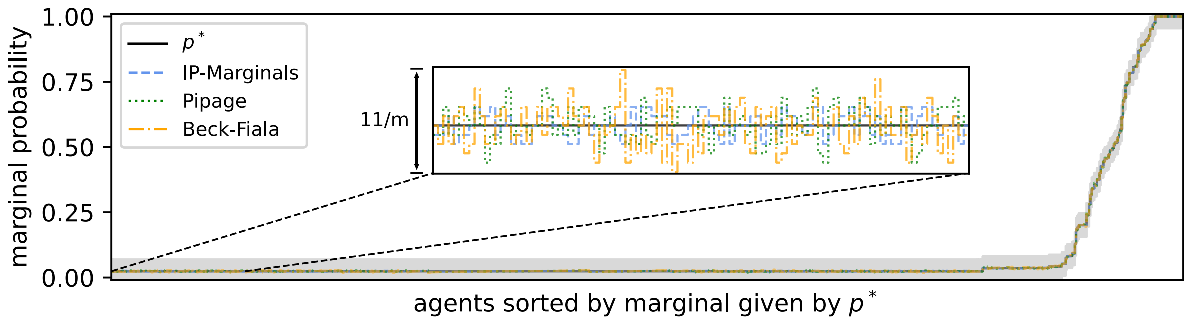


Figure 5.9: cca

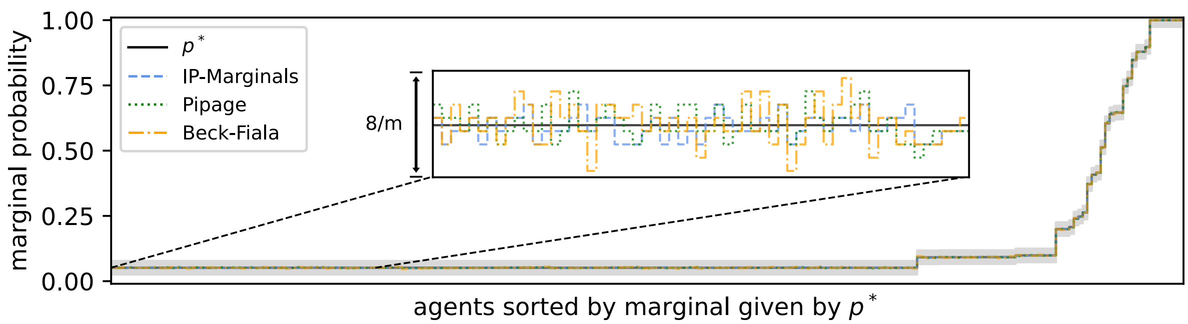


Figure 5.10: hd

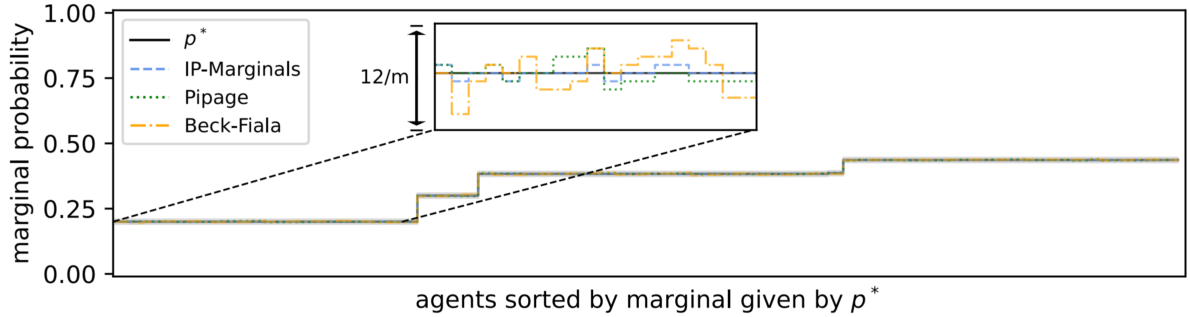


Figure 5.11: mass

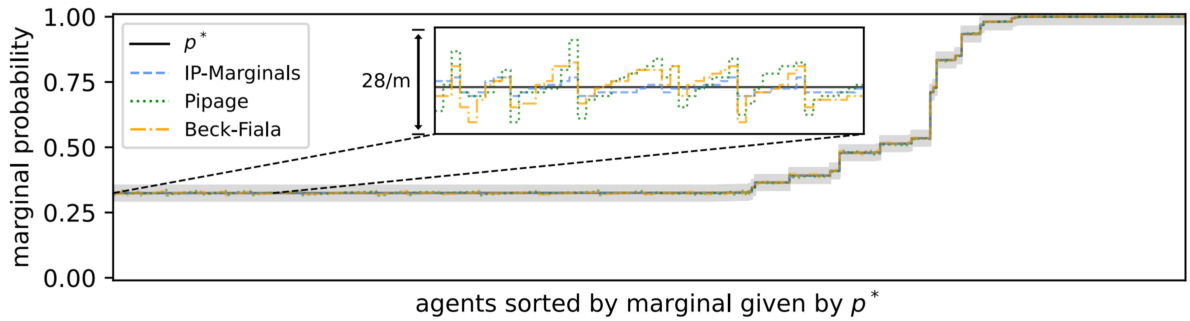


Figure 5.12: nexus

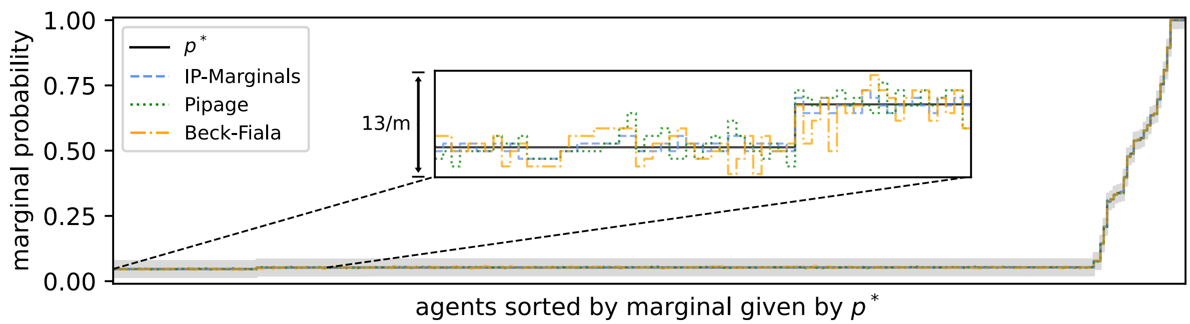


Figure 5.13: obf

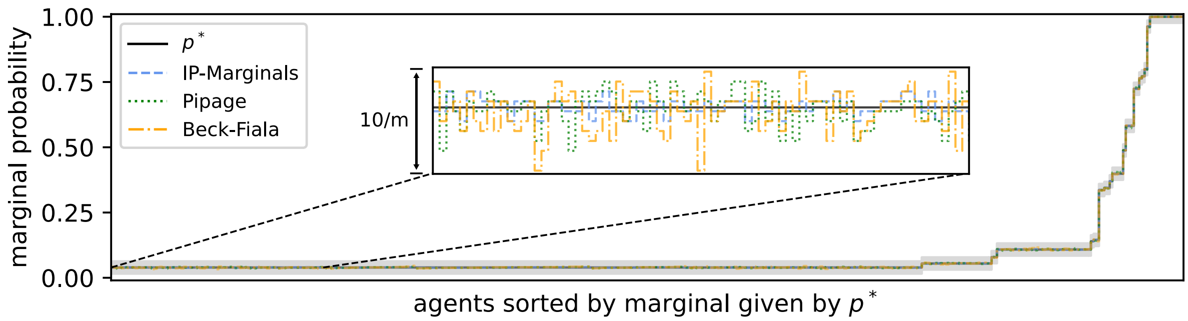


Figure 5.14: ndem

Chapter 6

Sortition via Multi-Objective Optimization

6.1 Introduction

Most people think of democracy as synonymous with elections. But that has not always been the case: from the inception of democracy in ancient Athens until the American and French revolutions, democracy had typically been associated with random selection of representatives [Van16], a paradigm known as *sortition*.

These days, sortition is mainly practiced in the form of *citizens' assemblies* — randomly selected groups of people who deliberate on central questions, with the goal of generating recommendations and informing policy. The impact and prevalence of citizens' assemblies around the world have motivated computational work on how to fairly and transparently select assembly members [Fla+21; Fla+20a; FKP21]. But there are signs that sortition is becoming even more widely accepted, including its recent institutionalization in Belgium, where permanent sortition-based bodies are now working alongside the parliaments of the German-speaking region and the Brussels region. In light of this progress, it may only be a matter of time until one of the many blueprints for sortition-based democracy [GW19] is implemented at the level of an entire country.

The excitement about sortition is driven by several appealing qualities, which are seen as providing solutions to some of the problems plaguing electoral democracy. We briefly present two of them in the context of *uniform selection*, which selects a uniformly random panel and is considered to be the ideal sortition method [Eng89].

- *Descriptive representation*: A panel selected uniformly at random is likely to reflect the composition of the population from which it was drawn. Representation lends *legitimacy* to the process [Par11; Fis18], as individuals are able to identify some panelists who are similar to themselves.

- *Fairness*: Under uniform selection, each individual has an equal chance to participate. Political theorists have argued that this quality realizes philosophical ideals like equality of opportunity and allocative justice [Sto11].

By any reasonable measure of the fairness of selection probabilities — e.g., the minimum selection probability of any individual [Fla+21] — uniform selection achieves perfect fairness, as selection probabilities are equalized. We ask: *Is uniform selection also representative in a rigorous sense?* If we had an analogous measure of representation, we would be able to explore whether this is the case. But quantifying representation poses a conceptual challenge.

6.1.1 Our approach

We address this challenge by assuming that there exists a *representation metric* on individuals, which measures to what degree one individual represents another (smaller distance means better representation). Readers familiar with the algorithmic fairness literature will no doubt make the connection to the *similarity metric* of Dwork et al. [Dwo+12], which has been criticized on the grounds that it is difficult to explicitly construct [Bal+19]; a major obstacle is that the question of whether certain features should be used to determine similarity is domain-specific and tied to legal interpretation. By contrast, the prospect of constructing a representation metric is perhaps more tractable for practitioners. This is because the organizers of a citizens assembly (say) can choose it to be a function of a common set of features which are already routinely used by practitioners for the purpose of guaranteeing the representation of selection rules. Moreover, some of our main results, which pertain to uniform selection, are fully independent of the metric — for these results it suffices to suppose that such a metric *exists*.

Note, however, that a distance metric (on individuals) does not directly define a measure of the degree to which an individual is represented by a panel. Following recent work by Caragiannis, Shah, and Voudouris [CSV22], we assume that the *cost* of a panel for an individual is determined by the q-th closest member of the panel, and our results are parameterized by q.

We can now define representation by taking a page from the literature on *distortion* in social choice [Ans+21a]. Specifically, for a given selection algorithm, we measure its representation via the ratio between the social cost (sum of costs) of the optimal panel and that of the panel chosen by the algorithm, *in the worst case over underlying representation metrics.*

6.1.2 Our results

Returning to the question of whether uniform selection is also representative, and, more generally, the eponymous question of whether sortition is both representative and fair, our answer is that "it depends" — on the value of q.

When $k/2 < q \leqslant k - \Omega(k)$, where k is the size of the panel, we show that uniform selection (which is perfectly fair) achieves constant representation. Qualitatively, we view this as providing positive answers to our questions in the regime of $q > k/2$. Note that this regime has a natural interpretation: each individual wants a majority of the panel to be representative of themselves. This is especially justifiable when the panel makes decisions or recommendations through voting, which is often the case in citizens' assemblies.

By contrast, for the regime of $q \leqslant k/2$, we prove that any selection algorithm that chooses each individual with probability somewhat higher than q/n, where n is the number of individuals, must in the worst case have representation of precisely 0. This result clearly applies to uniform selection, where the minimum selection probability is k/n, and it motivates us to consider weaker fairness guarantees. We design an algorithm, RandomReplace, which selects each individual with probability at least q/n and has a nontrivial representation guarantee of $1/(q+1)$ for any value of q.

Finally, we evaluate the average-case representation of uniform selection and RandomReplace on inferred metrics derived from two demographic datasets. For $q > k/2$, our experiments show that uniform selection achieves representation consistently greater than 55%. Even for $q \leqslant k/2$, in contrast to our worst-case result, uniform selection achieves good average-case representation. These results suggest that, in practice, uniform selection may be a good choice for all q-cost distance functions. While RandomReplace outperforms uniform selection across the board in terms of representation (at the expense of fairness), its advantage is small for values of q around $k/2$ and close to k, which means it does not offer a good representation-fairness tradeoff in those regimes. Lastly, for both algorithms, we observe a spike in representation at $q = k/2$, which demonstrates that the chasm between provable guarantees for $q \leqslant k/2$ versus $q > k/2$ is not merely a theoretical curiosity.

6.1.3 Related work

The design of practical, fair and transparent algorithms for selecting citizens' assemblies was explored in several previous papers [Fla+21; Fla+20a; FKP21] — two of which appeared in previous NeurIPS conferences. Assemblies are required to be representative of the population with respect to features like gender, age, ethnicity, education and geography. This is generally done by setting quotas on individual features; for example, a panel of 100 people might be required to include at least 48 men, at least 48 women, and at least two people who identify as non-binary. The challenge is that an assembly is selected from a pool of volunteers, and this pool is typically unrepresentative of the population due to self-selection bias. Uniformly random selection from the pool, therefore, would likely itself result in an unrepresentative panel. Instead, the primary selection algorithm advocated by Flanigan et al. [Fla+21] computes a distribution over quota-compliant panels that (roughly speaking) maximizes the minimum selection probability of any volunteer, thereby maximizing fairness subject to these hard demographic constraints. By contrast, like Benadè et al. [BGP19] and the political theory literature, we take a longer-term view: We are interested in random selection directly from the population, which is a hallmark of some plans for sortition-based democracy [GW19]. In addition, we take a fundamentally different, and arguably more nuanced, view of representation. Indeed, our framework can theoretically accommodate considerations of intersectionality (to what degree is a rural, college-educated man represented by a rural, college-educated woman?) and could also capture a more holistic analysis of the composition of the panel (to what degree is a rural, college-educated man represented by a panel that includes 10 rural, college-educated men and 90 urban women with no college education?). Lang and Skowron [LS18], Celis, Huang, and Vishnoi [CHV18], and Do et al. [Do+21] also study the problem of creating committees or assemblies in which the features of the participants satisfy some desired quotas, but in the absence of an underlying representation metric and any fairness constraints.

Our approach to evaluating representation through a metric is rooted in spatial theories of voting from political theory [EH84; Arr90]. The idea of measuring how poorly a panel represents an individual by the distance of the q-th closest panel member to the individual was introduced by Caragiannis, Shah, and Voudouris [CSV22] in the context of committee elections. Finally, aiming to minimize the total misrepresentation across all people and comparing that to the optimal panel makes our representation measure the inverse of *distortion* in voting theory [PR06; Ans+18]. In voting, it is assumed that we only have partial access to the metric in the form of voters' ranked preferences over the candidates, which are induced by distance comparisons. In contrast, our results for uniform selection use no knowledge of the underlying metric, while our other algorithms assume complete access. When selecting a single candidate as the winner, it is known that the best distortion achievable by deterministic selection is 3 (which maps to 1/3 representation in our formulation) [GHS20]. Our setting is closer to committee selection, where a committee of k candidates is selected. Here, Caragiannis, Shah, and Voudouris [CSV22] show that when each voter measures her distance to the q-th closest committee member, there is a trichotomy: the best possible distortion is infinite when $q \leqslant k/3$, linear in the number of voters when $q \in (k/3, k/2]$, and 3 for deterministic selection when $q > k/2$. Sortition may be viewed as a special case of randomized committee elections, in which the set of candidates is the same as the set of voters. Hence, all the positive results from Caragiannis, Shah, and Voudouris [CSV22] carry over in the absence of any fairness constraints. However, our results show that when (perfect) fairness in selection probabilities is sought in conjunction with representation, the distortion becomes infinite (zero representation) for all $q \leqslant k/2$ but constant distortion can still be achieved for $q > k/2$. The idea of the set of voters acting as the set of candidates was explored by Cheng, Dughmi, and Kempe [CDK17; CDK18]. However, they model infinitely many voters using a continuous distribution over the metric space.

6.2 Preliminaries

For all $t \in \mathbb{N}$, define $[t] = \{1, \ldots, t\}$. Let $N = [n]$ be the set that indexes the underlying population. A *panel* P is a subset of the population. Let $\mathcal{S}_k(N)$ denote the set of all subsets of N of size k. (We omit N when it is clear form the context.) The population lies in a metric space with distance d, which we view as the *representation metric* discussed earlier. For each $i, j \in N$, $d(i, j)$ denotes the distance between i and j; d is a metric if the following properties are satisfied: (a) $d(i,j) \geqslant 0$, and $d(i,j) = 0$ if and only if $i = j$, (b) $d(i,j) = d(j,i)$, and (c) for each $i, j, \ell \in N$, $d(i,\ell) + d(\ell,j) \geqslant d(i,j)$. The last property is known as the triangle inequality. An instance of our problem is given by the underlying population along with distances as defined by d; hereinafter, we simply denote such an instance by d.

Given a panel P of size k and a positive integer $q \in [k]$, the *q-cost* of individual i for P, denoted by $c_q(i, P; d)$, is equal to the distance of i from her q-th closest representative in P. Note that for $q = 1$, we have $c_1(i, P; d) = \min_{j \in P} d(i,j)$ and for $q = k$ we get $c_k(i, P; d) = \max_{j \in P} d(i,j)$. Let $\mathsf{top}_q(i, P; d)$ be the set of q closest members of P to i (with ties broken arbitrarily). The *q-social cost* of panel P is given by $\mathrm{SC}_q(P; d) = \sum_{i \in N} c_q(i, P; d)$, i.e., the sum of the q-costs over all individuals. (Observe that if $q = 1$ then $\mathrm{SC}_q(P; d)$ is the standard k-medians clustering objective evaluated for centers P.) We omit d from the notation when it is clear from the context.

In this setting, a *selection algorithm* $\mathcal{A}_{k,q}$ parameterized by k and q takes as input the metric d and outputs a distribution over all panels of size k. We are especially interested in the *uniform selection* algorithm, denoted by $\mathcal{U}_k$, that always outputs a uniform distribution over $\mathcal{S}_k$, regardless of q. In other words, it does not take into account the underlying metric space or q, but instead outputs a committee of size k chosen uniformly at random. We now formally introduce the fairness and representation of a selection algorithm.

6.2.1 Fairness

One appealing property of uniform selection is that each individual is selected to be part of the panel with probability exactly equal to k/n, i.e., $\Pr[i \in \mathcal{U}_k] = \frac{k}{n}$. In particular, all individuals have an equal chance of being chosen. We call this property *perfect fairness*; in general an algorithm $\mathcal{A}_{k,q}$ provides perfect fairness when for each instance d, it ensures that $\min_{i \in N} \Pr[i \in \mathcal{A}_{k,q}] = k/n$.

When perfect fairness is too restrictive, we relax this constraint by allowing individuals to be selected with probability less than k/n. In this case, the fairness of an algorithm $\mathcal{A}_{k,q}$ is the worst-case ratio of the minimum probability of an individual to be selected by the algorithm and the ideal selection probability of k/n. Formally,

$$\mathsf{fairness}_q(\mathcal{A}_{k,q}) = \inf_d \frac{\min_{i\in N} \Pr[i \in \mathcal{A}_{k,q}(d)]}{k/n}.$$

6.2.2 Representation

The other key property we would like to measure is representation. To do so, we consider a panel to be a good representative of the whole population when its q-social cost is not much larger than the best possible. In other words, a selection algorithm $\mathcal{A}_{k,q}$ that outputs a distribution over the different committees of size k provides good representation when the expected q-social cost of the panel is similarly small. More formally, we define the *representation* of a selection algorithm $\mathcal{A}_{k,q}$ as the worst-case ratio of the minimum possible social q-cost of any panel and the expected q-social cost of the panel chosen by $\mathcal{A}_{k,q}$ over all possible instances; i.e.,

$$\mathsf{repr}_q(\mathcal{A}_{k,q}) = \inf_d \frac{\min_{P'\in\mathcal{S}_k(N)} \mathrm{SC}_q(P';d)}{\mathbb{E}[\mathrm{SC}_q(\mathcal{A}_{k,q}(d))]}.$$

6.3 Representation with Perfect Fairness for Large q

We begin by considering the case that $q > k/2$. We show that in this case uniform selection is asymptotically optimal with respect to representation among all selection algorithms that are perfectly fair. Moreover, the representation of uniform selection is constant for any $q = c \cdot k$ for $1/2 < c < 1$.

Theorem 45. *For $q > k/2$, uniform selection satisfies* $\mathsf{repr}_q(\mathcal{U}_k) \geqslant \frac{1}{2} \cdot \frac{k-q+1}{k}$.

A crucial property that we exploit in this section is that the q-costs of the individuals satisfy the triangle inequality when $q > k/2$. This observation was first made by Caragiannis et al. [CSV22]; we present the lemma below for completeness.

Lemma 18. *For $q > k/2$, individuals $i, j \in N$, and a panel P, $c_q(i, P; d) + c_q(j, P; d) \geqslant d(i, j)$.*

Proof. Let $T_i = \mathsf{top}_q(i, P)$ and $T_j = \mathsf{top}_q(i, P)$ be the q closest neighbors of i and j, respectively, in the panel P. As $|T_i| = |T_j| > k/2$, there exists an individual $k \in T_i \cap T_j$. Therefore,

$$d(i, j) \leqslant d(i, k) + d(k, j) \leqslant c_q(i, P) + c_q(j, P). \qquad \square$$

We use this observation to lower bound the social cost of the optimal committee.

Lemma 19. *For $q > k/2$, the q-social cost of the optimal panel P^* is at least*

$$\mathrm{SC}_q(P^*; d) \geqslant \frac{1}{2(n-1)} \cdot \sum_{i \in N} \sum_{j \in N \setminus \{i\}} d(i, j).$$

Proof. By applying Lemma 18 for all pairs of individuals (i, j), we get

$$\sum_{i \in N} \sum_{j \in N \setminus \{i\}} \Big(c_q(i, P^*) + c_q(j, P^*) \Big) \geqslant \sum_{i \in N} \sum_{j \in N \setminus \{i\}} d(i, j).$$

The q-cost of each person appears exactly $2(n-1)$ times on the left hand side. Thus,

$$\mathrm{SC}_q(P^*; d) = \sum_{i \in N} c_q(i, P^*) \geqslant \frac{1}{2(n-1)} \sum_{i \in N} \sum_{j \in N \setminus \{i\}} d(i, j). \qquad \square$$

We are now ready to prove the theorem.

Proof of Theorem 45. For any committee P of size k,

$$\begin{aligned} c_q(i, P) = \min_{q' \in \{q, \ldots, k\}} c_{q'}(i, P) &\leqslant \frac{1}{k - q + 1} \sum_{q' \in [q, k]} c_{q'}(i, P) \\ &\leqslant \frac{1}{k - q + 1} \sum_{q' \in [1, k]} c_{q'}(i, P) = \frac{1}{k - q + 1} \sum_{j \in P} d(i, j), \end{aligned}$$

where in the first inequality the minimum is upper bounded by the average. Therefore, the expected social cost of uniform selection is at most

$$\begin{aligned}
\mathbb{E}[\mathrm{SC}_q(\mathcal{U}_k(N))] &= \sum_{i \in N} \mathbb{E}_{P \sim \mathcal{U}_k}\left[c_q(i, P)\right] \\
&\leqslant \frac{1}{k-q+1} \sum_{i \in N} \mathbb{E}_{P \sim \mathcal{U}_k}\Big[\sum_{j \in P} d(i,j)\Big] \\
&= \frac{1}{k-q+1} \sum_{i \in N} \sum_{j \in N \setminus \{i\}} d(i,j) \cdot \Pr_{P \sim \mathcal{U}_k}[j \in P] \\
&= \frac{1}{k-q+1} \sum_{i \in N} \sum_{j \in N \setminus \{i\}} d(i,j) \cdot \frac{k}{n}.
\end{aligned}$$

By Lemma 19 and the upper bound shown above, we have

$$\mathsf{repr}_q(\mathcal{U}_k) \geqslant \frac{\frac{1}{2(n-1)} \sum_{i \in N} \sum_{j \in N \setminus \{i\}} d(i,j)}{\frac{1}{k-q+1} \cdot \frac{k}{n} \cdot \sum_{i \in N} \sum_{j \in N \setminus \{i\}} d(i,j)} = \frac{1}{2} \cdot \frac{n}{n-1} \cdot \frac{k-q+1}{k} \geqslant \frac{1}{2} \cdot \frac{k-q+1}{k}. \qquad \square$$

In the proof above, the only property of uniform selection we use is that the marginal probabilities are equal to k/n. Hence, this lower bound also holds for any perfectly fair selection algorithm.

We next establish an upper bound on the representation of any perfectly fair selection algorithm. It shows that the lower bound of Theorem 45 is tight up to a factor of 4.

Theorem 46. *For any $q > k/2$, every selection algorithm $\mathcal{A}_{k,q}$ with $\mathsf{fairness}(\mathcal{A}_{k,q}) = 1$ satisfies $\mathsf{repr}_q(\mathcal{A}_{k,q}) \leqslant 2 \cdot \frac{k-q+1}{k+1}$.*

Proof. First, note that if $q = \frac{k+1}{2}$ the statement holds, since $\mathsf{repr}_q(\mathcal{A}_{k,q}) \leqslant 2 \cdot \frac{k-q+1}{k+1} = 1$ which is true for any algorithm. Thus, we assume $q > \frac{k+1}{2}$. Consider an instance with $n = k+1$ individuals where $k-q+1$ individuals are located at 0 and q individuals are at 1, denoted by N_0 and N_1, respectively. Any committee of size k leaves one person out of the committee, and for each individual $i \in N$, this happens with probability of

$$\Pr_{P \sim \mathcal{A}_{k,q}}[i \notin P] = 1 - \Pr_{P \sim \mathcal{A}_{k,q}}[i \in P] = (1 - \tfrac{k}{k+1}) = \tfrac{1}{k+1}.$$

Individuals in N_0 will always have a q-cost of 1, because $k+1-q < \frac{k+1}{2} < q$ individuals are located there. Therefore, $\mathbb{E}_{P \sim \mathcal{A}_{k,q}}[\sum_{i \in N_0} c_q(i, P)] = |N_0|$. For individuals in N_1, their q-cost is

1 if and only if less than q individuals are selected from N_1, i.e., the single person left out of the committee is located at 1. This event happens with probability

$$\Pr_{P\sim\mathcal{A}_{k,q}}[\textstyle\bigcup_{i\in N_1}(i\notin P)] = \sum_{i\in N_1}\Pr_{P\sim\mathcal{A}_{k,q}}[i\notin P] = \frac{|N_1|}{k+1},$$

where the first equality comes from the fact that the events are disjoint (which holds because any committee leaves out exactly one individual). Therefore, $\mathbb{E}_{P\sim\mathcal{A}_{k,q}}[\sum_{i\in N_1} c_q(i,P)] = |N_1|\cdot\frac{|N_1|}{k+1}$, and the expected social cost of any perfectly fair algorithm is

$$\mathbb{E}_{P\sim\mathcal{A}_{k,q}}[\mathrm{SC}_q(P;d)] = |N_0| + |N_1|\cdot\frac{q}{k+1} \geqslant |N_0| + |N_1|\cdot\frac{1}{2}.$$

The optimal committee would leave out a person from N_0 and achieve a social cost of $|N_0|$. Therefore, the representation of any algorithm with perfect fairness is at most

$$\frac{|N_0|}{|N_0|+\frac{1}{2}\cdot|N_1|} \leqslant \frac{|N_0|}{\frac{1}{2}\cdot|N_0|+\frac{1}{2}\cdot|N_1|} = 2\cdot\frac{k-q+1}{k+1}. \qquad \square$$

6.4 Representation with Relaxed Fairness for Small q

In stark contrast to the case of $q > k/2$, uniform selection and, more generally, any perfectly fair selection algorithm, cannot obtain bounded representation when $q \leqslant k/2$. In fact, the following theorem shows that selection algorithms with fairness strictly more than $\frac{q+(k \bmod q)}{k}$ (which itself is upper bounded by $(2q-1)/k$) suffer from unbounded representation. The proof is in Section 6.4.

Theorem 47. *For $q \leqslant k/2$, $\epsilon > 0$, and any selection algorithm $\mathcal{A}_{k,q}$ with $\mathsf{fairness}(\mathcal{A}_{k,q}) \geqslant \frac{q+(k \bmod q)}{k} + \epsilon$, $\mathsf{repr}_q(\mathcal{A}_{k,q})$ is 0.*

Proof. Assume that $n > 2\cdot\max\{\sqrt{kq/\epsilon}, k+1\}$, and let $m = \lfloor k/q\rfloor - 1$. Consider the real line, and suppose there are sets of q individuals at each position in $\{1,2,\ldots,m\}$, denoted by $X_1,\ldots,X_m$, respectively, and the set of remaining $n-mq$ individuals, denoted by X_{m+1}, is at position $m+1$. The optimal panel P^* would have at least q people from each position, i.e.,

$|P^* \cap X_i| \geqslant q$ for all $i \in [m+1]$. The q-cost of each person for P^* is 0 as at least q people are selected from her own position. Hence, $\mathrm{SC}_q(P^*) = 0$.

Turning to the analysis of $\mathcal{A}_{k,q}$, we claim that

$$\mathbb{E}_{P\sim\mathcal{A}_{k,q}}[|X_{m+1} \cap P|] \geqslant q + (k \bmod q) + \epsilon. \tag{6.1}$$

To prove this, note that since each individual is included with a marginal probability of at least $\frac{q+(k \bmod q)}{n} + \epsilon$, we have

$$\begin{aligned}\mathbb{E}_{P\sim\mathcal{A}_{k,q}}[|X_{m+1} \cap P|] &\geqslant \left(\frac{q + (k \bmod q)}{n} + \epsilon\right)(n - mq)\\ &= q + (k \bmod q) - \frac{mq \cdot (q + (k \bmod q))}{n} + (n - mq) \cdot \epsilon\end{aligned}$$

We will show that the right hand side is at least $q + (k \bmod q) + \epsilon$. Because $mq < k$, $q + k \bmod q < 2q$, and $n - mq > n - k \geqslant n/2 + 1$, the right hand side is at least

$$q + (k \bmod q) - \frac{qk}{n/2} + n\epsilon/2 + \epsilon \geqslant q + (k \bmod q) + \epsilon,$$

where the inquality follows from our choice of $n > 2\sqrt{kq/\epsilon}$. This establishes Equation (6.1).

Now, as the panel size is k, it holds that $\sum_{i\in[m+1]} \mathbb{E}_{P\sim\mathcal{A}_{k,q}}[|X_i \cap P|] = k$. By Equation (6.1),

$$\sum_{i\in[m]} \mathbb{E}_{P\sim\mathcal{A}_{k,q}}[|X_i \cap P|] < k - (q + (k \bmod q)) - \epsilon = mq - \epsilon.$$

Therefore, there exists $i \in [m]$ such that $\mathbb{E}_{P\sim\mathcal{A}_{k,q}}[|X_i \cap P|] \leqslant q - \epsilon/q$. Using Markov's inequality,

$$\Pr_{P\sim\mathcal{A}_{k,q}}(|X_i \cap P| \geqslant q) \leqslant \tfrac{q-\epsilon/q}{q} \leqslant 1 - \epsilon.$$

Thus, with probability at least ϵ, less than q people are selected from position i, in which case the q-cost of each person in X_i will be at least 1. Hence, $\mathbb{E}_{P\sim\mathcal{A}_{k,q}}[\mathrm{SC}_q(\mathcal{A}_{k,q})] \geqslant q\epsilon$ while $\mathrm{SC}_q(P^*) = 0$. □

Although bounded representation is not feasible with fairness slightly larger than $\frac{q}{k}$, we design a selection algorithm that, given an α-representative panel, can achieve representation

Algorithm 8 RandomReplace$_q$

Require: Panel P with $\mathsf{repr}_q(P) = \alpha$
Ensure: Randomly selected panel by replacing at most q individuals of P
1: Pick $S \in \mathcal{S}_q$ uniformly at random
2: Set $P_S \leftarrow P$ and $\overline{S} \leftarrow S \setminus P$
3: **for** $i \in \overline{S}$ **do**
4: Pick an arbitrary $j_i \in \mathsf{top}_q(i, P) \setminus S$
5: $P_S \leftarrow P_S \cup \{i\} \setminus \{j_i\}$
6: **return** P_S

of at least $\frac{\alpha}{q+1}$ with fairness of $\frac{q}{k}$. In particular, starting from the optimal panel ($\alpha = 1$), it achieves $\frac{1}{q+1}$ representation. Although finding the optimal panel given the metric space is an NP-hard problem for most metric spaces, computing a constant-factor approximation (i.e., constant α) is feasible in polynomial time [KR13], as we explain in Section 7.5.

Our Algorithm, RandomReplace$_q$, is given in Algorithm 8. It starts from a panel P, randomly selects a panel S of q individuals, and replaces individuals in P with individuals in S as follows. First, individuals in $S \cap P$ remain in the final panel. Then, for $i \in S \setminus P$, swap i with one of its q-closest neighbors in the optimal panel $\mathsf{top}_q(i, P)$ that has not been replaced by the algorithm yet. The next theorem establishes the fairness and representation guarantees of RandomReplace.

Theorem 48. *For any $q \in [k]$ and any panel P with $\mathsf{repr}_q(P) = \alpha$, we have that $\mathsf{repr}_q(\text{RandomReplace}_q(P)) \geqslant \frac{\alpha}{q+1}$ and $\mathsf{fairness}(\text{RandomReplace}_q) \geqslant \frac{q}{k}$.*

Proof. Let $S \subseteq N$ be a set of size q chosen uniformly at random. We denote with P_S the panel that is returned from the algorithm. First, we show that Line 4 of the algorithm is valid. The algorithm reaches this line when it considers $i \in S \setminus P_S$, meaning that i is not included in the panel P_S and therefore is not included in P. Hence, $\mathsf{top}_q(i, P) \setminus S$ cannot be empty since $|\mathsf{top}_q(i, P)| = q$, $|S| = q$ and there exists i in S but not in $\mathsf{top}_q(i, P)$.

We see that every individual in S is included in P_S as in Line 5 the algorithms ensures that each such individual is included and is never excluded afterwards. Hence, as each individual is chosen in S with probability at least q/n, we can see that $\mathsf{fairness}(\textsc{RandomReplace}_q) \geqslant q/k$.

Now, we prove that for any $S \in \mathcal{S}_q$ and any individual $i' \in N$,

$$c_q(i', P_S) \leqslant c_q(i', P) + \max_{i \in S} c_q(i, P) \leqslant c_q(i', P) + \sum_{i \in S} c_q(i, P). \tag{6.2}$$

The second inequality holds as the maximum is at most the sum. Therefore, we focus on the first inequality. If $c_q(i', P_S) \leqslant c_q(i', P)$, then it trivially holds. Otherwise, $c_q(i', P_S) > c_q(i', P)$. In this case, we can show that there exists some $i \in S \setminus P$ such that $d(i', i) \geqslant c_q(i', P_S)$ and $j_i \in \mathsf{top}_q(i', P)$. First, note if for each $i \in S \setminus P^*$, j_i does not belong in $\mathsf{top}_q(i', P)$, then it is not possible that $c_q(i', P_S) > c_q(i', P)$, since $\mathsf{top}_q(i', P) \subseteq P_S$. Next, suppose for contradiction that for every i that was included in P_S when $j_i \in \mathsf{top}_q(i', P)$ was excluded from it, it holds that $d(i', i) < c_q(i', P)$. Then, in P_S there are $|\,\mathsf{top}_q(i', P) \setminus \cup_{i \in S \setminus P}(\{j_i\} \cap \mathsf{top}_q(i', P))|$ individuals that have distance at most $c_q(i', P) < c_q(i', P_S)$ from i' and $|\{i \in S \setminus P : j_i \in \mathsf{top}_q(i', P)\}|$ individuals that have distance less than $c_q(i', P_S)$ from i'. Note that

$$|\cup_{i \in S \setminus P} (\{j_i\} \cap \mathsf{top}_q(i', P))| = |\{i \in S \setminus P : j_i \in \mathsf{top}_q(i', P)\}|,$$

and hence we get that there are at least $|\,\mathsf{top}_q(i', P)| = q$ individuals in P_S with distance strictly less than $c_q(i', P_S)$ from i', which is a contradiction.

From the above observation , we have that

$$c_q(i', P_S) \leqslant d(i', i) \leqslant d(i', j_i) + d(j_i, i) \leqslant c_q(i', P) + c_q(i, P) \leqslant c_q(i', P) + \max_{i \in S} c_q(i, P)$$

where the penultimate inequality holds because $j_i \in \mathsf{top}_q(i', P)$ and $j_i \in \mathsf{top}_q(i, P)$ from the definition of j_i. This proves (6.2).

Summing (6.2) over all $i' \in N$, we have $\mathrm{SC}_q(P_S) \leqslant \mathrm{SC}_q(P) + n \cdot \sum_{i \in S} c_q(i, P)$. Taking the expectation of this equation with respect to S, and using the fact that $\Pr[i \in S] = q/n$, we have

$$\mathbb{E}[\mathrm{SC}_q(P_S)] \leqslant \mathrm{SC}_q(P) + n \cdot \sum_{i \in N} \frac{q}{n} \cdot c_q(i, P) = (q+1) \cdot \mathrm{SC}_q(P),$$

that is, $\frac{1}{q+1} \cdot \mathbb{E}[\mathrm{SC}_q(P_S)] \leqslant \mathrm{SC}_q(P)$. Combined with the assumption that P is α-representative, i.e., $\alpha \cdot \mathrm{SC}_q(P) \leqslant \min_{P' \in S_k(N)} \mathrm{SC}_q(P')$, we get

$$\mathsf{repr}_q(\textsc{RandomReplace}_q(P)) \leqslant \frac{\alpha}{k+1}. \qquad \square$$

The next theorem shows that when $q = \Omega(k)$, $\textsc{RandomReplace}_q$ attains an asymptotically optimal representation-fairness tradeoff.

