Algorithms for Public Decision Making

by

Brandon Fain

Department of Computer Science
Duke University

Date: ________________________
Approved:

____
Kamesh Munagala, Supervisor

____
Vincent Conitzer

____
Pankaj K. Agarwal

____
Ashish Goel

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Computer Science
in the Graduate School of Duke University
2019

Abstract

Algorithms for Public Decision Making

by

Brandon Fain

Department of Computer Science
Duke University

Date: ______________________
Approved:

____
Kamesh Munagala, Supervisor

____
Vincent Conitzer

____
Pankaj K. Agarwal

____
Ashish Goel

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science in the Graduate School of Duke University
2019

Abstract

In public decision making, we are confronted with the problem of aggregating the conflicting preferences of many individuals about outcomes that affect the group. Examples of public decision making include allocating shared public resources and social choice or voting. We study these problems from the perspective of an algorithm designer who takes the preferences of the individuals and the constraints of the decision making problem as input and efficiently computes a solution with provable guarantees with respect to fairness and welfare, as defined on individual preferences.

Concerning fairness, we develop the theory of group fairness as core or proportionality in the allocation of public goods. The core is a stability based notion adapted from cooperative game theory, and we show extensive algorithmic connections between the core solution concept and optimizing the Nash social welfare, the geometric mean of utilities. We explore applications in public budgeting, multi-issue voting, memory sharing, and fair clustering in unsupervised machine learning.

Regarding welfare, we extend recent work in implicit utilitarian social choice to choose approximately optimal public outcomes with respect to underlying cardinal valuations using limited ordinal information. We propose simple randomized algorithms with strong utilitarian social cost guarantees when the space of outcomes is metric. We also study many other desirable properties of our algorithms, including approximating the second moment of utilitarian social cost. We explore applications in voting for public projects, preference elicitation, and deliberation.

Contents

Acknowledgements

The work represented in this document would have been impossible without the support of many teachers and mentors. I should especially thank my doctoral adviser Kamesh Munagala. He had faith in me as a young graduate student that I would only later develop in myself, and I owe much of my scholarly expertise to him. I admire his determination to seek and work on problems that are interesting and relevant in society as well as the scholarly community, and to emphasize constructive algorithmic approaches that prefer simplicity.

Debmalya Panigrahi, Pankaj Agarwal, Vincent Conitzer, and Rong Ge were all important teachers in my graduate education at Duke; I am indebted to them for my training as a computer scientist. Rong taught me much about algorithms for machine learning and optimization. From Vincent, I learned a great deal about algorithmic game theory, mechanism design, and computational social choice. Pankaj was my first and very excellent teacher of algorithms in graduate school. In addition to teaching approximation and graph algorithms, Debmalya has always provided clear and useful feedback. I should also note that I had the privilege of teaching with Carlo Tomasi, John Reif, Ashwin Machanavajjhala, and Xi He, and that teaching with these individuals formed me just as significantly as the classes I myself took from others.

I have also been fortunate to have many wonderful collaborators in my research over the years. In addition to my adviser, these collaborators include Ashish Goel, Nisarg Shah, Nina Prabhu, Xingyu Chen, Liang Lyu, and Mayuresh Kunjir. Mayuresh was a more senior graduate student who worked graciously with me on my first research project during my first year of graduate school. Xingyu and Liang are both undergraduates at Duke who worked diligently despite their academic obligations, and I am grateful for having had the opportunity to mentor and assist them. Nina was a high school senior at the North Carolina School of Science and Math when she began working on a project with me, and is now an undergraduate computer science student at Stanford. Nisarg Shah was finishing his PhD at Carnegie Mellon when we met, and is now a professor at the University of Toronto. We have many shared interests in research, and talking with Nisarg has always encouraged me about the possibilities of our work, as well as clarified the problems at hand. Ashish was been deeply involved in establishing many of the fundamental questions in this dissertation as research interests for me (namely participatory budgeting and deliberation), and brings a rich breadth of experience to these issues that spans theory and practice. The contributions of these collaborators to specific chapters are noted at the beginning of each chapter.

A dissertation requires more than intellectual support. Marilyn Butler deserves special thanks for her role in supporting me as a staff person through the entire program from admission to graduation. I must also thank my family, the other students in the computer science department at Duke, the Graduate Christian Fellowship at Duke, Blacknall Presbyterian Church, and especially Seaver Wang and Olivia Brown for forming the community that grounded, supported, and encouraged me. The ideas of these individuals and communities are not present in this dissertation, but it would surely have been impossible without them.

Finally, my work was financially supported by Duke University and by NSF grants CCF-1408784, CCF-1637397, and IIS-1447554.

1

Introduction

"If a house is divided against itself, that house cannot stand."[1]

Any collection of individuals seeking to work together must make decisions that affect the whole, even when they disagree about the right decision. This tension is endemic to all modes of human society, and is increasingly relevant to collections of artificially intelligent agents as well as automated algorithmic decision making via machine learning. In this thesis, I design and analyze principled algorithms for making these kinds of public decisions.

There are many aspects to public decision making. The processes by which we come to decisions are deeply rooted in traditions and institutions. The individuals involved in decision making may be complex human beings with desires, anxieties, relationships, fears, and hopes. While I do not intend to dispute the significance of these aspects of lived decision-making, I do wish to suggest that there are deep problems in addition to these historical, emotional, and social concerns. The aspects of public decision making I consider in this thesis are normative and algorithmic

[1] Jesus of Nazareth according to the Gospel of Mark

in nature: even if we could somehow cleanly separate out the preferences of the individual agents, how should we aggregate those preferences into a single decision, and how would we justify our procedure in terms of those preferences?

I will tend to consider problems as follows. We will have a set of agents and a set of possible outcomes. Feasible outcomes may be constrained in some way. For example, perhaps we only have so many dollars to spend. Agents will have preferences over possible outcomes, and we will generally represent these preferences as utilities.[2] We will design efficient algorithms that take as input the preferences of the agents, and output a feasible outcome that provides some provable guarantee in terms of the preferences of the agents. In so doing, we will develop principled solutions for public decision making.

This thesis considers two broad approaches. In the first chapters, we will consider fair resource allocation of public goods. Here, we consider guarantees that fairly respect the relative entitlements of the agents, ensuring that no group of agents is poorly treated. More specifically, we adapt a game theoretic notion of stability called the *core* as a novel fair solution concept in public decision making or public resource allocation. Subsequently, we study the computation of core solutions under a variety of settings including divisible and indivisible public goods, linear and concave utilities, and metric centroid clustering. Along the way, we consider applications in public budgeting, voting, memory sharing, and fair machine learning. Resource allocation is a fundamental problem in economics and computer science, but most of

[2] The use of the word "utility" connotes many things to many people. In the philosophical tradition, it may indicate a historically Anglo-American approach to ethics. I will generally use utilities to quantify the benefit an individual derives from a particular outcome of public decision making. I am not ideologically committed to the notion that an individual is ethical in so far as she makes decisions that maximize the arithmetic mean of societal utilities. Also, the fact that I generally think of utilities as quantifying the benefit to an individual of an outcome does not imply that I assume the agents are selfish or malicious. If one prefers, one can think of utilities as quantifying the extent to which an agent perceives a given outcome as good in the broad sense. In so far as agents disagree, deep problems remain.

the prior work in the fair resource allocation literature focused on private goods, for which the solution concepts tend to be either ill-defined or too-weak when considering public goods. Nevertheless, we find interesting connections with prior and ongoing work in terms of our algorithmic approaches, some of which are based on the *Nash social welfare* function.

In the second part of the thesis, we consider public decision making as social choice. We study making a single public decision in the *implicit utilitarian model* where we want to minimize the utilitarian social cost but we only know ordinal preferences of agents (i.e., we are not given utilities, only rankings over outcomes). We design simple randomized mechanisms that approximately minimize the utilitarian social cost when the space of possible outcomes forms a metric. We extend prior work in this setting by providing mechanisms that have additional desirable properties like approximating the second moment of the distribution of social cost (where the randomness comes from the randomization in the mechanism) and using only a constant number of queries.

Though our treatment will often be mathematical and necessarily abstract, the broad problem I have outlined is not. Americans are often prone to wonder whether conducting presidential elections by the popular vote or the electoral college would be better. Every year, tax dollars are allocated at the federal, state, and local levels to fund public projects and resources. Algorithms and machine learning are increasingly used for making public decisions about granting loans and bail, matching employers and employers, school admissions, and more. The average family or group of friends may be familiar with the inscrutable difficulties of agreeing on a place to eat dinner. Public decisions are made one way or another, and through mathematical analysis, we can hopefully come to see the tradeoffs more clearly across diverse scenarios.

It is important to clarify at the outset that my goal is constructive, not descriptive or rhetorical. I seek to provide clean and precise abstract descriptions of problems in public decision making, to present reasonable and precise notions of a good solution, and to provide efficient algorithms for constructing such solutions. I will not answer two other extremely interesting kinds of questions. First, I will not consider the development of a broad theory of justice from which particular statements of rights or entitlements may be derived. While I argue for the reasonableness of various normative guarantees in this thesis, I make no attempt to justify them as necessary. Second, I will not consider the normative question for individuals. That is, I will not ask questions of the sort "what should some particular agent do when engaged in making a public decision?" I will always take the normative perspective with respect to the mediator or designer of a mechanism, and only ever assume strategic or sincere behavior on the part of agents.

1.1 Fair Resource Allocation Overview

In this section, I describe some of the most important concepts in the area of fair resource allocation in general terms as a reference and a gentle introduction that will relate to our discussion of both fair resource allocation for public goods and implicit utilitarian social choice. Individual chapters further refine notation and definitions for specific applications; this introduction is intended to clarify something of the breadth of possible models for related problems. In particular, many of these notions are stated for private goods as is standard practice, and we will need to slightly adapt the notions for public goods.

A setting (or model) for a fair resource allocation problem consists of a set N of $|N| = n$ agents and a set M of $|M| = m$ goods. If the goods are divisible then an allocation is constituted by values $x_{ij} \in \mathbb{R}_{\geqslant 0}$ for every agent $i \in N$ and good $j \in M$ denoting the amount of good j that agent i receives. If the goods are indivisible then an allocation is a partition $X = \{X_1, ..., X_n\}$ such that for all $i \in N$, $X_i \subseteq M$, and for all $i, i' \in N$ such that $i \neq i'$, $X_i \cap X_{i'} = \emptyset$.

Denote the set of generic (divisible or indivisible) feasible allocations as $Y(M)$. An agent has ordinal preferences if she has a complete and transitive ordering $\succeq_i$ over $Y(M)$ where $y_1 \succeq_i y_2$ for $y_1, y_2 \in Y(M)$ should be read as "agent i weakly prefers allocation y_1 to allocation y_2." An agent has cardinal preferences if she has ordinal preferences and a utility function $u_i : Y(M) \to \mathbb{R}$ satisfying $u_i(y_1) \geqslant u_i(y_2)$ if and only if $y_1 \succeq_i y_2$. In my work, I generally focus on cardinal preferences. Finally, a good is private if it can be allocated to at most one agent, i.e., all other agents receive 0 utility from the good. A good is public if it is allocated generally such that every agent may receive nonzero utility from the good.

Given a specific model, a typical fair allocation problem defines an objective or certain desiderata that constitute a "good" allocation and proceeds to devise an algorithm that optimizes that objective or satisfies those desiderata. Some of the most standard objectives are called social welfare functions, which are functions of the utilities of the agents (note that such functions are only well defined for cardinal preferences). Let $\mathbf{u} \in \mathbb{R}^n$ be a vector of the utilities of agents. The three most common social welfare functions are defined below.

- The utilitarian social welfare (USW) is defined as

  $$USW(\mathbf{u}) = \frac{1}{n} \sum_{i \in N} u_i.$$

  Informally, maximizing the utilitarian social welfare provides "the greatest good for the greatest number" and maximizes the efficiency of an allocation.

- The Nash social welfare (NSW) is defined as

  $$NSW(\mathbf{u}) = \left(\prod_{i \in N} u_i \right)^{1/n}.$$

Informally, maximizing the Nash social welfare balances the efficiency and the fairness of an allocation.

- The egalitarian social welfare (ESW) is defined as

$$ESW(\mathbf{u}) = \min_{i \in N} u_i.$$

Informally, maximizing the egalitarian social welfare provides the most good possible for the worst off individual, and is typically considered the strictest fairness objective among social welfare functions.

Much of the fair allocation literature attempts to guarantee certain desiderata, rather than maximize an objective function. We list some common guarantees. Let y_i be the generic allocation to agent i.

- Envy-Freeness. An allocation is envy-free if for all $i, i' \in N, y_i \succeq_i y_{i'}$.
- Proportionality. An allocation is proportional if for all $i \in N, u_i(y_i) \geqslant \frac{u_i(M)}{n}$.
- Pareto efficiency or Pareto optimality. A generic allocation $y \in Y(M)$ is Pareto efficient or Pareto optimal if there does not exist another feasible allocation $z \in Y(M)$ such that $z \succeq_i y$ for all $i \in N$ and $z \succ_{i'} y$ for some $i' \in N$ (where $z \succ_{i'} y$ means $z \succeq_{i'} y$ but $y \not\succeq_{i'} z$).[3]

Finally, the following concepts are the most important for analyzing the strategic behavior of rational agents. Suppose that we want to design an algorithm or mechanism[4] for fair allocation, $A(\cdot)$, and we elicit information in the form of a type vector

[3] Note that while Pareto efficiency is a weak notion of efficiency, it can be nontrivial to combine with envy-freeness or proportionality.

[4] Some authors distinguish between an algorithm and a mechanism. Those who do typically use mechanism for algorithms that specifically account for strategic behavior. In generally, I do not make this distinction, and use algorithm, protocol, and mechanism interchangeably and synonymously in most contexts.

$\mathbf{t}$. Every agent $i \in N$ has a true type τ_i (typically these types will be some form of utility information for different outcomes) and reports a type t_i (not necessarily equal to τ_i) to the mechanism. Let $\mathbf{t}_{-i}$ be the reported types of all agents apart from agent i. Let $u_i(\tau_i, A(\mathbf{t}))$ be agent i's utility given their true type τ_i under the outcome of the mechanism given type vector $\mathbf{t}$ (or the expected utility if $A(\cdot)$ is randomized).

- Truthfulness. $A(\cdot)$ is truthful or strategyproof or dominant-strategy incentive-compatible if $\forall i \in N$, $\forall \tau_i$, and $\forall \mathbf{t}$, $u_i(\tau_i, A(\tau_i, \mathbf{t}_{-i})) \geqslant u_i(\tau_i, A(\mathbf{t}))$. Informally, if a mechanism is truthful, no agent has an incentive to misreport their type regardless of the reports of the other agents.

- Truthful Equilibrium. $A(\cdot)$ has truth telling as a Nash equilibrium (also called ex-post equilibrium) if $\forall i \in N$, all agents reporting τ_i and receiving payoff $u_i(\tau_i, A(\tau_1, \ldots, \tau_n))$ is a Nash equilibrium. Informally, a mechanism has truth telling as a Nash equilibrium if no agent has an incentive to misreport their type, assuming no other agents are misreporting their types.

1.2 High Level Survey of Related Work and Applications

An application to which we will repeatedly return is called *participatory budgeting* [1]. In many cities around the world, a certain amount of public tax money is allocated directly according to votes elicited from citizens. Some other canonical examples of fair division like splitting taxi fare, splitting rent [73], and dividing up indivisible goods have been implemented in software by a recent web deployment `http://www.spliddit.org/`. Fair allocation algorithms have also been deployed for allocating course slots to students [37]. A significantly scaled up application of fair resource allocation in recent years is the allocation of resources in a datacenter environment (historically, fairness was of little concern and efficiency was obviously the right objective for these kinds of problems, but now data centers are often populated with tasks from different users with priorities and rights). For examples, see [76, 139, 100, 135, 72].

1.2.1 Fair Allocation.

A comprehensive technical survey of the fair allocation (also known as fair division) literature would require its own book, as the field has matured substantially over the last two decades, bringing together substantial contributions from computer science and economics. Rather than attempting such a Herculean task, it is my intent in this section to provide the reader with a high level overview of some of the topics that are most related to this thesis. Individual chapters will provide additional related work that is specific to and relevant for the content of the chapter. The interested reader can also find a survey of topics in fair resource allocation in Part 2 of [34].

Cake Cutting. The canonical problem of fairly allocating divisible private goods is called *cake cutting*. The model is usually described by an interval $[0, 1]$ over which n agents have some utility function. The problem was outlined in a survey by Ariel Procaccia in 2013 [141]. Generally, one seeks an algorithm within a discrete query model for computing an envy-free allocation. For two agents, this is always possible and the algorithm is trivial: ask one agent to cut the "cake" into two pieces such that each has equal value; ask the second agent to choose among these two pieces. For three agents, a slightly less trivial algorithm has been known, but a recent breakthrough showed (i) a discrete envy-free protocol for 4 agents [19] and (ii) a discrete envy-free protocol for n agents [18].

Heterogeneous Resources. The fair allocation problem may become substantially more complicated when users have non-additive demands over multiple heterogeneous goods. This case has been highlighted in the context of algorithms for resource allocation in data centers [76, 139, 100, 135], where the utility jobs receive for memory, processor time, and bandwidth interact in a way typically described by Leontief preferences, rather than additive. The crucial idea behind the breakthrough work of dominant resource fairness [76] is to consider fairness with respect to dominant resources, which one can think of as the bottleneck resource for a given job.

Integral Allocations: Envy-freeness. Fair allocation also becomes more difficult in the case where the goods are integral, in the sense that they cannot be fractionally allocated to the agents. In this context, it is straightforward to see that an envy-free allocation is not always possible: simply consider the case where there is a single good and two agents who desire it.[5] For this reason, there is an extensive hierarchy of fair solution concepts developed in the literature. The most straightforward relaxation of envy-freeness is envy-freeness up to one good, which says that an allocation would be envy-free if, for every pair of agents, we considered removing one good from the allocation to the envied agent. Such solutions always exist for monotone utility functions [119], and in the special case of additive utility, finding the integral allocation that maximizes the Nash social welfare satisfies the property [41] and is also Pareto efficient. Unfortunately, finding the integral allocation that maximizes the Nash social welfare is APX-hard [115], although constant factor approximation algorithms are known [50, 49, 6, 24].

A solution concept that lies between envy-freeness and envy-freeness up to one good is called envy-freeness up to *any* good. Such a solution would be envy-free if, for every pair of agents, we considered removing *any* good of the envied agent.

[5] This assumes, of course, some minimal efficiency guarantee such as we must allocate all of the goods or slightly stronger, Pareto optimality. Also, envy-freeness in expectation is still achievable with a randomized mechanism, but such solutions are often unsatisfactory in the context of fair division, as we are often performing a single allocation of goods that may not be repeated. For example, we might be dividing goods among inheritors of an estate. For this reason, the literature tends to focus on ex-post or deterministic guarantees, rather than guarantees in expectation.

The leximin solution, which maximizes the egalitarian social welfare and breaks ties lexicographically by the least, second least, and so forth utility, computes an envy-free up to any good solution for two agents [136]. Existence for more agents remains an open question.

Integral Allocations: Maximin Share Guarantee. Another extensively studied solution concept is the *maximin share guarantee*, which was introduced by Budish in [37] along with the notion of approximate competitive equilibrium from equal incomes. The maximin share of an agent is the utility they would receive from their least favorite bundle of goods under the partition of the goods into n bundles that maximizes their utility from that least favorite bundle. Procaccia and Wang initiated the computational study of maximin share guarantees [143], which has since seen extensive follow up work in the literature [23, 5, 77]. The state of the art is that a 3/4-multiplicative approximation to the maximin share guarantee always exists and can be efficiently computed for additive valuations. For submodular valuations, the best known approximation factor is 1/3.

Fair Allocation of Public Goods. As we will argue in this thesis, allocating private goods is a special case of the more general problem of allocating public goods. Therefore, while solution concepts for public goods do translate to the private goods case, the converse is not necessarily true, and new approaches are necessary. In particular, public goods can be shared, and traditional fair solution concepts for private goods do not take this into account. Instead, we advocate for considering group fairness that considers guarantees to more general subsets of agents than just singleton individuals. In addition to our own work, such group guarantees have been considered in many contexts, including private goods allocation [22, 54] and multi-winner approval voting in social choice [16, 149].

1.2.2 Other Related Topics in Computer Science and Economics

Our treatment will necessarily involve some tools and techniques from outside of the fair allocation literature itself, and in this section I briefly survey some of the more relevant such topics at a high level. Again, individual chapters will provide citations for any references necessary for the chapter.

Market Equilibrium. The culmination of neoclassical economic theory is Walrasian market equilibrium [99]. Traditionally, economists have been interested in the existence and properties of equilibrium. More recently however, computer scientists have become interested in the computability of equilibrium, showing both hardness [158, 44] and efficient methods [98, 97]. This study is partially motivated by wanting to understand market dynamics: if an equilibrium is computationally hard to compute, should we expect it to be reached under simple dynamics? However, there is also interest in computing a market equilibrium outright for other applications in fair allocation, where the "money" in the market is virtual. The most common example is a market based notion of fairness called competitive equilibrium from equal incomes [37] where the idea is that an allocation is fair if it is a market equilibrium computed from agents with equal incomes. Fisher market equilibrium proved invaluable in the analysis that Cole and Gkatzelis provided of an algorithm to approximately maximize the Nash social welfare for indivisible goods [50]. We crucially use a line of work developed for public goods equilibrium, called the Lindahl equilibrium [118, 101, 71], to compute fair allocations of public goods subject to a budget constraint in Chapter 2.

Game Theory and Bargaining. Game theory is a normative study of how rational and self-interested (or strategic) agents should behave in a situation where the actions of the agents determine the payoffs, or utilities, the players receive. The fundamental solution concept for noncooperative games (those in which the players cannot coordinate their actions) is the Nash equilibrium: a state such that no agent has an incentive to change their action given the actions of the other agents. A fundamental computational result in algorithmic game theory is the proof that computing a Nash equilibrium is a hard problem, more specifically, it is the canonical problem for the PPAD complexity class (polynomial parity arguments on directed graphs) [58].

Cooperative game theory studies situations where the agents *can* coordinate their behavior; a canonical problem with two agents is the bargaining problem, first phrased in game theoretic terms by Nash [130], who showed that under a set of four axioms, the Nash social welfare maximizer is the unique solution to such a game. Rubenstein provided additional justification for Nash's solution to the bargaining problem by giving a mechanism that implements the solution as a Nash equilibrium in a noncooperative game [146]. We use Nash's solution concept extensively in Chapter 5. Kalai and Smorodinsky provided alternative solution concepts under related sets of axioms [104, 103]. The study of the bargaining problem has also led to an attempt to justify the use of cardinal valuations [154, 89]. Myerson expanded the model to account for incomplete information [128]. A stronger solution concept that we consider extensively in Chapter 2 is the core, first expounded in game theoretic terms by Scarf [150]. Informally, an outcome is in the core if no subset of agents can deviate and select for themselves a preferable outcome. Scarf provides sufficient, but not necessary, conditions for the existence of the core.

Mechanism Design. Mechanism design attempts to create functions or rules for allocation problems where the agents who report information to the algorithm have preferences over the outcomes of the algorithm and must be incentivized to provide correct information. Two classic notions of such incentives are truthfulness and Nash equilibrium, described in Section 1.1. Designing truthful mechanisms is often difficult. Two classic impossibility results include the Gibbard-Satterthwaite impossibility theorem for truthful social choice with ordinal valuations [78] and the Myerson-Satterthwaite impossibility theorem for truthful bilateral trade [129]. As a more recent example, Cole et. al examine the fair allocation of divisible private goods in a strategyproof context without money [51]. They design an algorithm that, for every agent, provides at least a 1/e fraction of the utility they would have received under the proportionally fair allocation, while being strategyproof, but they also provide a 1/2 upper bound for the same problem.

There are, however, several novel techniques and approaches. In 2007, McSherry and Talwar presented the exponential mechanism, a way to use differential privacy to construct approximately truthful randomized mechanisms for optimization problems [93]. We use this technique in Chapter 2. In 2013, Cai, Daskalakis, and Weinberg provided a black box reduction from designing Bayesian incentive compatible, ex post individually rational, arbitrarily feasibility constrained mechanisms optimizing arbitrary objectives to algorithm design for the underlying optimization plus virtual revenue and virtual welfare [39]. Subsequently, they expanded the work to include ex-post constraints [57]. A separate approach to mechanism design considers showing that the incentive to misreport often (though not always) diminishes in large games or large markets. Azevedo and Budish defined the notion of strategyproofness in the large along these lines [14] and Feldman et. al showed how to bound the Price of Anarchy (a notion of the worst case quality of Nash equilibrium introduced in [109]) in large games [68].

Utilitarian Social Choice. Informally, social choice theory studies how to pick a point from a decision space that represents the common good of society. The canonical example is voting. In that case, the decision space is usually a small set of candidates, and the agents comprising the society express only ordinal valuations over these candidates. Typically, social choice is studied under ordinal preferences, despite the classic impossibility result from Arrow [11]. In *utilitarian social choice*, we assume that agents have cardinal valuations (i.e., utility functions) over the possible decisions. Some recent work in this area takes an *implicit* utilitarian perspective, assuming that agents cannot report their full utility functions but that their reported ordinal valuations are generated by an underlying cardinal utility function [31, 7, 25, 46]. Such work tries to bound the *distortion*, the worst case utilitarian social welfare over all utility functions that could have generated the ordinal reports of agents. Our work in Chapter 5 and Chapter 6 is in this model.

Much of utilitarian social choice focuses on specific structures for the decision space and the utilities agents have over it. For combinatorially structured decision spaces, there is a long line of work [151, 134, 36, 60] on combinatorial public projects. These results focus on truthful mechanism design and the winner determination problem. There is ongoing related work on candidate selection even for simple analytical models like points in $\mathbb{R}$ [67]. Median graphs and their ordinal generalization, median spaces, have been extensively studied in the context of social choice. The special cases of trees and grids have been examined as structured models for voter preferences [153, 21]. For general median spaces, it is known that the Condorcet winner (that is an alternative that pairwise beats any other alternative in terms of voter preferences) is strongly related to the generalized median [20, 147, 159] – if the former exists, it coincides with the latter. [132] shows that any single-peaked domain which admits a non-dictatorial and neutral strategy-proof social choice function is a median space. [48] also showed that any set of voters and alternatives on a median graph will have a Condorcet winner. In a sense, these are the largest class of structured and spatial preferences where ordinal voting over alternatives leads to a "clear winner" even by pairwise comparisons. Goel and Lee [82] recently worked with median graphs in devising a sequential process for reaching concsensus. We follow up on these ideas with our work in Chapter 5.

The notion of *democratic equilibrium* [75] considers social choice mechanisms in continuous spaces where individual agents with complex utility functions perform update steps inspired by gradient descent, instead of ordinal voting on the entire space. Several works have considered *iterative voting* where the current alternative is put to vote against one proposed by different random agent chosen each step [4, 116, 144], or other related schemes [124]. The analysis tends to focus on convergence to an equilibrium instead of welfare or efficiency guarantees.

Textbooks and References. The now classic text on applications of game theory within an algorithmic context is [133]. A recent book on computational social choice serves as an excellent survey and reference [34]. My preferred text on approximation algorithms is [160]. I have found [117] useful in the study of Markov chains, and [10] for its analysis of the multiplicative weights method. My favorite book on game theory is an old one by John Harsanyi [88], but I have also used the more recent book by Yoav Shoham and Kevin Leyton-Brown [155].

1.3 Outline of Results.

Chapter 2. In this chapter, we introduce the fair participatory budgeting problem as a canonical example of public decision making. There are multiple public projects that we can fund subject to a budget constraint. We introduce the notion of the core from cooperative game theory as a fair solution concept that generalizes the idea of proportionality to all subsets of agents. We show that for additive and linear utility, maximizing the Nash social welfare is a core solution. This can be implemented efficiently as the Proportional Fairness convex program, and we briefly discuss application to memory allocation. We go on to show that we can adapt the exponential mechanism from differential privacy to design an approximately truthful mechanism for computing approximately core solutions with high probability. We then generalize the class of utility functions under consideration, and show that we can still compute core solutions via a generalization of the Proportional Fairness program for a fairly broad class of separable concave utility functions.

Chapter 3. Here, we consider the problem of public decision making where the outcomes must be chosen integrally and deterministically. We use a broad model that generalizes much of the prior literature for indivisible fair allocation: there is a set of public goods, from which we must choose a feasible subset. We work with additive valuations, and our results are differentiated based on the type of feasibility constraints that can be imposed.

It is easy to see that core solutions and their multiplicative approximations may not exist for integral allocations. Nevertheless, for matroid and matching constraints, we show that any local maximum of a smoothed version of the Nash social welfare function is a constant approximate additive core solution, and that such local maxima can be efficiently computed via local search techniques. The matroid setting is sufficient to model the problem where there are a number of binary issues, and we must choose yes or no for each (i.e., multiple simultaneous referendums). Matroid constraints are also sufficient to model the multi-winner election problem, where there are some number of candidates, and we want to elect some subset of the candidates of constrained size (also known as committee selection). For general packing constraints, where each good has an associated multi-dimensional cost, and feasibility is defined by packing constraints on these costs, we show that the previous approach is inadequate. Instead, we compute the fractional maximum Nash social welfare solution and a certain maximally fair fractional solution that we define. We mix these fractional solutions and design a hierarchical rounding algorithm that yields a nearly logarithmic additive approximate core solution in polynomial time.

Chapter 4. In this chapter, we consider centroid clustering as a problem in public decision making. In particular, we suppose that the data points to be clustered are individuals who prefer to be accurately clustered, in terms of small distance to their center in some metric space. We adapt the notion of core in this context and call it proportional clustering. We note that traditional approaches to clustering like optimizing the k-median, k-means, and k-center objectives need not produce proportional clusterings, and indeed, exactly proportional clusterings may not exist. Nevertheless, we give a greedy algorithm for efficiently computing a constant approximate proportional clustering. Furthermore, we formulate approximate proportionality as a linear constraint, and show how to adapt rounding techniques to approximately minimize the k-median objective subject to proportionality as a constraint. We show that proportionality is approximately preserved under random sampling, and that this idea can be used to more efficiently compute and audit proportional clustering solutions. Finally, we consider the empirical tradeoff between proportionality and the k-means objective on real data sets.

Chapter 5. This begins the second part of this thesis, wherein we consider implicit utilitarian social choice. We suppose that we want to choose a single outcome from a set of alternatives so as to minimize the utilitarian social cost, but we do not know the costs of the agents for the alternatives. We design a randomized sequential deliberation protocol based on iterative Nash bargaining and show that it has expected social cost within a constant factor of the social optimum when the space of alternatives are metric (that is, when costs satisfy the triangle inequality). Furthermore, we show that the second moment of social cost is also within a constant factor of the squared social cost of the social optimum. We call this concept of approximation on the second moment of cost *Squared Distortion*, and interpret it as a quantification of the risk associated with a randomized mechanism. We show stronger results (in terms of approximation of the social optimum, but also uniqueness of stationary distribution, Pareto efficiency, and Nash equilibria) in the special case where the space of alternatives lie on a median graph, which can be thought of as graphs or metric spaces that embed isometrically into hypercubes (potentially of high dimension).

Chapter 6. In this chapter, we continue working in the implicit utilitarian model of social choice using randomized mechanisms when the space of alternatives forms a metric. However, in this chapter we propose a simple query model and study mechanisms that are constrained to make a constant number of queries, motivated by the practical concerns of social choice in very large spaces of alternatives where lightweight mechanisms are desirable. We show that any mechanism that only queries about the top-k preferred alternatives of agents (for constant k) must necessarily have Squared Distortion that scales linearly with the number of agents. On the positive side, we design a simple mechanism we call Random Referee that incorporates a single comparison query, and just three queries in total, to obtain constant Squared Distortion, which also implies that the expected social cost is within a constant factor of the social optimum. We also show that Random Referee has better expected social cost in the worst case than any mechanism that only considers top-k preferred alternatives when the space of alternatives are embedded in the Euclidean plane. Finally, among mechanisms constrained to only consider the top preferences of agents, we give a mechanism called Random Oligarchy that is essentially best in class with respect to the expected social cost.

2

Fair Divisible Decisions

In this chapter, we introduce the problem of fair resource allocation for public goods, the general focus of the first part of this thesis. We begin by considering the problem of fairly allocating divisible public goods subject to a budget constraint, with the motivating application of participatory budgeting. We introduce the general problem, and discuss possible solution concepts. Ultimately, we argue that the core, a stability condition from cooperative game theory, can be interpreted as a fairness condition for the allocation of public goods by enforcing proportionality for all subsets of agents. We subsequently show how to compute core allocations by maximizing the Nash social welfare, when agents have linear and additive utilities. We describe an application in memory sharing, and show how to adapt the exponential mechanism from differential privacy to construct an approximately truthful mechanism for computing approximate core solution. Finally, we design a novel convex program for computing core allocations when agents have a more general class of utility functions that are separable and concave. This allows us to use concave utility functions to encode soft budget caps on public goods, which is important for our motivating problem of participatory budgeting.

Acknowledgements. These results are published in [63], which is joint work with Kamesh Munagala and Ashish Goel. The discussion of memory sharing in Section 2.3.1 draws on work published in [113], which is joint work with Mayuresh Kunjir, Kamesh Munagala, and Shivnath Babu.

2.1 Introduction

Classical fair resource allocation tends to focus on *private* goods. For example, one can imagine that we have a cake that can be cut into pieces and then allocated to agents, who perhaps have different utilities for different portions of the cake. In classical economic theory, a fair allocation is one that is Pareto-efficient and envy-free [157], meaning that there is no other way to divide up the cake that all agents prefer, and that no agent prefers the portion of cake allocated to any other agent.

In contrast, we focus on *public* goods. For example, one can think of shared resources like roads or parks. Economically, public goods can be characterized as having positive externalities. We will tend to view public goods as those that can grant utility to more than one agent simultaneously. More formally, suppose we have a set N of n agents and a set M of m public goods. We denote an allocation, or a solution, as a vector of values $\mathbf{x}$, where $x_j \in \mathbb{R}_{\geqslant 0}$ is the fraction of good $j \in M$ that is allocated. Let $\Delta(\mathbf{x})$ be the set of feasible allocations. We assume that we are constrained by some budget $B > 0$, and an allocation $\mathbf{x}$ is feasible if $\sum_{j \in M} x_j \leqslant B$. For every agent $i \in N$, there is a continuous and non-decreasing utility function $u_i : \Delta(\mathbf{x}) \to \mathbb{R}_{\geqslant 0}$.

This allows us to model our motivating application of Participatory Budgeting (PB) [38, 1]: a process by which a municipal organization (eg. a city or a district) puts a small amount of its budget to direct vote by its residents. PB is growing in popularity, with over 30 such elections conducted in 2015. Implementing participatory budgeting requires careful consideration of how to aggregate the preferences of community members into an actionable project funding plan. It is important to note that in participatory democracy, equitable and fair outcomes are an important systemic goal. Also, we are not restricted to modeling participatory budgeting in particular, and it should be noted that any divisible budgeting problem, such as that faced by federal, state, and local governments every year, can be modeled as above.

Unfortunately, the classical notion of fairness as a Pareto-efficient and envy-free allocation proves inadequate for public goods. In particular, the notion of envy-freeness relies on different agents receiving different allocations. When allocating public goods, the allocation is shared among all agents, and envy-freeness is vacuously true of *any* allocation. Other common fairness notions like *Max-min fair share*, *Min-max fair share*, and *competitive equilibrium from equal incomes* are also predicated on a shared assumption that the resources to be allocated are private. Therefore, our first contribution is to develop a new fair solution concept that provides a strong guarantee for allocating public resources.

2.2 Fairness as Core

One classical notion of fairness that still applies in the public goods setting is *proportionality*, also known as *sharing incentive*.[1] This concept derives from a straightforward baseline. Consider the following naive mechanism: allocate B/n budget to every agent, and then allow that agent to choose the allocation of those B/n dollars that maximizes her utility. Call this the partition algorithm.

It is not hard to see that the partition algorithm can be extremely inefficient. It is possible that the agents each have one public good from which they derive the

[1] Some might draw a distinction between these related concepts as follows: sharing incentive concerns the utility an agent would get by maximizing her utility with B/n budget, whereas proportionality concerns a $1/n$ fraction of the utility an agent would get by maximizing her utility with B budget. For linear additive utilities and divisible goods, the concepts coincide, but they need not in general. We mean the former here, and will specify in later chapters when dealing with the alternate notion. In general, we think of the former as *proportionality on endowments* and the latter as *proportionality on utilities*.

most utility but from which no other agents derive utility, while there is a common public good from which all agents derive a moderate amount of utility. Funding the shared good can lead to $\Omega(n)$ greater utility for all agents.[2] Therefore, the partition algorithm is not a viable solution to the problem of allocating public goods, but it does provide something of a baseline. Proportionality guarantees every agent at least as much utility as they would get under the partition algorithm, regardless of what the other $n-1$ agents prefer.

For private goods, proportionality provides a reasonably strong guarantee, although it relaxes the notion of envy-freeness. For public goods however, the guarantee is far too weak. The reason is that private goods cannot be shared, whereas public goods are shared by definition. Imagine the perspective of an agent in the partition algorithm: the choices of the other $n-1$ agents are irrelevant, as they will simply use their budget to allocate goods to themselves. In contrast, for public goods, the remaining $n-1$ agents may have shared interest in certain outcomes, and agents may feel justified in complaining that they have been treated unfairly as a group, even if each individual has received a proportional share of utility.

To see this in a dramatic example, suppose that there are one million agents in a given city considering two public goods $M = \{1, 2\}$ with a budget of 1 million dollars. The first is a public highway of use to the whole city, whereas the second is a mansion for the mayor. The utility functions might reasonably look like $u_i(\mathbf{x}) = x_1$ for all $i \in N$ who are not the mayor, and $u_i(\mathbf{x}) = x_2$ for the mayor. Proportionality guarantees that every agent should get at least 1 utility. So one possible proportional allocation is to spend 1 dollar on the highway, and 1 million minus 1 dollars on the mansion.

It is hopefully intuitive that this is unfair, but it is interesting to note *why* it

[2] Nevertheless, the partition algorithm is commonly implemented in many computer systems, for example it is commonly how memory is shared, despite the fact that memory is a public good.

appears unfair. In particular, the $n-1$ agents might rightfully complain that while no single one of them necessarily deserves that all of the money should be spent on the highway, their *collective* common interest in the shared highway should guarantee that it receives more attention. But proportionality as an individual fairness guarantee from private resource allocation cannot accommodate this reasoning. We resolve this tension by generalizing proportionality to all coalitions of agents. In particular, the concept of fairness with which we work is the core. This notion is borrowed from cooperative game theory and was first phrased in game theoretic terms in [150]. It has been studied subsequently in public goods settings [71, 126], but not from a computational perspective.

Definition 1. *An allocation* $\mathbf{x}$ *is a core solution if there is no subset* S *of agents who, given a budget of* $(|S|/n)B$*, could compute an allocation* $\mathbf{y}$ *where every user in* S *receives strictly more utility in* $\mathbf{y}$ *than* $\mathbf{x}$*, i.e.,* $\forall i \in S,\ u_i(\mathbf{y}) > u_i(\mathbf{x})$*.*

Note that when $S = \{1, 2, \ldots, n\}$, the above constraints encode (weak) Pareto-Efficiency. Further, when S is a singleton voter, the core captures proportionality. In general, the core captures a kind of group proportionality: No community of users should be able to come up with a justifiable complaint of the sort "if we had our proportional share of the budget, we could come up with an allocation we all prefer." However, the core is a guarantee on an exponential number of subsets of agents, and there is no obvious brute force algorithm for computing core solutions. Moreover, it is not even immediately clear whether core solutions necessarily exist. In the next section, we will prove existence constructively, by showing that core solutions can be computed by the celebrated proportional fairness program, which is a transformation of the maximum Nash social welfare program.

2.3 Linear Additive Utilities

Our computational results vary based on the utility model for the agents. Recall that for every agent $i \in N$, there is a continuous and non-decreasing utility function $u_i : \Delta(\mathbf{x}) \to \mathbb{R}_{\geqslant 0}$. In the simplest case, these functions may be linear in each public good, and additive across public goods. More formally, for each agent $i \in N$ and public good $j \in M$, there is a scalar $u_{ij} \geqslant 0$ such that $u_i(\mathbf{x}) = \sum_{j \in M} u_{ij} x_j$. In this section, we study core allocations with linear and additive utility. Our first result is that core allocations can be efficiently computed by the celebrated Proportional Fairness convex program, which is equivalent to maximizing the Nash social welfare.

Recall that the Nash social welfare (NSW) is defined as $NSW(\mathbf{u}) = \left(\prod_{i \in N} u_i\right)^{1/n}$. We will show that maximizing the Nash social welfare is sufficient for computing a core allocation (it is not necessary). Moreover, maximizing the Nash social welfare can be written as a convex program. In particular, taking the logarithm of the Nash social welfare yields the proportional fairness objective, and maximizing either objective leads to the same optimum. We define the proportional fairness convex program formally.

$$\text{Maximize} \sum_{i \in N} \log(u_i(\mathbf{x})) \text{ subject to: } \sum_{j \in M} x_j \leqslant B \tag{2.1}$$

The program is concave and can therefore be efficiently maximized. We give a simple proof using the KKT (Karush-Kuhn-Tucker) conditions [111] that the optimum is a core solution.

Theorem 1. *The solution to the Program 2.1 is a core solution.*

Proof. Let $\mathbf{x}$ denote the optimal solution. Let d denote the dual variable for the constraint $\sum_{j \in M} x_j \leqslant B$. By the KKT conditions, we have:

$$x_j > 0 \implies \sum_i \frac{u_{ij}}{u_i(\mathbf{x})} = d$$

$$x_j = 0 \implies \sum_i \frac{u_{ij}}{u_i(\mathbf{x})} \leqslant d$$

Multiplying the first set of identities by x_j and summing them, we have

$$d \cdot B = d(\sum_{j \in M} x_j) = \sum_i \frac{\sum_{j \in M} x_j u_{ij}}{u_i(\mathbf{x})} = n$$

This fixes the value of d to be n/B. Next, consider a subset S of users, along with some allocation $\mathbf{y}$ with $\sum_{j \in M} y_j = \frac{K}{n}B$. First note that the KKT conditions implied:

$$\sum_{i \in N} \frac{u_{ij}}{u_i(\mathbf{x})} \leqslant d \quad \forall j \in M$$

Multiplying by y_j and summing, we have:

$$\sum_{i \in N} \frac{u_i(\mathbf{y})}{u_i(\mathbf{x})} \leqslant d \sum_{j \in M} y_j = |S|$$

Since utilities are nonnegative

$$\sum_{i \in S} \frac{u_i(\mathbf{y})}{u_i(\mathbf{x})} \leqslant |S|$$

Therefore, if $u_i(\mathbf{y}) > u_i(\mathbf{x})$ for some $i \in S$, then there exists $j \in S$ for which $u_j(\mathbf{y}) < u_j(\mathbf{x})$. This shows that no subset S can deviate to improve their utility, so the solution $\mathbf{x}$ is a core solution. □

2.3.1 Application in Memory Sharing

In this section, we briefly consider an application of the core solution concept and it's computation for divisible goods in memory sharing. We have a paper that treats this topic in much greater deatil [113]. As the work in that paper will appear in the thesis of Mayuresh Kunjir (the lead student author), we only mention the model and results briefly here.

In a cluster copmuting environment, it is possible that many users are concurrently running tasks on the same physical machine. The processor on that machine is a private good, in the sense that only one of the users can use it at a time. The memory, or cache, however, is a public good in so far as multiple users can simultaneously benefit from access to the same file stored in memory. We call this a multi-tenant scenario.

Computation on data stored in the cache can be orders of magnitude faster than on the same data stored on disk, but there is generally a limited amount of cache available, so there is a strong incentive to allocate the available cache efficiently. However, simply optimizing for throughput irrespective of the multi-tenant setting can result in *starving* some users, which is undesirable from a fairness perspective, and impractical in terms of run time performance. Instead, many multi-tenant setups involve partitioning the cache equally among the users (*i.e.*, if there are four users and four GB of cache, each user gets exactly one GB of cache to do with as she pleases.)

This naive solution by partitioning is inefficient in at least two senses. First, if multiple users want to store the same files, it requires that the file be stored multiple times. Second, it may be impossible to store a large file under such a scheme, even if there is sufficient space. In the previous example, suppose that two users want to work with the same two GB file: this is impossible under the partitioning scheme, despite the fact that the two users together seem "entitled" to the two GB of space. In both senses, the naive solution is ignoring the possibility of sharing.

The core is a fair solution concept that is precisely intended to generalize the intuition of fair memory allocation via proportional partitioning to account for the possibility of sharing. In [113], we design algorithms based on the proportional fairness objective to compute core allocations in the multi-tenant memory allocation context. Note that memory files themselves are *not* divisible goods. However, we argue that randomization is acceptable in the memory allocation context due to the long time horizon of most jobs with respect to the batching interval at which we potentially re-allocate memory. We therefore model the allocation problem as that of finding probability distributions over integral allocations, where the utility of a user is linear and additive with respect to this distribution.

There is some difficulty in that the number of variables needed to represent such a distribution is potentially exponential, which could lead to an intractable optimization problem. Instead, we design simple first order heuristics for optimization based on a polynomial size sample of possible allocations of files into the cache. We implement these algorithms in Spark to allocate memory in batches across time, and show empirically that we are able to provide fairness guarantees while simultaneously competing on throughput.