Theorem 49. *For any $q \leqslant k/2$ such that $k \bmod q = 0$, every selection algorithm $\mathcal{A}_{k,q}$ with $\mathsf{fairness}(\mathcal{A}_{k,q}) \geqslant q/k$ satisfies $\mathsf{repr}_q(\mathcal{A}_{k,q}) \leqslant k/q^2$.*

Proof. Let $m = k/q$. Consider an instance with $n > 2k$ individuals on the real line, where one individual i_0 is located at 0, $\lceil n/m \rceil - 1$ people are at 1, and at least $\lfloor n/m \rfloor$ people are located at each position $j \in \{2, \ldots, m\}$. This way, there are at least $n/m - 1 = (n/k) \cdot q - 1 \geqslant 2q - 1 \geqslant q$ individuals located at each $j \in [m]$.

Any optimal panel would include q individuals from each position $j \in [m]$, which results in a q-cost of 1 for i_0 and a q-cost of 0 for the rest. Hence, $\mathrm{SC}_q(P^*) = 1$. However, any selection algorithm with fairness of at least q/k selects i_0 with probability of at least q/n. When i_0 is selected, there must exist a group $j \in [m]$ from which the algorithm selects at most $q - 1$ people, incurring a q-cost of 1 for at least n/m people (person i_0 and at least $n/m - 1$ people at position j). Hence,

$$\mathbb{E}[\mathrm{SC}_q(\mathcal{A}_{k,q})] \geqslant \frac{q}{n} \cdot \frac{nq}{k} = \frac{q^2}{k},$$

which completes the proof. $\square$

From the above theorem, it follows that Algorithm 8 achieves the highest possible fairness of q/k subject to positive representation.

Algorithm 9 $\textsc{RandomDictator}_{k,q}$

1: Pick $i \in N$ uniformly at random
2: Pick $S \in \mathcal{S}_{k-q}$ uniformly at random
3: **return** $i \cup \mathsf{top}_q(i, P; d) \cup S$

6.5 Tradeoffs between Representation and Fairness

In the previous section, we show that by relaxing the perfect fairness, in the case that $q \leqslant k/2$, we are able to improve the representation dramatically . In general, an interesting challenge is to measure the trade-off between representation and fairness, and in this section, we take some initial steps in this direction for the entire range of values of q.

We start with the case of $q > k/2$ and show that a simple algorithm, which is a variant of the natural *random dictatorship* rule, provides constant representation by sacrificing some quantity of perfect fairness. The high-level idea is that the outcome largely depends on only one individual from the underlying population. Specifically, the algorithm $\textsc{RandomDictator}_{k,q}$, presented as Algorithm 9, works as follows: Given an instance d, it chooses an individual i from the underlying population and returns the panel $P = i \cup \mathsf{top}_q(i, P; d) \cup S$ where S is a panel of size $k - q$ chosen uniformly at random.

Theorem 50. *For any $q > k/2$, it holds that*

$$\mathsf{repr}_q(\textsc{RandomDictator}_{k,q}) \geqslant \frac{1}{3} \quad \textit{and} \quad \mathsf{fairness}(\textsc{RandomReplace}_{k,q}) = \frac{k - q + 1}{k}.$$

Proof. We start by proving the fairness guarantee of the algorithm. Note that each an individual i is included in the panel P that is returned by $\textsc{RandomDictator}_{k,q}$ either if i is selected at the first step which happens with probability $1/n$ or if i is selected at the second step which happens with probability $(k - q)/n$. Hence, we get that $\mathsf{fairness}(\textsc{RandomReplace}_{k,q}) = (k - q + 1)/k$.

The representation guarantee follows from Corollary 2 in [CSV22]. □

Now, we turn our attention to the case that $q \leqslant k/2$. In Section 6.4, we introduce $\textsc{RandomReplace}_{k,q}$ with fairness q/k and representation $1/(q + 1)$. Actually, note that if we replace q with any $r \in [q]$, we can show that the algorithm provides representation of at least $1/(r+1)$ with fairness r/k, Essentially, $\textsc{RandomReplace}_{k,r}$ for any $r \in [q]$ in Line 1 of Algorithm 8 chooses a subset S of the underlying population with size r instead of q uniformly at random.

Proposition 1. *For any $q \in [k]$, it holds that*

$$\mathsf{repr}_q(\textsc{RandomReplace}_{r,q}) \geqslant \frac{1}{r+1} \quad \textit{and} \quad \mathsf{fairness}(\textsc{RandomReplace}_{r,q}) = r/k.$$

With identical arguments as in the proof of Theorem 48, the above proposition follows.

6.6 Experiments

Next, to complement the worst-case analysis of the selection algorithms we have considered so far, we conduct an empirical comparison between the average-case q-social cost and representation (repr_q) of the different selection methods.

Datasets

Data on the metric-structure preferences of groups in their full richness are difficult to come by, but it is reasonable to expect that the extent to which individuals feel well-represented by one another is at least partly a function of their relative characteristics along some observable features. We therefore begin with two datasets that express the characteristics of populations along a range of features, and randomly construct synthetic metric preferences from these feature signatures.

`Adult` *dataset.* Our first source of demographic data is the UCI `Adult` dataset, which was derived from the 1994 Current Population Survey of the US Census Bureau, and is made available by the UCI Machine Learning Repository under a CC BY 4.0 license [DG17]. It contains a range of demographic variables principally related to employment. Our experiments do not require `Adult` to be representative of any actual population, nor should this an assumption be made lightly [Din+21]. For `Adult` we choose the features `workclass`, `education`, `marital status` and `sex`. `Adult` contains $n = 30162$ individuals with values for each feature, who may be viewed as a distribution over the 721 unique feature vectors which they collectively hold.

`ESS` *dataset.* Our second source of demographic data is the *European Social Survey* (`ESS`) [Rep21], which is made available by the Norwegian Centre for Research Data under a CC BY 4.0 license. We use the `ESS` Round 9 (2018) data, which consists of 46,276 people in 27 countries, and contains ∼1450 features regarding socioeconomic demographic, political beliefs, geographical region, house-hold composition, personal values, media use and trust, etc. Most of the features are country-specific, which leaves roughly 250 features available per country, while each country has between 781 and 2745 entries (with a mean of 1713). Each entry is assigned an analysis weight which is aimed to correct the differential selection probabilities. In contrast to our experiments with `Adult`, we use all of the available features available in `ESS`. We report the experiment results based on the data of the United Kingdom (2204 entries). Similar results were obtained using the data from other countries.

Metric construction

We choose some set of features along which to evaluate the individuals in the population. These are the features which will inform our metric. For each feature s, F^s is the set of possible values that this feature can take. Each individual i is then represented as a vector of feature values, where f_i^s is the value that i has for feature s. For *categorical* features (e.g., marital status, sex), we define $d(i, j; s) = 1 - \mathbb{1}\{f_i^s = f_j^s\}$. For *range* (e.g., income) features we define $d(i, j; s) = |f_i^s - f_j^s|/(\max_{i',j'} |f_{i'}^s - f_{j'}^s|)$, where the normalization makes $d(i, j; s) \in [0, 1]$. Then, for each feature we sample a weight $w_t \sim U[0, 1]$ uniformly at random from the interval $[0, 1]$. Finally the distance between individuals i and j and our metric d is defined to be $d(i, j) := \sum_s w_s \cdot d(i, j; s)$.

For the empirical evaluation of our selection algorithms on these randomly generated metrics, we suppose that this is in fact the distribution for a population large enough that there are at least k individuals with any given feature vector. As these metrics represent populations, the q-social costs in Figures 6.1 and 6.2 are normalized by population size. Figures 6.1 to 6.3 depict data averaged over 100 random metrics constructed in this manner, and the error bars show the standard error of the mean.

Proxy for the optimum

In evaluating UniformSelection and RandomReplace, we use a proxy for the optimal q-social cost, since n and k are too large to support finding $SC_q(P^*, d)$ exactly. This selection algorithm OptProxy (indicated as 'opt_apx' in legends) is an implementation of the fault-tolerant metric k-medians algorithm of Kumar et al. [KR13], which guarantees a constant-factor approximation to $SC_q(P^*, d)$. This algorithm uses a constant-factor metric k-medians algorithm as a primitive; we implement the local search algorithm of Arya et al. [Ary+04] with single swaps. More specifically, OptProxy works as follows. It finds the (approximately) optimal $\lfloor k/q \rfloor$-median solution and picks q individuals from each of these $\lfloor k/q \rfloor$ positions, and then selects the remaining $k \bmod q$ people uniformly at random.

To evaluate the fidelity of OptProxy we compare it with $\mathrm{SC}_q(P^*, d)$ for 100 metrics d constructed by drawing 30 randomly chosen feature vectors from the supports of `Adult` instances described above. For panels of size $k = 10$ and all values of q we find that OptProxy recovers $\mathrm{SC}_q(P^*, d)$ exactly (Figure 6.2). The fact that the q-social cost attained by OptProxy for fixed q and $q \in (k/2, k]$ appears constant, on the other hand, is not universally true of $\mathrm{SC}_q(P^*, d)$. These plateaus are due to the way OptProxy selects panels by returning q copies of the optimal 1-median location. Since P_q^* need not be of this form, it is interesting that it exhibits this same step-like behavior in Figure 6.2.

Experimental results

In Figure 6.1a, we see UniformSelection behaving as expected. For a fixed k, the q-social cost of a uniformly random panel is reliably higher for larger q, and for a fixed q, it decays smoothly as we increase the panel size k. Under the same conditions, OptProxy behaves similarly (Figure 6.1b) but the decay is not as smooth due to the reasons explained above. Here, all of our approximately optimal panels P_q start at the same value of $\mathrm{SC}_q(P_q, d)$ when $q = k$ and decay as q remains fixed and the size of the panel k increases.[1]

Figure 6.3 shows the ratio between the q-social cost obtained by different selection algorithms and the q-social cost of OptProxy, which is an approximation of their representation, averaged over 100 randomly generated metrics. Specifically, we evaluate UniformSelection and RandomReplace$_r$ (indicated as 'unif_select' and 'rand_replace_r' in legends, respectively), the latter being a generalization of RandomReplace$_q$ that replaces r individuals instead of q in the input panel and achieves representation at least $\alpha/(r+1)$ (when the input panel has representation α) and fairness at least r/k. We compare these algorithms for a range of r and q, on the `Adult` dataset (Figure 6.3a) with panels of size $k = 20$ and on the `ESS` (Figure 6.3b) dataset with panels of size $k = 40$.

Representation of UniformSelection*.* As expected, the representation of UniformSelection is the lowest for all k and q. However, it still achieves consistently good representation across the board, with representation dipping below 50% only on the `Adult` dataset with q close to k.

We find it especially interesting that, on both datasets, UniformSelection exhibits good representation in the regime of $q \leqslant k/2$, as this contrasts with our worst-case result. We

[1] This first property is to be expected: when $q = k$ and points in the metric may appear on the panel multiple times (as is the case with these distributions), an optimal panel P_q^* consists of k copies of a (carefully chosen) single point in the metric. For these panels P_q, OptProxy is simply finding that point.

conclude that, in real-world datasets, UNIFORMSELECTION may be quite representative even with respect to the q-social cost measure when $q \leqslant k/2$. Overall, we view these results a providing empirical support for the use of UNIFORMSELECTION in practice.

Representation of RANDOMREPLACE*.* Figure 6.3 shows that RANDOMREPLACE$_r$ achieves significant improvements in representation compared to UNIFORMSELECTION for small values of r — at the cost of fairness, of course. This tradeoff may be justified for some values of q, but the advantage of RANDOMREPLACE is much smaller around $q = k/2$ and when q approaches k, so for these regimes RANDOMREPLACE should not be in the running.

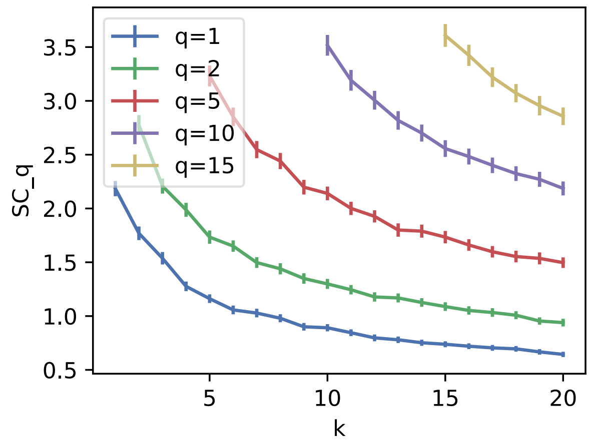
Shift at $q = k/2$*.* In both Figures 6.3a and 6.3b, there is a sudden improvement in the representation of UNIFORMSELECTION at $q = k/2$. This provides empirical support for the peculiar transition in the q-social cost occurring $q = k/2$, which is predicted by the theory and is the reason we separated Sections 6.3 and 6.4.

6.7 Discussion

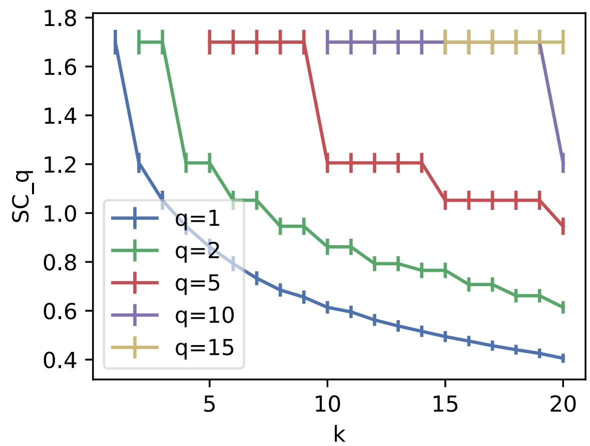
Our results in Section 6.4 show that in some cases, relaxing fairness requirements allows improving representation dramatically. More generally, it is interesting to understand the tradeoff between representation and fairness, and to chart the Pareto frontier. In Section 6.5, we take some first steps in this direction. One observation is the aforementioned generalization of RANDOMREPLACE that replaces r individuals instead of q. We also show that a random-dictatorship-like algorithm gives nontrivial fairness and representation guarantees in the regime of $q > k/2$. However, there remain several open questions: for example, when $q \leqslant k/2$, what level of fairness can be achieved if we seek constant representation?

We focused our attention on the q-cost formulation of Caragiannis, Shah, and Voudouris [CSV22], in which each individual measures their distance to the q-th closest panel member. One can analyze the representation-fairness tradeoff with other cost functions. For example, what if different individuals use different values of q? Another appealing choice is when each individual measures the *average* distance to all panel members?

More broadly, one can use measures of representation other than the *total cost* to all individuals. For example, one may wish to use a selection algorithm that minimizes the *maximum cost* to any individual, or strikes a balance between maximum cost and total cost. We hope that answers to some of these questions will lead to a better understanding of the strengths of sortition, and to new ways of realizing this democratic paradigm.



(a) UniformSelection



(b) OptProxy

Figure 6.1: The q-social cost of UniformSelection and OptProxy for $k \in [1, 20]$ and a selection of q, based on `Adult`.

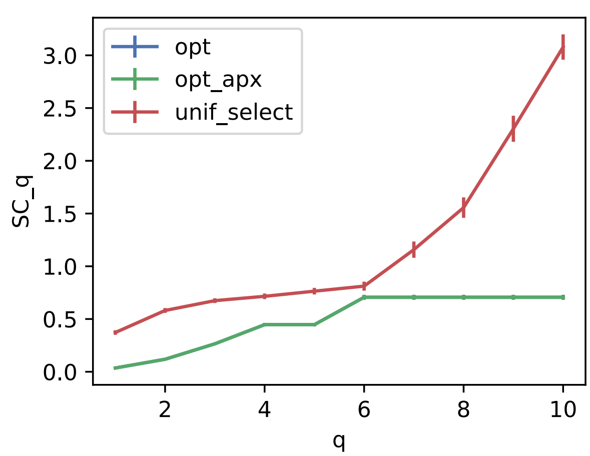
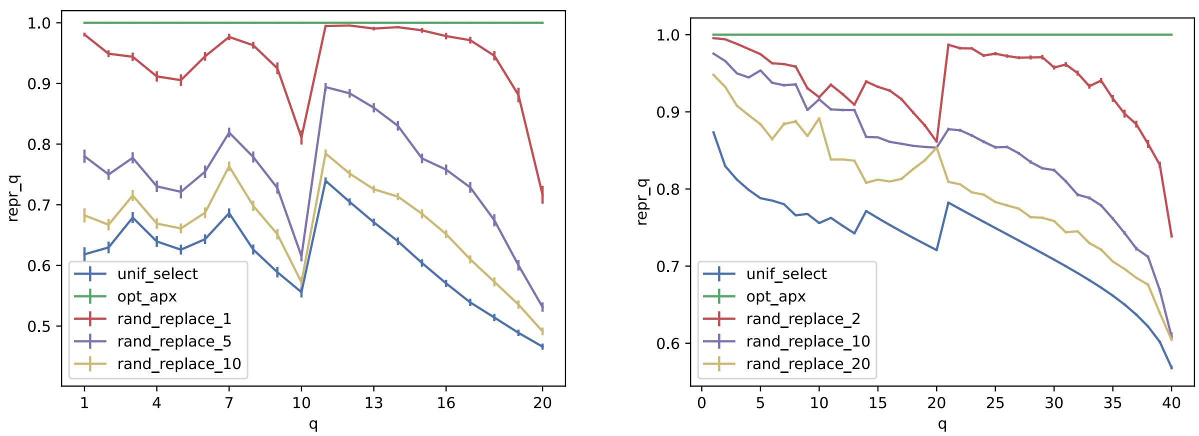


Figure 6.2: OptProxy finds the *optimal* panel $\mathrm{SC}_q(P^*, d)$ on small instances $k = 10$, based on `Adult`.



(a) `Adult` with $k = 20$.

(b) `ESS` UK with $k = 40$.

Figure 6.3: Comparison of different algorithms for fixed k, where RandomReplace$_r$ is applied to the panel selected by OptProxy. The y-axis shows the average ratio of the q-social cost of OptProxy to that of different algorithms.

Part III

Group Selection

Chapter 7

Strategyproof Mean Estimation from Multiple-Choice Questions

7.1 Introduction

Organizations often desire accurate estimates of population statistics (e.g., the mean of a set of values) in settings where eliciting exact values from agents is costly or impractical. For instance, suppose that you sit on an admissions committee, and your committee's task is to accurately estimate the number of candidates who will accept their admission offer, perhaps in order to decide how many more admissions offers to extend (concretely, consider the problem of admitting students in two waves corresponding to early action and regular admission). This is a consequential problem; there are significant direct and indirect costs associated with having many more or fewer matriculants than intended.

Without more information either about or from the admits, this is hopeless. One potential approach is to ask each admit to provide an estimate of the probability p_i that the admit matriculates, but this is problematic because each admit may not know their own exact probability of accepting the offer, and coming up with exact probabilities places nontrivial cognitive loads on participants. Therefore, it is more reasonable to ask a multiple-choice question of the form, "How likely are you to accept this offer?" with choices "High," "Medium," and "Low." The task for the university, then, is to reconstruct an accurate estimate of the number of students who will accept their offers based on the coarse-grained information yielded by these multiple-choice queries. Specifically, the question is

> "What collection of multiple-choice questions should you ask, and how should you interpret the answers so as to estimate the expected number of matriculants as accurately as possible?"

7.1.1 Our Approach and Results

Consider n agents, where each agent i has a value $p_i \in [0,1]$ and $p := (p_1, \dots, p_n)$ is the vector of all agents' values. We are interested in estimating $\|p\|_1 = \sum_{i=1}^{n} p_i$.

We interpret multiple-choice questions as forming a partition of $[0,1]$ into subsets $X_1, \dots, X_k$, and asking agent i for the index j such that $p_i \in X_j$. It is known [LS09] that in order for multiple-choice questions to elicit truthful responses, each X_j must itself be an interval. Moreover, intervals are easier to interpret than arbitrary subsets; for example, the choice "Low" can be defined as $p_i \in [0, 1/3]$. Therefore, we restrict our multiple-choice questions to this framework.

Our goal is to design an estimator that consists of a set of (possibly different) multiple-choice questions which are posed to the agents, together with a function that outputs an estimate of $\|p\|_1$ based on the agents' answers to the multiple-choice questions; we denote the output of the estimator by $q(p)$. We measure the accuracy of the estimator using the *mean squared error (MSE)* $\mathbb{E}[(\|p\|_1 - q(p))^2]$ or the *mean absolute error (MAE)* $\mathbb{E}[|\|p\|_1 - q(p)|]$.[1]

[1] Throughout, we evaluate additive error with respect to estimating the sum of the p_i, from which the additive error with respect to mean estimation can easily be derived.

We consider two settings corresponding to different levels of information about the agent's values. When no information about p is known (worst case), we consider the problem of designing a randomized estimator with good worst-case performance (when averaged over the estimator's randomness). We give a single randomized estimator $\bar{q}$ which guarantees

$$\mathsf{mse}(\bar{q}) = O\left(\frac{n}{k^2}\right), \quad \mathsf{mae}(\bar{q}) = O\left(\frac{\sqrt{n}}{k}\right)$$

and demonstrate that this is asymptotically optimal for both measures of error.

In the second setting, each p_i is drawn from a known distribution P_i; we consider the problem of designing a deterministic estimator which performs well on average (over the randomness of the p_i). We present an MSE-optimal estimator, and show that the problem of devising an MAE-optimal estimator is $\#\mathcal{P}$-hard.

Finally, we conduct experiments in the latter setting of known distributions, in which we aim to quantify the benefit of tailoring the estimator to the distributions. Focusing on the MSE due to our computational results, we show that the optimal estimator significantly outperforms a naïve estimator.

7.1.2 Related Work

Caragiannis, Procaccia, and Shah [CPS16] study strategyproof mean estimation in a related setting, where strategic agents supply samples in order to move the estimation of the mean close to their own value. In this setting, they ask if the sample median is the best truthful estimator of the population mean, which is not the case, and characterize worst-case optimal truthful estimators that provably outperform the median for distributions with bounded support.

More broadly, mechanism design for information elicitation has been widely studied in computer science and economics [ZR08; CK11; WC14]. Many prior works in mechanism design focus on eliciting truthful signals from agents, often through direct verification mechanisms like strictly proper scoring rules [GR07; Bri50; Goo52; Win69] and prediction markets [WZ04;

Ber+08]. In a related vein, Radanovic and Faltings [RF14] developed a mechanism for truthful elicitation of continuous signals, but we consider the problem of reconstructing a continuous value from discrete reports.

Additionally, Soloviev and Halpern [SH18] consider the problem of acquiring information with resource limitations, where budget constraints on the number of tests in their setting roughly map to constraints in our setting on the granularity of queries. However, their setting involves noisy tests of Boolean formula truth values, as opposed to estimating a population statistic.

Furthermore, by considering estimation error as it varies over a range of k, we can investigate the relationship between the elicitation of values and the accuracy of our estimate, a tradeoff which has been studied for intelligent decision-making systems in a sequential setting by Boutilier [Bou02].

Alternatively, the task of optimally estimating $\sum_i p_i$ in terms of MSE or MAE can be viewed as variants of k-means and k-medians clustering, respectively. On the one hand, it resembles a special case of clustering in that P is product distribution over $[0,1]^n$ (as opposed to a general distribution over a metric space), and our 'clusters' C correspond to vectors p which yield the same vector of multiple-choice answers, and the C are constrained to have a product structure as well. On the other hand, it is distinct from clustering in that r reports scalar representative ℓ_1 norms for the clusters, and so we take the ℓ_1 norm of our C before calculating error.

This line of inquiry is also related to the notion of approximate query processing (AQP) in the data mining and databases literature, which is the practice of answering expensive aggregation queries with limited resources. Multiple research groups have focused on the bounded-error estimation of aggregates (e.g., sums of values in a database), but mostly through sampling techniques as opposed to summarization approaches [Jag+98; CDN07; Cha+01; BCD03]. More recently, there has been some work on summarizing data distributions with histograms in order to minimize the ℓ_2 distance between the distribution and the histogram approximation [Ach+15; Din+16], but this only coincides with our MSE setting when we query exactly one agent.

7.2 The Model

We consider a set of agents $[n] = \{1, \ldots, n\}$, each with an associated number $p_i \in [0, 1]$.

Our goal is to devise a scheme for estimating the sum $\|p\|_1 = \sum_{i=1}^{n} p_i$ (or, equivalently, the mean) to minimize additive error. We may ask each agent i which of k intervals contains their p_i, and so our estimator chooses n partitions $\mathcal{B}_i := \{B_{i,1}, \ldots, B_{i,k}\}$ of $[0, 1]$ into intervals, and for each i the function $b_i : [0, 1] \to [k]$ poses the question to agent i and returns their response; $b_i(p_i) = j$ if $p_i \in B_{i,j}$. We refer to $b(p) := (b_1(p_1), \ldots, b_n(p_n))$ as the *classifier.* Next the *aggregator* $r : [k]^n \to \mathbb{R}$ takes the agents' responses and estimates $\sum_i p_i$; we refer to the output of r as the *report.*

In order to ensure agents truthfully report which interval their p_i falls in, we propose to pay agents based on their eventual decision and their report. In the setting of school admission, for instance, this payment may b e a (partial) refund of each student's application fee. To this end, we build on work by Lambert and Shoham [LS09] that studies the problem of eliciting truthful answers to multiple-choice questions. In their framework, payments based on agent reports and the observed outcome are used to induce truthful answers from agents, and they establish necessary and sufficient conditions on the structure of multiple-choice questions that ensure the existence of such payments. In our setting, their characterization implies that we can elicit truthful answers (i.e., there exists a payment scheme that incentivizes truthful reporting) if and only if the questions we ask are of the form, "To which of these intervals does your p_i belong?"

Let $c_i := b_i(p_i)$ denote the answer of each agent to this question. For $c \in [k]^n$, the box $C_c := \prod_i B_{i,c_i} = b^{-1}(c)$ is the set of $(p_1, \ldots, p_n)$ for which each agent i answers c_i. In terms of the classifier and aggregator, our goal may be restated as finding the *estimator*

$$q := r \circ b : [0,1]^n \to \mathbb{R}$$

that minimizes expected error.

We consider two natural measures of error, mean squared error (MSE) and mean absolute error (MAE), which are defined to be

$$\begin{aligned} \mathsf{mse}(q) &= \mathbb{E}\left[(\|p\|_1 - q(p))^2\right], \\ \mathsf{mae}(q) &= \mathbb{E}\left[|\|p\|_1 - q(p)|\right]. \end{aligned}$$

When the P_i are adversarially chosen, these expectations are taken over the randomness of the estimator; when each p_i is drawn from a known distribution P_i, these expectations are taken over the product distribution P.

Finally, throughout the paper we denote the centroid of $C \subset \mathbb{R}^n$ by $\mu(C)$. More formally,

$$\mu(C) := \frac{1}{P(C)} \int_C p \, dP,$$

where P is a measure on $\mathbb{R}^n$.

7.3 Worst-Case Guarantees

In this section we consider the case where no knowledge of the p_i is assumed, and establish upper and lower bounds on the performance of deterministic and randomized estimators.

First, suppose that the estimator $q = r \circ b$ is deterministic. For fixed b, it is clear that r should report the sum corresponding to the center point of each box $C_c = \prod_i B_{i,c_i}$, since this minimizes the worst-case error across all $p \in C_c$. Accordingly, an adversary will seek the box with the largest ℓ_1 diameter, which can be identified by finding the $B_{i,j}$ of maximum diameter for each i. Therefore the worst-case optimal deterministic estimator chooses equipartitions

$\mathcal{B}_i = \{[0, \frac{1}{k}], \ldots, [\frac{k-1}{k}, 1]\}$, reports the ℓ_1 norm of the center of each box, and satisfies

$$\max_p (\|p\|_1 - q(p))^2 = \frac{n^2}{4k^2}$$
$$\max_p |\|p\|_1 - q(p)| = \frac{n}{2k},$$

and this is clearly tight. This is the uniform estimator, and we will denote it $q_U := r_U \circ b_U$, where b_U partitions each $[0,1]$ into equal-size subintervals, and $r_U(c)$ is the ℓ_1 norm of the center of each C_c.

A randomized estimator, however, can perform significantly better over worst-case inputs. Indeed, consider the following randomized estimator, which we denote by $\bar{q} = r \circ \bar{b}$. We construct a randomized classifier by choosing "shifts" $s_i \in [0, \frac{1}{k-1})$ for each i uniformly and independently. Take

$$\bar{b}_i(p_i) := j \qquad \text{s.t.} \qquad \frac{j-1}{k-1} \leqslant p_i + s_i < \frac{j}{k-1}.$$

Intuitively, this partitions $[0,1]$ into k subintervals by taking the $k-1$ thresholds $1/(k-1), \ldots, 1$ and shifting them left by s_i. Then make the (deterministic) reports

$$r_i(j) := \frac{j-1}{k-1}$$

for $j \in [k]$, and define $\bar{b}(q) := (\bar{b}_1(p_1), \ldots, \bar{b}_n(p_n))$ and take the aggregator $r(c) := r_1(c_1) + \ldots + r_n(c_n)$. Putting these together yields the randomized estimator

$$\bar{q} := r \circ \bar{b}.$$

Our main result for this section is the following theorem.

Theorem 51. *In the worst-case setting, the randomized estimator $\bar{q}$ satisfies*

$$\mathsf{mse}(\bar{q}) = O\left(n/k^2\right)$$
$$\mathsf{mae}(\bar{q}) = O\left(\sqrt{n}/k\right).$$

Moreover, these bounds are asymptotically optimal for both measures of error.

The rest of the section is devoted to proving the theorem. We do so via several lemmas, starting with the upper bound.

Lemma 20. *The randomized estimator $\bar{q}$ satisfies*

$$\mathsf{mse}(\bar{q}) = O\left(n/k^2\right)$$
$$\mathsf{mae}(\bar{q}) = O\left(\sqrt{n}/k\right).$$

Proof. For fixed p, we begin by analyzing the estimator coordinate-by-coordinate. Take

$$X_i := p_i - (r_i \circ \bar{b}_i)(p_i)$$

to be the (signed) error of $\bar{q}$ in coordinate i, and let $j_i := \max\{j \in [k] : \frac{j-1}{k-1} \leqslant p_i\}$. If

$$z_i := p_i(k-1) - (j_i - 1)$$

is the proportion of the way between multiples of $1/(k-1)$ that p_i falls, then $X_i = (z_i - 1)/(k-1)$ with probability z_i and $X_i = z_i/(k-1)$ with probability $1 - z_i$. Note that $\mathbb{E}[X_i] = 0$. By the definition of r and the independence of the X_i,

$$\begin{aligned}
\mathsf{mse}(\bar{q}(p)) &= \mathbb{E}\left[\left(\sum_i X_i\right)^2\right] = \sum_i \mathrm{Var}\left[X_i\right] \\
&= \sum_i \frac{z_i^2(1-z_i) + z_i(1-z_i)^2}{(k-1)^2} \leqslant \frac{n/4}{(k-1)^2} \\
&= O\left(\frac{n}{k^2}\right).
\end{aligned}$$

We now turn to the MAE case. Note that X_i is bounded in some range of width $1/(k-1)$, and that

$$\mathsf{mae}(\bar{q}(p)) = \mathbb{E}[|\|p\|_1 - \bar{q}(p)|] = \mathbb{E}\left[\left|\sum_i X_i\right|\right].$$

Next we apply Hoeffding's inequality in order to upper bound this expectation. By Hoeffding [Hoe63],

$$\Pr\left[\left|\sum_i X_i\right| \geqslant t\right] \leqslant 2\exp\left\{\frac{-2t^2}{\frac{n}{(k-1)^2}}\right\},$$

and so choosing $t = m\sqrt{n}/(k-1)$ for $m \in \mathbb{N}$ yields

$$\Pr\left[\left|\sum_i X_i\right| \geqslant \frac{m\sqrt{n}}{(k-1)}\right] \leqslant \frac{2}{e^{2m^2}}. \tag{7.1}$$

Finally, let $X := \sum_i X_i$ and observe that for any $\sigma > 0$,

$$\begin{aligned}\mathsf{mae}(\bar{q}(p)) &= \mathbb{E}\left[|X|\right] \\ &\leqslant \sum_{m=1}^{\infty} \sigma m \Pr\left[|X| \in [\sigma(m-1), \sigma m]\right] \\ &\leqslant \sum_{m=1}^{\infty} \sigma m \Pr\left[|X| \geqslant \sigma(m-1)\right].\end{aligned}$$

Taking $\sigma = \sqrt{n}/(k-1)$ and applying Equation (8.72),

$$\leqslant \frac{\sqrt{n}}{k-1} \sum_{m=1}^{\infty} \frac{2m}{e^{2(m-1)^2}}.$$

This infinite series converges, which implies that $\mathsf{mae}(\bar{q}(p)) = O(\sqrt{n}/k)$, as desired. □

By Yao's minimax principle [Yao77], in order to derive a lower bound for all randomized algorithms, it suffices to fix a distribution over inputs and lower bound the average performance of any deterministic algorithm over this randomized input. To this end, we will consider the uniform distribution over $[0,1]^n$, which we denote D, and lower bound the performance of any deterministic estimator over it. In doing so, we will prove the intuitive fact that the uniform estimator $q_U = r_U \circ b_U$ is optimal for D.

First we present a structural lemma about the optimal aggregator r for any fixed classifier b. Let $S(C, P)$ denote the probability distribution over $\mathbb{R}$ derived by taking the ℓ_1 norm of C:

$$\Pr[S \leqslant x] = \Pr_P[\|p\|_1 \leqslant x | p \in C].$$

We will repeatedly make use of the following insight, which follows from calculus and reflects one of Lloyd's optimality conditions for k-means clustering [Llo82]:

Lemma 21. *For $p \sim P$ and for fixed b, the MSE- and MAE- optimal aggregators $r_b^1(c)$ and $r_b^2(c)$ (respectively) report the means and medians (respectively) of $S(C_c, P)$ for all $c \in [k]^n$.*

With this in hand, we are ready to analyze the performance of q_U over D, which (by Yao's minimax principle) establishes the lower bound of Theorem 51.

Lemma 22. *If $p \sim D$ is drawn uniformly at random then the uniform estimator $q_U = r_U \circ b_U$ is optimal in terms of both MSE and MAE, and*

$$\mathsf{mae}(q_U) = \Omega\left(n/k^2\right),$$
$$\mathsf{mae}(q_U) = \Omega\left(\sqrt{n}/k\right).$$

Proof. This follows in two steps. We will first prove that the uniform strategy $q_U = r_U \circ b_U$ is optimal for MSE and compute $\mathsf{mse}(q_U)$ directly. Then we will prove that q_U is optimal for MAE, and finally lower bound $\mathsf{mae}(q_U)$.

To begin, note that under the uniform distribution $P = D$ and for fixed b, Lemma 21 implies that for both MSE and MAE, the optimal aggregator is the r_b which reports the ℓ_1 norms of the centers of the C_c. Therefore we may assume that all estimators use reports r_b which are optimal for their b, and argue that b_U is the best partitioning.

The MSE case boils down to a question of variance, and it turns out we can compute the error directly. As before, let

$$X_i := \sum_{j=1}^{k} \mathbb{1}_{\{p_i \in B_{i,j}\}} \left(p_i - \mu(B_{i,j})\right)$$

be the (signed) error in coordinate i, the difference between their actual p_i and the center $\mu(B_{i,j})$ of the interval $B_{i,j}$ containing their p_i. Then because the p_i are independent and

$$r_b(c) = \mu\left(S\left(C_c, D\right)\right) = \sum_i \mu_D(B_{i,c_i}),$$

we have that for $q = r_b \circ b$:

$$\begin{aligned}
\mathsf{mse}(q) &= \mathbb{E}_D\left[(\|p\|_1 - q(p))^2\right] = \mathbb{E}_D\left[\left(\sum_i X_i\right)^2\right] \\
&= \mathrm{Var}\left[\sum_i X_i\right] = \sum_i \mathrm{Var}[X_i] = \sum_i \int_0^1 X_i^2\,dx \\
&= \sum_i \sum_j \int_{B_{i,j}} (x - \mu_D(B_{i,j}))^2\,dx \\
&= \sum_i \sum_j \frac{1}{12}\mathsf{diam}(B_{i,j})^3.
\end{aligned}$$

At this point, the method of Lagrange multipliers confirms that MSE is minimized when all $B_{i,j}$ are of equal diameter $1/k$, which yields precisely b_U. Therefore $q_U = r_U \circ b_U$ is optimal for D, and it has cost $\mathsf{mse}(q_U) = n/12k^2$.

We now turn to MAE, and use a differential argument to prove that q_U is MAE-optimal when $P = D$. We begin by showing that the MAE contributed by a box C_c is convex in each of the dimensions of C_c. This will let us argue that for any classifier b with partitions $\{\mathcal{B}_i\}$ in which one partition is unbalanced, meaning that for some i (say $i = 1$) and some j it is the case that $\mathsf{diam}(B_{1,j}) < \mathsf{diam}(B_{1,j+1})$, the $b^\star$ which equalizes their widths decreases the MAE: $\mathsf{mae}(r_{b^\star} \circ b^\star) < \mathsf{mae}(r_b \circ b)$. This then implies that $q_U = r_U \circ b_U$ is MAE-optimal, since it is the only b which cannot be equalized in this way.

To see that MAE contribution is convex in each dimension of C, let $C' := \prod_{i=2}^n [-w_i, w_i]$ and consider the box $C(t) := [-t, t] \times C'$. Then by Lemma 21 the optimal report for both C and C' is 0. Let the contribution to MAE by C with report r be denoted $e(C, r)$, and call the contribution with optimal report $e(C)$. Then by symmetry,

$$\begin{aligned}
e(C(t), 0) &= \int_{C(t)} \left|\sum_i p_i\right|\,dp \\
&= 2\int_0^t \int_{C'} \left|x + \sum_{i=2}^n p_i\right|\,dp\,dx \\
&= 2\int_0^t e(C', x)\,dx,
\end{aligned}$$

and therefore

$$\frac{de(C(t),0)}{dt} = 2e(C', t).$$

The (omitted) proof of Lemma 21 shows that $\frac{de(C,r)}{dr} > 0$ for r greater than the optimal report, and so

$$\frac{d^2e(C(t),0)}{dt^2} > 0,$$

as desired.

In order to show that $b^\star$ improves upon b, note that their boxes C_c differ only for those of the form

$$\widehat{C}_c := B_{1,j} \times \prod_{i=2}^{n} B_{i,c_i}, \qquad \tilde{C}_c := B_{1,j+1} \times \prod_{i=2}^{n} B_{i,c_i}$$

Pairing these up by c, it suffices to show that for each c,

$$e(\widehat{C}_c^\star) + e(\tilde{C}_c^\star) < e(\widehat{C}_c) + e(\tilde{C}_c),$$

where $C_c^\star$ are the boxes given by $b^\star$. This follows from the convexity of $e(C)$ in each dimension of C, established above. We conclude that q_U is MAE-optimal.

We next lower bound $\mathsf{mae}(q_U)$. Let $L \subseteq [n]$ and $H \subseteq [n]$ be the set of indices i with errors X_i that are negative and positive, respectively. It holds that

$$\begin{aligned}\mathsf{mae}(q_U) &= \mathbb{E}\left[|\sum_{i\in[n]}(p_i - q_U(p_i))|\right]\\ &= \mathbb{E}\left[|\sum_{i\in L} X_i + \sum_{i\in H} X_i|\right].\end{aligned}$$

We establish that L and H are, with constant probability, of sufficiently different sizes to lead to $\sqrt{n}/k$ error. First, note that because we are playing against a uniform adversary, the probability that each p_i is in L is $1/2$; the probability that each p_i is in H is symmetrically $1/2$. Because these are just Bernoulli random variables, applying the De Moivre-Laplace theorem (a version of the Central Limit Theorem) tells us that, as n becomes large, the sum of these Bernoulli random variables converges to a normal distribution with mean $n/2$ and standard deviation $\sqrt{n}/2$. Therefore, we know that with constant probability, the total number of agents in H is at least $\sqrt{n}$ from its mean of $n/2$, that is,

$$\Pr\left[||H| - \mathbb{E}[|H|]| \geqslant \sqrt{n}\right] = \beta$$

for a constant β. It follows that, with constant probability β, $||L| - |H|| \geqslant 2\sqrt{n}$; denote this event by $\mathcal{E}$. Therefore,

$$\begin{aligned}
\mathsf{mae}(q_U) &= \mathbb{E}\left[|\sum_{i=1}^{n} X_i|\right] \\
&\geqslant \mathbb{E}\left[|\sum_{i=1}^{n} X_i| \;\middle|\; \mathcal{E}\right] \cdot \Pr[\mathcal{E}] \qquad (7.2)\\
&= \mathbb{E}\left[|\sum_{i=1}^{n} X_i| \;\middle|\; \mathcal{E}\right] \cdot \beta.
\end{aligned}$$

Now assume that $\mathcal{E}$ occurred Without loss of generality, assume that $|L| \leqslant |H|$ and, in particular, randomly break H up into two sets, H_1 and H_2, such that $|H_1| = |L|$ and $|H_2| \geqslant 2\sqrt{n}$. By construction of H_1, the sum of the errors in indices $i \in L \sqcup H_1$ is symmetric with mean 0. It holds that

$$\begin{aligned}
&\mathbb{E}\left[|\sum_{i=1}^{n} X_i| \;\middle|\; \mathcal{E}\right] \\
&\quad = \mathbb{E}\left[|\sum_{i\in L\cup H_1} X_i + \sum_{i\in H_2} X_i| \;\middle|\; \mathcal{E}\right] \\
&\quad \geqslant \Pr\left[|\sum_{i\in L\cup H_1} X_i| \geqslant 0 \;\middle|\; \mathcal{E}\right] \mathbb{E}\left[|\sum_{i\in H_2} X_i| \;\middle|\; \mathcal{E}\right] \\
&\quad \geqslant \frac{1}{2} \cdot 2\sqrt{n} \cdot \frac{1}{4k} = \Omega\left(\frac{\sqrt{n}}{k}\right).
\end{aligned}$$

The desired bound follows by combining this with Equation (7.2). □

7.4 Estimation with Priors

In practice, it is useful to go beyond worst-case guarantees and ask to what degree knowing some additional information about the p_i can improve our ability to estimate their sum.

Specifically, suppose that we have access to distributions $P_1, \ldots, P_n$ from which the p_i are drawn; we make the standard assumption (for computational complexity results) that these distributions are discrete. Can we design a more accurate estimator that takes prior knowledge into account?

For instance, in our running example of admissions (say for a Ph.D. program), one could easily come up with priors by applying machine learning to historical admissions data. The prior for a candidate would, of course, depend on relevant features such as their alma mater and research interests.

It is important to note that, in this setting, the optimal estimator is deterministic, as the error of a randomized estimator is just a convex combination of deterministic estimators in its support.

7.4.1 An Efficient Estimator for MSE

When estimation is evaluated according to MSE, it turns out that we can answer the foregoing question in the positive:

Theorem 52. *Given discrete priors P_i for each p_i, $i \in [n]$, there is a polynomial-time estimator that is optimal with respect to MSE.*

A key component of our analysis is the following structural insight:

Lemma 23. *Given an estimator $q = r_b \circ b$, where r_b reports optimally for a given classifier b, $\mathsf{mse}_P(q) = \sum_i \mathsf{mse}_{P_i}(q_i)$, where $q_i = r_{b_i} \circ b_i$.*

We will prove this by employing a result from the vector quantization literature. In quantization, roughly speaking, the task is to compress some signal using only a representative subset of its values in such a way that compression error is minimized. Vector quantization performs this task for vector-valued signals, and does so using an n-dimensional partition [GN98]. Since it typically seeks an MSE-error-minimizing vector representative $\boldsymbol{r}(c)$ for each box C_c in its n-dimensional partition, vector quantization may be seen as an instance of k^n-means clustering in $\mathbb{R}^n$ subject to the constraint that all clusters obey this product-of-partitions structure. In this setting, it is known [JDS11] that if the partitions are made along independent axes, then the MSE of the optimal vector quantizer is additive:

Lemma 24. *If $\boldsymbol{q}$ is a vector quantizer as described above which reports the centroids $\mu(C_c)$ for every C_c, and is given by $\boldsymbol{r}(c) = (r_1(c_1), \ldots, r_n(c_n))$ and $b(p) = (b_1(p_1), \ldots, b_n(p_n))$, then*

$$\mathbb{E}_P\left[\|p - \boldsymbol{q}(p)\|_2^2\right] = \sum_i \mathbb{E}_{P_i}\left[(p_i - r_i(b_i(p_i)))^2\right], \tag{7.3}$$

where each r_i reports the centroid $\mu(B_{i,j})$.

This is a direct consequence of independence of the p_i, together with the fact that the X_i have mean 0.

The key difference between our problem and the vector quantization setting is that we measure error with respect to the aggregate $\sum_i p_i$, which means errors with respect to individual agents can "cancel out." Nevertheless, the same structural insight holds, as we show now.

Proof of Lemma 23. We proceed in two steps. First we argue that for fixed classifier b, the optimal report r_b reports the ℓ_1 norm $r_b(c) = \|\mu_P(C_c)\|_1$ of the centroid of each box C_c. Next, we argue that this coincides with the error given on the right-hand side of Equation (7.3); the theorem then follows.

We begin by showing that the optimal aggregator reports $r_b(c) = \|\mu_P(C_c)\|_1$. As in Lemma 22, let $e(C, r)$ denote the contribution of $C = \prod_i B_i$ to MSE under report r. We proceed via a differential argument:

$$\begin{aligned}
e(C, r) &= \int_C (\|p\|_1 - r)^2 \, dP \\
&= r^2 \int_C dP - 2r \int_C \|p\|_1 \, dP + \int_C \|p\|_1^2 \, dP \\
\frac{de(C, r)}{dr} &= 2r \int_C dP - 2 \int_C \|p\|_1 \, dP.
\end{aligned}$$

Setting $\frac{de(C,r)}{dr} = 0$ yields the optimal report,

$$\begin{aligned}
r^\star &= \frac{1}{P(C)} \int_C \|p\|_1 \, dP = \frac{1}{P(C)} \left(\sum_i \int_C p_i \, dP \right) \\
&= \frac{1}{\prod_i P_i(B_i)} \sum_i \left(\int_{B_i} p_i \, dP_i \prod_{j \neq i} \int_{B_j} dP_j \right) \\
&= \sum_i \frac{1}{P_i(B_i)} \int_{B_i} p_i \, dP_i = \sum_i \mu_{P_i}(B_i) \\
&= \|\mu_P(C)\|_1,
\end{aligned}$$

where this last step is a standard property of the centroid.

Therefore for $q = r_b \circ b$, the error $\mathsf{mse}(q)$ takes the form

$$\begin{aligned}
\mathsf{mse}_P(q) &= \sum_{c \in [k]^n} \int_{C_c} \left(\|p\|_1 - \|\mu(C_c)\|_1 \right)^2 \, dP \\
&= \sum_{c \in [k]^n} \int_{C_c} \left(\sum_i \left(p_i - \mu_{P_i}(B_{i,c_i}) \right) \right)^2 \, dP.
\end{aligned}$$

Since the centroid is an unbiased estimator,

$$\begin{aligned}
&= \sum_{c \in [k]^n} \int_{C_c} \sum_i \left(p_i - \mu_{P_i}(B_{i,c_i}) \right)^2 \, dP \\
&= \mathbb{E}_P \left[\|p - \boldsymbol{q}(p)\|_2^2 \right],
\end{aligned}$$

as in the right-hand side of Equation (7.3). Therefore by Lemma 24,

$$\begin{aligned}
&= \sum_i \mathbb{E}_{P_i} \left[(p_i - \mu(B_{i,c_i}))^2 \right] \\
&= \sum_i \mathsf{mse}_{P_i}(q_i).
\end{aligned}$$

□

Proof of Theorem 52. Our goal is to minimize $\mathsf{mse}_P(q)$, and Lemma 23 implies that this can be accomplished by individually minimizing $\mathsf{mse}_{P_i}(r_{b_i} \circ b_i)$ for each $i \in [n]$. Since the P_i are discrete distributions, finding the b_i which minimizes $\mathsf{mse}_{P_i}(r_{b_i} \circ b_i)$ is precisely an instance of one-dimensional Euclidean k-means. It is well-known that this can be solved efficiently via dynamic programming using a recurrence described by Jensen [Jen69]. Therefore, given priors P_i we can derive an estimator q which minimizes $\mathsf{mse}(q)$: we first find $b_1^\star, \ldots, b_n^\star$ which minimize $\mathsf{mse}_{P_i}(r^2_{b_i} \circ b_i)$ for each $i \in [n]$, where $r^2_{b_i}$ is the optimal aggregator for b_i, then take the optimal partitions

$$b(p) = (b_1^\star(p_1), \ldots, b_n^\star(p_n)) .$$

Lemma 23 then guarantees that this $q = r^2_b \circ b$ is optimal. □

7.4.2 Hardness for MAE

In contrast to the case of MSE with priors, it turns out that devising an MAE-optimal mean estimation strategy is $\#\mathcal{P}$-hard.

To be concrete, the computational problem is defined as follows: Given a collection of discrete prior distributions $P_1, \ldots, P_n$ and a positive integer k, we are asked for a collection of partitions $b_i : [0,1] \to [k]$ and an aggregator $r : [k]^n \to \mathbb{R}$ which together minimize

$$\mathsf{mae}(r \circ b) = \mathbb{E}_P \left[\left| \|p\|_1 - r \circ b(p) \right| \right] ,$$

where $P = \prod_i P_i$.

Theorem 53. *Given discrete priors P_i for each p_i, $i \in [n]$, the problem of computing an optimal estimator with respect to MAE is $\#\mathcal{P}$-hard.*

In a nutshell, we prove the theorem through a pair of reductions; beginning with the counting version of Knapsack, which is $\#\mathcal{P}$-complete, we show that the problem of finding a median of the sum of the (independent) Bernoulli random variables $\alpha_i \mathsf{Bernoulli}(p)$ is $\#\mathcal{P}$-complete. We then describe a way to derive distributions $P_1, \ldots, P_n$ from a collection of weighted Bernoulli random variables such that devising an MAE-optimal strategy for $P_1, \ldots, P_n$ finds a median of the weighted Bernoulli sum. We start by reducing from the following problem:

Definition 12. Given a rational $x \in [0,1]$ and nonnegative integer weights $\alpha_1, \ldots, \alpha_n$, Weighted-Binomial-Median (WBM) asks for a median of the random variable

$$Z := \sum_{i=1}^{n} \alpha_i \,\mathsf{Bernoulli}(x),$$

where the $\mathsf{Bernoulli}(x)$ are independent (and identically distributed).

This weighted binomial distribution (WBD) is comparable to the Poisson binomial distribution (PBD) in that both generalize the binomial distribution. However the PBD is an unweighted sum of Bernoulli random variables with distinct probabilities x_i, while the WBD is a sum of Bernoulli random variables with a common x but distinct integer weights.

Lemma 25. WBM *is* $\#\mathcal{P}$*-Hard.*

Proof. In order to show to show that WBM is $\#\mathcal{P}$-complete, we will reduce from the counting version of the knapsack problem, which is known to be $\#\mathcal{P}$-complete [GJ79]: Given a list of nonnegative integer weights $w_1, \ldots, w_n$ and an integer capacity W, #Knapsack asks how many sets $S \subseteq [n]$ exist such that $\sum_{i \in S} w_i \leqslant W$. And we will make use of a slight variant of counting knapsack: Given an integer k, a list of nonnegative integer weights $w_1, \ldots, w_n$ an integer capacity W, and an integer threshold N, k#Knapsack finds $|\mathcal{S}|$, where

$$\mathcal{S} := \{S \subseteq [n] : \sum_{i \in S} w_i \leqslant W \text{ and } |S| = k\}.$$

It can be seen that k#Knapsack is $\#\mathcal{P}$-complete via an easy reduction from #Knapsack: given an instance of #Knapsack, simply query k#Knapsack for all values of k and return the sum of the answers.

Turning to the hardness of WBM, we begin by arguing that WBM may be assumed to return the largest possible median. This is because, for an instance of WBM given by $(x, \alpha_1, \ldots, \alpha_n)$, we may instead take a perturbed probability $\bar{x} = x + \gamma$. By choosing γ small enough, we can ensure that the median $\bar{m}$ of $\bar{Z} := \sum_i \alpha_i \mathsf{Bernoulli}(\bar{x})$ is a median of Z, but that it is the largest possible such median. Informally, we may tweak x gently enough that we preserve the median but break any median ties.

Formally, let F_Z be the cumulative density function (CDF) of Z. Since Z is a distribution comprised solely of atoms of weight $x^k(1-x)^{n-k}$ for $k \in [n]$, it suffices to find some perturbation γ for which

$$F_Z(m) - F_{\bar{Z}}(m) < a,$$

where m is a median of Z and a is a lower bound on the size of an atom in both Z and $\bar{Z}$. To show that we may choose such an a, note we may assume that $(1-x)^n \leqslant 1/2$, since otherwise the largest possible m is 0, and similarly that $\bar{x}^n \leqslant 1/2$, since otherwise we may easily check if the largest possible m is $\sum_{i \in [n]} \alpha_i$. Among all Z for which $x^n \leqslant 1/2$ and $(1-x)^n \leqslant 1/2$, the smallest possible atom is of size $\frac{1}{2}(2^{1/n} - 1)^n$, and so $a := 1/n^n$ is a lower bound on the atom size in Z for any value of x that concerns us.

Since Z is atomic, we then have that

$$F_Z(y) = \sum_{z \leqslant y} \Pr[Z = z] \tag{7.4}$$

$$= \sum_{S \subseteq [n]} x^{|S|}(1-x)^{n-|S|} \, \mathbb{1}_{\{\sum_{i \in S} w_i \leqslant y\}} \tag{7.5}$$

and so

$$\frac{\partial F_Z(y)}{\partial x} \leqslant \sum_{S \subseteq [n]} \frac{\partial}{\partial x} x^{|S|}(1-x)^{n-|S|} \leqslant n2^n. \tag{7.6}$$

Therefore taking $\gamma = \frac{a}{n2^n}$ will suffice, and $\bar{x} = x + \gamma$ will have a binary representation which is polynomial in the number of input bits.

We now reduce from k#Knapsack. Given an instance of k#Knapsack described by $(k, w_1, \dots w_n, W)$, let $\Gamma := \langle k \rangle + \sum_i \langle w_i \rangle + \langle W \rangle$ be the length of the binary representation of these integers. For each i, let

$$\alpha_i := G + w_i,$$

where $G := (n+1) \sum_i w_i$. If $Z = \sum_i \alpha_i \mathsf{Bernoulli}(x)$ for some rational $x \in [0, 1]$, then since the w_i are positive, the support of Z is clustered to the left of the integers $0, G, \dots, nG$. Specifically, we have by Equation (7.5) that

$$\begin{aligned} F_Z(Gk) &= \sum_{S\subseteq[n]} x^{|S|}(1-x)^{n-|S|}\, \mathbb{1}_{\{\sum_{i\in S} w_i \leqslant Gk\}} \\ &= \sum_{j=0}^{k-1} \binom{n}{j} x^j (1-x)^{n-j}, \end{aligned}$$

and so $F_Z(Gk)$ can be computed in time polynomial in $\Gamma + \langle x\rangle$.

Next, with k given, consider a binary search over (rational) x which searches for the largest possible x for which $m \leqslant Gk + W$. Once the binary search is far enough along and the change in x is sufficiently small, $F_Z(m)$ approaches $1/2$ and the remaining change possible in $F_Z(m)$ will be small with respect to the atomic lower bound a. We may terminate our search, say, when $F_Z(m) \in [1/2, 1/2 + a/10]$. At this point m is the largest value of size at most $Gk + W$ in the support of Z, and so by this maximality of $m \leqslant Gk + W$,

$$\begin{aligned} F_Z(m) &= \sum_{S\subseteq[n]} x^{|S|}(1-x)^{n-|S|}\, \mathbb{1}_{\{\sum_{i\in S} w_i \leqslant m\}} \\ &= \sum_{j=0}^{k-1} \binom{n}{j} x^j (1-x)^{n-j} + |\mathcal{S}_k|\, x^k (1-x)^{n-k}. \end{aligned}$$

At this point a is much smaller than the other terms, and we may solve for $|\mathcal{S}_k|$, round, and solve k#Knapsack:

$$|\mathcal{S}_k| \in \frac{1/2 \pm a/10 - \sum_{j=0}^{k-1} \binom{n}{j} x^j (1-x)^{n-j}}{x^k (1-x)^{n-k}}$$

It remains only to justify that this binary search for x terminates sufficiently quickly. By Equation (7.6) in order to guarantee that $F_Z(m)$ is within $a/10$ of $1/2$ it suffices to guarantee that the binary search step for x has size at most $\frac{a}{10n2^n}$. This requires $\log\left(10n^{n+1}2^n\right)$ steps, which is polynomial in n. □

We are now ready to establish the hardness of our problem.

Proof of Theorem 53. We reduce from WBM. If $k = 1$ then the reduction is immediate: if each of the P_i is a scaled down copy of $\alpha_i \mathsf{Bernoulli}(x)$, then finding the optimal report for the random variable $\sum_i P_i$ amounts to finding the (scaled down) median of $\sum_i \alpha_i \mathsf{Bernoulli}(x)$.

More generally, given an instance of WBM described by $(x, \alpha_1, \ldots, \alpha_n)$, we will construct an instance of our problem, MAE-Estimator, for any $k \geqslant 2$ for which determining optimal partitions and reporting scheme will solve our instance of WBM.

Our P_i will be discrete distributions given by

$$\Pr\left[p_i = \frac{1}{2k}\right] = \frac{1-x}{k} \tag{7.7}$$

$$\Pr\left[p_i = \frac{1 + \delta \frac{\alpha_i}{\sum_t \alpha_t}}{2k}\right] = \frac{x}{k} \tag{7.8}$$

$$\Pr\left[p_i = \frac{2j-1}{2k}\right] = \frac{1}{k} \qquad \text{for } j = 2, \ldots, k\ . \tag{7.9}$$

We will choose δ small enough such that the optimal partition of each of the P_i necessarily groups the atoms described in Equation (7.7) and Equation (7.8) together, and gives each of the atoms of Equation (7.9) its own interval in the partition. To find such a δ, first consider the "good" case when the partitions are of this form. In this case, there are k^n total boxes, each with weight $1/k^n$. Within each box C, the distribution of ℓ_1 norms has range upper bounded by $\delta/(2k)$. Within each C, the range of this distribution is an upper bound on the ℓ_1 distance between any atom in C and the optimal report for C. Therefore, a loose upper bound on total MAE is

$$\sum_{c \in [k]^n} P(C_c) \frac{\delta}{2k} = \frac{\delta}{2k}. \tag{7.10}$$

On the other hand, consider the "bad" case when at least one of the partitions groups either two of the Equation (7.9) atoms together or the Equation (7.8) atom together with at least one of the Equation (7.9) atoms. Assume without loss of generality that the $i = 1$ partitioning is "bad". We will focus on the case when an Equation (7.8) and at least one Equation (7.9) atom are grouped together (because it is an interval, necessarily $j = 2$ is included), since in the best case it is the least costly scenario. Because of the product structure of the boxes induced by the partitions, for every pair of vectors u and u' in the support of P of the form

$$u = \left(\frac{1 + \delta \frac{\alpha_i}{\sum_j \alpha_j}}{2k}, u^- \right) \qquad\qquad u' = \left(\frac{3}{2k}, u^- \right),$$

where $u^- \sim \prod_{j=2}^n P_j$, necessarily u and u' are contained in the same box. Therefore among each pair of u and u', at least $M_{u^-} = \frac{\min\{x,1-x\}}{k} \prod_{j=2}^n P_j(u_j^-)$ mass must travel $\|u'\|_1 - \|u\|_1$ to the estimate for their shared box, which yields a lower bound on the error of

$$\sum_{u^-} \left(\frac{1-\delta}{k} M_{u^-} \right) = \frac{(1-\delta)\min\{x, 1-x\}}{k^2}. \tag{7.11}$$

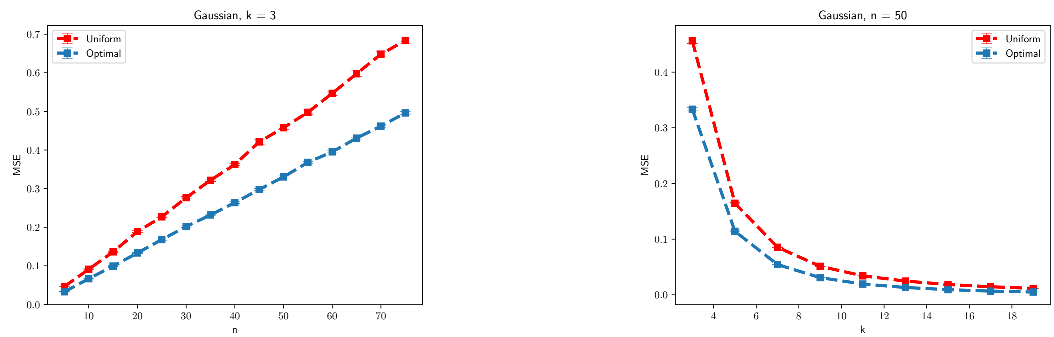
By Equations (7.10) and (7.11), choosing a $\delta < \frac{\min\{x,1-x\}}{k}$ guarantees that the optimal partitioning for our instance is the "good" partitioning, and so all of the Equation (7.7) and Equation (7.8) atoms appear in the same box $C^\star := \prod_i B_{i,1}$.

Recall that by Lemma 21, the MAE-minimizing estimate for a fixed box C is a median of the distribution of ℓ_1 norms of the vectors $u \in C$ according to P. Therefore MAE-Estimator finds some MAE-optimal report $r^\star$ for the box $C^\star$, which by Equation (7.7) and Equation (7.8) implies that $\frac{r^\star - n/2k}{\delta}$ is a median of $\sum_i \alpha_i \mathsf{Bernoulli}(x)$, solving the given instance of WBM. □

7.5 Experiments

In Section 7.4 we showed that when prior distributions are known, the optimal estimator with respect to the MSE can be computed in polynomial time. However, it is reasonable to ask to what degree incorporating these prior distributions leads to more accurate estimation schemes for plausible families of prior distributions. In this section we aim to answer this question. We focus on the MSE as our measure of error, because Theorem 53 shows that an optimal estimator with respect to MAE is hard to compute.

In more detail, we compare the MSE-optimal prior-sensitive estimator of Theorem 52 to the deterministic worst-case optimal strategy described in Section 7.3, which does not incorporate knowledge of prior distributions. This is the uniform estimator, which for all $i \in [n]$ partitions



(a) Fixed $k = 3$, ranging over n

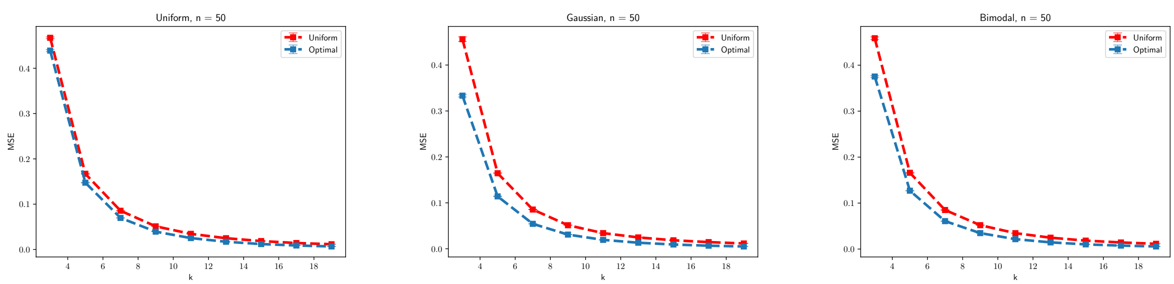
(b) Fixed $n = 50$, ranging over k

Figure 7.1: MSE of the uniform and optimal estimators, averaged over 100 distributions sampled from the Gaussian family. Bars are standard error of the mean.

the ith interval into equal intervals of length $1/k$. It is pointless to use the more elaborate randomized estimator of Theorem 51, because when instances are drawn from a distribution, randomization does not help: for any randomized estimator there is a deterministic estimator that performs at least as well.

We compare the two estimation schemes (optimal and uniform) on discrete distributions drawn from three families: uniform, Gaussian, and bimodal. Recall that in our model prior distributions are comprised of m atoms. For each choice of number of agents n, number of intervals k, and family of distributions, we generate instances as follows: we sample n discrete prior distributions of m uniformly weighted atoms, one distribution per agent. For the family of uniform distributions, each sampled distribution is formed by drawing m samples from $U(0, 1)$; for the Gaussian family, each sampled distribution is formed by drawing m samples from a truncated Gaussian over the domain $[0, 1]$ with mean μ and standard deviation σ drawn i.i.d. from $U(0, 1)$; and for the bimodal family, each sampled distribution is formed by drawing m samples from an equal mixture of two truncated Gaussians over the domain $[0, 1]$, each again with μ and σ drawn i.i.d. from $U(0, 1)$. We then independently sample a number of points from each of the agents' discrete distributions to form a collection of draws $p \sim P$, and evaluate the performance of both the uniform and MSE-optimal estimators on these draws.

Figure 7.1 shows sample averages of MSE for both the uniform and optimal estimators applied to distributions from the Gaussian family, for a range of n and for fixed $k = 3$, or for a range of k and fixed $n = 50$. The MSE for each pair of generated distribution P and estimator is measured as an average over 1000 draws $p \sim P$. For a fixed value of k, as n increases, the optimal estimator significantly outperforms the uniform estimator, suggesting that knowledge of the distribution gives a significant benefit in practice.



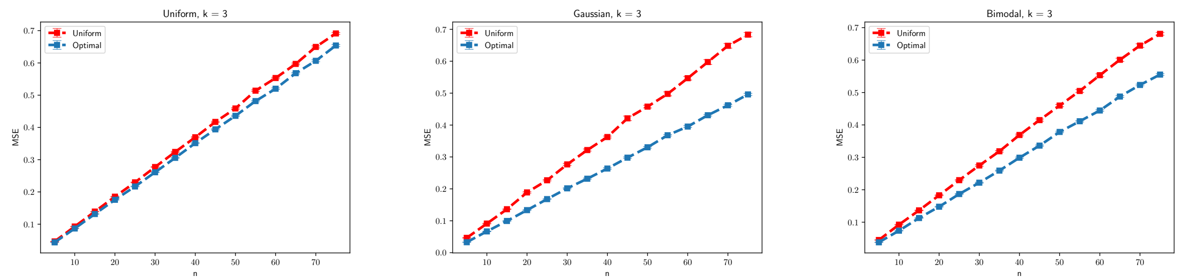
(a) Uniform (b) Gaussian (c) Bimodal

Figure 7.2: MSE of the uniform and optimal algorithms for fixed $n = 50$ and a range of k, averaged over 100 distributions sampled from the various families. Bars are standard error of the mean.

Figure 7.2 shows the estimation schemes' relative performance for a fixed value of $n = 50$ and a range of possible k, with constant sample distribution support of size 100. On the other hand, Figure 7.3 shows their relative performance for a fixed value of $k = 3$ and a range of possible n, again with constant sample distribution support of size 100.

7.6 Discussion

Although we have assumed throughout that $p_i \in [0, 1]$, the results in Section 7.3 can be generalized to general bounded p_i. Similarly, the optimal strategy described in Section 7.4.1



(a) Uniform (b) Gaussian (c) Bimodal

Figure 7.3: MSE of the uniform and optimal algorithms for fixed $k = 3$ and a range of n, averaged over 100 distributions sampled from various families. Bars are standard error of the mean.

holds for prior distributions over unbounded $p_i \in \mathbb{R}$.

There are also promising avenues for future work to extend our results to richer settings. For instance, in Section 7.4 we assume that the prior distributions P_i are given to us. However, it would be interesting to consider the setting in which the P_i are initially unknown but gradually discovered over rounds of questions; i.e., a learning setting where the elicitation scheme learns the prior distributions P_i in the course of accurately estimating the mean of the p_i.

Moreover, while we have focused on estimating $\|p\|_1$, one may ask if it is possible to estimate other functions of p. For instance, can the median of the p_i be efficiently and accurately estimated in this multiple-choice question setting?

Future work may also explore how our results carry over to settings in which the error metric is asymmetric: For instance, it may be more costly to underestimate than overestimate the size of a matriculating class due to space and resource constraints, and an optimal estimator would take this cost asymmetry into account.

Chapter 8

Recruitment Strategies that Take a Chance

8.1 Introduction

Anyone who has served on a faculty hiring committee or a PhD admissions committee knows that a successful outcome requires resolving the tension between two competing goals. On the one hand, some candidates are (perceived to be) better qualified than others, and the aim is to recruit the best candidates. On the other hand, there are a given number of positions to be filled, and while there is typically some flexibility, there is a real cost to recruiting too many or too few people. The tension arises in part because the stronger a candidate is, the more likely they are to receive multiple attractive offers and the less likely they are to accept any particular offer. In order to manage uncertainty, a good strategy may involve a mix of offers to stellar candidates and "safer" candidates.

To formalize this problem, we assume that a recruiting entity (academic or otherwise) has access to two numbers for each candidate i: their value x_i and their probability p_i of accepting an offer. We acknowledge that in current practice, these numbers are not always explicitly estimated. However, committees typically rank or assign numerical scores to candidates based on their strength or fit, and savvy committees roughly estimate recruitment chances by classifying candidates as, say, "high yield," "low yield" or "extremely low yield", for example, by past experience or assistive computational tools [SG06; Abe+15]. Therefore, we believe that the gap between current practice and explicit value and probability estimates is not large.

Our approach builds on the work of Purohit, Gollapudi, and Raghavan [PGR19], who cast hiring under uncertainty as a stochastic optimization problem. In their basic model, there are n candidates (each associated with a value and probability), k positions, and t time steps. In each time step, the algorithm (i.e., recruitment strategy) may make an offer to a single candidate and receive a response; that is, at most t sequential offers can be made, and the budget of k cannot be exceeded. The goal is to maximize the expected value of candidates who accept offers. Purohit, Gollapudi, and Raghavan [PGR19] also consider the setting where the algorithm may make parallel offers in each round. For both problems, they develop polynomial-time, constant-factor approximation algorithms (with approximations ratios of 2 and 8, respectively).

This problem formulation captures key aspects of recruitment, but, in our view, it does have two shortcomings. First, in a sense it is overcomplicated, as computational challenges stem from the assumption that offers are made sequentially. But, in our experience of faculty hiring and PhD admissions in several universities, offers are typically made in one batch. Indeed, delayed offers (in the case of faculty hiring) and waitlists (in the case of PhD admissions) are usually avoided as they negatively impact yield.

The second, and more crucial, shortcoming is that Purohit, Gollapudi, and Raghavan [PGR19] consider the constraint of hiring k candidates as firm. Again, this is inconsistent with our experience: offers are made so that the expected yield roughly matches a desired target, but some faculty hiring or PhD admission cycles are "too successful," in the sense that the number of candidates who accept their offers is much larger than expected. This has a real cost: in the case of PhD students, it creates difficulties in finding funding and advisors, and in the case of faculty hiring, it may precipitate a shortage of resources with long-term impacts on future hiring and even tenure. For example, in one of our institutions, a faculty hiring cycle with yield that was much higher than expected led to the cancellation of the subsequent year's search.

Let us, therefore, reformulate the problem of hiring under uncertainty in a way that avoids both issues. We assume that offers are made in a single batch, and the target number of positions k is a soft constraint. Specifically, a penalty is incurred for deviating from the target number of positions; we consider several different options for this penalty function. The optimization problem is this:

> *Select a subset S of candidates that maximizes overall expected reward $\sum_{i \in S} p_i \cdot x_i$, minus expected penalty for deviating from the target number of positions.*

Enumerating all possible subsets S may be practicable for small instances, for example in the case of faculty hiring in small departments. However, a brute-force approach will not work for this purpose in larger departments, or at the scale of PhD admissions even in smaller programs, which motivates our search for good algorithms.

Our results

We first consider a simplified case where the goal is to solely minimize the penalty term of our objective (irrespective of the rewards), and show that the greedy algorithm that selects candidates in decreasing order of their probabilities is optimal (Section 8.3.1).

The full objective is considerably more complex, and we analyze it under two natural penalty functions. When the penalty function is the squared error from the target, we show that the optimization problem is weakly NP-hard, and provide a fully polynomial-time approximation scheme (FPTAS). When the penalty is linear in the extent to which the target is exceeded (that is, a linear penalty is incurred by overshooting, but not by undershooting), we show that two greedy heuristics — picking in the decreasing order of the value x_i and the expected value $p_i x_i$ — provide approximations to the optimal solution that are polynomial in the minimum probability $p_{\min}$. We then present a constant-factor approximation algorithm that runs in polynomial time for fixed $p_{\min}$ and candidate value relative to overshooting penalty, thereby improving upon the greedy heuristics.

Finally, we carry out experiments on synthetically generated data (Section 8.4), focusing on the linear penalty incurred by overshooting. We observe that the two greedy heuristics perform reasonably well, especially if the values and probabilities are positively correlated. At the same time, compared to the greedy heuristics, our constant-factor approximation algorithm better adapts to specific instances, especially when this correlation is negative. These numerical experiments corroborate our theoretical results that the greedy heuristics provide reasonable guarantees, justifying their use in practice for both simplicity and good performance. However, in many regimes the additional flexibility of the constant-factor approximation algorithm is likely worth its complexity overhead.

Related work

In practice, the challenge of uncertainty in admissions is mitigated by practices such as admitting students in multiple rounds, using a waitlist, and using a rolling process [LW20]. There are also assistive computational tools that predict student yield rate with machine learning [SG06; Abe+15]. There are a few theoretical formulations that model and address the uncertainty in such problems. As previously mentioned, Purohit, Gollapudi, and Raghavan [PGR19] consider an online setting where in each time step if a candidate is given an offer then their decision is revealed immediately, and analyze the optimal ordering to give offers to the candidates subject to a hard constraint on the total number of acceptances. Ganguly, Basu, and Nagarajan [GBN22] consider a setting with multiple rounds where the yield rate in each round is either q_L or q_H (with $q_L < q_H$). To model the negative correlation between the candidate quality and the probability of acceptance, they assume that the probability of the yield rate being q_H is linear in the number of students given offers. They subsequently derive a decision tree that computes the number of offers to make in each round. For single batch selection, Zhang and Pippins [ZP21] analyze the optimal number of applicants to admit using techniques from yield management, under the assumption that each applicant has identical value and probability. A distinct line of work casts the admissions problem as a decentralized matching market [CK16; DJ21], where the uncertainty in acceptance is modeled by the students' stochastic preferences over multiple schools. An objective combining the utility and the accepted size is considered, but the penalty is in terms of the expected size, and does not consider variance. Another line of work analyzes various metrics for the secretary problem under uncertain offer acceptance [Smi75; Tam91; PST21], where candidates appear sequentially in a uniformly random ordering in a fleeting nature. Here the reward is defined based on the ranking of the candidate accepting the offer relative to other candidates in the sequence, and the goal is to decide whether to make an offer to the candidate appearing in each time step.

Our proposed formulation is closely related to the knapsack problem and its many variants. In the stochastic knapsack problem, each item has a deterministic value and an independent stochastic size; the actual size of an item is revealed only after it is selected [KRT00; DGV04; BGK11]. For the one-shot version of this problem, the aim is to choose a subset maximizing the expected value of the realized items, such that the probability that these realizations violate the knapsack constraint is below some threshold. In contrast, our "items" have equal size, but we pay a penalty which is some function of our realized distance from our knapsack target size, exchanging the constraint for a mixed-sign objective. In this spirit, the one-sided loss functions we consider are similar to objectives which arise in the penalty method for solving constrained optimization problems [Ols94], though our setting is stochastic and we do not introduce the penalty term in service of ultimately satisfying a hard constraint.