2.3.2 Mechanism Design

In this section we consider developing a truthful mechanism for computing core allocations without payments under linear and additive utilities. Truthfulness has long been considered a serious problem for the allocation of public goods [86, 126]. We study *asymptotic approximate truthfulness* [120]. Our notion is asymptotic in the sense that $n \gg k$, which is reasonable in practice. We note at the outset that computing the core of public goods with no private goods (*e.g.*, money transfers) satisfies the property of *strategy-proof in the large* [14], meaning that if agents know a distribution from which the preferences of other agents are drawn, then the market is truthful in expectation in the limit as the number of agents grows large.

To go further, we show that when agents' utilities are linear (and more generally, homogeneous of degree 1), there is an efficient randomized mechanism that implements an ϵ-approximate core solution as a dominant strategy for large n. We use the Exponential Mechanism [123] from differential privacy to achieve this. The application of the Exponential Mechanism is not straightforward since the proportional fairness objective is not separable when used as a scoring function; the allocation variables are common to all agents. Furthermore, this objective varies widely when one agent misreports utility. We define a scoring function directly based on the *gradient condition* of proportional fairness to circumvent this hurdle.

We develop a randomized mechanism that finds an approximately core solution with high probability while ensuring approximate dominant-strategy truthfulness for all agents. In the spirit of [120], we assume the large market limit so that $n \gg k$; in particular, we assume $k = o(\sqrt{n})$. We present the mechanism for linear utility functions where $U_i(\mathbf{x}) = \sum_{j=1}^{k} u_{ij} x_j$, noting that it easily generalizes to degree one homogeneous functions. The values of u_{ij} are reported by the agents. Without loss of generality, these are normalized so that $\|\mathbf{u_i}\|_1 = 1$. Also without loss of generality, let B be normalized to 1. Recall that for linear utility functions, the proportional fairness algorithm that maximizes $\sum_i \log U_i(\mathbf{x})$ subject to $\|\mathbf{x}\|_1 \leqslant 1$ and $\mathbf{x} \geqslant 0$ computes the Lindahl equilibrium.

We will design additive approximations to the core (see Definition 2) that achieve approximate truthfulness in an additive sense. We use the Exponential Mechanism [123] to achieve approximate truthfulness.

Definition 2. *For $\alpha > 1$, an allocation $\mathbf{x}$ lies in the α-approximate multiplicative (resp. additive) core if for any subset S of agents, there is no allocation $\mathbf{y}$ using a budget of $(|S|/n)B$, s.t. $U_i(\mathbf{y}) > \alpha U_i(\mathbf{x})$ (resp. $U_i(\mathbf{y}) > U_i(\mathbf{x}) + \alpha$) for all $i \in S$.*

Fix a constant $\gamma \in (0, 1)$ to be chosen later. We first define the convex set of feasible allocations as $\mathcal{P} := \{\mathbf{x} : \mathbf{x} \geqslant n^{-\gamma}, \|\mathbf{x}\|_1 \leqslant 1\}$. Note that all such allocations are restricted to allocating at least $n^{-\gamma}$ to each project. Since the utility vector of any agent is normalized so $\|\mathbf{u_i}\|_1 = 1$, this implies that every agent gets a baseline utility of at least $n^{-\gamma}$, a fact we use frequently. We define the following scoring function, which is based on the gradient optimality condition of Proportional Fairness:

$$q(\mathbf{x}) := n - n^{-\gamma} \max_{\mathbf{y} \in \mathcal{P}} \left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})} \right)$$

We will approximately and truthfully maximize this scoring function. The trade off in defining the scoring function is between reducing the sensitivity of the function to the report of an individual agent and thus improving the approximation to truthfulness, and having just enough sensitivity so that the mechanism defined in terms of the scoring function provides a good approximation to the core. This scoring function, derived from the gradient condition of the proportional fairness program, provides this balance. Using this scoring function, we can define our formal mechanism, and show that it can be sampled efficiently.

Definition 3. *Define μ to be a uniform probability distribution over all feasible allocations $\mathbf{x} \in \mathcal{P}$. For a given set of utilities, let the mechanism ζ_q^ϵ be given by the rule:*

$$\zeta_q^\epsilon := \text{ choose } \mathbf{x} \text{ with probability proportional to } e^{\epsilon q(\mathbf{x})} \mu(\mathbf{x})$$

First we argue that we can efficiently sample from this distribution.

Lemma 1. *ζ_q^ϵ can be sampled in polynomial time with small additive error in truthfulness.*

Proof. The sampling procedure is based on Markov Chain Monte Carlo techniques, and thus there is a small additive error introduced by the total variation distance between the approximate and target distribution. To show that sampling from ζ_q^ϵ can be done in polynomial time, we show that it is sampling a log-concave function over a convex set [121]. Clearly, the allocation space is convex. Therefore, it is sufficient to show that $e^{\epsilon q(\mathbf{x})}\mu(\mathbf{x})$ is log-concave. Note that $\ln\left(e^{\epsilon q(\mathbf{x})}\mu(\mathbf{x})\right) = \epsilon q(\mathbf{x}) + \ln\left(\mu(\mathbf{x})\right)$ Since μ is only uniform, this is just an affine transformation of $q(\mathbf{x})$, therefore we need to show that $q(\mathbf{x})$ is concave.

Recall the definition $q(\mathbf{x}) := n - n^{-\gamma} \max_{\mathbf{y}\in\mathcal{P}} \left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right)$. Each individual utility function U_i is concave, since it is a linear function $(\mathbf{u_i} \cdot \mathbf{x})$. Thus, $U(\mathbf{x})^{-1}$ is convex because it is the composition of the convex and non-increasing scalar function $1/x$ with the concave multivariate (but scalar valued) $U_i(x)$. The sum is still convex, as a linear combination of convex functions. Finally, the max operator preserves convexity, so that $q(\mathbf{x})$ is concave in $\mathbf{x}$. □

Finally, the main result of this section demonstrates that ζ_q^ϵ can find an approximate core solution while providing approximate truthfulness. The proof relies on technical details from [123] and is in multiple parts; we present the proof in Appendix 8.1.

Theorem 2. *ζ_q^ϵ is $(e^{2\epsilon} - 1)$-approximately truthful. Furthermore, if k is $o(\sqrt{n})$ and $\frac{1}{\epsilon} > \frac{kn}{(n-k^2)\ln n}$ then ζ_q^ϵ can be used to choose an allocation $\mathbf{x}$ that is an $O\left(\frac{k \ln n}{\epsilon\sqrt{n}}\right)$-approximate additive core solution w.p. $1 - \frac{1}{n}$.*

2.4 Concave Utilities

In this section, we explore computing core allocations for more general concave utility functions, which introduces substantial complications algorithmic complications. These challenges will force us to develop an algorithmic understanding of the Lindahl Equilibrium; a market based notion we will define shortly. Before proceeding however, we define the more general utility functions we consider more precisely.

Utility Functions We consider utility functions that generalize linear utility functions. These utility functions, which we term Scalar Separable, have the form

$$U_i(\mathbf{x}) = \sum_j u_{ij} f_j(x_j)$$

for every agent i where $\{f_j\}$ are smooth, non-decreasing, and concave, and $\mathbf{u_i} \geqslant 0$. By $\sum_j$ we always mean the sum over the k projects. Scalar Separable utilities are fairly general and well-motivated. First, this concept encompasses linear utilities and several other canonical utility functions (see below). Secondly, if voters express scalar valued preferences (such as up/down approval voting), Scalar Separable utilities provide a natural way of converting these votes into cardinal utility functions. In fact, as we discuss below, we will do precisely this when handling real data. We consider two subclasses that we term Non-satiating and Saturating utilities respectively. Each arises naturally in settings related to participatory budgeting.

Non-satiating Functions: For our main computational result, we consider a subclass of utility functions that we term Non-satiating.

Definition 4. *A differentiable, strictly increasing, concave function f is called Non-Satiating if $xf_j'(x)$ is monotonically increasing and equal to 0 when $x = 0$.*

This is effectively a condition that the functions grow at least as fast as $\ln x$. Several utility functions used for modeling substitutes and complements fall in this class. For instance, constant elasticity of substitution (CES) utility functions where

$$U_i(\mathbf{x}) = \left(\sum_j u_{ij} x_j^{\rho}\right)^{\frac{1}{\rho}} \quad \text{for } \rho \in (0, 1]$$

can be monotonically transformed into Non-satiating utilities.[3] CES functions are also *homogeneous of degree 1*, meaning that $U_i(\alpha\mathbf{x}) = \alpha U_i(\mathbf{x})$ for any scalar $\alpha \geqslant 0$.

[3] Note that the core remains unchanged if utilities undergo a monotone transform.

When $\rho = 1$, this captures linear utilities. When $\rho \to 0$, these are Cobb-Douglas utilities which for $\alpha_{ij} > 0$ such that $\sum_j \alpha_{ij} = 1$, can be written as $U_i(\mathbf{x}) = \prod_j x_j^{\alpha_{ij}}$.

Saturating Functions: Note that Scalar Separable utilities assume goods are divisible. Fractional allocations make sense in their own right in several scenarios: Budget allocations between goals such as defense and education at a state or national level are typically fractional, and so are allocations to improve utilities such as libraries, parks, gyms, roads, etc. However, in the settings for which we have real data, the projects are indivisible and have a monetary cost s_j, so that we have the additional constraint $x_j \in \{0, s_j\}$ on the allocations. We therefore need utility functions that model budgets in individual projects. These utility functions must also be simple to account for the limited information elicited in practice. For example, in the voting data that we use in our experiments, each voter receives an upper bound on how many projects she can select, and the ballot cast by a voter is simply the subset of projects she selects. A related voting scheme implemented in practice, called Knapsack Voting [81], has similar elicitation properties. For modeling these two considerations, we consider Saturating utilities.

Definition 5. *Saturating utility functions $U_i(\mathbf{x})$ have the form $\sum_j u_{ij} \min\left(\frac{x_j}{s_j}, 1\right)$.*

For converting our voting data into a Saturating utility, we set s_j to be the budget of project j, and set u_{ij} to 1 if agent i votes for project j and 0 otherwise. Note that if $x_j = s_j$, then the utility of any agent who voted for this item is 1. This implies the total utility of an agent is the number (or total fraction) of items that he voted for that are present in the final allocation. Clearly, Saturating utilities do not satisfy Definition 4. However, in a full paper we connect Non-satiating and Saturating utilities by developing an approximation algorithm and heuristic for computing core allocations in the saturating model using results developed for the Non-satiating model [63].

2.4.1 Computing the Lindahl Equilibrium

In a fairly general public goods setting, there is a market based notion of fairness due to Lindahl [118] and Samuelson [148] termed the *Lindahl equilibrium*, which is based on setting different prices for the public goods for different agents. The market on which the Lindahl equilibrium is defined is a mixed market of public and private goods. We present a definition below that is specialized to just a public goods market relevant for participatory budgeting.

Definition 6. *In a public goods market with budget B, per-voter prices $\mathbf{p_1}, \mathbf{p_2}, ..., \mathbf{p_n}$ each in $\mathbb{R}_+^k$ and allocation $\mathbf{x} \in \mathbb{R}_+^k$ constitute a Lindahl equilibrium if the following two conditions hold:*

1. *For every agent i, the utility $U_i(\mathbf{y}_i)$ is maximized subject to $\mathbf{p_i} \cdot \mathbf{y}_i \leqslant B/n$ when $\mathbf{y}_i = \mathbf{x}$; and*
2. *The profit defined as $(\sum_i \mathbf{p_i}) \cdot \mathbf{z} - \|\mathbf{z}\|_1$, subject to $\mathbf{z} \geqslant 0$ is maximized when $\mathbf{z} = \mathbf{x}$.*

The price vector for every agent is traditionally interpreted as a tax. However, unlike in private goods markets, in our case these prices (or taxes) are purely hypothetical; we are only interested in the allocation that results at equilibrium (in fact, we eliminate the prices from our characterization of the equilibrium). Under innocuous conditions for the mixed public and private goods market, Foley proved that the Lindahl equilibrium exists and lies in the core [71]. This remains true in our specialized instance of the problem; the omitted proof is a trivial adaption from [71]. Thus, computing a Lindahl equilibrium is sufficient for the purpose of computing a core allocation. However, Foley only proves existence of the equilibrium via a fixed point argument that does not lend itself to efficient computation.

Our first result is a characterization of the Lindahl equilibrium that uses the optimality conditions to eliminate the price variables entirely.

Theorem 3. *An allocation* $\mathbf{x} \geqslant 0$ *corresponds to a Lindahl equilibrium if and only if*

$$\sum_i \left(\frac{u_{ij} f_j'(x_j)}{\sum_m u_{im} x_m f_m'(x_m)} \right) \leqslant \frac{n}{B} \tag{2.2}$$

for all items j*, where this inequality is tight when* $x_j > 0$*.*

Proof. We prove the statement for more general utility functions, $\{U_i(\mathbf{x})\}$. Recall the definition of Lindahl equilibrium from Definition 6: Condition (1) implies that there is a dual variable λ_i for every agent such that

$$\forall j \qquad \frac{\partial}{\partial x_j} U_i(\mathbf{x}) \leqslant \lambda_i p_{ij} \tag{2.3}$$

with the inequality being tight if $x_j > 0$. Condition (1) in Definition 6 also implies that $\sum_j p_{ij} x_j = B/n$ for all i. Multiplying Equation (2.3) by x_j, noting that the inequality is tight when $x_j > 0$, and summing,

$$\forall i, \quad \sum_j x_j \frac{\partial}{\partial x_j} U_i(\mathbf{x}) = \sum_j \lambda_i p_{ij} x_j = \lambda_i (B/n)$$

Rearranging

$$\forall i, \quad \lambda_i = \frac{n}{B} \sum_j x_j \frac{\partial}{\partial x_j} U_i(\mathbf{x}) \tag{2.4}$$

Similarly, Condition (2) in Definition 6 implies that $\forall x_j, \sum_i p_{ij} \leqslant 1$ where the inequality is tight when $x_j > 0$. Substituting into Inequality (2.3) and summing, we have:

$$\forall x_j, \quad \sum_i \frac{\frac{\partial}{\partial x_j} U_i(\mathbf{x})}{\lambda_i} \leqslant 1 \tag{2.5}$$

with the inequality being tight when $x_j > 0$. Using Equation (2.4) to eliminate λ_i in Inequality (2.5), we finally obtain

$$\forall x_j, \quad \sum_i \frac{\frac{\partial}{\partial x_j} U_i(\mathbf{x})}{\sum_m x_m \frac{\partial}{\partial x_m} U_i(\mathbf{x})} \leqslant \frac{n}{B}$$

with the inequality being tight when $x_j > 0$. Taking the appropriate partial derivatives in the scalar separable utility model yields the theorem statement. □

We now present our main computational result that builds on the characterization above to give the first non-trivial polynomial time method for computing the Lindahl equilibrium. We need the non-satiation assumption on the functions $\{f_j\}$ given in Definition 4.

Theorem 4. *When $U_i(\mathbf{x}) = \sum_j u_{ij} f_j(x_j)$ where $\{f_j\}$ satisfy Definition 4, the Lindahl equilibrium (and therefore a core solution) is the solution to a convex program.*

Proof. Recall the characterization of the Lindahl equilibrium from Theorem 3. Define $z_j = x_j f_j'(x_j)$. Note that $x_j = 0$ iff $z_j = 0$. Since f_j satisfies non-satiation, this function is continuous and monotonically increasing, and hence invertible. Let h_j be this inverse such that $h_j(z_j) = x_j$. Let $r_j(z_j) = h_j(z_j)/z_j = 1/f_j'(x_j)$. The Lindahl equilibrium characterization therefore simplifies to:

$$\sum_i \left(\frac{u_{ij}}{\sum_m u_{im} z_m} \right) \leqslant \frac{n}{B} r_j(z_j)$$

with the inequality being tight when $z_j > 0$. Let $R_j(z_j)$ be the indefinite integral of r_j (with respect to z_j). Define the following potential function

$$\Phi(\mathbf{z}) = \sum_i \log \left(\sum_j u_{ij} z_j \right) - \left(\frac{n}{B} \right) \sum_j R_j(z_j) \tag{2.6}$$

We claim that $\Phi(\mathbf{z})$ is concave in $\mathbf{z}$. The first term in the summation is trivially concave. Also, since $f_j'(x_j)$ is a decreasing function, $1/f_j'(x_j)$ is increasing in x_j. Since $r_j(z_j) = 1/f_j'(x_j)$, this is increasing in x_j and hence in z_j. This implies $R_j(z_j)$ is convex, showing the second term in the summation is concave as well. It is easy to check that the optimality conditions of maximizing $\Phi(\mathbf{z})$ subject to $\mathbf{z} \geqslant 0$ are exactly the conditions for the Lindahl equilibrium. This shows that the Lindahl equilibrium corresponds to the solution to the convex program maximizing $\Phi(\mathbf{z})$. □

It is not difficult to argue that an approximately optimal solution to the convex program gives an approximate core solution, which implies polynomial time computation to arbitrary accuracy. We note that the non-satiation condition essentially implies that $f_j(x_j)$ should grow faster than $\ln x_j$. In combination with the assumption that $f_j(x_j)$ is concave, this leaves us with a broad class of concave functions for which the Lindahl equilibrium and hence the core can be efficiently computed. Furthermore, it is easy to see that for linear utility functions, maximizing the potential in equation 2.6 reduces to the proportional fairness program.

2.5 Conclusion and Open Directions

In this chapter, we addressed the problem of allocating divisible public resources according to budget constraints. We adapted a stability notion from game theory called the core as a novel solution concept for group fairness in the allocation of public goods. We demonstrated that for linear utilities, the celebrated proportional fairness program computes a core solution. Furthermore, for linear utilities, we were able to adapt techniques from differential privacy to give an approximately truthful mechanism for computing an approximate core solution. For more general utilities, we turned to the Lindahl equilibrium for existence, and we initiated the computational study of the Lindahl equilibrium. We expressed the equilibrium (and hence the core) purely in terms of common allocation variables. In a sense, this is a mirror image of the role common prices play in private good markets. Having common allocations instead of prices poses some challenges in computation and mechanism design, but has the advantage that we can compute equilibria for a wider class of utility functions.

Our work is just the first step towards understanding these equilibria. We do not yet understand the computational complexity for more general utility functions. Are they as hard as general concave games? Our truthfulness result uses a scoring function based on the gradient; it would be nice to generalize this result to other utility functions, as well as design simpler, more practical schemes.

3

Fair Indivisible Decisions

In this chapter, we continue examining the core as a fair solution concept for public goods. However, we no longer assume that we can allocate fractional goods. Instead, we assume that we must integrally choose a set of goods to allocate. Our computational results vary based on the constraints imposed on such a solution set. Our algorithms are quite distinct from those for divisible goods. We provide constant additive approximation algorithms for matroid and matching constraints that operate by local search, and then give a rounding algorithm for packing constraints. Matroid constraints allow us to model voting, in particular, multi-winner elections and multiple referendums. Packing constraints allow us to model participatory budgeting when projects must be completely funded or completely rejected.

Acknowledgements. These results are published in [66], which is joint work with Kamesh Munagala and Nisarg Shah. Conversations with Rupert Freeman, Vincent Connitzer, and Paul Gölz were also very helpful.

3.1 Introduction

Consider an example to highlight what an indivisible public decision problem might look like, and why fairness might be a concern. Suppose that the next time you vote, you see that there are four referendums for your consideration on the ballot, all of which concern the allocation of various public goods in your city: A = a new school, B = enlarging the public library, C = renovating the community college, and D = improving a museum. In 2016, residents of Durham, North Carolina faced precisely these options [122]. Suppose the government has resources to fund only two of the four projects, and the (hypothetical) results were as follows: a little more than half of the population voted for (A, B), a little less than half voted for (C, D), and every other combination received a small number of votes. Which projects should be funded?

If we naïvely tally the votes, we would fund A and B, and ignore the preferences of a very large minority. In contrast, funding A and C seems like a reasonable compromise. Of course, it is impossible to satisfy *all* voters, but given a wide enough range of possible outcomes, perhaps we can find one that fairly reflects the preferences of large subsets of the population. This idea is not captured by fairness axioms like proportionality or their approximations [53], which view fairness from the perspectives of *individual* agents. Indeed, in the aforementioned example, *every* allocation gives zero utility to *some* agent, and would be deemed equally good according to such fairness criteria.

In this chapter, we consider a fairly broad model for public goods allocation that generalizes much of the prior work in the literature [53, 63, 16, 70, 41, 136]. There is a set of voters (or agents) $N = [n]$. Public goods are modeled as elements of a ground set W. We denote $m = |W|$. An *outcome* $\mathbf{c}$ is a subset of W. Let $\mathcal{F} \subseteq 2^W$ denote the set of feasible outcomes.

The utility of agent i for element $j \in W$ is denoted $u_{ij} \in \mathbb{R}_{\geqslant 0}$. We assume that agents have additive utilities, i.e., the utility of agent i under outcome $\mathbf{c} \in \mathcal{F}$ is $u_i(\mathbf{c}) = \sum_{j \in \mathbf{c}} u_{ij}$. Since we are interested in scale-invariant guarantees, we assume without loss of generality that $\max_{j \in W} u_{ij} = 1$ for each agent i, so that $u_{ij} \in [0, 1]$ for all i, j. Crucially, this does not restrict the utility of an agent for an outcome to be 1: $u_i(\mathbf{c})$ can be as large as m. Specifically, let $V_i = \max_{\mathbf{c} \in \mathcal{F}} u_i(\mathbf{c})$, and $V_{\max} = \max_{i \in N} V_i$. Our results differ by the feasibility constraints imposed on the outcome. We consider three types of constraints, special cases of which have been studied previously in literature.

Matroid Constraints. In this setting, we are given a matroid $\mathcal{M}$ over the ground set W, and the feasibility constraint is that the chosen elements must form a basis of $\mathcal{M}$ (see [112] for a formal introduction to matroids).

This generalizes the *public decision making* setting introduced by [53]. In this setting, there is a set of issues T, and each issue $t \in T$ has an associated set of alternatives $A^t = \{a_1^t, \ldots, a_{k_t}^t\}$, exactly one of which must be chosen. Agent i has utility $u_i^t(a_j^t)$ if alternative a_j^t is chosen for issue t, and utilities are additive across issues. An outcome $\mathbf{c}$ chooses one alternative for every issue. It is easy to see that if the ground set is $\cup_t A^t$, the feasibility constraints correspond to a partition matroid. We note that public decision making in turn generalizes the classical setting of *private goods allocation* [119, 41, 136] in which private goods must be divided among agents with additive utilities, with each good allocated to exactly one agent.

Matroid constraints also capture multi-winner elections in the voting literature (see, e.g. [16]), in which voters have additive utilities over candidates, and a committee of at most k candidates must be chosen. This is captured by a uniform matroid over the set of candidates.

Matching Constraints. In this setting, the elements are edges of an undirected graph $G(V, E)$, and the feasibility constraint is that the subset of edges chosen must form a *matching*. Matchings constraints in a bipartite graph can be seen as the intersection of two matroid constraints. Matching constraints are a special case of the more general packing constraints we consider below.

Packing Constraints. In this setting, we impose a set of packing constraints $A\mathbf{x} \leqslant \mathbf{b}$, where $x_j \in \{0, 1\}$ is the indicator denoting whether element j is chosen in the outcome. Suppose A is a $K \times m$ matrix, so that there are K packing constraints. By scaling, we can assume $a_{kj} \in [0, 1]$ for all k, j. Note that even for one agent, packing constraints encode independent set. Thus, to make the problem tractable, we assume $\mathbf{b}$ is sufficiently large, in particular, $b_k = \omega(\log K)$ for all $k \in \{1, 2, \ldots, K\}$. This is in contrast to matroid and matching constraints, for which single-agent problems are polynomial time solvable. A classic measure of how easy it is to satisfy the packing constraints is the *width* ρ [138]:

$$\rho = \max_{k \in [K]} \frac{\sum_{j \in [m]} a_{kj}}{b_k}. \tag{3.1}$$

Packing constraints capture the general KNAPSACK setting, in which there is a set of m items, each item j has an associated size s_j, and a set of items of total size at most B must be selected. This setting is motivated by participatory budgeting applications [1, 81, 80, 63, 75, 70, 25], in which the items are public projects, and the sizes represents the costs of the projects. KNAPSACK uses a single packing constraint. Multiple packing constraints can arise if the projects consume several types of resources, and there is a budget constraint for each resource type. For example, consider a statewide participatory budgeting scenario where each county has a budget than can only be spent on projects affecting that county, the state has some budget that can be spent in any county, and projects might affect multiple counties. In such settings, it is natural to assume a small width, i.e., that the budget for each resource is such that a large fraction (but not all) of the projects can be funded. We note that the aforementioned multi-winner election problem is a special case of the KNAPSACK problem with unit sizes.

3.1.1 Fairness Properties

As with divisible goods, we define fairness desiderata for indivisible public goods by generalizing appropriate desiderata from the private goods setting such as Pareto optimality, which is a weak notion of efficiency, and proportionality, which is a per-agent fair share guarantee. For convenience, we define these concepts explicitly for outcomes consisting of indivisible goods.

Definition 7. *An outcome* $\mathbf{c}$ *satisfies Pareto optimality if there is no other outcome* $\mathbf{c}'$ *such that* $u_i(\mathbf{c}') \geqslant u_i(\mathbf{c})$ *for all agents* $i \in N$*, and at least one inequality is strict.*

Recall that V_i is the maximum possible utility agent i can derive from a feasible outcome.

Definition 8. *The proportional share of an agent* $i \in N$ *is* $Prop_i := \frac{V_i}{n}$*. For* $\beta \in (0, 1]$*, we say that an outcome* $\mathbf{c}$ *satisfies* β*-proportionality if* $u_i(\mathbf{c}) \geqslant \beta \cdot Prop_i$ *for all agents* $i \in N$*. If* $\beta = 1$*, we simply say that* $\mathbf{c}$ *satisfies proportionality.*

The difficulty in this setting stems from requiring integral outcomes, and not allowing randomization. In the absence of randomization, it is reasonably straightforward to show that we cannot guarantee β-proportionality for any $\beta \in (0, 1]$. Consider a problem instance with two agents and two feasible outcomes, where each outcome gives a positive utility to a unique agent. In any feasible outcome, one agent has zero utility, which violates β-proportionality for every $\beta > 0$.

To address this issue, [53] introduced the novel relaxation of proportionality up to one issue in their public decision making framework, inspired by a similar relaxation called envy-freeness up to one good in the private goods setting [119, 41]. They say that an outcome $\mathbf{c}$ of a public decision making problem satisfies *proportionality up to one issue* if for all agents $i \in N$, there exists an outcome $\mathbf{c}'$ that differs from $\mathbf{c}$ only on a single issue and $u_i(\mathbf{c}') \geqslant Prop_i$. Proportionality up to one issue is a reasonable fairness guarantees only when the number of issues is larger than the number of agents; otherwise, it is vacuous and is satisfied by all outcomes. Thus, it is perfectly reasonable for some applications (e.g., three friends choosing a movie list to watch together over the course of a year), but not for others (e.g., when thousands of residents choose a handful of public projects to finance). In fact, it may produce an outcome that may be construed as unfair if it does not reflect the wishes of large groups of voters. Thus, in this chapter we address the following question posed by [53]:

> *Is there a stronger fairness notion than proportionality in the public decision making framework...? Although such a notion would not be satisfiable by deterministic mechanisms, it may be satisfied by randomized mechanisms, or it could have novel relaxations that may be of independent interest.*

Our primary contributions are twofold. First, we define a fairness notion for public goods allocation that is stronger than proportionality, ensures fair representation of groups of agents, and in particular, provides a meaningful fairness guarantee even when there are fewer goods than agents. Second, we provide polynomial time algorithms for computing integer allocations that approximately satisfy this fairness guarantee for a variety of settings generalizing the public decision making framework and participatory budgeting. Our fairness notion is an adaptation of the core. Below, we define the notion of notion of a core outcome for indivisible goods. Note that we study the core defined on proportional shares of utility rather than proportional shares of endowments.

Definition 9. *Given an outcome* $\mathbf{c}$*, we say that a set of agents* $S \subseteq N$ *form a blocking coalition if there exists an outcome* $\mathbf{c}'$ *such that* $(|S|/n) \cdot u_i(\mathbf{c}') \geqslant u_i(\mathbf{c})$ *for all* $i \in S$ *and at least one inequality is strict. We say that an outcome* $\mathbf{c}$ *is a core outcome if it admits no blocking coalitions.*

Again, note that non-existence of blocking coalitions of size 1 is equivalent to proportionality, and non-existence of blocking coalitions of size n is equivalent to Pareto optimality. Hence, a core outcome is both proportional and Pareto optimal. However, the core satisfies a stronger property of being, in a sense, Pareto optimal for coalitions of *any size*, provided we scale utilities based on the size of the coalition.

Another way of thinking about the core is to view it as a fairness property that enforces a proportionality-like guarantee for coalitions: *e.g.*, if half of all agents have identical preferences, they should be able to get at least half of their maximum possible utility. It is important to note that the core provides a guarantee for every possible coalition. Hence, in satisfying the guarantee for a coalition S, a solution cannot simply make a single member $i \in S$ happy and ignore the rest as this would likely violate the guarantee for the coalition $S \backslash \{i\}$.

Unfortunately, since a proportional outcome is not guaranteed to exist (even allowing for multiplicative approximations), the same is true for the core. However, an additive approximation to the core still provides a meaningful guarantee, even when there are fewer elements than agents because it provides a non-trivial guarantee to large coalitions of like-minded agents.

Definition 10. *For* $\delta, \alpha \geqslant 0$*, an outcome* $\mathbf{c}$ *is a* (δ, α)*-core outcome if there exists no set of agents* $S \subseteq N$ *and outcome* $\mathbf{c}'$ *such that*

$$\frac{|S|}{n} \cdot u_i(\mathbf{c}') \geqslant (1+\delta) \cdot u_i(\mathbf{c}) + \alpha$$

for all $i \in S$, and at least one inequality is strict.

A $(0,0)$-core outcome is simply a core outcome. A $(\delta,0)$-core outcome satisfies δ-proportionality. Similarly, a $(0,1)$-core outcome $\mathbf{c}$ satisfies the following relaxation of proportionality that is slightly weaker than proportionality up to one issue: for every agent $i \in N$, $u_i(\mathbf{c}) + 1 \geqslant Prop_i$. We note that this definition, and by extension, our algorithms satisfy scale invariance, i.e., they are invariant to scaling the utilities of any individual agent. Because we normalize utilities of the agents, the true additive guarantee is α times the maximum utility an agent can derive from a single element. Since an outcome can have many elements, an approximation with small α remains meaningful.

The advantage of an approximate core outcome is that it fairly reflects the will of a like-minded subpopulation relative to its size. An outcome satisfying approximate proportionality only looks at what *individual* agents prefer, and may or may not respect the collective preferences of sub-populations. We present such an instance in Example 1 (Section 3.2.1), in effect showing that an approximate core outcome is arguably more fair.

In our results, we will assume $\delta < 1$ to be a small constant, and focus on making α as small as possible. In particular, we desire guarantees on α that exhibit sub-linear or no dependence on n, m, or any other parameters. Deriving such bounds is the main technical focus of our work.

3.1.2 Results

We present algorithms to find approximate core outcomes under matroid, matching, and general packing constraints. Our first result (Section 3.3) is the following:

Theorem 5. *If feasible outcomes are constrained to be bases of a matroid, then a $(0,2)$-core outcome is guaranteed to exist, and for any $\epsilon > 0$, a $(0, 2+\epsilon)$-core outcome can be computed in time polynomial in n, m, and $1/\epsilon$.*

In particular, for the public decision making framework, the private goods setting, and multi-winner elections (a.k.a. KNAPSACK with unit sizes), there is an outcome whose guarantee for *every coalition* is close to the guarantee that Conitzer et al. provide to individual agents [53].

In Section 3.4, we consider matching constraints. Our result now involves a tradeoff between the multiplicative and additive guarantees.

Theorem 6. *If feasible outcomes are constrained to be matchings in an undirected graph, then for constant $\delta \in (0,1]$, a $(\delta, 8 + 6/\delta)$-core outcome can be computed in time polynomial in n and m.*

Our results in Section 3.5 are for general packing constraints. Here, our guarantee depends on the width ρ from Equation (3.1), which captures the difficulty of satisfying the constraints. In particular, the guarantee improves if the constraints are easier to satisfy. This is the most technical result of this chapter, and involves different techniques than those used in proving Theorems 5 and 6; we present an outline of the techniques in Section 3.5.1.

Theorem 7. *For constant $\delta \in (0,1)$, given K packing constraints $A\mathbf{x} \leqslant \mathbf{b}$ with width ρ and $b_k = \omega\left(\frac{\log K}{\delta^2}\right)$ for all $k \in [K]$, there exists a polynomial time computable (δ, α)-core solution, where*

$$\alpha = O\left(\frac{1}{\delta^4} \cdot \log\left(\frac{\min(V_{\max}, n, \rho) \cdot \log^* V_{\max}}{\delta}\right)\right).$$

Here, $\log^*$ is the iterated logarithm, which is the number of times the logarithm function must be iteratively applied before the result becomes less than or equal to 1. Recall that $V_{\max}$ is the maximum utility an agent can have for an outcome (thus $V_{\max} \leqslant m$); our additive error bound is a vanishing fraction of this quantity. Our bound is also small if the number of agents n is small. Finally, the guarantee improves for small ρ, i.e., as the packing constraints become easier to satisfy. For instance, in participatory budgeting, if the total cost of all projects is only a constant times more than the budget, then our additive guarantee is close to a constant.

Note that $V_{\max}$ (which is bounded by m), n, and ρ are all unrelated quantities — either could be large with the other two being small. In fact, in Section 3.5, we state the bound more generally in terms of what we call the *maximally proportionally fair value* R, which informally captures the (existential) difficulty of finding a proportionally fair allocation. The quantity $\min(V_{\max}, n, \rho)$ stems from three different bounds on the value of R.

Also, note that the lower bound on $\mathbf{b}$ in the above theorem is necessary: if $\mathbf{b} = O(1)$, then no non-trivial approximation to the core can be guaranteed, even when ρ is a constant. To see this, consider a complete bipartite graph $G(L, R, E)$, where $|L| = |R| = m/2$. The vertices are the elements of the ground set W, and the constraints ensure that feasible outcomes are independent sets, which can be encoded with constraints of the form $A\mathbf{x} \leqslant \mathbf{b}$ where $\mathbf{b} = \mathbf{1}$ (and the width is $\rho = 2$). There are two agents. Agent 1 has unit utility for each vertex in L, and zero utility for each vertex in R, while agent 2 has unit utility for each vertex in R. A feasible outcome is forced to choose either vertices from L or vertices from R, and hence gives zero utility to at least one agent. But this agent can deviate and choose an outcome with utility $m/2$, which is then scaled down to $m/4$. Hence, no feasible outcome is $(\delta, m/4)$-core for any $\delta > 0$.

3.1.3 Related Work

A simpler property than the core is proportionality, which like the core, is impossible to satisfy to any multiplicative approximation using integral allocations. To address this problem, [53] defined proportionality up to one issue in the public decision making framework, inspired by related notions for private goods. This guarantee is satisfied by the integral outcome maximizing the *Nash welfare* objective, which is the *geometric* mean of the utilities to the agents. For public goods, this objective is not only NP-Hard to approximate to any multiplicative factor, but approximations to the objective also do not retain the individual fairness guarantees.

We extend the notion of additive approximate proportionality to additive approximate core outcomes, which provides meaningful guarantees even when there are fewer goods than agents. Unlike proportionality, we show in Section 3.2.1 that the approach of computing the optimal integral solution to the Nash welfare objective fails to provide a reasonable approximation to the core. Therefore, for our results about matroid constraints (Theorem 5) and matching constraints (Theorem 6), we slightly modify the integer Nash welfare objective and add a suitable constant term to the utility of each agent. We show that maximizing this smooth objective function achieves a good approximation to the core. However, maximizing this objective is still NP-hard [70], so we devise local search procedures that run in polynomial time and still give good approximations of the core. In effect, we make a novel connection between appropriate *local optima* of smooth Nash Welfare objectives and the core.

Fairness on Endowments. Classically, the core is defined in terms of agent endowments, not scaled utilities. This is the notion we explored in Chapter 2. The notion of core with endowments logically implies a number of fairness notions considered in multi-winner election literature, such as justified representation, extended justified representation [16], and proportional justified representation [149]. Approval-based multi-winner elections are a special case of packing constraints, in which voters (agents) have binary utilities over a pool of candidates (elements), and we must select a set of at most B candidates. The idea behind proportional representation is to define a notion of large cohesive groups of agents with similar preferences, and ensure that such coalitions are proportionally represented. The core on endowments represents a more general condition that holds for all coalitions of agents, not just those that are large and cohesive. Nevertheless, our local search algorithms for Theorems 5 and 6 are similar to local search algorithms for proportional approval voting (PAV) [156, 17] that achieve proportional representation. It would be interesting to explore the connection between these various notions in greater depth.

Strategyproofness. In this work, we will not consider game-theoretic incentives for manipulation for two reasons. First, even for the restricted case of private goods allocation, preventing manipulation leads to severely restricted mechanisms. For instance, [152] shows that the only strategyproof and Pareto efficient mechanisms are dictatorial, and thus highly unfair, even when there are only two agents with additive utilities over divisible goods. Second, our work is motivated by public goods settings with a large number of agents, such as participatory budgeting, wherein individual agents often have limited influence over the final outcome. It would be interesting to establish this formally, using notions like *strategyproofness in the large* [14].

3.2 Prelude: Nash Social Welfare

Our approach to computing approximate core solutions revolves around the Nash social welfare, which is the product (or equivalently, the sum of logarithms) of agent utilities. This objective is commonly considered to be a natural tradeoff between the fairness-blind utilitarian social welfare objective (maximizing the sum of agent utilities) and the efficiency-blind egalitarian social welfare objective (maximizing the minimum agent utility). This function also has the advantage of being *scale invariant* with respect to the utility function of each agent, and in general, preferring more equal distributions of utility.

3.2.1 Integer Nash Welfare and Smooth Variants

The integer *Max Nash Welfare* (MNW) solution [41, 53] is an outcome $\mathbf{c}$ that maximizes $\sum_{i \in N} \ln u_i(\mathbf{c})$. More technically, if every integer allocation gives zero utility to at least one agent, the MNW solution first chooses a largest set S of agents that can be given non-zero utility simultaneously, and maximizes the product of utilities to agents in S.

[53] argue that this allocation is reasonable by showing that it satisfies proportionality up to one issue for public decision making. A natural question is whether it also provides an approximation of the core. *We answer this question in the negative.* The example below shows that even for public decision making (a special case of matroid constraints), the integer MNW solution may fail to return a (δ, α)-core outcome, for any $\delta = o(m)$ and $\alpha = o(m)$.

Example 1. Consider an instance of public decision making [53] with m issues and two alternatives per issue. Specifically, each issue t has two alternatives $\{a_1^t, a_2^t\}$, and exactly one of them needs to be chosen. There are two sets of agents $X = \{1, \ldots, m\}$ and $Y = \{m+1, \ldots, 2m\}$. Every agent $i \in X$ has $u_i^i(a_1^i) = 1$, and utility 0 for all other alternatives. Every agent $i \in Y$ has $u_i^t(a_2^t) = 1$ and $u_i^t(a_1^t) = 1/m$ for all issues $i \in \{1, 2, \ldots, m\}$. Visually, this is represented as follows.

	a_1^1	a_2^1	a_1^2	a_2^2	$\ldots$	a_1^m	a_2^m
$u_{1 \in X}$	1	0	0	0	$\ldots$	0	0
$u_{2 \in X}$	0	0	1	0	$\ldots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$u_{m \in X}$	0	0	0	0	$\ldots$	1	0
$u_{i \in Y}$	$1/m$	1	$1/m$	1	$\ldots$	$1/m$	1

The integer MNW outcome is $\mathbf{c} = (a_1^1, a_1^2, \ldots, a_1^m)$ because any other outcome gives zero utility to at least one agent. However, coalition Y can deviate, choose outcome $\mathbf{c}' = (a_2^1, a_2^2, \ldots, a_2^m)$, and achieve utility m for each agent in Y. For $\mathbf{c}$ to be a (δ, α)-core outcome, we need

$$\exists i \in Y : (1+\delta) \cdot u_i(\mathbf{c}) + \alpha \geqslant \frac{|Y|}{|Y|+|X|} \cdot u_i(\mathbf{c}') \quad \Rightarrow \quad 1+\delta+\alpha \geqslant \frac{m}{2}.$$

Hence, $\mathbf{c}$ is not a (δ, α)-core outcome for any $\delta = o(m)$ and $\alpha = o(m)$. In contrast, it is not hard to see that $\mathbf{c}'$ is a $(0, 1)$-core outcome because each agent in X gets utility at most one in any outcome.

Further, note that outcome $\mathbf{c}$ gives every agent utility 1. Since $Prop_i \leqslant 1$ for each agent i, $\mathbf{c}$ satisfies proportionality, and yet fails to provide a reasonable approximation to the core. One may argue that $\mathbf{c}'$, which is a $(0, 1)$-core outcome, is indeed fairer because it respects the utility-maximizing choice of half of the population; the other half of the population cannot agree on what they want, so respecting their top choice is arguably a less fair outcome. Hence, the example also shows that outcomes satisfying proportionality (or proportionality up to one issue) can be very different from and less fair than approximate core outcomes.

Smooth Nash Welfare. One issue with the Nash welfare objective is that it is sensitive to agents receiving zero utility. We therefore consider the following smooth Nash welfare objective:

$$F(\mathbf{c}) := \sum_{i \in N} \ln\left(\ell + u_i(\mathbf{c})\right) \tag{3.2}$$

where $\ell \geqslant 0$ is a parameter. Note that $\ell = 0$ coincides with the Nash welfare objective. The case of $\ell = 1$ was considered by [70], who showed it is NP-Hard to optimize. Recall that we normalized agent utilities so that each agent has a maximum utility of 1 for any element, so when we add ℓ to the utility of agent i, it is equivalent to adding $\ell \max_j u_{ij}$ to the utility of agent i when utilities are not normalized.

We show that local search procedures for the smooth Nash welfare objective, for appropriate choices of ℓ, yield a $(0, 2)$-core outcome for matroid constraints (Section 3.3) and a $\left(\delta, O\left(\frac{1}{\delta}\right)\right)$-core outcome for matching constraints (Section 3.4). In contrast, we present Example 2 to show that optimizing any fixed smooth Nash welfare objective cannot guarantee a good approximation to the core, even with *a single* packing constraint, motivating the need for a different algorithm.

3.2.2 Fractional Max Nash Welfare Solution

For general packing constraints, we use a fractional relaxation of the Nash welfare program. A *fractional outcome* consists of a vector $\mathbf{w}$ such that $w_j \in [0,1]$ measures the fraction of element j chosen. The utility of agent i under this outcome is $u_i(\mathbf{w}) = \sum_{j=1}^{m} w_j u_{ij}$.

The fractional *Max Nash Welfare* (MNW) solution is a fractional allocation that maximizes the Nash welfare objective (without any smoothing). Define the packing polytope as:

$$\mathcal{P} = \left\{ \mathbf{w} \in [0,1]^m \mid \sum_{j=1}^{m} a_{kj} w_j \leqslant b_k, \forall k \in [K] \right\}$$

Then the fractional MNW solution is $\arg\max_{\mathbf{c} \in \mathcal{P}} \sum_i \ln u_i(\mathbf{c})$.

It is easy to show that the fractional MNW allocation lies in the core using the same rough proof idea as in the previous chapter. Unfortunately, for the allocation of public goods, it can be shown that the fractional MNW outcome can be irrational despite rational inputs [3], preventing an exact algorithm. For our approximation results, a fractional solution that approximately preserves the utility to each agent would suffice, and we prove the following theorem in Appendix 8.2.

Theorem 8. *For any $\epsilon, \delta > 0$, we can compute a fractional (δ, ϵ)-core outcome in time polynomial in the input size and $\log \frac{1}{\epsilon\delta}$.*

3.3 Matroid Constraints

We now consider public goods allocation with matroid constraints. In particular, we show that when the feasibility constraints encode independent sets of a matroid

$\mathcal{M}$, maximizing the smooth Nash welfare objective in Equation (3.2) with $\ell = 1$ yields a $(0, 2)$-core outcome. However, optimizing this objective is known to be NP-HARD [70]. We also show that given $\epsilon > 0$, a local search procedure for this objective function (given below) yields a $(0, 2 + \epsilon)$-core outcome in polynomial time, which proves Theorem 5.

Algorithm. Fix $\epsilon > 0$. Let $\gamma = \frac{\epsilon}{4m}$, where $m = |W|$ is the number of elements. Recall that there are n agents.

1. Start with an arbitrary basis $\mathbf{c}$ of $\mathcal{M}$.
2. Compute $F(\mathbf{c}) = \sum_{i \in N} \ln(1 + u_i(\mathbf{c}))$.
3. Let a *swap* be a pair (j, j') such that $j \in \mathbf{c}$, $j' \notin \mathbf{c}$, and $\mathbf{c}' = \mathbf{c} \backslash \{j\} \cup \{j'\}$ is also a basis of $\mathcal{M}$.
4. Find a swap such that $F(\mathbf{c}') - F(\mathbf{c}) \geqslant \frac{n\gamma}{m}$.
   - If such a swap exists, then perform the swap, i.e., update $\mathbf{c} \leftarrow \mathbf{c}'$, and go to Step (2).
   - If no such swap exists, then output $\mathbf{c}$ as the final outcome.