8.2 Problem Formulation

Taking a knapsack perspective, consider some n items (corresponding to candidates) with associated values $x_1, \ldots, x_n \in \mathbb{R}$. If we select an item $i \in [n]$, there is a probability $p_i \in [0, 1]$ that we receive this item (the candidate accepts the offer). We write these values and probabilities as vectors $x \in \mathbb{R}^n$ and $p_i \in [0, 1]^n$. Let $Z_i \in \{0, 1\}$ be the indicator variable that we receive item i if it is selected, so that $Z_i \sim \mathrm{Ber}(p_i)$. We assume the events that we receive each individual item are independent. Let $S_Z \subseteq S$ denote the random realization of chosen items; that is, $S_Z \coloneqq \{i \in S : Z_i = 1\}$. Our goal is to select a subset $S \subseteq [n]$. First, we consider the reward for a subset S as the expected total value obtained:

$$R(S) \coloneqq \mathbb{E}\left[\sum_{i \in S} Z_i x_i\right] = \sum_{i \in S} p_i x_i.$$

At the same time, let $M \in \mathbb{N}_+$ denote a target size that we want the realized set S_Z to achieve. We want to control the expected deviation of the realized size of S_Z, which is $|S_Z| = \sum_{i \in S} Z_i$, from the target size M. We consider this penalty as

$$V(S) \coloneqq \mathbb{E}\left[\rho\left(|S_Z|,\ M\right)\right],$$

where $\rho \colon \mathbb{N} \times \mathbb{N}_+ \to \mathbb{R}_{\geqslant 0}$ is a loss function, to be specified later. Combining the two parts, we define the overall objective as

$$U(S) \coloneqq R(S) - \lambda \cdot V(S), \tag{8.1}$$

where $\lambda \in \mathbb{R}_+$ is a hyperparameter that governs the importance of the penalty relative to the reward. Our goal is to find the subset that maximizes the overall expected utility:

$$S^* \in \operatorname*{arg\,max}_{S \subseteq [n]} U(S).$$

We denote a problem instance by $\mathcal{I} \coloneqq (x, p, M, \lambda)$, and the solution S^* is thus a function of the problem instance and the loss function ρ. It is worth noting that since the overall utility U is a mixed-sign objective, the optimal value of $U(S^*)$ may be negative depending on the instance and choice of ρ.

We consider a range of choices for the loss function ρ. Given a target size M, it is natural for ρ to be a convex function minimized at M, which penalizes any deviation from M, or alternatively a monotone convex function that is nonzero above M, which can be seen as penalizing violation of a budget constraint. We focus in particular on one- and two-sided linear and quadratic losses, which are formally introduced in Section 8.3.2 below.

8.3 Theoretical Results

To begin we note that if we only consider the reward term and set the penalty term to be $V(S) := 0$, then the solution is to trivially select all items. In what follows, we first discuss the other extremal case, taking $R(S) := 0$ and considering the penalty V in isolation. These may be viewed as the extreme cases when $\lambda = 0$ and $\lambda \to \infty$. We will then turn to the general objective and consider both terms jointly.

8.3.1 Warm-Up: Penalty Only

To gain intuition for this problem, we start with the simplified case in which our goal is only to minimize the penalty term. Note that in this case our objective is strictly nonpositive.

Algorithm 10 pGreedy

Require: $p \in [0,1]^n$

$S \leftarrow \emptyset$
Sort $\{p_i\}_{i\in[n]}$ in decreasing order and re-index the items such that $p_1 \geqslant \ldots \geqslant p_n$
for $i = 1, 2, \ldots, n$ **do**
  if $U(S \cup \{i\}) \geqslant U(S)$ **then**
    $S \leftarrow S \cup \{i\}$
  else
    break
return S

We consider pGreedy, the greedy algorithm with respect to p_i (Algorithm 10). In words, pGreedy selects items in their decreasing order of probabilities, with ties broken arbitrarily if there are multiple items with the same probability.[1] The algorithm keeps selecting the next item defined by this order, and terminates when adding the next item would decrease the objective. This greedy algorithm is computationally efficient, since the stopping criterion in algorithm 10 can be checked in polynomial time given access to ρ (see Lemma 30 in Section 8.7.8 for details). Surprisingly, pGreedy is optimal for minimizing $V(S)$ in isolation.

Proposition 2. *Let $M \in \mathbb{N}_+$ be any target size. If the loss function $\rho(\,\cdot\,, M)$ is convex, then* pGreedy *(Algorithm 10) yields an optimal solution to minimizing the penalty* $\min_{S \subseteq [n]} V(S)$.

The proof of this proposition is provided in Section 8.7.3. This result is not obvious, as one might expect that as the sum of probabilities of all selected items so far approaches the target size, it may be better to select an item with lower probability than an item with higher probability to "fill the gap." This is not true. Intuitively, it is because the realization of each acceptance Z_i is binary, so the outcome of adding another item i into the selection is either we add this item (with probability p_i) or not (with probability $1 - p_i$). If adding this item gives lower penalty, then we desire to add the item with the highest probability possible.

8.3.2 The General Objective

We now turn to the general objective. At the outset it bears noting that $U(S)$ is submodular in S whenever $\rho(\cdot, M)$ is convex, as is the case for the loss functions we consider. Unfortunately the existing body of work on (non-monotone) submodular maximization cannot be leveraged to obtain a general-purpose approximation to $U(S)$, since U is mixed-sign and may be negative even at optimality, and applying an affine transformation in order to engineer nonnegativity will generally destroy any approximation guarantees.

We focus on a few natural choices of the loss function. First, we consider ρ given by linear

[1] In practice it is natural to break ties in favor of items of higher value, though this does not affect our results.

and quadratic losses, which we denote by L_1 and L_2 respectively. These yield penalty terms $V(S)$ which are equal to the mean average error (MAE) and mean squared error (MSE) for the realized size of the subset S_Z. We also consider the corresponding one-sided losses, defined by

$$L_1^+(|S_Z|, M) := \begin{cases} |S_Z| - M & \text{if } |S_Z| \geqslant M \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad L_2^+(|S_Z|, M) := L_1^+(|S_Z|, M)^2.$$

All of these losses considered penalize the case where the realized size is greater than the target size M. In applications such as admissions and hiring, there is a limited, pre-specified amount of resources allocated to the newly admitted or hired people. Hence, having more people than the target size is not desired. At the same time, the two-sided losses give explicit preference that the realized size should also not be smaller than the target size. This explicit penalty for undershooting could represent a hit to morale (an unsuccessful recruitment cycle really is demoralizing) or insufficient staffing for required tasks, such as teaching certain courses. The one-sided loss functions may still to some extent capture these considerations, as there is an implicit opportunity cost described by the reward term when fewer candidates accept.

An FPTAS for L_2 Loss

Given that we understand our problem in both extremal cases (when only considering the reward term or the penalty term), one might hope that some interpolation between them could solve the general case. However, the general case is more complicated. Recall that for the penalty-only objective, pGreedy in Algorithm 10 attains the optimal selection, by adding items in decreasing order of p_i, and terminates once the next item strictly decreases the objective. But pGreedy is clearly ill-suited to the general objective, since it does not take values into consideration. We now present two more natural greedy heuristics analogous to pGreedy, and show that they are provably not optimal for the general objective. Specifically, we consider:

- xGreedy: adds items in decreasing order of their value x_i, and terminates once the next item in this order strictly decreases the objective.
- xpGreedy: adds items in decreasing order of their expected reward $x_i p_i$, and terminates once the next item in this order strictly decreases the objective.

Despite these heuristics appearing intuitive, they perform in a certain sense arbitrarily poorly even for the (squared) L_2 loss, as formalized by the following result.

Proposition 3. *Consider the two-sided loss $\rho = L_2$ and any $\lambda > 0$. Then for* pGreedy*,* xGreedy *and* xpGreedy*, there exists an instance such that the algorithm selects $S \subseteq [n]$ for which $U(S) \leqslant 0$, while $U(S^*) > 0$.*

The proof of this proposition is provided in Section 8.7.4, and we now provide an informal description of the instances constructed. Since pGreedy does not take into account the item values x_i at all, it is natural to expect that pGreedy is not suitable for the general objective. Specifically, we consider two items where one item has probability 1 and value 0, and the other item has a "good" probability less than 1 and a "good" value. Then pGreedy selects the first item, whereas selecting the second item yields a positive objective value. For xGreedy and xpGreedy, we consider two items that have almost the same expected reward. We let item 1 has a probability of 1. We let item 2 have a slightly greater expected reward for tie-breaking, and let item 2 have a smaller probability. In this case, xGreedy and xpGreedy start by picking item 2, which introduces nontrivial variance. When λ becomes large, this variance drives the overall objective negative. On the other hand, picking item 1 yields a strictly positive objective.

Despite the failure of heuristic approaches, when the chosen loss function is $\rho = L_2$ our problem can in fact be approximated up to negligible additive error. For this loss, our full objective (8.1) may be written as

$$U(S) = \sum_{i \in S} p_i x_i - \lambda \cdot \mathbb{E} \left(\sum_{i \in S} Z_i - M \right)^2 . \tag{8.2}$$

Letting $b_i \in \{0, 1\}$ be the binary decision variable for whether $i \in S$, that is, $b_i := \mathbb{1}\{i \in S\}$, the optimization problem then becomes

$$\operatorname*{arg\,max}_{S \in [n]} U(S) = \operatorname*{arg\,max}_{b \in \{0,1\}^n} \quad \sum_{i \in [n]} b_i p_i x_i - \lambda \cdot \mathbb{E} \left(\sum_{i \in [n]} b_i Z_i - M \right)^2 . \tag{8.3}$$

Expanding (8.3) yields a collection of terms which are constant, linear, and quadratic in b_i, and so the objective can be reformulated as an unconstrained binary quadratic program (UBQP). Although UBQP is strongly NP-hard, Çela, Klinz, and Meyer [ÇKM06] present a pseudo-polynomial time algorithm for UBQP when the coefficient matrix for the quadratic form in the objective has constant rank. By proving that the objective (8.2) is sufficiently insensitive to small changes in our problem parameters, we leverage this pseudo-polynomial time algorithm to derive a FPTAS for approximating the optimal objective value for our problem. A standard search-to-decision reduction then yields the following result.

Theorem 54. *For $\rho = L_2$, Algorithm 13 identifies some $S \subseteq [n]$ satisfying $U(S) \geqslant U(S^*) - \epsilon$ in time poly$(1/\epsilon, n, M, \lambda)$.*

Pseudocode describing Algorithm 13 and the proof of this theorem are provided in Section 8.7.5. On the other hand, by a reduction from equipartition we have the following hardness:

Theorem 55. *For $\rho = L_2$, optimizing $U(S)$ is weakly NP-hard.*

The proof of this theorem is provided in Section 8.7.6. The hardness landscape of our problem when $\rho = L_2$ is therefore similar to that of the knapsack problem, which is heartening since the knapsack problem is relatively tractable in practice. However, unlike the knapsack problem, we should only hope for additive rather than multiplicative guarantees; this is because for $\rho = L_2$ our optimal value is not bounded away from zero and may be strictly negative, even if all values x_i are nonnegative.

In contrast to L_2, we find that the one-sided loss L_2^+ is not straightforward to analyze. In this case the objective does not admit a quadratic factorization in terms of decision variables, and although the objective is nonnegative, it is difficult to analyze the performance of the greedy algorithm or contend with the nonlinearity of the loss function in a principled way. We instead turn to L_1^+ loss, where surprisingly these obstacles can be overcome.

Approximations for L_1^+ Loss

The loss $\rho = L_1^+$ enables the possibility of a multiplicative approximation because the optimum value $U(S^*)$ of the mixed-sign objective is nonnegative. More generally, any ρ with $\rho(0, M) = 0$ has a nonnegative optimal value, since in this case $U(\emptyset) = 0$. For $\rho = L_1^+$, choosing any single item i with positive x_i and p_i has strictly positive objective, since $M \geqslant 1$ implies that $L_1^+(1, M) = 0$ and so $U(\{i\}) = p_i x_i$. More generally, for this loss it suffices to consider only i for which $x_i > 0$, since under this loss the marginal contribution of any i with $x_i \leqslant 0$ is nonpositive.

The L_1^+ loss may appear amenable to a greedy algorithmic approach, since early items incur no penalty and the marginal penalty of adding a later item i simply turns out to be proportional to the probability that the current solution exceeds the target M. However, as in the case for the L_2 loss, these natural heuristics fail to consider the relation between items in the selection. While efficient, these greedy algorithms perform arbitrarily badly compared to the optimal solution in the worst case. The failure of pGreedy is again apparent as in the case for the L_2 loss. We now provide some informal "bad" instances for xGreedy and xpGreedy for intuition.

Recall that xGreedy chooses items in decreasing order by value. Consider $M = 1$, and consider two types of items with $(x_1, p_1) = (1, p)$ and $(x_2, p_2) = (0.5, 1)$. xGreedy picks item 1 and yields an objective of $O(p)$, whereas picking item 2 yields a constant objective. The other greedy algorithm xpGreedy chooses items i in decreasing order of their expected value

$x_i p_i$. We consider the instance with $M = 1$, and two types of items $(x_1, p_1) = (1 + \epsilon, 1)$ and $(x_2, p_2) = (1/p, p)$ with some tiny $\epsilon > 0$ so that xpGreedy chooses an item from type 1 and yields a constant objective. On the other hand, choosing $\frac{1}{p}$ copies of item 2 yields an objective of $\Omega(1/p)$.

For both xGreedy and xpGreedy, the constructed instances yield an upper bound on the approximation ratio, scaling as the minimum probability $p_{\min} := \min_{i \in [n]} p_i$ associated with the items. It suggests that $p_{\min}$ is a natural parameter for measuring the complexity of an instance with respect to the L_1^+ loss. Another natural parameter is $x_{\max} := \max_{i \in [n]} x_i$, the maximum value among all items. Surprisingly, the performance of these greedy algorithms can be lower-bounded in terms of $p_{\min}$ as well, under an additional assumption about the values of items relative to λ.

Theorem 56. *Consider the one-sided loss $\rho = L_1^+$. If there is some fixed constant $c > 0$ such that $x_{max} \leqslant (1 - c) \cdot \lambda$, then*

(a) There exist instances for which pGreedy *selects S with $U(S) = 0$, while $U(S^*) > 0$.*

(b) The worst-case approximation ratio for xpGreedy *is $\Theta(p_{min})$.*

(c) The worst-case approximation ratio for xGreedy *is $\Omega(p_{min}^2)$ and $O(p_{min})$.*

The proof of this theorem is provided in Section 8.7.7. As a consequence, the approximation ratios of these greedy algorithms can be arbitrarily small as $p_{\min} \to 0$. Also note that the upper bounds do not require the assumption that $x_i \leqslant (1 - c)\lambda$; in all three cases there exist bad instances irrespective of this assumption. The reason we can derive better lower bounds for xpGreedy than xGreedy is intuitive; this is because xpGreedy measures the expected reward conferred by an item, and so can be more directly related to the optimal solution S^*.

This is in notable contrast to when $\rho = L_2$, where we saw that multiplicative guarantees are inapt and all three greedy algorithms may incur arbitrarily large additive loss. Theorem 56 also raises a natural question: is this $p_{\min}$ upper bound tight, or is it possible to efficiently attain better approximations to $U(S^*)$ which do not depend on $p_{\min}$? We address this by introducing $\textsc{OneSided}\mathrm{L}_1^+$, presented in Algorithm 11, which attains a constant-factor approximation to the optimal solution. In this pseudocode, for any vector $v \in \mathbb{R}^n$ and set $S \subseteq [n]$, we use $v|_S$ to denote the $|S|$-dimensional vector obtained by restricting v to its coordinates indexed by S. Its runtime is parameterized by the minimum probability $p_{\min}$ and the ratio of the maximum value $x_{\max}$ to the penalty parameter λ; in particular for fixed $p_{\min}$ and $\lambda/x_{\max}$ it is polynomial in n.

At a high level, $\textsc{OneSided}\mathrm{L}_1^+$ proceeds first by dividing the items into three groups according to their values x_i, and considering each group in turn. Since U is submodular (see Lemma 26 in Section 8.7.1), the optimal solution within at least one of these groups is also constant-competitive with $U(S^*)$. We obtain a constant-factor approximation for each group in a different way.

For the items with low values bounded away from λ, $\textsc{LowValue}\mathrm{L}_1^+$ (Algorithm 14 in Section 8.7.8) checks all small subsets, which succeeds if the optimal subset for this group is small. It also computes rounded probabilities and values for each item in this group, and then efficiently computes the optimal solution according to this rounded instance. If the optimal subset is large, we then prove that this search over rounded solutions necessarily identifies a subset with objective value comparable to that of the optimal subset. This is the technical crux of proving that $\textsc{OneSided}\mathrm{L}_1^+$ is a constant-factor approximation.

For the items with values just below λ, $\textsc{MediumValue}\mathrm{L}_1^+$ (Algorithm 15 in Section 8.7.8) returns the optimal subset if the group is small. If the group is large, it tries to choose a subset such that the expected number of realizations is about M; if there are not enough items, it chooses a subset with approximately half the expected number of realizations of the group overall. Finally, for the group of items with values above λ, it is straightforward to see that choosing the entire group is optimal. The pseudocode and related proofs for these algorithms appear in Section 8.7.8. The following result provides a theoretical guarantee for Algorithm 11.

Theorem 57 (Constant-factor approximation for L_1^+)**.** *Algorithm 11 is a constant-factor*

Algorithm 11 $\textsc{OneSidedL}_1^+$

Require: Problem instance $\mathcal{I} = (x, p, M, \lambda)$
Ensure: $S \subseteq [n]$ for which $U(S) \geqslant c \cdot U(S^*)$ for universal constant c
1: $N_L \leftarrow \{i \in [n] \colon x_i \leqslant (1 - \frac{p_{\min}}{4}) \cdot \lambda\}$
2: $N_M \leftarrow \{i \in [n] \colon (1 - \frac{p_{\min}}{4}) \cdot \lambda < x_i < \lambda\}$
3: $N_H \leftarrow \{i \in [n] \colon x_i \geqslant \lambda\}$
4: $S_L \leftarrow \textsc{LowValueL}_1^+(x|_{N_L}, p|_{N_L}, \lambda, M)$
5: $S_M \leftarrow \textsc{MediumValueL}_1^+(x|_{N_M}, p|_{N_M}, \lambda, M)$
6: $S_H \leftarrow N_H$
7: Compute $U(S_L)$, $U(S_M)$, and $U(S_H)$
8: **return** $S \in \{S_L, S_M, S_H\}$ maximizing $U(S)$

approximation to $U(S^*)$ which runs in time $n^{O\left(\frac{1}{p_{min}^2}\max\left\{1, \log\left(\frac{1}{p_{min}}\right), \log\left(\frac{\lambda}{x_{max}}\right)\right\}\right)}$.

The proof of this theorem is provided in Section 8.7.8. Intuitively, the reason Algorithm 11 divides the items into cases depending on their values is to handle the case when the reward portion of $U(S^*)$ is almost equal to the penalty portion. This presents an impediment to the performance of solving a rounded version of the instance, since in this case the magnitude and even the sign of $U(S^*)$ is potentially quite sensitive to changes in p_i and x_i. By restricting attention to items with low values, we prove that the expected number of realized items in S^* is not much more than the target M. This then allows us to argue that there exist good rounded solutions that can be efficiently identified.

We conclude our theoretical results with a surprising equivalence between one- and two-sided linear losses L_1^+ and L_1. In what follows, we use U_{L_1} and $U_{L_1^+}$ to denote the objective with $\rho = L_1$ and $\rho = L_1^+$, respectively. We use $U(S; \mathcal{I})$ to denote the evaluation of $U(S)$ specifically with respect to the instance $\mathcal{I} = (x, p, \lambda, M)$.

Theorem 58 (Equivalence between L_1 and L_1^+)**.** *For any instance $\mathcal{I} = (x, p, \lambda, M)$, construct $\mathcal{I}' = (x', p, \lambda', M)$ given by $x_i' := x_i - \lambda$ and $\lambda' := 2\lambda$. Then for all $S \subseteq [n]$,*

$$U_{L_1}(S; \mathcal{I}) = U_{L_1^+}(S; \mathcal{I}') - \lambda \cdot M.$$

The proof of this theorem is provided in Section 8.7.9. In particular, since λ and M do not depend on S, this implies that S maximizes U_{L_1} on instance $\mathcal{I}$ if and only if it maximizes $U_{L_1^+}$

on instance $\mathcal{I}'$.

Although Theorem 58 establishes a correspondence between the solutions to our problem for $\rho = L_1$ and $\rho = L_1^+$, it is not approximation preserving, so it does not convert $\textsc{OneSided}\mathrm{L}_1^+$ into an approximation algorithm for the two-sided setting. Indeed, as in the $\rho = L_2$ setting, the optimal value when $\rho = L_1$ can be strictly negative.

8.4 Experiments

Having established worst-case theoretical guarantees, we wish to test how well our algorithms perform empirically. We focus on L_1^+ loss because our result for L_2 is an FPTAS, so we know its performance can be made arbitrarily close to optimal. Specifically, the experiments benchmark the subroutine $\textsc{LowValue}\mathrm{L}_1^+$ (part of $\textsc{OneSided}\mathrm{L}_1^+$) against $\textsc{xGreedy}$ and $\textsc{xpGreedy}$ for the regime where $x_i \leqslant (1-c)\lambda$. This is the regime which $\textsc{LowValue}\mathrm{L}_1^+$ was developed to handle for $\textsc{OneSided}\mathrm{L}_1^+$, and it is the regime for which we prove performance guarantees for $\textsc{xpGreedy}$ and $\textsc{xGreedy}$. In Section 8.5, we also provide comparison to the optimal solution for smaller instances (Section 8.5.1), and compare $\textsc{xGreedy}$ to $\textsc{xpGreedy}$ under other losses (Section 8.5.2). All error bars shown in the plots represent standard error of the mean. The code to reproduce our simulation results is available at this github repository.

8.4.1 Experimental Setting

In constructing instances we follow the approach of Purohit, Gollapudi, and Raghavan [PGR19] in their use of beta distributions to orchestrate different kinds of correlation between x_i and p_i. We therefore first draw $x_i \sim \mathrm{Unif}[0,1]$, and then produce three types of correlation as follows:

- *Negative correlation:* $p_i \sim p_{\min} + (1-p_{\min}) \cdot \mathrm{Beta}(10(1-x_i),\ 10x_i)$.

- *Positive correlation:* $p_i \sim p_{\min} + (1-p_{\min}) \cdot \mathrm{Beta}(10x_i,\ 10(1-x_i))$.

- *No correlation*: $p_i \sim \mathrm{Unif}[p_{\min}, 1]$.

This construction differs from the sampling paradigm of Purohit, Gollapudi, and Raghavan [PGR19] only in that we re-normalize the probabilities $\{p_i\}$ so that they are bounded in $[p_{\min}, 1]$. We consider $n = 50$ and $p_{\min} = 0.01$ throughout, and explore the greedy heuristics xGreedy and xpGreedy, as well as the constant-factor approximation algorithm OneSidedL$_1^+$ (Algorithm 11), for a range of M and λ. This lower bound on $p_{\min}$ ensures that the performance of the greedy heuristics and runtime of our algorithm are reasonable; a value of 0.01 (say) is realistic because in practice, if a candidate takes the time and effort to apply, it is reasonable to assume that they at least have some nontrivial probability to accept if they were given an offer. We also focus on the regime where $x_i < \lambda$, which is assumed by Theorem 56 and handled in Algorithm 11 by the subroutine LowValueL$_1^+$. We believe that this is the main regime of practical interest: candidates with $x_i \geqslant \lambda$ are beneficial regardless of how many candidates have already accepted offers, and one might suppose that such candidates are rare.

Note that our theoretical guarantees in Theorem 57 necessitate that all candidate solutions up to size $\tau = \widetilde{O}(1/p_{\min}^2)$ are checked by brute force. In this implementation of LowValueL$_1^+$ we take $\tau = 0$ and isolate its search over rounded solutions. As we consider small target sizes M, this prevents LowValueL$_1^+$ from outperforming the greedy algorithms simply by virtue of having considered every relevant solution. This only hinders the performance of LowValueL$_1^+$. This implementation additionally only considers rounded solutions S satisfying $\sum_{i \in S} x_i p_i \leqslant 2M$, which improves its runtime and theoretically only hinders its performance relative to LowValueL$_1^+$. So long as $x_i < \lambda$ and small solutions are checked, such a stopping condition can be implemented without hindering the performance of LowValueL$_1^+$; for details, see Section 8.7.8 and Lemma 27 in Section 8.7.1. We believe this provides a favorable tradeoff between runtime and accuracy, and illustrates a lower bound on the performance of LowValueL$_1^+$ as written.

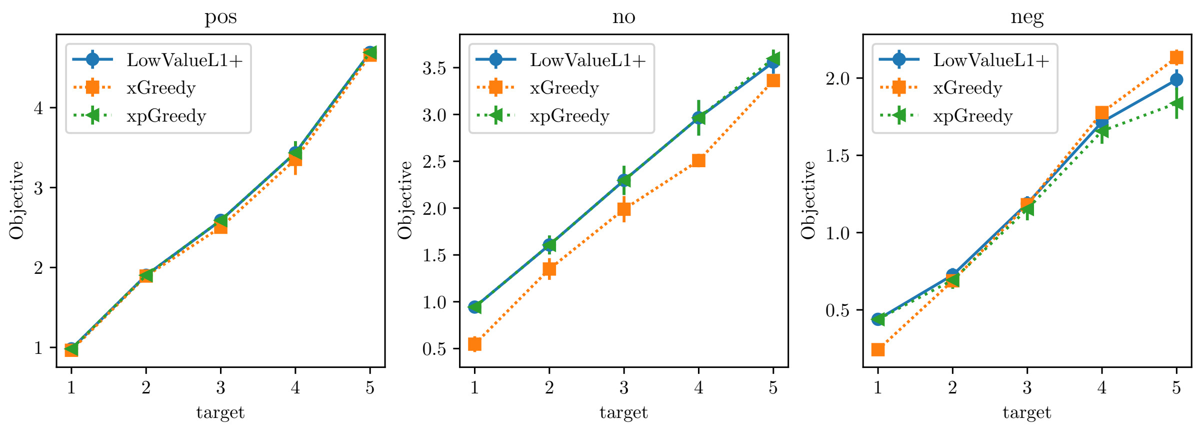


Figure 8.1: Sampling from the beta distribution with positive, no, and negative correlation. Here $n = 50$ and $\lambda = 3$.

8.4.2 Experimental Results

The objective values that our algorithms of interest attain for these distributions are shown in Figure 8.1. Note that positive correlation leads xGreedy and xpGreedy to pursue very similar (and optimal) strategies, as expected. This is intuitively the easier setting, and here LowValueL_1^+ performs on par with the greedy heuristics. In the no-correlation and negative-correlation settings, there are regimes where one of the two greedy heuristics performs better than the other one, whereas LowValueL_1^+ appears to perform as well as the better of two depending on the regimes, showing its better adaptivity across these instances in practice as well as in theory.

Negative correlation between x_i and p_i is of particular interest to us, since it seems most relevant for the setting of faculty hiring and PhD admissions, and in fact hiring and recruitment more broadly. In Figure 8.1, we also observe that negative correlation is the setting that displays the most heterogeneity in algorithm behavior. We therefore turn to this negative-correlation setting and explore the effect of increasing the penalty regularizer λ in Figure 8.2. In general, LowValueL_1^+ appears comparable to the better of the two greedy

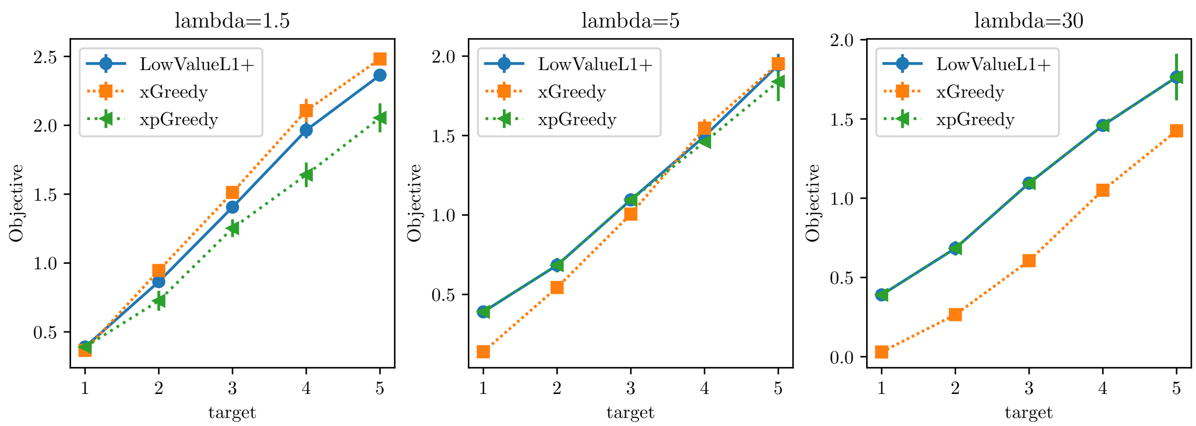


Figure 8.2: Performance for increasing penalty regularizer λ. Here $n = 50$ and sampling is via the negatively correlating beta distribution.

heuristics across the values of M and λ that we examine, though there is a small gap between the objectives achieved by LowValueL$_1^+$ and xGreedy when $\lambda = 1.5$.

This is also good news for xGreedy and xpGreedy, because it suggests that the two of them together remain competitive across a wide range of instances. To the extent that LowValueL$_1^+$ falls short of the objective achieved, it is due to systematically rounding the probabilities p_i up by a constant factor when computing the prospective utility of solutions. Because its rounding preserves the reward term, such a systematic overestimate in p_i leads it to overestimate the penalty term of any set under consideration. The impact of rounding may be small on each individual item but collectively large on the objective, and therefore explains the extent to which LowValueL$_1^+$ lags behind xGreedy in Figure 8.2; the latter chooses many such items while the former judges their influence on the penalty to be too large. However, this can be mitigated by choosing smaller multiplicative bucket sizes for LowValueL$_1^+$ in rounding, which is particularly effective in the case where the probabilities $\{p_i\}$ of an instance fall in a small number of clusters or exhibit other structure.

8.5 Additional Experiments

In this section, we present additional experiments which shed more light on the performance of xGreedy, xpGreedy, and OneSidedL_1^+ relative to one another and to the optimal solution, for a broader range of objectives. The family of distributions from which we sample instances is the same as the one described in Section 8.4.1.

8.5.1 Comparison to Optimal

First, we recreate Figure 8.1 and Figure 8.2, now including the objective value of the optimal solution S^* as a benchmark for the three algorithms considered above. Since determining $U(S^*)$ by brute force is computationally costly, this comparison is undertaken for smaller instances $(n = 20)$. Following Section 8.4.1, we consider $\lambda = 3$.

Here Figure 8.3 shows the performance of xGreedy, xpGreedy, and OneSidedL_1^+ relative to the objective $U(S^*)$ of the optimal solution, when values and probabilities are positively correlated, uncorrelated, and negatively correlated, for a range of target sizes M. Figure 8.4 shows the performance of xGreedy, xpGreedy, and OneSidedL_1^+ relative to $U(S^*)$ as the penalty regularizer increases, for negatively correlated x_i and p_i and again for a range of target sizes M.

It is noteworthy that in both Figure 8.3 and Figure 8.4, the best algorithms in each setting nearly attain the optimal objective value. It is unclear the extent to which we should expect that this continues to hold for larger instances, where solving the optimal solution by brute force is computationally infeasible.

8.5.2 Other Objectives

In Section 8.4.2 and Section 8.5.1, we examine the performance of different algorithms for the L_1^+ loss, since this is the loss function for which we derive worst-case multiplicative guarantees and for which the algorithm ONESIDEDL_1^+ was designed.

We now investigate how these algorithms perform with respect to other loss functions, despite the absence of worst-case theoretical guarantees for the greedy heuristics. Figure 8.5 compares the greedy heuristics between the L_1^+ and L_2^+ loss objectives across different correlation regimes. Figure 8.6 does the same for the two-sided L_1 and L_2 loss objectives.

In Figure 8.5 the performance of both greedy heuristics is very similar under the two one-sided losses. For the two-sided losses L_1 and L_2, Figure 8.6 suggests that xpGreedy dramatically outperforms xGreedy across the choice of the two-sided losses. We observe that the objective values are no longer uniformly positive, and are no longer monotonically increasing in the target size. This is because the problem under the two-sided losses is fundamentally more difficult: under one-sided losses, only selecting over the target is penalized; it is straightforward to observe that selecting M items always yields a penalty of 0 and hence a positive objective value. Under two-sided losses, selecting under the target and selecting over the target is both penalized; there is also non-zero variance towards achieving the exact target M, and hence the objective is negative when the regularizer λ is large.

Comparing the two-sided losses L_1 and L_2 in Figure 8.6, the problem under the L_2 loss is more difficult due to its higher penalty (the quadratic function always attains a higher value than the linear function on integers). The objective starts decreasing as a function of the target M: if we are aiming at a larger target M, more items are selected, leading to an inevitable increase in the variance and hence a lower objective.

We provide an informal explanation for the superior performance of xpGreedy over xGreedy for the two-sided losses, using the two-sided L_2 loss as an example. Under this loss, a candidate i contributes $x_i p_i$ to the reward term of the objective, while contributing $p_i(1 - p_i)$ to the variance of the realized size. When faced with two candidates of equal value x_i, we should therefore at the margin prefer the candidate with the higher probability, since this candidate contributes less to the variance per contribution to the reward. Note that for sufficiently large λ, two-sided losses encourage algorithms to choose solutions whose expected size is very close M, meaning that the variance and the penalty term are nearly equal. Here xpGreedy prefers this higher-probability candidate, while xGreedy is indifferent, explaining the superior performance of xpGreedy.

8.6 Discussion

One of the takeaways from our theoretical and empirical results is that, in addition to xpGreedy, the greedy algorithm xGreedy, which makes offers to a subset of candidates with the highest values, is practicable for L_1^+ loss. This is intriguing because the algorithm is quite similar to how faculty hiring and admissions committees typically think: they want to make offers to the best candidates. The difference is that xGreedy carefully selects the *number* of offers to be made, in a way that greedily maximizes the objective. Since xGreedy amounts to a relatively small tweak to current practice, we believe committees would find the algorithm to be especially palatable.

An issue our results do not address is which penalty function best matches the needs of a specific recruitment process. For example, is there a rigorous way to argue that a particular choice of penalty function is more broadly applicable than another? That said, the choice between one-sided and two-sided penalty is rather intuitive, depending on the application. And our results provide computational arguments in favor of L_2 when two-sided penalty is desired, and L_1^+ for one-sided penalty.

From an ethical viewpoint, a potential concern is that our proposal may ultimately have unintended negative consequences. For example, if many faculty hiring committees adopted our optimization-based approach, might candidates have fewer opportunities? We believe, however, that the opposite is true. Currently the academic job market is strikingly inefficient, as committees often converge on a few candidates who are inundated with interviews and offers, while comparably strong candidates are left with nothing. If our approach is adopted widely, it is likely to widen the pool of candidates who receive appealing offers. Granted, a centralized matching market (in the style of the National Resident Matching Program) may be an even better solution, but creating such a market requires a huge — and often impractical — degree of coordination; by contrast, our approach can be adopted independently by institutions and even by individual departments or committees.

8.7 Proofs

In this section, we present the proofs of all theoretical results.

8.7.1 Preliminaries

For any set or event S, we use $\overline{S}$ to denote its complement. We use the notation $f(x) \lesssim g(x)$ to denote that there exists some universal positive constant $c > 0$, such that $f(x) \leqslant c \cdot g(x)$, and use the notation $f(x) \gtrsim g(x)$ when $g(x) \lesssim f(x)$.

For any vector $x \in \mathbb{R}^n$ and set $S \subseteq [n]$, we use the shorthand $x_S \coloneqq \{x_i\}_{i \in S}$. Let $\mu_S \coloneqq \mathbb{E}[\sum_{i \in S} Z_i] = \sum_{i \in S} p_i$. We also denote by $\mu^* \coloneqq \mu_{S^*}$ the expected size of the optimal subset.

The following lemma shows the submodularity of the objective U in the selection S.

Lemma 26. *If $\rho(\cdot, M)$ is convex then $U(S)$ is submodular in S.*

Proof. We write out the objective $U(S)$ over all possible realizations of $Z \in \{0,1\}^n$ as:

$$\begin{aligned} U(S) &\coloneqq R(S) - \lambda \cdot \mathbb{E}\left[\rho\left(|S_Z|,\, M\right)\right] \\ &= \sum_{i \in S} p_i x_i - \lambda \sum_{z \in \{0,1\}^n} \mathbb{P}(Z = z) \cdot \rho\left(\sum_{i \in S} z_i,\, M\right). \end{aligned} \tag{8.4}$$

The first term in (8.6) is additive, and hence submodular. Since ρ is convex, it can be verified that the loss $\rho\left(\sum_{i \in S} z_i,\, M\right)$ is supermodular in S for each fixed realization z. Taking linear combinations of these terms yields the submodularity of $U(S)$. □ □

The proof of this lemma is provided in Section 8.7.2. The submodularity is used in the proof of Theorem 57 (in Section 8.7.8).

For the L_1^+ loss, the following lemma allows us to reason about the cardinality of S^* in the case when the penalty λ is larger than any of the item values.

Lemma 27 (Mean Bound)**.** *Consider the L_1^+ objective. There exists a universal constant $c_0 > 0$ such that the following is true. For any $\epsilon \in (0, \frac{3}{4})$, if $x_{max} \leqslant (1-\epsilon)\cdot\lambda$, then either*

$$|S^*| \leqslant \frac{c_0 \log(\frac{1}{\epsilon})}{p_{min}} \tag{8.5a}$$

or

$$\mu^* \leqslant \frac{101}{100} M. \tag{8.5b}$$

The proof of this lemma is provided in Section 8.7.2. It is used in the proofs of Theorem 56 and Theorem 57. Intuitively, this lemma says that when all values are less than and bounded away from λ, one of the following two caeses is true: either the number of items in the optimal solution is small (Eq. (8.5a)), or when the number of items in the optimal solution is large, then the realized size of the optimal solution concentrates to the expected size with a small variance. Since the item values are relatively small, in order to minimize the L_1^+ loss, the expected size is at most on the order of M (Eq. (8.5b)).

8.7.2 Proof of Lemma 26

We write out the objective $U(S)$ over all possible realizations of $Z \in \{0,1\}^n$ as:

$$\begin{aligned} U(S) &\coloneqq R(S) - \lambda \cdot \mathbb{E}\left[\rho\left(|S_Z|,\ M\right)\right] \\ &= \sum_{i \in S} p_i x_i - \lambda \sum_{z \in \{0,1\}^n} \mathbb{P}(Z = z) \cdot \rho\left(\sum_{i \in S} z_i,\ M\right). \end{aligned} \tag{8.6}$$

The first term in (8.6) is additive, and hence submodular. Since ρ is convex, it can be verified that the loss $\rho\left(\sum_{i\in S} z_i,\ M\right)$ is supermodular in S for each fixed realization z. Taking linear combinations of these terms yields the submodularity of $U(S)$. $\square$

Proof of Lemma 27

Recall that S^* denotes the optimal solution under the L_1^+ objective. Denote by $i^* \coloneqq \operatorname{argmin}_{i\in S^*} p_i$ the item in the optimal selection with the minimal probability, and denote $S \coloneqq S^* \setminus \{i^*\}$. In what follows, we prove claim (8.5) by deriving a lower bound and an upper bound on $\mathbb{P}(\sum_{i\in S} Z_i \leqslant M)$.

Lower bounding $\mathbb{P}(\sum_{i\in S} Z_i \leqslant M)$ by the optimality of S^*. Recall that the L_1^+ loss penalizes the case when the total number of accepted items exceeds M. Intuitively, adding item i^* to the set S is only beneficial if $\mathbb{P}(\sum_{i\in S} Z_i \geqslant M)$ is small. Formally, we have

$$\begin{aligned} U(S^*) - U(S) &= x_{i^*}p_{i^*} - \lambda\mathbb{E}\left[(\sum_{i\in S} Z_i + Z_{i^*} - M)_+ - (\sum_{i\in S} Z_i - M)_+\right] \\ &= x_{i^*}p_{i^*} - \lambda p_{i^*}\cdot\mathbb{P}\Big(\sum_{i\in S} Z_i \geqslant M\Big) \end{aligned}$$

By the optimality of S^*, we have $U(S^*) \geqslant U(S)$, and hence

$$\mathbb{P}\Big(\sum_{i\in S} Z_i \geqslant M\Big) \leqslant \frac{x_{i^*}}{\lambda} \overset{\text{(i)}}{\leqslant} 1-\epsilon,$$

where step (i) is true by the assumption that $x_{\max} \leqslant (1-\epsilon)\lambda$. Hence, we have

$$\mathbb{P}\Big(\sum_{i\in S} Z_i \leqslant M\Big) \geqslant \mathbb{P}\Big(\sum_{i\in S} Z_i < M\Big) \geqslant \epsilon. \tag{8.7}$$

Upper bounding $\mathbb{P}(\sum_{i\in S} Z_i \leqslant M)$ by concentration. Let the universal constant c_0 satisfy $c_0 \geqslant \frac{200}{\log(\frac{4}{3})}$. If $|S^*| < 200$, then we have

$$|S^*| < 200 \leqslant \frac{c_0 \log(\frac{1}{\epsilon})}{p_{\min}},$$

satisfying (8.5a). Hence, it remains to consider the case when $|S^*| \geqslant 200$. In what follows, we assume that condition (8.5b) does not hold. That is, we assume $\mu^* > \frac{101}{100}M$. Then we prove that condition (8.5a) holds. We derive a multiplicative Chernoff bound to upper bound $\mathbb{P}(\sum_{i\in S} Z_i \leqslant M)$. We first establish a relation between μ_S and M. Using the definition that i^* is the item with the smallest probability in the optimal selection S^*, we have

$$\mu^* = \sum_{i \in S} p_i + p_{i^*} \leqslant \frac{|S^*|}{|S^*|-1}\mu_S \overset{\text{(i)}}{\leqslant} \frac{200}{199}\mu_S, \tag{8.8}$$

where step (i) uses the assumption that $|S^*| \geqslant 200$. Combining (8.8) with the assumption that condition (8.5b) does not hold and hence $\mu^* > \frac{101}{100}M$, we have

$$\begin{aligned}\frac{101}{100}M < \mu^* &\leqslant \frac{200}{199}\mu_S \\ M &< \frac{200}{199}\cdot\frac{100}{101}\mu_S \leqslant (1-c)\mu_S,\end{aligned}$$

where $c > 0$ is a universal constant. Since $\mu_S > M$, by the multiplicative Chernoff bound,

$$\mathbb{P}\Big(\sum_{i\in S} Z_i \leqslant M\Big) \leqslant \mathbb{P}\Big(\sum_{i\in S} Z_i \leqslant (1-c)\,\mu_S\Big) \leqslant \mathrm{Exp}\left(-\frac{c^2\mu_S}{2}\right) \leqslant \mathrm{Exp}\left(-\frac{c^2\cdot|S|\cdot p_{\min}}{2}\right). \tag{8.9}$$

Combining the lower and the upper bounds. Combining the lower bound (8.7) and the upper bound (8.9) on $\mathbb{P}(\sum_{i\in S} Z_i \leqslant M)$, we have

$$\epsilon \leqslant \mathbb{P}\Big(\sum_{i\in S} Z_i \leqslant M\Big) \leqslant \mathrm{Exp}\left(-\frac{c^2\cdot|S|\cdot p_{\min}}{2}\right)$$

and so

$$|S| \leqslant \frac{2}{c^2}\cdot\frac{\log(\frac{1}{\epsilon})}{p_{\min}}.$$

By the assumption that $\epsilon \leqslant \frac{3}{4}$, we have $|S^*| = |S| + 1 \leqslant \frac{c_0\log(\frac{1}{\epsilon})}{p_{\min}}$ for some universal constant $c_0 > 0$, satisfying condition (8.5a). $\square$

8.7.3 Proof of Proposition 2

We fix any target size $M > 0$. For notational simplicity, we use the shorthand $\rho(\cdot) := \rho(\cdot, M)$ for the loss function. For the set $\{1, 2, \ldots, n\}$, we say that the subset $\{1, 2, \ldots, k\}$ for each

$k \in \{0, \ldots, n\}$ is a prefix. Recall that Algorithm 10 sorts the items in decreasing order of probability as $p_1 \geqslant \ldots \geqslant p_n$, with ties broken arbitrarily. In what follows, we first show that there exists a prefix of items that is an optimal selection. Then we show that using the stopping criterion in Line 4 of Algorithm 10 achieves the minimum variance among all prefixes, and hence is an optimal selection.

Showing that a prefix of items in decreasing order of p_i achieves an optimal selection. Assume that there exists an optimal selection, denoted by $S^* \subseteq [n]$, that is not a prefix in decreasing order of p_i. By the assumption that S^* is not a prefix of items in decreasing order of p_i, there must exist items $i \in S^*$ and $i' \notin S^*$, such that $p_{i'} \geqslant p_i$. We now show that $S^{*\prime} := S^* \cup \{i'\} \setminus \{i\}$, namely removing i from S^* and then adding i', also yields an optimal selection.

If $p_i' = p_i$, it is straightforward to see that the variance remains the same, and hence $S^{*\prime} = S^* \cup \{i'\} \setminus \{i\}$ is optimal. Now we consider the case $p_{i'} > p_i$. For any subset $S \subseteq [n]$, we consider the additional variance induced by adding any item $k \notin S$ to the subset S:

$$\begin{aligned} V\left(S \cup \{k\}\right) - V\left(S\right) &= \mathbb{E}_{Z_{S\cup\{k\}}}\left[\rho\left(\sum_{i\in S\cup\{k\}} Z_i\right) - \rho\left(\sum_{i\in S} Z_i\right)\right] \\ &\overset{\text{(i)}}{=} \mathbb{E}_{Z_S}\left[\mathbb{E}_{Z_k}\,\rho\left(\sum_{i\in S\cup\{k\}} Z_i\right) - \rho\left(\sum_{i\in S} Z_i\right)\right] \\ &\overset{\text{(ii)}}{=} p_k \cdot \underbrace{\mathbb{E}_{Z_S}\left[\rho\left(\sum_{i\in S} Z_i + 1\right) - \rho\left(\sum_{i\in S} Z_i\right)\right]}_{T(S)}, \end{aligned} \tag{8.10}$$

where (i) is true by the assumption that the random variables $\{Z_i\}_{i=1}^n$ are independent, and (ii) is true by taking an expectation over Z_k. Setting $S = S^* \setminus \{i\}$ and $k \in \{i, i'\}$ in (8.10), we have

$$V\left(S^*\right) - V(S^* \setminus \{i\}) = p_i \cdot T(S^* \setminus \{i\}), \tag{8.11a}$$

$$V\left(S^{*\prime}\right) - V(S^* \setminus \{i\}) = p_{i'} \cdot T(S^* \setminus \{i\}), \tag{8.11b}$$

Combining (8.11a) with the assumption that S^* is an optimal selection, we have $T(S^* \setminus \{i\}) \geqslant 0$. Combining (8.11) with the assumption that $p_{i'} > p_i$, we have

$$V\left(S^{*\prime}\right) \geqslant V\left(S^*\right). \tag{8.12}$$

Since by assumption S^* is an optimal selection, equality holds in (8.12) and $S^{*\prime}$ is also an optimal selection.

If $S^{*\prime}$ is not a prefix, we keep repeating the same modification, until the resulting selection is a prefix. Since $p_{i'} \geqslant p_i$, we have $i' \leqslant i$, and hence in each modification, the sum of the indices in the selection decreases, namely $\sum_{k \in S^{*\prime}} k < \sum_{k \in S^*} k$. Hence, the sequence of modifications terminates, yielding an optimal selection that is a prefix.

Showing that the stopping criterion obtains a best prefix among all prefixes. We now show that the stopping criterion in Line 4 of Algorithm 10 obtains a prefix with the minimum variance among all prefixes. Since we have showed that there exists a prefix that is an optimal selection, this prefix obtained by the stopping criterion is optimal.

We consider the term T in (8.10) when adding to a selection $S \subseteq [n]$ some new item $k \notin S$. We have

$$\begin{aligned}
T(S \cup \{k\}) &= \mathbb{E}_{Z_S}\mathbb{E}_{Z_k}\left[\rho\left(\sum_{i \in S \cup \{k\}} Z_i + 1\right) - \rho\left(\sum_{i \in S \cup \{k\}} Z_i\right)\right] \\
&\overset{\text{(i)}}{=} p_k \mathbb{E}_{Z_S}\left[\rho\left(\sum_{i \in S} Z_i + 2\right) - \rho\left(\sum_{i \in S} Z_i + 1\right)\right] \\
&\qquad\qquad + (1 - p_k) \cdot \mathbb{E}_{Z_S}\left[\rho\left(\sum_{i \in S} Z_i + 1\right) - \rho\left(\sum_{i \in S} Z_i\right)\right] \\
&\overset{\text{(ii)}}{\geqslant} \mathbb{E}_{Z_S}\left[\rho\left(\sum_{i \in S} Z_i + 1\right) - \rho\left(\sum_{i \in S} Z_i\right)\right] = T(S),
\end{aligned} \tag{8.13}$$

where step (i) takes an expectation over Z_k, and step (ii) uses the property that $\rho(t+2) - \rho(t+1) \geqslant \rho(t+1) - \rho(t)$ for any $t \in \mathbb{R}$, due to the convexity of ρ. Due to the stopping criterion, Algorithm 10 yields a prefix $\{1, 2, \ldots, i^*\}$ such that $T([i]) \leqslant 0$ for all $i \leqslant i^*$, and $T([i^* + 1]) > 0$. By (8.13), it can be verified that $T([i]) > 0$ for all $i > i^*$. Hence, the variance decreases or stays the same for adding each item up to item i^*, and then strictly increases for adding each of item $(i^* + 1)$ through item n. Hence, the prefix $[i^*]$ attains the minimal variance among all prefixes, and hence is an optimal selection. □

8.7.4 Proof of Proposition 3

Consider any instance (x, p, λ, M) and any constant $c > 0$. It is straightforward to verify that the optimal solution and the solution given by any of the three greedy algorithms is identical for the instance (x, p, λ, M) and the instance $(cx, p, c\lambda, M)$. Hence, it suffices to construct an instance for a fixed value of $\lambda > 0$. We now construct instances for the greedy algorithms separately.

Instance for pGreedy. Let $M = 1$. We consider an instance consisting of two items:

$$(x_1, p_1) = (0, 1)$$

$$(x_2, p_2) = (1, p),$$

for some $p \in (0, 1)$ whose value is specified later. It is straightforward to derive that pGreedy selects item 1, attaining an objective of 0. On the other hand, the objective of only picking item 2 is:

$$p - \lambda(1 - p).$$

We take p to be sufficiently close to 1 such that $p/(1 - p) > \lambda$. Then the objective of only picking item 2 is strictly positive, and hence the objective of the optimal solution is strictly positive.

Instance for xGreedy and xpGreedy. Let $M = 1$. We consider an instance consisting of two items:

$$(x_1, p_1) = (1, 1)$$
$$(x_2, p_2) = \left(2 + \epsilon, \frac{1}{2}\right),$$

for some $\epsilon > 0$ whose value is specified later. The objective for the four possible selections is computed as:

$$U(\emptyset) = -\lambda$$
$$U(\{1\}) = 1$$
$$U(\{2\}) = 1 + \frac{\epsilon - \lambda}{2}$$
$$U(\{1, 2\}) = 2 + \frac{\epsilon - \lambda}{2}.$$

It is straightforward to derive that both xGreedy and xpGreedy pick item 2 first followed by item 1, attaining an objective of $2 + \frac{\epsilon-\lambda}{2}$. We set any value of λ such that $\lambda > 4$, and set $\epsilon = \frac{\lambda}{2} - 2 > 0$. The objective becomes $1 - \frac{\lambda}{4} < 0$. On the other hand, the optimal selection is $S^* = \{1\}$, with a strictly positive objective of 1. $\square$

8.7.5 Proof of Theorem 54

We first describe a pseudo-polynomial time algorithm proposed by Çela et al. [ÇKM06] for solving a specific form of rank-1 binary quadratic programming. Then we describe our algorithm, which operates by rounding the parameters and using the pseudo-polynomial time algorithm as a sub-routine.

Pseudo-polynomial time algorithm of [ÇKM06]. Çela et al. [ÇKM06] study unconstrained binary quadratic programming problems of the form

$$\min_{x \in \{0,1\}^n} \langle x, Ax \rangle + \langle b, x \rangle.$$

where $A \in \mathbb{R}^{n \times n}$ is symmetric and $a \in \mathbb{R}^n$. When A has rank one, this can be reformulated as

$$\min_{x \in \{0,1\}^n} \langle a, x \rangle + \gamma(\beta + \langle u, x \rangle)^2 \tag{8.14}$$

for some $a, u \in \mathbb{R}^n$ and $\gamma, \beta \in \mathbb{R}$. We note that the representation (a, u, γ, β) for the problem (8.14) is not unique. Çela et al. [ÇKM06] propose an algorithm to solve (8.14) exactly with a run time dependent on the magnitude of the representation.

Proposition 4 (Proposition 1 of [ÇKM06] with $d = 1$)**.** *Consider any instance of* (8.14) *with $u \in \mathbb{Z}^n$, $\beta \in \mathbb{Z}$, $a \in \mathbb{Z}^n$, and $\gamma \in \mathbb{Q}$. Let $K := 2\max(\|u\|_\infty, \|a\|_\infty)$. Then the minimum objective attained by* (8.14) *can be computed in $O(K^4 n^5)$ time.*

We refer to the algorithm satisfying Proposition 4 as R1UBQPSolver, which is described in the proof of Proposition 1 in Çela et al. [ÇKM06]. R1UBQPSolver takes as inputs (a, u, γ, β), and outputs the minimum objective attained by (8.14). We also note that while R1UBQPSolver as stated requires $\gamma \in \mathbb{Q}$, this serves only as a sufficient condition for arguing that arithmetic operations involving γ can be performed efficiently. Since our runtime analysis is in terms of the number of arithmetic operations performed, we remain agnostic to the exact representation of the numbers in our problem instance.

Modified R1UBQP Solver. For convenience of the presentation, we use a slightly more general version of this binary quadratic programming algorithm, which essentially serves as a rescaling of R1UBQPSolver. We will call this R1UBQPSolver2 (Algorithm 12).

Algorithm 12 R1UBQPSolver2

Require: $u \in \mathbb{Z}^n$, $\beta \in \mathbb{Z}$, $a \in \mathbb{Q}^n$ such that $a = \frac{a'}{E}$ for some $a' \in \mathbb{Z}^n$ and $E \in \mathbb{N}$, and $\gamma \in \mathbb{R}$
Ensure: $\min_{b \in \{0,1\}^n} \langle a, b \rangle + \gamma(\beta + \langle u, b \rangle)^2$
1: $a' \leftarrow Ea$
2: $\gamma' \leftarrow E\gamma$
3: **return** $\frac{1}{E}$R1UBQPSolver(a', γ', β, u)

Proposition 5. *Consider an instance of* (8.14) *with $u \in \mathbb{Z}^n$, $\beta \in \mathbb{Z}$, $a \in \mathbb{Q}^n$ such that $a = \frac{a'}{E}$ for some $a' \in \mathbb{Z}^n$ and $E \in \mathbb{N}$, and $\gamma \in \mathbb{R}$. Let $K' := 2\max(\|u\|_\infty, \|a'\|_\infty)$. Then*

R1UBQPSolver2 *computes the minimum objective attained by* (8.14) *in* $O(K'^4 E^4 n^5)$ *time.*

Proof. The correctness of R1UBQPSolver2 follows immediately from Proposition 4. For the runtime guarantee, note that the invocation of R1UBQPSolver is with $a'_i = Ea_i$, so the guarantee from R1UBQPSolver holds with $K \geqslant EK'$. □

Proposed FPTAS. We now derive the following FPTAS for our problem, which uses R1UBQPSolver2 as a subroutine:

Algorithm 13 ApproxL$_2$

Require: Problem instance $\mathcal{I} = (x, p, \lambda, M)$; additive error $\epsilon > 0$
Ensure: $S \subseteq [n]$ for which $U(S) \geqslant U(S^*) - \epsilon$

```
D ← ⌈2nλ(2M + 3(n + 1))/ε⌉
E ← ⌈2n/ε⌉
ā ← (1/E)⌊E(−p ∘ x + λ · p − λ · (p ∘ p))⌋
γ′ ← λ/D^2
β′ ← DM
u′ ← ⌊Dp⌋
OPT‾ ← −R1UBQPSolver2(ā, γ′, β′, u′)
S ← [n]
while ∃i ∈ S such that −R1UBQPSolver2(ā|_{S\{i}}, γ′|_{S\{i}}, β′|_{S\{i}}, u′|_{S\{i}}) = OPT‾ do
    S ← S \ {i}
return S
```

Given an instance of our problem, ApproxL$_2$ generates a rounded instance and then runs R1UBQPSolver2 on this rounded instance. The objective is guaranteed to be close to optimal, and so ApproxL$_2$ first finds the optimal rounded value, and then identifies a set which attains this rounded value.

We prove that ApproxL$_2$ (Algorithm 13) is a FPTAS for our problem when $\rho = L_2$.

Rewriting the objective in the form of (8.14)**.** We begin by establishing that if $\rho = L_2$ then $U(S)$ can be written in the form of Eq. (8.14). Recall from (8.2) that our objective can be written as

$$U(S) = \sum_{i \in S} p_i x_i - \lambda \cdot \mathbb{E} \left(\sum_{i \in S} Z_i - M \right)^2 .$$

Recall from (8.3) the notation $b \in \{0,1\}^n$ with $b_i \coloneqq \mathbb{1}\{i \in S\}$. The vector b is a representation of the set S, and we slightly abuse notation to let $U(b) \coloneqq U(S)$). We have

$$U(b) = \sum_{i \in [n]} b_i p_i x_i - \lambda \cdot \mathbb{E} \left(\sum_{i \in [n]} b_i Z_i - M \right)^2 . \tag{8.15}$$

For any two vectors $u, v \in \mathbb{R}^n$, let $u \circ v$ denote the entrywise product. Expanding the squared term in (8.15) yields

$$\begin{aligned}
U(b) &= \sum_{i \in [n]} b_i p_i x_i - \lambda \cdot \left(\mathbb{E} \left(\sum_{i \in [n]} b_i Z_i \right)^2 - 2M \sum_{i \in [n]} b_i p_i + M^2 \right) \\
&\overset{\text{(i)}}{=} (p \circ x)^T b - \lambda \cdot \mathbb{E} \left[\sum_{i \in [n]} b_i Z_i + \sum_{(i,j) \in [n]^2, i \neq j} b_i b_j Z_i Z_j \right] + 2\lambda M \cdot p^T b - \lambda M^2 \\
&= (p \circ x)^T b - \lambda \cdot p^T b - \lambda \cdot \sum_{(i,j) \in [n]^2, i \neq j} b_i b_j p_i p_j + 2\lambda M \cdot p^T b - \lambda M^2 \\
&= (p \circ x)^T b - \lambda \cdot p^T b - \lambda \cdot \left(\left(\sum_{i \in [n]} b_i p_i \right)^2 - \sum_i b_i^2 p_i^2 \right) + 2\lambda M \cdot p^T b - \lambda M^2,
\end{aligned}$$

where step (i) uses the fact that b_i and Z_i are binary, and hence $b_i^2 = b_i$ and $Z_i^2 = Z_i$. Using the fact that b_i is binary again, we have

$$\begin{aligned}
U(b) &= (p \circ x)^T b - \lambda \cdot p^T b - \lambda \cdot (p^T b)^2 + \lambda \cdot (p \circ p)^T b + 2\lambda M \cdot p^T b - \lambda M^2 \\
&= (p \circ x - \lambda \cdot p + \lambda \cdot (p \circ p))^T b - \lambda \left(-M + p^T b \right)^2 . \qquad (8.16)
\end{aligned}$$

Negating (8.16) yields

$$\min_{S \subseteq [n]} -U(b) = \min_{b \in \{0,1\}^n} (-p \circ x + \lambda \cdot p - \lambda \cdot (p \circ p))^T b + \lambda \left(-M + p^T b \right)^2 , \tag{8.17}$$

which matches the form of (8.14) with

$$a \coloneqq -p \circ x + \lambda \cdot p - \lambda \cdot (p \circ p), \tag{8.18a}$$

$$\gamma \coloneqq \lambda, \tag{8.18b}$$

$$\beta \coloneqq -M, \tag{8.18c}$$

$$u \coloneqq p. \tag{8.18d}$$

Rounding the parameters. We argue that rounding the parameters of an instance does not significantly affect the objective value. Consider rounded probabilities $\bar{p}$, with $|p_i - \bar{p}_i| \leqslant 1/D$ and rounded a_i with $|a_i - \bar{a}_i| \leqslant 1/E$, for some integers D and E to be specified later. How much does the value of (8.14) change for the input (8.18) if these $u_i = p_i$ are changed to $\bar{u} \coloneqq \bar{p}_i$ and a_i to $\bar{a}_i$, regardless of b? For notational simplicity, we write the objective as

$$\Psi(b, a, \gamma, \beta, u) = \langle a, b \rangle + \gamma(\beta + \langle u, b \rangle)^2,$$

with the choice of variables specified in (8.18). Letting $p'_i := p_i - \bar{p}_i$ (for compactness), the difference before and after rounding is bounded by

$$\begin{aligned} \Delta &= |\Psi(b, a, \gamma, \beta, u) - \Psi(b, \bar{a}, \gamma, \beta, \bar{u})| \\ &\leqslant \left|(a - \bar{a})^T b\right| + \gamma\left|(p^T b)^2 - (\bar{p}^T b)^2 - 2M(p - \bar{p})^T b\right| \\ &\leqslant \frac{n}{E} + \gamma \underbrace{\left(\left|(p^T b)^2 - (\bar{p}^T b)^2\right|\right)}_{T} + \gamma \left(2M \frac{n}{D}\right) \end{aligned} \tag{8.19}$$

We bound the term T by

$$
\begin{aligned}
T &= \sum_{i\neq j} b_i b_j (p_i p_j - \bar{p}_i \bar{p}_j) + \sum_i b_i^2 (p_i^2 - \bar{p}_i^2) + \\
&\leqslant \sum_{i\neq j} |p_i p_j - \bar{p}_i \bar{p}_j| + \sum_i |p_i^2 - \bar{p}_i^2| \\
&\leqslant \sum_{i\neq j} (\bar{p}_i p'_j + \bar{p}_j p'_i + p'_i p'_j) + \sum_i (2\bar{p}_i p'_i + p_i'^2) \\
&\leqslant \sum_{i\neq j} (p'_j + p'_i + p'_i p'_j) + \sum_i (2p'_i + p_i'^2) \\
&\leqslant 2n^2 \frac{1}{D} + n^2 \frac{1}{D^2} + n\frac{2}{D} + n\frac{1}{D^2} \\
&= \frac{n}{D}\left(2(n+1) + \frac{n+1}{D}\right) \qquad (8.20)
\end{aligned}
$$

Plugging (8.20) back to (8.19), we have

$$
\Delta \leqslant \frac{n}{E} + \frac{\gamma n}{D}\left(2(n+1) + \frac{n+1}{D} + 2M\right). \qquad (8.21)
$$

Recalling that $\gamma = \lambda$ for our problem. It can be verified by (8.21) that choosing $D \geqslant 2n\lambda(2M + 3(n+1))/\epsilon$ and $E \geqslant 2n/\epsilon$ ensures that $|\Psi(b, a, \gamma, \beta, u) - \Psi(b, a, \gamma, \beta, \bar{u})| \leqslant \epsilon$, for any arbitrary binary vector b.

We now define rounded versions of the problem parameters, which are rounded to increments of D. For all i, let

$$
\begin{aligned}
P_i &\coloneqq \lfloor D p_i \rfloor \\
\bar{u}_i &\coloneqq \frac{P_i}{D} = \frac{1}{D}\lfloor D p_i \rfloor \\
u'_i &\coloneqq P_i = \lfloor D p_i \rfloor,
\end{aligned}
$$

and $\beta' \coloneqq -DM$ and $\gamma' \coloneqq \frac{\gamma}{D^2} = \frac{\lambda}{D^2}$. Then letting $S(b)$ be the set indicated by b,

$$
\Psi(b, a, \gamma, \beta, u) = \langle a, b \rangle + \gamma(\beta + \langle u, b \rangle)^2 = -U(S(b))
$$

by (8.17), while

$$\begin{aligned}\Psi(b, \bar{a}, \gamma', \beta', u') &= \langle \bar{a}, b\rangle + \gamma'(\beta' + \langle u', b\rangle)^2 \\ &= \langle \bar{a}, b\rangle + \frac{\gamma}{D^2}(\beta D + \langle u', b\rangle)^2 \\ &= \langle \bar{a}, b\rangle + \gamma\left(\beta + \left\langle \frac{u'}{D}, b\right\rangle\right)^2 \\ &= \Psi(b, \bar{a}, \gamma, \beta, \bar{u}).\end{aligned} \tag{8.22}$$

We have just argued that $|U(S(b)) - \Psi(b, \bar{a}, \gamma, \beta, \bar{u})| \leqslant \epsilon$ when $D \geqslant 2n\lambda(2M + 3(n+1))/\epsilon$ and $E \geqslant 2n/\epsilon$. Observe also that the parameters $\bar{a}$, γ', β', u' are such that (8.22) satisfies the requirements for Proposition 5. Therefore $\overline{OPT} \leftarrow -\text{R1UBQPSolver2}(\bar{a}, \gamma', \beta', u')$ is some objective value within $\pm\epsilon$ of our optimal value $U(S^*)$.

Algorithm 13 therefore begins by finding some value $\overline{OPT}$ which is the optimal value of the rounded instance of the problem realized by some $b \in \{0,1\}^n$ and within ϵ of $-U(S(b)f)$. Since whatever value the b^* corresponding to S^* attains on the rounded instance is within ϵ of $-U(S^*)$, it follows that this b is an additive ϵ-approximation to $U(S^*)$.

The remainder of ApproxL$_2$ is dedicated to reconstructing the set $S(b)$ itself. It does this by iteratively removing candidate components of the solution $i \in [n]$, determining whether or not each is necessesarily part of some such $S(b)$ (subject to the i already discarded).

Runtime. R1UBQPSolver2 has runtime $O(K^4 E^4 n^5)$, and our reduction takes K to be the maximum of D and $\max(\lambda, x_{\max})/E$. In our reduction, $E = O((1+\lambda)n(M+n)/\epsilon)$. Since ApproxL$_2$ makes one call to R1UBQPSolver and all other steps are negligible, it therefore runs in time $O(\frac{n^9(M+n)^4(1+\lambda)^4}{\epsilon^4})$.

In the case that $x_i \geqslant 0$ for all $i \in n$, we assume that $M \leqslant n$, since for $M \geqslant n$ it is optimal to take $S = [n]$; in this case the runtime guarantee is therefore $O(\frac{n^{13}(1+\lambda)^4}{\epsilon^4})$. □

8.7.6 Proof of Theorem 55

We follow the reduction outlined in Section 2 of Çela et al. [ÇKM06]. In what follows, we reduce any instance of SubsetSum to an instance of our problem with the L_2 objective. An instance of SubsetSum is given by a set of positive integers $(t_1, \ldots, t_n)$ and a target sum T that is also a positive integer. We assume without loss of generality that $t_i \leqslant T$ for each $i \in [n]$.

Given any instance of SubsetSum, we now construct an instance of our problem $\mathcal{I} = (x, p, M, \lambda)$ with the L_2 objective. Recall the notation $b \in \{0,1\}^n$ with $b_i \coloneqq \mathbb{1}\{i \in S\}$. Recall from (8.17) that our optimization problem can be written as:

$$\min_{b \in \{0,1\}^n} \left(-p \circ x + \lambda \cdot p - \lambda \cdot (p \circ p)\right)^T b + \lambda \left(-M + p^T b\right)^2,$$

where the first term is linear in b, and the second term is quadratic in b. We choose the problem parameters, such that the linear term becomes zero, and the quadratic term becomes the SubsetSum problem. For the linear term, we set $x_i \coloneqq \lambda(1 - p_i)$ for each $i \in [n]$. It can be verified that $x_i \geqslant 0$, and that the linear term becomes zero. For the quadratic term, we set $M \coloneqq 1$, and set the regularizer to be any value such that $\lambda > 0$. Then we set $p_i \coloneqq \frac{t_i}{T}$, and we have $p_i \in [0,1]$ by the assumption that $t_i \leqslant T$ for every $i \in [n]$. The optimization problem then becomes

$$\min_{b \in \{0,1\}^n} \lambda \left(-1 + \sum_{i \in [n]} b_i \frac{t_i}{T}\right)^2 = \min_{S \subseteq [n]} \frac{\lambda}{T^2} \left(-T + \sum_{i \in S} t_i\right)^2. \tag{8.23}$$

Note that the optimization problem (8.23) always attains a non-negative objective value, and attains an objective of zero if and only if $\sum_{i \in S} t_i = T$, that is, there exists a solution to the SubsetSum instance. Therefore, any algorithm for finding the optimal subset S^* for our problem can be used to solve SubsetSum, completing the reduction. □

8.7.7 Proof of Theorem 56

We prove the three parts of the proposition separately. To prove upper bounds on the approximation ratio, we construct "bad" instances. Using the same argument as in the proof of Proposition 3 in Section 8.7.4, the values $\{x_i\}_{i\in[n]}$ can be rescaled according to λ, and it suffices to construct instances for a fixed value of $\lambda > 0$.

To prove lower bounds for xGreedy and xpGreedy (across all instances), it is without loss of generality to assume that all $x_i \geqslant 0$. Recapp that since the loss $\rho = L_1^+$ is monotonic, adding an item never decreases the penalty term, and thus adding an item with negative value always decreases the overall objective. Moreover, all items with negative values appear at the end in the orders used by xGreedy and xpGreedy, so there are no more items with positive values once the two greedy algorithms reach the first negative item. Therefore, xGreedy and xpGreedy never choose solutions containing any item with negative value, and hence such items can be ignored for the purposes of these proofs.

Proof of Theorem 56(a)

Let $M = 1$ and $\lambda > 1$. Consider an instance consisting of two items:

$$\begin{aligned}(x_1, p_1) &= (0, 1)\\ (x_2, p_2) &= \left(1, \frac{1}{2}\right).\end{aligned}$$

It can be verified that pGreedy only selects item 1, attaining an objective of 0. On the other hand, selecting item 2 attains a strictly positive objective of $\frac{1}{2}$.

Proof of Theorem 56(b)

We separately prove the upper and lower bounds for xpGreedy. We denote by S_{XP} the selection found by xpGreedy.

Upper bound for xpGreedy. First, suppose that $p_{\min}$ is such that $1/p_{\min} \in \{2, 3, 4, \ldots\}$. We assume this without loss of generality in order to show that xpGreedy is $O(p_{\min})$ for any $p_{\min} \in (0, 1]$. We may make this assumption because, first, for any $p_{\min} \in (1/2, 1]$ xpGreedy must be an $O(1)$ approximation, and for $p_{\min}$ in this range it is therefore also $O(p_{\min})$. On the other hand, for any choice $p_{\min} \in (0, 1/2)$, consider the instance outlined below with $\frac{1}{\lceil 1/p_{\min} \rceil}$ as the minimum probability, together with an additional item for which $(p_i, x_i) = (p_{\min}, 0)$. Then xpGreedy never chooses this last item, and the instance below demonstrates an upper bound of $O(\frac{1}{\lceil 1/p_{\min} \rceil}) = O(p_{\min})$.

We now construct the instance as follows. Let $M = 1$ and $\lambda = \frac{1+2c}{p_{\min}}$, and consider an unlimited number of items from two types:

$$(x_1, p_1) = \left(1 + \frac{c}{2}, 1\right)$$
$$(x_2, p_2) = \left(\frac{1}{p_{\min}}, p_{\min}\right).$$

We assume without loss of generality that $c \leqslant 1/2$, since if $x_i \leqslant (1-c) \cdot \lambda$ is true for some $c > 1/2$ then it is true for $c \leqslant 1/2$ also. Then it can be verified that $x_i \leqslant (1-c) \cdot \lambda$ for both types of items. Specifically

$$(1-c) \cdot \lambda = \frac{(1-c)(1+2c)}{p_{\min}} = \frac{1+c-2c^2}{p_{\min}} \geqslant 1 + c - 2c^2 \geqslant 1 + \frac{c}{2} = x_1$$

and

$$(1-c) \cdot \lambda = \frac{1+c-2c^2}{p_{\min}} \geqslant \frac{1}{p_{\min}} = x_2$$

both hold for any $c \in (0, 1/4]$.

xpGreedy adds items one-by-one from the first type. The objective after adding one item is $(1 + c/2)$. If a second item is added, the marginal change in the objective is $1 + \frac{c}{2} - \lambda = 1 + \frac{c}{2} - \frac{1+2c}{p_{\min}} < 0$. Hence, xpGreedy selects exactly one item from the first type, attaining an objective of

$$U(S_{\mathrm{XP}}) = 1 + c/2. \tag{8.24}$$

Now consider an alternative selection S consisting of t items of the second type. We now show that for some choice of t, the objective attained by the selection S satisfies $U(S) \gtrsim \frac{1}{p_{\min}}$. We write the objective as

$$\begin{aligned} U(S) &= R(S) - \lambda \cdot V(S) \\ &= t - \lambda \cdot \mathbb{E}\Big(|S_Z| - 1\Big)_+ \\ &= t - \lambda\Big(\mathbb{E}|S_Z| - 1 + \mathbb{P}(|S_Z| = 0)\Big) \\ &= t - \lambda\left(t \cdot p_{\min} - 1 + (1 - p_{\min})^t\right). \end{aligned}$$

Choosing $t = \frac{1}{p_{\min}}$, we have

$$\begin{aligned} U(S) &= \frac{1}{p_{\min}} - \lambda\left(\frac{1}{p_{\min}} \cdot p_{\min} - 1\right) - \lambda\left(1 - p_{\min}\right)^{\frac{1}{p_{\min}}} \\ &= \frac{1}{p_{\min}} - \lambda\left(1 - p_{\min}\right)^{\frac{1}{p_{\min}}} \\ &\geqslant \frac{1}{p_{\min}} - \frac{\lambda}{e}. \end{aligned}$$

Substituting in $\lambda = \frac{1+2c}{p_{\min}}$, the objective is lower bounded by

$$\begin{aligned} U(S) &\geqslant \frac{1}{p_{\min}} - \frac{1}{e} \cdot \frac{1+2c}{p_{\min}} \\ &= \frac{1}{p_{\min}} \cdot \frac{e-1-2c}{e} \overset{\text{(i)}}{\geqslant} \frac{1}{p_{\min}}, \end{aligned} \tag{8.25}$$

where step (i) is true by the assumption that $c \leqslant 1/4$.

Combining (8.24) and (8.25), we have an instance for which $U(S_{\mathrm{XP}}) \leqslant c' \cdot p_{\min} U(S)$ for some constant c', establishing an upper bound of $O(p_{\min})$ on the approximation ratio.

Lower bound for xpGreedy. First, if the total number of items is at most M, then it can be verified that selecting all items is optimal.

We denote by S_M be the M items with highest expected values $x_i p_i$, and denote by S_{XP} the solution that xpGreedy finds. Moreover, we have $S_M \subseteq S_{\mathrm{XP}}$, because all values are assumed nonnegative, and the penalty term for the L_1^+ loss is zero when adding the first M items. By definition, xpGreedy only improves the objective in each step. Hence, we have

$$U(S_M) \leqslant U(S_{\mathrm{XP}}). \tag{8.26}$$

We now provide a lower bound for the selection S_M. Applying the Mean Bound (Lemma 27) with $\epsilon = \min(c, 1/2)$, we have either

$$|S^*| \leqslant \frac{c'}{p_{\min}} \tag{8.27a}$$

where c' is a positive constant, or

$$\mu^* \leqslant \frac{101}{100} M. \tag{8.27b}$$

If (8.27a) holds, we have $|S^*| \leqslant \frac{cM}{p_{\min}}$ because $M \geqslant 1$. If (8.27b) holds, we have $p_{\min} \cdot |S^*| \leqslant \mu^* \leqslant \frac{101}{100} M$, and hence $|S^*| \leqslant \frac{101}{100} \cdot \frac{M}{p_{\min}}$. Combining the two cases, we have

$$|S^*| \lesssim \frac{M}{p_{\min}}. \tag{8.28}$$

Next, we consider the expected reward $R(S_M)$ for the selection S_M. We note that $|S^*| \geqslant |S_M| = M$ by the optimality of S^*. This is because $x_i \geqslant 0$, so adding any item to a selection containing less than M items only increases the objective. Recall that the selection S_M consists of the M items with the maximum expected reward $p_i x_i$. Hence, the mean expected reward $p_i x_i$ for the set S_M (over all items in this set) is greater than or equal to the mean expected reward for the set S^*. Namely,

$$\begin{aligned} \frac{1}{M} R(S_M) &= \frac{1}{|S_M|} \sum_{i \in S_M} p_i x_i \\ &\geqslant \frac{1}{|S^*|} \sum_{i \in S^*} p_i x_i = \frac{1}{|S^*|} R(S^*). \end{aligned}$$

Hence, we have

$$\begin{aligned} U(S_M) = R(S_M) &\geqslant \frac{M}{|S^*|} R(S^*) \\ &\overset{\text{(i)}}{\gtrsim} p_{\min} \cdot R(S^*) \geqslant p_{\min} \cdot U(S^*), \end{aligned} \tag{8.29}$$

where step (i) is true by plugging in (8.28). Combining (8.29) with (8.26), we have

$$U(S_{\text{XP}}) \geqslant U(S_{\text{XP}}) \gtrsim p_{\min} \cdot U(S^*),$$

completing the proof of the lower bound $\Omega(p_{\min})$ of the approximation ratio for xpGreedy. $\square$

Proof of Theorem 56(c)

We separately prove the upper and lower bounds for xGreedy.