Analysis. First, we show that the local search algorithm runs in time polynomial in n, m, and $1/\epsilon$. Note that $F(\mathbf{c}) = O(n \ln m)$ because in our normalization, each agent can have utility at most m. Thus, the number of iterations is $O\left(m^2 \log m/\epsilon\right)$. Finally, each iteration can be implemented in $O(n \cdot m^2)$ time by iterating over all pairs and computing the change in the smooth Nash welfare objective.

Next, let $\mathbf{c}^*$ denote the outcome maximizing the smooth Nash welfare objective with $\ell = 1$, and $\widehat{\mathbf{c}}$ denote the outcome returned by the local search algorithm. We show that $\mathbf{c}^*$ is a $(0, 2)$-core outcome, while $\widehat{\mathbf{c}}$ is a $(0, 2 + \epsilon)$-core outcome.

For outcome $\mathbf{c}$, define $F_i(\mathbf{c}) = \ln(1 + u_i(\mathbf{c}))$. Fix an arbitrary outcome $\mathbf{c}$. For an agent i with $u_i(\mathbf{c}) > 0$, we have that for every element $j \in \mathbf{c}$:

$$F_i(\mathbf{c}) - F_i(\mathbf{c}\backslash\{j\}) \leqslant \frac{u_{ij}}{u_i(\mathbf{c}) + 1 - u_{ij}} \leqslant \frac{u_{ij}}{u_i(\mathbf{c})}.$$

This holds because $\ln(x+h) - \ln x \leqslant \frac{h}{x}$ for $x > 0$ and $h \geqslant 0$. Summing this over all $j \in \mathbf{c}$ gives

$$\sum_{j \in \mathbf{c}} F_i(\mathbf{c}) - F_i(\mathbf{c}\backslash\{j\}) \leqslant \sum_{j \in \mathbf{c}} \frac{u_{ij}}{u_i(\mathbf{c})} = \frac{u_i(\mathbf{c})}{u_i(\mathbf{c})} = 1.$$

For an agent i with $u_i(\mathbf{c}) = 0$, we trivially have $\sum_{j \in \mathbf{c}} F_i(\mathbf{c}) - F_i(\mathbf{c}\backslash\{j\}) = 0$. Summing over all agents, we have that for every outcome $\mathbf{c}$:

$$\sum_{j \in \mathbf{c}} F(\mathbf{c}) - F(\mathbf{c}\backslash\{j\}) = \sum_{i \in N} \sum_{j \in \mathbf{c}} F_i(\mathbf{c}) - F_i(\mathbf{c}\backslash\{j\}) \leqslant n. \tag{3.3}$$

We now use the following result:

Lemma 2 ([112]). *For every pair of bases $\mathbf{c}$ and $\mathbf{c}'$ of a matroid $\mathcal{M}$, there is a bijection $f : \mathbf{c} \to \mathbf{c}'$ such that for every $j \in \mathbf{c}$, $\mathbf{c}\backslash\{j\} \cup \{f(j)\}$ is also a basis.*

Using the above lemma, combined with the fact that $\ln(x+h) - \ln x \geqslant \frac{h}{x+h}$ for $x > 0$ and $h \geqslant 0$, we have that for all $\mathbf{c}, \mathbf{c}'$:

$$\begin{aligned}\sum_{j \in \mathbf{c}} F(\mathbf{c}\backslash\{j\} \cup \{f(j)\}) - F(\mathbf{c}\backslash\{j\}) &\geqslant \sum_{i} \sum_{j \in \mathbf{c}} \frac{u_{if(j)}}{u_i(\mathbf{c}) + 1 - u_{ij} + u_{if(j)}} \\ &\geqslant \sum_{i \in S} \sum_{j' \in \mathbf{c}'} \frac{u_{ij'}}{u_i(\mathbf{c}) + 2} = \sum_{i \in S} \frac{u_i(\mathbf{c}')}{u_i(\mathbf{c}) + 2}.\end{aligned} \tag{3.4}$$

We now provide almost similar proofs for the approximations achieved by the global optimum $\mathbf{c}^*$ and the local optimum $\hat{\mathbf{c}}$.

Global optimum. Suppose for contradiction that $\mathbf{c}^*$ is not a $(0,2)$-core outcome. Then, there exist a subset S of agents and an outcome $\mathbf{c}'$ such that for all $i \in S$,

$$\frac{|S|}{n} \cdot u_i(\mathbf{c}') \geqslant u_i(\mathbf{c}^*) + 2,$$

and at least one inequality is strict. Rearranging the terms and summing over all $i \in S$, we obtain:

$$\sum_{i \in S} \frac{u_i(\mathbf{c}')}{u_i(\mathbf{c}^*) + 2} > \sum_{i \in S} \frac{n}{|S|} = n.$$

Combining this with Equation (3.4), and subtracting Equation (3.3) yields:

$$\sum_{j \in \mathbf{c}^*} \left(F(\mathbf{c}^* \backslash \{j\} \cup \{f(j)\}) - F(\mathbf{c}^*)\right) > 0.$$

This implies existence of a pair $(j, f(j))$ such that $F(\mathbf{c}^* \backslash \{j\} \cup \{f(j)\}) - F(\mathbf{c}^*) > 0$, which contradicts the optimality of $\mathbf{c}^*$ because $\mathbf{c}^* \backslash \{j\} \cup \{f(j)\}$ is also a basis of $\mathcal{M}$.

Local optimum. Similarly, suppose for contradiction that $\widehat{\mathbf{c}}$ is not a $(0, 2+\epsilon)$-core outcome. Then, there exist a subset S of agents and an outcome $\mathbf{c}'$ such that for all $i \in S$,

$$\frac{|S|}{n} \cdot u_i(\mathbf{c}') \geqslant u_i(\widehat{\mathbf{c}}) + 2 + \epsilon > (1+\gamma)\left(u_i(\widehat{\mathbf{c}}) + 2\right).$$

Here, the final transition holds because $\gamma < \epsilon/(m+2) \leqslant \epsilon/(u_i(\widehat{\mathbf{c}}) + 2)$. Again, rearranging and summing over all $i \in S$, we obtain:

$$\sum_{i \in S} \frac{u_i(\mathbf{c}')}{u_i(\widehat{\mathbf{c}}) + 2} > (1+\gamma) \sum_{i \in S} \frac{n}{|S|} \geqslant n \cdot (1+\gamma).$$

Once again, combining this with Equation (3.4), and subtracting Equation (3.3) yields:

$$\sum_{j \in \widehat{\mathbf{c}}} \left(F(\widehat{\mathbf{c}} \backslash \{j\} \cup \{f(j)\}) - F(\widehat{\mathbf{c}})\right) > n\gamma.$$

This implies existence of a pair $(j, f(j))$ such that $F(\widehat{\mathbf{c}} \backslash \{j\} \cup \{f(j)\}) - F(\widehat{\mathbf{c}}) > n\gamma/m$, which violates local optimality of $\widehat{\mathbf{c}}$ because $\widehat{\mathbf{c}} \backslash \{j\} \cup \{f(j)\}$ is also a basis of $\mathcal{M}$.

Lower Bound. While a $(0,2)$-core always outcome exists, we show in the following example that a $(0,1-\epsilon)$-core outcome is not guaranteed to exist for any $\epsilon > 0$.

Lemma 3. *For $\epsilon > 0$ and matroid constraints, $(0,1-\epsilon)$-core outcomes are not guaranteed to exist.*

Proof. Consider the following instance of public decision making where we have several issues and must choose a single alternative for each issue, a special case of matroid constraints. There are n agents, where n is even. There are $m = (n-2)+n/2$ issues. The first $n-2$ issues correspond to unit-value private goods, *i.e.*, each such issue has n alternatives, and each alternative gives utility 1 to a unique agent and utility 0 to others. The remaining $n/2$ issues are "pair issues"; each such issue has $\binom{n}{2}$ alternatives, one corresponding to every pair of agents that gives both agents in the pair utility 1 and all other agents utility 0.

It is easy to see that every integer allocation gives utility at most 1 to at least two agents. Consider the deviating coalition consisting of these two agents. They can choose the alternative that gives them each utility 1 on every pair issue, and split the $n-2$ private goods equally. Thus, they each get utility $n/2 + (n-2)/2 = n-1$. For the outcome to be a $(0,\alpha)$-core outcome, we need $1 + \alpha \geqslant (2/n) \cdot (n-1)$. As $n \to \infty$, this requires $\alpha \to 1$. Hence, for any $\epsilon > 0$, a $(0,1-\epsilon)$-core outcome is not guaranteed to exist. □

Note that Theorem 5 shows existence of a $(0,2)$-core outcome, which is therefore tight up to a unit additive relaxation. Whether a $(0,1)$-core outcome always exists under matroid constraints remains an important open question. Interestingly, we show that such an outcome always exists for the special case of private goods allocation, and, in fact, can be achieved by maximizing the smooth Nash welfare objective.

Lemma 4. *For private goods allocation, maximizing the smooth Nash welfare objective with $\ell = 1$ returns a $(0,1)$-core outcome.*

Proof. There is a set of agents N and a set of private goods M. Each agent $i \in N$ has a utility function $u_i : 2^M \to \mathbb{R}_{\geqslant 0}$. Utilities are additive, so $u_i(S) = \sum_{g \in S} u_i(\{g\})$ for all $S \subseteq M$. For simplicity, we denote $u_{ig} \triangleq u_i(\{g\})$. Without loss of generality, we normalize the utility of each agent such that $\max_{g \in M} u_{ig} = 1$ for each i. An allocation A is a *partition* of the set of goods among the agents; let A_i denote the bundle of goods received by agent i. We want to show that an allocation maximizing the objective $\prod_{i \in N}(1 + u_i(A_i))$ is a $(0,1)$-core outcome.

Let A denote an allocation maximizing the smooth Nash welfare objective with $\ell = 1$. We assume without loss of generality that every good is positively valued by at least one agent. Hence, $u_j(A_j) = 0$ must imply $A_j = \varnothing$.

For agents $i, j \in N$ with $A_j \neq \varnothing$ (hence $u_j(A_j) > 0$), and good $g \in A_j$, moving g to A_i should not increase the objective function. Hence, for each $g \in A_j$, we have

$$\big(1 + u_i(A_i \cup \{g\})\big) \cdot \big(1 + u_j(A_j \backslash \{g\})\big) \leqslant \big(1 + u_i(A_i)\big) \cdot \big(1 + u_j(A_j)\big).$$

Using additivity of utilities, this simplifies to

$$\frac{u_{ig}}{1 + u_i(A_i)} \leqslant \frac{u_{jg}}{1 + u_j(A_j) - u_{jg}} \leqslant \frac{u_{jg}}{u_j(A_j)}. \tag{3.5}$$

For every agent $j \in N$ with $A_j \neq \varnothing$ and good $g \in A_j$, define $p_g = u_{jg}/u_j(A_j)$. Abusing the notation a little, for a set $T \subseteq M$ define $p_T = \sum_{g \in T} p_g$. Then, from Equation (3.5), we have that for all players $i \in N$ and goods $g \in M$,

$$(1 + u_i(A_i)) \cdot p_g \geqslant u_{ig}. \tag{3.6}$$

Suppose for contradiction that A is not a $(0,1)$-core outcome. Then, there exists a set of agents $S \subseteq N$ and an allocation B of the set of all goods to agents in S such that $(|S|/n) \cdot u_i(B_i) \geqslant 1 + u_i(A_i)$ for every agent $i \in S$, and at least one inequality is strict. Rearranging the terms and summing over $i \in S$, we have

$$\sum_{i \in S} \frac{u_i(B_i)}{1 + u_i(A_i)} > \sum_{i \in S} \frac{n}{|S|} = n. \tag{3.7}$$

We now derive a contradiction. For agent $i \in S$, summing Equation (3.6) over $g \in B_i$, we get

$$(1 + u_i(A_i)) \cdot p_{B_i} \geqslant u_i(B_i) \Rightarrow \frac{u_i(B_i)}{1 + u_i(A_i)} \leqslant p_{B_i}.$$

Summing this over $i \in S$, we get

$$\sum_{i \in S} \frac{u_i(B_i)}{1 + u_i(A_i)} \leqslant \sum_{i \in S} p_{B_i} = \sum_{g \in M} p_g = \sum_{\substack{j \in N \text{ s.t.} \\ A_j \neq \varnothing}} \sum_{g \in A_j} \frac{u_{jg}}{u_j(A_j)} = \sum_{\substack{j \in N \text{ s.t.} \\ A_j \neq \varnothing}} \frac{u_j(A_j)}{u_j(A_j)} \leqslant n.$$

However, this contradicts Equation (3.7). □

3.4 Matching Constraints

We now present the algorithm proving Theorem 6. We show that if the elements are edges of an undirected graph $G(V, E)$, and the feasibility constraints encode a matching, then for constant $\delta \in (0, 1]$, a $(\delta, 8 + \frac{6}{\delta})$-core always exists and is efficiently computable. The idea is to again run a *local search* on the smooth Nash welfare objective in Equation (3.2), but this time with $\ell \approx 1 + \frac{4}{\delta}$.

Algorithm. Recall that there are n agents. Let $|V| = r$ and $|E| = m$. Let $\kappa = \frac{2}{\delta}$. For simplicity, assume $\kappa \in \mathbb{N}$. Our algorithm is inspired by the PRAM algorithm for approximate maximum weight matchings due to [92], and we follow their terminology. Given a matching $\mathbf{c}$, an *augmentation* with respect to $\mathbf{c}$ is a matching $T \subseteq E \backslash \mathbf{c}$. The *size* of the augmentation is $|T|$. Let $M(T)$ denote the subset of edges of $\mathbf{c}$ that have a vertex which is matched under T. Then, the matching $(\mathbf{c} \backslash M(T)) \cup T$ is called the *augmentation* of $\mathbf{c}$ using T.

1. Start with an arbitrary matching $\mathbf{c}$ of G.

2. Compute $F(\mathbf{c}) = \sum_i \ln(1 + 2\kappa + u_i(\mathbf{c}))$.

3. Let $\mathcal{C}$ be the set of all augmentations with respect to $\mathbf{c}$ of size at most κ.

   - If there exists $T \in \mathcal{C}$ such that $F((\mathbf{c} \backslash M(T)) \cup T) - F(\mathbf{c}) \geqslant \frac{n}{\kappa r}$, perform this augmentation (i.e., let $\mathbf{c} \leftarrow (\mathbf{c} \backslash M(T)) \cup T$) and go to Step (2).

   - Otherwise, output $\mathbf{c}$ as the final outcome.

Analysis. The outline of the analysis is similar to the analysis for matroid constraints. First, we show that the algorithm runs in polynomial time. Again, recall that each agent has utility at most m. Thus, $F(\mathbf{c}) = O(n \cdot \ln m)$. Because each improvement increases the objective value by at least $n/(\kappa r)$, the number of iterations is $O(\kappa r \ln m) = O(m^2/\delta)$. Each iteration can be implemented by naïvely going over all $O(m^\kappa)$ subsets of edges of size at most κ, checking if they are valid augmentations with respect to $\mathbf{c}$, and whether they improve the objective function by more than $n/(\kappa r)$. The local search therefore runs in polynomial time for constant $\delta > 0$.

Let $\mathbf{c}$ denote the outcome returned by the algorithm. We next show that $\mathbf{c}$ is indeed a $(\delta, 8 + 3\kappa)$-core outcome. Suppose for contradiction that this is not true. Then, there exist a subset of agents S and a matching $\mathbf{c}'$ such that for all $i \in S$,

$$\frac{|S|}{n} \cdot u_i(\mathbf{c}') \geqslant (1 + \delta) \cdot u_i(\mathbf{c}) + 8 + 3\kappa \geqslant (1 + \delta) \cdot (u_i(\mathbf{c}) + 3\kappa + 1),$$

and at least one inequality is strict (the last inequality is because $\delta \in (0, 1]$). Rearranging and summing over all $i \in S$, we obtain

$$\sum_{i \in S} \frac{u_i(\mathbf{c}')}{u_i(\mathbf{c}) + 3\kappa + 1} > (1 + \delta) \cdot \sum_{i \in S} \frac{n}{|S|} = n \cdot (1 + \delta). \tag{3.8}$$

For $j \in E$, define $w_j = \sum_{i \in N} \frac{u_{ij}}{u_i(\mathbf{c})+1}$ and $w'_j = \sum_{i \in N} \frac{u_{ij}}{u_i(\mathbf{c})+3\kappa+1}$. Let $W = \sum_{j \in \mathbf{c}} w_j$, and $W' = \sum_{j \in \mathbf{c}'} w'_j$. It is easy to check that

$$W \leqslant n \quad \text{and} \quad W' \geqslant n \cdot (1+\delta), \tag{3.9}$$

where the latter follows from Equation (3.8). Further note that $w_j \geqslant w'_j$ for all j.

For an augmentation T with respect to $\mathbf{c}$, define $\text{gain}(T) = \sum_{j \in T} w'_j - \sum_{j \in M(T)} w_j$. The next lemma is a simple generalization of the analysis in [92]; we give the adaptation here for completeness.

Lemma 5. *Assuming weights $w_j \geqslant w'_j$ for all edges j, for any integer $\kappa \geqslant 1$ and matchings $\mathbf{c}$ and $\mathbf{c}'$, there exists a multiset OPT of augmentations with respect to $\mathbf{c}$ such that:*

- *For each $T \in OPT$, $T \subseteq \mathbf{c}'$ and $|T| \leqslant \kappa$;*
- *$|OPT| \leqslant \kappa r$; and*
- *$\sum_{T \in OPT} \text{gain}(T) \geqslant \kappa \cdot W' - (\kappa+1) \cdot W$.*

Proof. We follow [92] in the construction the multiset OPT of augmentations with respect to $\mathbf{c}$ out of edges in $\mathbf{c}'$. Let $\mathbf{c} \triangle \mathbf{c}'$ be the symmetric difference of matchings $\mathbf{c}$ and $\mathbf{c}'$ consisting of alternating paths and cycles. For every cycle or path $\mathbf{d} \in \mathbf{c} \triangle \mathbf{c}'$, let $T_{\mathbf{d}}$ be be set of edges $\mathbf{d} \cap \mathbf{c}'$. For all $T_{\mathbf{d}}$ with $|T_{\mathbf{d}}| \leqslant \kappa$, just add T_S to OPT κ times (note that OPT is a multiset, not a set). For $T_{\mathbf{d}}$ with $|T_{\mathbf{d}}| > \kappa$, we break up $T_{\mathbf{d}}$ into multiple smaller augmentations. To do so, index the edges in $T_{\mathbf{d}}$ from 1 to $|T_{\mathbf{d}}|$ and add $|T_{\mathbf{d}}|$ different augmentations to OPT by considering starting at every index in $T_{\mathbf{d}}$ and including the next κ edges in $T_{\mathbf{d}}$ with wrap-around from $|T_{\mathbf{d}}|$ to 1.

Now we must argue that OPT as we have constructed it satisfies the conditions of the lemma. The first point, that $\forall T \in OPT, T \subseteq \mathbf{c}'$ and $|T| \leqslant \kappa$, follows trivially from the construction. The second point also follows easily from observing that we add κ augmentations to OPT for every $\mathbf{d} \in \mathbf{c} \cap \mathbf{c}'$, and graph G has r vertices.

To see the third point, note that every edge in $\mathbf{c}' \backslash \mathbf{c}$ is contained in at least κ augmentations in OPT. On the other hand, for every edge $e \in \mathbf{c} \backslash \mathbf{c}'$, there are no more than $\kappa + 1$ augmentations $T \in OPT$ such that $e \in M(T)$ (recall $M(T)$ are the edges of $\mathbf{c}$ with a vertex matched under T). This can happen, for example, if T_S happens to be a path of length $\kappa + 1$. Finally, for the edges $j \in \mathbf{c}' \cap \mathbf{c}$, the weight $w'_j \leqslant w_j$. Putting these facts together, the third point of the lemma follows. □

Consider the set of augmentations OPT from Lemma 5. For $T \in OPT$,

$$
\begin{aligned}
& F((\mathbf{c} \backslash M(T)) \cup T) - F(\mathbf{c}) \\
&= \Big(F((\mathbf{c} \backslash M(T)) \cup T) - F(\mathbf{c} \backslash M(T)) \Big) - \Big(F(\mathbf{c}) - F(\mathbf{c} \backslash M(T)) \Big) \\
&\geqslant \sum_{i \in N} \left(\frac{\sum_{T \in S} u_{ij}}{u_i(\mathbf{c}) + 2\kappa + 1 + \sum_{j \in T} u_{ij}} - \frac{\sum_{j \in M(T)} u_{ij}}{u_i(\mathbf{c}) + 2\kappa + 1 - \sum_{j \in M(T)} u_{ij}} \right) \\
&\geqslant \sum_{i \in N} \left(\frac{\sum_{j \in T} u_{ij}}{u_i(\mathbf{c}) + 3\kappa + 1} - \frac{\sum_{j \in M(T)} u_{ij}}{u_i(\mathbf{c}) + 1} \right) \\
&= \sum_{j \in T} w'_j - \sum_{j \in M(T)} w_j = \mathrm{gain}(T).
\end{aligned}
$$

Here, the second transition holds because $h/(x+h) \leqslant \ln(x+h) - \ln x \leqslant h/x$ for all $x \geqslant 1$ and $h \geqslant 0$, and the third transition holds due to $|T| \leqslant \kappa$ and $|M(T)| \leqslant 2|T| \leqslant 2\kappa$. Therefore, we have:

$$
\begin{aligned}
\sum_{T \in OPT} F((\mathbf{c} \backslash M(T)) \cup T) - F(\mathbf{c}) \geqslant \sum_{T \in OPT} \mathrm{gain}(T) &\geqslant \kappa \cdot W' - (\kappa + 1) \cdot W \\
&\geqslant \kappa \cdot n \cdot (1 + \delta) - (\kappa + 1) \cdot n = n,
\end{aligned}
$$

where the second transition follows from Lemma 5, and the third transition follows from Equation (3.9). Since $|OPT| \leqslant \kappa r$, there exists an augmentation $T \in OPT$

with $F((\mathbf{c}\backslash M(T)) \cup T) - F(\mathbf{c}) \geqslant {}^{n}\!/\!{\kappa r}$, which violates local optimality of $\mathbf{c}$. This completes the proof of Theorem 6.

Lower Bound. We give a stronger lower bound for matchings than the lower bound for matroids in Lemma 3.

Lemma 6. *A (δ, α)-core outcome is not guaranteed to exist for matching constraints, for any $\delta \geqslant 0$ and $\alpha < 1$.*

Proof. This example shows that Consider the graph $K_{2,2}$ (the complete bipartite graph with two vertices on each side). This graph has four edges, and two disjoint perfect matchings.

Let there be two agents. Agent 1 has unit utility for the edges of one matching, while agent 2 has unit utility for the edges of the other matching. Any integer outcome gives zero utility to one of these agents. This agent can deviate and obtain utility 2. Hence, for an outcome to be a (δ, α)-core outcome, we need $(1+\delta) \cdot 0 + \alpha \geqslant (1/2) \cdot 2$, which is impossible for any $\delta \geqslant 0$ and $\alpha < 1$. □

3.5 General Packing Constraints

In this section, we study approximation to the core under general packing constraints of the form $A\mathbf{x} \leqslant \mathbf{b}$. Recall that there are m elements, V_i is the maximum possible utility that agent i can receive from a feasible outcome, and $V_{\max} = \max_{i \in N} V_i$. We prove a statement slightly more general than Theorem 7. We first need the following concept. Given an instance of public goods allocation subject to packing constraints, we define the notions of an r-proportionally fair (r-PF) outcome, a maximally proportionally fair (MPF) outcome, and the MPF value of the instance.

Definition 11 (MPF Outcome)**.** *For $r > 0$, we say that a fractional outcome $\mathbf{w}$ is*

r-proportionally fair *(r-PF) if it satisfies:*

$$u_i(\mathbf{w}) \geqslant \frac{V_i}{r} - 1, \quad \forall i \in N.$$

The maximally proportionally fair *(MPF) value R of an instance is the least value r such that there exists an r-PF outcome. For simplicity, we say that an R-PF outcome is a* maximally proportionally fair *(MPF) outcome.*

This concept is crucial to stating and deriving our approximation results. In words, an r-PF outcome gives each agent an r fraction of its maximum possible utility V_i (which can be thought of as the fair share guarantee of the agent), if the agent is given 1 unit of utility for free. Thus, a smaller value of r indicates a better solution. The MPF value R denotes the best possible guarantee. The additive 1 in Def. 11 can be replaced by any positive constant; we choose 1 for simplicity.

We now show an upper bound for R that holds for all instances. Recall from Equation (3.1) that ρ is the *width* of the instance.

Lemma 7. *$R \leqslant \min(V_{\max}, n, \rho)$, and an MPF outcome is computable in polynomial time.*

Proof. To show that R is well-defined, note that for $r = V_{\max}$, an r-PF outcome $\mathbf{w}$ simply requires $u_i(\mathbf{w}) \geqslant 0$, which is trivially achieved by every outcome. Therefore, R is well-defined, and $R \leqslant V_{\max}$. Next, $R \leqslant n$ follows from the fact that there exist fractional outcomes satisfying proportionality (e.g., the outcome $\mathbf{w}$ obtained by taking the uniform convex combination of the n outcomes that give optimal to each individual agent). Finally, to show $R \leqslant \rho$, consider the outcome $\mathbf{w}$ in which $w_j = \frac{1}{\rho}$ for each element j. Clearly, $u_i(\mathbf{w}) \geqslant \frac{V_i}{\rho}$ for all i. Further, $A\mathbf{w} \leqslant \mathbf{b}$ is satisfied trivially due to the fact that ρ is the width of the packing constraints.

To compute the value of R as well as an MPF outcome, we first note that the value of V_i for each agent i can be computed by solving a separate LP. Then, we consider the following LP:

$$\text{Maximize } \hat{r} \tag{3.10}$$

$$\begin{array}{rcll} \sum_{j \in W} u_{ij} w_j & \geqslant & V_i \cdot \hat{r} - 1 & \forall i \in [n] \\ A\mathbf{w} & \leqslant & \mathbf{b} & \\ w_j & \in & [0,1] & \forall j \in W \end{array}$$

Here, $A\mathbf{w} \leqslant \mathbf{b}$ are the packing constraints of the instance, and $\hat{r}$ is a variable representing $1/r$. Thus, maximizing $\hat{r}$ minimizes r, which yields an MPF outcome. This can be accomplished by solving $n+1$ linear programs, which can be done in polynomial time. □

Our main result in this section uses any r-PF outcome, and provides a guarantee in terms of $\log r$. Thus, we do not need to necessarily compute an exact MPF outcome. We note that an MPF outcome can be very different from a core outcome. Yet, an MPF outcome gives each agent a large fraction of its maximum possible utility, subject to a small additive relaxation. As we show below, this helps us find integral outcomes that provide good approximations of the core.

3.5.1 Result and Proof Idea

Our main result for this section (Theorem 7) can be stated more precisely. Recall that $\log^*$ is the iterated logarithm, which is the number of times the logarithm function must be iteratively applied before the result becomes less than or equal to 1.

Theorem 9. *Fix constant $\delta \in (0,1)$. Suppose we are given a set of K packing constraints $A\mathbf{x} \leqslant \mathbf{b}$ such that $b_k = \omega\left(\frac{\log K}{\delta^2}\right)$ for all $k \in [K]$. Let R be the MPF value of this instance. Then there exists a polynomial time computable (δ, α)-core outcome, where*

$$\alpha = O\left(\frac{1}{\delta^4} \cdot \log\left(\frac{R \cdot \log^* V_{\max}}{\delta}\right)\right).$$

We first note that the above result cannot be obtained by maximizing the smooth Nash welfare objective; we present Example 2, which demonstrates this using only one packing constraint. To be precise, the example shows that no single value of parameter ℓ in the smooth Nash welfare objective can provide a polylog additive guarantee for all instances. While it may be possible to choose the value of ℓ based on the instance, it does not seem trivial.

Example 2. Recall that in the KNAPSACK setting, we are given a set of elements of different sizes, and our goal is to select a subset of elements with total size at most a given budget B. We show that for any $\ell > 0$, there exists a KNAPSACK instance in which maximizing the smooth Nash welfare objective $F(\mathbf{c}) = \sum_{i \in N} \ln(\ell + u_i(\mathbf{c}))$ returns an outcome that is not a $(O(m^{1/2-\epsilon}), O(m^{3/4-\epsilon}))$-core outcome. This is in contrast to Theorem 9, which provides a (δ, α)-core guarantee where δ is constant and α is logarithmic in the number of elements.

Fix $\ell > 0$. Set a large budget $B \geqslant \ell^4$. There are $m = B^{1/4} + B$ elements, of which $B^{1/4}$ are *large* elements of size $B^{3/4}$ and the remaining B are *small* elements of unit size. There are $n \geqslant 4B^{1/4} \log(2B)$ agents. Each agent has unit utility for each large element. A subset of αn agents are *special* (we determine α later). These special agents have unit utility for each small element, while the remaining agents have zero utility for the small elements.

The idea is to show that when α is sufficiently small, the smooth Nash welfare objective will choose only the large elements. However, α can still be large enough so that the special agents can deviate, and get a large amount of utility.

Note that maximizing the smooth Nash welfare objective returns a Pareto efficient solution, and hence can be one of two types: it either chooses all large elements (which gives utility $B^{1/4}$ to each agent), or it chooses $B^{1/4} - r$ large elements and $rB^{3/4}$ small elements. For the former to have a larger smooth Nash welfare objective value, we need that for each $1 \leqslant r \leqslant B^{1/4}$,

$$\ln(B^{1/4} + \ell) > \alpha \ln\left(rB^{3/4} + (B^{1/4} - r) + \ell\right) + (1 - \alpha)\ln(B^{1/4} - r + \ell).$$

This holds true if

$$\ln\left(\frac{B^{1/4} + \ell}{B^{1/4} + \ell - r}\right) > \alpha \ln\left(rB^{3/4} + (B^{1/4} - r) + \ell\right).$$

Since $0 < \ell \leqslant B^{1/4}$, the above is true for each $1 \leqslant r \leqslant B^{1/4}$ if

$$\ln\left(\frac{2B^{1/4}}{2B^{1/4} - 1}\right) > \alpha \ln\left(B^{1/4}B^{3/4} + B^{1/4}\right).$$

This is true when

$$\ln\left(1 + \frac{1}{2B^{1/4}}\right) \geqslant \alpha \ln(2B).$$

Since $\ln(1 + x) \geqslant x/2$ for $x \in [0, 1]$, the above holds when:

$$\alpha \leqslant \frac{1}{4B^{1/4}\ln(2B)}.$$

Let us set $\alpha = \frac{1}{4B^{1/4}\ln(2B)}$. Choosing all large elements maximizes the smooth Nash welfare objective. Since $n \geqslant 1/\alpha$, there is at least one special agent. The special agents get utility $B^{1/4}$ each. If they deviate and choose all the small elements, they get (scaled down) utility

$$\alpha B = \frac{B}{4B^{1/4}\ln(2B)} = \frac{B^{3/4}}{4\ln(2B)}.$$

Hence, for the solution to be a (δ, α)-core outcome, we need $(1 + \delta) \cdot B^{1/4} + \alpha \geqslant B^{3/4}/(4\ln(2B))$. Since $m = \Theta(B)$, this shows that the outcome is not a $(O(m^{1/2-\epsilon}), O(m^{3/4-\epsilon}))$-core outcome for any constant $\epsilon > 0$, as required.

We take a different approach. Our idea is to start with a fractional core solution $\mathbf{x}$. Suppose it assigns utility U_i^* to agent i. Fix $\delta > 0$, and consider the following program.

$$\text{Minimize} \quad \alpha \tag{3.11}$$

$$\begin{array}{rcll} \alpha + (1+\delta) \cdot \sum_{j \in W} u_{ij} w_j & \geqslant & U_i^* & \forall i \in [n] \\ A\mathbf{w} & \leqslant & \mathbf{b} & \\ w_j & \in & \{0,1\} & \forall j \in W \\ \alpha & \geqslant & 0 & \end{array}$$

For the optimum value α^*, we obtain an outcome that is (δ, α')-core for every $\alpha' > \alpha$. To see this, take a subset of agents S and a feasible utility vector $\mathbf{U}'$ under any other (even fractional) outcome. Because $\mathbf{x}$ is a core outcome, there exists $i \in S$ such that $U_i^* \geqslant (|S|/n) \cdot U_i'$. For $\alpha' > \alpha$, the ILP solution implies

$$\alpha' + (1+\delta) \cdot \sum_{j \in W} u_{ij} w_j > U_i^* \geqslant \frac{|S|}{n} \cdot U_i',$$

which implies that the solution is (δ, α')-core according to Definition 10.

However, α^* obtained from this program can be rather large, as illustrated in the following example. Consider the KNAPSACK setting with m unit-size projects. There is an overall budget $B = m/2$. For every feasible integral outcome $\mathbf{c}$, let there be an agent with utility 1 for every project in $\mathbf{c}$ and 0 for all other projects. Thus, there are $\binom{m}{m/2}$ agents. The fractional core outcome gives weight 1/2 to each project, thus giving utility $V_i/2 = m/4$ to each agent i. However, every integral outcome gives utility 0 to at least one agent, which implies $\alpha^* = \Omega(m)$.

This example shows that when there are a large number of agents, we cannot achieve Theorem 9 by hoping to approximately preserve the utilities to *all* agents with respect the fractional core solution. However, note that in the above example, though there is one agent who gets very little utility, this agent has no incentive to deviate if she is given one unit of utility for free. This insight leads us to our analysis below, which is based on rounding the fractional core solution $\mathbf{x}$.

Let us apply randomized rounding to $\mathbf{x}$. Instead of using Chernoff bounds to ensure that there are no "violations" (i.e., that no agent receives utility that is too far from its utility under the core outcome $\mathbf{x}$), we hope to bound the expected number of such violations. If there are few such agents, we still have an approximate core outcome because if this small coalition of agents deviates, its utility under a new outcome will be scaled down by a large factor. Unfortunately, it can be shown that bounding the expected number of deviations by a sufficiently small number forces $\alpha = \Omega(\log V_{\max})$. This is better than $\alpha = \Omega(m)$ from our previous approach, but still *much* larger than the bound we want to achieve in Theorem 9 when the width ρ is small.

This brings up the main technical idea. We observe that an MPF outcome, though not in the core, provides a reasonably large utility to each agent. We add a small amount of this outcome to the fractional core before applying randomized rounding. We are now ready to present our algorithm.

3.5.2 Algorithm

Fix $\delta \in (0, 1)$, and let $\gamma = \frac{\delta}{8}$.

1. Compute the (approximate) fractional core solution $\mathbf{x}$ as in Theorem 8, where x_j is the fraction of element j chosen.

2. Let $\mathbf{y}$ be an MPF outcome as in Definition 11.

3. Let $\mathbf{z} = (1 - \gamma)\mathbf{x} + \gamma\mathbf{y}$.

4. For each $j \in W$, choose j to be in the outcome $\mathbf{c}$ independently with probability $\hat{z}_j = (1 - \gamma)z_j$.

3.5.3 Analysis

We show that this algorithm yields, with at least a constant probability, a feasible outcome that satisfies the guarantee in Theorem 9. This directly shows the existence of such an outcome. Note that the fractional Max Nash Welfare solution $\mathbf{x}$ can be irrational, but we can compute an approximation in polynomial time (see Theorem 8 for details), which does not change our guarantee asymptotically. Further, $\mathbf{y}$ can be computed in polynomial time (Lemma 7). Hence, the algorithm runs in expected polynomial time.

We first show that the packing constraints are satisfied. Since we scale down $\mathbf{z}$ by a factor $(1-\gamma)$ before rounding, we have $A\hat{z} \leqslant (1-\gamma)\mathbf{b}$. Since $\mathbf{b} = \omega\left(\frac{\log K}{\delta^2}\right)$, a simple application of Chernoff bounds shows that with probability at least 0.99, the rounded solution $\mathbf{c}$ satisfies $A\mathbf{c} \leqslant \mathbf{b}$. Therefore, if we show that the algorithm also yields the desired approximation of the core with at least a constant probability (1/6 to be precise), we will be done by applying union bound to the two events, feasibility and good approximation to the core.

For the ease of presentation, we suppress constants throughout the proof and use asymptotic notation liberally. We also assume that $V_{\max} = \omega(1)$ since otherwise there is a trivial $(0, O(1))$-core outcome that chooses a null outcome, giving zero utility to each agent.

Grouping Agents. In order to analyze our algorithm, we partition the agents into groups with exponentially decreasing values of V_i. Recall that V_i is the maximum utility that agent i can get from any outcome. Set $Q_0 = \log V_{\max}$, and for $\ell = 0, 1, \ldots, L-1$, define group G_ℓ as:

$$G_\ell = \{i \in N \mid Q_\ell \geqslant \log V_i \geqslant Q_{\ell+1}\}.$$

Here, for $\ell = 0, 1, \ldots, L-1$, we define: $Q_{\ell+1} = 2 \log Q_\ell$.

We call $G_0, \ldots, G_{L-1}$ the *heavy groups*. We choose L so that $Q_L = \Theta\left(\log \frac{R \log^* V_{\max}}{\gamma^3}\right)$. This implies $L = \Omega(\log^* V_{\max}) = \omega(1)$, since $V_{\max} = \omega(1)$. For agent i in a heavy group, $V_i \geqslant e^{Q_L} \geqslant \frac{2RL}{\gamma^3} > 2R$. Thus, the utility that the MPF solution provides to agent i is at least $\frac{V_i}{R} - 1 \geqslant \frac{V_i}{2R}$.

Finally, we put the remaining agents (with a small V_i) in a *light group* defined as follows:

$$G_L = \{i \in N \mid \log V_i \leqslant Q_L\}.$$

The MPF solution may not provide any guarantee for the utility of agents in this group.

Bounding Violations of Utility Preservation. We want to bound the number of agents whose utilities are far from those under the core outcome. First, we need a specialized Chernoff bound.

Lemma 8. *Let $X_1, X_2, \ldots, X_q$ be independent random variables in $[0,1]$, and let $X = \sum_{j=1}^{q} X_j$. For $\gamma \in (0, 1/2)$, suppose $\mathbf{E}[X] = (1-\gamma) \cdot A + \gamma \cdot B$ for $A, B \geqslant 0$. Then*

$$\Pr[X < (1-2\gamma) \cdot A] \leqslant e^{-\frac{\gamma^3}{2} \max(B, A/2)}$$

Proof. We first state the standard theorem for Chernoff bounds. Let $X_1, X_2, \ldots, X_q$ be independent random variables in $[0,1]$, and let $X = \sum_{j=1}^{q} X_j$. For any $\epsilon \in (0,1)$, we have:

$$\Pr\left[X < (1-\epsilon)\mathbf{E}[X]\right] \leqslant e^{-\frac{\epsilon^2}{2}\mathbf{E}[X]}$$

Equivalently, for any $\eta < \mathbf{E}[X]$,

$$\Pr\left[X < \mathbf{E}[X] - \eta\right] \leqslant e^{-\frac{\eta^2}{2\mathbf{E}[X]}}$$

Lemma 8 follows from considering two cases. First, suppose $(1-\gamma)A \geqslant B$.

$$\Pr[X < (1-\gamma)^2 A] \leqslant \Pr[X < (1-\gamma)\mathbf{E}[X]] \leqslant e^{-\frac{\gamma^2}{2}\mathbf{E}[X]}$$
$$\leqslant e^{-\frac{\gamma^2}{2}\max(\gamma B,(1-\gamma)A)} \leqslant e^{-\frac{\gamma^3}{2}\max(B,A)}$$

In the other case, if $(1-\gamma)A < B$, then $\gamma B \leqslant \mathbf{E}[X] \leqslant (1+\gamma)B$. Then

$$\begin{aligned} \Pr[X < (1-\gamma)A] &\leqslant \Pr[X \leqslant \mathbf{E}[X] - \gamma B] \\ &\leqslant e^{-\gamma^2 \frac{B^2}{2\mathbf{E}[X]}} \leqslant e^{-\gamma^2 \frac{B^2}{2(1+\gamma)B}} \\ &\leqslant e^{-\frac{\gamma^3}{2}B} \leqslant e^{-\frac{\gamma^3}{2}(1-\gamma)A} \leqslant e^{-\frac{\gamma^3}{4}A} \end{aligned}$$

□

Recall that $\mathbf{x}$ is the fractional MNW solution, $\mathbf{y}$ is the fractional MPF solution, and our algorithm applies randomized rounding to their scaled down mixture $(1-\gamma)\mathbf{z} = (1-\gamma)^2\mathbf{x} + \gamma(1-\gamma)\mathbf{y}$. Let $\widehat{U}_i$ denote the utility of agent i under the final integral outcome obtained by randomly rounding $(1-\gamma)\mathbf{z}$. Recall that U_i^* is the utility of agent i under the core outcome $\mathbf{x}$. We want to show that $\widehat{U}_i$ is either multiplicatively or additively close to U_i^* for most agents. For a heavy group G_ℓ, where $\ell \in \{0, 1, \ldots, L-1\}$, define

$$F_\ell = \left\{ i \in G_\ell \;\middle|\; \widehat{U}_i < (1-3\gamma)U_i^* \right\}.$$

Simiarly, for the light group G_L, define

$$F_L = \left\{ i \in G_L \;\middle|\; \widehat{U}_i < \min\left((1-3\gamma)U_i^*, U_i^* - \frac{4Q_L}{\gamma^4} \right) \right\}.$$

We will use Lemma 8 to bound the sizes of F_ℓ for $\ell \in \{0, 1, \ldots, L\}$ as follows.

Theorem 10. *We have that:*

1. *With probability at least* $2/3$*, we have* $|F_\ell| \leqslant \frac{1}{2Le^{Q_\ell}} \cdot |G_\ell|, \quad \forall \ell \in \{0, 1, \ldots, L-1\}$.

2. *With probability at least* $1/2$*, we have* $|F_L| \leqslant \frac{1}{2e^{Q_L}} \cdot |G_L|$.

Thus, with probability at least $1/6$*, both the above inequalities hold simultaneously.*

Proof. We prove the first and the second part of Theorem 10 separately by considering the heavy groups and the light group in turn. The combined result follows from the union bound.

Case 1: Heavy Groups Consider a heavy group G_ℓ for $0 \leqslant \ell < L$. Recall that the MPF solution provides utility at least $V_i/(2R)$ to each agent in a heavy group. Hence, we have:

$$\mathbf{E}\left[\widehat{U}_i/(1-\gamma)\right] = u_i(\mathbf{z}) \geqslant (1-\gamma) \cdot U_i^* + \gamma \cdot \frac{V_i}{2R}. \tag{3.12}$$

The key point is that even if U_i^* is small, the expected utility is at least a term that is proportional to V_i. This will strengthen our application of Chernoff bounds. Using Lemma 8 with $A = U_i^*$ and $B = V_i/(2R)$, we have:

$$\begin{aligned}\Pr\left[\widehat{U}_i < (1-3\gamma) \cdot U_i^*\right] &\leqslant \Pr\left[\frac{\widehat{U}_i}{1-\gamma} < (1-2\gamma) \cdot U_i^*\right] \\ &\leqslant e^{-\frac{\gamma^3}{4}\frac{V_i}{2R}} \leqslant e^{-\frac{\gamma^3}{8R}Q_\ell^2} \\ &\leqslant e^{-Q_\ell \cdot \log L} \leqslant e^{-(Q_\ell + 2\log L + \log 6)} \leqslant \frac{1}{6L^2 e^{Q_\ell}},\end{aligned} \tag{3.13}$$

where the second inequality holds because $\log V_i \geqslant 2 \log Q_\ell$, the third holds because $Q_\ell = \Omega\left(\frac{RL}{\gamma^3}\right)$, and the fourth holds because $L = \omega(1)$.

We are now ready to prove the first part of Theorem 10. Let $\eta_\ell = \frac{1}{6L^2 e^{Q_\ell}}$. Recall that F_ℓ consists of agents in G_ℓ for which $\widehat{U}_i < (1-3\gamma) \cdot U_i^*$. Using the linearity of expectation in Equation (3.13), we have $\mathbf{E}[F_\ell] \leqslant \eta_\ell \cdot |G_\ell|$. By Markov's inequality,

$\Pr[|F_\ell| > 3L \cdot \eta_\ell \cdot |G_\ell|] \leqslant \frac{1}{3L}$. Applying the union bound over the L heavy groups, we have that with probability at least $2/3$,

$$|F_\ell| \leqslant 3L \cdot \eta_\ell \cdot |G_\ell| = \frac{1}{2Le^{Q_\ell}} \cdot |G_\ell|, \quad \forall \ell \in \{0, 1, \ldots, L-1\},$$

which proves the first part of Theorem 10.

Case 2: Light Group For the light group, note that $\log V_i \leqslant Q_L$. For this group, the MPF solution may not provide any non-trivial guarantee on the utility to the agents. Since the expected utility can now be small, we have to allow additive approximation as well. Recall that F_L consists of agents in G_L for whom $\widehat{U}_i < (1-3\gamma) \cdot U_i^*$ *as well as* $\widehat{U}_i < U_i^* - 4Q_L/\gamma^4$. We again consider two cases.

Case 1. If $U_i^* \leqslant \frac{4}{\gamma^4} Q_L$, then $\widehat{U}_i \geqslant U_i^* - \frac{4Q_L}{\gamma^4}$ trivially.

Case 2. Otherwise, $U_i^* \geqslant \frac{4Q_L}{\gamma^4}$, and using Lemma 8, we have:

$$\Pr\left[\widehat{U}_i < (1-3\gamma)U_i^*\right] \leqslant \Pr\left[\frac{\widehat{U}_i}{1-\gamma} < (1-2\gamma)U_i^*\right] \leqslant e^{-\frac{\gamma^3}{4}U_i^*} \leqslant e^{-\frac{\gamma^3}{4}\cdot\frac{4Q_L}{\gamma^4}} \leqslant \frac{1}{4e^{Q_L}}.$$

It is easy to check that the final transition holds because $\gamma < 1$ is a constant and $Q_L = \omega(1)$.

Note that none of the agents in F_L are in Case 1. Hence, by Markov's inequality, we again have:

$$\Pr\left[|F_L| \geqslant \frac{1}{2e^{Q_L}} \cdot |G_L|\right] \leqslant \frac{1}{2},$$

which proves the second part of Theorem 10.

□

Approximate Core. We showed that with probability at least $1/6$, our algorithm returns a solution that satisfies conditions in both parts of Theorem 10. We now show that such a solution is the desired approximate core solution. The main idea is that when a set of agents deviate, the fraction of agents in a group G_ℓ that are in F_ℓ is small enough such that even if they receive their maximum possible utility, which is e^{Q_ℓ}, their scaled down utility is at most a constant.

Theorem 11. *For every coalition S and every possible outcome $\mathbf{h}$, there exists an agent $i \in S$ s.t.*

$$\frac{|S|}{n} \cdot u_i(\mathbf{h}) \leqslant (1+8\gamma) \cdot \widehat{U}_i + \frac{5Q_L}{\gamma^4}.$$

Proof. Let $W = N \setminus \cup_{\ell=0}^{L} F_\ell$. In other words, W is the set of agents who either receive a good multiplicative approximation to their expected utility in the core (for the heavy groups), or a good additive approximation to their expected utility in the core (for the light group). In particular, for every $i \in W$, we have $\widehat{U}_i \geqslant \min\left((1-3\gamma) \cdot U_i^*, U_i^* - \frac{4Q_L}{\gamma^4}\right)$, which implies

$$U_i^* \leqslant \frac{1}{1-3\gamma} \cdot \widehat{U}_i + \frac{4Q_L}{\gamma^4}. \tag{3.14}$$

Consider a set of agents S that may want to deviate, and let $\mathbf{h}$ be any (even fractional) outcome. There are two cases:

Case 1. Suppose $|S \cap W| \geqslant (1-\gamma) \cdot |S|$. Then, due to the fractional core optimality condition (see Section 3.2.2), we have:

$$\sum_{i \in S \cap W} \frac{U_i(\mathbf{h})}{U_i^*} \leqslant n.$$

Note that in polynomial time, Theorem 8 only finds an approximate solution whose utilities $\{\tilde{U}_i\}$ with $\sum_{i \in S \cap W} \frac{U_i(\mathbf{h})}{\tilde{U}_i} \leqslant n(1+\eta)$ for small $\eta > 0$. It is easy to check this does not alter the rest of the proof and adds a small multiplicative factor of $(1+\eta)$

to the final approximation bound. We ignore this factor for simplicity and simply assume $\{U_i^*\}$ are the optimal MNW utilities. The above implies $\frac{|S|}{n} \cdot \sum_{i \in S \cap W} \frac{u_i(\mathbf{h})}{U_i^*} \leqslant |S| \leqslant \frac{1}{1-\gamma} \cdot |S \cap W|$.

Therefore, there exists an agent $i \in S \cap W$ such that

$$\frac{|S|}{n} \cdot u_i(\mathbf{h}) \leqslant \frac{1}{1-\gamma} \cdot U_i^* \leqslant \frac{1}{1-\gamma} \cdot \left(\frac{1}{1-3\gamma} \cdot \widehat{U}_i + \frac{4Q_L}{\gamma^4} \right),$$

where the last transition is due to Equation (3.14) and the fact that $i \in W$. Finally, it is easy to check that for $\gamma = \delta/8 \leqslant 1/8$, we have $\frac{1}{(1-\gamma)\cdot(1-3\gamma)} \leqslant 1+8\gamma$ and $4/(1-\gamma) \leqslant 5$, which yields:

$$\frac{|S|}{n} \cdot u_i(\mathbf{h}) \leqslant (1+8\gamma) \cdot \widehat{U}_i + \frac{5Q_L}{\gamma^4}. \tag{3.15}$$

Case 2. Otherwise, $|S \backslash W| \geqslant \gamma |S|$. In this case, we want to show that there exists an agent $i \in S \backslash W$ such that $(|S|/n) \cdot u_i(\mathbf{h}) \leqslant 1/\gamma$. Because $\widehat{U}_i \geqslant 0$ and $Q_L = \omega(1)$, such an agent will also satisfy Equation (3.15). We show this by taking two sub-cases.

First, suppose the light group satisfies $|S \cap F_L| \geqslant \frac{\gamma}{2}|S|$. Then: $|S| \leqslant \frac{2}{\gamma} \cdot |S \cap F_L| \leqslant \frac{2}{\gamma} \cdot |F_L|$. Thus, for any agent $i \in F_L$, we have

$$\frac{|S|}{n} \cdot u_i(\mathbf{h}) \leqslant \frac{2}{\gamma n} \cdot |F_L| \cdot V_i \leqslant \frac{2}{\gamma n} \cdot \frac{|G_L|}{2e^{Q_L}} \cdot V_i \leqslant \frac{1}{\gamma}.$$

Here, the second transition follows from Theorem 10. To see why the third transition holds, note that $|G_L| \leqslant n$, and that $\log V_i \leqslant Q_L$ because $i \in G_L$.

Similarly, in the other sub-case, suppose $|S \cap F_L| \leqslant \frac{\gamma}{2}|S|$. Then, there exists a heavy group $\ell \in \{0, 1, \ldots, L-1\}$ such that $|S \cap F_\ell| \geqslant \frac{\gamma}{2L}|S|$. This means $|S| \leqslant \frac{2L}{\gamma} \cdot |S \cap F_\ell| \leqslant \frac{2L}{\gamma} \cdot |F_\ell|$.

Again, for an arbitrary agent $i \in F_\ell$, we have:

$$\frac{|S|}{n} \cdot u_i(\mathbf{h}) \leqslant \frac{2L}{\gamma n} \cdot |F_\ell| \cdot V_i \leqslant \frac{2L}{\gamma n} \cdot \frac{|G_\ell|}{2Le^{Q_\ell}} \cdot V_i \leqslant \frac{1}{\gamma}.$$

Once again, the third transition follows from Theorem 10, and the fourth transition holds because $|G_\ell| \leqslant n$ and $\log V_i \leqslant Q_\ell$ as $i \in G_\ell$. Putting everything together, the theorem follows. $\square$

Since $\gamma = \frac{\delta}{8}$ and $Q_L = \Theta\left(\log \frac{R \log^* V_{max}}{\gamma}\right)$, Theorem 11 implies $\frac{|S|}{n} \cdot u_i(\mathbf{h}) \leqslant (1+\delta) \cdot \widehat{U}_i + \alpha^*$, where $\alpha^* = O\left(\frac{1}{\delta^4} \cdot \log\left(\frac{R \cdot \log^* V_{\max}}{\delta}\right)\right)$. The existence of such an agent implies that a solution satisfying Theorem 10 is a (δ, α)-core solution for every $\alpha > \alpha^*$, which completes the proof of Theorem 9.

3.6 Conclusion and Open Directions

We considered the problem of fairly allocating public goods. We argued that the *core*, which is a generalization of proportionality and Pareto efficiency, and approximations of the core provide reasonable fairness guarantees in this context. Given that no integral outcome may be in the core, we presented efficient algorithms to produce integral outcomes that are constant or near-constant approximations of the core, thereby also establishing the non-trivial existence of such outcomes. Note that our algorithms for matroid and matching constraints that globally optimize the smooth Nash welfare objective achieve *exact* rather than approximate Pareto efficiency, in addition to an approximation of the core. An interesting question is whether the same guarantee can be provided (regardless of computation time) for general packing constraints.

Another natural question following our work is to tighten our upper bounds, or to establish matching lower bounds. For instance, we show the existence of a $(0, 2)$-core outcome for matroid constraints (Theorem 5), but our lower bound only shows that a $(0, 1-\epsilon)$-core outcome may not exist. This leaves open the question of whether a $(0, 1)$-core outcome always exists. Existence of $(0, 1)$-core outcome is also an open question for matching constraints. For packing constraints, it is unknown if even a

(δ, α)-core outcome exists for constant $\delta > 0$ and $\alpha = O(1)$. This also remains an open question for the endowment-based notion of core.

At a higher level, we established connections between approximating the core in our multi-agent environment and the problem of finding the optimal (i.e., utility-maximizing) outcome for a *single* agent. For instance, given matching constraints, our algorithm uses the idea of *short* augmenting paths from fast PRAM algorithms. This hints at the possibility of a deeper connection between efficiency results and existence results.

4

Fair Clustering

In this chapter, we consider a very different kind of problem in public decision making: centroid clustering. One can think of the task of clustering a group of data points as a public resource allocation problem where the data points are the agents and the possible centers are the public goods. Now we think of the agents as having unit demand utilities (or costs, more precisely), in the sense that they only care about their favorite center, in the sense of the center that is closest to them in some metric (this is a standard model for facility location problems, for example). We will continue with this basic approach of modeling the costs of public decisions in metric spaces in the subsequent chapters. We adapt our notion of core as a fairness guarantee for clustering so that we ensure that we do not ignore large cohesive groups of agents for the sake of a global objective. Due to the different model of valuations, we term our fairness notion proportionality instead of core in this context, but it is in fact an application of the core under a particular utility model. Note that this is a different notion of proportionality than is employed for resource allocation problems. Our work introduces this form of proportionality as a novel solution concept in clustering and systematically studies its computation.

Acknowledgements. These results are from joint work with Xingyu Chen, Charles Lyu, and Kamesh Munagala, and is under review at the 36th International Conference on Machine Learning (ICML) at the time of this writing [45].

4.1 Introduction

The data points in machine learning are often real human beings. There is legitimate concern that traditional machine learning algorithms that are blind to this fact may inadvertently exacerbate problems of bias and injustice in society [102]. Motivated by concerns ranging from the granting of bail in the legal system to the quality of recommender systems, researchers have devoted considerable effort to developing fair algorithms for the canonical supervised learning tasks of classification and regression [61, 107, 87, 106, 161, 56, 137, 162, 105, 83, 91].

We extend this work to a canonical problem in unsupervised learning: centroid clustering. In centroid clustering, we want to partition data into k clusters by choosing k "centers" and then matching points to one of the centers. This is a classic context for clustering work [84, 43, 13], and is perhaps best known as the setting for the celebrated k-means heuristic (independently discovered many times, see [94] for a brief history). We provide a novel group based notion of fairness as proportionality, inspired by recent related work on the fair allocation of public resources [16, 53, 66, 74]. We suppose that data points represent the individuals to whom we wish to be fair, and that these agents prefer to be clustered accurately (that is, they prefer their cluster center to be representative of their features). A solution is fair if it respects the entitlements of groups of agents, where we assume that a subset of agents is entitled to choose a center for themselves if they constitute a sufficiently large fraction of the population with respect to the total number of clusters (*e.g.*, $1/k$ of the population, if we are clustering into k groups). The guarantee must hold for *all* subsets of sufficient size, and therefore does not hinge on any particular a priori knowledge about which points should be protected. This is in line with other recent observations that information about which individuals should be protected may not be available in practice [91].

Consider an example of the problem of proportional clustering. We want to assign incoming college students to one of k themed communities. Each student fills out a survey providing information about their preferred community, *e.g.*, quiet hours, community activities, etc. We choose the k themes in light of this information, and we want to do so in a way that is fair to the incoming students. The idea is that any $1/k$ fraction of the incoming students should be allowed to form their own community if they prefer. A proportional solution is one for which there is no such large group that prefers to choose their own theme over the theme or themes to which we assigned them. In that sense, one can also see proportionality as a no justified complaint guarantee, a stability condition, and a fairness condition.

For an example in a different context, consider medical data. Each data point may represent an individual, and clustering is a basic data analysis tool that can influence further research and development. One approach to fairness is to ensure that the clustering obscures protected demographic information by ensuring that all clusters have comparable ratios of all relevant demographics [47]. While this makes sense when clustering is used to generate features for downstream decision making like loan applications, it is less clearly appropriate in an application like medicine, where protected information like sex, race, and age may all be highly medically relevant. Instead, we are inspired by the idea of precision medicine [52], to tailor research, development, and treatment to account for variability in the base population. Toward that end, proportional clustering guarantees that the data analyst will not inadvertently ignore large groups of data while clustering for the sake of some global objective. It does this in a robust fashion that does not require the a priori definition of specific protected groups.

4.1.1 Preliminaries and Definition of Proportionality

We have a set $\mathcal{N}$ of $|\mathcal{N}| = n$ individuals or data points, and a set $\mathcal{M}$ of $|\mathcal{M}| = m$ feasible cluster centers. We will sometimes consider the important special case where $\mathcal{M} = \mathcal{N}$ (i.e., where one is only given a single set of points as input), but most of our results are for the general case where we make no assumption about $\mathcal{M} \cap \mathcal{N}$. For all $i, j \in \mathcal{N} \cup \mathcal{M}$, we have a distance $d(i, j)$ satisfying the triangle inequality. Our task is centroid clustering as treated in the classic k-median, k-means, and k-center problems. We wish to open a set $X \subseteq \mathcal{M}$ of $|X| = k$ centers (assume $|\mathcal{M}| \geqslant k$), and then match all points in $\mathcal{N}$ to their closest center in X. For a particular solution X and agent $i \in \mathcal{N}$, let $D_i(X) = \min_{x \in X} d(i, x)$. In general, a good clustering solution X will have small values of $D_i(X)$, although the aforementioned objectives differ slightly in how they measure this. In particular, the k-median objective is $\sum_{i \in \mathcal{N}} D_i(X)$, the k-means objective is $\sum_{i \in \mathcal{N}} (D_i(X))^2$, and the k-center objective is $\max_{i \in \mathcal{N}} D_i(X)$.[1]

To define proportional clustering, we assume that individuals prefer to be closer to their center in terms of distance (i.e., ensuring that the center is more representative of the point). Any subset of at least $r\lceil \frac{n}{k} \rceil$ individuals is entitled to choose r centers. We call a solution proportional if there does not exist any such sufficiently large set of individuals who, using the number of centers to which they are entitled, could produce a clustering among themselves that is to their mutual benefit in the sense of Pareto dominance. More formally, a *blocking coalition* is a set $S \subseteq \mathcal{N}$ of at least $r\lceil \frac{n}{k} \rceil$ points and a set $Y \subseteq \mathcal{M}$ of at most r centers such that $D_i(Y) < D_i(X)$ for all $i \in S$. It is easy to see that because $D_i(X) = \min_{x \in X} d(i, x)$, this is functionally equivalent to Definition 12; a larger blocking coalition necessarily implies a blocking

[1] We note that this is very similar to the related facility location problem, although in that context one typically thinks of the number of facilities to open (i.e., the value of k) as an optimization term rather than a constraint.

coalition with a single center.

Definition 12. *Let $X \subseteq \mathcal{M}$ with $|X| = k$. $S \subseteq N$ is a blocking coalition against X if $|S| \geqslant \lceil \frac{n}{k} \rceil$ and $\exists y \in \mathcal{M}$ such that $\forall i \in S$, $d(i, y) < D_i(X)$. $X \subseteq \mathcal{N}$ is proportional if there is no blocking coalition against X.*

Equivalently, X is proportional if $\forall S \subseteq \mathcal{N}$ with $|S| \geqslant \lceil \frac{n}{k} \rceil$ and for all $y \in \mathcal{M}$, there exists $i \in S$ with $d(i, y) \geqslant D_i(X)$. It is important to note that this quantification is over *all* subsets of sufficient size, and not a particular partition into groups.

It is instructive to briefly consider an example. In Figure 4.1, $\mathcal{N} = \mathcal{M}$, $k = 2$, and there are 12 individuals, represented by the embedded points. Suppose we want to minimize the k-center objective in the pursuit of fairness: we would then choose the red points. However, this is not a proportional solution because the middle six points constitute half of the points, and would all prefer to be matched to the central blue point. Furthermore, choosing the blue point (and any other center) *is* a proportional solution, because for any arbitrary group of six points and new proposed center, at least one of the six points will be closer to the blue point than the proposed center.

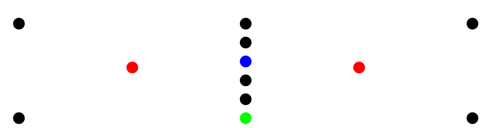


FIGURE 4.1: Proportionality Example

Proportionality has many advantages as a notion of fairness in clustering, beyond the intuitive appeal of groups being entitled to a proportional share of centers. We name a few of these advantages explicitly.

- Proportionality implies (weak) *Pareto optimality*: namely, for any proportional solution X, there does not exist another solution X' such that $D_i(X') < D_i(X)$ for all $i \in \mathcal{N}$.

- Proportionality is *oblivious* in the sense that it does not depend on the definition of sensitive attributes or protected sub-groups.
- Proportionality is *robust* to outliers in the data, since only groups of points of sufficient size are entitled to their own center.
- Proportionality is *scale invariant* in the sense that a multiplicative scaling of all distances does not affect the set of proportional solutions.
- Approximately proportional solutions can be *efficiently computed*, and one can optimize a secondary objective like k-median subject to proportionality *as a constraint*, as we show in Section 4.2 and Section 4.3.
- Proportionality can be *efficiently audited*, in the sense that one does not need to compute the entire pairwise distance matrix in order to check for violations of proportionality, as we show in Section 4.4.

In the worst case, proportionality is incompatible with all three of the classic k-center, k-means, and k-median objectives; *i.e.*, there exist instances for which any proportional solution has an arbitrarily bad approximation to all objectives.

Claim 1. *There exist instances for which any proportional clustering has an unbounded approximation to the optimal k-center, k-means, and k-median objectives.*

Proof. The simplest example to see this has $\mathcal{N} = \mathcal{M}$, $n = 6$, and $k = 3$ (i.e., we want to choose 3 centers from six individuals, all of which are possible cluster centers). There are two points at position a, two at position b, and one each at positions c and d. The pairwise distances are given in the following matrix.

	a	b	c	d
a	0	1	∞	∞
b	1	0	∞	∞
c	∞	∞	0	∞
d	∞	∞	∞	0

Because $n = 6$ and $k = 3$, and there are $6/3 = 2$ points at a and b, any proportional solution must include a and b, else the two points at the position not included in the solution could form a blocking coalition. Therefore, any proportional solution can only allocate one center to either c or d. The point at the position among c and d that is not allocated a center will have an $D_i(X) = \infty$. By contrast, the optimal solution on any of the three objectives is to instead choose a center at c at d, and without loss of generality, at a, for a total cost of 1 under all objectives. □

Furthermore, as we show in Section 4.2 and observe empirically in Section 4.5, proportional solutions may not always exist. We therefore consider the natural approximate notion of proportionality that relaxes the Pareto dominance condition by a multiplicative factor.

Definition 13. *$X \subseteq \mathcal{M}$ with $|X| = k$ is ρ-approximate proportional (hereafter ρ-proportional) if $\forall S \subseteq \mathcal{N}$ with $|S| \geqslant \lceil \frac{n}{k} \rceil$ and for all $y \in \mathcal{M}$, there exists $i \in S$ with $\rho \cdot d(i, y) \geqslant D_i(X)$.*

Recall the example in Figure 4.1. Although choosing the red points is not a proportional solution, it is an approximate proportional solution. To see this, suppose the middle six agents wish to deviate to the blue point as before. The green point would decrease its distance to a center by deviating, but not by more than a constant factor, say 3, so the red points would constitute a 3-proportional solution.

4.1.2 Results and Outline

In Section 4.2 we show that proportional solutions may not always exist. In fact, one cannot get better than a 2-proportional solution in the worst case. In contrast, we give a greedy algorithm (Algorithm 1) and prove Theorem 12: The algorithm yields a $\left(1 + \sqrt{2}\right)$-proportional solution in the worst case.

In Section 4.3, we treat proportionality as a constraint and seek to optimize the k-median objective subject to that constraint. We show how to write approximate proportionality as m linear constraints. Incorporating this into the standard linear programming relaxation of the k-median problem, we show how to use the rounding from [43] to find an $O(1)$-proportional solution that is an $O(1)$-approximation to the k-median objective of the optimal proportional solution.

In Section 4.4, we show that proportionality is approximately preserved if we take a random sample of the data points of size $\tilde{O}(k^3)$, where the $\tilde{O}$ hides low order terms. This immediately implies that for constant k, we can check if a given clustering is proportional as well as compute approximately proportional solutions in near linear time, comparable to the time taken to run the classic k-means heuristic.

In Section 4.5, we provide a local search heuristic that efficiently searches for a proportional clustering. Our heuristic is able to consistently find nearly proportional solutions in practice. We test our heuristic and Algorithm 1 empirically against the celebrated k-means heuristic in order to understand the tradeoff between proportionality and the k-means objective. We find that the tradeoff is highly data dependent: Though these objectives are compatible on some datasets, there exist others on which these objectives are in conflict.

4.1.3 Related Work

Unsupervised Learning. Metric clustering is a well studied problem. There are constant approximation polynomial time algorithms for both the k-median [96, 43, 13, 125] and k-center objective [84]. Proportionality is a constraint on the *centers* as opposed to the data points; this makes it difficult to adapt standard algorithmic approaches for k-medians and k-means such as local search [13], primal-dual [96], and greedy dual fitting [95]. For instance, our greedy algorithm in Section 4.2 grows balls around potential centers, which is very different from how balls are grown in the primal-dual schema [96, 125]. Somewhat surprisingly, in Section 4.2 we show that for the problem of minimizing the k-median objective subject to proportionality as a constraint, we can extend the linear program rounding technique of [43] to get a constant approximation algorithm. However, the additional constraints we add in the linear program formulation render the primal-dual and other methods inapplicable.

In [47], the authors consider fair clustering in terms of balance: There are red and blue points, and a balanced solution has roughly the same ratio of blue to red points in every cluster as in the overall population. The authors are motivated to extract features that cannot discriminate between status in different groups. This ensures that subsequent regression or classification on these features will be fair between these groups. In contrast, we assume that our data points prefer to be accurately clustered, and that an unfair solution provides accurate clusters for some groups, while giving other large groups low quality clusters. Finally, we note that there is a line of work in fair unsupervised learning concerned with constructing word embeddings that avoid bias [30, 40], but these problems seem orthogonal to our concerns in clustering.

Supervised Learning. In the extensive work on fair supervised learning [61, 107, 106, 162, 161], the standard model has a set of *protected agents* given as input to a classification algorithm which must classify agents into a positive and negative group. Most of these notions of fairness do not apply in any natural way to unsupervised learning problems. Our work further differs from the supervised learning literature in that we do not assume information about which agents are to be protected. Instead, we provide a fairness guarantee to arbitrary groups of agents, including protected groups even if we do not know their identity, similar to the ideas considered in [105] and [91].

Fair Resource Allocation. Our notion of proportionality is derived from the notion of core in economics [150, 71]. The core has been adapted as a natural generalization of the idea of fairness as proportionality: that every agent should get a proportional share (*e.g.*, $1/n$ where there are n agents) of resources. The core generalizes this notion in that it considers proportionality with respect to groups of agents rather than just individuals, which is important when resources are public and can be shared [63, 53, 16]. In clustering, the "resources" are the centers themselves, and they are shared by definition. Much of this prior work in resource allocation requires randomization (or lotteries) to achieve a core outcome, but in our work, we find that the metric property in clustering allows us to make progress without resorting to randomized outcomes.

4.2 Existence and Computation of Proportional Solutions

We begin with a negative result: in the worst case, there may not be an exact proportional solution. Claim 2 is stated for arbitrary $\mathcal{N}$ and $\mathcal{M}$, but the impossibility remains even when $\mathcal{N} = \mathcal{M}$ via a similar argument.

Claim 2. *For all $\rho < 2$, a ρ-proportional solution is not guaranteed to exist.*

Proof. Consider the following instance with $\mathcal{N} = \{a_1, a_2, \ldots, a_6\}$, $\mathcal{M} = \{x_1, x_2, \ldots, x_6\}$ and $k = 3$. Distances are specified in the following table. For distances not explicitly specified in the table (e.g., $d(x_1, x_2)$), the distance is the shortest path distance on the weighted graph defined by the table weights (so, $d(x_1, x_2) = d(x_1, a_3) + d(a_3, x_2) = 3$). It is important to note that the resulting metric satisfies the triangle inequality overall.

	x_1	x_2	x_3	x_4	x_5	x_6
a_1	4	1	2	∞	∞	∞
a_2	2	4	1	∞	∞	∞
a_3	1	2	4	∞	∞	∞
a_4	∞	∞	∞	4	1	2
a_5	∞	∞	∞	2	4	1
a_6	∞	∞	∞	1	2	4

Notice that the data is separate into two areas. Since $k = 3$, in a feasible solution, we only open one center in one of these two areas. Without loss of generality, suppose that we open exactly one center among $\{x_1, x_2, x_3\}$. The instance is symmetric, so again suppose without loss of generality that we open x_1. Then consider the individuals in $\{a_1, a_2\}$. This coalition is of size $\lceil \frac{n}{k} \rceil = 2$, and both individuals would reduce their distance by a factor of 2 by switching to x_3. Thus, any solution is only 2-proportional. $\square$

Computing a $(1 + \sqrt{2})$*-Approximate Proportional Clustering.* Claim 2 establishes that we should focus our attention on designing an efficient approximation algorithm. We give a simple and efficient algorithm that achieves a $(1 + \sqrt{2})$-proportional solution, very close to the existential lower bound of 2. For notational ease, let $B(x, \delta) = \{i \in \mathcal{N} : d(i, x) \leqslant \delta\}$. That is, $B(x, \delta)$ is the ball (defined on $\mathcal{N}$) of distance δ about center x. For simplicity of exposition, we present Algorithm 1 as a *continuous* algorithm where a δ parameter is smoothly increasing. The algorithm can be easily discretized using priority queues.

In essence, Algorithm 1 grows balls continuously around the centers, and when the ball around a center has "captured" $\lceil \frac{n}{k} \rceil$ points, we greedily open that center and disregard all of the captured points. Open centers continue to greedily capture points as their balls continue to expand. Though [96, 125] similarly expand balls about points to compute approximately optimal solutions to the k-median problem, there is a crucial difference: They grow balls around data points rather than centers.

Algorithm 1 Greedy Capture

$\delta \leftarrow 0$; $X \leftarrow \varnothing$; $N \leftarrow \mathcal{N}$
while $N \neq \varnothing$ **do**
  Smoothly increase δ
  while $\exists x \in X$ s.t. $|B(x, \delta) \cap N| \geqslant 1$ **do**
    $N \leftarrow N \backslash B(x, \delta)$
  while $\exists x \in (\mathcal{M} \backslash X)$ s.t. $|B(x, \delta) \cap N| \geqslant \lceil \frac{n}{k} \rceil$ **do**
    $X \leftarrow X \cup \{x\}$
    $N \leftarrow N \backslash B(x, \delta)$
return X

Theorem 12. *Algorithm 1 yields a* $(1+\sqrt{2})$*-proportional clustering, and there exists an instance for which this bound is tight.*

Proof. Let X be the solution computed by Algorithm 1. First note that X uses at most k centers, since it only opens a center when $\lceil \frac{n}{k} \rceil$ *unmatched* points are absorbed by the ball around that center, and this can happen at most k times. Now, suppose for a contradiction that X is not a $(1+\sqrt{2})$-proportional clustering. Then there exists $S \subseteq \mathcal{N}$ with $|S| \geqslant \lceil \frac{n}{k} \rceil$ and $y \in \mathcal{M}$ such that

$$\forall i \in S,\ (1+\sqrt{2}) \cdot d(i, y) < D_i(X). \tag{4.1}$$

Let r_y be the distance of the farthest agent from y in S, that is, $r_y := \max_{i \in S} d(i, y)$, and call this agent i^*. There are two cases. In the first case, $B(x, r_y) \cap S = \varnothing$ for all $x \in X$. This immediately yields a contradiction, because it implies that Algorithm 1 would have opened y. In particular, note that $S \subseteq B(y, r_y)$, so if $S \cap B(x, r_y) = \varnothing$ for all $x \in X$, then $B(y, r_y)$ would have had at least $\lceil \frac{n}{k} \rceil$ unmatched points.

In the second case, $\exists x \in X$ and $\exists i \in N$ such that $i \in B(x, r_y) \cap S$. This case is drawn below in Figure 4.2. By the triangle inequality, $d(x, y) \leqslant d(i, x) + d(i, y)$. Therefore, $d(i^*, x) \leqslant r_y + d(i, x) + d(i, y)$. Also, $d(i, x) \leqslant r_y$, since $i \in B(x, r_y)$. Consider the minimum multiplicative improvement of i and i^*:

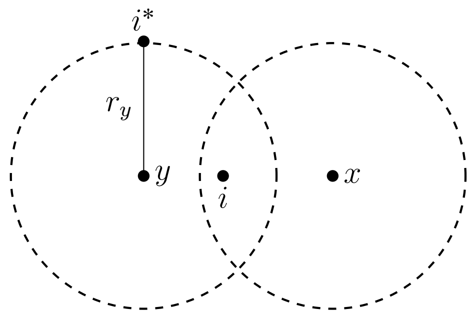


FIGURE 4.2: Diagram for Proof of Theorem 12

$$\begin{aligned}
&\min\left(\frac{d(i,x)}{d(i,y)},\ \frac{d(i^*,x)}{d(i^*,y)}\right)\\
&\leqslant \min\left(\frac{d(i,x)}{d(i,y)},\ \frac{r_y + d(i,x) + d(i,y)}{r_y}\right)\\
&\leqslant \min\left(\frac{r_y}{d(i,y)},\ 2 + \frac{d(i,y)}{r_y}\right)\\
&\leqslant \max_{z \geqslant 0}\left(\min\left(z,\ 2 + 1/z\right)\right) = 1 + \sqrt{2}
\end{aligned}$$

which violates equation 4.1. To see that there exists an instance for which Algorithm 1 yields exactly this bound, consider the following instance with $\mathcal{N} = \{a_1, a_2, \ldots, a_6\}$, $\mathcal{M} = \{x_1, x_2, x_3, x_4\}$ and $k = 3$. Distances are specified in the following table, where $\epsilon > 0$ is some small constant.

	x_1	x_2	x_3	x_4
a_1	1	$1+\sqrt{2}$	∞	∞
a_2	$\sqrt{2}-1$	$1-\epsilon$	∞	∞
a_3	$1+\sqrt{2}$	$1-\epsilon$	∞	∞
a_4	∞	∞	1	$1+\sqrt{2}$
a_5	∞	∞	$\sqrt{2}-1$	$1-\epsilon$
a_6	∞	∞	$1+\sqrt{2}$	$1-\epsilon$

The distances satisfy the triangle inequality. Note that Algorithm 1 will open x_2 and x_4. The coalition $\{a_1, a_2\}$ can each reduce their distance by a multiplicative factor approaching $1+\sqrt{2}$ as $\epsilon \to 0$ by deviating to x_1. □

Theorem 12 establishes that constant approximate proportional clusterings always exist, and can be computed efficiently. However, Algorithm 1 may prove unsatisfactory, in the sense that it may find an approximate proportional clustering with poor global objective (e.g., k-median or k-means), even when exact proportional clusterings with good global objectives exist. For example, suppose $k = 2$ and there are two easily defined clusters, one containing 40% of the data and the other containing 60% of the data. It is possible that Algorithm 1 will only open centers inside of the larger cluster. This is proportional, but clearly bad on the k-median and k-means objectives, even with respect to proportional solutions (notice that the "correct" clustering of such an example is still proportional). In Section 4.3, we show how to address this concern by optimizing the k-median objective subject to proportionality as a constraint. Later, in Section 4.5, we empirically study the tradeoff between the k-means objective and proportionality on real data.

4.3 Proportionality as a Constraint

We consider the k-median and k-means objectives to be reasonable measures of the global quality of a solution. We see minimizing the k-center objective more as a competing notion of fairness, and so we focus on optimizing the k-median objective subject to proportionality.[2]

Minimizing the k-median objective without proportionality is a well studied problem in approximation algorithms, and several constant approximations are known [43, 13, 125]. Most of this work is in the model where $\mathcal{N} \subseteq \mathcal{M}$, and we follow suit in this section. We show the following.

Theorem 13. *Suppose there is a ρ-proportional clustering with k-median objective*

[2] A constant approximation algorithm for minimizing the k-median objective immediately implies a constant approximation algorithm for minimizing the k-means objective by running the algorithm on the squared distances [125].

$$\text{Minimize} \quad \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{M}} d(i,j) z_{ij} \tag{4.2}$$

$$\text{Subject to} \quad \sum_{j \in \mathcal{M}} z_{ij} = 1 \qquad \forall i \in \mathcal{N} \tag{4.3}$$

$$z_{ij} \leqslant y_j \qquad \forall j \in \mathcal{M}, \forall i \in \mathcal{N} \tag{4.4}$$

$$\sum_{j \in \mathcal{M}} y_j \leqslant k \tag{4.5}$$

$$\sum_{j' \in B(j, \gamma R_j)} y_{j'} \geqslant 1 \qquad \forall j \in \mathcal{M} \tag{4.6}$$

$$z_{ij}, y_j \in [0,1] \qquad \forall j \in \mathcal{M}, \forall i \in \mathcal{N} \tag{4.7}$$

FIGURE 4.3: Proportional k-median Linear Program

c. In polynomial time in m and n, we can compute an $O(\rho)$-proportional clustering with k-median objective at most $8c$.

In particular, if there is a proportional clustering of minimum k-median objective, we can compute a constant approximate proportional clustering with k-median objective within a constant factor. In the remainder of this section, we will sketch the proof of Theorem 13. We begin with the standard linear programming relaxation of the k-median minimization problem, and then add a constraint to encode proportionality. The final linear program is shown in Figure 4.3. Let $B(x, \delta) = \{i \in \mathcal{M} : \ d(i,x) \leqslant \delta\}$. Let R_j be the minimum value such that $|B(j, R_j)| \geqslant \lceil \frac{n}{k} \rceil$. In other words, R_j is the distance of the $\lceil \frac{n}{k} \rceil$ farthest point in $\mathcal{N}$ from j.

In this LP, z_{ij} is an indicator variable equal to 1 if $i \in \mathcal{N}$ is matched to $j \in \mathcal{M}$. y_j is an indicator variable equal to 1 if $j \in X$, that is, if we want to use center $j \in \mathcal{M}$ in our clustering. Objective 4.2 is simply the k-median objective. Constraint 4.3 requires that every point be matched, and constraint 4.4 only allows a point to be matched to an open center. Constraint 4.5 allows at most k centers to be opened, and constraint 4.7 relaxes the indicator variables to be real values between 0 and 1.

Constraint 4.6 is the new constraint that we introduce. Our crucial lemma argues that constraint 4.6 approximately encodes proportionality.

Lemma 9. *Let X be a clustering, and let $\gamma \geqslant 1$. If $\forall j \in \mathcal{M}$ there exists some $x \in X$ such that $d(j, x) \leqslant \gamma R_j$, then X is $(1+\gamma)$-proportional. If X is γ-proportional, then $\forall j \in \mathcal{M}$ there exists some $x \in X$ such that $d(j, x) \leqslant (1+\gamma) R_j$.*

Proof. Suppose that $\forall j \in \mathcal{M}$ there exists some $x \in X$ such that $d(j, x) \leqslant \gamma R_j$. Suppose for a contradiction that X is not $(1+\gamma)$-proportional. Then there exists $S \subseteq \mathcal{N}$ with $|S| \geqslant \lceil \frac{n}{k} \rceil$ and $j \in \mathcal{M}$ such that $\forall i \in S$, $(1+\gamma) \cdot d(i, j) < D_i(X)$. By assumption, $\exists x \in X$ such that $d(j, x) \leqslant \gamma R_j$, so by the triangle inequality $D_i(X) \leqslant d(i, j) + d(j, x) \leqslant d(i, j) + \gamma R_j$. Therefore, $\forall i \in S$, $\gamma \cdot d(i, j) < D_i(X) - d(i, j) \leqslant \gamma R_j$. However, by definition of R_j, since $|S| = \lceil \frac{n}{k} \rceil$, there must exist some $i \in S$ such that $d(i, j) \geqslant R_j$.

Suppose that X is γ-proportional. Let $j \in \mathcal{M}$. Consider the set S of the closest $\lceil \frac{n}{k} \rceil$ points in $\mathcal{N}$ to j. By definition of proportionality $\exists i \in S$ and $x \in X$ such that $\gamma d(i, j) \geqslant d(i, x)$. Therefore, by the triangle inequality, $d(j, x) \leqslant d(i, j) + d(i, x) \leqslant (1+\gamma) d(i, j)$. By definition of S, $d(i, j) \leqslant R_j$, so there exists $x \in X$ such that $d(j, x) \leqslant (1+\gamma) R_j$. □

Now, suppose there is a ρ-proportional clustering X with k-median objective c. Then we write the linear program shown in Figure 4.3 with $\gamma = \rho + 1$ in constraint 4.6. Lemma 9 guarantees that X is feasible for the resulting linear program, so the fractional solution has k-median objective at most c. We then round the resulting fractional solution. In [43], the authors give a rounding algorithm for the the linear program in Figure 4.3 without Constraint 4.6. We show that a slight modification to this rounding algorithm also preserves Constraint 4.6 to a constant approximation.

Lemma 10. *(Proved in Appendix 8.3) Let* $\{y_j\}, \{z_{ij}\}$ *be a fractional solution to the linear program in Figure 4.3. Then there is an integer solution* $\{\hat{y_j}\}, \{\hat{z_{ij}}\}$ *that is an* 8*-approximation to the objective, and that opens* k *centers. Furthermore, for all* $j \in \mathcal{M}$, $\sum_{j' \in B(j, 27\gamma R_j)} \hat{y_{j'}} \geqslant 1$.

Given Lemma 10, applying Lemma 9 again implies that the result of the rounding is $(27(1+\rho)+1)$-proportional, since we set $\gamma = 1 + \rho$. Since the k-median objective of the fractional solution is at most c, the fact that the k-median objective of the rounded solution is at most $8c$ follows directly from the proof from [43]. We note that the constant factor of 27 can be improved to 13 in the special case where $\mathcal{N} = \mathcal{M}$.

It is surprising that such a black box application is successful. The ostensibly similar primal-dual approach of [96] does not appear amenable to the added constraint of proportionality (in particular, the reduction to facility location from [96] is no longer straightforward).

4.4 Sampling for Linear-Time Implementations and Auditing

In this section, we study proportionality under uniform random sampling (i.e., draw $|N|$ individuals i.i.d. from the uniform distribution on $\mathcal{N}$). In particular, we show that proportionality is well preserved under random sampling. This allows us to design an efficient implementation of Algorithm 1 and other heuristics discussed in Section 4.5, and introduce an efficient algorithm for auditing proportionality. We first present the general property and then demonstrate its various applications.

Proportionality Under Random Sampling. For any $X \subseteq \mathcal{M}$ of size k and center $y \in \mathcal{M}$, define $R(\mathcal{N}, X, y) = \{i \in \mathcal{N} : D_i(X) > \rho \cdot d(i, y)\}$. Note that solution X is not ρ-proportional with respect to $\mathcal{N}$ if and only if there is some $y \in \mathcal{M}$ such that $\frac{|R(\mathcal{N},X,y)|}{|\mathcal{N}|} \geqslant \frac{1}{k}$. A random sample approximately preserves this fraction for all solutions X and deviating centers y. The following theorem is a straightforward consequence of Hoeffding's inequality. The important idea in the proof is that we take a union bound over all possible solutions and deviations, and there are only $k\binom{m}{k}$ such combinations.

Theorem 14. *Given $\mathcal{N}$, $\mathcal{M}$ and parameter $\rho \geqslant 1$, fix parameters $\epsilon, \delta \in [0,1]$. Let $N \subseteq \mathcal{N}$ of size $\Omega\left(\frac{k^3}{\epsilon^2} \log \frac{m}{\delta}\right)$ be chosen uniformly at random. Then, with probability at least $1-\delta$, the following holds for all (X, y):*

$$\left|\frac{|R(N,X,y)|}{|N|} - \frac{|R(\mathcal{N},X,y)|}{|\mathcal{N}|}\right| \leqslant \frac{\epsilon}{k}$$

Proof. Recall that N is a random sample of $\mathcal{N}$. Hoeffding's inequality implies that for any fixed (X, y), a sample of size $|N| = O\left(\frac{1}{\hat{\epsilon}^2} \log \frac{1}{\hat{\delta}}\right)$ is sufficient to achieve

$$\left|\frac{|R(N,X,y)|}{|N|} - \frac{|R(\mathcal{N},X,y)|}{|\mathcal{N}|}\right| \leqslant \hat{\epsilon}$$

with probability at least $1-\hat{\delta}$. Note that there are $q = m\binom{m}{k}$ possible choices of (X, y) over which we take the union bound. Setting $\delta = \frac{\hat{\delta}}{q}$, and $\epsilon = \frac{\hat{\epsilon}}{k}$ is sufficient for the union bound to yield the theorem statement. □

In order to apply the above theorem, we say that a solution X is ρ-proportional to $(1+\epsilon)$-deviations if for all $y \in \mathcal{M}$ and for all $S \subseteq \mathcal{N}$ where $|S| \geqslant (1+\epsilon)\frac{n}{k}$, there exists some $i \in S$ such that $\rho \cdot d(i,y) \geqslant D_i(X)$. Note that if X is ρ-proportional to 1-deviations, it is simply ρ-proportional. We immediately have the following:

Corollary 1. *Let $N \subset \mathcal{N}$ be a uniform random sample of size $|N| = \Omega\left(\frac{k^3}{\epsilon^2} \ln \frac{m}{\delta}\right)$. Suppose $X \subseteq M$ with $|X| = k$ is ρ-proportional with respect to N. Then with probability at least $1-\delta$, X is ρ-proportional to $(1+\epsilon)$-deviations with respect to $\mathcal{N}$.*

Linear Time Implementation. We now consider how to take advantage of Theorem 14 to optimize Algorithm 1. To state running times simply, we use the convention that $f(n)$ is $\tilde{O}(g(n))$ if $f(n)$ is $O(g(n))$ up to poly-logarithmic factors. First, note that Algorithm 1 takes $\tilde{O}(mn)$ time, which is quadratic in input size. A corollary of Theorem 14 is that we can approximately implement Algorithm 1 in nearly linear time, comparable to the running time of the standard k-means heuristic.

Corollary 2. *Algorithm 1, when run on $\mathcal{M}$ and a random sample $N \subseteq \mathcal{N}$ of size $|N| = \tilde{\Theta}\left(\frac{k^3}{\epsilon^2}\right)$, provides a solution that is $(1+\sqrt{2})$-proportional to $(1+\epsilon)$-deviations with high probability in $\tilde{O}\left(\frac{k^3}{\epsilon^2}m\right)$ time.*

Efficient Auditing. Alternatively, one might still want to run one's favorite clustering algorithm, e.g., the k-means heuristic, and ask whether the solution produced happens to be proportional. We call this the *Audit Problem.* Given $\mathcal{N}, \mathcal{M}$, and $X \subset \mathcal{N}$ with $|X| \leqslant k$, find the minimum value of ρ such that X is ρ-proportional.

It is not too hard to see that one can solve the Audit Problem exactly in $O((k+m)n)$ time by computing for each $y \in \mathcal{M}$, the quantity ρ_y, the $\lceil \frac{n}{k} \rceil$ largest value of $\frac{D_i(X)}{d(i,y)}$. We subsequently find the y that maximizes ρ_y. Again, this takes quadratic time, which can be worse than the time taken to find the clustering itself.

Consider a slightly relaxed (ϵ, δ)-Audit Problem where we are asked to find the minimum value of ρ such that X is ρ-proportional to $(1+\epsilon)$-deviations with probability at least $1-\delta$. This problem can be efficiently solved by using a random sample $N \subseteq \mathcal{N}$ of points to conduct the audit.

Corollary 3. *The (ϵ, δ)-Audit Problem can be solved in $\tilde{O}\left((k+m)\frac{k^3}{\epsilon^2}\right)$ time.*

4.5 Implementations and Empirical Results

In this section, we study proportionality on real data taken from the UCI Machine Learning Repository [59]. In particular, we consider three qualitatively different data sets used for clustering: Iris, Diabetes, and KDD. For each data set, we only have a single set of points given as input, so we take $\mathcal{N} = \mathcal{M}$ to be the set of all points in the data set. We use the standard Euclidean L2 distance.