Upper bound for xGreedy. Let $M = 1$ and $\lambda = \frac{2}{p}$. Consider an instance consisting of an unlimited number of items from two types:

$$(x_1, p_1) = (1, p_{\min})$$

$$(x_2, p_2) = (1 - \epsilon, 1),$$

where $\epsilon \in (0, 1)$ is a constant, and $p_{\min} \in (0, 1)$. We again suppose without loss of generality that $c \leqslant 1/2$, and it can be verified that $x_i \leqslant (1 - c) \cdot \lambda$ for both types of items.

xGreedy adds items one-by-one from the first type. The objective after adding one item is $p_{\min}$. If a second item is added, the marginal change in the objective is $p_{\min} - \lambda p_{\min}^2 = -p_{\min} < 0$. Hence, xGreedy selects exactly one item from the first type, attaining an objective of $p_{\min}$.

On the other hand, choosing a single item from the second type attains an objective value of $(1 - \epsilon)$. Therefore xGreedy has a worst-case approximation ratio of at most $\frac{p_{\min}}{1-\epsilon}$, namely $O(p_{\min})$.

Lower bound for xGreedy. We modify the construction of S_M in the proof of part (b) to be the set of M items with the highest values x_i, Then we apply similar arguments as in part (b), and outline the steps as follows.

Denote by S_M the set of M items with the highest values x_i, and denote by S_{X} the solution that xGreedy finds. Then again we have $S_M \subseteq S_{\text{X}}$ and hence

$$U(S_M) \leqslant U(S_{\text{X}}). \tag{8.30}$$

We now provide a lower bound for the selection S_M. Using the same argument as in part (b), we have (cf. (8.28)):

$$|S^*| \lesssim \frac{M}{p_{\min}}. \tag{8.31}$$

We note that $|S^*| \geqslant |S_M| = M$ by the optimality of S^*. Since the selection S_M consists of the M items with the maximum values x_i, the mean value for the set S_M is greater than or equal to the mean value for the set S^*. Namely,

$$\frac{1}{M}\sum_{i \in S_M} x_i \geqslant \frac{1}{|S^*|}\sum_{i \in |S^*|} x_i.$$

Next note that for any $i \in S_M$, $\frac{1}{p_{\min}}(p_i x_i)$ is larger than $p_j x_j$ for any $j \in S^* \setminus S_M$, since for such i and j we have $\frac{1}{p_{\min}}(p_i x_i) \geqslant x_i \geqslant p_j x_j$. Therefore,

$$\begin{aligned}
U(S_M) = R(S_M) = \sum_{i \in S_M} p_i x_i &\geqslant p_{\min} \sum_{i \in S_M} x_i \\
&\geqslant p_{\min} \cdot \frac{M}{|S^*|} \sum_{i \in S^*} x_i \\
&\geqslant p_{\min} \cdot \frac{M}{|S^*|} \sum_{i \in S^*} p_i x_i \\
&\overset{\text{(i)}}{\gtrsim} p_{\min}^2 \cdot R(S^*) \geqslant p_{\min}^2 \cdot U(S^*),
\end{aligned} \tag{8.32}$$

where step (i) is true by plugging in (8.31). Combining (8.30) with (8.32) completes the proof of the lower bound $\Omega(p_{\min}^2)$ of the approximation ratio for xGreedy. □

8.7.8 Proof of Theorem 57

Notation. We begin with some notation that is used in the proofs in this section. Given any instance $\{x_i, p_i\}_{i \in [n]}$, we construct a *rounded* instance $\{y_i, q_i\}_{i \in [n]}$ as follows. First we round *up* the probabilities p_i to $q_i := 2^{\lceil \log_2 p_i \rceil}$, that is, the smallest power of two that is greater than or equal to p_i. Then we construct new values y_i such that the expected value of each item is preserved. Formally,

$$\begin{aligned}
q_i &:= \min\left\{\frac{1}{2^i}, i \in \mathbb{N} : \frac{1}{2^i} \geqslant p_i\right\}, \\
y_i &:= \frac{p_i}{q_i} x_i.
\end{aligned}$$

We slightly abuse the notation, and for any selection $S = \{x_i, p_i\}_{i\in[n]}$, we denote by $S' \coloneqq \{y_i, q_i\}_{i\in[n]}$ the corresponding set with rounded probabilities and values. The parameters M and λ for this rounded instance remain unchanged. Note that by construction, we have

$$R(S) = R(S') \tag{8.33a}$$

$$V(S) \leqslant V(S') \tag{8.33b}$$

$$U(S) \geqslant U(S'), \tag{8.33c}$$

Eq. (8.33a) holds by the definition of the rounded set S'; Eq. (8.33b) in fact holds for all nondecreasing loss function ρ, because $\sum_{i\in S'} Z_i$ stochastically dominates $\sum_{i\in S} Z_i$; Eq. (8.33c) follows by combining (8.33a) and (8.33b).

Finally, recall the observation from Section 8.3.2 that we assume without loss of generality that $x_i > 0$ for all $i \in [S]$, since the marginal contribution of any i for which $x_i \leqslant 0$ to any $U(S)$ is nonpositive.

Overview of Algorithm 11. We begin by reiterating the overview of Algorithm 11 presented in Section 8.3.2. At a high level, this algorithm proceeds first by dividing the items into three groups according to their values x_i.

$$N_{\mathrm{L}} \coloneqq \{i \in [n] : x_i \leqslant (1-\epsilon)\lambda\}$$

$$N_{\mathrm{M}} \coloneqq \{i \in [n] : (1-\epsilon)\lambda < x_i < \lambda\}$$

$$N_{\mathrm{H}} \coloneqq \{i \in [n] : x_i \geqslant \lambda\}.$$

Since U is submodular (see Lemma 26), the optimal solution within at least one of these groups is constant-competitive with $U(S^*)$. We consider each group separately, and obtain a constant-factor approximation for each group. We now provide an overview of the three cases. In particular, the small items in N_{L} are handled by Algorithm 14, and the medium items in set N_{M} are handled by Algorithm 15.

Algorithm 14 LowValueL$_1^+$ (with universal constant c)

Require: Problem instance $\mathcal{I} = (x, p, \lambda, M)$, with $x_i \leqslant (1 - \frac{p_{\min}}{4})\lambda$
Ensure: $S \subseteq [n]$ for which $U(S)/U(S^*) \gtrsim 1$
1: $\tau \leftarrow \frac{c}{p_{\min}^2} \max\left\{1, \log\left(\frac{1}{p_{\min}}\right), \log\left(\frac{\lambda}{x_{\max}}\right)\right\}$
2: $\mathcal{L} \leftarrow \{S \subseteq [n] \colon |S| \leqslant \tau\}$ // Brute-force small instances

3: **for** $S \in \mathcal{L}$ **do**
4: Calculate $U(S)$
5: $S_L \leftarrow \arg\max_{S \in \mathcal{L}} U(S)$
6: Let q be the rounded p and $Q \leftarrow \{q_i\}$ the distinct rounded probabilities; let t_r be the multiplicity of each rounded probability r in the vector q. // Round large instances
7: $\mathcal{H} \leftarrow \emptyset$
8: **for** $s \in \prod_{r \in Q} \{0, 1, \ldots, t_r\}$ **do**
9: Construct S from the s_r many $i \in [n]$ of highest x_i for which $q_i = r$, for each $r \in Q$
10: Calculate $U(S)$, using unrounded probabilities and values
11: $\mathcal{H} \leftarrow \mathcal{H} \cup \{S\}$
12: $S_H \leftarrow \arg\max_{S \in \mathcal{H}} U(S)$
13: **return** $S \in \{S_L, S_H\}$ maximizing $U(S)$

Algorithm 15 MediumValueL$_1^+$

Require: Problem instance $\mathcal{I} = (x, p, \lambda, M)$, with $(1 - \frac{p_{\min}}{4}) \cdot \lambda \leqslant x_i \leqslant \lambda$
Ensure: $S \subseteq [n]$ for which $U(S)/U(S^*) = \Omega(1)$
1: **if** $n \leqslant \frac{36}{p_{\min}^2}$ **then**
2: **return** $\arg\max_{S \subseteq [n]} U(S)$
3: **else if** $\mu_{[n]} \geqslant M$ **then**
4: Choose any $S \subseteq [n]$ such that $M \leqslant \mu_S < M + 1$
5: **else**
6: Choose any $S \subseteq [n]$ such that $\frac{\mu_{[n]}}{3} \leqslant \mu_S \leqslant \frac{\mu_{[n]}}{2}$
7: **return** S

- **Low-value items $N_{\mathbf{L}}$ (Algorithm 14):** LowValueL$_1^+$ presented in Algorithm 14 handles the case where items have low values. It consists of two parts: a search over small candidate solutions and a search over rounded candidate solutions. In the first part, we brute-force all small solutions whose size are at most τ (Line 3). This brute-force search succeeds if the optimal selection is small.

  The second part is the technical crux of proving the constant-factor approximation of LowValueL_1^+. In the second part, we compute rounded probabilities and values (q_i, y_i) for each item. This rounding procedure reduces the number of candidate solutions dramatically. We then brute-force over all rounded solutions (Line 8), select the rounded solution that maximizes the objective value, and prove that this solution is comparable to the (unrounded) optimal solution. Since the first part succeeds the case where the optimal solution is small, we may assume in this second part that the optimal solution is sufficiently large; this allows us to prove that our selection is robust to rounding.

As an aside, we take the rounding to be to powers of two, but our analysis generalizes to rounding to powers of $(1 + c)$ for any constant $c > 0$, and this parameter may be tuned in order to trade off between runtime and performance in practice.

- **Medium-value items $N_{\mathbf{M}}$ (Algorithm 15):** MediumValueL_1^+ presented in Algorithm 15 handles items with values close to λ. If the number of items is small, it brute-forces over all possible solutions (Line 2). If the number of items is large, the algorithm chooses any subset such that the expected number of accepted items is around M (Line 4). If no such subset exists, then the expected number of accepted items when choosing all items must be less than M. In this case, then we choose a subset with approximately half the expected realizations compared to that of all items (Line 6). We choose a proportion less than one in order to ensure that the penalty incurred is not too large relative to the reward. This subset (Line 6) along with the subset defined in (Line 4) and always exists, formalized in the proof of Lemma 29 in Section 8.7.8.

- **High-value items $N_{\mathbf{H}}$:** for the group of items with values above λ, it is easy to see that choosing the entire group is optimal.

We now prove the approximation ratio and runtime for OneSidedL_1^+.

Proof of Theorem 57. To begin, we split the items in the optimal set S^*, according to their values:

$$S_L^* := S^* \cap N_L,$$

$$S_M^* := S^* \cap N_M,$$

$$S_H^* := S^* \cap N_H.$$

By the submodularity of $U(S)$ in Lemma 26, we have $U(S_L^*) + U(S_M^*) + U(S_H^*) \geqslant U(S^*)$. In particular, this implies that $\max\{U(S_L^*), U(S_M^*), U(S_{\mathrm{H}}^*)\} \geqslant \frac{1}{3}\, U(S^*)$. In order to provide a constant-factor approximation to $U(S^*)$, it therefore suffices to identify sets which provide constant-factor approximations to $U(S_L^*)$, $U(S_M^*)$, and $U(S_H^*)$, and return the set with the highest objective value among them. We choose $\epsilon := p_{\min}/4$ to determine the boundary between N_L and N_M, and address each group separately. In each case we seek to find a subset which competes with the optimal subset of N_L (say), which in turn is an approximation to $U(S_L^*)$.

Low-value items $N_{\mathbf{L}}$. The following lemma provides the approximation guarantee of LowValueL$_1^+$.

Lemma 28 (Small x_i)**.** *Suppose that $x_i \leqslant (1 - \frac{p_{min}}{4}) \cdot \lambda$ for all $i \in [n]$. Then* LowValueL$_1^+$ *(Algorithm 14) is a constant-factor approximation to $U(S^*)$ which runs in time $n^{\frac{c}{p_{min}^2} \max\left\{1, \log\left(\frac{1}{p_{min}}\right), \log\left(\frac{\lambda}{x_{max}}\right)\right\}}$, where c is a universal constant.*

The proof of this lemma is provided in Section 8.7.8, and is arguably the heart of the analysis of OneSidedL$_1^+$. By applying this lemma to $[n] = N_L$ we obtain S_L with $U(S_L)$ within a constant factor to the optimal objective among selections within N_{L}, and hence a constant factor to $U(S_L^*)$.

Medium-value items $N_{\mathbf{M}}$. The following lemma provides the approximation ratio guarantee of MediumValueL$_1^+$.

Lemma 29 (Medium x_i)**.** *If $\lambda\left(1-\frac{p_{min}}{4}\right) \leqslant x_i \leqslant \lambda$ for all $i \in [n]$ then* MediumValueL_1^+ *(Algorithm 15) finds some $S \subseteq [n]$ which is a constant-factor approximation to $U(S^*)$ and runs in time $n^{O(1/p_{min}^2)}$.*

The proof of this lemma is provided in Section 8.7.8. By applying this lemma to $[n] = N_M$ we obtain S_M with $U(S_M)$ within a constant factor to the optimal objective among all selections within N_{M}, and hence a constant factor to $U(S_M^*)$.

High-value items $N_{\mathbf{H}}$. This case is simple: we select all items by taking $S_H = N_H$. It can be verified that adding every item strictly increases the objective, and hence N_{H} attains the optimal objective among all selections within N_{H}. By the optimality of S_H, we have $U(S_H) = U(S_H^*)$.

Putting the three cases together, we have S_L, S_M, and S_H, and by the argument provided above at least one of these is a constant-factor approximation to $U(S^*)$. Therefore choosing the one with highest objective value gives a constant-factor approximation.

Runtime. The algorithms for the cases above operate by identifying a collection of sets to test the objective value of, and then evaluating the objective. Fortunately this can be done efficiently.

Lemma 30 (Efficient Objective Evaluation)**.** *Suppose that $\rho(a, M)$ is known for all $a \in \{0, 1, \ldots, n\}$. For any set of items $\{x_i, p_i\}_{i\in[n]}$, the objective $U([n])$ can be computed in $O(n^2)$ arithmetic operations.*

This is proved in Section 8.7.8. Applying this lemma to any candidate subset S shows that the objective with respect $S \subseteq [n]$ can be computed in $O(|S|^2)$ arithmetic operations.

By Lemma 28, the runtime of LowValueL_1^+ is $n^{\frac{c}{p_{\min}^2}\max\left\{1,\log\left(\frac{1}{p_{\min}}\right),\log\left(\frac{\lambda}{x_{\max}}\right)\right\}}$. By Lemma 29 the runtime of MediumValueL_1^+ is $n^{O(1/p_{\min}^2)}$, which is less than that of LowValueL_1^+. Finally, the high-value items case entails evaluating the objective in of a single set; by Lemma 30 this can be done in $O(n^2)$.

The cost of combining these cases is polynomial in n, and so the brute force stage of $\textsc{LowValueL}_1^+$ dictates the runtime of $\textsc{OneSidedL}_1^+$, giving the claimed runtime of $n^{c\frac{1}{p_{\min}^2}\max\left\{1,\log\left(\frac{1}{p_{\min}}\right),\log\left(\frac{\lambda}{x_{\max}}\right)\right\}}$ for some universal constant $c > 0$. □

We now turn to the statements and proofs of the supporting lemmas.

The following lemma says that the solution can be downsampled so that its $\sum_i p_i$ is at most a constant factor from the original, while $\sum_i p_i x_i$ is at least a constant factor from the original. Informally, we simply select the appropriate number of items with the highest x_i.

Lemma 31 (Downsampling Lemma). *Consider an instance $\{p_i, x_i\}_{i\in[n]}$ with $x_i \geqslant 0$ for all $i \in [n]$. Then for any $S \subseteq [n]$ and any $\beta \in [0,1]$, there exists some $T \subseteq S$ that satisfies*

$$\mu_T \leqslant \beta \cdot \mu_S \tag{8.34a}$$

and

$$R(T) \geqslant \beta\left(1 - \frac{1}{\beta \cdot p_{min} \cdot |S|}\right) \cdot R(S). \tag{8.34b}$$

The proof of this lemma is provided in Section 8.7.8. If we were allowed to choose items to be in T fractionally, then condition (8.34b) would more closely mimic condition (8.34a) and the proof of this lemma would be even more straightforward; as it is, condition (8.34b) must be slightly weaker since we must sometimes leave the last item out of T in order to satisfy condition (8.34a).

This lemma supports the efficient search over rounded solutions which is conducted in $\textsc{LowValueL}_1^+$. Informally, it does this by proving that if some starting set satisfies certain properties, then either there is a small subset with good objective value, or the search over rounded solutions will identify a subset with good objective value.

Lemma 32 (Rounding Lemma). *Consider the one-sided $\rho = L_1^+$ loss. Consider any selection*

$S \subseteq [n]$ that simultaneously satisfies

$$U(S) \geqslant 0 \tag{8.35a}$$

$$\mu_S \leqslant \frac{3}{2}M \tag{8.35b}$$

$$\lambda V(S) \leqslant \frac{1}{15}R(S). \tag{8.35c}$$

Then there exists some subset $T \subseteq S$ that satisfies either

$$U(T) \geqslant \frac{1}{3}U(S) \qquad \textit{and} \qquad |T| \leqslant \frac{24}{p_{min}}, \tag{8.36a}$$

or

$$U(T) \geqslant U(T') \geqslant \frac{1}{24}U(S), \tag{8.36b}$$

where T' denotes the rounded instance of the set T.

The proof of this lemma is provided in Section 8.7.8.

This next lemma bounds the penalty of a subset with expected realized size smaller than M. It uses the independence of the events Z_i to apply tail bounds to the probability that the realized size of the subset exceeds M. When applied to a downsampled subset derived from Lemma 31, it will show that the penalty decreases exponentially while the reward decreases only linearly, yielding a subset which is within a small factor of the starting set's objective but is much less balanced.

Lemma 33 (Downsampling Penalty Bound)**.** *Consider the one-sided $\rho = L_1^+$ loss. Consider any selection $S \subseteq [n]$ such that $\mu_S \leqslant M$. Then for all $k \in \mathbb{N}_+$, the penalty term is bounded as*

$$V(S) \leqslant \lambda \cdot k \cdot e^{\frac{-2(M-\mu_S)^2}{|S|}} \cdot \left(1 - e^{\frac{-4(M-\mu_S)k}{|S|}}\right)^{-2}. \tag{8.37}$$

The proof of this lemma is provided in Section 8.7.8. Informally, under this loss the penalty increases linearly in the extent to which the realized size of S exceeds M, while the probability that such a violation occurs decreases exponentially. The parameter k is the size of the buckets for which the analysis of these competing influences is performed.

Proof of Lemma 28

Recall that we define $\tau := \frac{c}{p_{\min}^2} \max\left\{1, \log\left(\frac{1}{p_{\min}}\right), \log\left(\frac{\lambda}{x_{\max}}\right)\right\}$ in Line 2 of Algorithm 14. Let c_0 be the universal constant identified in Lemma 27. With $p_{\min} \leqslant 1$, it is straightforward to verify that there exists a universal constant c, such that τ is bounded by

$$\tau > \frac{c_0}{p_{\min}} \log\left(\frac{4}{p_{\min}}\right), \tag{8.38a}$$

$$\tau \geqslant \frac{24}{p_{\min}} \tag{8.38b}$$

$$\tau \geqslant \frac{9}{2}\left(\frac{1}{p_{\min}^2}\left(7 + \log\left(\frac{1}{p_{\min}}\right) + 3\log\left(\frac{\lambda}{x_{\max}}\right)\right)\right) \tag{8.38c}$$

We use these bounds in the remaining proof.

In Line 2-5, we evaluate the objective for each selection S with $|S| \leqslant \tau$ by brute-force. Hence, if $|S^*| \leqslant \tau$, then the optimal selection is correctly identified. It remains to consider the case when $|S^*| > \tau$.

When $|S^*| > \tau$, we apply the Mean Bound (Lemma 27) with $\epsilon = \frac{p_{\min}}{4}$. We have either

$$|S^*| \leqslant \frac{c_0}{p_{\min}} \log\left(\frac{4}{p_{\min}}\right) \tag{8.39a}$$

or

$$\mu^* \leqslant \frac{101}{100} M. \tag{8.39b}$$

The bound (8.38a) on τ contradicts (8.39a). Hence we have that (8.39b) holds, namely $\mu^* \leqslant \frac{101}{100} M$ from.

We call a set S "balanced" if $\lambda V(S) > \frac{1}{15} R(S)$, that is, the penalty term is a nontrivial portion of the reward term. Otherwise, we call the set "unbalanced". We consider the following two cases separately depending on whether the set S^* is balanced or not.

Case 1: $|S^*| > \tau$ **and** $\lambda V(S^*) \leqslant \frac{1}{15} R(S^*)$**.**

Note that the optimal selection always has a nonnegative objective for the L_1^+ loss. That is, $U(S^*) \geqslant 0$. Hence, conditions (8.35) are satisfied. Applying the Rounding Lemma (Lemma 32), there exists some $T \subseteq S^*$ such that

$$U(T) \geqslant \frac{1}{3}U(S^*) \qquad \text{and} \qquad |T| \leqslant \frac{24}{p_{\min}}, \tag{8.40a}$$

or

$$U(T) \overset{\text{(i)}}{\geqslant} U(T') \geqslant \frac{1}{24}U(S^*), \tag{8.40b}$$

In this first case (8.40a), by the bound (8.38b) on τ, we have

$$\tau \geqslant \frac{24}{p_{\min}} \geqslant |T|.$$

Hence, the selection T is included in the brute-force search. We obtain a constant-factor approximation to $U(S^*)$ in the brute-force search over small solutions (Line 5 of Algorithm 14).

In the second case (8.40b), if $|T| \leqslant \tau$, then again the selection T is included in the brute-force search in Line 5 of Algorithm 14, and the brute-force identifies a solution which is at least as good and hence a constant-factor approximation. If $|T| > \tau$, then Line 6 to Line 12 search over all possible rounded solutions, including T' which is a constant-factor approximation due to (8.40b). Hence, it identifies a solution which is at least as good and hence a constant-factor approximation. identifies some $\widehat{T}$ for which $U(\widehat{T}) \geqslant U(\widehat{T}') \geqslant U(T')$, which provides a constant-factor approximation to $U(S^*)$.

Case 2: $|S^*| > \tau$ and $\lambda V(S^*) > \frac{1}{15}R(S^*)$.

As an overview of this case, we appeal to the Downsampling Lemma (Lemma 31) with a small downsampling ratio in order to obtain some $T \subseteq S$, and then argue that $\lambda V(T) \leqslant \frac{1}{15}R(T)$. Then we obtain a constant-factor approximation to $U(T)$ by solving the rounded problem in Case 1.

Downsampling to an unbalanced set. Starting with the optimal selection S^*, we apply the Downsampling Lemma (Lemma 31) construction some $T \subseteq S^*$ such that this T is unbalanced but still yields a large objective. Specifically, applying Lemma 31 with $\beta = \frac{1}{2}$, there exists some $T \subseteq S^*$ that satisfies $\mu_T \leqslant \frac{\mu^*}{2}$ and

$$R(T) \geqslant \left(\frac{1}{2} - \frac{1}{|S^*| \cdot p_{\min}}\right) R(S^*). \tag{8.41}$$

We assume that the selection T is sufficiently unbalanced as:

$$V(T) \leqslant \frac{R(T)}{15}. \tag{8.42}$$

We now identify a constant-factor approximation by similar arguments as in Case 1. Specifically, under the assumption (8.42), the objective of T satisfies

$$\begin{aligned} U(T) \overset{\text{(i)}}{\geqslant} \frac{14}{15} R(T) \overset{\text{(ii)}}{\geqslant} \frac{14}{15} \cdot \left(\frac{1}{2} - \frac{1}{|S^*| \cdot p_{\min}}\right) R(S^*) \\ \geqslant \frac{14}{15} \cdot \left(\frac{1}{2} - \frac{1}{|S^*| \cdot p_{\min}}\right) U(S^*), \end{aligned} \tag{8.43}$$

where step (i) is due to the assumption (8.42), and step (ii) is due to (8.41). By the assumption $|S^*| \geqslant \tau$ and the bound (8.38b) on τ, we have

$$|S^*| \geqslant \tau \geqslant \frac{24}{p_{\min}}. \tag{8.44}$$

Applying (8.44) to inequality (8.43), the selection T is a constant-factor approximation to S^* with $U(T) \geqslant \frac{1}{4} U(S^*)$. Since T is sufficiently unbalanced by assumption (8.42), using the same arguments as in Case 1 to the set T identifies a selection that is a constant-factor approximation to T, and therefore to $U(S^*)$. It now remains to prove (8.42).

Proving (8.42)**.** Recall from (8.39b) that $\mu^* \leqslant \frac{101}{100} M$. Hence, we have $\mu_T \leqslant \frac{\mu^*}{2} \leqslant \frac{101}{200} M < M$. Then we provide an upper bound on the variance term $V(T)$ by applying Lemma 33 to the set T. Applying Lemma 33 with $k = 1$, we have

$$V(T) \leqslant \lambda \cdot \underbrace{\operatorname{Exp}\left(\frac{-2(M - \mu_T)^2}{|T|}\right)}_{T_1} \cdot \underbrace{\left(1 - e^{\frac{-4(M-\mu_T)}{|T|}}\right)^{-2}}_{T_2}. \tag{8.45}$$

We bound the two terms in (8.45) separately.

Recall from (8.39b) that $\mu^* \leqslant \frac{101}{100}M$. We then have

$$M - \mu_T \geqslant \left(\frac{100}{101} - \frac{1}{2}\right)\mu^* > \frac{1}{3}\mu^*. \tag{8.46}$$

Using (8.46), we bound the term T_1 as

$$\begin{aligned} T_1 = \operatorname{Exp}\left(\frac{-2(M-\mu_T)^2}{|T|}\right) \leqslant \operatorname{Exp}\left(-\frac{2(\mu^*)^2}{9|T|}\right) &\overset{\text{(i)}}{\leqslant} \operatorname{Exp}\left(-\frac{2}{9}p_{\min}^2 \cdot |S^*|\right) \\ &\overset{\text{(ii)}}{\leqslant} \frac{1}{720}\frac{p_{\min}^3 x_{\max}}{\lambda} \end{aligned}$$

where step (i) is true by plugging in $\mu^* \geqslant |S^*| \cdot p_{\min}$, and $|S^*| \geqslant |T|$ due to $T \subseteq S^*$; step (ii) is true by the fact that $|S^*| \geqslant \tau$ with (8.38c). Let i_{max} be the item with the highest value. With $M \geqslant 1$, the utility for selecting the item with the highest value is $U(\{i_{\max}\}) = R(\{i_{\max}\}) = p_{i_{\max}} x_{\max} \geqslant p_{\min} x_{\max}$. Hence, the reward of the optimal selection is bounded by $R(S^*) \geqslant U(S^*) \geqslant U(\{i_{\max}\}) \geqslant p_{\min} x_{\max}$. Hence,

$$T_1 \leqslant \frac{p_{\min}^2}{45}\frac{R(S^*)}{\lambda}, \tag{8.47}$$

Using again $M - \mu_T \geqslant \frac{1}{3}\mu^* \geqslant \frac{1}{3} \cdot |S^*| \cdot p_{\min}$ and $|S^*| \geqslant |T|$, we bound the term T_2 as

$$T_2 = \left(1 - e^{\frac{-4(M-\mu_T)}{|T|}}\right)^{-2} \leqslant \left(1 - e^{\frac{-4\mu^*}{9|S^*|}}\right)^{-2} \leqslant (1 - e^{-\frac{4}{9}p_{\min}})^{-2} \leqslant \frac{16}{p_{\min}^2}, \tag{8.48}$$

where step (i) is true because it can be shown by algebra that

$$1 - e^{-\frac{4}{9}p} - \frac{1}{4}p \geqslant 0 \qquad \text{for every } p \in [0,1].$$

Plugging term T_1 from (8.47) and term T_2 from (8.48) back to (8.45) yields

$$V(T) \leqslant \lambda \cdot \frac{1}{45}R(S^*) \overset{\text{(i)}}{\leqslant} \frac{1}{15\lambda}\left(\frac{1}{2} - \frac{1}{|S^*| \cdot p_{\min}}\right)R(S^*) \overset{\text{(ii)}}{\leqslant} \frac{R(T)}{15\lambda}, \tag{8.49}$$

where step (i) is true by $|S^*| \geqslant \tau \geqslant \frac{24}{p_{\min}}$ due to (8.38b), and step (ii) is true due to (8.41), proving (8.42).

Runtime. We conclude by analyzing the runtime of $\textsc{LowValueL}_1^+$.

The number of sets S such that $|S| \leqslant \tau$ is bounded by $|\mathcal{L}| = 2^\tau \leqslant n^\tau$. For each $S \in \mathcal{L}$, we compute $U(S)$ in $O(\tau^2)$. By Lemma 30, the objective of each such set may be evaluated in $O(\tau^2)$ operations, and so the runtime of evaluating the objective for all of these small subsets is $n^{O(\tau)}$.

We also compute the objective for the $O(n^{|Q|})$ rounded sets identified in Algorithm 14, where $|Q| \leqslant \left\lceil \log_2(\frac{1}{p_{\min}}) \right\rceil$, which again by Lemma 30 can be done in $O(n^2)$ operations per set. This is therefore $n^{O(\tau)}$ also.

All of the other simple steps of $\textsc{LowValueL}_1^+$ are also polynomial in n or τ. Therefore, its overall runtime is $n^{O(\tau)}$. $\square$

Proof of Lemma 29

First, we observe that when $n \leqslant \frac{36}{p_{\min}^2}$, we find the optimal solution exactly by brute forcing over all possible solutions $S \subseteq [n]$ (Line 2 of Algorithm 15). Hence, in the rest of the proof we assume that $n \geqslant \frac{36}{p_{\min}^2}$. We discuss the two cases of $\mu_{[n]} \leqslant M$ (Line 6) and $\mu_{[n]} > M$ (Line 4) separately.

We start by establishing a reformulation of the objective under L_1^+ penalty which is convenient when all items have value close to λ, and a pair of upper and lower bounds.

Bounding the objective. Recall that our objective is of the form $U(S) = \mathbb{E}_Z\left[F_S(Z)\right]$, where the random vector $Z \in \{0,1\}^n$ is the Bernoulli realization of each item, and $F_S(Z)$ is the realized utility:

$$\begin{aligned} F_S(Z) &\coloneqq \sum_{i\in S} Z_i x_i - \lambda\left(\sum_{i\in S} Z_i - M\right)_+ \\ &= \sum_{i\in S} Z_i x_i - \lambda\cdot\max\left(0,\ \sum_{i\in S} Z_i - M\right) \\ &= \min\left(\sum_{i\in S} Z_i x_i,\ \sum_{i\in S} Z_i x_i - \lambda\cdot\left(\sum_{i\in S} Z_i - M\right)\right). \end{aligned} \tag{8.50}$$

Plugging in the assumption that $(1-\epsilon)\cdot\lambda \leqslant x_i \leqslant \lambda$ to (8.50), and recalling that $|S_Z| := \sum_{i\in S} Z_i$, we derive the upper bound

$$\begin{aligned} F_S(Z) &\leqslant \min\left\{\lambda\cdot|S_Z|, \lambda\cdot M + \sum_{i\in S} Z_i(x_i-\lambda)\right\} \\ \frac{F_S(Z)}{\lambda} &\leqslant \min\left\{|S_Z|,\ M\right\}, \end{aligned} \tag{8.51a}$$

and the lower bound

$$\begin{aligned} F_S(Z) &\geqslant \min\left\{(1-\epsilon)\lambda\cdot|S_Z|, (1-\epsilon)\lambda\cdot|S_Z| - \lambda(|S_Z|-M)\right\} \\ \frac{F_S(Z)}{\lambda} &\geqslant \min\left\{(1-\epsilon)\cdot|S_Z|,\ M-\epsilon\cdot|S_Z|\right\}. \end{aligned} \tag{8.51b}$$

We denote $\epsilon = \frac{p_{\min}}{4}$ for notational simplicity. As an overview, for the two cases to be presented below, we apply the upper bound (8.51a) to the optimal set S^*, and the lower bound (8.51b) to our candidate sets which we are proving are competitive with S^*.

Case 1: $\mu_{[n]} > M$ **and** $n \geqslant \frac{36}{p_{\mathbf{min}}^2}$**.** Consider any arbitrary set $S \subseteq [n]$ that satisfies (cf. Line 4 of Algorithm 15):

$$M \leqslant \sum_{i\in S} p_i < M+1. \tag{8.52}$$

Such a set S always exists because each $p_i \leqslant 1$. Furthermore, such a set in S can be found efficiently, by greedily adding items to the set in an arbitrary order one-by-one until condition (8.52) is satisfied. We denote by $\mathbb{E}$ the event that at most half of the Bernoulli random variables from S are 1. Formally, $\mathbb{E} \coloneqq \{|S_Z| \leqslant M/2\}$. By the multiplicative Chernoff bound,

$$\mathbb{P}(\mathbb{E}) = \mathbb{P}\Big(|S_Z| \leqslant \frac{M}{2}\Big) \overset{\text{(i)}}{\leqslant} \Pr\left(|S_Z| \leqslant \frac{\mu_S}{2}\right) \leqslant e^{-\frac{\mu_S}{8}} \overset{\text{(ii)}}{\leqslant} e^{-\frac{M}{8}}, \tag{8.53}$$

where steps (i) and (ii) are true by the construction of S in (8.52). We derive a lower bound on $F_S(Z)$ depending on $\mathbb{E}$. Conditional on $\mathbb{E}$, the penalty term is 0 and we have $F_S(Z) \geqslant 0$. We now consider the case conditional on $\overline{\mathbb{E}}$. By the assumption of $\mu_S < M+1$ from (8.52), we have the deterministic relation $|S_Z| \leqslant \frac{M+1}{p_{\min}}$. Applying the lower bound in (8.51b), conditional on $\overline{\mathbb{E}}$,

$$\begin{aligned}
F_S(Z) &\geqslant \lambda \cdot \min\left\{(1-\epsilon)\cdot\frac{M}{2},\ M - \epsilon\cdot|S_Z|\right\}\\
&\geqslant \lambda \cdot \min\left\{\frac{(1-\epsilon)M}{2},\ M - \frac{\epsilon(M+1)}{p_{\min}}\right\}\\
&= \lambda M \cdot \min\left\{\frac{1-\epsilon}{2},\ 1 - \frac{\epsilon(M+1)}{M\cdot p_{\min}}\right\}\\
&\overset{\text{(i)}}{=} \lambda M \cdot \frac{1-\epsilon}{2}
\end{aligned}$$

where step (i) holds because by the assumption $\epsilon \leqslant \frac{p_{\min}}{4}$ and $M \geqslant 1$ (recall that $M \in \mathbb{N}_+$), we have $\frac{\epsilon}{p_{\min}} \leqslant \frac{1}{4}$ and $\frac{M+1}{M} \leqslant 2$. Therefore, we have $1 - \frac{\epsilon(M+1)}{M\cdot p_{\min}} \geqslant \frac{1}{2}$. Taking an expectation over Z, we then have

$$\begin{aligned}
U(S) &= \mathbb{E}[F_S(Z)]\\
&\geqslant 0\cdot\Pr(\mathbb{E}) + \frac{(1-\epsilon)\lambda M}{2}\cdot\Pr\left(\overline{\mathbb{E}}\right)\\
&\overset{\text{(i)}}{\geqslant} \frac{(1-\epsilon)(1-e^{-\frac{M}{8}})}{2}\cdot\lambda M\\
&\overset{\text{(ii)}}{\geqslant} \frac{(1-\epsilon)(1-e^{-\frac{M}{8}})}{2}\cdot U(S^*),
\end{aligned}$$

where step (i) is true by (8.53), and step (ii) is true by applying (8.51a) to the optimal selection S^*. Since $\epsilon = \frac{p_{\min}}{4} \leqslant \frac{1}{4}$ and $M \geqslant 1$ by assumption, this guarantees a constant-factor approximation of S to the optimal subset S^*.