- Iris. This data set contains information about the petal dimensions of three different species of iris flowers. There are 50 samples of each species.
- Diabetes. This is the Pima Indians Diabetes data set, and contains information about 768 diabetes patients, recording features like glucose, blood pressure, age and skin thickness.
- KDD. This is the KDD cup 1999 data set, containing information about connections of sequences of TCP Packets. Each TCP Packet can classified as normal or one of twenty-two types of intrusions. Of these 23 classes, normal, "neptune", and "smurf" account for 98.3% of the data.[3]

4.5.1 Local Capture Heuristic

We observe that while our Greedy Capture algorithm (Algorithm 1) always produces an approximately proportional solution, it may not produce an exactly proportional solution in practice, even on instances where such solutions exist (see Figure 4.4a and Figure 4.4b). We therefore introduce a Local Capture heuristic that is able to consistently outperform Greedy Capture in practice.

[3] The data set is extremely large; we work with a subsample of 100,000 points. We run k-means++ on this entire 100,000 point sample. For efficiency, we run our Local Capture algorithm by further sampling 5,000 points uniformly at random to treat as $\mathcal{N}$ and sampling 400 points via the k-means++ initialization to treat as $\mathcal{M}$. For the sake of a fair comparison, we generate a different sample of 400 centers using the k-means++ initialization that we use to determine the value of ρ we

Algorithm 2 Local Capture Heuristic

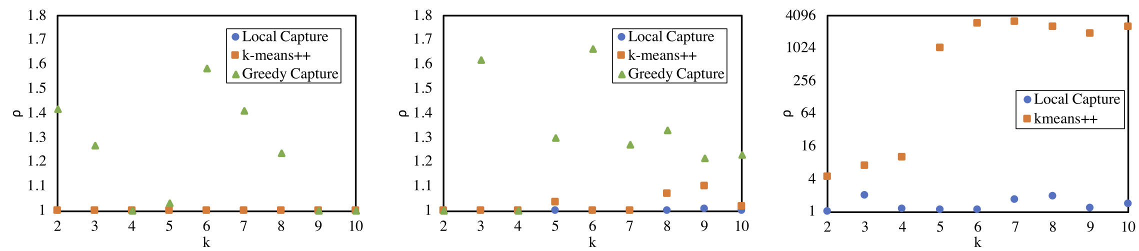
Input ρ
Initialize X as a random subset of k centers from $\mathcal{M}$.
repeat
  for $y \in \mathcal{M}$ **do**
    $S_y \leftarrow \{i \in N : \rho \cdot d_{iy} < D_i(X)\}$
    if $|S_y| \geqslant \lceil \frac{n}{k} \rceil$ **then**
      $x^* \leftarrow \text{argmin}_{x \in X} |\{i \in N : d_{ix} = D_i(X)\}|$
      $X \leftarrow (X \backslash \{x^*\}) \cup \{y\}$
until no changes occur
return X

Algorithm 2 searches for violations of ρ-proportionality (note that ρ is a parameter), and if it finds one, swaps the least demanded center for the center of the proposed violation. Every iteration of Algorithm 2 (the entire inner for loop) runs in $\tilde{O}(mn^2)$ time, and n can be reduced via sampling as discussed in Section 4.4. In our experiments, we try to find the minimum ρ for which the algorithm terminates in a small number of iterations via binary search over possible input of ρ. If Algorithm 2 terminates, then it returns a ρ-proportional solution. In [13], the authors also evaluate a local search swapping procedure for the k-median problem, but their swap condition is based on the relative k-median objective of two solutions, whereas our swap condition is based on violations to proportionality.

4.5.2 Proportionality and k-means Objective Tradeoff

We compare Greedy Capture (Algorithm 1) and Local Capture (Algorithm 2) with the k-means++ algorithm (Lloyd's algorithm for k-means minimization with the k-means++ initialization [12]) for a range of values of k. For the Iris data set, Local Capture and k-means++ always find an exact proportional solution (Figure 4.4a), and have comparable k-means objectives (Figure 4.5a). The Iris data set is very simple with three natural clusters, and validates the intuition that proportionality

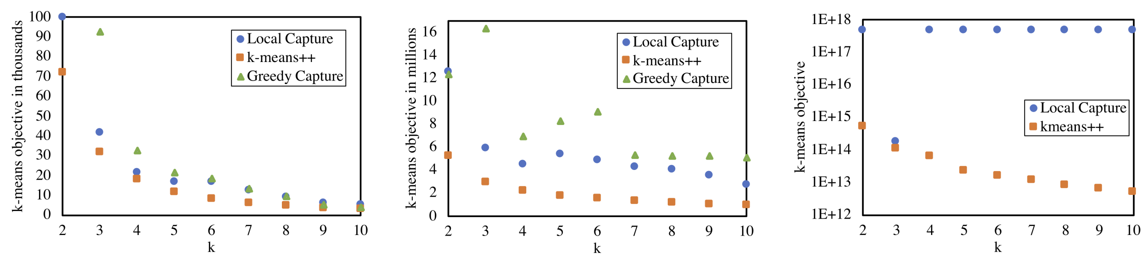
report for both Local Capture and the k-means++ algorithm. The k-means objective is measured on the original 100,000 points for both algorithms.



(a) Iris (b) Diabetes (c) KDD, geometric scale

FIGURE 4.4: Minimum ρ such that the solution is ρ-proportional

.



(a) Iris (b) Diabetes (c) KDD, geometric scale

FIGURE 4.5: k-means objective

and the k-means objective are not always opposed.

The Diabetes data set is larger and more complex. As shown in Figure 4.4b, k-means++ no longer always finds an exact proportional solution. Local Capture always finds a better than 1.01-proportional solution. As shown in Figure 4.5b, the k-means objectives of the solutions are separated, although generally on the same order of magnitude.

The contrast between the k-means objective and proportionality is pronounced for the KDD data set. Greedy Capture's performance on this data set is comparable to Local Capture, so we omit it for clarity. In Figure 4.4c, note that for many values of k, the k-means++ algorithm only finds a roughly 4,000-proportional clustering. On the other hand, Local Capture is still consistently able to find a roughly 2-proportional clustering or better. There is a similarly dramatic difference of between three and five orders of magnitude in the k-means objectives achieved by the algorithms, as seen in Figure 4.5c. This behavior appears to be due to the presence of significant outliers in the KDD data set. This is in keeping with the theoretical impossibility of simultaneously approximating the optima on both objectives, and demonstrates that this tension arises in practice as well as theory.

Note that if one is allowed to use $2k$ centers when k is given as input, one can trivially achieve the proportionality of Local Capture and the k-means objective of the k-means++ algorithm by taking the union of the two solutions. Thinking in this way leads to a different way of quantifying the tradeoff between proportionality and the k-means objective: Given an approximately proportional solution, how many *extra* centers are necessary to get comparable k-means objective as the k-means++ algorithm? For a given data set, the answer is a value between 0 and k, where larger numbers indicate more incompatibility, and lower numbers indicate less incompatibility.

We consider this question by first computing the union of centers computed by Local Capture and the k-means++ algorithm. We then greedily remove centers as long as doing so does not increase the minimum ρ such that the solution is ρ-proportional (defined on k, not $2k$) by more than a multiplicative factor of α, and does not increase the k-means objective by more than a multiplicative factor β.

On the KDD dataset, we set $\alpha = 1.2$ and $\beta = 1.5$, so the proportionality of the result is within 1.2 of Local Capture in Figure 4.4c, and the k-means objective is within 1.5 of k-means++ in Figure 4.5c. We observe that this heuristic uses at most 3 extra centers for any $k \leqslant 10$. So while there is real tension between proportionality and the k-means objective, this tension is still not maximal. In the worst case, one might need to add k centers to a proportional solution to compete with the k-means objective of the k-means++ algorithm, but in practice we find that we need at most 3 for $k \leqslant 10$.

4.6 Conclusion and Open Directions

We have introduced proportionality as a fair solution concept for centroid clustering. Although exact proportional solutions may not exist, we gave efficient algorithms for computing approximate proportional solutions, and considered constrained optimization and sampling for further applications. Finally, we studied proportionality on real data and observed a data dependent tradeoff between proportionality and the k-means objective. While this tradeoff is in some sense a negative result, it also demonstrates that proportionality as a fairness guarantee matters in the sense that it meaningfully constrains the space of solutions.

We have shown that ρ-proportional solutions need not exist for $\rho < 2$, and always exist for $\rho \geqslant 1 + \sqrt{2}$. Closing this approximability gap is one outstanding question. Another is whether there is a more efficient and easily interpretable algorithm for optimizing total cost subject to proportionality, as our approach in Section 4.3 requires solving a linear program on the entire data set. We would ideally like a more efficient and easily interpretable primal-dual or local search type algorithm. More generally, what other fair solution concepts for clustering should be considered alongside proportionality, and can we characterize their relative advantages and disadvantages? Finally, can the idea of proportionality as a group fairness concept be adapted for supervised learning tasks like classification and regression?

5

Metric Implicit Utilitarian Voting via Bargaining

Here, we study social choice: informally, how to pick a point from a decision space that represents the common good of society. The canonical example is voting. In that case, the decision space is usually a small set of candidates, and the agents comprising the society express only ordinal preferences over these candidates. In contrast, we will study large and complex decision spaces under the implicit utilitarian assumption that agents have cardinal preferences over decision points, even if they cannot concisely express those preferences. The parallel to previous chapters should now be clear: decision points are public goods and our budget allows us to choose exactly one good. Intuitively this seems simpler, but we will make severe restrictions on the kind of input information that agents can provide (note that we elicited full utility functions in previous sections) in order to develop a practical mechanism for social choice. Technically, we provide a novel connection between the concept of *Nash bargaining* in cooperative game theory and the median of a set of points in a *median graph*, and use this connection to design a simple randomized mechanism that approximately maximizes the utilitarian social welfare. The analysis is over a Markov chain on a hypercube. We also provide several lower bounds, analyze higher moments of the utilitarian social welfare, and show how the results change in general metric spaces.

Acknowledgements. These results are published in [65], which is joint work with Ashish Goel, Kamesh Munagala, and Sukolsak Sakshuwong.

5.1 Introduction

Suppose a university administrator plans to spend millions of dollars to update her campus, and she want to elicit the input of students, staff, and faculty. In a typical social choice setting, she could first elicit the favorite decisions, or *bliss points*, of the students, say “new gym,” “new library,” and “new student center.” However, voting on these options need not find the social optimum, because it is not clear that the social optimum is even on the ballot. In such a setting, *deliberation* between individuals would find entirely new alternatives, for example “replace gym equipment plus remodeling campus dining plus money for scholarship”. This leads to finding a social optimum over a wider space of semi-structured outcomes that the system/mechanism designer was not originally aware of, and the participants had not initially articulated.

We therefore start with the following premise: The mechanism designer may not be able to enumerate the outcomes in the decision space or know their structure, and this decision space may be too big for most ordinal voting schemes. (For instance, ordinal voting is difficult to implement in complex combinatorial spaces [114] or in continuous spaces [75].) However, we assume that agents can still reason about their preferences and small groups of agents can negotiate over this space and collaboratively propose outcomes that appeal to all of them. Our goal is to design protocols based on such a primitive by which small group negotiation can lead to an aggregation of societal preferences without a need to formally articulate the entire decision space and without every agent having to report ordinal rankings over this space.

The need for small groups is motivated by a practical consideration as well as a theoretical one. On the practical side, there is no online platform, to the best of our knowledge, that has a successful history of large scale deliberation and decision making on complex issues; in fact, large online forums typically degenerate into vitriol and name calling when there is substantive disagreement among the participants. Thus, if we are to develop practical tools for decision making at scale, a sequence of small group deliberations appears to be the most plausible path. On the theoretical side, we understand the connections between sequential protocols for deliberation and axiomatic theories of bargaining for small groups, e.g. for pairs [146, 28], but not for large groups, and we seek to bridge this gap.

Summary of Contributions. Our main contributions are two-fold:

- A simple and practical sequential protocol that only requires agents to negotiate in pairs and collaboratively propose outcomes that appeal to both of them.
- A canonical analytic model in which we can precisely state properties of this protocol in terms of approximation of the social optimum, Pareto-efficiency, and incentive-compatibility, as well as compare it with simpler protocols.

5.1.1 Background: Bargaining Theory

Before proceeding further, we review bargaining, the classical framework for two player negotiation in Economics. Two-person bargaining, as framed in [130], is a game wherein there is a disagreement outcome and two agents must cooperate to reach a decision; failure to cooperate results in the adoption of the disagreement outcome. Nash postulated four axioms that the bargaining solution ought to satisfy assuming a convex space of alternatives: Pareto optimality (agents find an outcome that cannot be simultaneously improved for both of them), symmetry between agents, invariance with respect to affine transformations of utility (scalar multiplication or additive translation of any agent's utility should not change the outcome), and independence of irrelevant alternatives (informally that the presence of a feasible outcome that agents do not select does not influence their decision). Nash proved that the solution maximizing the Nash product (that we describe later) is the unique solution satisfying these axioms. To provide some explanation of how two agents might find such a solution, [146] shows that Nash's solution is the subgame perfect equilibrium of a simple repeated game on the two agents, where the agents take turns making offers, and at each round, there is an exogenous probability of the process terminating with no agreement.

The two-person bargaining model is therefore clean and easy to reason about. As a consequence, it has been extensively studied. In fact, there are other models and solutions to two-person bargaining, each with a slightly different axiomatization [103, 104, 127], as well as several experimental studies [145, 131, 29]. In a social choice setting, there are typically many more than two agents, each agent having their own complex preferences. Though bargaining can be generalized to n agents with similar axiomatization and solution structure, such a generalization is considered impractical. This is because in reality it is difficult to get a large number of individuals to negotiate coherently; complexities come with the formation of coalitions and power structures [90, 110]. Any model for simultaneous bargaining, even with three players [26], needs to take these messy aspects into account.

5.1.2 A Practical Compromise: Sequential Pairwise Deliberation

We take a middle path, avoiding both the complexity of explicitly specifying preferences in a large decision space (fully centralized voting), and that of simultaneous n-person bargaining (a fully decentralized cooperative game). We term this approach *sequential deliberation*. We use 2-person bargaining as a basic primitive, and view deliberation as a sequence of pairwise interactions that refine good alternatives into better ones as time goes by. Our framework has the advantage of being simple with low cognitive overhead, and is easy to implement and reason about.

More formally, there is a decision space $\mathcal{S}$ of feasible alternatives (these may be projects, sets of projects, or continuous allocations) and a set $\mathcal{N}$ of agents. We assume each agent has a hidden cardinal utility for each alternative. We encapsulate deliberation as a sequential process. Let $\mathcal{H}$ denote the distribution over $\mathcal{S}$ generated by sequential deliberation. We define a bargaining function $\mathcal{B}(u, v, a)$ that takes as input a pair of agents u and v and a default alternative a. The agents negotiate and output a consensus alternative $o \in \mathcal{S}$; if no consensus is reached then a is output. We assume $\mathcal{B}(u, v, a)$ satisfies the axioms – symmetry, scale-invariance, independence of irrelevant alternatives, and Pareto-efficiency – that correspond to Nash bargaining. Our mechanism is defined in Figure 5.1.

Let a^1 be the bliss point of an agent chosen uniformly at random.

1. In each round $t = 1, 2, \ldots, T$:
   (a) A pair of agents u^t and v^t are chosen independently and uniformly at random with replacement.
   (b) These agents are presented with a disagreement alternative a^t, and perform $\mathcal{B}(u_t, v_t, a_t)$ as described below.
   (c) Agents u_t and v_t are asked to output a consensus alternative; if they fail to reach a consensus within a specified time frame, the alternative a^t is output.
   (d) Let o^t denote the alternative that is output in round t. We set $a^{t+1} = o^t$, where we assume a^1 is the bliss point of an arbitrary agent.
2. The final social choice is a^T. Note that this is equivalent to drawing an outcome at random from the distribution generated by repeating this protocol several times.

FIGURE 5.1: A framework for sequential pairwise deliberation.

5.1.3 Analytical Model: Median Graphs and Sequential Nash Bargaining

The framework in Figure 5.1 is well-defined and practical irrespective of an analytical model. However, we provide a simple analytical model for specifying the preferences of the agents in which we can precisely quantify the behavior of this framework as justification. Assuming that disutility is some metric over the space [7] seems reasonable enough, and our tightest results are for a median graph metrics specifically. Assume that the set $\mathcal{S}$ of alternatives are vertices of a *median graph*. A median graph has the property that for each triplet of vertices u, v, w, there is a unique point that is common to the three sets of shortest paths, those between u, v, between v, w, and between u, w. This point is the unique *median* of u, v, w. We assume each agent u has a bliss point $p_u \in \mathcal{S}$, and his disutility for an alternative $a \in \mathcal{S}$ is simply $d(p_u, a)$, where $d(\cdot)$ is the shortest path distance function on the median graph. (Note that this disutility can have an agent-dependent scale factor.)

Several natural graphs are median graphs, including trees, points on the line, hypercubes, and grid graphs in arbitrary dimensions [108]. Median graphs have been extensively studied as structured models for spatial preferences in voting theory. They are appealing for a very simple reason [48]: For any set of agents and alternatives on a median graph, if the agents could perform ordinal voting on the alternatives, then the Condorcet winner (an alternative that beats all others in pairwise voting) always exists, and coincides with the generalized 1-median (the alternative that minimizes the sum of distances to the agents' bliss points). In a sense, we consider spaces where had the agents been able to vote, there would have been a clear winner even using pairwise comparisons. Our premise is that the space is so large that (a) voting over all possible alternatives is often impractical, and (b) it is not obvious how to find a small set of "good" alternatives. We therefore ask the natural questions:

> How close can we come to the overall Condorcet winner or median by asking agents to propose alternatives collaboratively, instead of by explicit voting? Can such a method also be truthful and Pareto-efficient?

Nash Bargaining. The model for two-person bargaining is simply the classical *Nash bargaining* solution described before. Given a disagreement alternative a, agents u and v choose that alternative $o \in \mathcal{S}$ that maximizes:

$$\text{Nash product } = (d(p_u, a) - d(p_u, o)) \times (d(p_v, a) - d(p_v, o))$$

The Nash product maximizer need not be unique; in the case of ties we postulate that agents select the outcome that is closest to the disagreement outcome. As mentioned before, the Nash product is a widely studied axiomatic notion of pairwise interactions, and is therefore a natural solution concept in our framework.

Social Cost and Distortion. The *social cost* of an alternative $a \in \mathcal{S}$ is given by $SC(a) = \sum_{u \in \mathcal{N}} d(p_u, a)$. Let $a^* \in \mathcal{S}$ be the minimizer of social cost, *i.e.*, the *generalized median.* We measure the Distortion of outcome a as

$$\text{Distortion}(a) = \frac{SC(a)}{SC(a^*)} \tag{5.1}$$

where we use the expected social cost if a is the outcome of a randomized algorithm.

Note that our model is fairly general. First, the bliss points of the agents in $\mathcal{N}$ form an arbitrary subset of $\mathcal{S}$. Further, the alternative chosen by bargaining need not correspond to any bliss point, so that pairs of agents are exploring the space of alternatives when they bargain. This is essentially *discovery of new options via deliberation*, and is not captured by simply voting on bliss points. We emphasize that while we present analytical results for sequential deliberation in specific decision spaces, the framework in Figure 5.1 is well defined regardless of the underlying decision space and the mediator's understanding of the space. At a high level, this flexibility and generality in practice are its key advantages.

5.2 Median Graphs and Nash Bargaining

In this section we will use the notation $\mathcal{N}$ for a set of agents, $\mathcal{S}$ for the space of feasible alternatives, and $\mathcal{H}$ for a distribution over $\mathcal{S}$. Most of our results are for the analytic model given earlier wherein the set $\mathcal{S}$ of alternatives are vertices of a *median graph*; see Figure 5.2 for some examples.

Definition 14. *A median graph $G(\mathcal{S}, E)$ is an undirected graph with the following property: For each triplet of vertices $u, v, w \in \mathcal{S} \times \mathcal{S} \times \mathcal{S}$, there is a unique point that is common to the shortest paths (which need not be unique between a given pair) between u, v, between v, w, and between u, w. This point is the unique* median *of u, v, w.*

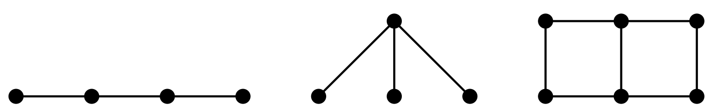


FIGURE 5.2: Examples of Median Graphs

In the framework of Figure 5.1, we assume that at every step, two agents perform Nash bargaining with a disagreement alternative. The first results characterize Nash bargaining on a median graph. In particular, we show that Nash bargaining at each step will select the median of bliss points of the two agents and the disagreement alternative. After that, we show that we can analyze the Distortion of sequential deliberation on a median graph by looking at the embedding of that graph onto the hypercube.

Lemma 11. *For any median graph $G = (\mathcal{S}, E)$, any two agents u, v with bliss points $p_u, p_v \in \mathcal{S}$, and any disagreement outcome $a \in \mathcal{S}$, let M be the median. Then M maximizes the Nash product of u and v given a, and is the maximizer closest to a.*

Proof. Since G is a median graph, M exists and is unique; it must by definition be the intersection of the three shortest paths between $(p_u, p_v), (p_u, a), (p_v, a)$. Note that we can therefore write $d(p_u, a) = d(p_u, M) + d(M, a)$ and similarly for $d(p_v, a)$. Let $\alpha = d(p_u, M)$; $\beta = d(p_v, M)$; and $\gamma = d(a, M)$. Suppose Nash bargaining finds an outcome $o^* \in \mathcal{S}$. Let $x = d(o^*, p_u)$ and $y = d(o^*, p_v)$. Observing that M lies on the shortest path between p_u and p_v, and using the triangle inequality, we obtain that $x + y \geqslant \alpha + \beta$.

Noting that $d(p_u, a) = \alpha + \gamma$ and $d(p_v, a) = \beta + \gamma$, the Nash product of the point o^* is:

$$(\alpha + \gamma - x) \times (\beta + \gamma - y) \leqslant (\alpha + \gamma - x)(\gamma - (\alpha - x)) = \gamma^2 - (\alpha - x)^2$$

This is maximized when $x = \alpha$ and $y = \beta$. One possible maximizer is therefore the point $o^* = M$. Suppose $d(o^*, a) < \gamma$, then by triangle inequality, $d(p_u, o^*) > \alpha$, and similarly $d(p_v, o^*) > \beta$. Therefore, there cannot be a closer maximizer of the Nash product to a than the point M. □

Hypercube Embeddings. For any median graph $G = (\mathcal{S}, E)$, there is an isometric embedding $\phi : G \to Q$ of G into a hypercube Q [108]. This embedding maps vertices $\mathcal{S}$ into a subset of vertices of Q so that all pairwise distances between vertices in $\mathcal{S}$ are preserved by the embedding. A simple example of this embedding for a tree is shown in Figure 5.3. We use this embedding to show the following result.

Lemma 12. *Let $G(\mathcal{S}, E)$ be a median graph, and let ϕ be its isometric embedding into hypercube $Q(V, E')$. For any three points $t, u, v \in \mathcal{S}$, let M_G be the median*

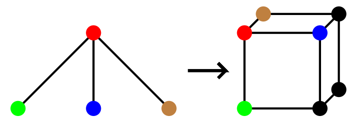


FIGURE 5.3: The hypercube embedding of a 4-vertex star graph

of vertices t, u, v *and let* M_Q *be the median of vertices* $\phi(t), \phi(u), \phi(v) \in V$*. Then* $\phi(M_G) = M_Q$*.*

Proof. By definition, since ϕ is an isometric embedding [108],

$$d(x, y) = d(\phi(x), \phi(y)) \text{ for all } x, y \in \mathcal{S} \tag{5.2}$$

Since G is a median graph, M_G is the unique median of $t, u, v \in \mathcal{S}$, which by definition satisfies the equalities $d(t, u) = d(t, M_G) + d(M_G, u)$, $d(t, v) = d(t, M_G) + d(M_G, v)$, and $d(u, v) = d(u, M_G) + d(M_G, v)$.

$Q(V, E')$ is a hypercube, and is thus also a median graph, so M_Q is the unique median of $\phi(t), \phi(u), \phi(v) \in V$, which by definition satisfies the equalities

I $d(\phi(t), \phi(u)) = d(\phi(t), M_Q) + d(M_Q, \phi(u))$

II $d(\phi(t), \phi(v)) = d(\phi(t), M_Q) + d(M_Q, \phi(v))$

III $d(\phi(u), \phi(v)) = d(\phi(u), M_Q) + d(M_Q, \phi(v))$

Applying Equation (5.2) to the first set of equalities shows that $\phi(M_G)$ satisfies equalities I,II, and III respectively. But $\phi(M_G) \in V$ and M_Q is the unique vertex in V satisfying equalities I,II, and III. Therefore, $\phi(M_G) = M_Q$. □

5.3 The Efficiency of Sequential Deliberation

In this section, we show that the Distortion of sequential deliberation is at most 1.208. We then show that this bound is significant, meaning that mechanisms from simpler classes are necessarily constrained to have higher Distortion values.

Upper Bounding Distortion Recall the framework for sequential deliberation in Figure 5.1 and the definition of Distortion in Equation (5.1). We first map the problem into a problem on hypercubes using Lemma 12.

Corollary 4. *Let $G = (\mathcal{S}, E)$ be a median graph, let $\phi : G \to Q$ be an isometric embedding of G onto a hypercube $Q(V, E')$, and let $\mathcal{N}$ be a set of agents such that each agent u has a bliss point $p_u \in \mathcal{S}$. Then the Distortion of sequential deliberation on G is at most the Distortion of sequential deliberation on $\phi(G)$ where each agent's bliss point is $\phi(p_u)$.*

Proof. Fix an initial disagreement outcome $a^1 \in \mathcal{S}$ and an arbitrary list of T pairs of agents $(u^1, v^1), (u^2, v^2), \ldots, (u^T, v^T)$. In round 1 bargaining on G, Lemma 11 implies that sequential deliberation will select $o^1 = \text{median}(a^1, p_u^1, p_v^1)$. Furthermore, Lemma 12 implies that if we had considered $\phi(G)$ and bargaining on $\phi(a^1), \phi(p_u^1), \phi(p_v^1)$ instead, sequential deliberation would have selected $\phi(o^1)$. Suppose at some round t that we have a disagreement outcome a^t. Then the same argument yields that if o^t is the bargaining outcome on G, $\phi(o^t)$ would have been the bargaining outcome on $\phi(G)$. Thus, by induction, we have that if the list of outcomes on G is $o^1, ..., o^T$ then the list of outcomes on $\phi(G)$ is $\phi(o^1), ..., \phi(o^T)$. But recall that $\phi(\cdot)$ is an isometric embedding and the social cost of an alternative is just its sum of distances to all points in $\mathcal{N}$, so a^T and $\phi(a^T)$ have the same social cost.

Furthermore, let $a^* \in \mathcal{S}$ denote the generalized median of $\mathcal{N}$. Then, $\phi(a^*)$ has the same social cost as a^*. This means the median of the embedding of $\mathcal{N}$ into Q has at most this social cost, which in turn means that the Distortion of sequential deliberation in the embedding cannot decrease. □

Our main result in this section shows that as $t \to \infty$, the Distortion of sequential deliberation approaches 1.208, with the convergence rate being exponentially fast in t and independent of the number of agents $|\mathcal{N}|$, the size of the median space $|\mathcal{S}|$, and the initial disagreement point a^1. In particular, the Distortion is at most 1.22 in at most 9 steps of deliberation, which is indeed a very small number of steps!

Theorem 15. *Sequential deliberation among a set $\mathcal{N}$ of agents, where the decision space $\mathcal{S}$ is a median graph, yields $\mathbb{E}[Distortion(a^t)] \leqslant 1.208 + \frac{6}{2^t}$.*

Proof. By Corollary 4, we can assume the decision space is a D-dimensional hypercube Q so that decision points (and thus bliss points) are vectors in $\{0,1\}^D$. For every dimension k, let f_k be the fraction of agents whose bliss point has a 1 in the kth dimension, and let $p_{u,k}$ be the 0 or 1 bit in the kth dimension of the bliss point p_u for agent u. Let $a^* \in \{0,1\}^D$ be the minimum social cost decision point, i.e., $a^* := \text{argmin}_{a \in Q} SC(a)$. Clearly, a^*_k is 1 if $f_k > 1/2$ and 0 otherwise, so for every dimension k, the total distance to a^*_k, summed over $\mathcal{N}$ is:

$$\sum_{u \in \mathcal{N}} |a^*_k - p_{u,k}| = |\mathcal{N}| \min\{f_k,\, 1 - f_k\}$$

Now, note that sequential deliberation defines a Markov chain on Q. The state in a given step is just a^t, and the randomness is in the random draw of the two agents. Let $\mathcal{H}^*$ be the stationary distribution of the Markov chain. Then we can write

$$\lim_{t \to \infty} \mathbb{E}[\text{Distortion}(a^t)] = \mathbb{E}_{a \in \mathcal{H}^*}\left[\text{Distortion}(a)\right]$$

To write down the transition probabilities, we assume this random draw is two independent uniform random draws from $\mathcal{N}$, with replacement. We also note that Lemma 11 implies that on Q, sequential deliberation will pick the median in every step, *i.e.*, given a disagreement outcome a^t and two randomly drawn agents with bliss points p_u, p_v, the new decision point will be $o^t = \text{median}(p_u, p_v, a^t)$. On a hypercube, the median of three points is just the dimension-wise majority. Thus, we get a 2-state Markov chain in each dimension k, with transition probabilities

$$\Pr[o^t_k = 1 | a^t_k = 1] = f_k^2 + 2 f_k (1 - f_k) \text{ and } \Pr[o^t_k = 1 | a^t_k = 0] = f_k^2$$

Let $\mathcal{P}_k^* = \lim_{t\to\infty} \Pr[o_k^t = 1]$, and let $\mathcal{H}_k^*$ denote this stationary distribution for the corresponding 2-state Markov chain. Then,

$$\mathcal{P}_k^* = \left(2f_k - f_k^2\right)\mathcal{P}_k^* + \left(f_k^2\right)\left(1 - \mathcal{P}_k^*\right) \qquad \Rightarrow \qquad \mathcal{P}_k^* = \frac{f_k^2}{1 + 2f_k^2 - 2f_k}$$

By linearity of expectation, the total expected distance for every dimension k, summed over $u \in \mathcal{N}$ to the final outcome is given by

$$\begin{aligned}
&\mathbb{E}_{a_k \in \mathcal{H}_k^*}\left[\sum_{u\in\mathcal{N}} |a_k - p_{u,k}|\right] \\
&= |\mathcal{N}|\left(\left(\frac{f_k^2}{f_k^2 + (1-f_k)^2}\right)(1-f_k) + \left(1 - \frac{f_k^2}{f_k^2 + (1-f_k)^2}\right) f_k\right) \\
&= |\mathcal{N}|\left(\frac{f_k(1-f_k)}{f_k^2 + (1-f_k)^2}\right)
\end{aligned}$$

Without loss of generality, let $f_k \in [0, 1/2]$ so that for dimension k, the total distance to a_k^* is $|\mathcal{N}| f_k$. Then the ratio of the expected total distance to $\mathcal{H}_k^*$ to the total distance to a_k^* is at most:

$$\frac{\mathbb{E}_{a_k \in \mathcal{H}_k^*}\left[\sum_{u\in\mathcal{N}} |p_{u,k} - a_k|\right]}{\sum_{u\in\mathcal{N}} |p_{u,k} - a_k^*|} \leqslant \max_{f_k \in [0,1/2]} \frac{1 - f_k}{f_k^2 + (1-f_k)^2} \leqslant 1.208$$

Since the above bound holds in each dimension of the hypercube, we can combine them as:

$$\mathbb{E}_{a\in\mathcal{H}^*}\left[\text{Distortion}(a)\right] = \frac{\sum_{k=1}^{D} \mathbb{E}_{a_k\in\mathcal{H}_k^*}\left[\sum_{u\in\mathcal{N}} |p_{u,k} - a_k|\right]}{\sum_{k=1}^{D}\sum_{u\in\mathcal{N}} |p_{u,k} - a_k^*|} \leqslant 1.208$$

Now that we have bounded the Distortion of the stationary distribution, we need to consider the convergence rate. We will not bound the mixing time of the overall Markov chain. Rather, note that in the preceding analysis, we only used the *marginal* probabilities $\mathcal{H}_k^*$ for each dimension k. Furthermore, the Markov chain defined by sequential deliberation need not walk along edges on Q, so we can consider separately the convergence of the chain to the stationary marginal in each dimension.

After t steps, let $P_{kt} = \Pr[o_k^t = 1]$ and let $\mathcal{H}_k^t$ denote this distribution. Assume $f_k \in [0, 1/2]$. If the total variation distance between $\mathcal{H}_k^t$ and $\mathcal{H}_k^*$ is $\epsilon \frac{f_k^2}{f_k^2 + (1-f_k)^2}$, then it is easy to check that the expected total distance to $\mathcal{H}_k^t$ is within a $(1+\epsilon)$ factor of the expected distance to $\mathcal{H}_k^*$, which implies a Distortion of at most $1.208 + \epsilon$ in that dimension. We therefore bound how many steps it takes to achieve total variation distance $\epsilon \frac{f_k^2}{f_k^2 + (1-f_k)^2}$ in any dimension k; if this bound holds uniformly for all dimensions k, this would imply the overall Distortion is at most $1.208 + \epsilon$, completing the proof.

For any dimension k, two executions of the 2-state Markov chain along that dimension couple if the agents picked in a time step have the same value in that dimension. At any step, this happens with probability at least $(f_k^2 + (1-f_k)^2)$. Therefore, the probability that the chains have not coupled in t steps is at most

$$\left(1 - \left(f_k^2 + (1-f_k)^2\right)\right)^t = \left(2f_k(1-f_k)\right)^t$$

We therefore need T large enough so that

$$\begin{aligned}
(2f_k(1-f_k))^T &\leqslant \epsilon \frac{f_k^2}{f_k^2 + (1-f_k)^2} \\
\Rightarrow \qquad T &= \max_{f_k \in [0,1/2]} \left(\frac{\log \frac{1}{\epsilon}}{\log \frac{1}{2f_k(1-f_k)}} + \frac{\log \frac{1}{f_k^2} + \log(f_k^2 + (1-f_k)^2)}{\log \frac{1}{2f_k(1-f_k)}} \right) \\
\Rightarrow \qquad T &\leqslant \log_2 \frac{1}{\epsilon} + 2.575
\end{aligned}$$

Since this bound of T holds uniformly for all dimensions, this directly implies the theorem. □

5.3.1 Lower Bounds on Distortion

We will now show that the Distortion bounds of sequential deliberation are significant, meaning that mechanisms from simpler classes are constrained to have higher

Distortion values. We present a sequence of lower bounds for social choice mechanisms that are restricted to use increasingly richer information about the space of alternatives on a median graph $G = (\mathcal{S}, E)$ with a set of agents $\mathcal{N}$ with bliss points $V_{\mathcal{N}} \subseteq \mathcal{S}$. We first consider mechanisms that are constrained to choose outcomes in $V_{\mathcal{N}}$. For instance, this captures the random dictatorship that chooses the bliss point of a random agent as the final outcome. It shows that the compromise alternatives found by deliberation do play a role in reducing Distortion.

Lemma 13. *Any mechanism constrained to choose outcomes in $V_{\mathcal{N}}$ has Distortion at least 2.*

Proof. It is easy to see that the k-star graph (the graph with a central vertex connected to k other vertices none of which have edges between themselves) is a median graph. Consider an $|\mathcal{N}|$-star graph where $V_{\mathcal{N}}$ are the non central vertices; that is, each and every agent has a unique bliss point on the periphery of the star. Then any mechanism constrained to choose outcomes in $V_{\mathcal{N}}$ must choose one of these vertices on the periphery of the star. The social cost of such a point is $(|\mathcal{N}| - 1) \times 2$, whereas the social cost of the optimal central vertex is clearly just $|\mathcal{N}|$. The Distortion goes to 2 as $|\mathcal{N}|$ grows large. □

We next consider mechanisms that are restricted to choosing the median of the bliss points of some three agents in $\mathcal{N}$. In particular, this captures sequential deliberation run for $T = 1$ steps, as well as mechanisms that generalize dictatorship to an oligarchy composed of at most 3 agents. This shows that iteratively refining the bargaining outcome has better Distortion than performing only one iteration.

Lemma 14. *Any mechanism constrained to choose outcomes in $V_{\mathcal{N}}$ or a median of three points in $V_{\mathcal{N}}$ must have Distortion at least 1.316.*

Proof. Let G be a median graph; in particular let G be the D-dimensional hypercube $\{0, 1\}^D$. For every dimension, an agent has a 1 in that dimension of their bliss point independently with probability p. In expectation $p|\mathcal{N}|$ agents' bliss points have a 1 in any given dimension. We assume $0 < p < 1/2$ is an absolute constant. For a^* being the all 0's vector, $\mathbb{E}[SC(a^*)] = D|\mathcal{N}|p$, where the randomness is in the construction. Now, suppose $D = \text{polylog}(|\mathcal{N}|)$. Then, for any $\beta < 1$, with probability at least $1 - \frac{1}{\text{poly}(|\mathcal{N}|)}$ every three points in $V_{\mathcal{N}}$ has at least $\beta(3p^2 - 2p^3)D$ ones. By union bounds, for some $\alpha \in (\beta, 1)$, the social cost of any median of three points in $V_{\mathcal{N}}$ is at least:

$$\begin{aligned}\mathbb{E}[SC(a)] &\geqslant D|\mathcal{N}|\alpha\left[(3p^2 - 2p^3)(1-p) + p(1 - 3p^2 + 2p^3)\right]\\ &= D|\mathcal{N}|\alpha\left(4p^4 - 8p^3 + 3p^2 + p\right)\end{aligned}$$

where again, the randomness is in the construction. Then there is nonzero probability that

$$\frac{\mathbb{E}[SC(a)]}{\mathbb{E}[SC(a^*)]} \geqslant \alpha\left(4p^3 - 8p^2 + 3p + 1\right) \tag{5.3}$$

If we choose the argmax of Equation 5.3, we get nonzero probability over the construction that the Distortion is at least $(1.316)\alpha$. Letting α grow close to 1 and noting that the nonzero probability over the construction implies the existence of one such instance completes the argument. □

We finally consider a class of mechanisms that includes sequential deliberation as a special case. We show that any mechanism in this class cannot have Distortion arbitrarily close to 1. This also shows that sequential deliberation is close to best possible in this class.

Lemma 15. *Any mechanism constrained to choose outcomes on shortest paths between pairs of outcomes in $V_{\mathcal{N}}$ must have Distortion at least $9/8 = 1.125$.*

Proof. The construction of the lower bound essentially mimics that of Lemma 14. In this case however, we get that each point on a shortest path between two agents has at least $\alpha(p^2 \times (1-p) + (1-p^2) \times p)D$ ones, so

$$\mathbb{E}[SC(a)] \geqslant D|\mathcal{N}|\alpha\left[p^2 \times (1-p) + (1-p^2) \times p\right] = D|\mathcal{N}|\alpha\left(-2p^3 + p^2 + p\right)$$

So there is nonzero probability over the construction that

$$\frac{\mathbb{E}[SC(a)]}{\mathbb{E}[SC(a^*)]} \geqslant \alpha\left(-2p^2 + p + 1\right)$$

where the maximum of $-2p^2 + p + 1$ over $p < 1/2$ is $9/8$ when $p = 1/4$. The rest of the argument follows as in Lemma 14. □

The significance of the lower bound in Lemma 15 should be emphasized: though there is always a Condorcet winner in median graphs, it need not be any agents bliss point, nor does it need to be Pareto optimal for any pair of agents. The somewhat surprising implication is that any local mechanism (in the sense that the mechanism chooses locally Pareto optimal points) is constrained away from finding the Condorcet winner. At first glance, it might seem like [82] violates this lower bound; but it is crucial to note that in their model, the decision space consists only of the bliss points of agents, rather than being a potentially much larger superset.

5.4 Properties of Sequential Deliberation

In this section, we study some natural desiderata that should be satisfied by reasonable social choice functions, and ask if Sequential Deliberation satisfies them.

Pareto-Efficiency We first show that the outcome of sequential deliberation is ex-post Pareto-efficient on a median graph. In other words, in any realization of the random process, suppose the final outcome is o; then there is no other alternative a such that $d(a, v) \leqslant d(o, v)$ for every $v \in \mathcal{N}$, with at least one inequality being strict. This is a weak notion of efficiency, but it is not trivial to show; while it is easy to see that a one shot bargaining mechanism using only bliss points is Pareto efficient by virtue of the Pareto efficiency of bargaining, sequential deliberation defines a potentially complicated Markov chain for which many of the outcomes need not be bliss points themselves.

Theorem 16. *Sequential deliberation among a set $\mathcal{N}$ of agents, where the decision space $\mathcal{S}$ is a median graph and the initial disagreement point a^1 is the bliss point of some agent, yields an ex-post Pareto Efficient alternative.*

Proof. Let $G = (\mathcal{S}, E)$ be a median graph, let $\mathcal{N}$ be a set of agents, and let $V_{\mathcal{N}} \subseteq \mathcal{S}$ be the set of bliss points of the agents in $\mathcal{N}$. It follows from the proof of Corollary 4 that without loss of generality we can suppose $\mathcal{S}$ is a hypercube embedding.

Consider some realization $(x^1, y^1), (x^2, y^2), ..., (x^T, y^T)$ of sequential bargaining, where $(x^t, y^t) \in \mathcal{S}$ are the bliss points of the agents drawn to bargain in step t. Let o^T denote the final outcome. For the sake of contradiction assume there is an alternative z that Pareto-dominates o^T, i.e., $d(z, v) \leqslant d(o^T, v)$ for each $v \in \mathcal{N}$, with at least one inequality being strict.

Recall that on the hypercube, the median of three points has the particularly simple form of the dimension wise majority. Let $\mathcal{K}^t$ be the set of dimensions of the hypercube $\mathcal{S}$ that are "decided" by the agents in round t in the sense that these agents agree on that dimension (and can thus ignore the outside alternative in that dimension) and all future agents disagree on that dimension (and thus keep the value decided by bargaining in round t). Formally, $\mathcal{K}^t = \left\{k : x_k^t = y_k^t \text{ and } \forall t' > t, x_k^{t'} \neq y_k^{t'}\right\}$. Then by the majority property of the median on the hypercube, for any dimension k such that $k \in \mathcal{K}^t$ for some $t \in \{1, \ldots, T\}$, it must be that $o_k^T = x_k^t$. An example is shown in Figure 5.4.

Consider the final round T. It must be that $\forall k \in \mathcal{K}^T, z_k = o_k^T$. If this were not the case, z would not be pairwise efficient (i.e., on a shortest path from x^T to y^T),

$\mathcal{V}_{\mathcal{N}}$

01100
10100
00001
11000

t	x^t / y^t	$a^t \downarrow o^t$	$\mathcal{K}^t$
1	10100 / 00001	01100 ↓ 00100	{2}
2	10100 / 11000	00100 ↓ 10100	{5}
3	00001 / 01100	10100 ↓ 00100	∅
4	01100 / 00001	00100 ↓ 00100	{1,4}

FIGURE 5.4: An example of sequential deliberation with $\mathcal{K}^t$ labeled. The dimensions are numbered $1, 2, \ldots, 5$.

whereas o^T is pairwise efficient by definition of the median, so one of the agents in round T would strictly prefer o^T to z, violating the dominance of z over o^T.

Next, consider round $T-1$. Partition the dimensions of $\mathcal{S}$ into $\mathcal{K}^T, \mathcal{K}^{T-1}$ and all others. Suppose for a contradiction that $\exists k \in \mathcal{K}^{T-1}$ such that $z_k \neq o_k^T$, where $x_k^{T-1} = y^{T-1} = o_k^T$ by definition. Then the agents in round $T-1$ must strictly prefer o^T to z on the dimensions in $\mathcal{K}^{T-1}$. But for $k \in \mathcal{K}^T$, we know that $z_k = o_k^T$, so the agents are indifferent between z and o^T on the dimensions in $\mathcal{K}^T$. Furthermore, for $k \notin \mathcal{K}^T \cup \mathcal{K}^{T-1}, x_k^{T-1} \neq y_k^{T-1}$, so at least one of the two agents at least weakly prefers o^T to z on the remaining dimensions. But then at least one agent must strictly prefer o^T to z, contradicting the dominance of z over o^T.

Repeating this argument yields that for all $k \in \mathcal{K}^1 \cup \mathcal{K}^2 \cup \ldots \cup \mathcal{K}^T$, $z_k = o_k^T$. For all other dimensions, o_k^T takes on the value a^1, which is the bliss point of some agent. Since that agent must weakly prefer z to o^T, z must also take the value of her bliss point on these remaining dimensions. But then $z = o^T$, so z does not Pareto dominate o^T, a contradiction. □

Uniqueness of Stationary Distribution Next, we show that the Markov chain corresponding to sequential deliberation converges to a unique stationary distribution on the actual median graph, rather than just showing that the marginals and thus the expected distances from the perspectives of the agents converge.

Theorem 17. *The Markov chain defined in Theorem 15 has a unique stationary distribution.*

Proof. It is helpful to note that the Markov chain defined by sequential bargaining on G by $\mathcal{N}$ only puts nonzero probability mass on points in the median closure $M(V_{\mathcal{N}})$ of $V_{\mathcal{N}}$ (see Definition 15 and Figure 5.5 for an example). This is the state space of the Markov chain, and there is a directed edge (i.e., nonzero transition probability) from x to v if there exist $u, w \in V_{\mathcal{N}}$ such that $v = M(x, u, w)$ (where $M(x, u, w)$ is the median of x, u, w by a slight abuse of notation).