Case 2: $\mu_{[n]} \leqslant M$ and $n \geqslant \frac{36}{p^2_{\min}}$. In this case n is large enough to apply concentration bounds, so we downsample the set $[n]$ by a factor of two and discount the probability that $|S_Z|$

exceeds M. This differs from Case 1 in that we cannot compare our objective against $\lambda \cdot M$. In particular, we consider any arbitrary set $S \subseteq [n]$ that satisfies (cf. Line 6 of Algorithm 15):

$$\frac{\mu_{[n]}}{3} \leqslant \mu_S \leqslant \frac{\mu_{[n]}}{2} \leqslant \frac{M}{2}. \tag{8.54}$$

We first show that such a set S always exists. Note that we have $\mu_{[n]} \geqslant np_{\min} \geqslant np_{\min}^2 \geqslant 36$ by the assumption of $n \geqslant \frac{36}{p_{\min}^2}$. Hence, $\frac{\mu_{[n]}}{2} - \frac{\mu_{[n]}}{3} \geqslant 1$. Since each $p_i \leqslant 1$, a set S always exists. Moreover, it can be found efficiently, by greedily adding items one-by-one in any arbitrary order until condition (8.54) is satisfied. In what follows, we separately bound the values of $U(S)$ and $U(S^*)$. We use an intermediate quantity of the expectation of the random variable $|S_Z|$ truncated at $2\mu_S$, defined by

$$G(|S_Z|) \coloneqq \mathbb{E}\Big[|S_Z| \cdot \mathbb{1}\{|S_Z| \leqslant 2\mu_S\}\Big].$$

Lower bound on $U(S)$. Due to the condition (8.54) that $\mu_S \leqslant \frac{M}{2}$, we have

$$\begin{aligned} G(|S_Z|) &\coloneqq \mathbb{E}\Big[|S_Z| \cdot \mathbb{1}\{|S_Z| \leqslant 2\mu_S\}\Big] \\ &\leqslant \mathbb{E}\left[|S_Z| \cdot \mathbb{1}\{|S_Z| \leqslant M\}\right]. \end{aligned} \tag{8.55}$$

We claim the deterministic relation

$$\min\left\{|S_Z|,\ \frac{M - \epsilon|S_Z|}{(1-\epsilon)}\right\} \geqslant |S_Z| \cdot \mathbb{1}\{|S_Z| \leqslant M\}. \tag{8.56}$$

To see (8.56), we observe that when $|S_Z| \leqslant M$, the left-hand side has

$$\frac{1}{1-\epsilon}\left(M - \epsilon|S_Z|\right) \geqslant M \geqslant |S_Z|. \tag{8.57a}$$

When $|S_Z| > M$, the right-hand side is zero, and the left-hand side is nonnegative, because

$$M \geqslant 2\mu_S \geqslant 2p_{\min}|S| \geqslant 2p_{\min}|S_Z| \geqslant \epsilon|S_Z|. \tag{8.57b}$$

Plugging (8.56) to (8.55), we have

$$\begin{aligned}
G(|S_Z|) &\leqslant \mathbb{E}\min\left\{|S_Z|,\ \frac{M-\epsilon|S_Z|}{1-\epsilon}\right\} \\
&\leqslant \frac{1}{1-\epsilon}\mathbb{E}\min\left\{|S_Z|,\ \frac{M-\epsilon|S_Z|}{1-\epsilon}\right\} \\
&= \frac{1}{(1-\epsilon)^2}\mathbb{E}\Big[\min\left\{(1-\epsilon)\cdot|S_Z|,\ M-\epsilon\cdot|S_Z|\right\}\Big] \\
&\overset{\text{(i)}}{\leqslant} \frac{1}{(1-\epsilon)^2}\cdot\frac{U(S)}{\lambda},
\end{aligned} \tag{8.58}$$

where step (i) follows from (8.51b).

Upper bound on $U(S^*)$. We decompose the expectation of $|S_Z|$ as

$$\mathbb{E}|S_Z| = G(|S_Z|) + \mathbb{E}\left[|S_Z|\cdot\mathbb{1}\{|S_Z|>2\mu_S\}\right],$$

and hence

$$\begin{aligned}
G(|S_Z|) &= \mu_S - \mathbb{E}\Big[|S_Z|\cdot\mathbb{1}\{|S_Z|>2\mu_S\}\Big] \\
&\overset{\text{(i)}}{\geqslant} \mu_S - \mathbb{E}\Big[|S_Z|\cdot\mathbb{1}\{|S_Z|>M\}\Big] \\
&= \mu_S - \mathbb{E}\Big[M\cdot\mathbb{1}\{|S_Z|>M\} + (|S_Z|-M)\cdot\mathbb{1}\{|S_Z|>M\}\Big] \\
&= \mu_S - \underbrace{M\cdot\mathbb{P}\Big(|S_Z|>M\Big)}_{T_1} - \underbrace{\mathbb{E}\Big[(|S_Z|-M)_+\Big]}_{T_2},
\end{aligned} \tag{8.59}$$

where step (i) is true by the condition (8.54) that $\mu_S \leqslant \frac{M}{2}$. We now analyze the two terms T_1 and T_2 separately. We define δ such that $M=(1+\delta)\mu_S$, and we have $\delta \geqslant 1$ by the construction of S in (8.54).

For the term T_1, we apply the multiplicative Chernoff bound. We have

$$\mathbb{P}\Big(|S_Z|>M\Big) \leqslant \mathbb{P}\Big(|S_Z|>2\mu_S\Big) \leqslant e^{-\frac{\mu_S}{3}}.$$

Then

$$T_1 \leqslant 2\mu_S\cdot e^{-\frac{\mu_S}{3}} \overset{\text{(i)}}{\leqslant} \frac{e}{6}, \tag{8.60}$$

where it can verified that step (i) holds for any $\mu_S \in \mathbb{R}$.

For the term T_2, note that $T_2 = \frac{V(S)}{\lambda}$, and by condition (8.54) we have $\mu_S \leqslant \frac{M}{2} \leqslant M$. Applying Lemma 33 with $k = \lceil \frac{1}{p_{\min}} \rceil$ yields

$$T_2 \leqslant \left\lceil \frac{1}{p_{\min}} \right\rceil \cdot \operatorname{Exp}\left(\frac{-2(M-\mu_S)^2}{|S|} \right) \cdot \left(1 - e^{\frac{-4(M-\mu_S)}{|S|} \left\lceil \frac{1}{p_{\min}} \right\rceil} \right)^{-2}.$$

Using the relations $|S| \leqslant \frac{\mu_S}{p_{\min}}$ and $M - \mu_S \geqslant \mu_S$, we have

$$\begin{aligned} T_2 &\leqslant \left\lceil \frac{1}{p_{\min}} \right\rceil \cdot \operatorname{Exp}\left(\frac{-2\mu_S^2}{\mu_S/p_{\min}} \right) \cdot \left(1 - e^{\frac{-4\mu_S}{\mu_S/p_{\min}} \left\lceil \frac{1}{p_{\min}} \right\rceil} \right)^{-2} \\ &= \left\lceil \frac{1}{p_{\min}} \right\rceil \cdot \operatorname{Exp}\left(-2p_{\min}\mu_S \right) \cdot \left(1 - e^{-4} \right)^{-2} \\ &\leqslant \left\lceil \frac{1}{p_{\min}} \right\rceil \cdot \left(1 - e^{-4} \right)^{-2}. \end{aligned} \tag{8.61}$$

Plugging term T_1 from (8.60) and term T_2 from (8.61) back to (8.59) yields

$$G(|S_Z|) \geqslant \mu_S - \frac{6}{e} - 1.04 \cdot \left\lceil \frac{1}{p_{\min}} \right\rceil.$$

Recall from the construction of S in (8.54) that $\frac{\mu_{[n]}}{3} \leqslant \mu_S \leqslant \frac{\mu_{[n]}}{2}$. Furthermore, by the assumption that $n \geqslant \frac{36}{p_{\min}^2}$, we have

$$\mu_{[n]} \geqslant np_{\min} \geqslant \frac{36}{p_{\min}} \geqslant \max\left\{ 36, 18 \left\lceil \frac{1}{p_{\min}} \right\rceil \right\}. \tag{8.62}$$

Hence, we have

$$G(|S_Z|) \geqslant \frac{\mu_{[n]}}{3} - \frac{\mu_{[n]}}{12} - \frac{\mu_{[\mu_S]}}{12} \geqslant \frac{\mu_{[n]}}{6}.$$

Applying inequality (8.51a) with the fact that $\mathbb{E}[|S_Z|] \leqslant \mu_{[n]}$, we have

$$U(S^*) \leqslant \lambda \cdot \mu_{[n]}$$

and hence

$$G(|S_Z|) \geqslant \frac{\mu_{[n]}}{6} \geqslant \frac{U(S^*)}{6\lambda}. \tag{8.63}$$

Finally, combining (8.58) and (8.63) yields

$$\frac{U(S)}{U(S^*)} \geqslant \frac{\lambda(1-\epsilon)^2 \cdot G(|S_Z|)}{6\lambda \cdot G(|S_Z|)} = \frac{(1-\epsilon)^2}{6},$$

yielding a constant-factor approximation with $\epsilon = \frac{p_{\min}}{4} \leqslant \frac{1}{4}$.

Runtime. MediumValueL$_1^+$ (Algorithm 15) begins by brute forcing over small sets, and there are $2^n \leqslant 2^{36/p_{\min}^2} \leqslant n^{36/p_{\min}^2}$ such sets. By Lemma 30, the objective value for each such set can be evaluated in polynomial time, and so the runtime in this case is $n^{3+36/p_{\min}^2}$.

For the other two cases (Line 3 and Line 5), the chosen set can be identified in $O(n)$. Therefore the overall runtime of MediumValueL$_1^+$ is $n^{O(1/p_{\min}^2)}$. □

Proof of Lemma 30

Recall from (8.1) that the objective is computed as $U([n]) = R([n]) - \lambda \cdot V([n])$, with

$$R([n]) := \sum_{i\in[n]} p_i x_i$$
$$V([n]) := \mathbb{E}\, \rho\left(\sum_{i\in[n]} Z_i,\ M\right).$$

It is clear that computing the reward term R may be done in $O(n)$ operations. We now show that the penalty term V can be computed in $O(n^2)$ operations.

We start by rewriting the term V as:

$$V([n]) = \sum_{k=0}^{n} \mathbb{P}\Big(\sum_{i\in[n]} Z_i = k\Big) \cdot \rho(k, M) \tag{8.64}$$

For any integer $m \in \{0, 1, \ldots, n\}$, we define the $(m+1)$-dimensional vector $\{w^{(m)}\}_{k=0}^m$ by

$$w_k^{(m)} := \mathbb{P}\Big(\sum_{i\in[m]} Z_i = k\Big).$$

Since we assume that the relevant values of ρ are known at the outset, it suffices to show that the probabilities involved in (8.64), or equivalently the $(n+1)$-dimensional vector $\{w_k^{(n)}\}_{k=0}^n$, can be computed in $O(n^2)$ operations.

We iteratively compute the vector of $\{w_k^{(m)}\}_{k=0}^m$ for $m \in \{0, 1, \ldots, n\}$. First, we observe that $w^{(0)} = 0$. Then we observe the iterative relation that for each $m \in [n]$ and $k \in \{0, \ldots, m\}$, we have

$$w_k^{(m)} = p_m \cdot w_{k-1}^{(m-1)} + (1 - p_m) \cdot w_k^{(m-1)}.$$

Hence, given the values of the m-dimensional vector $\{w_k^{(m-1)}\}_{k=0}^{m-1}$, computing each term $w_k^{(m)}$ takes c operations, where c is a universal constant. Hence, given the values of the m-dimensional vector $w_\cdot^{(m-1)}$, it takes $c(m+1)$ operations to compute the $(m+1)$-dimensional vector $w_\cdot^{(m)}$. Hence, the number of operations for computing the vector $w_\cdot^{(n)}$, by iteratively taking $m \in \{1, 2, \ldots, n\}$, is

$$c \sum_{m=1}^{n} (m+1) = O(n^2),$$

completing the proof. □

Proof of Lemma 31

We re-index the items $\{p_i, x_i\}_{i \in [n]}$ in decreasing order of the value x_i, such that $x_1 \geqslant \ldots \geqslant x_n$.

First, note that if $p_1 > \beta \cdot \mu_S$, then $T = \emptyset$ satisfies the lemma. Clearly for this T (8.34a) holds. Then because $R(S) \geqslant 0$ we also have

$$0 > \beta - \frac{p_i}{\mu_S} \geqslant \beta - \frac{1}{\mu_S} \geqslant \beta \left(1 - \frac{1}{\beta \cdot p_{\min}|S|}\right),$$

and so multiplying by $R(S)$ yields

$$R(T) = 0 > \beta \left(1 - \frac{1}{\beta \cdot p_{\min}|S|}\right) \cdot R(S),$$

satisfying (8.34b).

Otherwise we assume that $p_1 \leqslant \beta \cdot \mu_S$. We construct a set T be selecting as many items as possible in the decreasing order of the value x_i, subject to the constraint that (8.34a) is satisfied. Formally, we consider the set $T \coloneqq \{1, \ldots, t\}$, where

$$t \coloneqq \max \left\{ m \in [n] : \sum_{i=1}^{m} p_i \leqslant \beta \cdot \mu_S \right\}.$$

By the definition of t, the set T satisfies (8.34a). It remains to show that the set T also satisfies (8.34b).

If $\beta = 1$ then the resulting $T = S$ clearly suffices. Otherwise $\beta < 1$, and so we have $t < n$ (we assume that each item has strictly positive probability without loss of generality. By the definition of t, we have $\sum_{i=1}^{t+1} p_i > \beta \cdot \mu_S$. Equivalently,

$$\mu_T > \beta \cdot \mu_S - p_{t+1}. \tag{8.65}$$

In what follows, we use the following inequality that holds for any for $\{a_i\}_{i \in [n]}$ and $\{b_i\}_{i \in [n]}$ with $b_i \geqslant 0$:

$$\min_i \frac{a_i}{b_i} \leqslant \frac{\sum_i a_i}{\sum_i b_i} \leqslant \max_i \frac{a_i}{b_i}. \tag{8.66}$$

To see why this is true, note that for any $\{r_i\}_{i \in [n]}$ and $\{w_i\}_{i \in [n]}$, with $w_i \geqslant 0$ and $\sum_i w_i = 1$, we have

$$\min_i r_i \leqslant \sum_i w_i r_i \leqslant \max_i r_i.$$

We recover (8.66) by setting $r_i = \frac{a_i}{b_i}$ and $w_i = \frac{b_i}{\sum_i b_i}$.

Applying (8.66) yields

$$\frac{R(T)}{\mu_T} = \frac{\sum_{i \in T} p_i x_i}{\sum_{i \in T} p_i} \overset{\text{(i)}}{\geqslant} x_t \overset{\text{(ii)}}{\geqslant} \frac{\sum_{i \in S \setminus T} p_i x_i}{\sum_{i \in S \setminus T} p_i} = \frac{R(S \setminus T)}{\mu_S - \mu_T}, \tag{8.67}$$

where steps (i) and (ii) hold because the items are sorted in the decreasing order of x_i. Plugging $R(S) = R(T) + R(S \setminus T)$ into (8.67) and rearranging yields

$$\begin{aligned} R(T) &\geqslant \frac{\mu_T}{\mu_S} R(S) \\ &\overset{\text{(i)}}{\geqslant} \left(\beta - \frac{1}{\mu_S} \right) R(S) \\ &\overset{\text{(ii)}}{\geqslant} \left(\beta - \frac{1}{p_{\min} \cdot |S|} \right) R(S), \end{aligned}$$

where step (i) is true by (8.65), and step (ii) follows again from the fact that $\mu_S \geqslant p_{\min}|S|$. Hence, the set T satisfies (8.34b), completing the proof. $\square$

Proof of Lemma 32

To begin, we partition S into "high," "bucketable," and "leftover" items according to their p_i so that $S = H \sqcup B \sqcup L$ in Algorithm 16.

Algorithm 16 PARTITION

Require: $S \in [n]$, $p \in [0,1]^n$
Ensure: A partition $S = H \sqcup B \sqcup L$ with $B = D_1 \sqcup D_2 \sqcup D_3$
$H \leftarrow \{i \in S : p_i \geqslant \frac{1}{4}\}$
$L \leftarrow \{\}$
$D_1, D_2, D_3 \leftarrow \{\}$
for $\ell = 2, \ldots, \left\lceil \log_2(\frac{1}{p_{\min}}) \right\rceil - 1$ **do**
  $B^\ell \leftarrow \{i \in S : 2^{-(\ell+1)} \leqslant p_i < 2^{-\ell}\}$
  for $j = 0, \ldots, \left\lfloor \frac{|B^\ell|}{3} \right\rfloor - 1$ **do**
    $B^\ell_j \leftarrow \{b^\ell_{3j+1}, b^\ell_{3j+2}, b^\ell_{3j+3}\}$
    $D_1 \leftarrow D_1 \cup \{b^\ell_{3j+1}\}$
    $D_2 \leftarrow D_2 \cup \{b^\ell_{3j+2}\}$
    $D_3 \leftarrow D_3 \cup \{b^\ell_{3j+3}\}$
  $L \leftarrow L \cup \{b^\ell_{\left\lfloor 3\frac{|B^\ell|}{3} \right\rfloor+1}, \ldots, b^\ell_{|B^\ell|}\}$
return $S = H \sqcup B \sqcup L$ with $B = D_1 \sqcup D_2 \sqcup D_3$

First let $H = \{i \in S : p_i \geqslant \frac{1}{4}\}$, the high-probability items. Next consider the collection of buckets $B^\ell = \{i \in S \setminus H : 2^{-(\ell+1)} \leqslant p_i < 2^{-\ell}\}$. Note that the number of buckets is at most $\log_2(\frac{1}{p_{\min}})$. Form the contents of each B^ℓ into groups of three, $\{B^\ell_j\}_j$ (that is, the set $\left|B^\ell_j\right| = 3$ for each j. If the number of items in B^ℓ is not divisible by 3, we leave them to L). Let $B = \cup_\ell \cup_j B^\ell_j$, and let L be the leftover $L := S \setminus (H \cup B)$ which do not belong to groups of three.

Next note that $R(H) + R(B) + R(L) = R(S)$ and that $V(H), V(B), V(L) \leqslant V(S)$. We handle the cases when each of these is large separately.

Case 1: $R(H) \geqslant \frac{R(S)}{3}$. If $|H| \leqslant \frac{24}{p_{\min}}$ then the set H satisfies (8.36a). We now consider the case $|H| > \frac{24}{p_{\min}}$, and construct a set $T \subseteq H$ that satisfies (8.36b).

Applying Lemma 31 with $k = 6$ yields a set $T \subseteq H$ such that

$$\mu_T \leqslant \frac{1}{6}\mu_H \tag{8.68a}$$

and

$$R(T) \overset{\text{(i)}}{\geqslant} \frac{1}{6} \cdot \left(1 - \frac{6}{p_{\min} \cdot |H|}\right) \cdot R(H) \overset{\text{(ii)}}{\geqslant} \frac{1}{8}R(H), \tag{8.68b}$$

where step (i) follows from Lemma 31 and step (ii) is true by the assumption that $|H| > \frac{24}{p_{\min}}$. By the definition of H, we have $p_i \geqslant 1/4$ for each $i \in H$, and hence

$$|T| \leqslant 4\mu_T \overset{\text{(i)}}{\leqslant} \frac{2}{3}\mu_H \leqslant \frac{2}{3}\mu_S \overset{\text{(ii)}}{\leqslant} M,$$

where step (i) is true by (8.68a), and step (ii) is true by the assumption that $\mu_S \leqslant \frac{3}{2}M$. Hence, we have $V(T) = V(T') = 0$. By the rounding procedure, we have $R(T') = R(T)$. Therefore,

$$U(T') = U(T) = R(T) \overset{\text{(i)}}{\geqslant} \frac{1}{8}R(H) \overset{\text{(ii)}}{\geqslant} \frac{1}{24}R(S) \geqslant \frac{1}{24}U(S), \tag{8.69}$$

where step (i) is due to (8.68b) and step (ii) is true by the assumption of this case. Hence, the set T satisfies the condition (8.36b).

Case 2: $R(L) \geqslant \frac{R(S)}{3}$. Recall that the number of buckets is at most $\log_2(\frac{1}{p_{\min}})$. Since there are at most two elements in L from each bucket, the number of items in L is at most $|L| \leqslant 2\log_2(\frac{1}{p_{\min}}) < \frac{2}{p_{\min}}$, satisfying condition (8.36a).

Case 3: $R(B) \geqslant \frac{R(S)}{3}$. Further partition B into three equal-sized sets $B = D_1 \sqcup D_2 \sqcup D_3$ by arbitrarily assigning each member of each bucket-group B_j^ℓ to a distinct D_ℓ, as performed in each iteration of line 6. Without loss of generality, assume that D_1 has the maximum reward among these three sets, namely $R(D_1) \geqslant \max\{R(D_2), R(D_3)\}$, so that $R(D_1) \geqslant \frac{R(B)}{3} \geqslant \frac{R(S)}{9}$.

In what follows, we first show that the set D_1 satisfies (8.36b) under the assumption

$$V(D_1') \leqslant V(S). \tag{8.70}$$

Then we show that assumption (8.70) always holds.

Proving (8.36b) **for set** D_1**.** For the reward term, we have

$$R(D_1') = R(D_1) \geqslant \frac{1}{9}R(S).$$

For the penalty term, recall that we assume $\lambda \cdot V(S) \leqslant \frac{1}{15}R(S)$. Combining the reward term and the penalty term, we have

$$\begin{aligned} U(D_1') &= R(D_1') - \lambda \cdot V(D_1') \\ &\overset{\text{(i)}}{\geqslant} \frac{1}{9}R(S) - \lambda \cdot V(S) \\ &\geqslant \frac{1}{9}R(S) - \frac{1}{15}R(S) \\ &\geqslant \frac{1}{24}U(S), \end{aligned}$$

where step (i) uses assumption (8.70). Hence, the set D_1' satisfies condition (8.36b). It remains to prove assumption (8.70).

Proving (8.70)**.** For any sets S_1 and S_2, we say that S_1 stochastically dominates S_2, if the random variable $\sum_{i\in S_1} Z_i$ stochastically dominates the random variable $\sum_{i\in S_2} Z_i$. Namely, for any $t \in \mathbb{R}$, we have

$$\mathbb{P}\Big(\sum_{i\in S_1} Z_i \geqslant t\Big) \geqslant \mathbb{P}\Big(\sum_{i\in S_2} Z_i \geqslant t\Big)$$

Since the one-sided loss L_1^+ is nondecreasing, it can be verified that if S_1 stochastically dominates S_2, then $V(S_1) \geqslant V(S_2)$.

By construction we have $B \subseteq S$, and hence S stochastically dominates B. If B stochastically dominates D_1', then we have

$$V(D_1') \leqslant V(B) \leqslant V(S),$$

proving (8.70). It remains to prove that B stochastically dominates D_1'.

For each bucket group $B_z = B_j^\ell$, let $B_z = \{p_1, p_2, p_3\}$ with $p_1 \in D_1$. Then let the associated random variables be

$$X_z \coloneqq \mathrm{Ber}(q_1) \qquad \text{and} \qquad Y_z \coloneqq \mathrm{Ber}(p_1) + \mathrm{Ber}(p_2) + \mathrm{Ber}(p_3),$$

where q_i is obtained by rounding p_i up to the nearest power of two. Note that $\sum_{i \in D_1'} Z_i' = \sum_z X_z$, where the Z_i' are the realizations derived from the rounded probabilities, and $\sum_{i \in B} Z_i = \sum_z Y_z$. Moreover, $\{X_z\}_z$ are independent, and $\{Y_z\}_z$ are independent. It suffices to show the stochastic dominance of Y_z over X_z for each bucket group b_z, and then the stochastic dominance of B over D_1' follows.

To show the stochastic dominance of Y_z over X_z, we consider the probabilities

$$\begin{aligned}\mathbb{P}(X_z = 0) &= 1 - q_1 \\ \mathbb{P}(Y_z = 0) &= (1-p_1)(1-p_2)(1-p_3),\end{aligned}$$

and show that $\mathbb{P}(X_z = 0) \geqslant \mathbb{P}(Y_z = 0)$. By the construction of each bucket group B_ℓ, we have $p_1, p_2, p_3 \in [2^{-(l+1)}, 2^{-l})$ and hence $q_1 = 2^{-l}$. Consequently, we have $p_1, p_2, p_3 \geqslant \frac{q_1}{2}$. We have

$$\mathbb{P}(X_z = 0) = 1 - q_1 \overset{\text{(i)}}{\geqslant} \Big(1 - \frac{q_1}{2}\Big)^3 \geqslant (1-p_1)(1-p_2)(1-p_3) = \mathbb{P}(Y_z = 0).$$

where it can be verified that step (i) holds for every $q_1 \in [0, \frac{1}{4}]$. Hence, Y_z stochastically dominates X_z, completing the proof of (8.70). $\square$

Proof of Lemma 33

For notational simplicity, we assume $\lambda = 1$ without loss of generality, and denote the random variable $T \coloneqq \sum_{i \in S} Z_i$. We write the penalty term $V(S)$ as

$$\begin{aligned}V(S) &= \mathbb{E}\,(T - M)_+ \\ &= \sum_{i=1}^{\infty} i \cdot \mathbb{P}\Big(T = M + i\Big), \end{aligned} \tag{8.71}$$

We consider the probability that the value of T lies in each interval $(M+ik, M+(i+1)k]$, for each integer $i \geqslant 0$. We have

$$\begin{aligned} V(S) &\leqslant \sum_{i=0}^{\infty}(i+1)k \cdot \mathbb{P}\Big(M+ik < T \leqslant M+(i+1)k\Big) \\ &\leqslant k \cdot \sum_{i=0}^{\infty}(i+1) \cdot \mathbb{P}\Big(T > M+ik\Big). \end{aligned}$$

We bound each term $\mathbb{P}\Big(T > M+ik\Big)$ by Hoeffding's inequality. We have $\mathbb{E}[T] = \mu_S \leqslant M$ by assumption. Hence, by Hoeffding's inequality,

$$\begin{aligned} \mathbb{P}\Big(T > M+ik\Big) &\leqslant \operatorname{Exp}\left(-\frac{2(M+ik-\mu_S)^2}{|S|}\right) \\ &= \operatorname{Exp}\left(-\frac{2(M-\mu_S)^2}{|S|}\right) \cdot \operatorname{Exp}\left(-\frac{4(M-\mu_S)ik + 2(ik)^2}{|S|}\right) \\ &\leqslant \operatorname{Exp}\left(-\frac{2(M-\mu_S)^2}{|S|}\right) \cdot \operatorname{Exp}\left(-\frac{4(M-\mu_S)ik}{|S|}\right). \end{aligned} \tag{8.72}$$

Plugging (8.72) into (8.71) yields

$$\begin{aligned} V(S) &\leqslant k \cdot \operatorname{Exp}\left(-\frac{2(M-\mu_S)^2}{|S|}\right) \cdot \sum_{i=0}^{\infty}(i+1) \cdot \operatorname{Exp}\left(-\frac{4(M-\mu_S)ik}{|S|}\right) \\ &\leqslant k \cdot \operatorname{Exp}\left(-\frac{2(M-\mu_S)^2}{|S|}\right) \cdot \sum_{i=0}^{\infty}(i+1) \cdot \left(e^{-\frac{4(M-\mu_S)k}{|S|}}\right)^i \\ &\overset{\text{(i)}}{=} k \cdot \operatorname{Exp}\left(-\frac{2(M-\mu_S)^2}{|S|}\right) \cdot \left(1 - e^{\frac{-4(M-\mu_S)k}{|S|}}\right)^{-2}, \end{aligned}$$

where step (i) uses the fact that for any $0 < x < 1$, we have

$$\sum_{i=0}^{\infty}(i+1)x^i = \sum_{t=0}^{\infty}\sum_{i=t}^{\infty} x^i = \sum_{i=0}^{\infty}\frac{x^t}{1-x} = \frac{1}{1-x}\sum_{i=0}^{\infty} x^t = \frac{1}{(1-x)^2}.$$

□

8.7.9 Proof of Theorem 58

We fix any arbitrary problem instance (x, p, λ, M) for the L_1 loss, and problem instance (x', p, λ', M) for the L_1^+ loss, with

$$
\begin{aligned}
x'_i &:= x_i - \lambda \\
\lambda' &:= 2\lambda.
\end{aligned}
$$

We fix any arbitrary $S \subseteq [n]$, and demonstrate the desired equality

$$
U_{L_1}(S) = U_{L_1^+}(S') - \lambda \cdot M. \tag{8.73}
$$

by induction on the number of elements in S. First, we consider $|S| = 0$, or equivalently $S = \emptyset$. Then it can be verified that

$$
\begin{aligned}
U_{L_1}(S) &= -\lambda M \\
U_{L_1^+}(S') &= 0,
\end{aligned}
$$

satisfying (8.73).

Next suppose that (8.73) holds for all set S with $|S| \leqslant k$. We consider the marginal change to the objective when adding any item $j \notin S$ to the set S. Let $S_j =$ for some S with $|S| < j$. The marginal change to the objective with the L_1 loss is

$$
\begin{aligned}
U_{L_1}(S \cup \{j\}) - U_{L_1}(S) &= p_j x_j + \lambda \cdot \mathbb{E}_{Z_{S \cup \{j\}}} \left[\left| \sum_{i \in S} Z_i + Z_j - M \right| - \left| \sum_{i \in S} Z_i - M \right| \right] \\
&= p_j x_j + \lambda \cdot p_j \cdot \mathbb{E}_{Z_S} \underbrace{\left[\left| \sum_{i \in S} Z_i + 1 - M \right| - \left| \sum_{i \in S} Z_i - M \right| \right]}_{T}
\end{aligned} \tag{8.74}
$$

Note that the term T satisfies

$$
T = \begin{cases} 1 & \text{if } \sum_{i \in S} Z_i \geqslant M \\ -1 & \text{if } \sum_{i \in S} Z_i < M. \end{cases} \tag{8.75}
$$

Using the fact (8.75) in (8.74), we have

$$\begin{aligned}
U_{L_1}(S \cup \{j\}) - U_{L_1}(S) &= p_j x_j + \lambda \cdot p_j \left[\mathbb{P}\Big(\sum_{i \in S} Z_i \geqslant M\Big) - \mathbb{P}\Big(\sum_{i \in S} Z_i < M\Big)\right] \\
&= p_j x_j + \lambda \cdot p_j \left[2 \cdot \mathbb{P}\Big(\sum_{i \in S} Z_i \geqslant M\Big) - 1\right] \\
&= p_j(x_j - \lambda) + 2\lambda p_j \cdot \mathbb{P}\Big(\sum_{i \in S} Z_i \geqslant M\Big) \\
&= p_j x'_j + \lambda' p_j \cdot \mathbb{P}\Big(\sum_{i \in S} Z_i \geqslant M\Big). \qquad (8.76)
\end{aligned}$$

Using a similar analysis, the marginal change to the objective with the L_1^+ loss is

$$\begin{aligned}
U_{L_1^+}(S \cup \{j\}) - U_{L_1^+}(S) &= p_j x'_j + \lambda' \cdot \mathbb{E}_{Z_{S \cup \{j\}}} \left[\Big(\sum_{i \in S} Z_i + Z_j - M\Big)_+ - \Big(\sum_{i \in S} Z_i - M\Big)_+\right] \\
&= p_j x'_j + \lambda' p_j \cdot \mathbb{P}\Big(\sum_{i \in S} Z_i \geqslant M\Big). \qquad (8.77)
\end{aligned}$$

Combining (8.76) and (8.77) demonstrates that the marginal change is equal for the L_1 and L_1^+ losses, under their respective instances. Therefore, applying the induction hypothesis that (8.73) holds for all S with $|S| \leqslant k$ completes the induction step. $\square$

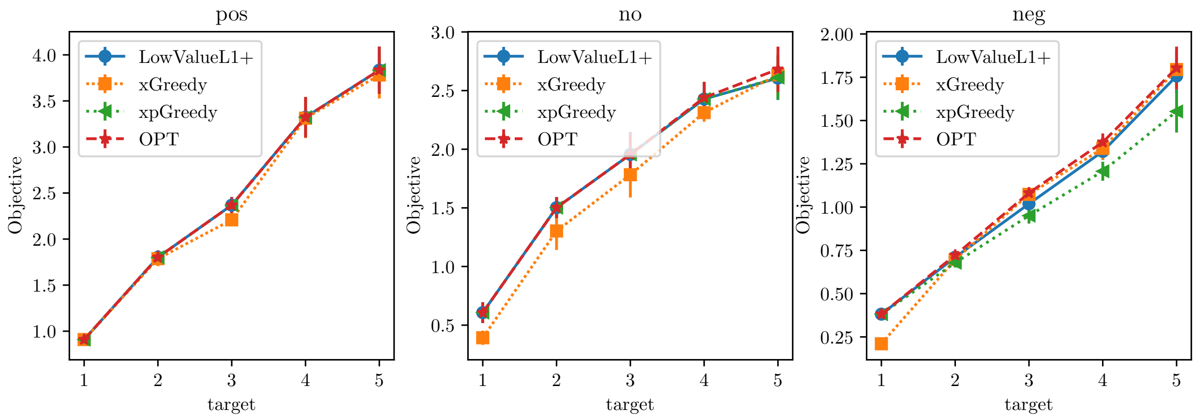


Figure 8.3: Sampling from the beta distribution with positive, no, and negative correlation. Here $n = 20$ and $\lambda = 3$, and OPT denotes $U(S^*)$.

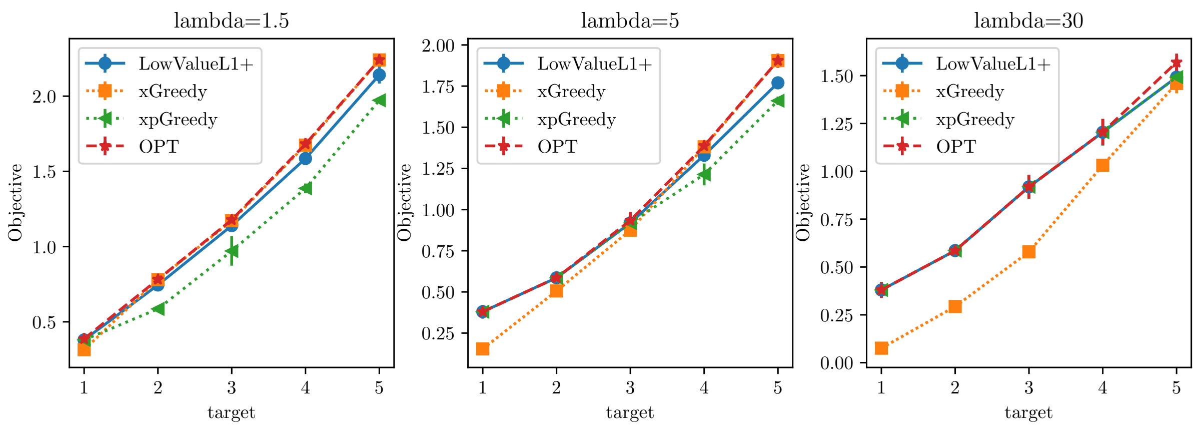


Figure 8.4: Performance for increasing penalty regularizer λ. Here $n = 20$ and sampling is via the negatively correlating beta distribution.

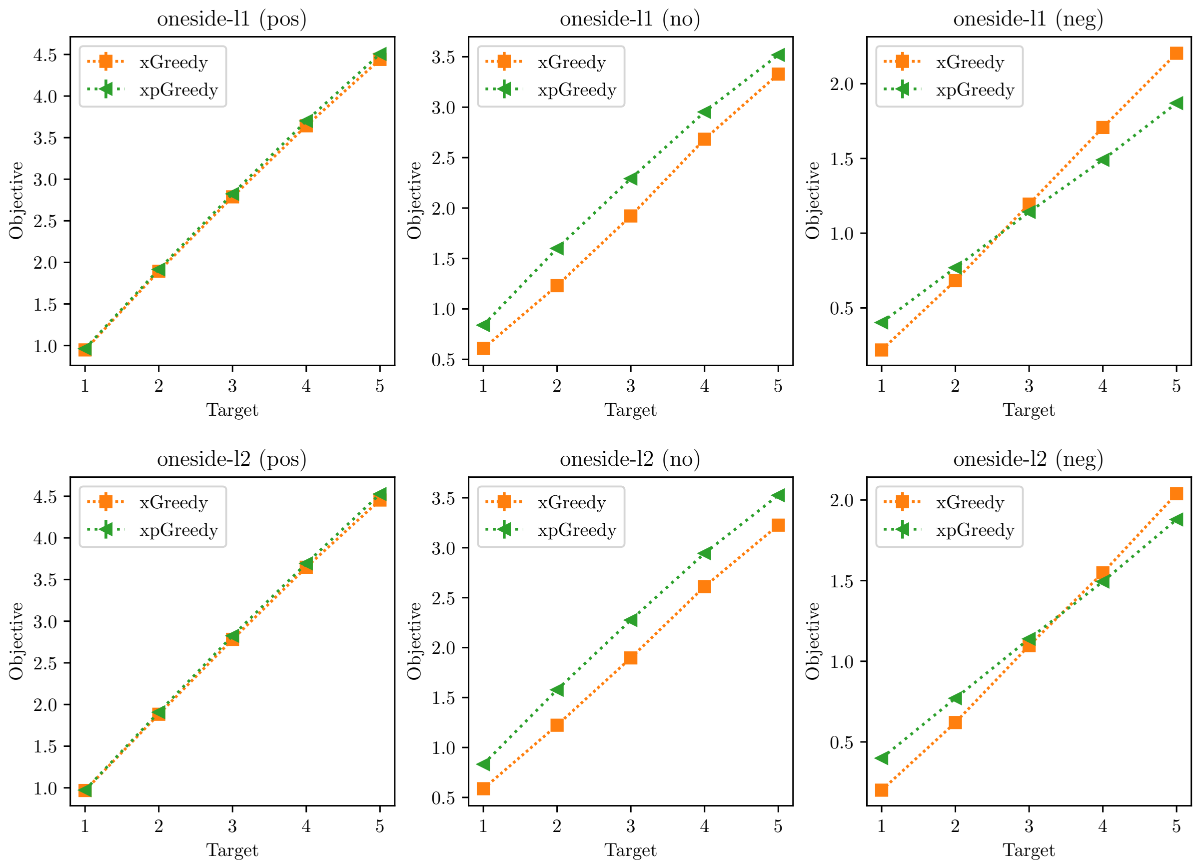


Figure 8.5: Evaluation of greedy heuristics for L_1^+ versus L_2^+ one-sided loss. Here $n = 50$ and $\lambda = 3$.