Definition 15. *Let $V_{\mathcal{N}} \subseteq \mathcal{S}$ be the set of bliss points of agents in $\mathcal{N}$. A point v is in $M(V_{\mathcal{N}})$ if $v \in V_{\mathcal{N}}$ or if there exists some sequence $(x^1, y^1), (x^2, y^2), \ldots, (x^t, y^t)$ such that every point in every pair in the sequence is in $V_{\mathcal{N}}$ and some $z \in V_{\mathcal{N}}$ s.t.*

$$v = M(x^t, y^t, M(x^{t-1}, y^{t-1}, M(\ldots M(x^1, y^1, z) \ldots)))$$

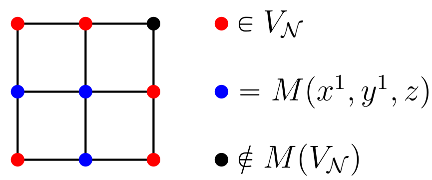


FIGURE 5.5: The median closure of the red points is given by the red and blue points.

Let $G = (\mathcal{S}, E)$ be a median graph, let $\mathcal{N}$ be a set of agents, and let $V_{\mathcal{N}} \subseteq \mathcal{S}$ be the set of bliss points of the agents in $\mathcal{N}$. The Markov chain will have a unique stationary distribution if it is aperiodic and irreducible. To see that the chain is aperiodic, note that for any state $o^t = v \in M(V_{\mathcal{N}})$ of the Markov chain at time t, there is a nonzero probability that $o^{t+1} = v$. This is obvious if $v \in V_{\mathcal{N}}$, as the agent corresponding to that bliss point might be drawn twice in round $t+1$ (remember, agents are drawn independently with replacement from $\mathcal{N}$). If instead $v \notin V_{\mathcal{N}}$, we know by definition of $M(V_{\mathcal{N}})$ that there exist $u, w \in V_{\mathcal{N}}$ and $x \in M(V_{\mathcal{N}})$ such that $v = \text{median}(u, w, x)$. But then $v = \text{median}(u, w, v)$. Clearly we can write $d(u, v) = d(u, v) + d(v, v)$ and $d(w, v) = d(w, v) + d(v, v)$, then the fact that $v = \text{median}(u, w, x)$ implies that $d(u, w) = d(u, v) + d(v, w)$. Taken together, these equalities imply that v is the median chosen in round $t+1$. So in either case, there is some probability that $o^{t+1} = v$. The period of every state is 1, and the chain is aperiodic.

To argue that the chain is irreducible, suppose by contradiction that there exist $t, v \in M(V_{\mathcal{N}})$ such that there is no path from t to v. Then $v \notin V_{\mathcal{N}}$, since all nodes in $V_{\mathcal{N}}$ clearly have an incoming edge from every other node in $M(V_{\mathcal{N}})$. Then by definition there exists some sequence $(x^1, y^1), (x^2, y^2), \ldots, (x^t, y^t) \in V_{\mathcal{N}}$ and some $z \in V_{\mathcal{N}}$ such that

$$v = M(x^t, y^t, M(x^{t-1}, y^{t-1}, M(\ldots M(x^1, y^1, z)\ldots))).$$

Since $z \in V_{\mathcal{N}}$, z must have an incoming edge from t. But then there is a path from t to v. This is a contradiction; the chain must be irreducible as well. Both properties together show that the Markov chain has a unique stationary distribution.

□

Truthfulness of Extensive Forms Finally, we show that sequential deliberation has truth-telling as a Nash equilibrium in its induced extensive form game. Towards this end, we formalize a given round of bargaining as a 2-person non-cooperative game between two players who can choose as a strategy to report any point v on a median graph; the resulting outcome is the median of the two strategy points chosen by the players and the disagreement alternative presented. The payoffs to the players are just the utilities already defined; i.e., the player wishes to minimize the distance from their true bliss point to the outcome point. Call this game the non-cooperative bargaining game (NCBG).

It is important to note the following subtlety: although this procedure generates the same *outcome* as Nash bargaining under Nash's axiomatic characterization, it is different from Nash's model in that we are dropping the assumption of perfect information. It now becomes meaningful to discuss the incentives of the bargaining model, and more specifically, to address the following question: From the perspective of an agent, is truthfully playing one's bliss point as a strategy every time one bargains a Nash Equilibrium of the extensive form game defined by sequential bargaining? In particular, we want to ensure that there is not necessarily an incentive for an agent to misrepresent their true preferences throughout bargaining.

The extensive form game tree defined by non-cooperative bargaining consists of $2T$ alternating levels: Nature draws two agents at random, then the two agents play NCBG and the outcome becomes the disagreement alternative for the next NCBG. The leaves of the tree are a set of points in the median graph; agents want to minimize their expected distance to the final outcome.

Theorem 18. *Sequential NCBG on a median graph has a sub-game perfect Nash equilibrium where every agent truthfully reports their bliss point at all rounds of bargaining.*

Proof. The proof is by backward induction. Let $G = (\mathcal{S}, E)$ be a median graph. In the base case, consider the final round of bargaining between agents u and v with bliss points p_u and p_v and disagreement alternative a. The claim is that u playing p_u and v playing p_v is a Nash equilibrium. By Lemma 12, we can embed G isometrically into a hypercube Q as $\phi : G \to Q$ and consider the bargaining on this embedding. Then for any point z that agent u plays

$$d(p_u, M(z, p_v, a)) = d(\phi(p_u), M(\phi(z), \phi(p_v), \phi(a)))$$

The median on the hypercube is just the bitwise majority, so if u plays some z where for some dimension $\phi_k(z) \neq \phi_k(p_u)$, it can only increase u's distance to the median. So playing p_u is a best response.

For the inductive step, suppose u is at an arbitrary subgame in the game tree with t rounds left, including the current bargain in which u must report a point, and assuming truthful play in all subsequent rounds. Let $\{(x^1, y^1), (x^2, y^2), ..., (x^t, y^t)\}$ represent (x^1, y^1) as the outside alternative and other agent bliss point against which u must bargain, (x^2, y^2) as the bliss points of the agents drawn in the next round, and so on. We want to show that it is a best response for agent u to choose p_u, i.e., to truthfully represent her bliss point. Define

$$M_u^t := M(x^t, y^t, M(x^{t-1}, y^{t-1}, M(\ldots M(x^1, y^1, p_u)\ldots)))$$

where $M(\cdot)$ indicates the median, guaranteed to exist and be unique on the median graph. Also, for any point z, similarly define

$$M_z^t := M(x^t, y^t, M(x^{t-1}, y^{t-1}, M(\ldots M(x^1, y^1, z)\ldots)))$$

Suppose by contradiction that p_u is not a best response for agent u, then there must exist $z \neq p_u$ and some $r > 0$ such that $d(p_u, M_u^r) > d(p_u, M_z^r)$. We embed G isometrically into a hypercube Q as $\phi : G \to Q$. Then by the isometry property, $d(\phi(p_u), \phi(M_u^r)) > d(\phi(p_u), \phi(M_z^r))$. By the proof of Corollary 4, we can pretend the process occurs on the hypercube.

Consider some dimension k. If $\phi_k(x^t) = \phi_k(y^t)$ for some $t \leqslant r$, then this point becomes the median in that dimension, so the median becomes independent of $\phi_k(w)$, where w is the initial report of agent u. Till that time, the bargaining outcome in that dimension is the same as $\phi_k(w)$. In either case, for all times $t \leqslant r$ and all k, we have:

$$|\phi_k(p_u) - \phi_k(M_u^t)| \leqslant |\phi_k(p_u) - \phi_k(M_z^t)|$$

Summing this up over all dimensions, $d(\phi(p_u), \phi(M_u^t)) \leqslant d(\phi(p_u), \phi(M_z^t))$, which is a contradiction. Therefore, p_u was a best response for agent u[1]. Therefore, every agent truthfully reporting their bliss points at all rounds is a subgame perfect Nash equilibrium of Sequential NCBG. □

5.5 General Metric Spaces

We now work in the very general setting that the set $\mathcal{S}$ of alternatives are points in a metric space equipped with a distance function $d(\cdot)$ satisfying the assumptions of a metric. Again, we assume each agent $u \in \mathcal{N}$ has a bliss point $p_u \in \mathcal{S}$. An agent's disutility for an alternative $a \in \mathcal{S}$ is simply $d(p_u, a)$. Let $a^* \in \mathcal{S}$ be the social cost minimizer, i.e., the generalized median. For convenience, let $Z_a = d(a, a^*)$ for $a \in \mathcal{S}$. By a slight abuse of notation, let $Z_i = d(p_i, a^*)$ for $i \in \mathcal{N}$, i.e., the distance from agent i's bliss point to a^*.

Theorem 19. *The Distortion of sequential deliberation is at most 3 when the space of alternatives and bliss points lies in some metric, and this bound is tight.*

Proof. We will only use the assumption that $\mathcal{B}(u, v, a)$ finds a Pareto efficient point for u and v, so rather than taking an expectation over the choice of the disagreement alternative, we take the worst case. Let a^* be the social cost minimizer with social cost OPT. We can write then write expected worst case social cost of a step of

[1] It is important to note that we are *not* assuming that agent u will not bargain again in the subgame; there are no restrictions on the values of $\{(x^1, y^1), (x^2, y^2), ..., (x^t, y^t)\}$.

deliberation as:

$$\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}}\frac{1}{|\mathcal{N}|^2}\max_{a\in\mathcal{S}}\sum_{k\in\mathcal{N}} d(\mathcal{B}(i,j,a),p_k)$$

$$\leqslant \sum_{i,j\in\mathcal{N}}\frac{1}{|\mathcal{N}|^2}\max_{a\in\mathcal{S}}\sum_{k\in\mathcal{N}} d(\mathcal{B}(i,j,a),a^*) + Z_k = OPT + \sum_{i,j\in\mathcal{N}}\frac{1}{|\mathcal{N}|}\max_{a\in\mathcal{S}}\left(d(\mathcal{B}(i,j,a),a^*)\right)$$

$$\leqslant OPT + \sum_{i,j\in\mathcal{N}}\frac{1}{2|\mathcal{N}|}\max_{a\in\mathcal{S}}\left(d(\mathcal{B}(i,j,a),p_i) + Z_i + d(\mathcal{B}(i,j,a),p_j) + Z_j\right)$$

$$\leqslant OPT + \sum_{i,j\in\mathcal{N}}\frac{1}{2|\mathcal{N}|}\left(2Z_i + 2Z_j\right)\ \left[\text{Pareto Efficiency of } \mathcal{B}(\cdot)\right]$$

$$= 3OPT$$

Since this holds for the worst case choice of disagreement alternative, it holds over the whole sequential process. The tight example is a weighted graph that can be constructed as follows: start with a star with a single agent's bliss point located at each leaf of the star (i.e., an $|N|$-star) where every edge in the star has weight (or distance) 1. Now, for every pair of agents, add a vertex connected to the bliss points of both of the agents with weight $1-\epsilon$ for $\epsilon > 0$. Then as $|N| \to \infty$, every round of Nash bargaining selects one of these pairwise vertices. But the central vertex of the star has social cost $|N|$ whereas the pairwise vertices all have social cost approaching $3|N|$ as $\epsilon \to 0$. □

A simple baseline algorithm for our problem is Random Dictatorship: Output the bliss point p_u of a random agent $u \in \mathcal{N}$. The Distortion of Random Dictatorship is at most 2, and Theorem 13 shows this bound is tight. At first glance, it seems that Random Dictatorship is superior to sequential deliberation for general metric spaces. However, this is no longer true when one considers the second moment, or the expected squared social cost of sequential deliberation. We study the Distortion on this metric. define:

$$\text{Squared-Distortion} = \frac{\mathbb{E}[(SC(a))^2]}{(SC(a^*))^2}$$

where the expectation is over the set of outcomes a produced by Sequential Deliberation.

It is easy to see that Random Dictatorship has an unbounded Squared-Distortion. Consider the simple graph with two nodes, a fraction f of the agents on one node and $1 - f$ on the other. Let $f < 1/2$. Then the expected squared social cost of Random Dictatorship is just $f(1-f)^2 + (1-f)f^2$ whereas the optimal solution has squared social cost f^2, so Random Dictatorship has Squared-Distortion $(1-f)^2/f + (1-f)$, which is unbounded as $f \to 0$.

On the other hand, we show that Sequential deliberation has constant Squared-Distortion even for general metric spaces. One way to interpret this result is that the deviation of the social cost around its mean falls off quadratically instead of linearly, which means the outcome of sequential deliberation is well-concentrated around its mean value[2]. We will need the following technical lemma to characterize the outcome of a round of bargaining.

Lemma 16. *For all $i, j, u \in \mathcal{N}$ and $a \in \mathcal{S}$ we have that*

$$d(\mathcal{B}(i,j,a), p_u) \leqslant Z_u + 2\min(Z_i, Z_j) + \min(Z_a, \max(Z_i, Z_j))$$

Proof. Assume w.l.o.g. that $Z_i \leqslant Z_j$. Then the lemma statement reduces to

$$d(\mathcal{B}(i,j,a), p_u) \leqslant Z_u + 2Z_i + \min(Z_a, Z_j) \tag{5.4}$$

Recall that Nash bargaining asserts that given a disagreement alternative a, agents i and j choose that alternative $o \in \mathcal{S}$ that maximizes:

$$\text{Nash product } = (d(p_i,a) - d(p_i,o)) \times (d(p_j,a) - d(p_j,o))$$

[2] See also recent work by [140] that considers minimizing the variance of randomized truthful mechanisms.

Maximizing this on a general metric yields that $d(\mathcal{B}(i,j,a)$ will be chosen on the p_i to p_j shortest path (that is, the Pareto efficient frontier) at a distance of $\frac{d(p_i,p_j)}{2} + \frac{d(p_i,a)-d(p_j,a)}{2}$ from p_i. Therefore, we have that:

$$d(\mathcal{B}(i,j,a),p_i) = \frac{d(p_i,p_j)}{2} + \frac{d(p_i,a) - d(p_j,a)}{2} \tag{5.5}$$

Now we can show equation 5.4 by repeatedly using the triangle inequality. We have

$$\begin{aligned}
&d(\mathcal{B}(i,j,a),p_u) \\
&\leqslant d(p_i,p_u) + d(\mathcal{B}(i,j,a),p_i) \\
&= d(p_i,p_u) + \frac{d(p_i,p_j)}{2} + \frac{d(p_i,a) - d(p_j,a)}{2} \; \big[\text{equation 5.5}\big] \\
&\leqslant d(p_i,a^*) + d(p_u,a^*) + \frac{d(p_i,a^*) + d(p_j,a^*) + d(p_i,a^*) + d(a,a^*)}{2} - \frac{d(p_j,a)}{2} \\
&= Z_u + 2Z_i + \frac{Z_j + Z_a - d(p_j,a)}{2}
\end{aligned}$$

Note that we can apply the triangle inequality again to get two bounds, since $Z_j \leqslant d(p_j,a) + Z_a$ and $Z_a \leqslant d(p_j,a) + Z_j$. Since both bounds must hold, we have that $d(\mathcal{B}(i,j,a),p_u) \leqslant Z_u + 2Z_i + \min(Z_a, Z_j)$. □

Now we can bound the Squared-Distortion of sequential deliberation.

Theorem 20. *The Squared-Distortion of sequential deliberation for is at most* 41 *when the space of alternatives and bliss points lies in some metric.*

Let OPT be the squared social cost of a^*. Note that OPT does not depend on T. In particular, $OPT = \left(\sum_{i\in\mathcal{N}} Z_i\right)^2 = \sum_{i,j\in\mathcal{N}} Z_i Z_j$.

Let ALG be the expected squared social cost of sequential deliberation with T steps, where the disagreement alternative $a \in \mathcal{S}$ is used in the final round of bargaining. i and j are the last two agents to bargain, and p_k is the bliss point of an arbitrary agent. Using Lemma 16 we can write ALG as

$$
\begin{aligned}
ALG &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \frac{1}{|\mathcal{N}|^2}\left(\sum_{k\in\mathcal{N}} d(\mathcal{B}(i,j,a),p_k)\right)^2 \\
&\leqslant \frac{1}{|\mathcal{N}|^2}\sum_{i,j\in\mathcal{N}}\left(\sum_{k\in\mathcal{N}} Z_k + 2\min(Z_i,Z_j) + \min(Z_a,\max(Z_i,Z_j))\right)^2 \\
&= \frac{1}{|\mathcal{N}|^2}\sum_{i,j\in\mathcal{N}}\left(2|\mathcal{N}|\min(Z_i,Z_j) + |\mathcal{N}|\min(Z_a,\max(Z_i,Z_j)) + \left(\sum_{k\in\mathcal{N}} Z_k\right)\right)^2
\end{aligned}
$$

Now we expand the square and analyze term by term, using the facts that $\min(x,y)^2 \leqslant x\times y$ and $\max(x,y)\leqslant x+y$.

$$
\begin{aligned}
&= \frac{1}{|\mathcal{N}|^2}\sum_{i,j\in\mathcal{N}}\Bigg(4|\mathcal{N}|^2\min(Z_i,Z_j)^2 + |\mathcal{N}|^2\min(Z_a,\max(Z_i,Z_j))^2 + \left(\sum_{k\in\mathcal{N}} Z_k\right)^2 \\
&\qquad + 4|\mathcal{N}|^2\min(Z_i,Z_j)\min(Z_a,\max(Z_i,Z_j)) + 4|\mathcal{N}|\min(Z_i,Z_j)\left(\sum_{k\in\mathcal{N}} Z_k\right) \\
&\qquad + 2|\mathcal{N}|\min(Z_a,\max(Z_i,Z_j))\left(\sum_{k\in\mathcal{N}} Z_k\right)\Bigg) \\
&\leqslant \frac{1}{|\mathcal{N}|^2}\sum_{i,j\in\mathcal{N}}\Bigg(4|\mathcal{N}|^2 Z_iZ_j + |\mathcal{N}|^2\min(Z_a,\max(Z_i,Z_j))^2 + OPT \\
&\qquad + 4|\mathcal{N}|^2 Z_iZ_a + 4|\mathcal{N}|Z_i\left(\sum_{k\in\mathcal{N}} Z_k\right) + 2|\mathcal{N}|Z_a\left(\sum_{k\in\mathcal{N}} Z_k\right)\Bigg)
\end{aligned}
$$

We can trivially sum out the terms not involving a, leaving us with inequality 5.6.

$$
ALG \leqslant 9OPT + \sum_{i,j\in\mathcal{N}}\left(Z_a(Z_i+Z_j) + 4Z_iZ_a + \frac{2}{|\mathcal{N}|}Z_a\left(\sum_{k\in\mathcal{N}} Z_k\right)\right) \tag{5.6}
$$

Note that the triangle inequality implies that for any $i\in\mathcal{N}$, $Z_a \leqslant d(a,p_i) + Z_i$.

Applying this repeatedly in inequality 5.6 and simplifying yields:

$$ALG \leqslant 9OPT + \sum_{i,j\in\mathcal{N}} \left(6Z_iZ_j + 5d(a,p_j)Z_i + d(a,p_i)Z_j + \frac{2}{|\mathcal{N}|}(d(a,p_i)+Z_i)\left(\sum_{k\in\mathcal{N}} Z_k\right)\right)$$

$$= 17OPT + \sum_{i\in\mathcal{N}} \left(2d(a,p_i)\left(\sum_{k\in\mathcal{N}} Z_k\right) + \sum_{j\in\mathcal{N}} 5d(a,p_j)Z_i + d(a,p_i)Z_j\right)$$

Recall that Theorem 19 implies that $\sum_{i\in\mathcal{N}} d(a,p_i) \leqslant 3\sum_{i\in\mathcal{N}} Z_i$.

$$ALG \leqslant 17OPT + 6OPT + 15OPT + 3OPT = 41OPT$$

5.6 Conclusion and Open Directions

In this chapter, we took a first step to developing a theory around practical deliberation schemes. We employed the implicit utilitarian model to evaluate the properties of our deliberation mechanism for approximating the utilitarian optimum under limited information. We defined a general protocol for deliberation, and demonstrated a host of strong and desirable properties under the assumption that the space of possible outcomes form a median graph, along with more general guarantees when the space of possible outcomes are merely constrained to be metric.

We suggest several future directions. First, we do not have a general characterization of the Distortion of sequential deliberation for metric spaces. We have shown that for general metric spaces there is a small but pessimistic bound on the Distortion of 3, but that for specific metric spaces the Distortion may be much lower. We do not have a complete characterization of what separates these good and bad regimes.

More broadly, an interesting question is extending our work to take opinion dynamics into account, *i.e.*, proving stronger guarantees if we assume that when two agents deliberate, each agent's opinion moves slightly towards the other agent's opinion and the outside alternative. Furthermore, though we have shown that all agents deliberating at the same time does not improve on dictatorship, it is not clear how to extend our results to more than two agents negotiating at the same time. This runs into the challenges in understanding and modeling multiplayer bargaining [90, 110, 26].

6

Metric Implicit Utilitarian Voting with Constant Sample Complexity

In the previous chapter, we explored a novel connection between the cooperative game theoretic solution concept of Nash bargaining and social choice protocols in a metric implicit utilitarian model. However, Nash bargaining itself requires agents to reason over their cardinal preferences or utilities. In this chapter, we continue investigating social choice in the metric implicit utilitarian model by defining purely ordinal mechanisms, in the sense that they require neither communication of nor reasoning over cardinal preferences. We focus on designing practical and lightweight mechanisms that use at most a constant number of simple ordinal queries to provide guarantees with respect to Distortion and Squared Distortion.

Acknowledgements. These results are published [64], and are joint work with Ashish Goel, Kamesh Munagala, and Nina Prabhu.

6.1 Introduction

Consider the social choice problem of deciding on an allocation of public tax dollars to public projects. This is a voting problem over budgets. Clearly, the number of voters in such situations can be large. More interestingly, unlike in traditional social choice theory, there is no reason to believe that the number of alternatives (budgets) is small. It is therefore unreasonable to assume that we can elicit full ordinal preferences over alternatives from every agent. For a voting mechanism to be practical in such a setting, one would ideally like it to require only an absolute constant number of simple queries, regardless of the number of voters and alternatives. We call this property *constant sample complexity*, and we explore mechanisms of this sort in this chapter.

We define our model more formally in Section 6.1.3, but at a high level, we have a set N of agents (or voters) and a set of alternatives $\mathcal{S}$, from which we must choose a single outcome. We assume that N and $\mathcal{S}$ are both large, and that eliciting the full ordinal rankings may be prohibitively difficult. Instead, we work with an ordinal query model, and a constant sample complexity mechanism uses only a constant number of these queries.

- Top-k Query. "What are your k favorite alternatives, in order?" (We call a top-1 query a favorite query); and
- Comparison Query. "Which of two given alternatives do you prefer?"

Query models are not just of theoretical interest. They can be used to reduce cognitive overload in voting. For example, in the context of Participatory Budgeting [1], the space of possible budget allocations is large, and one mechanism is to ask voters to compare two proposed budgets. Similarly, in a context like transportation policy for a city, a single alternative can be an entire transportation plan. In such examples, not only are there many alternatives, but it may be infeasible to expect voters to compare more than two alternatives at the same time. We stress that constant sample complexity is particularly important in settings where there may be a large number of possibly complex alternatives.

To evaluate the quality of our mechanisms, we adopt the implicit utilitarian perspective with metric constraints [31, 46, 8, 79, 7, 67]. That is, we assume that agents have cardinal costs over alternatives, and these costs are constrained to be metric, but asking agents to work with or report cardinal costs is impractical or impossible. We want to design social choice mechanisms to minimize the total social cost, but our mechanism is constrained only to use ordinal queries, that is, those that can be answered given a total order over alternatives. We therefore measure the efficiency of a mechanism as its *Distortion* (see Section 6.1.3), the worst case approximation to the total social cost.

6.1.1 Results

The starting point for our inquiry is the constant sample complexity *Random Dictatorship* mechanism. The algorithm asks a single favorite query from an agent chosen uniformly at random, and has a tight Distortion bound of 3 [8]. In this chapter, we provide two new mechanisms (Random Referee and Random Oligarchy) that improve on this simple baseline in three different ways, outlined in each of our three technical sections. We hope that our work inspires future research on similarly lightweight mechanisms for social choice in large decision spaces.

Random Referee: Comparison Queries and Squared Distortion. In Section 6.2, we show that one disadvantage of Random Dictatorship lies not in its Distortion, but in its variance. For randomized mechanisms, Distortion is measured as the expected approximation to the first moment of social cost. However, in many social choice problems, we might want a bound on the risk associated with a given mechanism. We capture this via *Squared Distortion*, as suggested in [65]. The Squared Distortion (see Definition 17 in Section 6.1.3) is the expected approximation to the second moment of social cost. A mechanism with constant Squared Distortion has both constant Distortion and constant coefficient of variation of the Distortion.

We show that mechanisms using only top-k queries (including Random Dictatorship) have Squared Distortion $\Omega(|\mathcal{S}|)$. This motivates us to expand our query model to incorporate information about the relative preferences of agents between alternatives, *i.e.* use comparison queries. We define a novel mechanism called *Random Referee* (RR) that uses a random voter as a referee to compare the favorite alternatives of two other random voters (see Definition 18 in Section 6.2). Our main result in Section 6.2 is Theorem 22: The Squared Distortion of RR is at most 21. This also immediately implies that the Distortion of RR is at most 4.583.

Random Referee: Euclidean Plane. In Section 6.3, we show that top-k only mechanisms (again, including Random Dictatorship) achieve their worst case Distortion even on benign metrics such as low dimensional Euclidean spaces. We analyze a special case on the Euclidean plane and prove that the Distortion of Random Referee beats that of any top-k only mechanism. While the improvement we prove in Section 6.2 is quantitatively small, it is qualitatively interesting: we demonstrate that by using a *single* comparison query, Random Referee can exploit the structure of the metric space to improve Distortion, whereas Random Dictatorship or any other top-k only mechanism cannot. We conjecture this result extends to Euclidean spaces in any dimension, and present some evidence to support this conjecture in Section 6.5.

Random Oligarchy: Favorite Only Mechanisms. In Section 6.4, we consider mechanisms that are restricted to favorite queries and show that constant complexity mechanisms are nearly optimal. We present a mechanism that uses only three favorite queries that has Distortion at most 3 for arbitrary $|\mathcal{S}|$; however, it also has Distortion for small $|\mathcal{S}|$ that improves upon the best known favorite only mechanism from [85] that uses at most $\min(|N|+1, |\mathcal{S}|+1)$ favorite queries. Comparing with a lower bound for favorite only mechanisms, Random Oligarchy has nearly optimal distortion and constant sample complexity. Though this mechanism does not have constant Squared Distortion like Random Referee, we present it to demonstrate again the surprising power of constant sample complexity randomized social choice mechanisms in general, and of queries to just three voters in particular.

Techniques. We use different techniques to prove our different positive results. The proof of Squared Distortion (Theorem 22 in Section 6.2) relies heavily on Lemma 17, in which we prove (essentially) that Random Referee chooses a low social cost alternative as long as at least two of the three agents chosen at random are near the social optimum.

The proof of Distortion for Euclidean spaces (Theorem 25 in Section 6.3) is the most technical result. We show that we can upper bound the Distortion of a mechanism by the worst case "pessimistic distortion," of just a constant size tuple of points, where the "pessimistic distortion" considers all permutations of the points as participating in Random Referee and allows OPT to choose the optimal point on just this tuple. This allows us to employ a computer assisted analysis by arguing that if a high Distortion instance exists, we can detect it as an instance with a small constant number of points on a sufficiently fine (but finite) grid in the Euclidean plane. This approach may be of independent interest for providing tighter Distortion bounds for mechanisms in specific structured metric spaces.

6.1.2 Related Work

Distortion of Randomized Social Choice Mechanisms in Metrics. The Distortion of randomized social choice mechanisms in metrics has been studied in [31, 8, 79, 85]. Of particular interest to us are the Random Dictatorship mechanism that uses a single favorite query and the 2-Agree mechanism [85] that uses at most most $\min(|N| + 1, |\mathcal{S}| + 1)$ favorite queries. Random Dictatorship has an upper bound on Distortion of 3 [8], and 2-Agree provides a strong guarantee on Distortion when $|\mathcal{S}|$ is small (better than Random Dictatorship for $|\mathcal{S}| \leqslant 6$). There is ongoing work on analyzing the Distortion of randomized ordinal mechanisms for other classic optimization problems like graph optimization [2] and facility location [9].

Squared Distortion and Variance. We are aware of two papers in mechanism design that consider the variance of mechanisms for facility location on the real line [140] and kidney exchange [62]. Our work is more related to the former, but is not restricted to the real line, and does not focus on characterizing the tradeoff between welfare and variance. Using Squared Distortion as a proxy for risk was introduced in [65] along with the sequential deliberation protocol. Unlike sequential deliberation, Random Referee makes a constant number of ordinal queries. The most important baseline for Squared Distortion is the deterministic Copeland rule, which has Distortion 5 [7] and therefore Squared Distortion 25. However, Copeland requires the communication of $\Omega(|N||\mathcal{S}|\log(|\mathcal{S}|))$ bits [55], essentially the entire preference profile. Our Random Referee mechanism has constant sample complexity, and has better bounds on Squared Distortion (21) and Distortion (4.583) than the Copeland mechanism.

Communication Complexity. For a survey on the complexity of eliciting ordinal preferences to implement social choice rules, we refer the interested reader to [34]. Of particular interest to us is [55], in which the authors comprehensively characterize the *communication complexity* (in terms of the number of bits communicated) of common deterministic voting rules. A favorite query requires $O(\log(|\mathcal{S}|))$ bits of communication, so our mechanisms have constant sample complexity, but logarithmic communication complexity. [32] and [42] design social choice mechanisms with low communication complexity when there are a small number of voters, but potentially a large number of alternatives. All of our mechanisms have guarantees that are independent of the number of voters.

Strategic Incentives. We do not consider truthfulness chapter, and we do not use the term mechanism to imply any such property. While strategic incentives are not the focus of this work, we note that any truthful mechanism must have Distortion at least 3 [67]. Random Dictatorship has a Distortion of 3, and is therefore in some sense optimal among exactly truthful mechanisms. Other works suggest that truthfulness is also incompatible with the weaker notion of Pareto efficiency in randomized social choice [33, 15]. Still other authors have considered the problem of truthful welfare maximization under range voting [69] and threshold voting [27].

6.1.3 Preliminaries

We have a set N of agents (or voters) and a set $\mathcal{S}$ of alternatives, from which we must choose a single outcome. For each agent $u \in N$ and alternative $a \in \mathcal{S}$, there is some underlying dis-utility $d_u(a)$. Let $p_u = \text{argmin}_{a \in \mathcal{S}} d_u(a)$, that is, p_u is the most preferred alternative for agent u. Ordinal preferences are specified by a total order σ_u consistent with these dis-utilities (i.e., an alternative is ranked above another only if it has lower dis-utility). A preference profile $\sigma^{(N)}$ specifies the ordinal preferences of all agents. A deterministic social choice rule is a function f that maps a preference profile $\sigma^{(N)}$ to an alternative $a \in \mathcal{S}$. A randomized social choice rule maps a preference profile $\sigma^{(N)}$ to a distribution over $\mathcal{S}$.

We consider mechanisms that implement a randomized social choice rule using a constant number of queries of two types. A *top-k query* asks an agent $u \in N$ for the first k preferred alternatives according to the order σ_u (ties can be broken arbitrarily). We refer to a top-1 query as a *favorite query*, that asks an agent $u \in N$ for her most preferred alternative p_u. A *top-k only mechanism* f uses only top-k queries, for some constant k (constant with respect to $|N|$ and $|\mathcal{S}|$). Most of our lower bounds or impossibilities will be for any top-k only mechanism (for constant k), whereas our positive results will only need favorite and comparison queries. A *comparison query* with alternatives $a \in \mathcal{S}$ and $b \in \mathcal{S}$ asks an agent $u \in N$ for $\text{argmin}_{x \in \{a,b\}} d_u(x)$.

We use the term mechanism to clarify that our algorithms are in a query model. However, it is important to note that mechanisms so defined are still randomized social choice rules in the formal sense as long as they do not make queries based on exogenous information (e.g., names of participants). Our mechanisms will in fact be randomized social choice rules, and thus can be appropriately compared to other such rules in the literature that do not explicitly use a query model. By using the term mechanism, we do not mean to imply any strategic properties.

Distortion and Sample Complexity. We measure the quality of an alternative $a \in \mathcal{S}$ by its *social cost*, given by $SC(a) = \sum_{u \in N} d_u(a)$. Let $a^* \in \mathcal{S}$ be the minimizer of social cost. We define the commonly studied approximation factor called Distortion [142], which measures the worst case approximation to the optimal social cost of a given mechanism. We use the expected social cost if a is the outcome of a randomized mechanism, and we seek to minimize Distortion.

Definition 16. *The Distortion of an alternative a is $Distortion(a) = \frac{SC(a)}{SC(a^*)}$. The Distortion of a social choice mechanism f is*

$$Distortion(f) = \sup_{\{d_u(a)\}} \mathbb{E}_{f\left(\sigma^{(N)}\right)} Distortion(a)$$

where $\sigma^{(N)}$ is a preference profile consistent with $\{d_u(a)\}$.

We assume that $\mathcal{S}$ is a set of points in a metric space such that dis-utility can be measured by the distance from an agent. Specifically, we assume there is a distance function $d : (N \cup \mathcal{S}) \times (N \cup \mathcal{S}) \to \mathbb{R}_{\geqslant 0}$ satisfying the triangle inequality such that $d_u(a) = d(u, a)$. The metric assumption is common in the implicit utilitarian literature [8, 79, 65, 85, 46, 7, 67]. It is also a natural assumption for capturing social choice problems for which there is a natural notion of distance between alternatives. For example, in our original motivating example of public budgets, there is are natural notions of distance between alternatives in terms of dollars.

We do not assume access to $\sigma^{(N)}$ directly, which may be prohibitively difficult to elicit when there are many alternatives. Instead, we work with a query model. The queries are ordinal in the sense that they can be answered given only the information in $\sigma^{(N)}$. A mechanism f has *constant sample complexity* if there is an absolute constant c such that for all $\mathcal{S}$, N, and $\sigma^{(N)}$, f can be implemented using at most c queries. In this chapter, we consider top-k (and the special case of favorite) and comparison queries, and explore mechanisms with constant sample complexity.

Squared Distortion. It is easy to see that randomization is necessary for constant sample complexity mechanisms to achieve constant Distortion. As N grows large, any deterministic mechanism with constant sample complexity deterministically ignores (asks no queries of and receives no information from) an arbitrarily large fraction of N. An adversary can therefore place an alternative with 0 dis-utility for arbitrarily many agents; this gives a lower bound for Distortion approaching N as N becomes large.

This naturally leads us to ask: If we look at the distribution of outcomes produced by the mechanism, is this distribution well behaved? Following [65], we capture this notion via *Squared Distortion*: essentially the approximation to the optimal second moment of social cost.

Definition 17. *The Squared Distortion of an alternative* $a \in \mathcal{S}$ *is* $Distortion^2(a) = \left(\frac{SC(a)}{SC(a^*)}\right)^2$. *The Distortion of a social choice mechanism* f *is*

$$Distortion^2(f) = \sup_{\{d_u(a)\}} \mathbb{E}_{f\left(\sigma^{(N)}\right)} Distortion^2(a)$$

where $\sigma^{(N)}$ *is a preference profile consistent with* $\{d_u(a)\}$.

Note that a mechanism with constant Squared Distortion has both constant Distortion (by Jensen's inequality), and constant coefficient of variation. One way to interpret having constant Squared Distortion is that the deviation of the social cost around its mean falls off quadratically instead of linearly, which means that the social cost of such a mechanism is well concentrated around its mean value. We note that this interpretation gives randomized social choice mechanisms with constant Squared Distortion an interesting application in candidate selection. In particular, one can imagine running such a mechanism (like our Random Referee) to generate a candidate list on which one can use a deterministic (but potentially complex) voting mechanism like Copeland.

A related approach to understanding the distributional properties of a randomized mechanism is to characterize the tradeoff between Distortion (approximation to the first moment) and the variance of randomized mechanisms [140]. Our specific goal in this chapter is to develop a mechanism that achieves constant Distortion *and* constant variance, and this combination is captured by having constant Squared Distortion. We leave characterizing the exact tradeoff between the quantities as an interesting open direction.

6.2 Random Referee and Squared Distortion

Our first result is Theorem 21: Mechanisms that only elicit top-k preferences, for constant k, must necessarily have Squared Distortion that grows linearly in the size of the instance. This holds even for mechanisms that elicit the top-k preferences of all of the voters, mechanisms which would not have constant sample complexity.

Theorem 21. *Any top-k only social choice mechanism has Squared Distortion* $\Omega(|\mathcal{S}|)$.

Proof. We want to construct a single ordinal profile over top-k preferences such that for all randomized social choice mechanisms, there is some instantiation of these top-k preferences in a metric space on which the mechanism has Squared Distortion $\Omega(|\mathcal{S}|)$. The top-k ordinal profile is simple: Each agent $u \in N$ has completely unique top-k preferred feasible alternatives $S_u \subseteq \mathcal{S}$. There are no other alternatives, so $|\mathcal{S}| = k|N| = \Theta(|N|)$ (note that a top-k only mechanism must have constant k).

Any randomized social choice mechanism must choose a distribution on $\mathcal{S}$, so in particular, there will be some $u^* \in N$ such that the mechanism puts at least $1/|N|$ probability mass on the top-k preferred alternatives of u^*. Given this, consider a metric space, consistent with the top-k preference profile, where S_u^* forms a small clique well separated from all of the other alternatives which form a large clique. More formally, let $0 < \epsilon \leqslant 1/|N|$. All pairwise distances are ϵ except those between the small and large clique. Instead, $d(u, a) = 1$ for all $u \neq u^*, a \in S_{u^*}$, and similarly $d(u^*, a) = 1$ for all $a \in \mathcal{S} \backslash S_{u^*}$.

Then the Squared social cost of choosing $a \in S_{u^*}$ is simply $(|N|-1)^2$, whereas the optimal solution chooses any $a \in \mathcal{S} \backslash S_{u^*}$ for Squared social cost of $(1 + \epsilon(|N| - 1))^2 \leqslant (1 + \epsilon|N|)^2$. Since the mechanism chooses $a \in S_{u^*}$ with probability at least $1/|N|$, we can lower bound the Squared Distortion of any top-k only mechanism f as follows.

$$\begin{aligned} Distortion^2(f) &\geqslant \frac{1}{|N|}\frac{(|N|-1)^2}{(1+\epsilon|N|)^2} \\ &\geqslant \frac{(|N|-1)^2}{4|N|} = \Omega(|\mathcal{S}|) \end{aligned}$$

□

The problem with top-k only mechanisms, exploited in the proof of Theorem 21, is that they treat agents as indifferent between their $(k+1)$st favorite alternative and their least favorite alternative. This motivates the expansion of our query model to include *comparison queries*. Recall that a comparison query with alternatives $a \in \mathcal{S}$ and $b \in \mathcal{S}$ asks an agent $u \in N$ for $\text{argmin}_{x \in \{a,b\}} d_u(x)$. We use a single comparison query in our Random Referee mechanism.

Definition 18. *The Random Referee (RR) mechanism samples three agents $u, v, w \in N$ independently and uniformly at random with replacement. u and v are asked for their favorite feasible alternatives p_u and p_v in $\mathcal{S}$, and then w is asked to compare p_u and p_v. Output whichever of the two alternatives w prefers.*

Upper Bound for Random Referee. Our main result in this section is Theorem 22: The Squared Distortion of RR is at most 21. This also implies that the Distortion of RR is at most 4.583. As far as we are aware, no randomized mechanism has lower Squared Distortion in general.

We will need the following technical lemma. Let $a^* \in \mathcal{S}$ be the social cost minimizer. The basic intuition behind Random Referee is that it will choose a low social cost alternative as long as any two out of the three agents selected are near the optimal alternative a^*. Lemma 17 makes this intuition formal. For convenience, let $Z_u = d(u, a^*)$ for $u \in N$, i.e., the dis-utility of a^* for agent u. Let

$C(\{u,v\},w) = \text{argmin}_{y \in \{p_u,p_v\}} d(w,y)$, that is, the alternative Random Referee outputs when u and v are selected to propose alternatives and w chooses between them.

Lemma 17. *For all $u, v, w, x \in N$,*

$$d(C(\{u,v\},w),x) \leqslant Z_x + 2\min \begin{cases} \max(Z_u, Z_v), \\ Z_w + \min(Z_u, Z_v) \end{cases} .$$

Proof. One should think of $u, v \in N$ as the agents drawn to present their favorite alternative, w as the referee, and x as the agent from whom we measure the dis-utility of the resulting outcome. The triangle inequality implies that

$$d(C(\{u,v\},w),x) \leqslant Z_x + d(a^*, C(\{u,v\},w)).$$

We give two separate upper bounds on $d(a^*, C(\{u,v\},w))$, thus yielding the min. We will frequently use the fact that for all $u \in N$, $d(u,a^*) \leqslant d(p_u,u) + d(u,a^*) \leqslant 2Z_u$. This follows from the the definition of p_u: the favorite alternative of u, and thus no greater in distance from u than a^*. The first bound is straightforward: $C(\{u,v\},w) \in \{p_u,p_v\}$ by definition of Random Referee, so

$$\begin{aligned} d(a^*, C(\{u,v\},w)) &\leqslant \max\left(d(a^*,p_u), d(a^*,p_v)\right) \\ &\leqslant 2\max\left(Z_u, Z_v\right) \end{aligned}$$

In a sense, this bound concerns the situation where both u and v are near a^*, but w is far away from a^*. Now we argue for the second bound. Suppose without loss of generality that $Z_u \leqslant Z_v$. w chooses either p_u or p_v. If w chooses p_u, then $d(a^*, C(\{u,v\},w)) = d(a^*,p_u) \leqslant 2d(a^*,u) = 2\min(Z_u,Z_v)$, and so the bound holds.

If w chooses p_v, then $d(w, p_v) \leqslant d(w, p_u)$, which implies

$$\begin{aligned} d(a^*, C(\{u,v\}, w)) &= d(a^*, p_v) \\ &\leqslant d(a^*, w) + d(w, p_u) \\ &\leqslant Z_w + d(w, a^*) + d(a^*, p_u) \\ &\leqslant 2Z_w + 2Z_u \\ &= 2\left(Z_w + \min(Z_u, Z_v)\right) \end{aligned}$$

and again the bound holds. In either case, $d(a^*, C(\{u,v\}, w))$ is at most $2\left(Z_w + \min(Z_u, Z_v)\right)$. Intuitively, this bound concerns the situation where w and u are close to a^*, but v is far away from a^*. Taking the better of this bound with the $d(a^*, C(\{u,v\}, w)) \leqslant 2\max(Z_u, Z_v)$ bound through the min and factoring out the 2 yields the lemma. □

Using Lemma 17, we can upper bound the Squared Distortion of Random Referee by 21. This is in contrast to the $\Omega(|\mathcal{S}|)$ Squared Distortion of any favorite only mechanism (see Theorem 21).

Theorem 22. *The Squared Distortion of Random Referee is at most 21.*

Proof. Let OPT be the optimal squared social cost, that is, the squared social cost of a^*, where $a^* \in \mathcal{S}$ is the social cost minimizer. Recall that $Z_u = d(u, a^*)$. Then

$$OPT = \left(\sum_{u \in N} Z_u\right)^2 = \sum_{u \in N} \sum_{v \in N} Z_u Z_v.$$

Let ALG be the expected squared social cost of Random Referee. The expectation can be written out as

$$ALG = \frac{1}{|N|^3} \sum_{u,v,w \in N} \left(\sum_{x \in N} d(C(\{u,v\}, w), x)\right)^2.$$

Let $\alpha = \min\left(\max(Z_u, Z_v), Z_w + \min(Z_u, Z_v)\right)$. We apply Lemma 17 and simplify.

$$ALG \leqslant \frac{1}{|N|^3} \sum_{u,v,w \in N} \left(\sum_{x \in N} (Z_x + 2\alpha) \right)^2$$

Noting that α does not depend on x, we can expand the square and simplify to find

$$ALG \leqslant OPT + \frac{4}{|N|^3} \sum_{u,v,w \in N} \left(\alpha^2 |N|^2 + \alpha |N| \sum_{x \in N} Z_x \right).$$

Now we sum each term separately. Let

$$T_1 = \frac{4}{|N|^3} \sum_{u,v,w \in N} \alpha^2 |N|^2$$

$$T_2 = \frac{4}{|N|^3} \sum_{u,v,w \in N} \alpha |N| \sum_{x \in N} Z_x$$

We will use the following basic facts: for any real numbers $a, b \geqslant 0$, $(\min(a,b))^2 \leqslant a \cdot b$, $\max(a,b) \leqslant a + b$, and $\min(a,b) \cdot \max(a,b) = a \cdot b$.

$$\begin{aligned} T_1 &\leqslant \frac{4}{|N|} \sum_{u,v,w \in N} \left(\max(Z_u, Z_v)\left(Z_w + \min(Z_u, Z_v)\right)\right) \\ &\leqslant \frac{4}{|N|} \sum_{u,v,w \in N} \left((Z_u + Z_v) Z_w + Z_u Z_v\right) \\ &= 12\, OPT \end{aligned}$$

Similarly, we analyze the second term using the fact that $\alpha \leqslant Z_u + Z_v$.

$$\begin{aligned} T_2 &= \frac{4}{|N|^3} \sum_{u,v,w \in N} \alpha |N| \sum_{x \in N} Z_x \\ &\leqslant \frac{4}{|N|^2} \sum_{u,v,w \in N} (Z_u + Z_v) \sum_{x \in N} Z_x \\ &= 8\, OPT \end{aligned}$$

Adding together all of the terms, $ALG \leqslant 21\, OPT$. □

This immediately yields the root mean square Distortion bound via Jensen's inequality.