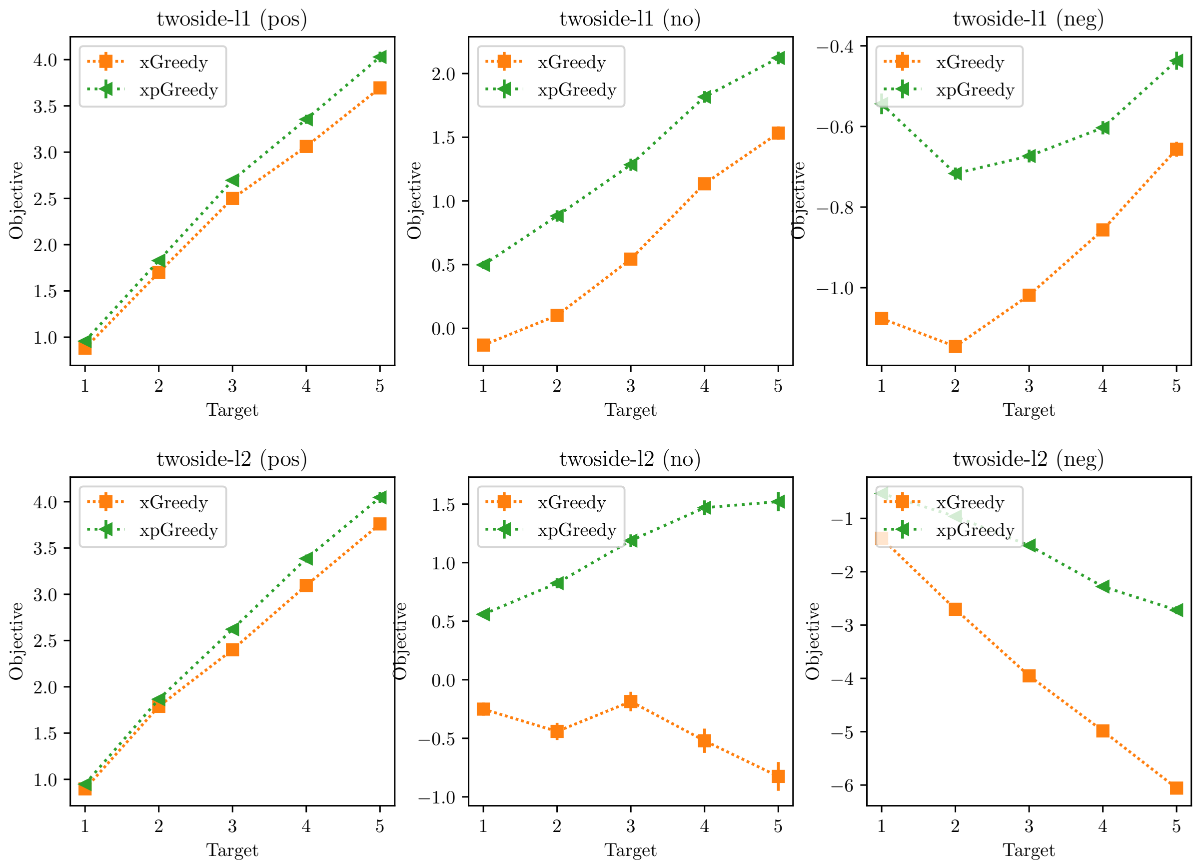


Figure 8.6: Evaluation of greedy heuristics for L_1 versus L_2 two-sided loss. Here $n = 50$ and $\lambda = 3$.

Chapter 9

Discussion

We conclude with more forward-looking reflections on the motivating challenges set forth in Chapter 1 and the progress the preceding chapters have endeavored to make towards meeting them.

The results established in Part I towards improving the robustness and scalability of preference and judgment aggregation rules and complementing current practice invite a number of avenues for further work. Some of these directions have already been identified and addressed. Brill and Peters [BP23] introduce a stronger notion of EJR, which they term EJR+, which is not only guaranteed to be satisfiable but (in contrast to EJR) can be efficiently verified for a given committee. Better yet, they demonstrate that this stronger condition can be efficiently attained in the online adaptive setting of Chapter 4, using an algorithm they term the Greedy Justified Candidate Rule.

Additional questions include what other elicitation formats and single- and multiwinner voting rules can be approximated in this model. For instance, there good evidence that general social choice rankings which take complete rankings cannot be approximated in this setting [BS20]. Another important consideration is what kinds of structural properties of large spaces of alternatives make this and related sparse adaptive query models more tractable. Fish et al.

[Fis+23] present interesting results in this direction, and also consider even larger latent spaces of alternatives to which query access is provided by oracles.

Related to these questions, participatory budgeting is a natural concrete setting into which to extend the work of Chapter 4 on querying procedures. Furthermore, the increasingly popular Method of Equal Shares of Peters, Pierczynski, and Skowron [PPS21] proceeds in a round-by-round manner which seems naturally suited to the online adaptive setting. Implementing this rule using sparse queries to voters could serve to improve the ease with which voters provide input and substantially extend the scope of practical instances to which the method can be applied. Furthermore, there are a number of possible choices of elicitation formats and models to consider in the participatory budgeting setting [Ben+21]. Since some elicitation formats lend themselves better to this setting than others (consider simple approval as compared to knapsack voting), it will be meaningful to establish which can be implemented at scale in this way, as well as which combinations of discrete versus continuous outcomes can be addressed [AS20].

The multi-objective model of sortition presented in Chapter 6 also raises a number of questions. Most immediately, the algorithmic results for attaining fairness-representation tradeoffs are established in the worst case over instance metrics. Given the simple and strict forms that the lower bound instances take, it would be highly interesting to establish potentially tighter results that are competitive on an instance-by-instance basis. Concretely: given a specific problem instance, is it possible to fix one objective value and attain a constant-factor approximation to the instance-specific optimal value of the other objective? Towards improving the practicality of the algorithms developed in Chapter 6, attaining finer-grained and even high-probability guarantees on the representation measures of chosen panels is also a priority. Panel selection algorithms which are likely to choose highly unrepresentative panels with significant probability will be untenable in practice.

More broadly, this framework seems potentially relevant to applications in robust network design and combinatorial optimization more broadly, under perhaps different measures of the equability of individual marginals. For $q = 1$ the model and results in Chapter 6 may be viewed as a robust $k-$medians problem, but this framework can be applied to other minimization and even maximization problems as well. There are likely further connections to be drawn to LP-based approaches for sampling solutions, for instance using maximum entropy distributions [KKG21].

Finally, is also optimistic but perhaps not unreasonable to expect that the combination of optimization and randomized elicitation strategies used in Part III to reduce the uncertainty of realized sizes in subset selection problems under simple (modular) objectives could be applicable to social choice problems. By way of example, consider the Michigan Citizens' Panel on COVID-19, which showcased the model of sortition which is considered in Chapter 5 [Of 20]. Their selection algorithm chose 30 participants from among the pool of volunteers, but three were unable to participate due to health and personal reasons. This disrupted the intended composition of their panel, affecting both participation probabilities for various groups and the reality of groups' shares of participating representatives in the deliberative process. This phenomenon could be at least partially mitigated by carefully augmenting the selection process with additional sampling rounds (or alternates) in the event of dropout. But modifying the practice of sortition to accommodate pool participation uncertainty in a more principled way could have other benefits as well.

In most modern instantiations of citizens' assemblies, randomly-selected members of the population are invited to enter the pool from which a panel is ultimately drawn, and this entails a commitment to participate for the duration of the citizens' assembly in the event that they are selected. A process with this form selects strongly against individuals who, due to their schedules or circumstances, cannot make commitments of this form. This bias could be potentially mitigated by a process which accepts pool participants who can only make probabilistic commitments to serve if selected, truthfully elicits these probabilities from the volunteers using the approach of Chapter 7, and uses them to design *adaptive* panel selection algorithms.

Bibliography

[AA18] B. Abramowitz and E. Anshelevich. “Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences”. In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018, pp. 894–901.

[Abe+15] J. Abelt, D. Browning, C. Dyer, M. Haines, J. Ross, P. Still, and M. Gerber. “Predicting likelihood of enrollment among applicants to the UVa undergraduate program”. In: *Systems and Information Engineering Design Symposium*. 2015, pp. 194–199.

[Ach+15] J. Acharya, I. Diakonikolas, C. Hegde, J. Z. Li, and L. Schmidt. “Fast and near-optimal algorithms for approximating distributions by histograms”. In: *Proceedings of the 34th*. 2015, pp. 249–263.

[Ans+18] E. Anshelevich, O. Bhardwaj, E. Elkind, J. Postl, and P. Skowron. “Approximating Optimal Social Choice under Metric Preferences”. In: *Artificial Intelligence* 264 (2018), pp. 27–51.

[Ans+21a] E. Anshelevich, A. Filos-Ratsikas, N. Shah, and A. A. Voudouris. “Distortion in Social Choice Problems: The First 15 Years and Beyond”. In: *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI) (Survey Track)*. 2021, pp. 4294–4301.

[Ans+21b] E. Anshelevich, A. Filos-Ratsikas, N. Shah, and A. A. Voudouris. “Distortion in social choice problems: an annotated reading list”. In: *ACM SIGecom Exchanges* 19.1 (2021), pp. 12–14.

[AP17] E. Anshelevich and J. Postl. “Randomized social choice functions under metric preferences”. In: *Journal of Artificial Intelligence Research (JAIR)* 58 (2017), pp. 797–827.

[Ara+17] P. Aragón, A. Kaltenbrunner, A. Calleja-López, A. Pereira, A. Monterde, X. E. Barandiaran, and V. Gómez. “Deliberative platform design: The case study of the online discussions in decidim Barcelona”. In: *Proceedings of the 9th International Conference on Social Informatics (SocInf)*. 2017, pp. 277–287.

[Arr90] K. Arrow. *Advances in the spatial theory of voting*. Cambridge University Press, 1990.

[Ary+04] V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit. "Local Search Heuristics for k-Median and Facility Location Problems". In: *SIAM J. Comput.* 33.3 (2004), pp. 544–562.

[AS16] 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)*. 2016, pp. 390–396.

[AS20] H. Aziz and N. Shah. "Participatory Budgeting: Models and Approaches". In: *CoRR* abs/2003.00606 (2020).

[AS98] A. Abdulkadiroğlu and T. Sönmez. "Random serial dictatorship and the core from random endowments in house allocation problems". In: *Econometrica* 66.3 (1998), pp. 689–701.

[Ash90] R. B. Ash. *Information Theory*. Dover Publications, 1990.

[AZ18] E. Anshelevich and W. Zhu. "Ordinal approximation for social choice, matching, and facility location problems given candidate positions". In: *Proceedings of the 14th Conference on Web and Internet Economics (WINE)*. Springer. 2018, pp. 3–20.

[Azi+15] H. Aziz, S. Gaspers, J. Gudmundsson, S. Mackenzie, N. Mattei, and T. Walsh. "Computational Aspects of Multi-Winner Approval Voting". In: *Proceedings of the 14th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 2015, pp. 107–115.

[Azi+17] H. Aziz, M. Brill, E. Elkind, R. Freeman, and T. Walsh. "Justified Representation in Approval-Based Committee Voting". In: *Social Choice and Welfare* 42.2 (2017), pp. 461–485.

[Azi+18] H. Aziz, E. Elkind, S. Huang, M. Lackner, L. Sánchez-Fernández, and P. Skowron. "On the Complexity of Extended and Proportional Justified Representation". In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018, pp. 902–909.

[Azi19] H. Aziz. "Justifications of welfare guarantees under normalized utilities". In: *ACM SIGecom Exchanges* 17.2 (2019), pp. 71–75.

[Bah+11] E. Baharad, J. Goldberger, M. Koppel, and S. Nitzan. "Distilling the wisdom of crowds: weighted aggregation of decisions on multiple issues". In: *Autonomous Agents and Multi-Agent Systems* 22.1 (Jan. 2011), pp. 31–42. ISSN: 1573-7454.

DOI: 10.1007/s10458-009-9120-y. URL: https://doi.org/10.1007/s10458-009-9120-y.

[Bah+12] E. Baharad, J. Goldberger, M. Koppel, and S. Nitzan. "Beyond Condorcet: optimal aggregation rules using voting records". In: *Theory and Decision* 72.1 (Jan. 2012), pp. 113–130. ISSN: 1573-7187.

[Bal+19] M.-F. Balcan, T. Dick, R. Noothigattu, and A. D. Procaccia. "Envy-Free Classification". In: *Proceedings of the 33rd Annual Conference on Neural Information Processing Systems (NIPS)*. 2019, pp. 1238–1248.

[BCD03] B. Babcock, S. Chaudhuri, and G. Das. "Dynamic sample selection for approximate query processing". In: *Proceedings of the 13th*. 2003, pp. 539–550.

[Ben+21] G. Benadè, S. Nath, A. D. Procaccia, and N. Shah. "Preference Elicitation for Participatory Budgeting". In: *Manag. Sci.* 67.5 (2021), pp. 2813–2827.

[Ber+08] J. Berg, R. Forsythe, F. Nelson, and T. Rietz. "Results from a dozen years of election futures markets research". In: *Handbook of Experimental Economics Results*. Ed. by C. R. Plott and V. L. Smith. Elsevier, 2008, pp. 742–751.

[Ber+18] R. Berk, H. Heidari, S. Jabbari, M. Kearns, and A. Roth. "Fairness in Criminal Justice Risk Assessments: The State of the Art". In: *Sociological Methods & Research* 50 (2018), pp. 3–44.

[BF81] J. Beck and T. Fiala. ""Integer-making" theorems". In: *Discrete Applied Mathematics* 3.1 (1981), pp. 1–8.

[BFT12] D. Bertsimas, V. F. Farias, and N. Trichakis. "On the Efficiency-Fairness Trade-Off". In: *Management Science* 58.12 (2012), pp. 2234–2250.

[BGK11] A. Bhalgat, A. Goel, and S. Khanna. "Improved Approximation Results for Stochastic Knapsack Problems". In: *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*. 2011, pp. 1647–1665.

[BGP19] G. Benadè, P. Gölz, and A. D. Procaccia. "No Stratification Without Representation". In: *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*. 2019, pp. 281–314.

[BJ12] P. Brändén and J. Jonasson. "Negative Dependence in Sampling". In: *Scandinavian Journal of Statistics* 39.4 (2012), pp. 830–838.

[Bou+12] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A. D. Procaccia, and O. Sheffet. "Optimal Social Choice Functions: A Utilitarian View". In: *Proceedings of the 13th ACM Conference on Economics and Computation (EC)*. 2012, pp. 197–214.

[Bou+15] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A. D. Procaccia, and O. Sheffet. "Optimal Social Choice Functions: A Utilitarian View". In: *Artificial Intelligence* 227 (2015), pp. 190–213.

[Bou02] C. Boutilier. "A POMDP Formulation of Preference Elicitation Problems". In: *Proceedings of the 18th AAAI Conference on Artificial Intelligence (AAAI)*. 2002, pp. 239–246.

[BP23] M. Brill and J. Peters. "Robust and Verifiable Proportionality Axioms for Multiwinner Voting". In: *Proceedings of the 24th ACM Conference on Economics and Computation (EC)*. Ed. by K. Leyton-Brown, J. D. Hartline, and L. Samuelson. ACM, 2023, p. 301.

[Bra+16] F. Brandt, V. Conitzer, U. Endress, J. Lang, and A. D. Procaccia, eds. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[Bri50] G. W. Brier. "Verification of forecasts expressed in terms of probability". In: *Monthly Weather Review* 78.1 (1950), pp. 1–3.

[BS20] M. Bentert and P. Skowron. "Comparing Election Methods Where Each Voter Ranks Only Few Candidates". In: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*. 2020, pp. 2218–2225.

[Bud+13] E. Budish, Y.-K. Che, F. Kojima, and P. Milgrom. "Designing random allocation mechanisms: theory and applications". en. In: *American Economic Review* 103.2 (2013), pp. 585–623.

[Car+17] I. Caragiannis, S. Nath, A. D. Procaccia, and N. Shah. "Subset Selection Via Implicit Utilitarian Voting". In: *Journal of Artificial Intelligence Research* 58 (2017), pp. 123–152.

[CDK17] Y. Cheng, S. Dughmi, and D. Kempe. "Of the people: voting is more effective with representative candidates". In: *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*. 2017, pp. 305–322.

[CDK18] Y. Cheng, S. Dughmi, and D. Kempe. "On the distortion of voting with multiple representative candidates". In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018, pp. 973–980.

[CDN07] S. Chaudhuri, G. Das, and V. Narasayya. "Optimized stratified sampling for approximate query processing". In: *ACM Transactions on Database Systems (TODS)* 32.2 (2007), article 9.

[Cha+01] S. Chaudhuri, G. Das, M. Datar, R. Motwani, and V. Narasayya. "Overcoming limitations of sampling for aggregation queries". In: *Proceedings of the 17th*. 2001, pp. 534–542.

[Che+07] Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. "A short introduction to computational social choice". In: *Proceedings of the 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM)*. 2007, pp. 51–69.

[Che+20] Y. Cheng, Z. Jiang, K. Munagala, and K. Wang. "Group Fairness in Committee Selection". In: *ACM Trans. Economics and Comput.* 8.4 (2020), 23:1–23:18.

[CHV18] L. E. Celis, L. Huang, and N. K. Vishnoi. "Multiwinner Voting with Fairness Constraints". In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. 2018, pp. 144–151.

[CK11] Y. Chen and I. A. Kash. "Information elicitation for decision making". In: *Proceedings of the 10th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 2011, pp. 175–182.

[CK16] Y.-K. Che and Y. Koh. "Decentralized College Admissions". In: *Journal of Political Economy* 124.5 (2016), pp. 1295–1338.

[ÇKM06] E. Çela, B. Klinz, and C. Meyer. "Polynomially solvable cases of the constant rank unconstrained quadratic 0-1 programming problem". In: *J. Comb. Optim.* 12.3 (2006), pp. 187–215.

[Con85] J.-A.-N. d. C. m. d. Condorcet. *Essai sur l'application de l'analyse á la probabilité des décisions rendus á la pluralité des voix*. Paris, Imprimerie Royale, 1785.

[CP11] I. Caragiannis and A. D. Procaccia. "Voting almost maximizes social welfare despite limited communication". In: *Artificial Intelligence* 175.9 (2011), pp. 1655–1671.

[CPS16] I. Caragiannis, A. Procaccia, and N. Shah. "Truthful univariate estimators". In: *Proceedings of the 33rd International Conference on Machine Learning (ICML)*. 2016, pp. 127–135.

[CS74] A. Cohen and H. B. Sackrowitz. "On Estimating the Common Mean of Two Normal Distributions". In: *Annals of Statistics* 2.6 (1974), pp. 1274–1282.

[CSV22] I. Caragiannis, N. Shah, and A. A. Voudouris. "The Metric Distortion of Multiwinner Voting". In: *Artificial Intelligence* 313 (2022), p. 103802.

[DG17] D. Dua and C. Graff. *UCI Machine Learning Repository*. 2017. URL: `http://archive.ics.uci.edu/ml`.

[DGV04] B. Dean, M. Goemans, and J. Vondrák. "Approximating the stochastic knapsack problem: The benefit of adaptivity". In: *45th Annual IEEE Symposium on Foundations of Computer Science*. 2004, pp. 208–217.

[Din+16] B. Ding, S. Huang, S. Chaudhuri, K. Chakrabarti, and C. Wang. "Sample + Seek: Approximating aggregates with distribution precision guarantee". In: *Proceedings of the 26th*. 2016, pp. 679–694.

[Din+21] F. Ding, M. Hardt, J. Miller, and L. Schmidt. "Retiring Adult: New Datasets for Fair Machine Learning". In: *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NIPS)*. 2021, pp. 6478–6490.

[DJ21] X. Dai and M. I. Jordan. "Learning Strategies in Decentralized Matching Markets under Uncertain Preferences". In: *Journal of Machine Learning Research* 22.260 (2021), pp. 1–50.

[DJR07] D. Dubhashi, J. Jonasson, and D. Ranjan. "Positive influence and negative dependence". In: *Combinatorics, Probability & Computing* 16.1 (2007), p. 29.

[Do+21] V. Do, J. Atif, J. Lang, and N. Usunier. "Online Selection of Diverse Committees". In: *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*. 2021, pp. 154–160.

[Doe07] B. Doerr. "Roundings respecting hard constraints". In: *Theory of Computing Systems* 40.4 (2007), pp. 467–483.

[DR96] D. Dubhashi and D. Ranjan. "Balls and Bins: A Study in Negative Dependence". In: *BRICS Report Series* 3.25 (1996), p. 30.

[Dwo+12] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. S. Zemel. "Fairness Through Awareness". In: *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS)*. 2012, pp. 214–226.

[Eba+22a] S. Ebadian, A. Kahng, D. Peters, and N. Shah. "Optimized Distortion and Proportional Fairness in Voting". In: *EC '22: The 23rd ACM Conference on Economics and Computation, Boulder, CO, USA, July 11 - 15, 2022*. ACM, 2022, pp. 563–600.

[Eba+22b] S. Ebadian, G. Kehne, E. Micha, A. D. Procaccia, and N. Shah. "Is Sortition Both Representative and Fair?" In: *NeurIPS*. 2022.

[EH84] J. M. Enelow and M. J. Hinich. *The spatial theory of voting: An introduction*. CUP Archive, 1984.

[EM23] S. Ebadian and E. Micha. "Boosting Sortition via Proportional Representation". In: *Working Paper* (2023).

[End16] U. Endriss. "Judgment Aggregation". In: *Handbook of Computational Social Choice*. Ed. by F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia. Cambridge University Press, 2016. Chap. 17.

[Eng89] F. Engelstad. "The assignment of political office by lot". In: *Social Science Information* 28.1 (1989), pp. 23–50.

[Fai+19] B. Fain, A. Goel, K. Munagala, and N. Prabhu. "Random Dictators with a Random Referee: Constant Sample Complexity Mechanisms for Social Choice". In: *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. 2019, pp. 1893–1900.

[Fer+17] L. S. Fernández, E. Elkind, M. Lackner, N. F. García, J. Arias-Fisteus, P. Basanta-Val, and P. Skowron. "Proportional Justified Representation". In: *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*. 2017, pp. 670–676.

[Fis+19] J. Fishkin, N. Garg, L. Gelauff, A. Goel, K. Munagala, S. Sakshuwong, A. Siu, and S. Yandamuri. "Deliberative democracy with the Online Deliberation platform". In: *Proceedings of the 7th AAAI Conference on Human Computation and Crowdsourcing (HCOMP)*. 2019, pp. 1–2.

[Fis+23] S. Fish, P. Gölz, D. C. Parkes, A. D. Procaccia, G. Rusak, I. Shapira, and M. Wüthrich. "Generative Social Choice". In: (2023).

[Fis18] J. Fishkin. *Democracy When the People Are Thinking: Revitalizing Our Politics Through Public Deliberation*. Oxford University Press, 2018.

[FKL16] B. Fish, J. Kun, and Á. D. Lelkes. "A confidence-based approach for balancing fairness and accuracy". In: *Proceedings of the 2016*. 2016, pp. 144–152.

[FKP21] B. Flanigan, G. Kehne, and A. D. Procaccia. "Fair Sortition Made Transparent". In: *Proceedings of the 35th Annual Conference on Neural Information Processing Systems (NIPS)*. 2021, pp. 25720–25731.

[Fla+20a] B. Flanigan, P. Gölz, A. Gupta, and A. D. Procaccia. "Neutralizing Self-Selection Bias in Sampling for Sortition". In: *Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NIPS)*. 2020, pp. 6528–6539.

[Fla+20b] B. Flanigan, P. Gölz, A. Gupta, A. D. Procaccia, and G. Rusak. *Panelot*. `http://www.panelot.org/`. 2020.

[Fla+21] B. Flanigan, P. Gölz, A. Gupta, B. Hennig, and A. D. Procaccia. "Fair Algorithms for Selecting Citizens' Assemblies". In: *Nature* 596 (2021), pp. 548–552.

[FSV20] R. Freeman, N. Shah, and R. Vaish. "Best of both worlds: ex-ante and ex-post fairness in resource allocation". In: *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*. 2020, pp. 21–22.

[Gan+06] R. Gandhi, S. Khuller, S. Parthasarathy, and A. Srinivasan. "Dependent rounding and its applications to approximation algorithms". In: *Journal of the ACM (JACM)* 53.3 (2006), pp. 324–360.

[GAX17] 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)*. 2017, pp. 544–550.

[GBN22] S. Ganguly, R. Basu, and V. Nagarajan. "A binomial decision tree to manage yield-uncertainty in multi-round academic admissions processes". In: *Naval Research Logistics (NRL)* 69.2 (2022), pp. 303–319.

[GHS20] V. Gkatzelis, D. Halpern, and N. Shah. "Resolving the optimal metric distortion conjecture". In: *Proceedings of the 61st Symposium on Foundations of Computer Science (FOCS)*. 2020, pp. 1427–1438.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: a Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.

[GKM11] A. Ghosh, S. Kale, and P. McAfee. "Who Moderates the Moderators? Crowdsourcing Abuse Detection in User-Generated Content". In: *Proceedings of the 12th ACM Conference on Economics and Computation (EC)*. 2011, pp. 167–176.

[GKM17] 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)*. 2017, pp. 287–304.

[GN98] R. Gray and D. Neuhoff. "Quantization". In: *IEEE Transactions on Information Theory* 44.6 (1998), pp. 2325–2383.

[Goo52] I. J. Good. "Rational decisions". In: *Journal of the Royal Statistical Society* 14.1 (1952), pp. 107–114.

[GR07] T. Gneiting and A. E. Raftery. "Strictly proper scoring rules, prediction, and estimation". In: *Journal of the American Statistical Association* 102.477 (2007), pp. 359–378.

[Gri04] G. Grimmett. "Stochastic apportionment". In: *The American Mathematical Monthly* 111.4 (2004), pp. 299–307.

[GW19] J. Gastil and E. O. Wright, eds. *Legislature by Lot: Transformative Designs for Deliberative Governance*. Verso, 2019.

[Hal+21] D. Halpern, G. Kehne, D. Peters, A. D. Procaccia, N. Shah, and P. Skowron. "Aggregating Binary Judgments Ranked by Accuracy". In: *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*. 2021, pp. 5456–5463.

[Hal+22] D. Halpern, G. Kehne, A. D. Procaccia, J. Tucker-Foltz, and M. Wüthrich. "Representation with Incomplete Votes". In: *CoRR* abs/2211.15608 (2022).

[HG20] B. Hennig and P. Gölz. *Sortition Foundation Stratification Application.* `https://github.com/sortitionfoundation/stratification-app`. 2020.

[Hoe63] W. Hoeffding. "Probability inequalities for sums of bounded random variables". In: *Journal of the American Statistical Association* 58.301 (1963), pp. 13–30.

[IKZ09] L. Iandoli, M. Klein, and G. Zollo. "Enabling On-Line Deliberation and Collective Decision-Making through Large-Scale Argumentation: A New Approach to the Design of an Internet-Based Mass Collaboration Platform". In: *International Journal of Decision Support System Technology (IJDSST)* 1.1 (Jan. 2009), pp. 69–92.

[Iri19] T. Irish Citizens' Assembly Project. `http://www.citizenassembly.ie/work/`. 2019.

[Ito+20] T. Ito, S. Suzuki, N. Yamaguchi, T. Nishida, K. Hiraishi, and K. Yoshino. "D-Agree: Crowd Discussion Support System Based on Automated Facilitation Agent". In: *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI).* 2020, pp. 13614–13615.

[Jag+98] H. V. Jagadish, N. Koudas, S. Muthukrishnan, V. Poosala, K. C. Sevcik, and T. Suel. "Optimal histograms with quality guarantees". In: *Proceedings of the 24th.* 1998, pp. 24–27.

[JDS11] H. Jégou, M. Douze, and C. Schmid. "Product Quantization for Nearest Neighbor Search". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.1 (2011), pp. 117–128.

[Jen69] R. E. Jensen. "A Dynamic Programming Algorithm for Cluster Analysis". In: *Operations Research* 17.6 (1969), pp. 1034–1057.

[JK96] S. M. Jordan and K. Krishnamoorthy. "Exact Confidence Intervals for the Common Mean of Several Normal Populations". In: *Biometrics* 52.1 (1996), pp. 77–86.

[Kem20a] D. Kempe. "An analysis framework for metric voting based on LP duality". In: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI).* 2020, pp. 2079–2086.

[Kem20b] D. Kempe. "Communication, Distortion, and Randomness in Metric Voting". In: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI).* 2020, pp. 2087–2094.

[KK22] A. Kahng and G. Kehne. "Worst-Case Voting When the Stakes Are High". In: *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI).* 2022, pp. 5100–5107.

[KKG21] A. R. Karlin, N. Klein, and S. O. Gharan. "A (slightly) improved approximation algorithm for metric TSP". In: *Proceedings of the 53rd Annual ACM Symposium on Theory of Computing (STOC)*. ACM, 2021, pp. 32–45.

[KKP20] A. Kahng, G. Kehne, and A. D. Procaccia. "Strategyproof Mean Estimation from Multiple-Choice Questions". In: *Proceedings of the 37th International Conference on Machine Learning (ICML)*. Vol. 119. PMLR, 2020, pp. 5042–5052.

[KPW22] G. Kehne, A. D. Procaccia, and J. Wang. "Recruitment Strategies That Take a Chance". In: *NeurIPS*. 2022.

[KR13] N. Kumar and B. Raichel. "Fault Tolerant Clustering Revisited". In: *Proceedings of the 25th Canadian Conference on Computational Geometry, CCCG*. Carleton University, Ottawa, Canada, 2013.

[KRT00] J. Kleinberg, Y. Rabani, and É. Tardos. "Allocating Bandwidth for Bursty Connections". In: *SIAM Journal on Computing* 30 (Jan. 2000), pp. 191–217.

[Lan19] J. Lang. "In Praise of Nonanomymity: Nonsubstitutable Voting". In: *The Future of Economic Design*. Ed. by J.-F. Laslier, H. Moulin, R. Sanver, and W. S. Zwicker. Springer, 2019, pp. 97–102.

[Llo82] S. Lloyd. "Least squares quantization in PCM". In: *IEEE Transactions on Information Theory* 28.2 (1982), pp. 129–137.

[LS09] N. Lambert and Y. Shoham. "Eliciting truthful answers to multiple-choice questions". In: *Proceedings of the 10th ACM Conference on Economics and Computation (EC)*. 2009, pp. 109–118.

[LS18] J. Lang and P. Skowron. "Multi-attribute proportional representation". In: *Artificial Intelligence* 227 (2018), pp. 74–106.

[LS93] A. L. Liestman and T. C. Shermer. "Additive graph spanners". In: *Networks* 23.4 (1993), pp. 343–363.

[LW20] Y. Luo and Y. Wang. *Dynamic Decision Making Under Rolling Admissions: Evidence from U.S. Law School Applications*. Tech. rep. tecipa-681. University of Toronto, Department of Economics, 2020.

[Mil69] J. C. Miller. "A Program for Direct and Proxy Voting in the Legislative Process". In: *Public Choice* 7 (1969), pp. 107–113. (Visited on 10/19/2022).

[Mou03] H. Moulin. *Fair division and collective welfare*. MIT press, 2003.

[MP16] S. J. Maglio and E. Polman. "Revising probability estimates: Why increasing likelihood means increasing impact." In: *Journal of Personality and Social Psychology* 111.2 (2016), p. 141.

[MW13] N. Mattei and T. Walsh. "Preflib: A library for preferences http://www. preflib. org". In: *International Conference on Algorithmic Decision Theory*. Springer. 2013, pp. 259–270.

[MW19] K. Munagala and K. Wang. "Improved metric distortion for deterministic social choice rules". In: *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*. 2019, pp. 245–262.

[New21] New York City Department of Education. *Random Selection in Admissions*. `https://www.schools.nyc.gov/enrollment/enroll-grade-by-grade/how-students-get-offers-to-doe-public-schools/random-numbers-in-admissions`. 2021.

[Noo+18] R. Noothigattu, S. S. Gaikwad, E. Awad, S. Dsouza, I. Rahwan, P. Ravikumar, and A. D. Procaccia. "A Voting-Based System for Ethical Decision Making". In: *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*. AAAI Press, 2018, pp. 1587–1594.

[NP82] S. Nitzan and J. Paroush. "Optimal Decision Rules in Uncertain Dichotomous Choice Situations". In: *International Economic Review* 23.2 (1982), pp. 289–297. ISSN: 00206598, 14682354.

[OEC20] OECD. *Innovative Citizen Participation and New Democratic Institutions*. OECD Publishing, 2020.

[Of 20] Of by for *. *Final report of the citizens' panel on COVID-19*. en. 2020. URL: `https://joinofbyfor.org/wp-content/uploads/2020/11/Final-Report-of-the-Citizens-Panel-on-COVID-19.pdf`.

[Ols94] A. L. Olsen. "Penalty Functions and the Knapsack Problem". In: *Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence, June 27-29, 1994*. IEEE, 1994, pp. 554–558.

[Par11] J. M. Parker. "Randomness and Legitimacy in Selecting Democratic Representatives". PhD thesis. University of Texas at Austin, 2011.

[PGR19] M. Purohit, S. Gollapudi, and M. Raghavan. "Hiring Under Uncertainty". In: *Proceedings of the 36th International Conference on Machine Learning*. Vol. 97. 2019, pp. 5181–5189.

[PPS21] D. Peters, G. Pierczynski, and P. Skowron. "Proportional Participatory Budgeting with Additive Utilities". In: *NeurIPS*. 2021, pp. 12726–12737.

[PR06] 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)*. 2006, pp. 317–331.

[PST21] S. Perez-Salazar, M. Singh, and A. Toriello. “Robust Online Selection with Uncertain Offer Acceptance”. In: *CoRR* abs/2112.00842 (2021). arXiv: `2112.00842`.

[RB08] V. F. Reyna and C. J. Brainerd. “Numeracy, ratio bias, and denominator neglect in judgments of risk and probability”. In: *Learning and Individual Differences* 18.1 (2008), pp. 89–107.

[Rep21] E. R. 9. E. S. S. (E.-9. 2. D. Report. “Edition 3.1. Bergen, European Social Survey Data Archive, NSD - Norwegian Centre for Research Data for ESS ERIC.” In: (2021). URL: `https://www.europeansocialsurvey.org/data/`.

[RF14] G. Radanovic and B. Faltings. “Incentives for truthful information elicitation of continuous signals”. In: *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*. 2014, pp. 770–776.

[RV97] P. Resnick and H. R. Varian. “Recommender Systems”. In: *Communications of the ACM* 40.3 (1997), pp. 56–58.

[SE17] P. K. Skowron and E. Elkind. “Social choice under metric preferences: Scoring rules and STV”. In: *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*. 2017, pp. 706–712.

[SG06] L. Siefert and F. Galloway. “A New Look at Solving the Undergraduate Yield Problem: The Importance of Estimating Individual Price Sensitivities”. In: *College and University* 81.3 (2006), pp. 11–17.

[SH18] M. Soloviev and J. Y. Halpern. “Information Acquisition Under Resource Limitations in a Noisy Environment”. In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018, pp. 6443–6450.

[Shi+19] D. Shibata, A. Moustafa, T. Ito, and S. Suzuki. “On Facilitating Large-Scale Online Discussions”. In: *PRICAI 2019: Trends in Artificial Intelligence*. Springer International Publishing, 2019, pp. 608–620.

[Sko+17] P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. “Proportional Rankings”. In: *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*. 2017, pp. 409–415.

[Sko21] P. Skowron. “Proportionality Degree of Multiwinner Rules”. In: *Proceedings of the 22nd ACM Conference on Economics and Computation (EC)*. 2021, pp. 820–840.

[SL15] M. J. Salganik and K. E. C. Levy. “Wiki surveys: open and quantifiable social data collection”. In: *PloS one* 10.5 (May 2015), e0123483.

[SLG21] M. Seddighin, M. Latifian, and M. Ghodsi. "On the Distortion Value of Elections with Abstention". In: *Journal of Artificial Intelligence Research (JAIR)* 70 (2021), pp. 567–595.

[Sma+21] C. Small, M. Bjorkegren, T. Erkkilä, L. Shaw, and C. Megill. "Polis: Scaling Deliberation by Mapping High Dimensional Opinion Spaces". In: *Revista De Pensament I Anàlisi* 26.2 (2021).

[Smi75] M. H. Smith. "A Secretary Problem with Uncertain Employment". In: *Journal of Applied Probability* 12.3 (1975), pp. 620–624. ISSN: 00219002.

[Spe85] J. Spencer. "Six Standard Deviations Suffice". In: *Transactions of the American Mathematical Society* 289.2 (1985), pp. 679–706.

[Sri97] A. Srinivasan. "Improving the Discrepancy Bound for Sparse Matrices: Better Approximations for Sparse Lattice Approximation Problems". In: *Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1997, pp. 692–701.

[Sto11] P. Stone. *The Luck of the Draw: The Role of Lotteries in Decision Making*. Oxford University Press, 2011.

[Tam91] M. Tamaki. "A Secretary Problem with Uncertain Employment and Best Choice of Available Candidates". In: *Operations Research* 39.2 (1991), pp. 274–284.

[Tra+20] A. Trask, E. Bluemke, B. Garfinkel, C. G. Cuervas-Mons, and A. Dafoe. *Beyond Privacy Trade-offs with Structured Transparency*. arXiv 2012.08347. 2020.

[TZ92] A. D. Taylor and W. S. Zwicker. "A Characterization of Weighted Voting". In: *Proceedings of the American Mathematical Society* 115.4 (1992), pp. 1089–1094.

[TZ99] A. D. Taylor and W. S. Zwicker. *Simple Games: Desirability Relations, Trading, Pseudoweightings*. Princeton University Press, 1999.

[Van16] D. Van Reybrouck. *Against Elections: The Case for Democracy*. Random House, 2016.

[WC14] B. Waggoner and Y. Chen. "Output agreement mechanisms and common knowledge". In: *Proceedings of the 2nd AAAI Conference on Human Computation and Crowdsourcing (HCOMP)*. 2014, pp. 220–226.

[Win69] R. L. Winkler. "Scoring rules and the evaluation of probability assessors". In: *Journal of the American Statistical Association* 64.327 (1969), pp. 1073–1078.

[WML20] S. J. Westwood, S. Messing, and Y. Lelkes. "Projecting confidence: How the probabilistic horse race confuses and demobilizes the public". In: *The Journal of Politics* 82.4 (2020), pp. 1530–1544.

[WX12] T. Walsh and L. Xia. "Lot-based voting rules". In: *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2012, pp. 603–610.

[WZ04] J. Wolfers and E. Zitzewitz. "Prediction markets". In: *Journal of Economic Perspectives* 18.2 (2004), pp. 107–126.

[XC08] L. Xia and V. Conitzer. "Generalized scoring rules and the frequency of coalitional manipulability". In: *Proceedings of the 9th ACM Conference on Electronic Commerce (EC)*. 2008, pp. 109–118.

[Yao77] A. C. Yao. "Probabilistic Computations: Towards a Unified Measure of Complexity". In: *Proceedings of the 17th Symposium on Foundations of Computer Science (FOCS)*. 1977, pp. 222–227.

[You75] H. P. Young. "Social Choice Scoring Functions". In: *SIAM Journal on Applied Mathematics* 28.4 (1975), pp. 824–838.

[ZP21] L. Zhang and J. C. Pippins. "Application of overbooking model to college admission at The Citadel". In: *International Transactions in Operational Research* 28.5 (2021), pp. 2402–2413.

[ZR08] A. Zohar and J. S. Rosenschein. "Mechanisms for information elicitation". In: *Artificial Intelligence* 172.16-17 (2008), pp. 1917–1939.