Corollary 5. *The Distortion of Random Referee is at most* $\sqrt{21} \approx 4.583$.

6.3 Distortion of Random Referee on the Euclidean Plane

Though the upper bound on the Distortion of Random Referee is slightly worse than that of Random Dictatorship, we now show another advantage of using a comparison query: Such mechanisms can exploit structure in specific metric spaces that favorite-only mechanisms cannot. In other words, we show that the Distortion of Random Referee improves significantly for more structured metric spaces, while top-k only mechanisms do not share this property.

We examine the Distortion of Random Referee on a specific canonical metric of interest: the Euclidean plane when $d(u, p_u) = 0$ for every $u \in N$. The second assumption, functionally equivalent to assuming $N \subseteq \mathcal{S}$, simplifies our analysis considerably, and corresponds to the 0-decisive case in [8, 85]. We note that in examples where $|\mathcal{S}|$ is very large, the assumption becomes more innocuous. If we consider our opening example of public budgets, the assumption is something like this: every agent is allowed to propose their absolute favorite over all budgets, and we assume that this budget p_u has dis-utility of 0. We consider the Euclidean plane because the problem of minimizing distortion on the real line can be solved exactly [8], whereas we are unaware of any results for the Euclidean plane that are stronger than those for general metric spaces.

6.3.1 Lower Bounds for Distortion in the Restricted Model.

We begin by giving lower bounds to demonstrate that our simplifying assumptions still result in a nontrivial problem in two senses: (1) the Distortion of randomized social choice mechanisms are still bounded away from 1 and (2) any top-k only mechanism (one that elicits the top k preferred alternatives of agents) for constant k has Distortion at least 2 as $|N|$ and $|\mathcal{S}|$ become large.

Theorem 23. *The Distortion of a randomized social choice mechanism is at least 1.2 generally, and at least 1.118 for the Euclidean plane, even when* $d(u, p_u) = 0$ *for every* $u \in N$.

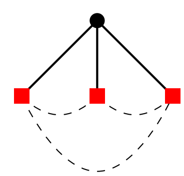


FIGURE 6.1: Example for Theorem 23. The red squares are agents as well as alternatives.

Proof. We begin with the more general claim. Consider a star graph with three leaves and three agents, one agent at each leaf, as depicted in Figure 6.1. All four points are also alternatives, and the distance between an agent and the colocated alternative is 0. Call the central vertex B and the set of outer vertices R. We want to construct a single ordinal preference profile such that for all randomized social choice mechanisms, there is some metric space that induces this preference profile and for which the mechanism has Distortion at least 1.2. The preference profile is simple: each of the three agents prefer their colocated alternative to the central alternative to either of the other leaf alternatives. We induce this profile via two metric spaces. In both spaces, $\forall v \in R,\ d(v, B) = 1$ (that is, the solid lines in Figure 6.1); the spaces will differ in the distances between points *inside* of R (that is, the dashed lines in Figure 6.1).

First, suppose that $\forall v, v' \in R,\ v \neq v',\ d(v, v') = 2$. Then the social cost of B is 3, and the social cost of every vertex in R is 4; clearly the distortion of B is 1 and the distortion of $v \in R$ is 4/3. Let $\Pr(R)$ be the probability that a randomized ordinal social mechanism chooses some $v \in R$, and let $\Pr(B)$ be the probability that the same mechanism chooses B. Then on this metric, the Distortion of the mechanism is (in expectation) just $\Pr(B) + (4/3)\Pr(R)$.

Second, suppose that $\forall v, v' \in R,\ v \neq v',\ d(v, v') = 1 + \epsilon$ for some $\epsilon \in (0, 0.5)$. Then the social cost of B is 3, and the social cost of every alternative in R is $2(1+\epsilon)$, so the distortion of B is $\frac{3}{2(1+\epsilon)}$ and the distortion of $v \in R$ is 1. Then on this metric, the Distortion of the mechanism is (in expectation) $\frac{3}{2(1+\epsilon)}\Pr(B) + \Pr(R)$.

A social choice mechanism with only ordinal information cannot distinguish the cases, and must set some values for $\Pr(R)$ and $\Pr(B)$ that sum to 1, so such a mechanism cannot have Distortion less than

$$\max\left(\Pr(B) + \frac{4}{3}\Pr(R),\ \frac{3}{2(1+\epsilon)}\Pr(B) + \Pr(R)\right).$$

This expression is minimized by setting $\Pr(R)$ to $\frac{2+2\epsilon}{5-4\epsilon}$ (and $\Pr(B) = 1 - \Pr(R)$) yielding a Distortion of

$$\frac{4}{3} - \left(\frac{1}{3}\right)\left(\frac{2+2\epsilon}{5-4\epsilon}\right) \longrightarrow 6/5 \text{ as } \epsilon \to 0$$

The argument for the lower bound of 1.118 on the Euclidean plane is similar. Now, the first case places the three points in R on a unit circle around the central point B, equidistant from one another. The law of cosines implies that $\forall v, v' \in R,\ v \neq v',\ d(v, v') = \sqrt{2 - 2\cos(120^\circ)} \approx 1.732$. In the second case, the angular distance between consecutive points in R is 60°, so that two pairwise distances are 1, and one pairwise distance is $\sqrt{2 - 2\cos(120^\circ)}$. Completing the same argument as before yields the bound of 1.11. □

Theorem 24. *The Distortion of any top-k only mechanism goes to* 2 *as* $|\mathcal{S}|$ *and* $|N|$ *become large, even on the Euclidean plane when* $d(u, p_u) = 0$ *for every* $u \in N$.

Proof. The construction for this lower bound follows that of Theorem 21. We briefly consider the argument for completeness. There are $|\mathcal{S}|/k$ agents, and for each agent u, there are k unique alternatives $S_u \subset S$ that are ϵ distance away from them. S_u are the top-k preferred alternatives for u. There are no other alternatives. Consider any top-k only mechanism (recall that k is constant). Such a mechanism chooses a distribution over alternatives, and thus $\exists u^* \in N$ with top-k preferred alternatives S_{u^*} such that the mechanism puts at least $1/|N|$ probability mass on S_{u^*}.

Suppose that the other $|N| - 1$ agents are arrayed in a circle such that the pairwise distance of diameter δ (clearly, if ϵ is small enough, such an arrangement need not violate our earlier construction). Further, suppose the distance from u^* to the center of this circle is 1. Then the social cost of any alternative in S_{u^*} is at least $(|N| - 1)(1 - \epsilon - \delta/2)$, and the social cost of any alternative not in S_{u^*} is at most $1 + \epsilon\delta/2 + (|N| - 2)(\delta + 2\epsilon)$. Therefore, the Distortion of choosing any alternative in S_{u^*} is at least

$$\frac{(|N| - 1)(1 - \epsilon - \delta/2)}{1 + \epsilon\delta/2 + (|N| - 2)(\delta + 2\epsilon)}.$$

As ϵ and δ approach 0, the above approaches $|N| - 1$. The Distortion of choosing any alternative not in S_{u^*} is clearly just 1. Since the mechanism chooses an alternative in S_{u^*} with probability at least $1/|N|$ (say $1/|N| + \alpha$, where $\alpha \geqslant 0$) and an alternative not in S_{u^*} with the remaining $1 - 1/|N| - \alpha$ probability, the over Distortion is lower bounded by

$$\begin{aligned}
&\left(\frac{1}{|N|} + \alpha\right)(|N| - 1) + \left(\frac{|N| - 1}{|N|} - \alpha\right) \\
&= 2 - 2/|N| + \alpha\,(|N| - 2) \\
&\geqslant 2 - 2/|N| \\
&\to 2 \text{ as } |N| \to \infty.
\end{aligned}$$

□

6.3.2 Upper Bound for Distortion of Random Referee in the Restricted Model.

Our positive result in this section demonstrates that a single comparison query is sufficient to construct a mechanism (Random Referee) with Distortion bounded below 1.97 for arbitrary $|\mathcal{S}|$ and $|N|$. We note that the bound in the theorem seems very slack; our goal is simply to show that using a comparison query provably decreases Distortion. We conjecture the actual bound is below 1.75 based on computer assisted search, but leave proving this stronger bound as an interesting open question.

Theorem 25. *The worst case distortion of Random Referee is less than 1.97 when $d(u, p_u) = 0$ for every $u \in N$, $\mathcal{S} \subseteq \mathbb{R}^2$, and for $x, y \in \mathbb{R}^2$ $d(x, y) = \|x - y\|_2$.*

In the remainder of this section, we sketch the proof of Theorem 25. Proofs of some of the technical lemmas are in the appendix. Suppose for a contradiction that there is a set N of agents in $\mathbb{R}^2$ such that the Distortion of Random Referee is at least 1.97 under the Euclidean metric. We will successively refine this hypothesis for a contradiction, finally arguing that it implies that some "bad" instance would appear on an exhaustive computer assisted search over a finite grid.

The crucial lemmas bound the *pessimistic distortion* of any set of five points in $\mathbb{R}^2$ and relate this to the actual distortion. The quantity is pessimistic because it allows "OPT" to choose a separate point for every 5-tuple; the numerator of each 5-tuple is just a rewriting of the expected social cost of Random Referee. In this section, since we assume $d(u, p_u) = 0$, it will not be necessary to refer to u and p_u separately, and it will be convenient to let $C(\{p_u, p_v\}, p_w) = \text{argmin}_{a \in \{p_u, p_v\}} d(p_w, a)$.

Definition 19. *The pessimistic distortion*[1] *of $\mathcal{P} = \{x_1, x_2, x_3, x_4, x_5\} \subset \mathbb{R}^2$ is defined as*

$$PD(\mathcal{P}) := \frac{SCRR_{avg}(\mathcal{P})}{OPT_{avg}(\mathcal{P})}$$

[1] Note that pessimistic distortion is technically a function of the mechanism used; we will use it exclusively in reference to Random Referee.

where $SCRR_{avg}(\mathcal{P})$ is the average social cost of Random Referee, which we can write as

$$\frac{1}{30}\sum_{i=1}^{5}\sum_{j>i}^{5}\sum_{k\neq i,j}\sum_{l\neq i,j,k} d(x_l, C(\{x_i, x_j\}, x_k)),$$

and the average cost of the optimal solution is

$$OPT_{avg}(\mathcal{P}) := \min_{y\in\mathbb{R}^2}\frac{1}{5}\sum_{r=1}^{5} d(x_r, y).$$

Finally, $PD(\mathcal{P}) = 1$ if $x_1 = x_2 = x_3 = x_4 = x_5$.

The first lemma relates this pessimistic distortion to the actual Distortion of Random Referee. The worst case (over 5-tuples) pessimistic distortion upper bounds the Distortion of Random Referee. Interestingly, this statement is not specific to the Euclidean plane, suggesting that our approach may be broadly applicable for proving stronger Distortion bounds on other specific metrics of interest.

Lemma 18. *If $PD(P) \leqslant \beta$ for all $\mathcal{P} = \{x_1, \ldots, x_5\} \subset \mathbb{R}^2$ then the Distortion of Random Referee is at most β on $\mathbb{R}^2$.*

Proof. Observe that we can rewrite the Distortion of Random Referee as a summation over all possible 5-tuples of points. Let $\mathcal{P}$ be a multiset of points (i.e., possibly with repeats). Let $\rho(\mathcal{P})$ be an ordering of $\mathcal{P}$. Let

$$C(\rho(\mathcal{P})) = C(\{\rho_1(\mathcal{P}), \rho_2(\mathcal{P})\}, \rho_3(\mathcal{P})).$$

We can rewrite the Distortion of Random Referee over permutations of 5-tuples as

$$\frac{\frac{1}{|N|^4}\sum_{i,j,k,l\in N} d(p_l, C(\{p_i, p_j\}, p_k)}{\frac{1}{|N|}\sum_{i\in N} d(p_i, a^*)}$$

$$= \frac{\sum_{\mathcal{P}\subset\mathcal{P}_N}\Pr(\mathcal{P})\frac{1}{5!}\sum_{\rho(\mathcal{P})}\sum_{l\in\{4,5\}} d(\rho_l(\mathcal{P}), C(\rho(\mathcal{P})))}{\frac{1}{|N|}\sum_{i\in N} d(p_i, a^*)}.$$

In words, we are considering all 5-tuples (with replacement) of agent points, and for each we consider the average distance over all orderings of the five agent points of the distance between the last two points and the outcome of Random Referee when the first three agents participate. This is in turn upper bounded by allowing OPT to choose a different a^* for every 5-tuple, so that the Distortion is at most

$$\sum_{\mathcal{P} \subset P_N} \Pr(\mathcal{P}) \frac{\frac{1}{5!} \sum_{\rho(\mathcal{P})} \sum_{l \in \{4,5\}} d(\rho_l(\mathcal{P}), C(\rho(\mathcal{P})))}{OPT_{avg}(\mathcal{P})}$$

$$\leqslant \max_{\mathcal{P} \subset P_N} \Pr(\mathcal{P}) \frac{\frac{1}{5!} \sum_{\rho(\mathcal{P})} \sum_{l \in \{4,5\}} d(\rho_l(\mathcal{P}), C(\rho(\mathcal{P})))}{OPT_{avg}(\mathcal{P})}.$$

To complete the proof, note that $SCRR_{avg}(\mathcal{P})$ is equal to the numerator. We start with all 120 orderings of the 5 points and avoid double counting the symmetric cases that arise from swapping the two points from which we take the argmin and swapping the two points from which we measure distance. □

Now we can refine our original hypothesis: without loss of generality, assume for a contradiction that there is a multiset $\mathcal{P} = \{x_1, \ldots, x_5\} \subseteq \mathbb{R}^2$ with $PD(\mathcal{P}) \geqslant 1.97$. We argue that we can further refine our hypothesis by assuming a canonical form in relation to a finite grid.

Lemma 19. *Let G be a δ-fine grid: $G = \{0, \delta, 2\delta, \ldots, \frac{\delta-1}{\delta}, 1\}^2$ where $\delta > 0$ and $1/\delta \in \mathbb{N}$. There is a set of points $\mathcal{Q} = \{y_1, y_2, y_3, y_4, y_5\}$ such that $PD(\mathcal{Q}) = PD(\mathcal{P})$ and $\mathcal{Q}$ has the properties:*

1. $\mathcal{Q} \subset [0, 1+\delta]^2$,
2. $\max_{i,j} d(y_i, y_j) = 1$,
3. *and these two maximally separated points y_i and y_j have the forms $y_i = (\alpha\delta, 0), y_j = (\alpha\delta, 1)$ for some $\alpha \in \{0, 1, 2, \ldots, 1/\delta\}$*

Proof. We note that the pessimistic distortion of $\mathcal{P}$ is invariant to scaling, rotation, and translation of $\mathcal{P}$. This follows because the pessimistic distortion is still defined in terms of euclidean distances. Given this, the argument for the Lemma is simply that we can construct $\mathcal{Q}$ by rotations, translations, and scaling.

First scale $\mathcal{P}$ so that the maximum distance separating two points is 1 (note that at least two points are not equal). Next, rotate the points so that the line between these maximally separated points is vertical. Finally, translate the points until these maximally separated points have the appropriate forms - that is to say, until they lie exactly at grid points, and all other points are within $[0, 1+\delta]^2$. To see that this is possible, note that since they are the maximally separated points, no other points can lie outside $[0, 1]$ in the vertical dimension. In the horizontal dimension, the total width spanned by the other points is at most 1, and we may need to expand by as much as δ in order to align the maximally separated points with grid points. □

The introduction of a δ-fine grid anticipates the computer analysis we employ. Lemma 19 allows us to further refine our assumption for a contradiction: Suppose without loss of generality[2] that there is some $\mathcal{P} = \{x_1, \ldots, x_5\} \subseteq [0, 1]^2$ such that $PD(\mathcal{P}) \geqslant \alpha$, $d(x_1, x_2) = 1$, and x_1, x_2 are of the form $(\alpha\delta, 0), y_j = (\alpha\delta, 1)$ for some $\alpha \in \{0, 1, 2, \ldots, 1/\delta\}$. We show the following contradiction.

Lemma 20. $PD(\mathcal{P}) < 1.97$.

Proof. Define $\phi : [0, 1]^2 \to G$ as $\phi(x) = \text{argmin}_{v \in G} d(x, v)$. In words, ϕ maps a general point in $[0, 1]$ to its nearest point on our δ-fine grid G. We will argue (roughly) that the pessimistic distortion of $\phi(\mathcal{P}) := \{\phi(x_1), \phi(x_2) \ldots, \phi(x_5)\}$ closely approximates that of $\mathcal{P}$, where the pessimistic distortion of of $\phi(\mathcal{P})$ is something we can compute

[2] For simplicity of notation, we will assume $\mathcal{P}$ lies in a $[0, 1]$ rather than $[0, 1+\delta]$, and will simply run our computer analysis for the expanded grid. Similarly, relabeling x_1 and x_2 as the maximally separated points is just a notational convenience.

directly in a brute force computer search. Recall Definition 19. We begin by bounding the numerator.

Let $C_\phi(\{x, y\}, z) = C(\{\phi(x), \phi(y)\}, \phi(z))$. Clearly we have from the definition of ϕ that $d(x, \phi(x)) \leqslant \left(\sqrt{2}/2\right)\delta$ for all x (and $\phi(x_1) = x_1$ and $\phi(x_2) = x_2$). However, it is not necessarily true that $d(C(\{x, y\}, z), C_\phi(\{x, y\}, z))$ is also small for all x, y, z. It is possible that z is slightly closer to x than y, but $\phi(z)$ is slightly closer to $\phi(y)$ than $\phi(x)$. We call such configurations indifferences.

Definition 20. *Call z indifferent with respect to x and y if* $|d(z, x) - d(z, y)| \leqslant \frac{3\sqrt{2}}{2}\delta$.

Then we have the following fact: If z is not indifferent with respect to x and y, then

$$d(C(\{x, y\}, z), C_\phi(\{x, y\}, z)) \leqslant \frac{\sqrt{2}}{2}\delta.$$

This follows from observing that if z is not indifferent, $\phi(C(\{x, y\}, z)) = C_\phi(\{x, y\}, z))$. Now, for our point set $\mathcal{P}$, we need to bound $d(x_l, C(\{x_i, x_j\}, x_k))$. We first consider the case where x_k is *not* indifferent with respect to x_i and x_j. Then we have can upper bound $d(x_l, C(\{x_i, x_j\}, x_k))$ by

$$\begin{aligned}
&d(x_l, \phi(x_l)) + d(\phi(x_l), C_\phi(\{x_i, x_j\}, x_k)) \\
&\quad + d(C_\phi(\{x_i, x_j\}, x_k), C(\{x_i, x_j\}, x_k)) \\
&\leqslant \begin{cases} \left(\sqrt{2}\right)\delta + d(\phi(x_l), C_\phi(\{x_i, x_j\}, x_k)) & l \in \{3, 4, 5\} \\ \left(\frac{\sqrt{2}}{2}\right)\delta + d(\phi(x_l), C_\phi(\{x_i, x_j\}, x_k)) & l \in \{1, 2\} \end{cases}
\end{aligned}$$

where the bound tightens for x_1 and x_2 because they are already at grid points. Suppose instead that x_k *is* indifferent with respect to x_i and x_j. Then we can upper bound $d(x_l, C(\{x_i, x_j\}, x_k))$ by

$$\begin{aligned}
&d(x_l, \phi(x_l)) + d(\phi(x_l), \phi(C(\{x_i, x_j\}, x_k))) \\
&\quad + d(\phi(C(\{x_i, x_j\}, x_k)), C(\{x_i, x_j\}, x_k))
\end{aligned}$$

which is at most

$$\begin{cases} \left(\sqrt{2}\right)\delta + d(\phi(x_l), \phi(C(\{x_i, x_j\}, x_k))) & l \in \{3,4,5\} \\ \left(\frac{\sqrt{2}}{2}\right)\delta + d(\phi(x_l), \phi(C(\{x_i, x_j\}, x_k))) & l \in \{1,2\} \end{cases}$$

Furthermore, since $C(\{x_i, x_j\}, x_k) \in \{x_i, x_j\}$, we know that

$$d(\phi(x_l), \phi(C(\{x_i, x_j\}, x_k))) \leqslant \max_{q \in \{i,j\}} d(\phi(x_l, \phi(x_q))).$$

Call the set of indifferent cases I. We can average over the indifferences, and over $\{x_1, x_2\}$ and $\{x_3, x_4, x_4\}$, to upper bound $SCRR_{avg}(\mathcal{P})$ by

$$\begin{aligned} &\left(\frac{3}{5}\right)(\sqrt{2})\delta + \left(\frac{2}{5}\right)\left(\frac{\sqrt{2}}{2}\right)\delta \\ &\quad + \frac{1}{30}\sum_{I}\sum_{l \neq i,j,k} d(\phi(x_l), \phi(C(\{x_i, x_j\}, x_k))) \\ &\quad + \frac{1}{30}\sum_{\overline{I}}\sum_{l \neq i,j,k} d(\phi(x_l), C_\phi(\{x_i, x_j\}, x_k)) \\ &\leqslant \frac{4\sqrt{2}}{5}\delta + \frac{1}{30}\sum_{I}\max_{q \in \{i,j\}}\sum_{l \neq i,j,k} d(\phi(x_l), \phi(x_q)) \\ &\quad + \frac{1}{30}\sum_{\overline{I}}\sum_{l \neq i,j,k} d(\phi(x_l), C_\phi(\{x_i, x_j\}, x_k)) \end{aligned}$$

Call the summations in this last line $GridSum$. Then

$$\frac{SCRR_{avg}(\mathcal{P})}{GridSum} \leqslant 1 + \frac{4\sqrt{2}\delta}{5\, GridSum}$$

$GridSum$ looks like a very involved quantity, but it is exactly what we will compute in our computer analysis. That is, when we are computing the pessimistic distortion on the grid for a set of five points, whenever we encounter an indifference, we take the worst case outcome over the two options. It is not hard to see that $GridSum \geqslant$

1/5 since there are least two points separated by a distance of 1, so we arrive at equation 6.1.

$$\frac{SCRR_{avg}(\mathcal{P})}{GridSum} \leqslant 1 + \left(4\sqrt{2}\right)\delta \tag{6.1}$$

Now we need to bound

$$\frac{OPT_{avg}(\phi(\mathcal{P}))}{OPT_{avg}(\mathcal{P})}.$$

Let

$$y_\phi^* = \text{argmin}_y \frac{1}{5}\sum_{l=1}^{5} d(\phi(x_l), y)$$

$$y^* = \text{argmin}_y \frac{1}{5}\sum_{l=1}^{5} d(x_l, y).$$

Since y^* was a feasible choice for the minimization over $\phi(\mathcal{P})$, the triangle inequality implies that

$$\begin{aligned}\frac{1}{5}\sum_{l=1}^{5} d(\phi(x_l), y_\phi^*) &\leqslant \frac{1}{5}\sum_{l=1}^{5} \left(d(\phi(x_l), x_l) + d(x_l, y^*)\right)\\ &\leqslant \left(\frac{3\sqrt{2}}{10}\right)\delta + \frac{1}{5}\sum_{l=1}^{5} d(x_l, y^*)\end{aligned}$$

This implies that

$$OPT_{avg}(\phi(\mathcal{P})) \leqslant \left(\frac{3\sqrt{2}}{10}\right)\delta + OPT_{avg}(\mathcal{P}).$$

Since we know that $d(x_1, x_2) = 1$, we have that $OPT_{avg}(\mathcal{P}) \geqslant 1/5$. So we get that

$$\frac{OPT_{avg}(\phi(\mathcal{P}))}{OPT_{avg}(\mathcal{P})} \leqslant 1 + \left(\frac{3\sqrt{2}}{2}\right)\delta. \tag{6.2}$$

Combining equation 6.1 and equation 6.2:

$$PD(\mathcal{P})\frac{OPT_{avg}(\phi(\mathcal{P}))}{GridSum} \leqslant \left(1 + (4\sqrt{2})\delta\right)\left(1 + \frac{3\sqrt{2}}{2}\delta\right).$$

Now, suppose for a contradiction that $PD(\mathcal{P}) \geqslant 1.97$. Then we know that

$$\frac{GridSum}{OPT_{avg}(\phi(\mathcal{P}))} \geqslant \frac{1.97}{\left(1+(4\sqrt{2})\delta\right)\left(1+\left(\frac{3\sqrt{2}}{2}\right)\delta\right)}$$

So for $\delta := 1/75$, we get that

$$\frac{GridSum}{OPT_{avg}(\phi(\mathcal{P}))} \geqslant 1.781.$$

But $\phi(\mathcal{P}) \subset G$, and computer analysis, searching over all possible configurations of points on G, reveals no such configuration. □

Lemma 20 provides the contradiction to our hypothesis and establishes Theorem 25.

Discussion. Our analysis is slack in two ways: (1) we have to interpolate between grid points, and (2) we consider pessimistic distortion over 5-tuples. Both are computational constraints: (1) because we cannot simulate an arbitrarily fine grid, and (2) because there is a combinatorial blow up in the search space when considering larger tuples (note that by considering 5-tuples, we implicitly allow OPT to choose a separate optimal solution for each 5-tuple). For Random Referee, the worst case example found by computer simulations for grid points is fairly simple: The 5 points lie on a straight line with pessimistic distortion 1.75. We conjecture the same example is the worst case even in the continuous plane, which suggests a distortion bound of at most 1.75. We leave finding the exact bound as an open question, as our result is sufficient to demonstrate that comparison queries can take advantage of structure that top-k queries cannot.

6.4 Favorite Only Mechanisms: Random Oligarchy

Recall that a favorite query asks an agent $u \in N$ for her favorite alternative p_u. In this section, we return to the general model (arbitrary metrics and not assuming that $d(u, p_u) = 0$) and study mechanisms that are restricted to only use favorite queries. We show that essentially optimal Distortion as a function of $|\mathcal{S}|$ is achieved by a simple mechanism that uses just 3 queries. We call this mechanism Random Oligarchy.

Definition 21. *The Random Oligarchy (RO) mechanism samples three agents $u, v, w \in N$ independently and uniformly at random with replacement. All three are asked for their favorite alternatives p_u, p_v, and p_w in $\mathcal{S}$. If the same alternative is reported at least twice, output that, else output one of the three alternatives uniformly at random.*

We prove that Random Oligarchy has the best of both worlds with respect to the other favorite only mechanisms of Random Dictatorship and 2-Agree. Unlike 2-Agree, Random Oligarchy has constant sample complexity and the same Distortion bound of 3 for large $|\mathcal{S}|$ as Random Dictatorship. However, like 2-Agree, it outperforms Random Dictatorship for small $|\mathcal{S}|$.

Theorem 26. *The Distortion of Random Oligarchy is upper bounded by 3 for arbitrary $|\mathcal{S}|$, and by the following expression for particular $|\mathcal{S}|$.*

$$1 + 2 \max_{p \in [0,1]} \left(1 + p^2(p-2) + \frac{(p-1)^3}{|\mathcal{S}| - 1} \right)$$

Proof. We analyze Random Oligarchy using a technical lemma proven in [85]. For clarity of exposition, we restate the lemma.

Lemma 21. *[85] The Distortion of any randomized mechanism f is less than or equal to*

$$1 + 2 \max_{p \in [0,1]} m_f(p) \frac{1-p}{p}$$

where $m_f(p)$ is the maximum probability of an alternative being output by f if exactly a p fraction of N consider this alternative their top choice.

To apply Lemma 21, we need to bound $m_f(p)$: the maximum probability of an alternative being output by f if exactly a p fraction of N consider this alternative their top choice (f, in this context, is just the Random Oligarchy mechanism). We will first consider a coarse bound that allows us to upper bound Distortion by 3 for arbitrary $|\mathcal{S}|$, and will then consider a fine grained bound that takes $|\mathcal{S}|$ into account.

It is straightforward from Definition 21 that Random Oligarchy outputs an alternative $x \in \mathcal{S}$ only if (i) it draws at least two agents who report x as their favorite, or if (ii) all three agents drawn report different favorite alternatives, and x is among them, in which case it is output with probability 1/3. Given that exactly a p fraction of N consider such an x as their top choice, it is straightforward to compute (i). A coarse upper bound on the probability of (ii) is $\binom{3}{1}p(1-p)^2$, the probability that the alternative in question is reported exactly once, and something else is reported for the other two draws. This is potentially an overestimate, since these two other alternatives reported must be distinct, but ignoring this detail gives the following coarse bound on $m_f(p)$.

$$\begin{aligned} m_f(p) &\leqslant \binom{3}{3}p^3 + \binom{3}{2}p^2(1-p) + \frac{1}{3}\binom{3}{1}p(1-p)^2 \\ &= -p^3 + p^2 + p \end{aligned}$$

Inserting this into Lemma 21 and simplifying shows that the Distortion of Random Oligarchy is at most 3.

$$\begin{aligned} Distortion(f) &\leqslant 1 + 2\max_{p\in[0,1]}\left(\left(-p^3+p^2+p\right)\frac{1-p}{p}\right) \\ &= 3 \end{aligned}$$

To obtain a more fine grained bound that considers $|\mathcal{S}|$, note that the probability of (ii) is maximized when the distribution over favorite points of all agents who do not report x as their favorite is uniform over the remaining $|\mathcal{S}| - 1$ alternatives. Taking into account that for event (ii) the two other reported alternatives must be unique yields the following bound on $m_f(p)$.

$$\begin{aligned} m_f(p) &\leqslant \binom{3}{3} p^3 + \binom{3}{2} p^2 (1-p) \\ &\quad + \frac{1}{3} \binom{3}{1} p(1-p) \left(1 - p - \frac{1-p}{|\mathcal{S}| - 1}\right) \\ &= -p^3 + p^2 + p - \frac{p(p-1)^2}{|\mathcal{S}| - 1} \end{aligned}$$

Substituting this expression into Lemma 21 allows us to upper bound the Distortion of Random Oligarchy by

$$1 + 2 \max_{p \in [0,1]} \left(\left(-p^3 + p^2 + p - \frac{p(p-1)^2}{|\mathcal{S}| - 1} \right) \frac{1-p}{p} \right).$$

Simplifying yields the theorem statement. □

Figure 6.2 shows the Distortion bounds of favorite only mechanisms. Comparing against the lower bound for any favorite only mechanism from [85] allows us to see that Random Oligarchy is essentially optimal among all favorite only mechanisms. In comparison to existing mechanisms, Random Oligarchy outperforms Random Dictatorship for small $|\mathcal{S}|$, and outperforms 2-Agree for large $|\mathcal{S}|$, while only using three favorite queries.

6.5 Conclusion and Open Directions

In this chapter, we studied constant sample complexity mechanisms for implicit utilitarian social choice. We proposed two mechanisms: Random Referee and Random Oligarchy, for achieving constant Distortion and Squared Distortion bounds using a constant number of simple ordinal queries. At a high level, we hope that our work

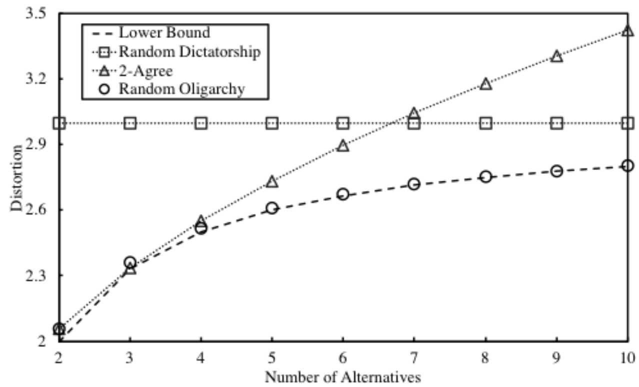


FIGURE 6.2: Distortion of Favorite Only Mechanisms. Upper bounds for Random Oligarchy are from Theorem 26. Lower bounds and 2-Agree upper bounds are from [85]. Random Dictatorship is analyzed in [8].

inspires future research on lightweight mechanisms for social choice in large decision spaces. We also mention some natural technical questions raised by our work.

Compared to Distortion, there is much less understanding of the Squared Distortion of mechanisms. The only universal lower bound for Squared Distortion we are aware of is 4, a consequence of the lower bound of 2 for Distortion [8]. Our Random Referee mechanism achieves Squared Distortion of at most 21 using just three ordinal queries. Closing this gap remains an interesting question even for mechanisms that elicit full ordinal information.

The analysis in Section 6.3 of Random Referee generalizes to higher dimensional Euclidean space, still only reasoning about 5-tuples of points. Computer search over a coarse grid in four dimensions again shows that the pessimistic distortion bound is better than 2. However, we have not been able to run the search on a fine enough grid to prove a Distortion bound formally. We leave proving this for higher dimensional Euclidean spaces as an interesting open question. Additionally, we hope that related methods may be of general interest for proving tighter Distortion bounds of mechanisms on restricted metric spaces.

It remains unclear whether using a constant number of queries greater than 3 would meaningfully improve our results. E.g., consider the natural extension of Random Referee: sample the favorite points of k agents and ask a random referee to choose their favorite from among these. Using $k > 2$ does not straightforwardly decrease the Squared Distortion bound of 21 for $k = 2$. Also, as k becomes large, this mechanism devolves to Random Dictatorship. Another natural question is whether $O(k)$ comparison queries are necessary/sufficient to bound the k'th moment of Distortion. We leave these as additional open questions.

7

Conclusion

We have studied public decision making as an algorithmic discipline from two broad perspectives: fairness in the allocation of public goods, and implicit utilitarian social choice. Our primary results in fair resource allocation have involved a deep exploration of the computation of core solutions to fair resource allocation problems, a novel fairness notion that generalizes the idea of proportionality by explicitly considering fairness with respect to subsets of agents in order to account for the potential for sharing public goods. These algorithms provide strong guarantees on fairness and efficiency for problems including budgeting, memory sharing, multiple referendums, multi-winner elections, and centroid clustering. In implicit utilitarian social choice, we designed a sequential protocol based on iterative Nash bargaining, and one shot protocols based on a simple and restricted query model, in order to provide constant Distortion guarantees (approximation to the social optimum). We also extended existing work by considering additional desiderata, most notably the Squared Distortion. These are lightweight protocols for social choice where there is a large space of alternatives.

In addition to many outstanding technical questions mentioned in specific chapters, there are some important broad questions that I wish to highlight in closing. First, we have only occasionally considered issues of strategic behavior, that is, the possibility that agents might misreport their true preferences in order to manipulate our algorithms to their advantage. The free rider problem is perhaps the most famous argument that public decision making is impeded by strategic concerns. In particular, for most of the problems considered in public decision making, the individual agents have an incentive to pretend to be disinterested in shared public goods, and instead feign interest only in more niche private goods.

This problem seems endemic in the allocation of public goods. There are classic impossibility results regarding strategic behavior, including the Gibbard-Satterthwaite impossibility theorem in ordinal voting [78], and the Myerson-Satterthwaite impossibility theorem in mechanism design [129]. These results are of the form that no voting rule or mechanism can simultaneously satisfy three or four desirable axioms, one of which is a form of truthfulness: that agents should not have an incentive to misreport their preferences.

Nevertheless, I believe there are important approaches for sidestepping some of these impossibilities. These ideas are based on approximation of two types. The first is to design mechanisms that may not be exactly truthful, but for which an agent cannot increase her expected utility by a large factor by misreporting her preferences. We have already seen this approach to mechanism design with the exponential mechanism of Section 2.3.2. Another approach is to accept that agents may behave strategically, and design algorithms for which other desirable properties (such as fairness, or welfare maximization) are still approximately obtained, even when agents so behave. This, for example, is the basic idea behind analysis of the price of anarchy [109]. I hope to more fully explore the application of these approaches to designing more strategically robust algorithms for public decision making.

Another relatively unexplored area of research concerns the true nature of many application domains in public decision making as problems of mixed public goods and "bads" (also sometimes called "chores" in the literature). In every chapter, we have either assumed that we are exclusively dealing with public goods (for which agents only have nonnegative utility) or exclusively dealing with public bads (in particular, in Chapters 4 through 6 we supposed that public outcomes had costs for agents). However, in reality, many public outcomes are good for some and bad for others. For example, building a highway through a neighborhood may be good to those who use the highway, but bad to those who live in the neighborhood.

This also seems to be a fundamental problem in much of the current literature on fair machine learning, including problems like predicting loan default, predictive policing, and recommender systems. I believe that algorithmic fairness in such a context demands considering a rich diversity of preferences over outcomes that may differ among the individuals and between the subjects and designers of the systems. Such an approach may allow one to go beyond statements about marginals of predictive distributions and instead make statements about the impact of algorithms on individuals. Normative utilitarian theory and corresponding algorithms may have much to contribute to this emerging field, but such work may need to account for mixed goods and bads present in many of these problems with fundamental developments in terms of solution concepts and computation.

At a higher level, our discussion has been characterized by a deep interaction between normative economics and developing efficient approximation algorithms with provable guarantees. This intersection between the normative and algorithmic (what is the right solution concept, and how can we compute it?) has obvious and traditional applications in voting and budgeting. But, this intersection is also significant in an age when machine intelligence is rapidly being deployed within regular human society. For better or for worse, algorithm designers will find themselves engaged in projects with substantial normative questions like machine learning driven algorithmic decision making, autonomous vehicles, social robotics, and many more projects that I cannot hope to predict. Thinking critically about the right solution concepts for these problems can either be done cursorily and after-the-fact, viewed as a distraction from the design of more efficient algorithms, or it can be seen as a fundamental part of algorithm design. It is my hope that this thesis contributes toward a future that takes the latter approach, for the public good of all.

8

Appendix

8.1 Proof of Theorem 2

The proof has many details and is divided into several technical points.

Lemma 22. *If* $\max_{\mathbf{y}\in\mathcal{P}}\left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right) = n$ *then the allocation* $\mathbf{x}$ *is a* $\frac{(k-1)n^{-\gamma}}{1-kn^{-\gamma}}$*-approximate additive core solution.*

Proof. Suppose by contradiction that $\max_{\mathbf{y}\in\mathcal{P}}\left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right) = n$, but $\mathbf{x}$ is not a $\frac{(k-1)n^{-\gamma}}{1-kn^{-\gamma}}$-approximate core solution. Then there is a subset S of agents who want to deviate to some allocation $\mathbf{z}$ where $\|\mathbf{z}\|_1 \leqslant \frac{|S|}{n}$ and for every agent i in S, $U_i(\mathbf{z}) > U_i(\mathbf{x}) + \frac{(k-1)n^{-\gamma}}{1-kn^{-\gamma}}$. Define the allocation $\mathbf{z}' := \frac{n\left(1-kn^{-\gamma}\right)}{|S|}\mathbf{z} + n^{-\gamma}\mathbf{1}$. Clearly $\mathbf{z}' \in \mathcal{P}$, so by assumption we have that $\sum_i \frac{U_i(\mathbf{z}')}{U_i(\mathbf{x})} \leqslant n$. Substituting for $\mathbf{z}'$ gives:

$$\sum_i \frac{\mathbf{u_i} \cdot \left(\frac{n\left(1-kn^{-\gamma}\right)}{|S|}\mathbf{z} + n^{-\gamma}\mathbf{1}\right)}{U_i(\mathbf{x})} \leqslant n$$

Thus, solving for $\sum_i \frac{\mathbf{u_i}\cdot\mathbf{z}}{U_i(\mathbf{x})}$ and simplifying:

$$\sum_i \frac{\mathbf{u_i}\cdot\mathbf{z}}{U_i(\mathbf{x})} \leqslant \frac{|S|}{n\left(1-kn^{-\gamma}\right)}\left(n-\sum_i \frac{n^{-\gamma}}{U_i(\mathbf{x})}\right) \leqslant \frac{1-n^{-\gamma}}{1-kn^{-\gamma}}|S|$$

However, recall that since S is a deviating coalition from the approximate core, it should be that $U_i(\mathbf{z}) > U_i(\mathbf{x}) + \frac{(k-1)n^{-\gamma}}{1-kn^{-\gamma}}$ for all agents $i \in S$. This implies that for all $i \in S$, it should be that $\frac{U_i(\mathbf{z})}{U_i(\mathbf{x})} > 1 + \frac{(k-1)n^{-\gamma}}{1-kn^{-\gamma}}$. Thus, we can bound the sum over i as $\sum_i \frac{U_i(\mathbf{z})}{U_i(\mathbf{x})} > \frac{1-n^{-\gamma}}{1-kn^{-\gamma}}|S|$. This is a contradiction, completing the proof. □

Using essentially the same argument, the following corollary follows easily.

Corollary 6. *If* $\max_{\mathbf{y}\in\mathcal{P}}\left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right) = n+\alpha$ *then the allocation* $\mathbf{x}$ *is an* $\frac{(k-1)n^{-\gamma}+\alpha n^{-1}}{1-kn^{-\gamma}}$*-approximate additive core solution.*

Lemma 23. $q(\mathbf{x}) \geqslant 0$ *and* $\max_{\mathbf{x}\in\mathcal{P}} q(\mathbf{x}) = n - n^{1-\gamma}$*. Further, if* Δq *is the largest possible difference in the scoring function between two sets of input differing only on the report of a single agent* i'*, (*i.e.*, the* sensitivity *of* q*), then* $\Delta q = 1$*.*

Proof. To see the first part, note that $U_i(\mathbf{x}) \geqslant n^{-\gamma}$ for all $\mathbf{x} \in \mathcal{P}$. Therefore,

$$q(\mathbf{x}) = n - n^{-\gamma}\max_{\mathbf{y}\in\mathcal{P}}\left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right) \geqslant n - n^{-\gamma}\sum_i \frac{1}{n^{-\gamma}} = 0$$

By the optimality condition of the Proportional Fairness convex program,

$$\min_{\mathbf{x}\in\mathcal{P}}\max_{\mathbf{y}\in\mathcal{P}}\left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right) = n$$

Therefore, $\max_{\mathbf{x}\in\mathcal{P}} q(\mathbf{x}) = n - n^{1-\gamma}$. Similarly, when agent i' misreports:

$$\begin{aligned}\Delta q &= \left|\left(n - n^{-\gamma}\max_{\mathbf{y}\in\mathcal{P}}\left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right)\right) - \left(n - n^{-\gamma}\max_{\mathbf{y}\in\mathcal{P}}\left(\frac{U_{i'}(\mathbf{y})}{U_{i'}(\mathbf{x})} + \sum_{i\neq i'}\frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right)\right)\right| \\ &\leqslant n^{-\gamma}\left(\max_{\mathbf{y}\in\mathcal{P}}\left(\frac{U_{i'}(\mathbf{y})}{U_{i'}(\mathbf{x})}\right) - 1\right) \leqslant 1\end{aligned}$$

The first inequality follows because we can assume w.l.o.g. that if the maximizing $\mathbf{y}$ changed for the misreported data, it yields a score no worse than the score of the original $\mathbf{y}$ on the misreported data, since that original $\mathbf{y}$ could have been chosen.

$\square$

Exponential Mechanism We now plug the above scoring function into the Exponential mechanism from [123]. We use $\epsilon > 0$ as the privacy approximation parameter, and thus as a parameter for the approximation of truthfulness.

Definition 22. *Define μ to be a uniform probability distribution over all feasible allocations $\mathbf{x} \in \mathcal{P}$. For a given set of utilities, let the mechanism ζ_q^ϵ be given by the rule:*

$$\zeta_q^\epsilon := \text{ choose } \mathbf{x} \text{ with probability proportional to } e^{\epsilon q(\mathbf{x})}\mu(\mathbf{x})$$

The following lemma follows by using the sensitivity bound from Lemma 23 in Theorem 6 from [123].

Lemma 24. *ζ_q^ϵ is $(e^{2\epsilon} - 1)$-approximately truthful.*

Finally, we demonstrate that ζ_q^ϵ can still find an approximate core solution while providing approximate truthfulness, establishing Theorem 2.

Claim 3. *If k is $o(\sqrt{n})$ and $\frac{1}{\epsilon} > \frac{kn}{(n-k^2)\ln n}$ then ζ_q^ϵ can be used to choose an allocation $\mathbf{x}$ that is an $O\left(\frac{k \ln n}{\epsilon\sqrt{n}}\right)$-approximate additive core solution w.p. $1 - \frac{1}{n}$.*

Proof. Let $t = \frac{k+1}{\epsilon}\ln n$. Lemma 7 in [123] states that

$$\Pr\left[n - n^{-\gamma}\max_{\mathbf{y}\in\mathcal{P}}\left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})}\right) \leqslant OPT - 2t\right] \leqslant \frac{e^{-\epsilon t}}{\mu(S_t)} \tag{8.1}$$

where OPT is the maximum value of $q(\mathbf{x})$ for feasible allocations $\mathbf{x}$ and $S_t = \{\mathbf{x} : q(\mathbf{x}) > OPT - t\}$. By Lemma 23, we have $OPT = n(1 - n^{-\gamma})$, but we need to bound

$\mu(S_t)$, the probability that $\mathbf{x}$ drawn uniformly at random from $\mathcal{P}$ is in S_t. We will show that $\mu(S_t) \geqslant n^{-k}$. Let $\mathbf{x}^* \in \mathcal{P}$ be the allocation such that $q(\mathbf{x}^*) = OPT$. Since $\|\mathbf{x}\|_1 = 1$, there is an item j' with $x^*_{j'} \geqslant 1/k$. Let $\delta = 1/n$. Define the set S_δ so that

$$S_\delta = \left\{ \mathbf{x} : \begin{array}{ll} x^*_j \leqslant x_j \leqslant x^*_j + \delta & j \neq j' \\ x^*_{j'} - k^2\delta x^*_{j'} \leqslant x_{j'} \leqslant x^*_{j'} - k^2\delta x^*_{j'} + \delta & j = j' \end{array} \right\}$$

It is not hard to see that since $x^*_{j'} \geqslant 1/k$, all $\mathbf{x} \in S_\delta$ are feasible. Furthermore, because there is a "width" of $1/n$ in possible choice of x_j for all j, $\mu(S_\delta) \geqslant n^{-k}$. Thus, to complete the argument that $\mu(S_t) \geqslant n^{-k}$, we just need to show that $S_\delta \subseteq S_t$. In our case,

$$S_t = \left\{ \mathbf{x} : \max_{y \in \mathcal{P}} \sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})} < n + \frac{k+1}{\epsilon} \ln n \right\}$$

Since $\frac{1}{\epsilon} > \frac{kn}{(n-k^2)\ln n}$, substituting shows that an allocation $\mathbf{x}$ is surely in S_t if the same sum is less than $\frac{n^2}{n-k^2}$. By construction, in the worst case for any agent i and allocation $\mathbf{x} \in S_\delta$ (namely, if $u_{ij'} = 1$), $U_i(\mathbf{x}) \geqslant \frac{n-k^2}{n} U_i(\mathbf{x}^*)$. Therefore, for all $\mathbf{x} \in S_\delta$

$$\max_{y \in \mathcal{P}} \sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})} \leqslant \frac{n}{n-k^2} \max_{y \in \mathcal{P}} \sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x}^*)} = \frac{n^2}{n-k^2}$$

Thus, we have that $S_\delta \subseteq S_t$ and therefore $\mu(S_t) \geqslant n^{-k}$. Substituting into equation 8.1 and simplifying yields

$$\Pr\left[\max_{\mathbf{y} \in \mathcal{P}} \left(\sum_i \frac{U_i(\mathbf{y})}{U_i(\mathbf{x})} \right) > n + 2\frac{k+1}{\epsilon} n^\gamma \ln n \right] \leqslant \frac{1}{n}$$

By applying Corollary 6, we get that $\mathbf{x}$ chosen according to ζ^ϵ_q is a $\frac{(k-1)n^{-\gamma} + 2(k+1)\epsilon^{-1} n^{\gamma-1} \ln n}{1 - kn^{-\gamma}}$-approximate core solution with probability $1 - \frac{1}{n}$. Plugging in $\gamma = 1/2$ and using the fact that k is $o(\sqrt{n})$ gives that $\mathbf{x}$ chosen according to ζ^ϵ_q is an $O\left(\frac{k \ln n}{\epsilon \sqrt{n}}\right)$-approximate core solution with probability $1 - \frac{1}{n}$.

□

8.2 Proof of Theorem 8

The fractional outcome maximizing the Nash welfare objective is the solution of the following program. For simplicity of presentation, we absorb the constraint that $w_j \leqslant 1$ for each $j \in W$ into the packing constraints.

$$\text{Maximize} \quad \sum_{i \in N} \ln U_i \tag{8.2}$$

$$\begin{array}{rcll} \sum_{j=1}^{m} a_{kj} w_j & \leqslant & b_k & \forall k \in [K] \\ U_i & = & \sum_{j=1}^{m} w_j u_{ij} & \forall i \in N \\ w_j & \geqslant & 0 & \forall j \in W \end{array}$$

Denote a vector of utilities by $\mathbf{U} = \langle U_1, U_2, \ldots, U_n \rangle$, and the polytope of feasible utility vectors by $\mathcal{U}$. Then, the fractional MNW outcome is obtained by the following maximization.

$$\max_{\mathbf{U} \in \mathcal{U}} \sum_{i \in N} \ln U_i \tag{8.3}$$

We want to compute a fractional (δ, ϵ)-approximate core outcome in time polynomial in $n, V_{\max}$, and $\log \frac{1}{\delta \epsilon}$. Assume that $\mathcal{U}$ is a convex polytope of feasible utility vectors. For any $\delta \geqslant 0, \epsilon > 0$, let $\epsilon' = \epsilon/(1+\delta)$. Define the following objective function. Note that in the absence of the ϵ' term, it would mimic the derivative of the Nash social welfare objective from Program (8.3).

$$\min_{\mathbf{U} \in \mathcal{U}} Q(\mathbf{U}), \text{ where } Q(\mathbf{U}) = \max_{\mathbf{U}' \in \mathcal{U}} \sum_{i \in N} \frac{U_i' + \epsilon'}{U_i + \epsilon'}. \tag{8.4}$$

Clearly, $Q(\mathbf{U}) \geqslant n$ for every $\mathbf{U}$. Thus, the objective value in Program (8.4) is at least n. In Section 3.2.2, we presented an argument showing that the fractional

MNW outcome is in the core. A similar argument using the first order optimality condition shows that if $\mathbf{U}^* \in \arg\max_{\mathbf{U}\in\mathcal{U}} \sum_{i\in N} \ln(U_i + \epsilon')$, then

$$\sum_{i\in N} \frac{U_i + \epsilon'}{U_i^* + \epsilon'} \leqslant n.$$

This implies the optimum of Program (8.4) is achieved at the fractional outcome maximizing the smooth Nash welfare objective $\sum_{i\in N} \ln(U_i + \epsilon')$, and this optimal value is exactly n.

Next, we turn to efficiently approximating Program (8.4), and show that if $Q(\mathbf{U}) \leqslant n(1+\delta)$, then $\mathbf{U}$ is a (δ, ϵ)-core outcome.

We want to use the Ellipsoid algorithm to approximately minimize the objective function $Q(\mathbf{U})$ over $\mathbf{U} \in \mathcal{U}$ in polynomial time. For this, all we need is that Q is a convex function, its subgradient is efficiently computable, the range of Q and the diameter of $\mathcal{U}$ are exponentially bounded, and polytope $\mathcal{U}$ is efficiently separable [35].

First, we claim that $Q(\mathbf{U})$ is a convex function of $\mathbf{U}$. To see this, note that for any fixed $\mathbf{U}'$, U_i'/U_i is convex in U_i. Since the sum and maximum of convex functions is convex, we conclude that $Q(\mathbf{U})$ is also convex.

Second, the subgradient of $Q(\mathbf{U})$ is efficiently computable for every $\mathbf{U} \in \mathcal{U}$. First, we find the $\mathbf{U}' \in \mathcal{U}$ that maximizes $\sum_{i\in N} \frac{U_i'+\epsilon'}{U_i+\epsilon'}$ by solving a linear program. Then, we fix $\mathbf{U}'$ and take the gradient of $\frac{U_i'}{U_i}$ with respect to U_i to obtain a subgradient of $Q(\mathbf{U})$.

Third, note that $U_i \in [0, V_{\max}]$ for each i. Hence, $Q(\mathbf{U}) \leqslant \frac{n\cdot(V_{\max}+\epsilon')}{\epsilon'}$, which is exponentially bounded in the input size. It is easy to see that the same holds for the diameter of the polytop $\mathcal{U}$.

Finally, polytope $\mathcal{U}$ is efficiently separable because it is a set of polynomially many linear inequalities.

Hence, we can efficiently obtain a solution $\hat{\mathbf{U}} \in \mathcal{U}$ that satisfies

$$\max_{\mathbf{U}' \in \mathcal{U}} \sum_{i \in N} \frac{U'_i + \epsilon'}{\hat{U}_i + \epsilon'} \leqslant n + \delta \leqslant n(1+\delta).$$

Finally, we show that $\hat{\mathbf{U}}$ must be a (δ, ϵ)-core outcome. Suppose for contradiction that it is not. Then, there exists a subset S of agents and an outcome $\mathbf{U}'$ such that

$$(1+\delta) \cdot \hat{U}_i + \epsilon \leqslant \frac{|S|}{n} \cdot U'_i$$

for all $i \in S$, and at least one inequality is strict. Rearranging the terms and summing over $i \in S$, we obtain

$$\sum_{i \in S} \frac{U'_i}{(1+\delta) \cdot \hat{U}_i + \epsilon} > |S| \cdot \frac{n}{|S|} = n.$$

However, we also have

$$\sum_{i \in S} \frac{U'_i}{(1+\delta) \cdot \hat{U}_i + \epsilon} \leqslant \sum_{i \in S} \frac{U'_i + \epsilon'}{(1+\delta) \cdot (\hat{U}_i + \epsilon')} = \frac{1}{1+\delta} \sum_{i \in S} \frac{U'_i + \epsilon'}{\hat{U}_i + \epsilon'} \leqslant n,$$

where the last inequality is due to approximate optimality of $\hat{\mathbf{U}}$. This is a contradiction. Hence, $\hat{\mathbf{U}}$ is a (δ, ϵ)-core outcome.

8.3 Proof of Lemma 10

First, we give a brief overview of the method from [43]. Note that our goal in this argument is to show that the new constraint we added (constraint 4.6) is approximately satisfied after this rounding. The authors work in a setting where points can have *demand*. In our setting, this just corresponds to points in $\mathcal{N}$ having a demand of 1, and moving or consolidating demand can be thought of as changing the instance by moving points in $\mathcal{N}$. Note that the original linear program requires $\mathcal{N} \subseteq \mathcal{M}$, and we follow suit in this proof.

Given a fractional solution to the linear program in Figure 4.3 as $\{y_j\}, \{z_{ij}\}$, let $\bar{C}_i = \sum_{j \in \mathcal{M}} d(i,j) z_{ij}$. That is, $\bar{C}_i$ is the contribution of point i to the k-median objective in the fractional optimum.

- Step 1: Consolidate all demands t_i to obtain $\{t'_i\}$ such that for all $i, j \in \mathcal{M}$ with $t'_i > 0, t'_j > 0$, they must be sufficiently far away such that $c_{ij} > 4 \max(\bar{C}_i, \bar{C}_j)$. Let $\mathcal{M}'$ be the set of centers with positive demand after this step, i.e. $\mathcal{M}' = \{j \in \mathcal{M} : t'_j > 0\}$.

- Step 2a: Consolidate open centers by moving each center not in $\mathcal{M}'$ to the nearest center in $\mathcal{M}'$. This gives a new solution $\{y'_j\}, \{z'_{ij}\}$ with $y'_j \geqslant \frac{1}{2}$ if $j \in \mathcal{M}'$ and $y'_j = 0$ if $j \notin \mathcal{M}'$. We call this a $\frac{1}{2}$-restricted solution.

- Step 2b: Modify the solution further to obtain $\{\bar{y_j}\}, \{\bar{z_{ij}}\}$ with $\bar{y_j} \in \{\frac{1}{2}, 1\}$ if $j \in \mathcal{M}'$ and $\bar{y_j} = 0$ if $j \notin \mathcal{M}'$. We call this a $\{\frac{1}{2}, 1\}$-integral solution.

- Step 3: Round $\{\bar{y_j}\}, \{\bar{z_{ij}}\}$ to obtain an integer solution $\{\hat{y_j}\}, \{\hat{z_{ij}}\}$.

We introduce constraint 4.6 in our linear program, and make a small modification to Step 2b as described in sections below.

8.3.1 Step 1: Consolidating Demands

Our first observation is that during the demand consolidation, it cannot be the case that all of the demand within a given $B(i, \gamma R_i)$ from constraint 4.6 is moved arbitrarily far away.

Lemma 25. *Fix $i \in \mathcal{M}$. For each $j \in B(i, \gamma R_i)$ that had its demand moved to $j' \in \mathcal{M}'$, $d(i, j') \leqslant 9\gamma R_i$.*

Proof. Let $j \in B(i, \gamma R_i)$. Let j' be the location to which the demand for j was moved in Step 1. Note that Step 1 is designed so that if demand at j is moved to another point j', then $\bar{C}_{j'} \leqslant \bar{C}_j$, and $c_{jj'} \leqslant 4\max(\bar{C}_j, \bar{C}_{j'}) = 4\bar{C}_j$. By constraint 4.6, we know that the demand of j could be completely satisfied by centers fractionally opened inside of $B(i, \gamma R_i)$, so $\bar{C}_j \leqslant 2\gamma R_i \implies c_{jj'} \leqslant 8\gamma R_i$. Since $j \in B(i, \gamma R_i)$, it follows that $d(i, j') \leqslant 9R_i$. □

8.3.2 Step 2: Consolidating Centers

Next, we argue that the consolidation of fractional centers in Step 2 approximately preserves constraint 4.6.

Lemma 26. *After Step 2a,* $\forall j \in \mathcal{M}$, $\sum_{j' \in B(j, 9\gamma R_j)} y'_{j'} \geqslant 1$.

Proof. Consider each location $j' \in B(j, \gamma R_j)$. Let j'' be the location in $\mathcal{M}'$ closest to j'. By Lemma 25, there exists such a j'' with $d(j, j'') \leqslant 9\gamma R_j$, so $j'' \in B(j, 9\gamma R_j)$. Step 2a will move the fractional center at j' to j'', so $y'_{j''} \geqslant \min(1, y_{j'} + y_{j''}) \geqslant y_{j'}$. Summing over all $j' \in B(j, \gamma R_j)$, we have $\sum_{j'' \in B(j, 9\gamma R_j)} y'_{j''} \geqslant \sum_{j' \in B(j, \gamma R_j)} y'_j \geqslant 1$ by constraint 4.6. □

For our algorithm, we slightly change Step 2b to the following: Let $\mathcal{M}'' = \{j \in \mathcal{M}' : y'_j < 1\}$, $m' = |\mathcal{M}'|$ and $m'' = |\mathcal{M}''|$. Sort the locations $j \in \mathcal{M}''$ in decreasing order of $t'_j d(s(j), j)$, where $s(j)$ is the location in $\mathcal{M}'$ closest to j (other than j). Set $\bar{y}_j = 1$ for the first $2k - 2m' + m''$ locations in $\mathcal{M}''$ or if $j \in \mathcal{M}' \backslash \mathcal{M}''$, and $\bar{y}_j = \frac{1}{2}$ otherwise. In other words, the only difference from the original Step 2b in [43] is that points with integral $y'_j = 1$ will not participate in the sorting; we simply set $\bar{y}_j = 1$ for such points, and then perform the standard rounding on M''. [43] use the following statement in their proof, and we show that it still holds after our modification to Step 2b.

Lemma 27. *For any $\frac{1}{2}$-restricted solution $\{y'_j\}, \{z'_{ij}\}$, the modified Step 2b givs a $\{\frac{1}{2}, 1\}$-integral solution with no greater cost.*

Proof. From Lemma 7 in [43], the cost of the $\frac{1}{2}$-restricted solution $\{y'_j\}, \{z'_{ij}\}$ is

$$\begin{aligned} &\sum_{j \in \mathcal{M}'} t'_j \left(1 - y'_j\right) d(s(j), j) \\ =& \sum_{j \in \mathcal{M}''} t'_j \left(1 - y'_j\right) d(s(j), j) \\ =& \sum_{j \in \mathcal{M}''} t'_j d(s(j), j)) - \sum_{j \in \mathcal{M}''} t'_j d(s(j), j) y'_j. \end{aligned}$$

where the second line follows because $y'_j = 1 \; \forall j \in \mathcal{M}' \backslash \mathcal{M}''$. Our algorithm in Step 2b maximizes $\sum_{j \in \mathcal{M}''} t'_j d(s(j), j) \bar{y_j}$ for the given set of $t'_j d(s(j), j)$, hence achieves a cost at most that of $\{y'_j\}, \{z'_{ij}\}$. □

Lemma 28. *After Step 2b, $\forall j \in \mathcal{M}$, there is either at least one $j' \in B(j, 9\gamma R_j)$ with $\bar{y_{j'}} = 1$ or at least two $j' \in B(j, 9\gamma R_j)$ with $\bar{y_{j'}} \geqslant \frac{1}{2}$.*

Proof. Given Lemma 26 and the constraints on y'_i after Step 2a, there must be at least one $j' \in B(j, 9\gamma R_j)$ with positive demand. If there is exactly one such j', Lemma 3 is equivalent to $y'_{j'} = 1$ and Step 2b will ensure $\bar{y_{j'}} = 1$. If there are at least two such j', all of them will have $\bar{y_{j'}} \geqslant \frac{1}{2}$ after Step 2b. □

8.3.3 Step 3: Rounding an Integer Solution

[43] gives two rounding schemes, and we use the first one that at most doubles the cost. The important observation is that any center with $\bar{y}_j = 1$ will be opened in the integral solution (that is, if $\bar{y_j} = 1$ then $\hat{y_j} = 1$, and for any center with $\bar{y_j} = \frac{1}{2}$, either j itself or another center in $\mathcal{M}'$ closest to j will be opened. This allows us to complete our argument.

Lemma 29. *For all $j \in \mathcal{M}$, $\sum_{j' \in B(j, 27\gamma R_j)} \hat{y_{j'}} \geqslant 1$.*

Proof. By Lemma 28, there are two cases: either there is some $j' \in B(j, 9\gamma R_j)$ with $\bar{y_{j'}} = 1$ or there are at least two $j' \in B(j, 9\gamma R_j)$ with $\bar{y_{j'}} \geqslant \frac{1}{2}$. In the first case,

$\hat{y_{j'}} = \bar{y_{j'}} = 1$, so the lemma statement clearly holds. In the second case, we are guaranteed that for each point j', either we set $\hat{y_{j'}} = 1$ (in which case the Lemma holds) or we set $\hat{y_{j''}} = 1$ where j'' is the closest other at least partially open center. But since in this case there are *two* points in $B(j, 9\gamma R_j)$ partially open, their pairwise distance is at most $18\gamma R_j$, so $d(j', j'') \leqslant 18\gamma R_j \implies d(j, j'') \leqslant 27\gamma R_j$. □

This concludes the proof of Lemma 10. The fact that $\{\hat{y}_j\}, \{\hat{z_{ij}}\}$ is an 8-approximation of the objective follows immediately from the proof from [43] given Lemma 27, as all other constraints are still satisfied. Finally, we note that the constant factor of 27 can be tightened to 13 in the special case where $\mathcal{N} = \mathcal{M}$. The argument is essentially the same. The crucial improvement comes from the guarantee that $\forall i \in \mathcal{M}$, there is demand at the *center* of each $B(i, \gamma R_i)$. Tracking this demand throughout the rounding leads to the tightened result.

Bibliography

[1] The participatory budgeting project. https://www.participatorybudgeting.org.

[2] Abramowitz, B., and Anshelevich, E. Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence* (2018), AAAI Press, pp. 894–901.

[3] Airiau, S., Aziz, H., Caragiannis, I., Kruger, J., and Lang, J. Positional social decision schemes: Fair and efficient randomized portioning. In *Proceedings of the 7th International Workshop on Computational Social Choice (COMSOC)* (2018).

[4] Airiau, S., and Endriss, U. Iterated majority voting. In *International Conference on Algorithmic Decision Theory* (2009), Springer, pp. 38–49.

[5] Amanatidis, G., Markakis, E., Nikzad, A., and Saberi, A. Approximation algorithms for computing maximin share allocations. *ACM Transactions on Algorithms 13*, 4 (Dec. 2017), 52:1–52:28.

[6] Anari, N., Oveis Gharan, S., Saberi, A., and Singh, M. Nash Social Welfare, Matrix Permanent, and Stable Polynomials. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference (ITCS)* (2017), pp. 36:1–36:12.

[7] Anshelevich, E., Bhardwaj, O., Elkind, E., Postl, J., and Skowron, P. Approximating optimal social choice under metric preferences. *Artificial Intelligence 264* (2018), 27 – 51.

[8] Anshelevich, E., and Postl, J. Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research 58*, 1 (2017), 797–827.

[9] Anshelevich, E., and Zhu, W. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. In *The 14th Conference on Web and Internet Economics (WINE)* (2018), Springer, pp. 3–20.

[10] Arora, S., Hazan, E., and Kale, S. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing 8*, 6 (2012), 121–164.

[11] Arrow, K. J. A difficulty in the concept of social welfare. *Journal of Political Economy 58*, 4 (1950), 328–346.

[12] Arthur, D., and Vassilvitskii, S. K-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (2007), pp. 1027–1035.

[13] Arya, V., Garg, N., Khandekar, R., Meyerson, A., Munagala, K., and Pandit, V. Local search heuristics for k-median and facility location problems. *SIAM Journal on Computing 33*, 3 (2004), 544–562.

[14] Azevedo, E. M., and Budish, E. Strategyproofness in the large as a desideratum for market design. In *EC* (2012), pp. 55–55.

[15] Aziz, H., Brandl, F., and Brandt, F. On the incompatibility of efficiency and strategyproofness in randomized social choice. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence* (2014), AAAI Press, pp. 545–551.

[16] Aziz, H., Brill, M., Conitzer, V., Elkind, E., Freeman, R., and Walsh, T. Justified representation in approval-based committee voting. *Social Choice and Welfare 48*, 2 (2017), 461–485.

[17] Aziz, H., Elkind, E., Huang, S., Lackner, M., Fernandez, L. S., and Skowron, P. On the complexity of extended and proportional justified representation. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence* (2018), AAAI Press, pp. 902–909.

[18] Aziz, H., and Mackenzie, S. A discrete and bounded envy-free cake cutting protocol for any number of agents. In *Proceedings of the 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (2016), pp. 416–427.

[19] Aziz, H., and Mackenzie, S. A discrete and bounded envy-free cake cutting protocol for four agents. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing (STOC)* (2016), ACM, pp. 454–464.

[20] Bandelt, H.-J., and Barthelemy, J.-P. Medians in median graphs. *Discrete Applied Mathematics 8*, 2 (1984), 131 – 142.

[21] Barbera, S., Gul, F., and Stacchetti, E. Generalized median voter schemes and committees. *Journal of Economic Theory 61*, 2 (1993), 262 – 289.

[22] Barman, S., Biswas, A., Krishnamurthy, S., and Narahari, Y. Groupwise maximin fair allocation of indivisible goods. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence* (2018), AAAI Press, pp. 917–924.

[23] Barman, S., and Krishna Murthy, S. K. Approximation algorithms for maximin fair division. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)* (2017), ACM, pp. 647–664.

[24] Barman, S., Krishnamurthy, S. K., and Vaish, R. Greedy algorithms for maximizing nash social welfare. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)* (2018), pp. 7–13.

[25] Benade, G., Nath, S., Procaccia, A. D., and Shah, N. Preference elicitation for participatory budgeting. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (2017), AAAI Press, pp. 376–382.

[26] Bennett, E., and Houba, H. Odd man out: Bargaining among three players. Working Papers 662, UCLA Department of Economics, May 1992.

[27] Bhaskar, U., Dani, V., and Ghosh, A. Truthful and near-optimal mechanisms for welfare maximization in multi-winner elections. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, pp. 925–932.

[28] Binmore, K., Rubinstein, A., and Wolinsky, A. The nash bargaining solution in economic modelling. *The RAND Journal of Economics 17*, 2 (1986), 176–188.

[29] Binmore, K., Shaked, A., and Sutton, J. Testing noncooperative bargaining theory: A preliminary study. *The American Economic Review 75*, 5 (1985), 1178–1180.

[30] Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., and Kalai, A. T. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. In *Advances in Neural Information Processing Systems (NIPS)*. 2016, pp. 4349–4357.

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

[32] Bouveret, S., Chevaleyre, Y., Durand, F., and Lang, J. Voting by sequential elimination with few voters. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)* (2017), pp. 128–134.

[33] Brandl, F., Brandt, F., Eberl, M., and Geist, C. Proving the incompatibility of efficiency and strategyproofness via smt solving. *Journal of the ACM 65*, 2 (2018), 6:1–6:28.

[34] Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. D. *Handbook of Computational Social Choice*, 1st ed. Cambridge University Press, 2016.

[35] Bubeck, S. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning 8*, 3-4 (2015), 231–357.

[36] Buchfuhrer, D., Schapira, M., and Singer, Y. Computation and incentives in combinatorial public projects. In *Proceedings of the 11th ACM Conference on Electronic Commerce* (2010), ACM, pp. 33–42.

[37] Budish, E. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy 119*, 6 (2011), 1061 – 1103.

[38] Cabannes, Y. Participatory budgeting: a significant contribution to participatory democracy. *Environment and Urbanization 16*, 1 (2004), 27–46.

[39] Cai, Y., Daskalakis, C., and Weinberg, S. M. Understanding incentives: Mechanism design becomes algorithm design. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (2013), pp. 618–627.

[40] Caliskan, A., Bryson, J. J., and Narayanan, A. Semantics derived automatically from language corpora contain human-like biases. *Science 356*, 6334 (2017), 183–186.

[41] Caragiannis, I., Kurokawa, D., Moulin, H., Procaccia, A. D., Shah, N., and Wang, J. The unreasonable fairness of maximum nash welfare. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)* (2016), ACM, pp. 305–322.

[42] Caragiannis, I., and Procaccia, A. D. Voting almost maximizes social welfare despite limited communication. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence* (2010), AAAI Press, pp. 743–748.

[43] Charikar, M., Guha, S., va Tardos, and Shmoys, D. B. A constant-factor approximation algorithm for the k-median problem. *Journal of Computer and System Sciences 65*, 1 (2002), 129 – 149.

[44] Chen, X., Dai, D., Du, Y., and Teng, S.-H. Settling the Complexity of Arrow-Debreu Equilibria in Markets with Additively Separable Utilities. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (2009), pp. 273–282.

[45] Chen, X., Fain, B., Lyu, L., and Munagala, K. Proportionally fair clustering. Manuscript under review at the the 36th International Conference on Machine Learning (ICML).

[46] Cheng, Y., Dughmi, S., and Kempe, D. Of the people: Voting is more effective with representative candidates. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)* (2017), ACM, pp. 305–322.

[47] Chierichetti, F., Kumar, R., Lattanzi, S., and Vassilvitskii, S. Fair clustering through fairlets. In *Advances in Neural Information Processing Systems (NIPS)*. 2017, pp. 5029–5037.

[48] Clearwater, A., Puppe, C., and Slinko, A. Generalizing the single-crossing property on lines and trees to intermediate preferences on median graphs. In *Proceedings of the 24th International Conference on Artificial Intelligence (IJCAI)* (2015), pp. 32–38.

[49] Cole, R., Devanur, N., Gkatzelis, V., Jain, K., Mai, T., Vazirani, V. V., and Yazdanbod, S. Convex program duality, fisher markets, and nash social welfare. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)* (2017), ACM, pp. 459–460.

[50] Cole, R., and Gkatzelis, V. Approximating the nash social welfare with indivisible items. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing (STOC)* (2015), ACM, pp. 371–380.

[51] Cole, R., Gkatzelis, V., and Goel, G. Mechanism design for fair division: Allocating divisible items without payments. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce (EC)* (2013), ACM, pp. 251–268.

[52] Collins, F. S., and Varmus, H. A new initiative on precision medicine. *New England Journal of Medicine 372*, 9 (2015), 793–795.

[53] Conitzer, V., Freeman, R., and Shah, N. Fair public decision making. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)* (2017), pp. 629–646.

[54] Conitzer, V., Freeman, R., Shah, N., and Vaughan, J. W. Group fairness for indivisible goods allocation. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence* (2019), AAAI Press.

[55] Conitzer, V., and Sandholm, T. Communication complexity of common voting rules. In *Proceedings of the 6th ACM Conference on Electronic Commerce (EC)* (2005), ACM, pp. 78–87.

[56] Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., and Huq, A. Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* (2017), pp. 797–806.

[57] Daskalakis, C., Devanur, N., and Weinberg, S. M. Revenue maximization and ex-post budget constraints. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation* (2015), ACM, pp. 433–447.

[58] Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. The complexity of computing a nash equilibrium. *SIAM Journal on Computing 39*, 1 (2009), 195–259.

[59] Dheeru, D., and Karra Taniskidou, E. Uci machine learning repository, 2017.

[60] Dughmi, S. A truthful randomized mechanism for combinatorial public projects via convex optimization. In *Proceedings of the 12th ACM Conference on Electronic Commerce (EC)* (2011), ACM, pp. 263–272.

[61] Dwork, C., Hardt, M., Pitassi, T., Reingold, O., and Zemel, R. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS)* (2012), pp. 214–226.

[62] Esfandiari, H., and Kortsarz, G. A bounded-risk mechanism for the kidney exchange game. In *LATIN 2016: Theoretical Informatics* (2016), E. Kranakis, G. Navarro, and E. Chávez, Eds., Springer, pp. 416–428.

[63] Fain, B., Goel, A., and Munagala, K. The core of the participatory budgeting problem. In *Proceedings of the 12th International Conference on Web and Internet Economics (WINE)* (2016), Springer, pp. 384–399.

[64] Fain, B., Goel, A., Munagala, K., and Prabhu, N. Random dictators with a random referee: Constant sample complexity mechanisms for social choice. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence* (2019), AAAI Press.

[65] Fain, B., Goel, A., Munagala, K., and Sakshuwong, S. Sequential deliberation for social choice. In *Proceedings of the 13th International Conference on Web and Internet Economics (WINE)* (2017), Springer, pp. 177–190.

[66] Fain, B., Munagala, K., and Shah, N. Fair allocation of indivisible public goods. In *Proceedings of the 2018 ACM Conference on Economics and Computation (EC)* (2018), ACM, pp. 575–592.

[67] Feldman, M., Fiat, A., and Golomb, I. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)* (2016), ACM, pp. 269–286.

[68] Feldman, M., Immorlica, N., Lucier, B., Roughgarden, T., and Syrgkanis, V. The price of anarchy in large games. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)* (2016), ACM, pp. 963–976.

[69] Filos-Ratsikas, A., and Miltersen, P. B. Truthful approximations to range voting. In *The 10th International Conference on Web and Internet Economics (WINE)* (2014), Springer, pp. 175–188.

[70] Fluschnik, T., Skowron, P., Triphaus, M., and Wilker, K. Fair Knapsack. *arXiv e-prints* (2017), arXiv:1711.04520.

[71] Foley, D. K. Lindahl's solution and the core of an economy with public goods. *Econometrica 38*, 1 (1970), pp. 66–72.

[72] Freeman, R., Zahedi, S. M., Conitzer, V., and Lee, B. C. Dynamic proportional sharing: A game-theoretic approach. *Proceedings of the ACM on Measurement and Analysis of Computing Systems 2*, 1 (2018), 3:1–3:36.

[73] Gal, Y. K., Mash, M., Procaccia, A. D., and Zick, Y. Which is the fairest (rent division) of them all? In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)* (2016), ACM, pp. 67–84.

[74] Garg, N., Goel, A., and Plaut, B. Markets for Public Decision-making. *arXiv e-prints* (2018), arXiv:1807.10836.

[75] Garg, N., Kamble, V., Goel, A., Marn, D., and Munagala, K. Collaborative optimization for collective decision-making in continuous spaces. In *Proceedings of the 26th International Conference on World Wide Web (WWW)* (2017), pp. 617–626.

[76] Ghodsi, A., Zaharia, M., Hindman, B., Konwinski, A., Shenker, S., and Stoica, I. Dominant resource fairness: Fair allocation of multiple resource types. In *Proceedings of the 8th USENIX Symposium on Networked System Design and Implementation (NSDI)* (2011), USENIX Association, pp. 323–336.

[77] Ghodsi, M., Hajiaghayi, M., Seddighin, M., Seddighin, S., and Yami, H. Fair allocation of indivisible goods: Improvements and generalizations. In *Proceedings of the 2018 ACM Conference on Economics and Computation (EC)* (2018), ACM, pp. 539–556.

[78] Gibbard, A. Manipulation of voting schemes: A general result. *Econometrica 41*, 4 (1973), 587–601.

[79] Goel, A., Krishnaswamy, A. K., and Munagala, K. Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)* (2017), ACM, pp. 287–304.

[80] Goel, A., Krishnaswamy, A. K., and Sakshuwong, S. Budget aggregation via knapsack voting: Welfare-maximization and strategy-proofness. *Collective Intelligence* (2016).

[81] Goel, A., Krishnaswamy, A. K., Sakshuwong, S., and Aitamurto, T. Knapsack voting. *Collective Intelligence* (2015).

[82] Goel, A., and Lee, J. Towards large-scale deliberative decision-making: Small groups and the importance of triads. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)* (2016), ACM, pp. 287–303.

[83] Goel, N., Yaghini, M., and Faltings, B. Non-discriminatory machine learning through convex fairness criteria. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence* (2018), AAAI Press, pp. 3029–3036.

[84] Gonzalez, T. F. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science 38* (1985), 293 – 306.

[85] Gross, S., Anshelevich, E., and Xia, L. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (2017), AAAI Press, pp. 544–550.

[86] Groves, T., and Ledyard, J. Optimal allocation of public goods: A solution to the "free rider" problem. *Econometrica 45*, 4 (1977), pp. 783–809.

[87] Hardt, M., Price, E., , and Srebro, N. Equality of opportunity in supervised learning. In *Advances in Neural Information Processing Systems (NIPS)*. 2016, pp. 3315–3323.

[88] Harsanyi, J. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge University Press, 1977.

[89] Harsanyi, J. C. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy 63*, 4 (1955), 309–321.

[90] Harsanyi, J. C. A simplified bargaining model for the n-person cooperative game. *International Economic Review 4*, 2 (1963), 194–220.

[91] Hashimoto, T., Srivastava, M., Namkoong, H., and Liang, P. Fairness without demographics in repeated loss minimization. In *Proceedings of the 35th International Conference on Machine Learning (ICML)* (2018), pp. 1929–1938.

[92] Hougardy, S., and Vinkemeier, D. E. Approximating weighted matchings in parallel. *Information Processing Letters 99*, 3 (2006), 119 – 123.

[93] Huang, Z., and Kannan, S. The exponential mechanism for social welfare: Private, truthful, and nearly optimal. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (2012), pp. 140–149.

[94] Jain, A. K. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters 31*, 8 (2010), 651 – 666.

[95] Jain, K., Mahdian, M., and Saberi, A. A new greedy approach for facility location problems. In *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing (STOC)* (2002), pp. 731–740.

[96] Jain, K., and Vazirani, V. V. Primal-dual approximation algorithms for metric facility location and k-median problems. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (1999), pp. 2–13.

[97] Jain, K., and Vazirani, V. V. Eisenberg-gale markets: Algorithms and structural properties. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing (STOC)* (2007), ACM, pp. 364–373.

[98] Jain, K., Vazirani, V. V., and Ye, Y. Market equilibria for homothetic, quasi-concave utilities and economies of scale in production. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium Discrete Algorithms (SODA)* (2005), pp. 63–71.

[99] Jehle, G., and Reny, P. *Advanced Economic Theory*, 3rd ed. Pearson, 2011.

[100] Joe-Wong, C., Sen, S., Lan, T., and Chiang, M. Multiresource allocation: Fairness-efficiency tradeoffs in a unifying framework. *IEEE/ACM Transactions on Networking 21*, 6 (2013), 1785–1798.

[101] Johansen, L. Some notes on the lindahl theory of determination of public expenditures. *International Economic Review 4*, 3 (1963), 346–358.

[102] Julia Angwin, Jeff Larson, S. M., and Kirchner, L. Machine bias, 2016.

[103] Kalai, E. Proportional solutions to bargaining situations: Interpersonal utility comparisons. *Econometrica 45*, 7 (1977), 1623–1630.

[104] Kalai, E., and Smorodinsky, M. Other solutions to nash's bargaining problem. *Econometrica 43*, 3 (1975), 513–518.

[105] Kearns, M., Neel, S., Roth, A., and Wu, Z. S. Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. In *Proceedings of the 35th International Conference on Machine Learning (ICML)* (2018), pp. 2569–2577.

[106] Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. Human decisions and machine predictions. Working Paper 23180, National Bureau of Economic Research, 2017.

[107] Kleinberg, J., Mullainathan, S., and Raghavan, M. Inherent Trade-Offs in the Fair Determination of Risk Scores. *ArXiv e-prints* (2016).

[108] Knuth, D. E. *The Art of Computer Programming: Combinatorial Algorithms, Part 1*, 1st ed. Addison-Wesley Professional, 2011.

[109] Koutsoupias, E., and Papadimitriou, C. Worst-case equilibria. In *Proceedings of the 16th Annual Conference on Theoretical Aspects of Computer Science* (1999), Springer-Verlag, pp. 404–413.

[110] Krishna, V., and Serrano, R. Multilateral bargaining. *The Review of Economic Studies 63*, 1 (1996), 61–80.

[111] Kuhn, H. W., and Tucker, A. W. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (1951), University of California Press, pp. 481–492.

[112] Kung, J. P. S. Basis-exchange properties. In *Theory of Matroids*, W. N., Ed. Cambridge University Press, 1986.

[113] Kunjir, M., Fain, B., Munagala, K., and Babu, S. Robus: Fair cache allocation for data-parallel workloads. In *Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD)* (2017), ACM, pp. 219–234.

[114] Lang, J., and Xia, L. Voting in combinatorial domains. In *Handbook of Computational Social Choice*, F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, Eds. Cambridge University Press, 2016.

[115] Lee, E. Apx-hardness of maximizing nash social welfare with indivisible items. *Information Processing Letters 122* (2017), 17 – 20.

[116] Lev, O., and Rosenschein, J. S. Convergence of iterative voting. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2* (2012), pp. 611–618.

[117] Levin, D., Peres, Y., and Wilmer, E. *Markov Chains and Mixing Times.* American Mathematical Society, 2009.

[118] Lindahl, E. Just taxation: A positive solution. In *Classics in the Theory of Public Finance*. Palgrave Macmillan, 1958.

[119] Lipton, R. J., Markakis, E., Mossel, E., and Saberi, A. On approximately fair allocations of indivisible goods. In *Proceedings of the 2004 ACM Conference on Electronic Commerce (EC)* (2004), ACM, pp. 125–131.

[120] Liu, Q., and Pycia, M. Ordinal efficiency, fairness, and incentives in large markets, 2012.

[121] Lovász, L., and Vempala, S. The geometry of logconcave functions and sampling algorithms. *Random Structures and Algorithms 30*, 3 (May 2007), 307–358.

[122] Luzum, N. Voters to decide fate of durham bond referenda. *The Chronicle* (2016).

[123] McSherry, F., and Talwar, K. Mechanism design via differential privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (2007), pp. 94–103.

[124] Meir, R., Polukarov, M., Rosenschein, J. S., and Jennings, N. R. Convergence to equilibria in plurality voting. In *Proceedings of Twenty-Fourth AAAI Conference on Artificial Intelligence* (2010), AAAI Press, pp. 823–828.

[125] Mettu, R. R., and Plaxton, C. G. Optimal time bounds for approximate clustering. *Machine Learning 56*, 1-3 (2004), 35–60.

[126] Muench, T. J. The core and the lindahl equilibrium of an economy with a public good: an example. *Journal of Economic Theory 4*, 2 (1972), 241 – 255.

[127] Myerson, R. B. Two-person bargaining problems and comparable utility. *Econometrica 45*, 7 (1977), 1631–1637.

[128] Myerson, R. B. Two-person bargaining problems with incomplete information. *Econometrica 52*, 2 (1984), 461–487.

[129] Myerson, R. B., and Satterthwaite, M. A. Efficient mechanisms for bilateral trading. *Journal of Economic Theory 29*, 2 (1983), 265 – 281.

[130] Nash, J. F. The bargaining problem. *Econometrica 18*, 2 (1950), pp. 155–162.

[131] Neelin, J., Sonnenschein, H., and Spiegel, M. An Experimental Test of Rubinstein's Theory of Bargaining. Working Papers 587, Princeton University, Department of Economics, Industrial Relations Section, 1986.

[132] Nehring, K., and Puppe, C. The structure of strategy-proof social choice. part i: General characterization and possibility results on median spaces. *Journal of Economic Theory 135*, 1 (2007), 269 – 305.

[133] Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. *Algorithmic game theory*, vol. 1. Cambridge University Press Cambridge, 2007.

[134] Papadimitriou, C., Schapira, M., and Singer, Y. On the hardness of being truthful. In *Proceedings of the 49th Annual IEEE Conference on Foundations of Computer Science* (2008), pp. 250–259.

[135] Parkes, D. C., Procaccia, A. D., and Shah, N. Beyond dominant resource fairness: Extensions, limitations, and indivisibilities. *ACM Transactions Economics and Computation 3*, 1 (2015), 3:1–3:22.

[136] Plaut, B., and Roughgarden, T. Almost envy-freeness with general valuations. In *Proceedings of the 29th Annual ACM/SIAM Symposium on Discrete Algorithms (SODA)* (2018), pp. 2584–2603.

[137] Pleiss, G., Raghavan, M., Wu, F., Kleinberg, J., and Weinberger, K. Q. On fairness and calibration. In *Advances in Neural Information Processing Systems (NIPS)*. 2017, pp. 5680–5689.

[138] Plotkin, S. A., Shmoys, D. B., and Tardos, É. Fast approximation algorithms for fractional packing and covering problems. *Mathematics of Operations Research 20*, 2 (1995), 257–301.

[139] Popa, L., Kumar, G., Chowdhury, M., Krishnamurthy, A., Ratnasamy, S., and Stoica, I. Faircloud: Sharing the network in cloud computing. In *Proceedings of the 2012 ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM)* (2012), ACM, pp. 187–198.

[140] Procaccia, A., Wajc, D., and Zhang, H. Approximation-variance tradeoffs in facility location games. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence* (2018), AAAI Press, pp. 1185–1192.

[141] Procaccia, A. D. Cake cutting: Not just child's play. *Communications of the ACM 56*, 7 (July 2013), 78–87.

[142] Procaccia, A. D., and Rosenschein, J. S. The distortion of cardinal preferences in voting. In *Cooperative Information Agents X* (2006), M. Klusch, M. Rovatsos, and T. R. Payne, Eds., Springer, pp. 317–331.

[143] Procaccia, A. D., and Wang, J. Fair enough: Guaranteeing approximate maximin shares. In *Proceedings of the 2014 ACM Conference on Economics and Computation (EC)* (2014), ACM, pp. 675–692.

[144] Rabinovich, Z., Obraztsova, S., Lev, O., Markakis, E., and Rosenschein, J. S. Analysis of equilibria in iterative voting schemes. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence* (2015), AAAI Press, pp. 1007–1013.

[145] Roth, A. E. Bargaining phenomena and bargaining theory. In *Laboratory Experiments in Economics: Six Points of View*. Cambridge University Press, 1987, pp. 14–41.

[146] Rubinstein, A. Perfect equilibrium in a bargaining model. *Econometrica 50*, 1 (1982), 97–109.

[147] Saban, D., and Stier-Moses, N. The competitive facility location problem in a duopoly: Connections to the 1-median problem. In *Proceedings of the 8th International Conference on Web and Internet Economics (WINE)* (2012), Springer, pp. 539–545.

[148] Samuelson, P. A. The pure theory of public expenditure. *The Review of Economics and Statistics 36*, 4 (1954), 387–389.

[149] Sánchez-Fernández, L., Elkind, E., Lackner, M., Fernández, N., Fisteus, J. A., Basanta Val, P., and Skowron, P. Proportional justified representation. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (2017), AAAI Press, pp. 670–676.

[150] Scarf, H. E. The core of an n person game. *Econometrica 35*, 1 (1967), pp. 50–69.

[151] Schapira, M., and Singer, Y. Inapproximability of combinatorial public projects. In *Proceedings of the 4th International Conference on Web and Internet Economics (WINE)* (2008), Springer, pp. 351–361.

[152] Schummer, J. Strategy-proofness versus efficiency on restricted domains of exchange economies. *Social Choice and Welfare 14* (1997), 47–56.

[153] Schummer, J., and Vohra, R. V. Strategy-proof location on a network. *Journal of Economic Theory 104*, 2 (2002), 405 – 428.

[154] Shapley, L. S. Utility comparison and the theory of games, 1967.

[155] Shoham, Y., and Leyton-Brown, K. *Multiagent Systems. Algorithmic, Game Theoretic, and Logical Foundations*. Cambridge University Press, 2009.

[156] Thiele, T. N. Om flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger* (1895), pp. 415–441.

[157] Varian, H. Two problems in the theory of fairness. *Journal of Public Economics 5*, 3-4 (1976), 249–260.

[158] Vazirani, V. V., and Yannakakis, M. Market equilibrium under separable, piecewise-linear, concave utilities. *Journal of the ACM 58*, 3 (2011), 10:1–10:25.

[159] Wendell, R. E., and McKelvey, R. D. New perspectives in competitive location theory. *European Journal of Operational Research 6*, 2 (1981), 174–182.

[160] Williamson, D. P., and Shmoys, D. B. *The Design of Approximation Algorithms*, 1st ed. Cambridge University Press, New York, NY, USA, 2011.

[161] Zafar, M. B., Valera, I., Gomez Rodriguez, M., and Gummadi, K. P. Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *Proceedings of the 26th International Conference on World Wide Web (WWW)* (2017), pp. 1171–1180.

[162] Zafar, M. B., Valera, I., Rodriguez, M., Gummadi, K., and Weller, A. From parity to preference-based notions of fairness in classification. In *Advances in Neural Information Processing Systems (NIPS)*. 2017, pp. 229–239.